



ATLAS NOTE

ATL-COM-PHYS-2016-1724

23rd June 2017



Draft version 1.5

Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a pair of b -quarks

Abstract

This supporting note describes a search for the SM Higgs boson produced in association with a vector boson, and decaying to a pair of b -quarks using pp collision data collected in 2015 and 2016 LHC Run-2 at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of $\approx 36.1 \text{ fb}^{-1}$. It includes description of object and event selections and signal and background modeling. Systematic uncertainty, statistical treatment used to extract signal strength and results are discussed. The observed signal strength and significance are shown in main analysis using multi-variate analysis. SM diboson (VZ) measurement and signal strength extraction of VH using m_{bb} shape are also performed as cross section of main analysis. The combination with the Run-1 result is also described.

15 Ahmadov, Faig^y, Argyropoulos, Spyridon^r, Arnold, Hannah^k, Amaral Coutinho, Yara^p,
 16 Sanchez Pineda, Arturo Rodolfo^q, Bell, Andrew Stuart^f, Benitez, Jose^r, Buzatu, Adrian^h,
 17 Calderini, Giovanni^v, Chan, Stephen Kam-wah^e, Chen, Chunhui^t, Buckley, Andrew^o, Buscher,
 18 Daniel^k, Delporte, Charles^l, De Lorenzi, Francesco^t, De Vivie De Regie, Jean-Baptiste^l, Dao,
 19 Valerio^a, Enari, Yuji^g, Francavilla, Paolo^v, Gargiulo, Simona^k, Gaycken, Goetzⁿ, Gray, Chloe^o,
 20 Grivaz, Jean-Francois^l, Hays, Jonathan Michael^j, Hesketh, Gavin Grant^f, Hobbs, John^a, Huth,
 21 John^e, Ishijima, Naoki^u, Kado, Marumi^l, Kato, Chikuma^g, Li, Changqiao^b, Li, Haifeng^a, Liu,
 22 Yanwen^b, Marchiori, Giovanni^v, Ma, Lianliangⁱ, Ma, Yanhuiⁱ, Mario Jose da Cunha Sargedas
 23 de Sousa^x, Masubuchi, Tatsuya^g, Mallik, Usha^r, Mehta, Andrew^s, Montalbano, Alyssa^a,
 24 Morange, Nicolas^l, Moser, Brian^k, Pandini, Carlo Enrico^d, Piacquadio, Giacinto^a, Prell,
 25 Soeren^t, Robson, Aidan^o, Scanlon, Tim^f, Schopf, Elisabethⁿ, Sekula, Stephen^w, Thompson,
 26 Paul^m, Tsybychev, Dmitri^a, Vergel Infante, Carlos Miguel^t, Wang, Song-Ming^h, Wang, Wei^{h,z},
 27 Weiser, Christian^k, Yamaguchi, Yohei^c, Yang, Siqi^r, Yu, Jie^t

²⁸ ^a*Stony Brook University*

²⁹ ^b*University of Science and Technology of China*

³⁰ ^c*Tokyo Institute of Technology*

³¹ ^d*INFN Pisa and Universita' di Pisa, Dipartimento di Fisica*

³² ^e*Harvard University*

³³ ^f*Department of Physics and Astronomy, University College London*

³⁴ ^g*International Center for Elementary Particle Physics and Department of Physics, The University of Tokyo*

³⁵ ^h*Academia Sinica, Taipei*

³⁶ ⁱ*Shandong University (CN)*

³⁷ ^j*Queen Mary University of London*

³⁸ ^k*Albert-Ludwigs-Universitaet Freiburg*

³⁹ ^l*LAL, Univ. Paris-Sud, IN2P3/CNRS, Universite Paris-Saclay*

⁴⁰ ^m*University of Birmingham*

⁴¹ ⁿ*University of Bonn*

⁴² ^o*University of Glasgow, SUPA - School of Physics and Astronomy*

⁴³ ^p*Universidade Federal do Rio De Janeiro, COPPE/EE/IF*

⁴⁴ ^q*INFN Gruppo Collegato di Udine and Universita' di Udine, Dipartimento di Chimica, Fisica e Ambiente*

⁴⁵ ^r*University of Iowa*

⁴⁶ ^s*University of Liverpool*

⁴⁷ ^t*Iowa State University*

⁴⁸ ^u*Osaka University*

⁴⁹ ^v*Laboratoire de Physique Nucleaire et de Hautes Energies (LPNHE), UPMC, Universite Paris-Diderot, CNRS/IN2P3*

⁵⁰ ^w*Southern Methodist University, Department of Physics*

⁵¹ ^x*LIP and FCUL*

⁵² ^y*JINR*

⁵³ ^z*Nanjing University, Department of Physics*

54 **Contents**

55	1 Introduction	7
56	2 Data and Simulated Samples	10
57	3 Event Reconstruction and Selection	13
58	3.1 Objects identification	13
59	3.2 Event selection	13
60	3.2.1 Event selection common to 0-, 1- and 2-lepton channels	14
61	3.2.2 0-lepton channel specific selection	15
62	3.2.3 1-lepton channel specific selection	16
63	3.2.4 2-lepton channel specific selection	18
64	3.2.5 Additional selection criteria for the dijet mass analysis cross-check	18
65	4 Signal and Background Modeling	20
66	4.1 MC Signal and Backgrounds	20
67	4.2 Truth tagging	20
68	4.3 Background Control Regions	21
69	4.3.1 1-lepton	21
70	4.3.2 2-lepton	22
71	4.4 Multijet and non-collisional backgrounds	25
72	4.4.1 0-lepton	25
73	4.4.2 1-lepton	29
74	4.4.3 2-lepton	34
75	4.5 Pre-fit data-MC comparison	39
76	4.5.1 0-lepton channel	39
77	4.5.2 1-lepton channel	41
78	4.5.3 2-lepton channel	44
79	5 Multi-Variate Analysis	51
80	5.1 Input variables	51
81	5.2 Setup and training	52
82	5.3 BDT Transformation	58
83	5.4 Optimization studies and performances	59
84	5.4.1 0-lepton MVA performance	59
85	5.4.2 1-lepton MVA performance	60
86	5.4.3 2-lepton MVA performance	62
87	6 Experimental Systematic Uncertainties	65
88	7 Statistical Treatment	67
89	7.1 Likelihood function definition	67
90	7.2 Fit inputs	68
91	7.3 Nuisance parameters: normalisation and systematic uncertainties	68
92	7.3.1 Smoothing of the Systematic Uncertainties	69
93	7.3.2 Pruning of the Systematic Uncertainties	72
94	7.4 Fit configuration validation	73

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

95	7.4.1	Nuisance parameter pulls and constraints	73
96	7.4.2	Correlation matrices	78
97	7.4.3	Nuisance parameter ranking	79
98	7.4.4	Breakdown of statistical and systematic uncertainties	80
99	7.4.5	Post-fit BDT distributions	81
100	7.4.6	Background-Subtracted Post-fit BDT distributions	85
101	7.4.7	Signal over Background Plots	88
102	7.5	Fit to diboson VZ resonance	90
103	7.5.1	Nuisance parameter pulls and constraints	90
104	7.5.2	Correlation matrices	95
105	7.5.3	Nuisance parameter ranking	96
106	7.5.4	Breakdown of statistical and systematic uncertainties	97
107	7.5.5	Post-fit BDT distributions	98
108	7.5.6	Signal over Background Plots	101
109	8	Run 2 Results	104
110	8.1	VH results	104
111	8.2	VZ Diboson results	107
112	9	Combination of Run 1 and Run 2 results	110
113	10	Conclusions	113
114	Appendices		114
115	A	0 Lepton Multijet distributions	115
116	B	2 Lepton Multijet studies	119
117	B.1	p_T^V dependence	119
118	B.2	Impact on buggy muon reconstruction and charge mis-assignment	120
119	C	Data-driven top background estimation	121
120	D	Kinematic Fitter Performance Study and Bias check	126
121	E	Prefit data-MC comparisons	130
122	E.0.1	0-lepton channel	130
123	E.0.2	1-lepton channel	135
124	E.0.3	2-lepton channel	145
125	F	Effect of MC stat uncertainties on the fit	157
126	G	Truth Tagging Closure	160
127	H	Continuous Tagging	162
128	I	BDT optimization studies	163
129	I.1	1-lepton	163

Not reviewed, for internal circulation only

130	I.1.1	MVA performance in 2-jet	163
131	I.1.2	MVA performance in 3-jet	166
132	I.2	2-lepton	169
133	I.2.1	MVA performance in 2jet, $150 \text{ GeV} < p_T^V$	169
134	I.2.2	MVA performance in 2jet, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$	171
135	I.2.3	MVA performance in 3+jet, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$	173
136	J	Fit performances in each leptonic channel	175
137	J.1	0-lepton channel	175
138	J.1.1	Nuisance parameter pulls and constraints	175
139	J.1.2	Correlation matrices	175
140	J.1.3	Nuisance parameter ranking	175
141	J.1.4	Post-fit distributions	175
142	J.1.5	Breakdown of statistical and systematic uncertainties	175
143	J.2	1-lepton channel	180
144	J.2.1	Nuisance parameter pulls and constraints	180
145	J.2.2	Correlation matrices	180
146	J.2.3	Nuisance parameter ranking	183
147	J.2.4	Post-fit distributions	183
148	J.2.5	Breakdown of statistical and systematic uncertainties	183
149	J.3	2-lepton channel	186
150	J.3.1	Nuisance parameter pulls and constraints	186
151	J.3.2	Correlation matrices	186
152	J.3.3	Nuisance parameter ranking	189
153	J.3.4	Post-fit distributions	189
154	J.3.5	Breakdown of statistical and systematic uncertainties	189
155	J.4	Nuisance parameter pulls and constraints in all channels	194
156	J.5	Detailed fit significances	198
157	K	Post-fit BDT input Distributions	199
158	K.1	0-lepton	199
159	K.2	1-lepton	204
160	K.3	2-lepton	216
161	L	Cross-check: Standard Model Diboson ($VZ \rightarrow b\bar{b}$) measurement	225
162	L.1	0-lepton	225
163	L.1.1	Prefit BDT distribution	225
164	L.1.2	Comparison to VH Fit	226
165	L.1.3	Results	227
166	L.2	1-lepton	231
167	L.2.1	Prefit BDT distribution	231
168	L.2.2	Comparison to VH Fit	232
169	L.2.3	Results	236
170	L.3	2-lepton	238
171	L.3.1	Prefit BDT distribution	238
172	L.3.2	Pull distribution, correlation and ranking check	238
173	L.3.3	Results	242

174	L.3.4 Comparison to <i>VH</i> Fit	243
175	L.4 Detailed fit significances	246
176	M Cross-check: dijet mass analysis	247
177	M.1 0-lepton	247
178	M.1.1 Prefit m_{bb} distribution	247
179	M.1.2 Comparison to <i>VH</i> Fit	248
180	M.1.3 Results	249
181	M.2 1-lepton	253
182	M.2.1 Prefit m_{bb} distribution	253
183	M.2.2 Comparison to <i>VH</i> Fit	254
184	M.2.3 Results	255
185	M.3 2-lepton	259
186	M.3.1 Prefit m_{bb} distribution	259
187	M.3.2 Pull distribution, correlation and ranking check	259
188	M.3.3 Results	264
189	M.4 Combined Fit	266
190	M.4.1 Pull Comparison with the unconditional fit to the data in <i>VH</i> Cut-based analysis	266
191	M.4.2 Results	267
192	M.4.3 Background subtracted post-fit m_{bb} distributions	275
193	M.5 Fit significances	278
194	N Detailed Studies on the combination of Run 1 and Run 2	279
195	N.1 Breakdown of statistical and systematic uncertainties	279
196	N.2 Correlation Studies	280
197	N.2.1 Jet Energy Scale	280
198	N.2.2 Modeling Systematics	281
199	N.2.3 Flavour Tagging	282

200 1. Introduction

201 In the Standard Model of particle physics (SM), the Higgs boson results from the Brout-Englert-Higgs
 202 mechanism [1–4], which is needed to give mass to the W and Z gauge bosons and to the fermions.
 203 During the Run-1 of the LHC, a Higgs boson-like particle was discovered by the ATLAS and CMS
 204 collaborations [5, 6]. Since the discovery, the ATLAS and CMS experiments have precisely measured the
 205 new boson properties. The mass, obtained by combining the ATLAS and CMS Run-1 measurements, is
 206 equal to 125.09 ± 0.24 GeV [7]. Studies of the spin-parity is in favor of the SM Higgs boson quantum
 207 numbers, $J^P = 0^+$ [8, 9]. The production, decay rates and coupling to bosons and fermions have also been
 208 measured by the ATLAS and CMS collaborations, and are in agreement with the predictions for a SM
 209 Higgs boson: significant Higgs boson signals decaying to a pair of photons, Z and W bosons, and τ leptons,
 210 were observed in Run-1 [10]. However, the Higgs boson has not yet been observed in the $H \rightarrow b\bar{b}$ decay
 211 channel, in spite of the large branching ratio ($\sim 58\%$) [11]. In this decay channel, a signal significance of
 212 2.1 standard deviations has been measured at the CMS experiment [12], and of 1.4 standard deviations
 213 at the ATLAS experiment [13]; when combining the ATLAS and CMS measurements, an observed
 214 significance of 2.6 standard deviations (and an expected significance of 3.7σ) is obtained at the end of
 215 Run-1 [10]. An observation of the decay to a pair of b -quarks is still one of the milestones of the LHC
 216 experiments physics programme during Run-2.

217 Figure 2 shows the production cross section and branching ratio of the Higgs boson as a function of
 218 its mass. Although the gluon fusion process is dominant at the LHC energies, the use of the $H \rightarrow b\bar{b}$
 219 decay channel is very challenging due to the large multijet background contamination. The sub-dominant
 220 VH production mode is one of the most promising to search for $H \rightarrow b\bar{b}$ at the LHC; this is due to the
 221 multijet background suppression achieved by the presence of the associated vector boson and its decay to
 222 leptons.

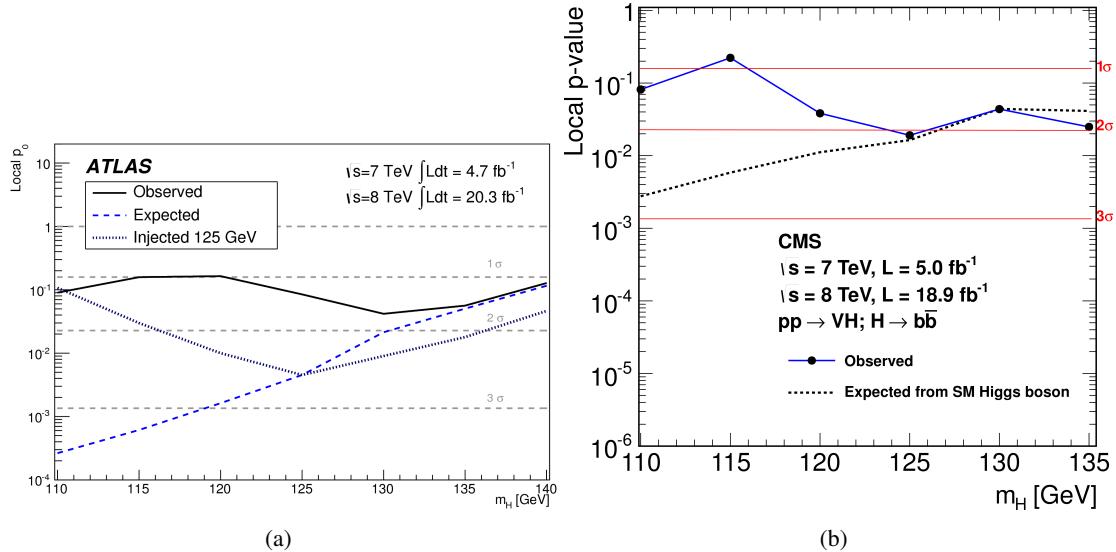


Figure 1: The observed and expected p_0 values as a function of m_H in Run-1 (the ATLAS experiment) (a). The observed and expected p_0 values as a function of m_H in Run-1 (the CMS experiment) (b).

223 In this document, an update of the first Run-2 search [14] based on Run-1 [13] for a SM Higgs boson

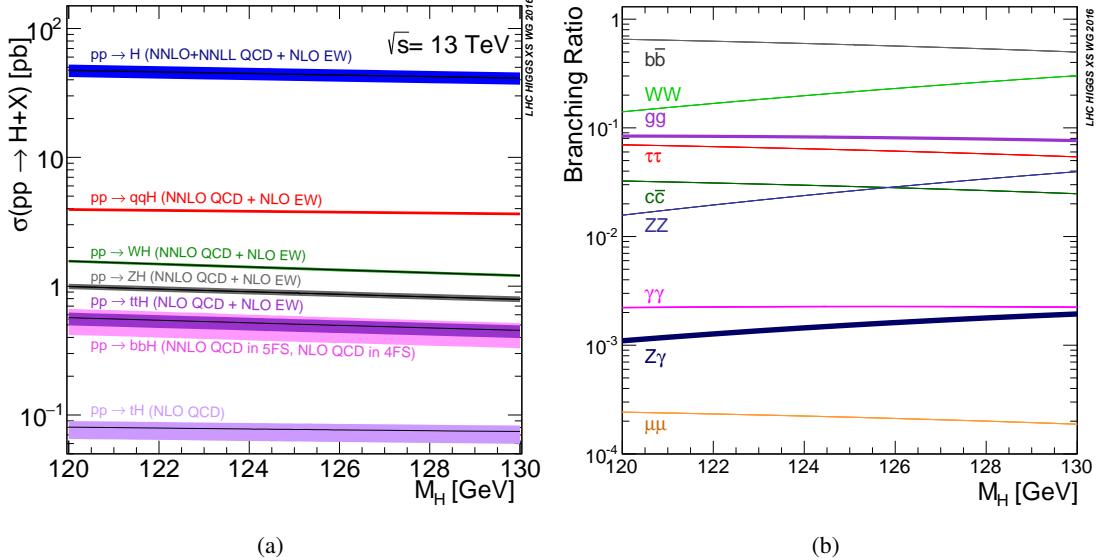


Figure 2: Higgs production cross section as a function of the Higgs boson mass for each production mode (a). Branching ratio as a function of the Higgs boson mass (b).

Not reviewed, for internal circulation only

produced in association with a vector boson, and decaying to a pair of b -quarks, is presented. The first Run-2 CONF note using 13.2 fb^{-1} was released for ICHEP 2016. The ratio of the measured signal strength to the SM expectation was found to be $\mu = 0.21^{+0.36}_{-0.35}(\text{stat.}) \pm 0.35(\text{syst.})$. This corresponds to an observed(expected) significance of $0.42(1.92)\sigma$. Since ICHEP 2016, the LHC has delivered more than 20 fb^{-1} . The Run-2 pp data delivered by the LHC in 2015 and 2016 at $\sqrt{s} = 13 \text{ TeV}$, corresponding to an integrated luminosity of $\approx 36 \text{ fb}^{-1}$, is used. The analysis has been performed in three channels, for events containing exactly zero, one and two charged leptons, which targets the following Higgs production and decay channels: $ZH \rightarrow \nu\nu b\bar{b}$, $WH \rightarrow \ell\nu b\bar{b}$ and $ZH \rightarrow \ell\ell b\bar{b}$, where ℓ indicates an electron or a muon. In the final state of the three channels, one or two leptons and/or missing transverse energy, and two b -jets, are observed. A multivariate analysis (MVA) is performed to separate the Higgs boson signal from the large $V+jets$ and top backgrounds, where the invariant dijet mass (m_{bb}) is the most discriminating variable. To enhance the sensitivity to signal, the selected events are subdivided in different categories according to the number of jets reconstructed in the event, and p_T^V categories. A binned likelihood fit is then performed to the MVA discriminants to extract the signal. A combination with the ATLAS Run-1 results is also performed, since the Run-1 statistical power is still significant. As a cross-check of the MVA analysis, the signal strength is measured with a fit to m_{bb} , and compatibility between the two analyses is checked. Additionally, the signal strength of the Standard Model $VZ \rightarrow b\bar{b}$ process is measured using an analogous MVA analysis trained for this process, and acts as an important cross-check to validate the background model and MVA analysis technique in VH search.

The note is organised as follows. The pp collision data and Monte Carlo (MC) simulated samples used in the analysis are described in Section 2. A brief description of the object identification and event selection is given in Section 3, while a more detailed description can be found in the *objects* supporting note [15]. An overview of the signal and background modelling is provided in Section 4, and a more detailed description can be found in the *modelling* supporting note [16]. Section 5 describes the multivariate analysis, and the final discriminant used in the nominal analysis. A brief description of the experimental systematic

249 uncertainties is provided in Section 6, whilst more details can be found in objects supporting note [15].
250 Section 7 describes the fit used to extract the SM Higgs boson signal. The main results obtained using the
251 multivariate discriminant, and several cross-checks and $VZ(\rightarrow b\bar{b})$ search that confirm observed results,
252 are summarised in Section 8. In Section 9 the main results combined with Run1 data are discussed.

[Not reviewed, for internal circulation only]

2. Data and Simulated Samples

This section presents an overview of the simulated and recorded data used in this analysis. More detailed information on the generation and studies of the simulated data are included in Ref. [16].

The proton-proton collision data used in this analysis was collected by the ATLAS detector during the 2015 and 2016 runs of the LHC, and corresponds to $\sim 36 \text{ fb}^{-1}$ at $\sqrt{s} = 13 \text{ TeV}$. Events have to pass a filter requirement given by the Good Run List (GRL) in order to ensure their quality and that all systems of the ATLAS detector were operating well when events were recorded.

Simulated $VH \rightarrow Vb\bar{b}$ quark induced signal samples were generated using Powheg MiNLO + Pythia 8 applying the AZNLO tune with NNPDF3 parton distribution functions (PDF). Gluon induced ZH signal samples were simulated using Powheg matrix element generator interfaced with Pythia 8 applying AZNLO tune with NNPDF3 PDFs for the $llb\bar{b}$ channel, while the CT10 PDFs are used for the $vvbb$ samples due to a generator bug in the new samples. All samples are normalised to the best available theoretical prediction for the cross section, for each process where the SM Higgs boson mass is fixed to 125 GeV; the $b\bar{b}$ branching fraction is fixed to 58%. WH signal samples are normalised to the production cross section at NNLO (QCD) and NLO (EW). The inclusive cross section of ZH production is calculated at NNLO (QCD) and NLO (EW), the cross section of gluon induced ZH production is then calculated at NLO (QCD), and quark induced production is taken as the difference of the two in order to avoid double counting [17]. A summary of the included signal samples is presented in Table 1.

Process	Generator	$\sigma \times \text{BR} [\text{fb}]$	N_{events}
$qq \rightarrow ZH \rightarrow vvbb$	Powheg MiNLO + Pythia 8 (NNPDF3)	89.08	2 M
$qq \rightarrow WH \rightarrow l^+vb\bar{b}$	Powheg MiNLO + Pythia 8 (NNPDF3)	164.6	1 M
$qq \rightarrow WH \rightarrow l^-vb\bar{b}$	Powheg MiNLO + Pythia 8 (NNPDF3)	104.5	2 M
$qq \rightarrow ZH \rightarrow llb\bar{b}$	Powheg MiNLO + Pythia 8 (NNPDF3)	44.84	3 M
$gg \rightarrow ZH \rightarrow vvbb$	Powheg + Pythia 8 (NNPDF3)	4.761	0.38 M
$gg \rightarrow ZH \rightarrow l^-l^+b\bar{b}$	Powheg + Pythia 8 (NNPDF3)	7.228	0.7 M

Table 1: Monte Carlo samples used for the signal processes and the cross section and branching ratio (BR) used to normalise the different processes at $\sqrt{s} = 13 \text{ TeV}$. Branching ratios correspond to the decays shown. l is inclusive of e, μ, τ leptons.

Monte Carlo samples are used to simulate the expected shape and normalisation of most SM background processes. One exception is the QCD multi-jet contribution in the 1-lepton channel which, due to the difficulty in modelling and limited statistics, is estimated using data driven methods, as discussed in Section 4.4.

The production of W and Z bosons is simulated with SHERPA 2.2.1 using the NNPDF3.0 NNLO PDFs with dedicated parton shower tuning developed by the SHERPA authors. In order to generate sufficient high $V(p_T)$ statistics, $V+\text{jets}$ samples are split according to $\max(H_T, p_T^V)$. Additionally, to obtain sufficient heavy-flavour final state statistics, the $V+\text{jets}$ samples are generated applying filters, summarised in Table 2. Samples are normalised using cross sections calculated at NNLO accuracy [18]. The $W + \text{jets}$ and $Z + \text{jets}$ simulated background samples are decomposed according to the true flavour of the dijet pair used to reconstruct the Higgs candidate, leading to the following twelve sub-samples:

- Zbb and Wbb : the two jets are labelled as b-jet;

- Zcc and Wcc : the two jets are labelled as c -jet;
- Zl and Wl : the two jets are labelled as light-jet;
- Zbc and Wbc : one of the two jets is labelled as b -jet and the others as c -jet;
- Zbl and Wbl : one of the two jets is labelled as b -jet and the others as light-jet;
- Zcl and Wcl : one of the two jets is labelled as c -jet and the others as light-jet.

The scheme used to define the jet flavour, described in Section 6 of Ref. [15], is based on a ΔR match between truth level b -hadrons, c -hadrons, τ -leptons and reconstructed jets.

Filter	Description
BFilter	at least 1 b -hadron with $p_T > 0$ GeV and $ \eta < 4$
CFilterBVeto(Z)	at least 1 c -hadron with $p_T > 4$ GeV and $ \eta < 3$ veto events which pass the BFilter
CFilterBVeto(W)	at least 1 c -hadron with $p_T > 4$ GeV and $ \eta < 4$ matched by $\Delta R < 0.5$ to a truth jet (<i>AntiKt4TruthJets</i>) with $p_T > 15$ GeV and $ \eta < 3$ veto events which pass the BFilter
CVetoBVeto	veto events which pass the BFilter or the CFilterBVeto

Table 2: Heavy flavour filters used for $V + \text{jets}$, along with a simple description of their application.

- Top-quark pair production is simulated using PowHEG with the PowHEG-BOX framework using the NNPDF 3.0 PDFs and interfaced with PYTHIA 8 using NNPDF 2.3 PDFs for parton showering, with A14 tune. The $t\bar{t}$ samples generated include a filter to require that at least one of the W bosons decays leptonically. Samples are normalised using cross sections calculated at NNLO+NNLL [19].
- Single top quark production (t , s and Wt channels) is simulated using PowHEG with CT10 PDFs interfaced with PYTHIA 6 using CTEQ6L1 PDFs for parton showering. Samples are normalised using cross sections calculated at NLO [20, 21].
- Diboson samples are generated using SHERPA 2.2.1 interfaced with NNPDF 3.0 NNLO PDFs and normalised to cross sections calculated at NLO [22].
- A summary of the included background samples is presented in Table 3. In this analysis, mc15c is used for all MC samples.
- All simulated samples include the effect of pile-up from multiple interactions in the same and neighbouring bunch crossing. This is achieved by overlaying minimum bias events, simulated using PYTHIA8, using A2 tune and interfaced with MSTW2008LO PDFs. All samples are processed through the Geant-4 based ATLAS detector simulation, and reconstructed with the standard ATLAS reconstruction software.

Process	Generator	$\sigma \times BR$ [pb]	N_{events}
Vector boson + jets			
$Z \rightarrow vv$	SHERPA 2.2.1	1914	146M
$W \rightarrow \ell\nu$	SHERPA 2.2.1	20080	410M
$Z\gamma^* \rightarrow \ell\ell$	SHERPA 2.2.1	2107 ($m_{\ell\ell} > 40$ GeV)	171M
Top-quark			
$t\bar{t}$	POWHEG + PYTHIA 8	831.76	60M
s - channel	POWHEG + PYTHIA 6	3.31	2M
t - channel	POWHEG + PYTHIA 6	66.51	10M
Wt - channel	POWHEG + PYTHIA 6	68.00	10M
Diboson			
WW	SHERPA 2.2.1	49.74	14.4M
WZ	SHERPA 2.2.1	21.65	28.6M
ZZ	SHERPA 2.2.1	6.53	10.8M

Table 3: Monte Carlo samples used for the background processes and the cross section times branching ratio (BR) used to normalise the different processes at $\sqrt{13}$ TeV. Branching ratios correspond to the decays shown. ℓ is inclusive of e, μ, τ leptons.

3. Event Reconstruction and Selection

In this section, the event selection applied in the 13 TeV data $VH \rightarrow Vb\bar{b}$ analysis is presented. Most of the selection criteria discussed in the following are inherited from the Run-1 analysis [23]; nevertheless, the values of the applied criteria have been revisited and additional ones have been defined due to the different p-p center-of-mass energy and LHC environment.

3.1. Objects identification

The definitions of the objects used in the analysis, i.e. VH-loose leptons, signal and forward jets, calorimeter-based (track-based) missing transverse energy E_T^{miss} ($E_{T,\text{trk}}^{\text{miss}}$) (and related quantities, i.e. E_T^{miss} HT), as well as the strategy for the removal of overlapping objects and the rebuilding of E_T^{miss} , are provided in Ref. [15]. Tables 4 and 5 summarize the electron and muon definitions used, respectively. The selection criteria for signal and forward jets are given in Table 6.

Electron Selection	p_T	η	ID	d_0^{sig}	$ \Delta z_0^{\text{BL}} \sin \theta $	Isolation
<i>VH – loose</i>	>7 GeV	$ \eta < 2.47$	LH Loose + B-layer cut	< 5	< 0.5 mm	LooseTrackOnly
<i>ZH – signal</i>	>27 GeV	$ \eta < 2.47$	LH Loose + B-layer cut	< 5	< 0.5 mm	LooseTrackOnly
<i>WH – signal</i>	>27 GeV	$ \eta < 2.47$	LH Tight	< 5	< 0.5 mm	FixedCutHighPtCaloOnly

Table 4: Electron selection requirements.

Muon Selection	p_T	η	ID	d_0^{sig}	$ \Delta z_0^{\text{BL}} \sin \theta $	Isolation
<i>VH – loose</i>	>7 GeV	$ \eta < 2.7$	Loose quality	< 3	< 0.5 mm	LooseTrackOnly
<i>ZH – signal</i>	>27 GeV	$ \eta < 2.5$	Loose quality	< 3	< 0.5 mm	LooseTrackOnly
<i>WH – signal</i>	>25 GeV	$ \eta < 2.5$	Medium quality	< 3	< 0.5 mm	FixedCutHighPtTrackOnly

Table 5: Muon selection requirements.

Jet Category	Selection Requirements
Forward Jets	jet cleaning $p_T > 30 \text{ GeV}$ $2.5 \leq \eta < 4.5$
Signal Jets	$p_T > 20 \text{ GeV}$ and $ \eta < 2.5$ jet cleaning $JVT \geq 0.59$ if $(p_T < 60 \text{ GeV}$ and $ \eta < 2.4)$

Table 6: `AntiKt4EMTopoJets` selection requirements. The jet cleaning is applied via the JetCleaningTool, that removes events in regions corresponding to hot calorimeter cells.

3.2. Event selection

This section summarizes the event selection in the 0-, 1- and 2-lepton channels.

318 First of all, events are required to pass the event pre-selection and cleaning, i.e. the GRL requirement (data
 319 only), primary vertex, bad-jets rejection, and calorimetric cleaning (data only), as detailed in Section 9.1
 320 of Ref. [15].

321 Then, the events are categorized according to the number of VH-loose leptons, defined in Table 4 and 5:
 322 the 0-lepton channel is composed of events with zero VH-loose leptons, the 1-lepton channel with exactly
 323 one, and the 2-lepton channel with exactly two. In the 0-lepton channel, events are recorded using E_T^{miss}
 324 triggers with online thresholds of 70 GeV for the data recorded in 2015, and of 90 and 110 GeV for the data
 325 recorded in 2016, due to the higher trigger rate, as described in more detail in Section 4.1 of Ref. [15].

326 In the 1-lepton channel, single electron events are required to be triggered by at least one of the lowest
 327 un-prescaled single-electron triggers. Events in the single muon channel are triggered using the same
 328 E_T^{miss} trigger as used in the 0-lepton channel to increase the trigger efficiency for single high p_T muons.
 329 Studies found this choice increased the single muon signal acceptance to $\sim 98\%$ over the use of a single
 330 muon trigger, offering a signal acceptance of only $\sim 79\%$. Furthermore, the combination of the E_T^{miss}
 331 trigger with the lowest un-prescaled muon trigger was found to offer only a small increase in the single
 332 muon signal acceptance, up to $\sim 99\%$, but without any increase in the signal significance. Therefore, the
 333 combination of triggers for the single muon channel is not considered in order to simplify the analysis,
 334 and only the E_T^{miss} trigger is used for the single muon channel [15].

335 In the 2-lepton channel, events are required to be triggered by at least one of the un-prescaled single-
 336 electron or single-muon triggers that are discussed in detail in Section 4.2 of Ref. [15], where one of the
 337 selected leptons must have fired the trigger.

338 In the 0- and 2-lepton channels, the transverse momentum of the Z boson (p_T^Z) is reconstructed as E_T^{miss}
 339 and as the transverse momentum of the two-lepton system, respectively. In the 1-lepton channel, the
 340 transverse momentum of the W boson (p_T^W) is reconstructed as the vectorial sum of the E_T^{miss} and the
 341 lepton transverse momentum. Events are categorized in bins of p_T^V to improve the analysis sensitivity to
 342 the signal. In particular, the regions $[150, \infty]$ GeV are considered in all three channels; in the 2-lepton
 343 channel, where the trigger allows to go to lower p_T^V the region $[75, 150]$ GeV is also considered. In the
 344 1-lepton channel, due to the modelling difficulties related to the estimate of the high multijet background
 345 contamination, the $[0, 150]$ GeV region is not considered. The inclusion of this region is planned for future
 346 updates to the analysis. Due to the slow E_T^{miss} trigger turn-on curve, the $[0, 150]$ GeV region is also not
 347 considered for the 0-lepton channel.

348 3.2.1. Event selection common to 0-, 1- and 2-lepton channels

349 In all three leptonic channels, events are required to contain at least two signal jets, defined in Table 6.
 350 Exclusive categories of events, depending on the number of selected jets they contain, are defined in
 351 order to maximize the signal significance: events containing two jets comprise the 2-jet category, three
 352 jets the 3-jet category, three or more jets 3+-jets category. In the 0- and 1-lepton channels, the 2- and
 353 3-jet categories are used, and events with four or more jets are rejected to reduce the high $t\bar{t}$ background
 354 contamination. In the 2-lepton channel, where the high jet multiplicity regions result in a 5-6% gain in
 355 expected signal significance, the 2-jet and 3+-jet categories are used.

356 In all three channels, b-tagging is applied to all signal jets selected using the MV2c10 algorithm at the 70 %
 357 efficiency working point [24]. The b-tagging strategy, and efficiency working point have been optimized
 358 to maximize the expected signal significance. Events are categorized according to the number of b-tagged

359 signal jets. Events with no b-tagged jets compose the 0-tag category, with exactly one b-tagged jet the
 360 1-tag category, with exactly two b-tagged jets the 2-tag category. Events with three or more b-tagged jets
 361 are rejected. Only the 2-tag region is considered in this analysis, as this is the region that has the largest
 362 signal sensitivity. The leading b-tagged jet in the 2-tag category is required to have $p_T > 45$ GeV.

363 The calibration of b-tagged jets is improved by applying the *muon-in-jet* and *PtReco* corrections described
 364 in Section 6.3.3. and 6.3.4. of Ref. [15]. These corrections are only applied to the jet derived variables
 365 after the events have passed the analysis selection (but before being passed to the MVA training), and
 366 are not propagated to the E_T^{miss} variables, following what was done for the Run-1 analysis. In 2 lepton,
 367 $ZH \rightarrow \ell\ell b\bar{b}$ can be fully reconstructed. 2 leptons from Z boson have better momentum/energy resolution
 368 than the resolution of jets. b -jet energy can be adjusted by using the balance of transverse momentum.
 369 The kinematic fit based on Run-1 [25, 26] has been developed for Run-2. Transfer functions were made
 370 in each case of muon inside jet and no-muon inside jet, separately. The m_{bb} mass resolution is improved
 371 by 20-30% with respect to the muon-in-jet correction mass resolution. The kinematic fit correction is
 372 applied to 2jet and 3jet events because the improvement is smeared out by the additional jets in the events
 373 with more jets.

374 The mass of the Higgs candidate, m_h is reconstructed using the two selected signal jets, referred to as
 375 *selected jets*, which correspond to the two b-tagged jets in the 2-tag category.

376 In the following paragraphs, a description of the channel-specific selections is provided. Table 8 summarizes
 377 the selection applied in each of the three channels.

378 3.2.2. 0-lepton channel specific selection

379 In the 0-lepton channel, a specific selection is defined to isolate events also containing a Z decaying to a
 380 pair of neutrinos, in addition to the Higgs boson selection. For this purpose, E_T^{miss} is required to be larger
 381 than 150 GeV. The triggers used are: HLT_xe70 for 2015 data, HLT_xe90_mht for 2016 data periods A-D3,
 382 and HLT_xe110_mht for 2016 data periods D4-. The trigger efficiency, for events with a reconstructed
 383 $E_T^{\text{miss}} = 150$ GeV, is of about 85% for xe70 and xe110 while about 90% for xe90; the efficiency plateau is
 384 reached at about $E_T^{\text{miss}} = 180$ GeV, as shown in Section 4.1 of Ref. [15]). Trigger scale factors between
 385 0.95 and 1.0 are applied to MC to correct for any residual turn-on effects.

386 Further requirements are applied on the scalar sum of the p_T (S_T) of up to three jets, referred to as
 387 *pre-selected jets*: the two or three, if present, leading jets, where at least two are required to be signal jets
 388 (see Table 6), and for the third, priority is given to signal, but also forward jets are allowed. For 2-jet
 389 events, S_T must be larger than 120 GeV and for 3-jet events, S_T must be larger than 150 GeV. This criterion
 390 is designed to remove a region which is mis-modelled in simulation due to a non-trivial dependence of
 391 the trigger efficiency on the jet activity, and is as well inherited from the Run-1 analysis.

392 The multijet background in the 0-lepton channel is estimated using a data-driven method described later.
 393 This background mainly enters due to jet energy mis-measurements. As a result, the fake missing
 394 transverse energy and momentum tend to be aligned with the mis-measured jet. A selection on the
 395 azimuthal difference with respect to the nearest jet is applied to remove these events.

396 The multijet background is thus reduced by imposing the four following requirements, also referred to as
 397 *anti-QCD cuts* in the text:

- 398 • $|\Delta\Phi(E_T^{\text{miss}}, E_{T,\text{trk}}^{\text{miss}})| < 90^\circ$

- 399 • $|\Delta\Phi(jet1, jet2)| < 140^\circ$
 400 • $|\Delta\Phi(\mathbf{E}_T^{\text{miss}}, h)| > 120^\circ$
 401 • $\min[|\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \text{pre-sel. jets})|] > 20^\circ$ for 2 jets, $> 30^\circ$ for 3 jets.

402 Here Φ is the azimuthal angle and $jet1$ and $jet2$ are the two selected jets forming the Higgs candidate
 403 h . $E_{T,\text{trk}}^{\text{miss}}$ is defined as the sum of the transverse momenta of the tracks associated to the primary vertex
 404 of the event and passing a set of quality cuts. In the $\min[\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \text{pre-sel. jets})]$ calculation, only the
 405 three leading signal jets are considered, but when no third signal jet passing the selection requirements
 406 is found, a forward jet is considered instead. The values of the cuts are tuned in such a way that the
 407 remaining fraction of multijet contamination is of the order of 1 % of the total background in the 1- and
 408 2-tag categories, and of the order of 10% of the signal in the 2-tag category in a window around the Higgs
 409 boson mass. For more details on the definition of the selection criteria, and to have an overview of the
 410 distributions before the selection is applied, see Section 4.4.

411 3.2.3. 1-lepton channel specific selection

412 In addition to the selections common to the 0-,1- and 2-lepton channels, additional event selections are
 413 implemented to confine events containing a W decaying to a lepton and a neutrino. In the 1-lepton channel,
 414 events are required to contain exactly 1 WH-signal lepton; events containing additional VH-loose leptons
 415 are vetoed. Due to the increased likelihood for a jet to be reconstructed incorrectly as an electron compared
 416 to a muon, separate anti-multijet selections have been optimised for the electron and muon channels. Due
 417 to the increased multi-jet contribution, only events with $p_T^W > 150$ GeV are considered. In order to further
 418 reduce the multijet contribution, events in the electron channel are required to have $E_T^{\text{miss}} > 30$ GeV.
 419 In the 1-lepton channel, m_{top} is calculated as the minimum invariant mass of the all W and signal jet
 420 combinations. The 4-vector of the W boson is calculated assuming all the E_T^{miss} in the event originates
 421 from the leptonic decay of the W boson, with further details given in Section 5.1.

422 Reduction of QCD Contamination

423

424 In addition to the standard $VH - \text{loose}$ lepton isolation cuts, outlined in Tables 4 and 5, studies have been
 425 carried out on the MC dijet samples listed in Table 15 to test the impact of further isolation cuts which
 426 can reduce the QCD background whilst maintaining a high signal efficiency for $WH - \text{signal}$ leptons. A
 427 number of isolation variables were tested, with the best performance being observed for `topoetcone20`
 428 and `ptcone20` in the electron and muon modes respectively. Figure 3 shows the signal and MJ efficiency
 429 as a function of the cut on respective isolation channels for each lepton mode, with a numerical summary
 430 presented in Table 7.

431 Based on the efficiencies presented in Table 7, the MJ background can be reduced by up to a factor of 5
 432 (4) in the electron (muon) modes, whilst maintaining a VH signal efficiency of greater than 95%.

433 Further details on the estimation of the 1-lepton QCD background using the template method are given in
 434 Section 4.4.2.

435 All selection criteria applied to define the 1-lepton analysis are summarized in Table 8.

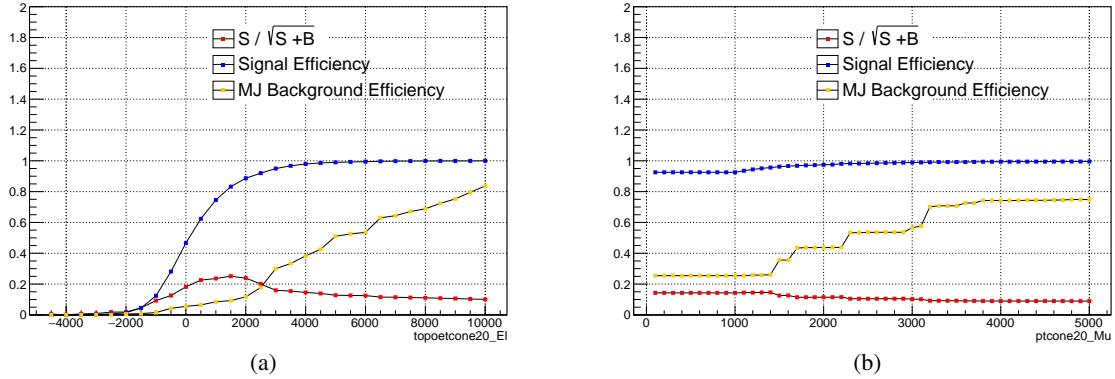


Figure 3: Multijet (yellow) and VH signal (blue) efficiency in the 1-lepton channel as a function of topoetcone20 in the electron mode (a) and as a function of ptcone20 in the muon mode (b).

Lepton Mode	Isolation Variable	Isolation Cut [GeV]	Multijet Efficiency [%]	Signal Efficiency [%]
Electron	<code>topoetcone20</code>	3.50	18	95
Muon	<code>ptcone20</code>	1.25	26	95

Table 7: Impact of lepton isolation cuts on the multijet and signal efficiencies in the electron and muon modes, in addition to the WH -signal lepton ID and p_T selections listed in Tables 4 and 5. Using these cuts, the MJ background can be expected to be reduced by up to a factor of 4 when applied in addition to the default lepton isolation cuts.

436 **3.2.4. 2-lepton channel specific selection**

437 In the 2-lepton channel, a Z decaying to two same flavour leptons ($ee, \mu\mu$) needs to be reconstructed in
 438 addition to the b-tagged jets coming from the Higgs boson decay. It is therefore required that the selected
 439 events have exactly two VH-loose leptons of the same flavour, at least one of which must also satisfy
 440 the ZH -signal requirements, as defined in Section 5.1.1 and 5.2.1 of Ref. [15]. Events with additional
 441 VH-loose leptons are discarded.

442 For selected di-muon events the two muons are further required to be of opposite charge; the requirement
 443 is not applied to di-electron events due to higher rate of charge misidentification. The invariant mass of the
 444 di-lepton system must be consistent with the Z boson mass: $81 < m(ee, \mu\mu) < 101$ GeV. This requirement
 445 suppresses backgrounds having a non-resonant lepton-pair, such as $t\bar{t}$ and multi-jet productions. In 2lepton
 446 channel, multi-jet background is estimated using side-band of dilepton mass. Also same-sign events are
 447 used to extract multi-jet model as discussed in 4.4.3.

448 **3.2.5. Additional selection criteria for the dijet mass analysis cross-check**

449 A cut-based analysis is performed as a cross-check to the main multivariate analysis, where the m_{bb}
 450 variable is used as a discriminant to separate signal from background. A number of additional selection
 451 criteria are applied to the events to further reduce the background contamination. Selections are based on
 452 p_T^V regions, as defined in Table 8.

453 In the 0-lepton channel, a cut on the ΔR between the two selected jets is applied:

- 454 • $|\Delta R(\text{jet1,jet2})| \leq 1.8$ ($150 \leq p_T^V < 200$ GeV), ≤ 1.2 ($200 \leq p_T^V$ GeV).

455 In the 2-lepton channel, the following additional selection criteria are applied:

- 456 • $|\Delta R(\text{jet1,jet2})| \leq 1.8$ ($p_T^V < 200$ GeV), ≤ 1.2 ($200 \leq p_T^V$ GeV);
 457 • $E_T^{\text{miss}}/\sqrt{HT} < 3.5\sqrt{\text{GeV}}$.

458 where HT is defined as the scalar S_T of 2 leptons and all reconstructed jets.

459 In the 1-lepton channel, the following additional selection criteria are applied:

- 460 • $|\Delta R(\text{jet1,jet2})| \leq 1.8$ ($150 \leq p_T^V < 200$ GeV), ≤ 1.2 ($200 \leq p_T^V$ GeV);
 461 • $m_T(W) < 120$ GeV.

462 In all three channels, for the dijet cross-check only, events with p_T^V between 150 GeV and 200 GeV, and
 463 events with p_T^V above 200 GeV, are separated in order to increase the sensitivity of the measurement. In
 464 2 lepton channel, the events with $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$ are also used.

Not reviewed, for internal circulation only

Common Selections	
Jets	≥ 2 central jets
b -jets	2 b -tagged signal jets
Leading jet p_T	> 45 GeV
$ \Delta R(\text{jet1}, \text{jet2}) $ (cut-based only)	≤ 1.8 ($p_T^V < 200$ GeV), ≤ 1.2 ($p_T^V > 200$ GeV)
0 Lepton	
Trigger	HLT_xe70, xe90_mht, and xe110_mht
Leptons	0 VH-loose lepton
E_T^{miss}	> 150 GeV
S_T	> 120 (2 jets), > 150 GeV (3 jets)
$ \min\Delta\phi(E_T^{\text{miss}}, \text{jet}) $	$> 20^\circ$ (2jet), $> 30^\circ$ (3jet)
$ \Delta\phi(E_T^{\text{miss}}, h) $	$> 120^\circ$
$ \Delta\phi(\text{jet1}, \text{jet2}) $	$< 140^\circ$
$ \Delta\phi(E_T^{\text{miss}}, E_{T, \text{trk}}^{\text{miss}}) $	$< 90^\circ$
p_T^V regions (BDT)	> 150 GeV
p_T^V regions (cut-based)	[150, 200] GeV, [200, ∞] GeV
1 Lepton	
Trigger	e channel: un-prescaled single electron Tables 5 and 6 of Ref. [15] μ channel: see 0-lepton triggers
Leptons	1 WH-signal lepton > 1 VH-loose lepton veto
E_T^{miss}	> 30 GeV (e channel)
m_{top}	< 225 GeV or $m_{bb} > 75$ GeV
m_T^W (cut-based only)	< 120 GeV
p_T^V regions (BDT)	> 150 GeV
p_T^V regions (cut-based)	[150, 200] GeV, [200, ∞] GeV
2 Lepton	
Trigger	un-prescaled single lepton Tables 5 and 6 of Ref. [15]
Leptons	2 VH-loose leptons (≥ 1 ZH-signal lepton) Same flavor, opposite-charge for $\mu\mu$
$m_{\ell\ell}$	$81 < m_{\ell\ell} < 101$ GeV
E_T^{miss} significance (cut-based)	$E_T^{\text{miss}} / \sqrt{HT} < 3.5\sqrt{\text{GeV}}$
p_T^V regions (BDT)	[75, 150], [150, ∞] GeV
p_T^V regions (cut-based)	[75, 150], [150, 200], [200, ∞] GeV

Table 8: Summary of the signal event selection in the 0-, 1- and 2-lepton analyses.

465 4. Signal and Background Modeling

466 The signal and background processes considered in the $VH \rightarrow b\bar{b}$ search have been modelled using several
 467 MC generators and control regions. In Section 4.1, the VH signals and the background processes modelled
 468 using MC are documented. In Section 4.2, details of the truth tagging procedure to boost MC statistics
 469 in b -tagged regions is discussed. The multijet and non-collisional backgrounds have been studied using
 470 both MC and data driven techniques, as described in Section 4.4. In Section 4.5, relevant distributions
 471 showing the data-MC pre-fit comparisons are provided.

472 4.1. MC Signal and Backgrounds

473 Quark and gluon induced $VH \rightarrow Vb\bar{b}$ signal processes are modelled in this analysis using a combination
 474 of PYTHIA 8 and Powheg + PYTHIA 8 MC samples, as described in Table 1. The dominant backgrounds
 475 that mimic such signal after the full selection is applied are: V +jets, top pair and single top production,
 476 and dibosons. The shapes for such processes are modelled using the MC samples listed in Table 3. The
 477 modelling and systematic uncertainties of signal and background MC samples are discussed in detail in
 478 Ref. [16].

479 For the $t\bar{t}$ background, a systematic uncertainty is applied to constrain the normalisation in the three (or
 480 plus) jets regions with respect to the one measured in the two jet regions. A shape systematic uncertainty is
 481 also applied, for the same background, to the m_{bb} distribution. Such systematic uncertainties are derived
 482 from MC generator comparisons at particle level, as described in Section 4.2.4 of Ref. [16].

483 For the single top background samples, normalisation systematic uncertainties are estimated from the
 484 theoretical scale, pdf, and α_s uncertainties provided by the LHC top group, as described in Section 4.3.3
 485 of Ref. [16]. Shape systematic uncertainties are applied on the m_{bb} distribution, derived from MC
 486 generator comparisons performed at particle level, as described in Section 4.3.3 of Ref. [16] (the SM VH
 487 specific sub-section).

488 For the V +jets samples, normalisation systematic uncertainties are evaluated at truth level using Sherpa
 489 2.2 samples. Shape systematic uncertainties are extracted, for the m_{bb} and p_T^V distributions, from a fit to
 490 data in V +jets dominated control regions. A description of the method used to extract normalisation and
 491 shape systematics, and the results obtained, can be found in Section 3.1.6 and 3.1.7 of Ref. [16].

492 For the diboson samples, acceptance systematic uncertainties and shape systematic uncertainties for the
 493 m_{bb} and p_T^V distributions are derived from MC generator comparisons performed at particle level, as
 494 described in Section 4.4 of Ref. [16].

495 An overview of the normalisation and shape systematic uncertainties, for all samples and regions con-
 496 sidered, is provided in Section 7.3, where the fit model is described.

497 4.2. Truth tagging

498 The tag rate function, or “truth-tagging” procedure, is used, for all samples and in all regions considered,
 499 for the MVA training in order to reduce effects of overtraining due to limited MC statistics after the two
 500 b -tagged jets requirement is applied. Truth-tagging is also used to improve the data-MC comparison plots
 501 (in the 2-tag regions only), and the final fit, for the $V + cc$, $V + c\ell$, $V + \ell$, and WW samples only.

502 When using truth-tagging, all events pass the 2-tag requirement by construction. A combination of two
 503 jets in the event are randomly selected to be “tagged”. The probability for a jet to be tagged is directly
 504 proportional to its b-tagging efficiency, which is a function of the jet’s “real” flavour in MC, p_T , and η .
 505 For a given tagging combination, a partial “truth-tagging” weight may be defined as the product of the
 506 b-tagging efficiencies of the two tagged jets times the product of one minus the efficiency of all untagged
 507 jets. The total truth-tagging event weight is taken to be the sum over all possible combinations, and the
 508 probability for selecting a given combination is directly proportional to its partial truth-tagging weight.

509 For example, in an event with three jets, labeling the efficiency of the i^{th} jet as ϵ_i , the total truth-tagging
 510 weight of the event is

$$\epsilon_{tot} = \epsilon_1 \epsilon_2 (1 - \epsilon_3) + \epsilon_1 (1 - \epsilon_2) \epsilon_3 + (1 - \epsilon_1) \epsilon_2 \epsilon_3, \quad (1)$$

511 and the probability of selecting jets 1 and 2 as the tagged jets is

$$\frac{\epsilon_1 \epsilon_2 (1 - \epsilon_3)}{\epsilon_{tot}}. \quad (2)$$

512 Since the b-tagging efficiency is much higher for b -quark-like jets than for c and light quark-like jets, the
 513 truth-tag event weight will preferentially suppress events with c - and light-like jets with respect to events
 514 with b -quark like jets to preserve the expected overall normalisation.

515 Under current settings, jets that are tagged are assigned a random b-tag score that is sampled from the
 516 MV2c10 cumulative distribution above the 70% efficiency working point cuts. All other jets in the event
 517 are assigned a random b-tag score below the 70% working point cut.

518 Further studies on the closure of truth tagging are presented in Appendix G.

519 4.3. Background Control Regions

520 To help better constrain the modelling of background processes, a number of control regions have been
 521 defined. These control regions are designed to be highly pure in one background process, and are defined
 522 using a series of selection cuts. These control regions are then orthogonal to the signal region, with
 523 negligible level of signal contamination.

524 4.3.1. 1-lepton

525 **$W + hf$ Control Region**

526 As an additional cross check of the $W + hf$ modelling, a control region has been defined to be highly
 527 pure in $W + hf$, with little signal contamination. This region is cut from the signal region, defined using
 528 a cut on the reconstructed top mass, m_{top} , as described in Section 5.1, and m_{bb} . This information is
 529 summarised in Table 9. The cut on m_{bb} is introduced to reduce VH signal contamination, whilst cutting
 530 on m_{top} greatly reduces the $t\bar{t}$ and single- t background contamination. Events passing these cuts are
 531 placed into the control region, otherwise they remain in the signal region, such that the two regions are
 532 fully orthogonal. The implementation of this additional region in the 1-lepton channel as a single bin has
 533 been shown to better constrain the $W + hf$ background normalisation in the 1-lepton fit.

Variable	Cut
m_{top}	> 225 GeV
m_{bb}	< 75 GeV

Table 9: Selection requirements for 1-lepton channel $W + hf$ control region. Events in this region are removed from the signal region, such that the two regions are fully orthogonal.

535 **4.3.2. 2-lepton**

536 **$e\mu$ Control Region for flavor symmetric background**

537

538 One of dominant backgrounds, $t\bar{t}$, is flavor symmetric. High purity $t\bar{t}$ control region can be obtained by
 539 requiring different flavor of a pair of dilepton ($e\mu$ or μe). Lepton flavor does not change the kinematics of
 540 $t\bar{t}$ between the SR (same flavor) and the CR(different flavor). $t\bar{t}$ kinematic distributions using the control
 541 region are estimated by selecting same kinematic selection as signal region.

$$N_{top}^{SR} = \frac{N_{top}^{CR,data}}{N_{top}^{CR,MC}} \times N_{top}^{SR,MC} = \frac{N_{top}^{SR,MC}}{N_{top}^{CR,MC}} \times N_{top}^{CR,data} \quad (3)$$

542 As described in 3, top background modeling in the signal region can be constrained in $e\mu$ control region.
 543 In the current fit scheme, four top control regions ($2p_T^V$ categories for 2jets and 3+jets) are set to use same
 544 phase space for each signal region. Thanks to two b-tagging requirement, very pure control regions are
 545 obtained. More than 99% in the control region is dominated by top ($t\bar{t}$ and Wt). The prefit yields and
 546 kinematic distributions of control region are shown in Table 26-27 and Figure 24-25.

547 Figures 4-7 show the shape comparison plots for several kinematic shapes in inclusive p_T^V region. Shape
 548 of $e\mu$ control region is same as that of same flavor signal region. Even BDT output shape, $e\mu$ control
 549 region provides good modeling of signal region. Therefore shape and normalization can be constrained by
 550 the control region without specific extrapolation factor and its uncertainty. To constrain top background
 551 yield and shape, m_{bb} shape which is one of most sensitive discriminants in this analysis is used. In the
 552 top MC shape, all experimental and theoretical systematic uncertainties are assigned as prior. potential
 553 mismodeling of yield and shape is constrained by fitting scheme and adequately propagated to top
 554 background in the signal region. The binning of control region has been optimized for each category since
 555 statistics is limited in high p_T^V category.

556 From rightmost of Eq.3, data-driven way using control region data is also under investigation. Detailed
 557 studies are in Appendix. C.

Not reviewed, for internal circulation only

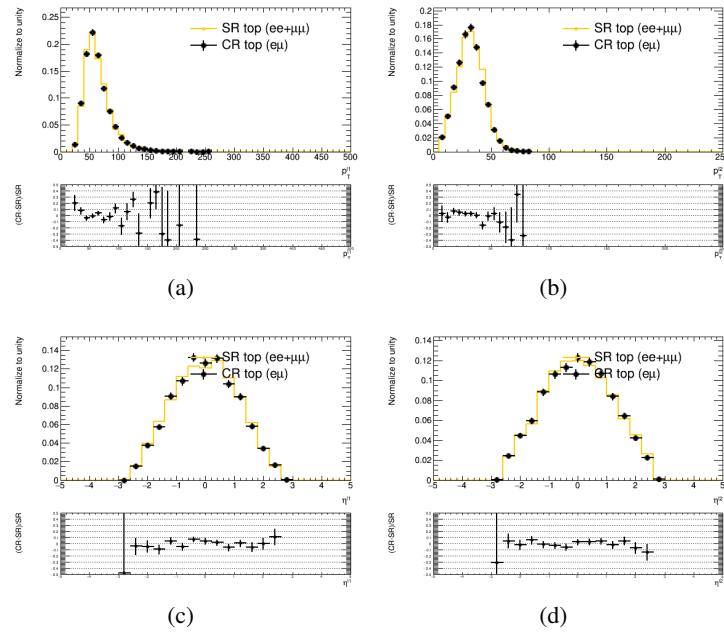


Figure 4: Shape comparison between same flavor signal region and different flavor control region in 2jet low- p_T region. The normalization for both shapes are scaled to unity.

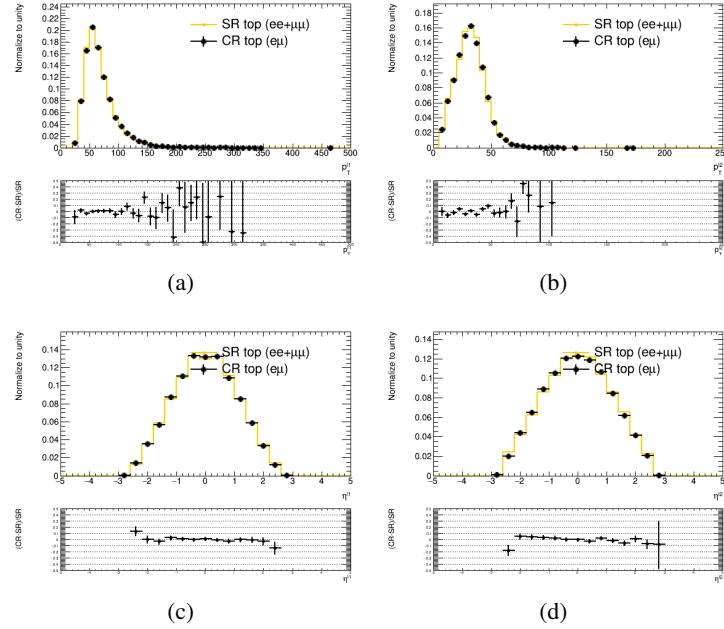


Figure 5: Shape comparison between same flavor signal region and different flavor control region in 2jet low- p_T region. The normalization for both shapes are scaled to unity.

Not reviewed, for internal circulation only

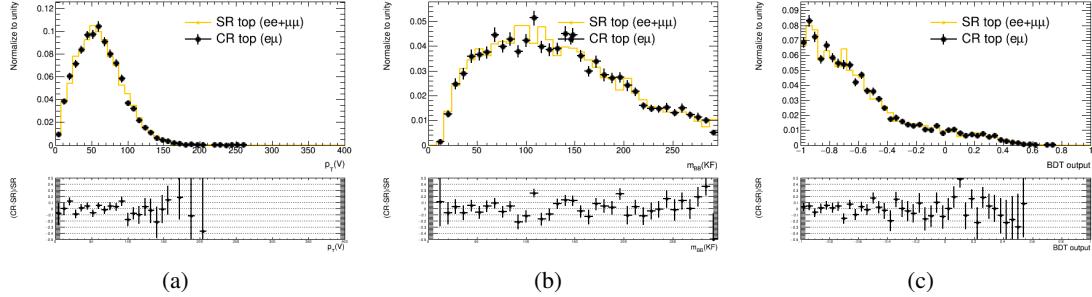


Figure 6: Shape comparison between same flavor signal region and different flavor control region in 2jet low- p_T region. The normalization for both shapes are scaled to unity.

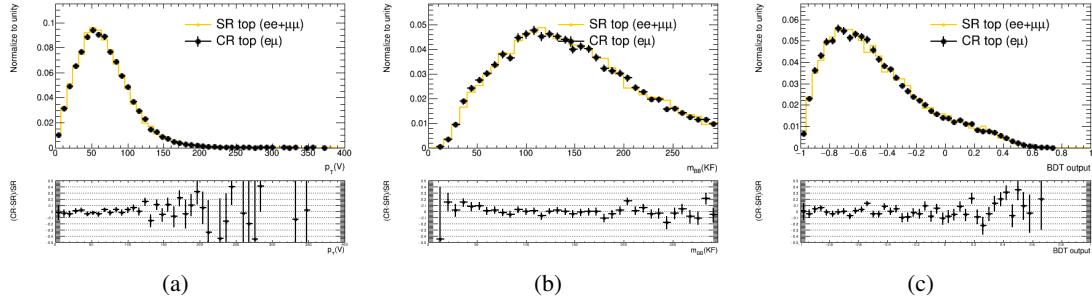


Figure 7: Shape comparison between same flavor signal region and different flavor control region in 3+jet low- p_T region. The normalization for both shapes are scaled to unity.

558 4.4. Multijet and non-collisional backgrounds

559 Multijet backgrounds arise from QCD production of jets, subsequently faking leptons or E_T^{miss} , or decaying
 560 to non-prompt leptons. Due to difficulties encountered to model this background, data driven approaches
 561 are used. This section describes the different methods used in each channel and provides the estimates
 562 used in the analysis.

563 4.4.1. 0-lepton

564 In the 0-lepton channel, as described in Section 3.2, anti-QCD selections are applied to suppress multijet
 565 and non-collisional background based on the azimuthal angles between the E_T^{miss} , p_T^{miss} , and the selected
 566 jets. These selections are inherited from the Run 1 effort, and have been re-optimized using the Run 2
 567 data and MC samples. In appendix A the data distributions in these variables are studied including also
 568 the dijet MC samples. From the MC dijet samples it is not possible to obtain a reliable estimate of the
 569 multijet contamination in the signal categories, due to limited statistics after all the selections only a few
 570 events remain giving large uncertainties in the yields. A data driven estimation is described below.

571 In order to estimate the remaining multi-jet contribution in the signal categories the $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$
 572 distribution is used after removing the selection on this variable. The pre-fit plots for this variable are
 573 shown in Figures 8. The MJ contribution is observed at low $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ and its yield is then
 574 extracted by a fit described below. From the shape of the multi-jet background predicted by the MC
 575 samples we expect an exponentially falling distribution in $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$, this is shown in Figure 9
 576 left plot. The pre-fit plots show there is a remaining multijet contribution in the 3jet category at low
 577 $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$. We estimate the multi-jet yield by performing a fit to the $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ as shown
 578 in Figure 9 (right) using an exponential model for the multijet motivated by the dijet MC with relaxed
 579 selections. In this fit the Z, W, and ttbar backgrounds estimated from MC are floated while the diboson
 580 and single-top are fixed to the MC prediction. From this fit the multijet yield in the signal region is estimated
 581 for different selections on $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ as tabulated in Table 10. A selection of 30 deg is chosen to
 582 give only about 10% contamination in the m(bb) signal region relative to the expected VH signal yield.

583 Furthermore, the shape of the multijet background is studied by selecting the events at low $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$,
 584 templates are extracted after subtracting the expected electro-weak and top backgrounds. The shapes are
 585 compared to the total Z+jets, W+jets, and $t\bar{t}$ background for both the m(bb) and BDT distributions in
 586 Figure 10. We find a similar shape between the multijet and those backgrounds showing that the remaining
 587 multijet contamination can be absorbed by these backgrounds in the nominal fit.

588 A similar approach cannot be applied to the 2-jet category due to the correlation with the other anti-QCD
 589 selections which remove all events at low $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$. However, the multijet shape is shown to
 590 have the same exponential behavior as in the 3-jet events as shown in Figure 9 left, from the events shown
 591 in Figure 8 it is then clear that the MJ yield is smaller for the 2jet category than in the 3-jet category while
 592 the VH signal yield is similar in both categories.

Table 10: Comparison of the multi-jet and VH yields for the 0 lepton 3jet category determined from the fit to the $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ distribution. The second and third columns show the yields for the given selection, the fourth column gives the ratio, and the last column shows the ratio after scaling each yield by the fraction determined from the $m(\text{bb})$ distributions in the range 80 to 160 GeV. A statistical uncertainty on the multijet normalization of 5% is determined from the fit described in the text, while a 15% systematic uncertainty is estimated by varying the fit range at low $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$.

Cut	VH	Multijet	Multijet/VH	Multijet/VH
				$80 \text{ GeV} < m(\text{bb}) < 160 \text{ GeV}$
20 deg	56.40	65.44	1.16	0.44
25 deg	55.26	33.23	0.60	0.23
30 deg	54.12	16.87	0.31	0.12
35 deg	52.94	8.57	0.16	0.06
40 deg	51.89	4.35	0.08	0.03

Table 11: Scale factors for the backgrounds from the fit for the multi-jet in $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ distribution in the 3jet category.

background	scale factor
Z+jets	1.2 ± 0.3
W+jets	1.32 ± 0.6
$t\bar{t}$	1.2 ± 0.3
VV, stop	1. (fixed)

Not reviewed, for internal circulation only

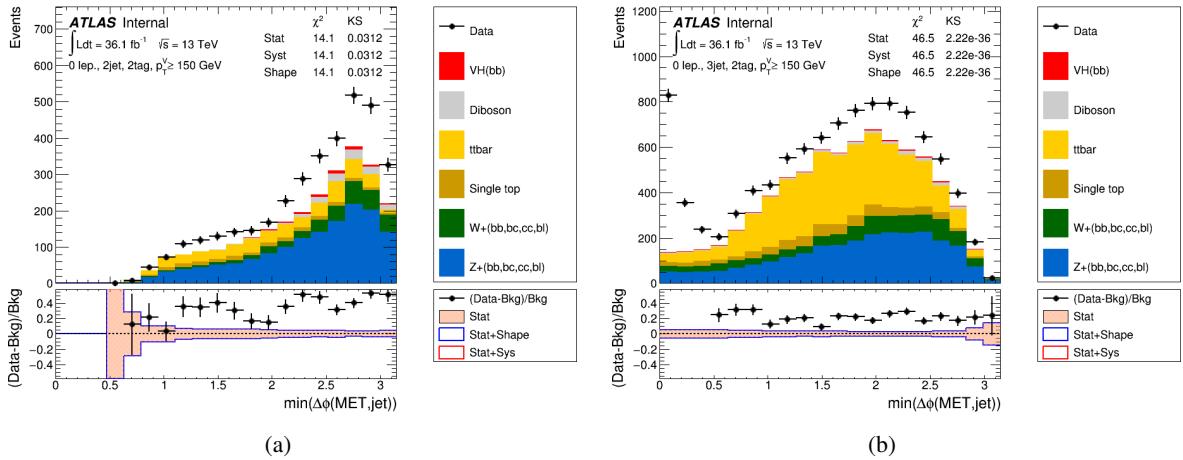


Figure 8: Pre-fit $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ distributions for the 2jet (left) and 3jet (right) categories excluding the selection on this variable.

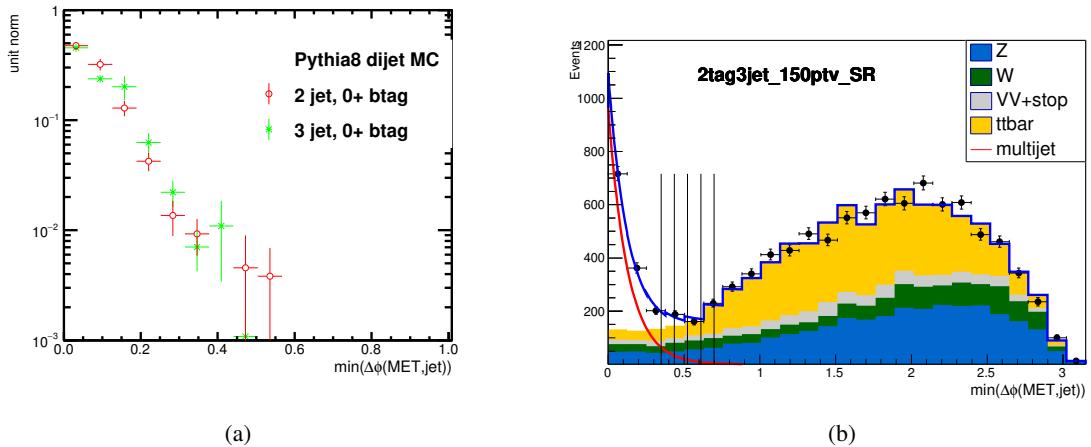


Figure 9: Left: shape of the multi-jet distribution in $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ as predicted by the di-jet MC samples with relaxed anti-QCD selections. Right: fit to the $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ distribution in the 3jet category. The multi-jet modeled using an exponential shape e^{-cx} and floated in the fit. The vertical lines are different selections tested to improve the multi-jet contamination, starting from 20 deg in 5 deg steps.

Not reviewed, for internal circulation only

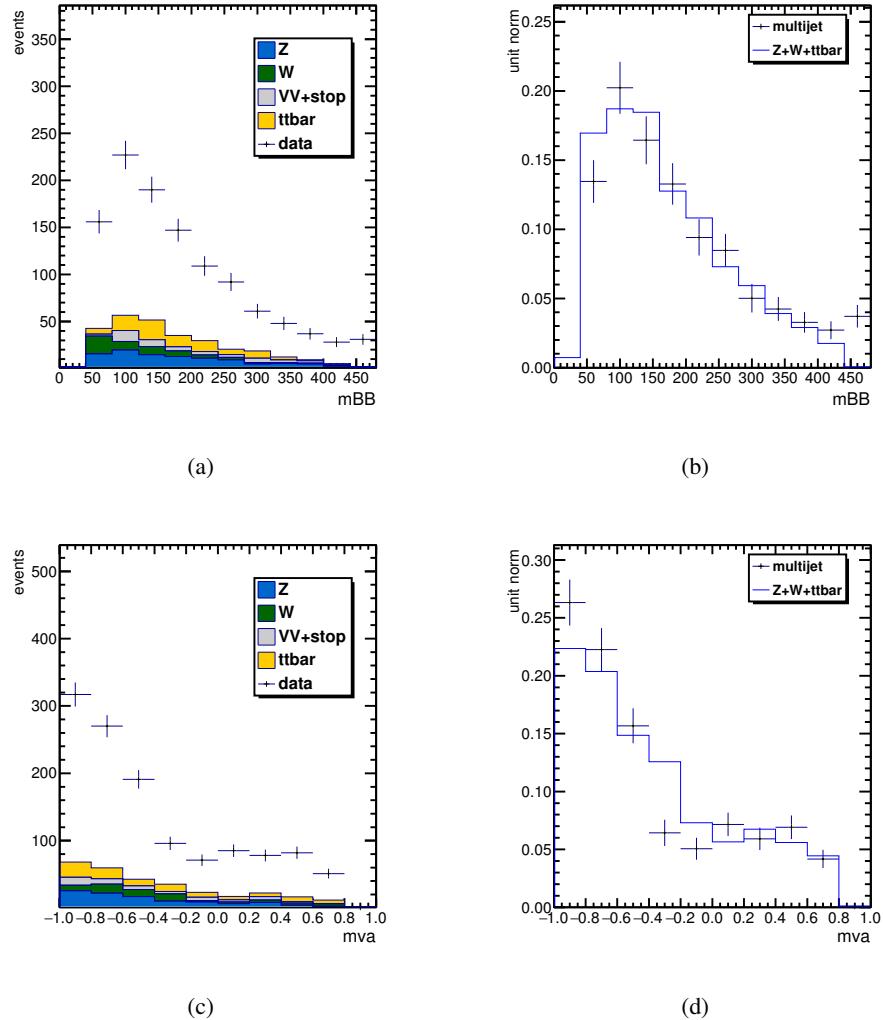


Figure 10: Comparisons of the multi-jet shapes in the m_{bb} (top) and BDT (bottom) distributions to the shape of the combined W , Z , and $t\bar{t}$ backgrounds. The multijet templates are determined by selecting data below 20 deg in $|\min\Delta\phi(E_T^{\text{miss}}, \text{jet})|$ (left) and subtracting the MC backgrounds. The templates are compared to the combined MC backgrounds in the right plots after normalizing to unity.

593 **4.4.2. 1-lepton**

594 In the 1-lepton channel, the contribution of the multijet background is greatly reduced by only considering
 595 the high p_T^V region. However, the multijet background still contributes a significant fraction of the
 596 background events in this region, and a robust procedure is necessary to estimate the contributions
 597 separately for the single e and single μ modes. In both modes, the template method is employed.

598 **Template Method**

599

600 The template method aims to determine the multijet contribution of fake leptons from photons and light
 601 jets and from non-prompt heavy flavour decays in the signal region using data in a multijet enriched control
 602 region. The MJ enriched control region is defined using inverted lepton isolation cuts, as summarised in
 603 Table 12. Furthermore, in order to reduce the extrapolation from the inverted isolation region, 1 b -tag is
 604 also required, as limited statistics prevent the direct use of the 2 b -tag region. The MJ background template
 605 is then extracted separately for the e/μ , 2/3-jet regions, using the corresponding inverted isolation region,
 606 taking into account the contamination from other background processes. Data and MC yields from the
 607 1-tag, inverted-isolation signal region are presented in Table 13.

	Isolated Region	Inverted Isolation Region
Electron	<code>IsLooseTrackOnlyIso</code> <code>TopoEtCone20<3.5 GeV</code>	<code>IsLooseTrackOnlyIso</code> <code>TopoEtCone20>3.5 GeV</code>
Muon	<code>IsLooseTrackOnlyIso</code> <code>PtCone20<1.25 GeV</code>	<code>IsLooseTrackOnlyIso</code> <code>PtCone20>1.25 GeV</code>

Table 12: Summary of differences between the isolated and inverted isolation regions used for the template method approximation of the MJ background.

608 In the inverted isolation region, there is a non-negligible contribution from electroweak background
 609 processes, such as $t\bar{t}$ and $W+jets$ processes. Due to uncertainties on the overall normalisations of
 610 these processes, a simultaneous fit of the normalisation of the top ($t\bar{t} + \text{single top}$) $W+jets$ and QCD
 611 backgrounds is conducted in the 2 b -tag, isolated lepton region to extract the normalisation factor of the
 612 QCD background. A simultaneous fit is conducted to the m_T^W distribution in both the e and μ modes,
 613 with the binning optimised such that the MC statistical uncertainty does not exceed 20% in any bin. A
 614 single bin is considered for the $W + hf$ control region for each of the e and μ modes. This choice is found
 615 to result in the best overall modelling of the MJ background, with the normalisation factors presented
 616 in Table 14, and the post-fit distributions presented in Figure 11. These normalisation factors are then
 617 applied to the electroweak background in the inverted isolation region as a systematic uncertainty of the
 618 effect of the normalisation of these processes.

619 The main assumption of the template method is that the template distribution used in the signal region
 620 and extracted in the multi-jet enhanced control region are the same. Since this assumption is not expected
 621 to be perfectly valid, this can introduce a bias in the template, and ultimately in the estimated multi-jet
 622 contribution in the signal region. The templates derived in the control regions are therefore corrected for
 623 such a bias before being used in the signal regions, by applying event-by-event extrapolation factors that
 624 depend on lepton pT and η , and, in the electron channel, also on the value of E_T^{miss} . These extrapolation

Not reviewed, for internal circulation only

Process	2 jet		3 jet	
	1-tag, e mode	1-tag, μ mode	1-tag, e mode	1-tag, μ mode
Zl	4.8 ± 2.2	5.8 ± 2.4	16.4 ± 4.1	8.1 ± 2.9
Zcl	3.4 ± 1.8	4.1 ± 2.0	16.3 ± 4.0	7.6 ± 2.8
Zhf	14.4 ± 3.8	21.8 ± 4.7	28.8 ± 5.4	32.5 ± 5.7
Wl	97.7 ± 9.9	117.8 ± 10.9	86.3 ± 9.3	185.3 ± 13.6
Wcl	251.0 ± 15.8	237.8 ± 15.4	296.4 ± 17.2	350.8 ± 18.7
Whf	270.2 ± 16.4	317.8 ± 17.8	456.8 ± 21.4	516.1 ± 22.7
single-top	155.6 ± 12.5	187.4 ± 13.7	370.9 ± 19.3	496.7 ± 22.3
ttbar	412.2 ± 20.3	508.5 ± 22.5	1683.2 ± 41.0	2141.5 ± 46.3
diboson	14.6 ± 3.8	15.6 ± 4.0	23.1 ± 4.8	29.0 ± 5.4
Total Bkg	1223.9 ± 35.0	1416.7 ± 37.6	2978.3 ± 54.6	3767.7 ± 61.4
Data	7693.0	3037.0	12324.0	6003.0
Data/MC	6.29	2.14	4.14	1.59

Table 13: Pre-fit yields for the 1-lepton analysis for $p_T^W > 150$ GeV, for an integrated luminosity of 36.07 fb^{-1} in the 1-tag, inverted isolation region. The uncertainty shown is of statistical nature only. The Vhf yield count includes events from the $V + bb$, $V + bc$, $V + bl$ and $V + cc$ categories.

Region	Top ($t\bar{t}$ + single top)	$W+\text{jets}$
2-tag, 2-jet	1.01	1.39
2-tag, 3-jet	0.98	1.35

Table 14: Summary of normalisation scale factors for $t\bar{t}$ and $W+\text{jets}$ derived in the isolated lepton region

625 factors are derived in additional control regions where the 2- and 3-jet requirement of the nominal selection
 626 is replaced by a 1-jet requirement, and any b-tagging requirement is removed. The extrapolation factors
 627 are computed as the ratio between events with an isolated and an isolation-reverted lepton, after removing
 628 the EW background contribution based on MC predictions. Further details on this method are presented
 629 in Appendix H of [15].

Generator	DSID	Leading jet p_T range [GeV]	σ [pb]
Sherpa 2.2.1	344715	30-50	283697.214
Sherpa 2.2.1	344716	50-80	67501.441
Sherpa 2.2.1	344717	80-130	13612.655
Sherpa 2.2.1	344718	130-200	1819.401
Sherpa 2.2.1	344719	$200-\sqrt{s}$	296.265

Table 15: MC di-jet samples used to evaluate lepton isolation cuts.

Not reviewed, for internal circulation only

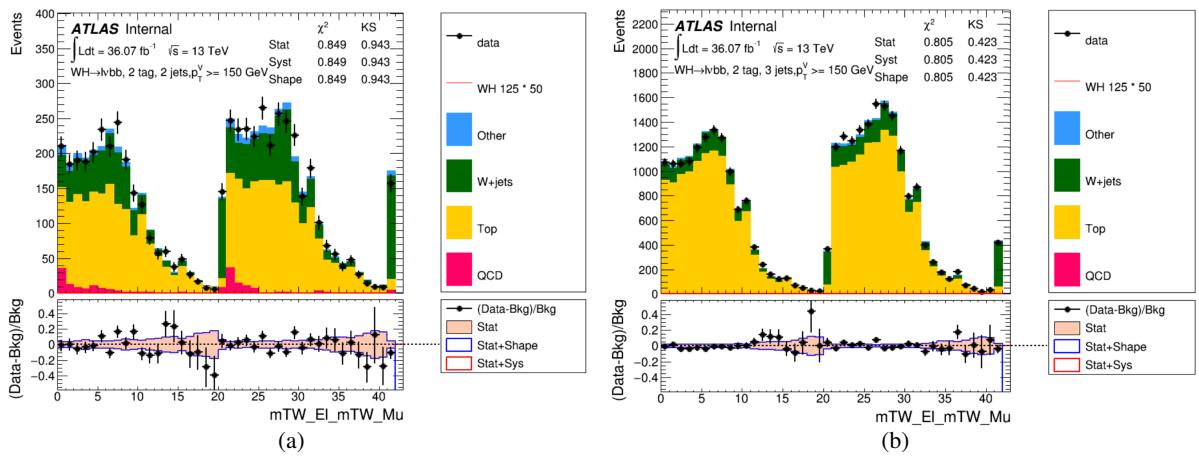


Figure 11: The m_T^W distribution in the isolated 1-lepton $p_T^W > 150$ GeV requiring exactly 2 b -tags with 2 signal jets (a), 3 signal jets (b), after applying $t\bar{t}$ and $W+jets$ normalisation factors. Bins 1-21 correspond to the e only channel, and bins 22 to 42 correspond to the μ only channel. Bin 21 (42) represents the $W + hf$ CR as a single bin in the e (μ) channel.

630 In both the electron and muon modes, a number of sources of systematic uncertainty are considered in the
 631 multijet estimation.

632 Shape Uncertainties

633

634 In order to evaluate the shape uncertainty of the MJ background, a number of shape systematic uncertainties
 635 are considered:

- 636 • **In the e mode**, the impact of the choice of electron trigger on the MJ estimate is evaluated, as this
 637 may introduce a bias in the inverted isolation region. Instead of using the combination of triggers,
 638 listed in Section 5.2.1 of Ref [15], simply the lowest p_T trigger is used. This corresponds to the
 639 trigger selections for each data period listed in Table 16. This uncertainty is correlated between the
 640 2 and 3 jet regions.
- 641 • **In the e and μ modes**, the impact of using the 1-tag region to extract MJ shape is evaluated. This is
 642 evaluated by comparing the shape of the 1-tag MJ template with the 2-tag MJ template, and taking
 643 the difference as a systematic uncertainty. This uncertainty is correlated between the 2 and 3 jet
 644 regions, and decorrelated between e/μ channels.
- 645 • **In the e and μ modes**, an evaluation of the uncertainty introduced by the use of the full inverted
 646 isolation region is considered. Instead, a reduced inverted-isolation region is defined, with the cuts
 647 in addition to the inverted isolation cuts in Table 12. In the e mode, this is defined with additionally
 648 requiring $3.5 < \text{topoetcone20} < 11$ GeV, and in the μ mode, $1.25 < \text{ptcone20} < 2.25$ GeV. This
 649 uncertainty is correlated between the 2 and 3 jet regions, and decorrelated between e/μ channels.
- 650 • **In the e and μ modes**, the impact of not applying the bias correction is evaluated. The MJ template
 651 shape is evaluated without applying the bias corrections, and the difference in shape to the nominal
 652 taken as a systematic uncertainty. This uncertainty is decorrelated between e/μ channels, and
 653 correlated between the 2 and 3 jet regions.
- 654 • **In the e and μ modes**, the impact of applying the top and $W + hf$ normalisation factors in the
 655 anti-isolated region are considered. The MJ template shape is evaluated when applying these
 656 normalisation corrections, and the difference in shape taken as a systematic uncertainty. This
 657 uncertainty is decorrelated between e/μ channels, and correlated between the 2 and 3 jet regions.

Dataset	Single e Trigger
2015	e24_1hmedium
2016	e26_1htight_nod0_ivarloose

Table 16: Reduced triggers used to evaluate possible trigger bias in inverted electron isolation region.

658 These systematic uncertainties are implemented as shape only systematic uncertainties on the BDT output
 659 by normalising the variation to the nominal MJ yield. Their individual impact, along with additional
 660 normalisation uncertainties are added in quadrature to give the overall normalisation uncertainty for each
 661 region and channel, as described below.

662 **Normalisation Uncertainty**

663

664 A normalisation uncertainty of the MJ background is estimated separately in the 2 and 3 jet regions, and
 665 the electron and muon modes. In addition to the impact on the MJ normalisation induced by the shape
 666 uncertainties described above, the following systematics are also considered:

- 667 • The impact of applying $t\bar{t}$ m_{bb} and p_T^V shape systematics in the inverted isolation region.
- 668 • Deriving the MJ template including the $E_T^{\text{miss}} < 30 \text{ GeV}$ region (e mode only)
- 669 • Deriving the MJ template from the E_T^{miss} distribution instead of m_T^W

670 These uncertainties are combined in quadrature to yield the total uncertainties presented in Table 17,
 671 with the central normalisation factors derived from the simultaneous fit to data in the signal region.

672 The corresponding fractions of the multi-jet contribution compared to the total background and their
 673 uncertainties are presented in Table 18.

Region	MJ Scaling Factor
2-tag, 2-jet, e	$2.07^{+3.74}_{-1.22} \times 10^{-2}$
2-tag, 2-jet, μ	$9.03^{+14.09}_{-9.03} \times 10^{-2}$
2-tag, 3-jet, e	$0.35^{+1.81}_{-0.35} \times 10^{-2}$
2-tag, 3-jet, μ	$4.09^{+12.19}_{-4.09} \times 10^{-2}$

Table 17: Summary of MJ normalisation scale factors, along with their associated uncertainty.

Region	MJ Fractions (%)	MJ norm. uncertainty
2-tag, 2-jet, e	$4.53^{+2.68}_{-2.79}$	-62% / +60%
2-tag, 2-jet, μ	$3.16^{+2.45}_{-3.16}$	-100% / +76%
2-tag, 3-jet, e	$0.56^{+1.09}_{-0.56}$	-100% / +194%
2-tag, 3-jet, μ	$0.38^{+1.99}_{-0.38}$	-100% / +520%

Table 18: Summary of MJ fractions, along with their associated uncertainty in the 2-jet and 3-jet 1-lepton regions
 (WCR and SR are combined) separate

674 4.4.3. 2-lepton

675 The requirements of two isolated leptons and the Z mass window greatly reduces multi-jet (MJ) back-
 676 ground. The remaining multi-jet background is evaluated using the Z mass peak side-band, as the multi-jet
 677 background will not produce any peak in di-lepton mass distribution. In the 2-lepton analysis, the multi-jet
 678 background is estimated using the template method.

679 Template Method

680

681 A background enriched region is defined to model the multi-jet background. The multi-jet background
 682 template is obtained from same-sign events to reduce the contamination of EW and top backgrounds.
 683 The multi-jet shape is modeled by exponential function, with SS-data fit to the non-MJ background and
 684 the exponential multi-jet model. The Z background normalization is a free parameter, and the shape
 685 is taken from MC since the $Z+jets$ background has quite a different shape from MJ background. The
 686 non- Z background (mainly top) shape is more similar to MJ shape, meaning that with a freely floating
 687 normalization parameter, the non- Z background can be absorbed by the MJ contribution. This results in an
 688 overestimation of MJ contribution. It is found that the dominant top background normalization is close to
 689 unity from other control regions, while the normalization of other non- Z backgrounds is fixed. Figure 12
 690 shows the di-electron mass distribution in same-sign events. Due to charge flipping of real electrons, a
 691 clear Z peak can still be observed. The multi-jet shape and normalization can then be obtained from a fit
 692 to SS events. The fitting is performed separately for 0-tag, 1-tag and 2-tag regions to accommodate for any
 693 dependence on b -tagging of the MJ contamination. In the 2-tag category, statistics are limited. Figure 13
 694 shows the di-muon mass distribution in same-sign events. Similarly, SS-data is fit with EW background
 695 and an exponential multi-jet model. The Z peak is also observed in the SS $\mu\mu$ channel. This strange
 696 property results from the bug of charge assignment in SiliconAssociatedForwardMuon. The impact of
 697 this bugged muon reconstruction has been assessed and found to be negligible. More detail can be found
 698 in appendix B

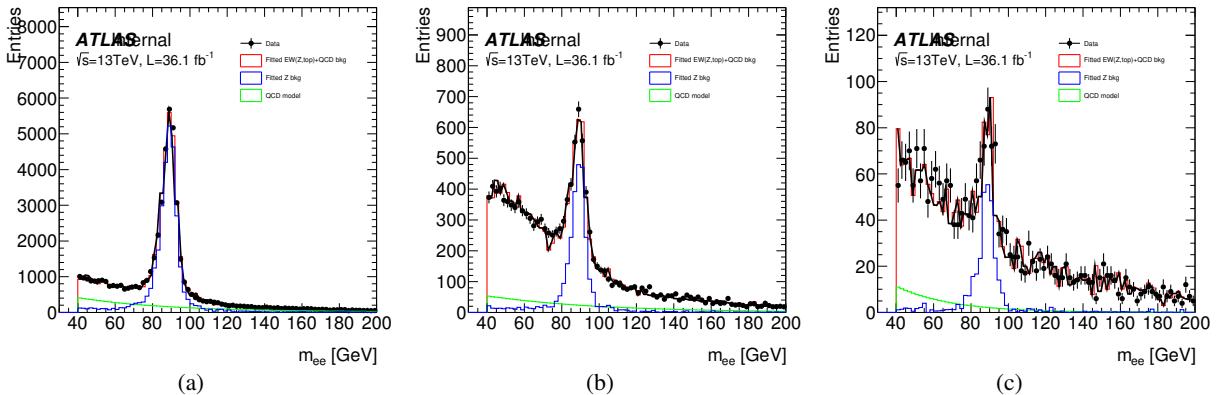


Figure 12: Same-sign dielectron mass distribution. 0-tag (a), 1-tag (b), 2-tag (b). Green histogram shows the estimated multi-jet contribution from same-sign events. The red histogram shows the fitted EW background modeled by MC plus multi-jet.

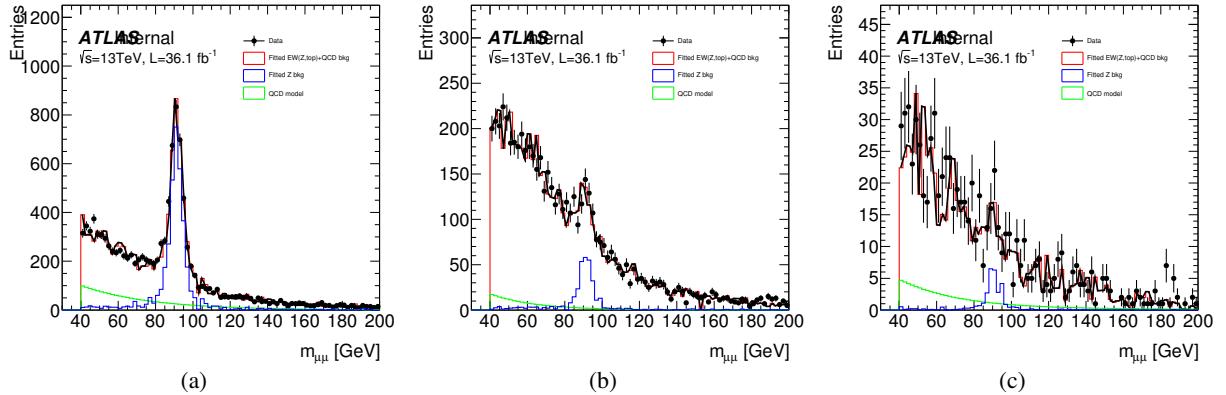


Figure 13: Same-sign dimuon mass distribution. 0-tag (a), 1-tag (b), 2-tag (c). Green histogram shows estimated multi-jet contribution from same-sign events. Red histogram shows fitted EW background modeled by MC plus multi-jet.

699 Figures 14 and 15 shows opposite sign dilepton mass distribution. Assuming OS/SS ratio is one for
 700 multi-jet background, It is found the multijet background is quite small compared to other backgrounds.
 701 Tables 19 and 20 show the multi-jet fraction of total number of data in OS signal region. The fraction of
 702 multi-jet background is less than 1% in signal enriched region, ($81 < m_{ll} < 101$ GeV).

Jet correction variation	Multi-jet fraction ($71 < m_{ll} < 121$ GeV)	Multi-jet fraction ($81 < m_{ll} < 101$ GeV)
0 b-tag	0.52%	0.25%
1 b-tag	0.64%	0.35%
2 b-tag	0.24%	0.17%

Table 19: Multi-jet background amount w.r.t. total background for 0,1,2 b -tagged events in di-electron channel, respectively. The numbers are shown in wider m_{ll} range (ICHEP selection) and tighter selection used in this analysis.

Jet correction variation	Multi-jet fraction ($71 < m_{ll} < 121$ GeV)	Multi-jet fraction ($81 < m_{ll} < 101$ GeV)
0 b-tag	0.04%	0.02%
1 b-tag	0.03%	0.02%
2 b-tag	0.05%	0.03%

Table 20: Multi-jet background amount w.r.t. total background for 0,1,2 b -tagged events in dimuon channel, respectively. The numbers are shown in wider m_{ll} range (ICHEP selection) and tighter selection used in this analysis.

703 Figure 16 shows m_{bb} distribution merged in all signal categories. To compare the shape of multi-jet and
 704 ZH signal, the normalization is set to unity. The multi-jet shape has been extracted from same-sign data
 705 subtracted EW background. The non- Z background has been scaled by 0.7 to obtain enough statistics in
 706 SS events. The multi-jet m_{bb} shape is quite different from signal and no-peak structure. In the signal
 707 mass window ($100 \text{ GeV} < m_{bb} < 140 \text{ GeV}$), the acceptance is 0.17 for multi-jet shape and 0.69 for signal
 708 shape. The signal fraction in this mass window with regard to total number of data in all signal regions is
 709 0.2%, while the multi-jet fraction is reduced to 0.02%. Even if the non- Z contribution is scaled down to

Not reviewed, for internal circulation only

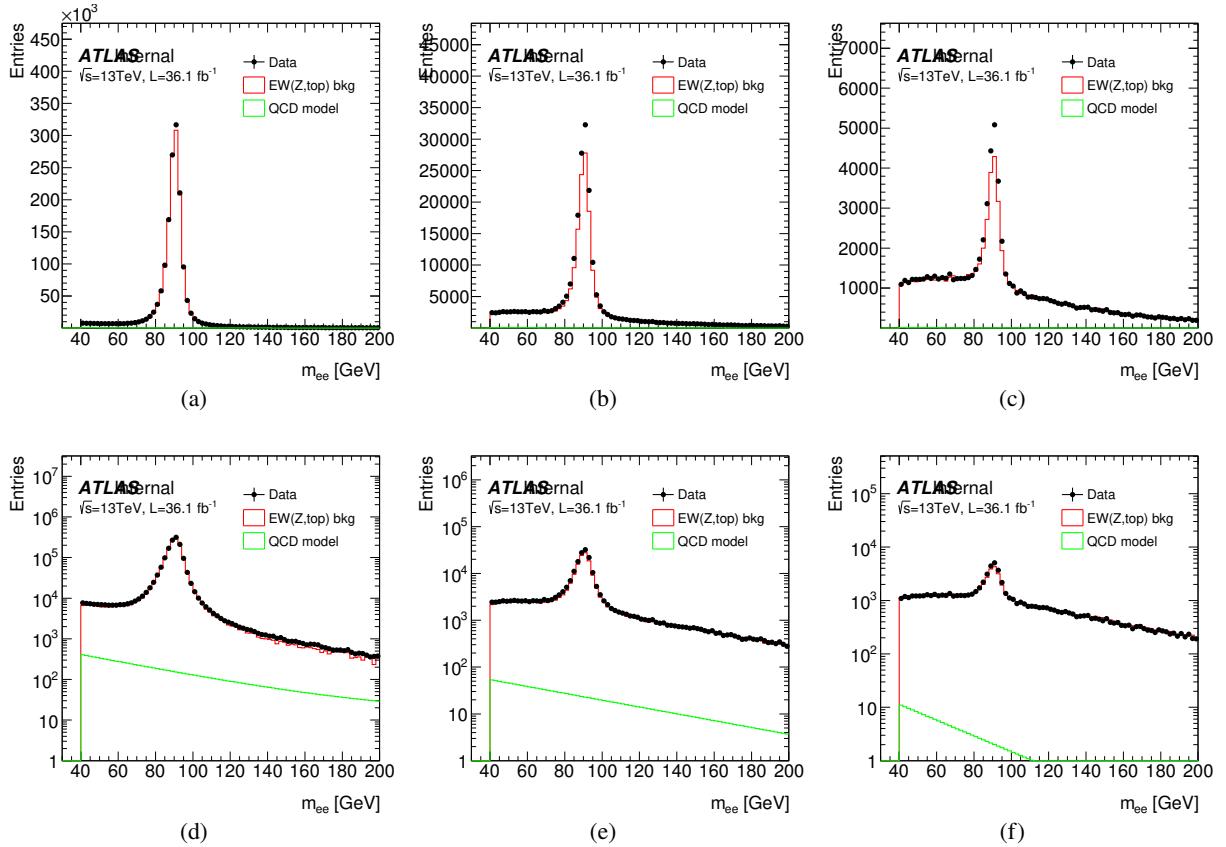


Figure 14: Opposite-sign dielectron mass distribution. 0-tag (a), 1-tag (b), 2-tag (b). Green histogram shows the estimated multi-jet contribution from same-sign events. The red histogram shows the fitted EW background modeled by MC. Bottom plots show in log-scale, respectively.

710 0.7 in SS events, the multi-jet fraction is still $\sim 0.05\%$, which is 4 times lower than signal yield within the
 711 mass window. The multi-jet background is therefore sufficiently small in the 2-lepton channel that it can
 712 be neglected. Figure 21 shows the yield comparison between signal and multi-jet background with and
 713 without m_{bb} mass window cut for 36.1 fb^{-1} .

Yield	Expected Signal Yield ($\mu = 1$)	Estimated MJ background
2 b-tag (all signal region)	165.73	55.28
2 b-tag ($100 < m_{bb} < 140$)	114.4	9.4

Table 21: Expected signal and multijet background yield with and without signal mass window cut.

714 Figure 17 shows the multi-jet contamination in the 2 b -tagged $e\mu$ CR using the template method. $e\mu$ events
 715 are dominated by the top contribution, with a very small $V+jets$ contamination. The non Z normalization
 716 has been fixed to estimate multi-jet modeling in same-sign $e\mu$ events. The multi-jet model has been applied
 717 to opposite-sign events, with an estimated multi-jet contamination of 0.3% in the $m_{e\mu}$ mass window. Even if
 718 the top background normalization in same-sign events are scaled down to 70%, the multi-jet contamination
 719 is still less than 1%. It is therefore concluded that the multi-jet background can be neglected in $e\mu$ CR as
 720 such a small contribution can be absorbed in the dominant top background normalization.

Not reviewed, for internal circulation only

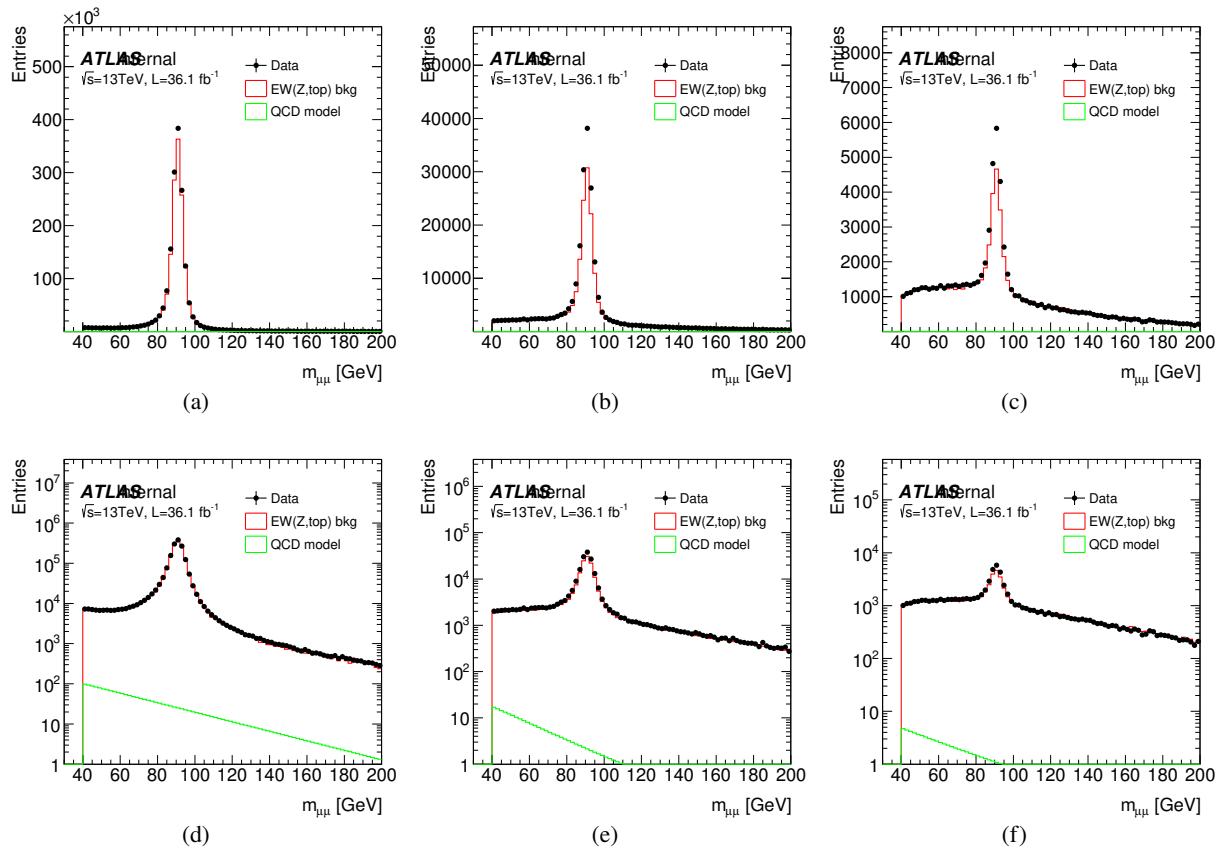
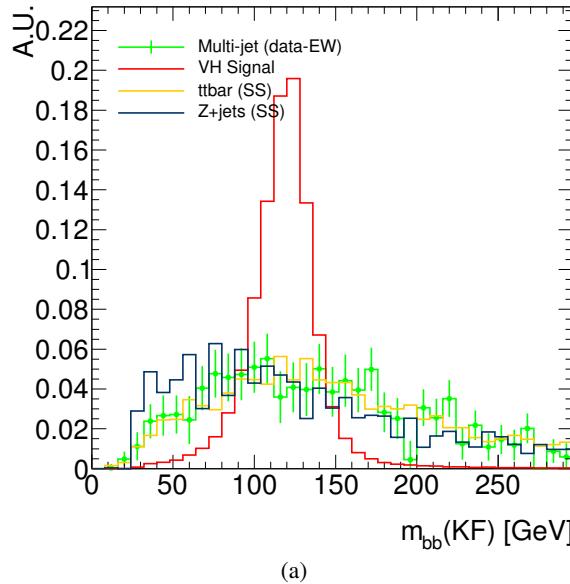


Figure 15: Opposite-sign dimuon mass distribution. 0-tag (a), 1-tag (b), 2-tag (b). Green histogram shows the estimated multi-jet contribution from same-sign events. The red histogram shows EW background modeled by MC. Bottom plots show in log-scale, respectively.

Not reviewed, for internal circulation only



(a)

Figure 16: Shape comparison among signal, $t\bar{t}$, Z+jets and multi-jet which is extracted from SS data.

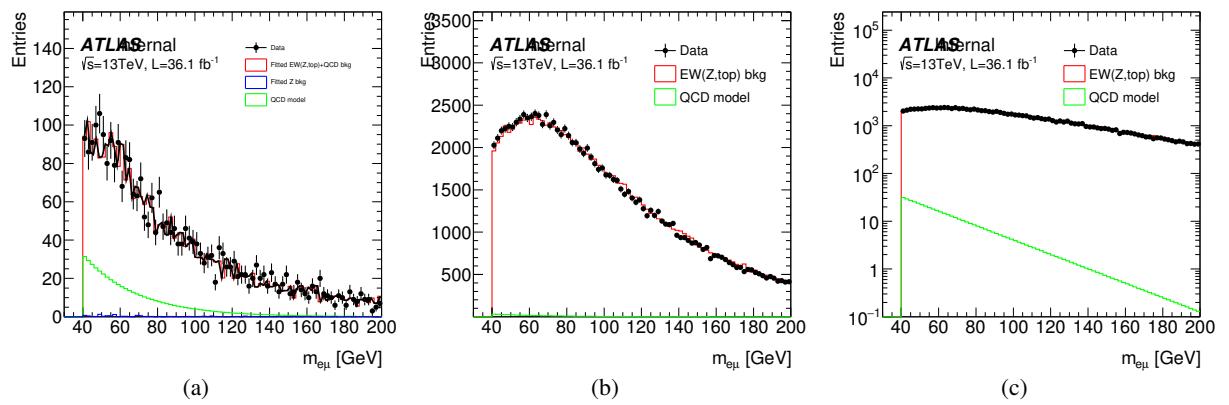


Figure 17:

721 4.5. Pre-fit data-MC comparison

722 In this section, shape and normalisation comparisons between data and expectations for the 0-, 1- and
 723 2-lepton analyses, after the selection described in Section 3 is applied, are presented. For these pre-fit plots
 724 the nominal simulations, without applying any post-fit scale factors, are used to compute the expectation.
 725 The multijet and non-collisional backgrounds are neglected in the 0- and 2-lepton channels since, as shown
 726 in Section 4.4, they are negligible.

727 4.5.1. 0-lepton channel

728 In Table 22, the data yield is compared to the pre-fit expectations, for the 0-lepton channel, after the full
 selection is applied.

Process	2 jet, 2 tag	3 jet, 2 tag
qqZvvH	31.07 ± 0.15	29.98 ± 0.14
qqZllH	0.25 ± 0.01	0.29 ± 0.01
ggZvvH	6.55 ± 0.13	11.41 ± 0.17
ggZllH	0.05 ± 0.01	0.08 ± 0.01
qqWH	8.76 ± 0.11	10.06 ± 0.15
Wl	10.3 ± 0.4	19.04 ± 0.16
Wcl	21.8 ± 0.4	43.5 ± 0.5
Whf	362.6 ± 9.5	886 ± 15.9
Zl	9.8 ± 0.03	17.0 ± 0.05
Zcl	22.8 ± 0.1	43.1 ± 0.3
Zhf	1549 ± 20	2490 ± 24
ttbar	476 ± 14.00	3400 ± 38
stop	144.7 ± 7.6	569 ± 16.2
diboson	118.1 ± 3.9	140 ± 4
Total Bkg	2715.9 ± 27.41	7611 ± 49.41
data	3520.00 ± 59.33	8634.00 ± 92.92

Table 22: Pre-fit yields for the 0-lepton analysis, for an integrated luminosity of 36.1 fb^{-1} . The uncertainty shown is of statistical nature only. The Vhf yield count includes events from the $V + bb$, $V + bc$, $V + bl$ and $V + cc$ categories.

729
 730 In this section, prefit data-MC comparisons are shown for a number of distributions, in the 2- and 3-jet
 731 regions. In Figure 18, and 19, the E_T^{miss} in the event, and the invariant mass of the two selected jets, are
 732 shown, respectively. Other prefit distributions used in the BDT are shown in Appendix E.0.1. In most of
 733 the distributions shown, MC and data shapes agree well, while a large overall normalisation difference is
 734 observed. The normalisations of some of the most important backgrounds, such as $t\bar{t}$ and $Zbb\bar{b}$, are badly
 735 described and will need to be rescaled in the fit later.

Not reviewed, for internal circulation only

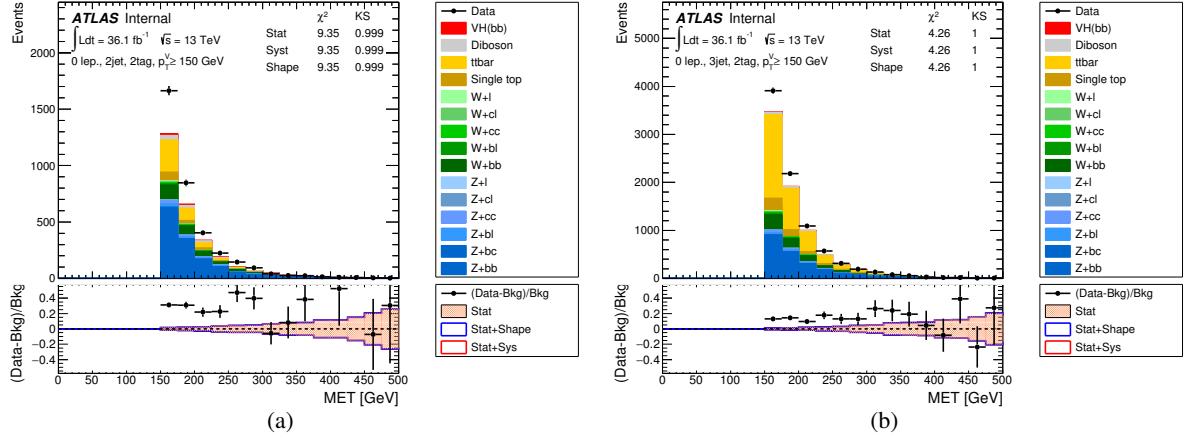


Figure 18: The pre-fit distributions in the 0-lepton channel for 2-tag 2-jet (a) and 2-tag 3-jet (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

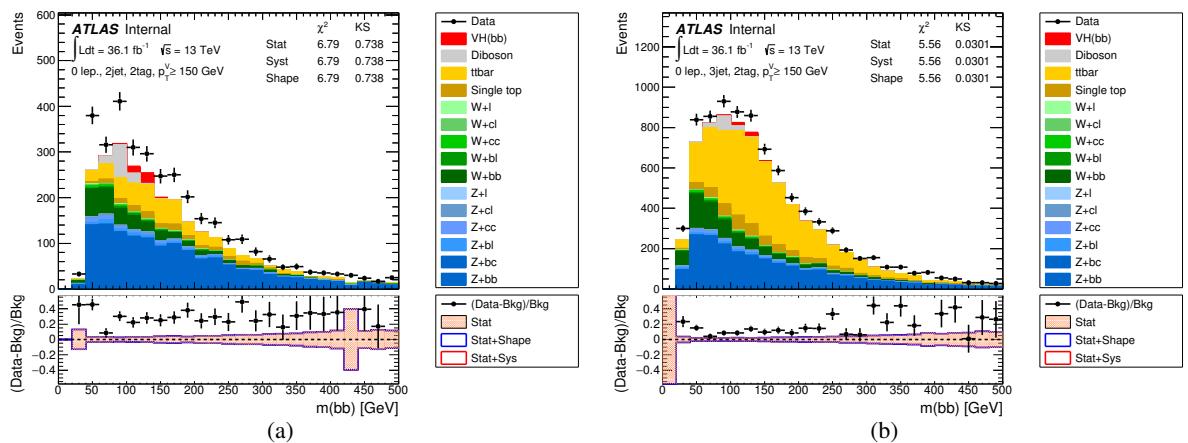


Figure 19: The pre-fit distributions in the 0-lepton channel for 2-tag 2-jet (a) and 2-tag 3-jet (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

⁷³⁶ **4.5.2. 1-lepton channel**

⁷³⁷ In Table 23, the data yields in the high p_T^W regions are compared to the pre-fit expectations, for the
⁷³⁸ 1-lepton channel, after the full selection is applied.

Process	2 jet		3 jet	
	2-tag, $W + hf$ CR	2-tag, SR	2-tag, $W + hf$ CR	2-tag, SR
$qqWlvH125$	0.3 ± 0.0	50.1 ± 0.3	1.0 ± 0.0	53.9 ± 0.3
$ggZH125$	0.0 ± 0.0	0.2 ± 0.0	0.0 ± 0.0	0.4 ± 0.0
$qqZH125$	0.0 ± 0.0	0.8 ± 0.0	0.0 ± 0.0	1.3 ± 0.0
Zl	0.1 ± 0.0	1.0 ± 0.3	0.1 ± 0.0	0.8 ± 1.2
Zcl	0.0 ± 0.0	2.2 ± 0.4	0.2 ± 0.1	4.2 ± 0.6
Zhf	5.3 ± 0.9	62.9 ± 3.0	15.4 ± 1.5	154.5 ± 4.6
Wl	1.1 ± 0.2	23.2 ± 4.2	2.9 ± 0.7	49.2 ± 5.7
Wcl	2.9 ± 0.9	75.1 ± 4.8	10.6 ± 1.6	135.8 ± 5.9
Whf	195.7 ± 7.3	1051.9 ± 20.5	453.5 ± 10.0	2480.4 ± 28.0
single-top	9.9 ± 2.4	863.8 ± 18.5	43.4 ± 4.5	3926.7 ± 41.0
ttbar	24.0 ± 3.1	2869.0 ± 34.2	116.2 ± 6.8	22395.1 ± 95.8
diboson	5.1 ± 1.0	79.4 ± 4.3	12.1 ± 1.5	147.9 ± 5.8
multijetEl	9.6 ± 0.6	119.9 ± 1.9	2.4 ± 0.1	29.0 ± 0.5
multijetMu	7.1 ± 1.4	138.2 ± 6.3	3.5 ± 0.8	86.3 ± 3.9
Total Bkg	260.8 ± 8.5	5286.5 ± 45.2	660.3 ± 13.2	29410.0 ± 108.6
Data	302.0	5307.0	790.0	28168.0
Data/MC	1.16	1.00	1.20	0.96

Table 23: Pre-fit yields for the 1-lepton analysis for $p_T^W > 150$ GeV, for an integrated luminosity of 36.07 fb^{-1} . The uncertainty shown is of statistical nature only. The Vhf yield count includes events from the $V + bb$, $V + bc$, $V + bl$ and $V + cc$ categories.

⁷³⁹ This section shows plots from 1-lepton channel BDT input variables. Each figure includes plots for the 1-
⁷⁴⁰ and 2-tag categories, for the 2- and 3-jet categories. The first set of figures referenced here correspond to
⁷⁴¹ $p_T^W > 150$ GeV. The invariant mass of the reconstructed Higgs candidate from the two leading jets, m_{bb}
⁷⁴² is presented in Figure 20. Figure 21 presents the p_T^W distribution. Additional prefit BDT input plots are
⁷⁴³ included in Appendix E.0.2.

Not reviewed, for internal circulation only

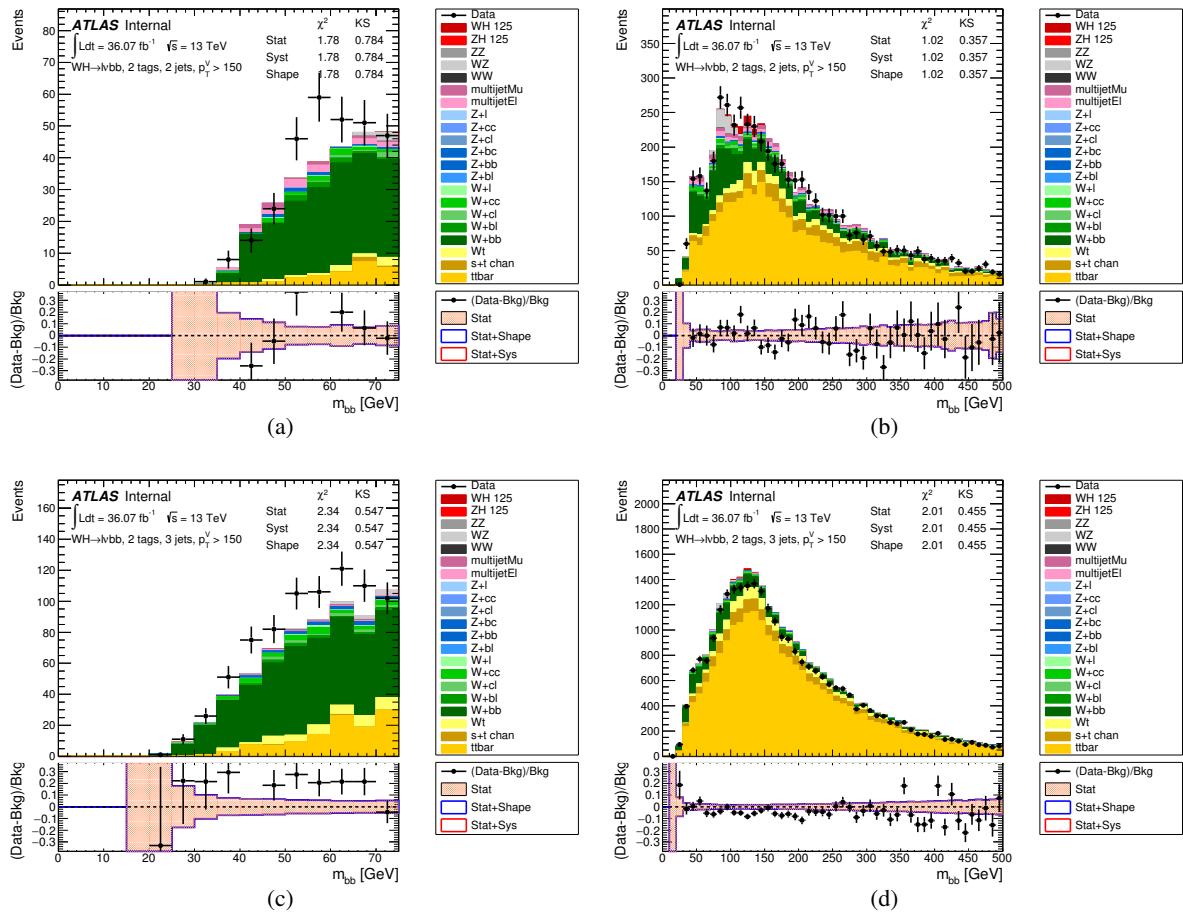


Figure 20: The pre-fit $p_T^W > 150$ GeV m_{bb} distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

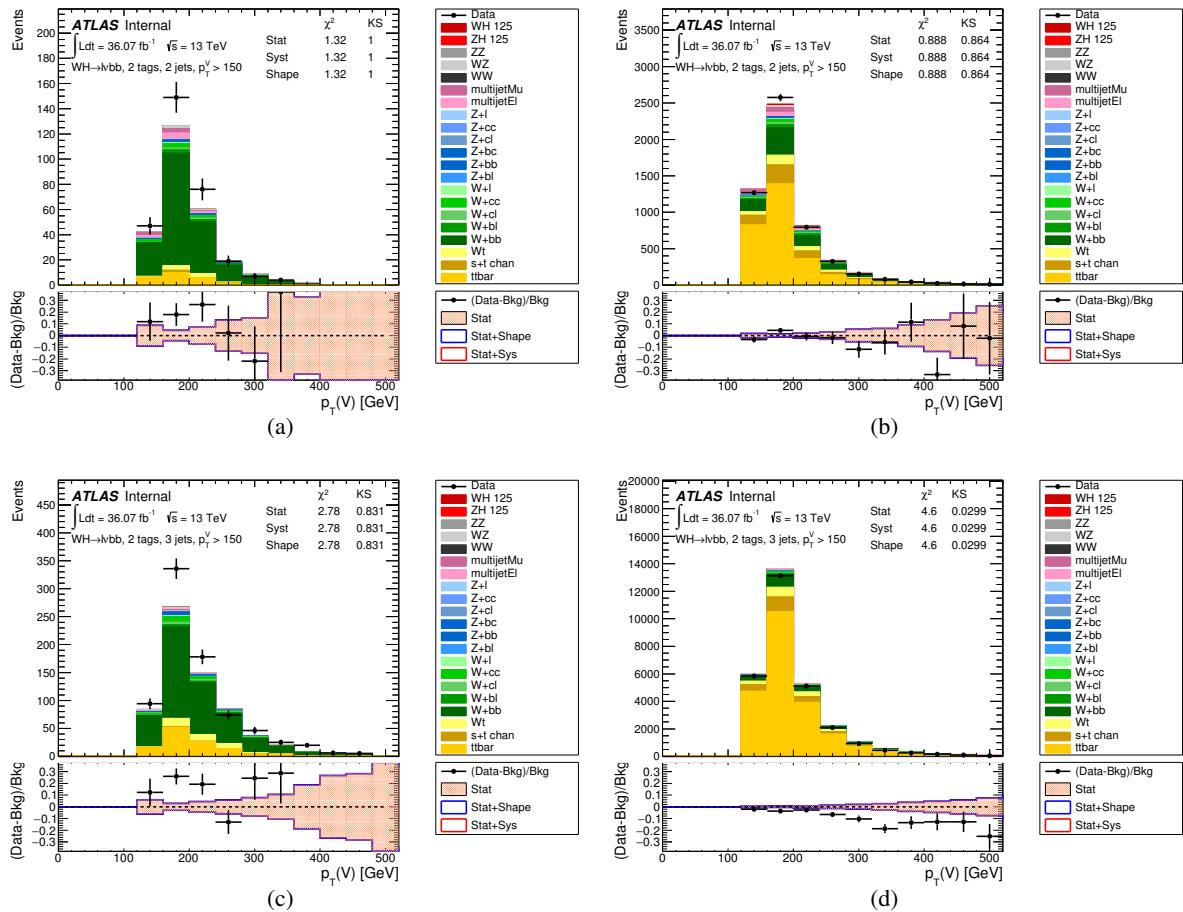


Figure 21: The pre-fit $p_T^W > 150$ GeV p_T^W distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

744 4.5.3. 2-lepton channel

745 In this section, yields and some sensitive kinematic distributions used in 2-lepton MVA are presented. The
 746 yields and kinematic distributions of pre-fit are compared with data. Each V +heavy flavor contribution
 747 ($V + bb, bc, bl, cc, cl$) is not well predicted in the MC simulation. Each normalization and shape will be
 748 corrected/profiled with using the statistical treatment described in Section 7. Post-fit distributions are
 749 presented in Section 7.4.5. As discussed in Section 4.4.3, the multi-jet background is negligible in the
 750 2-lepton analysis, and is not taken into account.

751 In Tables 24 - 25, the data yields in the two p_T^Z signal regions are compared to the pre-fit expectations
 752 after the pre-selection is applied. In Table 26- 27, the data yields in the two $p_T^Z e\mu$ control regions are
 753 compared to the pre-fit expectations after the pre-selection is applied.

Process	2 jet	3+ jet
	2 tag	2 tag
$qqZH(125)$	17.85 ± 0.156	29.49 ± 0.199
$ggZH(125)$	3.621 ± 0.0432	10.86 ± 0.0754
Zl	9.828 ± 0.0363	38.19 ± 0.179
Zcl	26.22 ± 0.208	109.4 ± 0.675
Zhf	2745 ± 38.4	7172 ± 40.5
W +jet	2.608 ± 1	4.813 ± 0.959
single top	59.45 ± 4.09	159.8 ± 5.6
$t\bar{t}$	1516 ± 14.8	4840 ± 27
Diboson	78.4 ± 2.2	172.5 ± 3.6
Ntotal bkg	4437.7 ± 41.5	12497 ± 49.1
Observed data	5113	13640
Ndata/NMC	1.15	1.09

Table 24: Pre-fit yields of signal region for the 2-lepton analysis for $75 \text{ GeV} < p_T^Z < 150 \text{ GeV}$, for an integrated luminosity of 36.1 fb^{-1} . The uncertainty shown is of statistical nature only. The Vhf yield count includes events from the $V + bb, V + bc, V + cc$ and $V + bl$ categories.

754 Figures 22-23 show the dijet mass distribution which is one of most sensitive discriminants in this analysis
 755 and p_T^V distributions for 2jet and 3+jet categories, respectively.

756 Figures 24-25 show m_{bb} and p_T^V distribution for $e\mu$ control regions. Control regions are divided into the
 757 two p_T^V categories to select same phase space as signal region.

758 Other kinematic plots used as input variables to the BDT are shown in Appendix E.0.3.

Process	2 jet	3+ jet
	2 tag	2 tag
$qqZH(125)$	8.991 ± 0.042	19.48 ± 0.0622
$ggZH(125)$	2.277 ± 0.0338	10.63 ± 0.0729
Zl	2.102 ± 0.00794	17.79 ± 0.0976
Zcl	5.754 ± 0.0421	50.75 ± 0.265
Zhf	522.7 ± 6.15	2668 ± 12.8
$W+jet$	0.2481 ± 0.176	1.763 ± 0.524
single top	7.348 ± 0.998	34.24 ± 2.25
$t\bar{t}$	52.96 ± 2.49	437.4 ± 7.99
Diboson	25.73 ± 1.32	99.08 ± 2.88
Ntotal bkg	616.82 ± 6.84	3309.5 ± 15.5
Observed data	724	3708
Ndata/NMC	1.17	1.12

Table 25: Pre-fit yields of signal region for the 2-lepton analysis for $p_T^Z > 150$ GeV, for an integrated luminosity of 36.1 fb^{-1} . The uncertainty shown is of statistical nature only. The Vhf yield count includes events from the $V + bb$, $V + bc$, $V + cc$ and $V + bl$ categories.

Process	2 jet	3+ jet
	2 tag	2 tag
$qqZH(125)$	0.001977 ± 0.000963	0.003839 ± 0.00186
$ggZH(125)$	0 ± 0	0.001631 ± 0.000865
Zl	0.001593 ± 0.000309	0.01267 ± 0.00496
Zcl	0.006685 ± 0.00215	0.03016 ± 0.00619
Zhf	1.574 ± 0.628	2.149 ± 0.385
$W+jet$	2.51 ± 1.93	1.026 ± 1.02
single top	58.96 ± 3.35	150.4 ± 5.57
$t\bar{t}$	1482 ± 13.9	4801 ± 26.5
Diboson	0 ± 0	0.2282 ± 0.142
Ntotal bkg	1545.1 ± 14.5	4955 ± 27.1
Observed data	1489	4967
Ndata/NMC	0.964	1

Table 26: Pre-fit yields of $e\mu$ control region for the 2-lepton analysis for $75 \text{ GeV} < p_T^Z < 150 \text{ GeV}$, for an integrated luminosity of 36.1 fb^{-1} . The uncertainty shown is of statistical nature only. The Vhf yield count includes events from the $V + bb$, $V + bc$, $V + cc$ and $V + bl$ categories.

Process	2 jet	3+ jet
	2 tag	2 tag
$qqZH(125)$	0.001799 ± 0.00105	0.002444 ± 0.000531
$ggZH(125)$	0 ± 0	0.001586 ± 0.000723
Zl	$0.0003399 \pm 7.95\text{e-}05$	0.003054 ± 0.000363
Zcl	0.0005207 ± 0.000401	0.3122 ± 0.304
Zhf	0.2711 ± 0.115	1.008 ± 0.245
$W+\text{jet}$	0.1605 ± 0.12	0.4662 ± 0.324
single top	7.318 ± 1.03	32.12 ± 2.32
$t\bar{t}$	55.45 ± 3.02	430 ± 8.05
Diboson	-0.02048 ± 0.0205	0 ± 0
Ntotal bkg	63.179 ± 3.2	463.94 ± 8.39
Observed data	50	470
Ndata/NMC	0.791	1.01

Table 27: Pre-fit yields of $e\mu$ control region for the 2-lepton analysis for $p_T^Z > 150$ GeV, for an integrated luminosity of 36.1 fb^{-1} . The uncertainty shown is of statistical nature only. The Vhf yield count includes events from the $V + bb$, $V + bc$, $V + cc$ and $V + bl$ categories.

Not reviewed, for internal circulation only

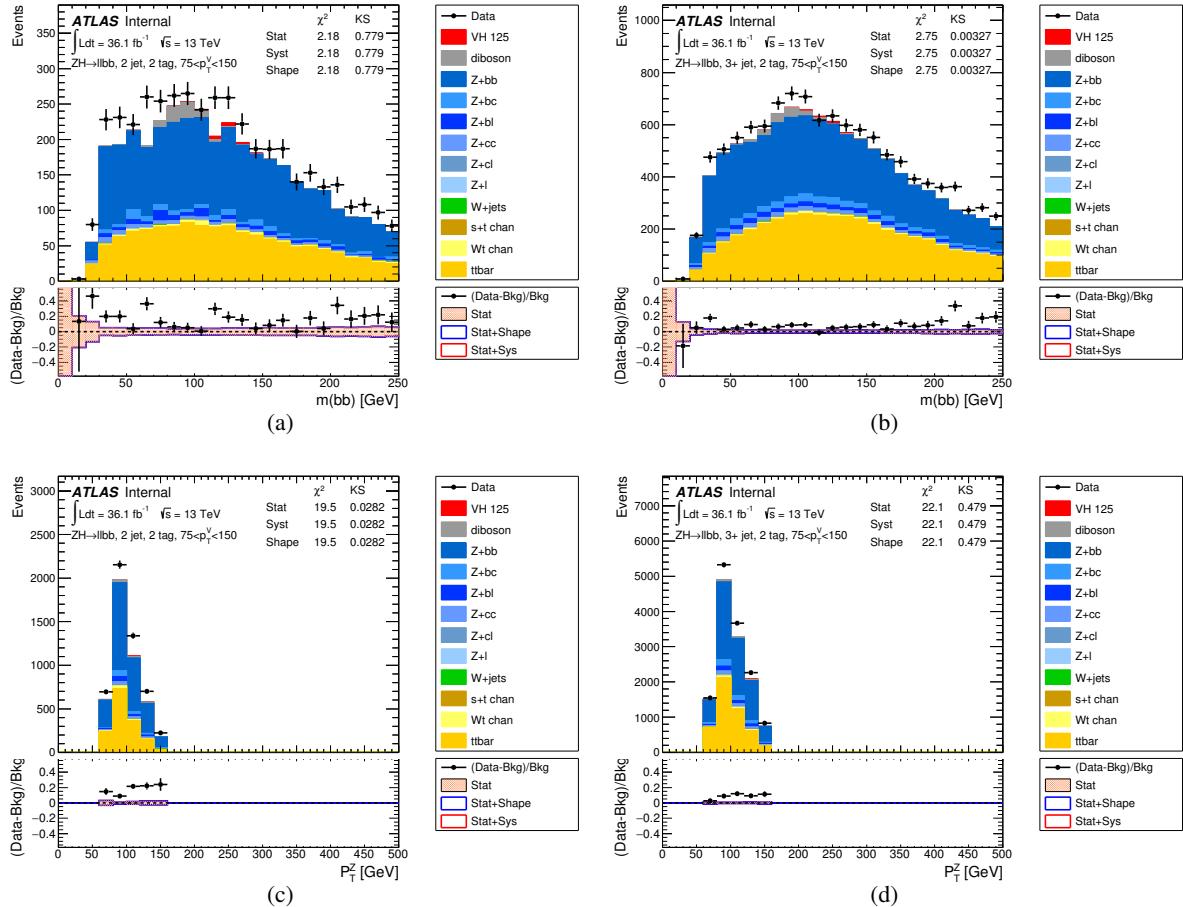


Figure 22: The pre-fit m_{bb} distributions in the 2-leptons channel, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$ category for 2-tag events, in the (a) 2jets, and (b) 3+jets are shown. The pre-fit p_T^Z distributions in the 2-leptons channel, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$ category for 2-tag events, in the (c) 2jets, and (d) 3+jets are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

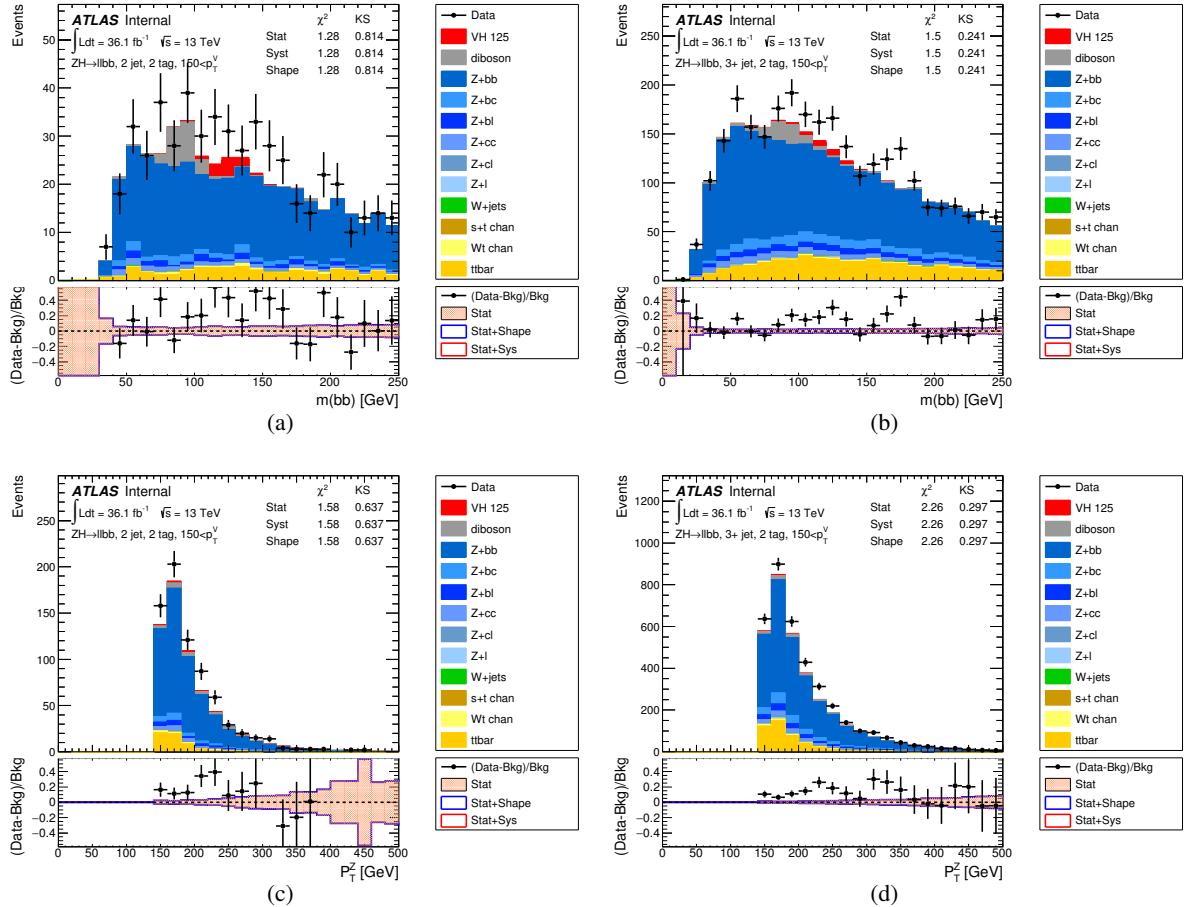


Figure 23: The pre-fit m_{bb} distributions in the 2-leptons channel, $150 \text{ GeV} < p_T^V$ category for 2-tag events, in the (a) 2jets, and (b) 3+jets are shown. The pre-fit p_T^Z distributions in the 2-leptons channel, $150 \text{ GeV} < p_T^V$ category for 2-tag events, in the (c) 2jets, and (d) 3+jets are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

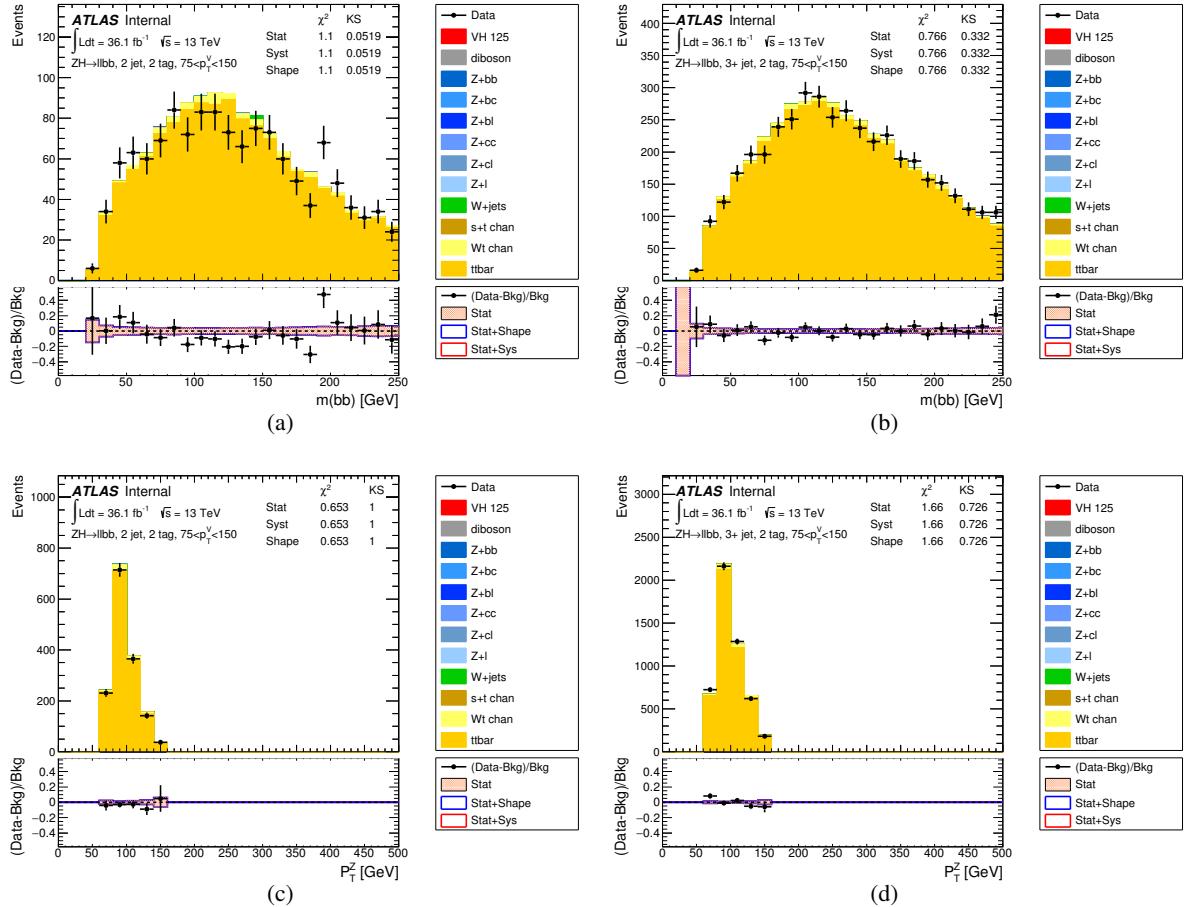


Figure 24: The pre-fit m_{bb} distributions in the $e\mu$ control region, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$ category for 2-tag events, in the (a) 2jets, and (b) 3+jets are shown. The pre-fit p_T^Z distributions in the 2-leptons channel, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$ for 2-tag events, in the (c) 2jets, and (d) 3+jets are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

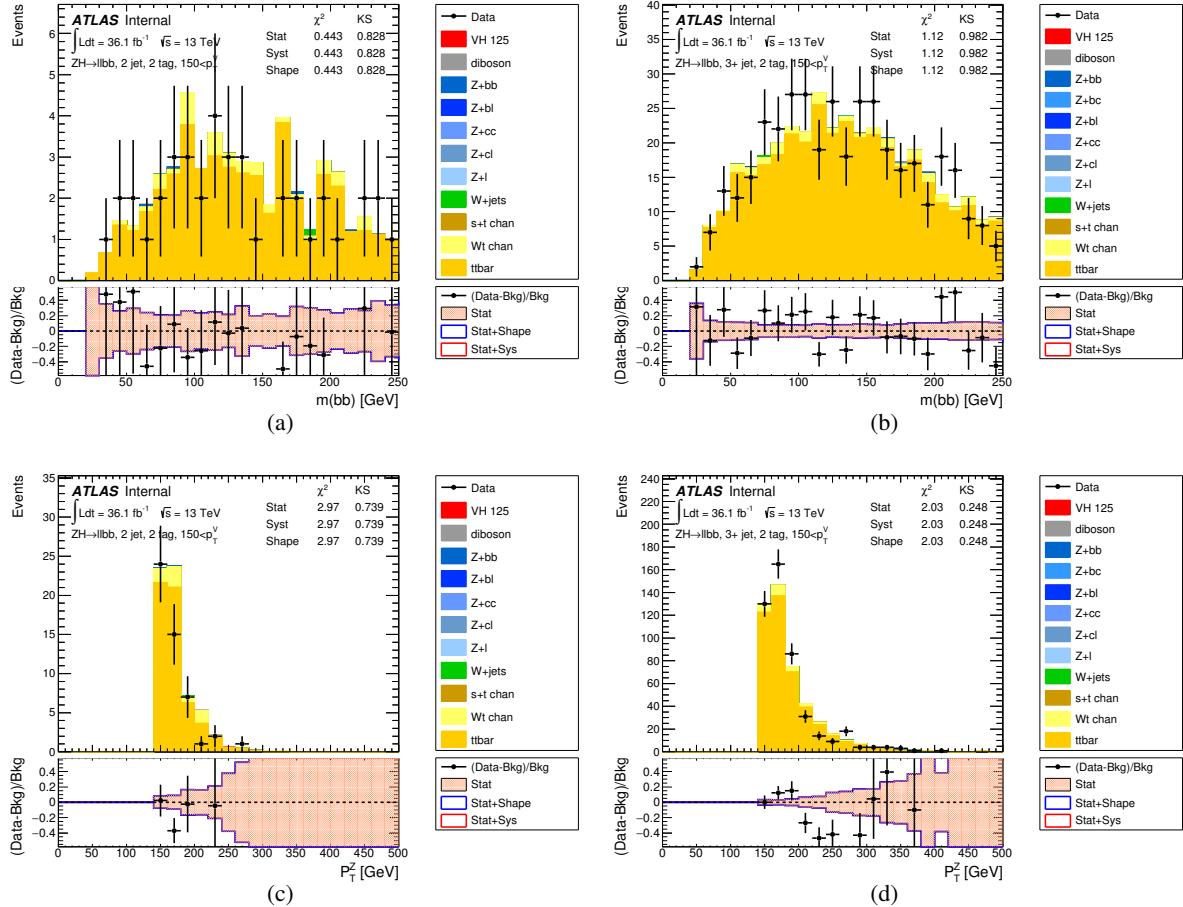


Figure 25: The pre-fit m_{bb} distributions in the $e\mu$ control region, $150 \text{ GeV} < p_T^V$ category for 2-tag events, in the (a) 2jets, and (b) 3+jets are shown. The pre-fit p_T^Z distributions in the 2-leptons channel, $150 \text{ GeV} < p_T^V$ category for 2-tag events, in the (c) 2jets, and (d) 3+jets are shown. The background and signal samples are normalized to the expected cross-section predictions.

759 5. Multi-Variate Analysis

760 MVAs are used in a variety of high energy physics analyses to offer increased signal purity and background
 761 rejection. This is achieved through the combination of a well-chosen set of discriminating input variables
 762 which the MVA is trained on, to construct a one dimensional discriminant. The algorithm should be set
 763 up taking into account the available MC statistics, so that the final result doesn't depend on the random
 764 statistical fluctuations in the input distributions. On the other hand, MVAs must be trained and evaluated
 765 on separate MC samples to ensure an unbiased result: in this analysis a 2-fold cross-validation of the
 766 MVA training is implemented. One training is performed using even (odd) event-numbered MC events,
 767 and then applied to odd (even) events, thereby ensuring orthogonality between the training and testing
 768 samples. In order to use all of the available statistics, all events are then summed up for the production of
 769 the final discriminant distribution. For this analysis, a Boosted Decision Tree (BDT) is used, similarly to
 770 what was done in the Run-1 analysis [13].

771 5.1. Input variables

772 The variables passed as an input to the BDT are defined after the selection requirements described in
 773 Section 3 are applied. The input variables used in each channel are summarised in Table 28. Definitions
 774 of the input variables are as follows:

- 775 • V : vector boson, constructed from E_T^{miss} , vectorial sum of E_T^{miss} and p_T^ℓ and vectorial sum of $p_T^{l_{1,2}}$
 776 in the 0-, 1- and 2-lepton channels respectively;
- 777 • H : Higgs candidate, constructed from the vectorial sum of the two b-tagged jets, as defined in
 778 Section 3.2.1 for the 0-, 1- and 2-lepton channels;
- 779 • M_{eff} : scalar sum of the p_T of all of the objects used in the event, specifically the signal jets,
 780 leptons, and E_T^{miss} . In the 0-lepton channel the third jet is considered in three jet events. This is
 781 clarified by the use of $M_{\text{eff}3}$ for the latter case;
- 782 • m_{jj} : dijet mass constructed from the two b-tagged jets, as defined in Section 3.2.1 for the 0-, 1-
 783 and 2-lepton channels;
- 784 • m_T^W : transverse mass of the W , defined as $m_T^W = \sqrt{2p_T^\ell E_T^{\text{miss}}(1 - \cos(\Delta\phi(\ell, E_T^{\text{miss}})))}$;

785 The overall signal and background input distributions passed to the BDT in the 0,2-lepton channel are
 786 shown also in Figures 26-27. For 2-lepton channel, BDT inputs for three or more than three jet, high p_T^V
 787 category are shown. Similar plots for other categories are shown in Appendix I.2.1. The list of input
 788 variables used for the BDT training are based on studies conducted during Run-1, where an iterative
 789 procedure was adopted to find the optimal set of variables and ranking to use. Initially, the BDT was
 790 constructed using simply m_{bb} and $\Delta R(j, j)$. Each candidate variable was then added to the MVA in
 791 turn, with the variable offering the best improvement in significance being added to the MVA as the third
 792 variable. The final optimal MVA is then constructed when all variables have been studied and no further
 793 improvement is seen [27].

794 Since Run-1, $\Delta Y(W, H)$ and m_{top} have been included in the 1-lepton channel MVA as this was found to
 795 offer a gain in the performance of the MVA. m_{top} is the reconstructed leptonically decaying top mass,
 796 $\Delta Y(W, H)$ is the reconstructed rapidity difference of the Higgs candidate and leptonically decaying W

Not reviewed, for internal circulation only

Variable	Name	0-lepton	1-lepton	2-lepton
p_T^V	pTV		✓	✓
E_T^{miss}	MET	✓	✓	✓
$p_T^{\text{jet}1}$	pTB1	✓	✓	✓
$p_T^{\text{jet}2}$	pTB2	✓	✓	✓
m_{jj}	mBB	✓	✓	✓
$\Delta R(\text{jet}_1, \text{jet}_2)$	dRBB	✓	✓	✓
$ \Delta\eta(\text{jet}_1, \text{jet}_2) $	dEtaBB	✓		
$\Delta\phi(V, H)$	dPhiVBB	✓	✓	✓
$\Delta\eta(V, H)$	dEtaVBB			✓
$M_{\text{eff}}(M_{\text{eff}3})$	HT	✓		
$\min(\Delta\phi(\ell, \text{jet}))$	dPhiLBmin		✓	
m_T^W	mTW		✓	
m_{ll}	mLL			✓
$\Delta Y(W, H)$	dYWH		✓	
m_{top}	mTop		✓	
Only in 3 Jet Events				
$p_T^{\text{jet}3}$	pTJ3	✓	✓	✓
m_{jjj}	mBBJ	✓	✓	✓

Table 28: Variables used to train the multivariate discriminant. Distributions of these inputs are included in Section 4.5.

797 boson. Both of these variables require p_z^γ to be determined, using m_W as a constraint to solve the quadratic
 798 equation

$$p_z^\gamma = \frac{1}{2(p_T^l)^2} \left[p_z^l X \pm E_l \sqrt{X^2 - 4(p_T^l)^2(E_T^{\text{miss}})^2} \right], \quad (4)$$

799 where

$$X = m_W^2 + 2p_x^l E_x^{\text{miss}} + 2p_y^l E_y^{\text{miss}}. \quad (5)$$

800 m_{top} is then reconstructed by selecting the jet from the two leading jets (as defined by the jet tagging
 801 strategy in Section 3.2.1), and solution to p_z^γ which minimises m_{top} . In the 2-tag case, both b -tagged jets
 802 are considered. If p_z^γ has an imaginary solution (i.e. $X^2 < 4(p_T^l)^2(E_T^{\text{miss}})^2$), the E_T^{miss} is shifted such that
 803 the discriminant is equal to zero.

804 The muon-in-jet and p_T -reco corrections are applied to the input variables based on jets p_T and directions.
 805 For the BDT training, in all three channels, pile-up reweighting is applied. Truth tagging is also applied
 806 to increase the MC statistics to get a more stable, optimal training.

807 5.2. Setup and training

808 The set of training parameters used for the BDT was optimized for the Run-1 analysis, a one-dimensional
 809 scan of each of the parameters was performed to obtain the optimal configuration shown in Table 29. It
 810 has been checked that this setup is nearly-optimal for the Run-2 analysis as well.

Not reviewed, for internal circulation only

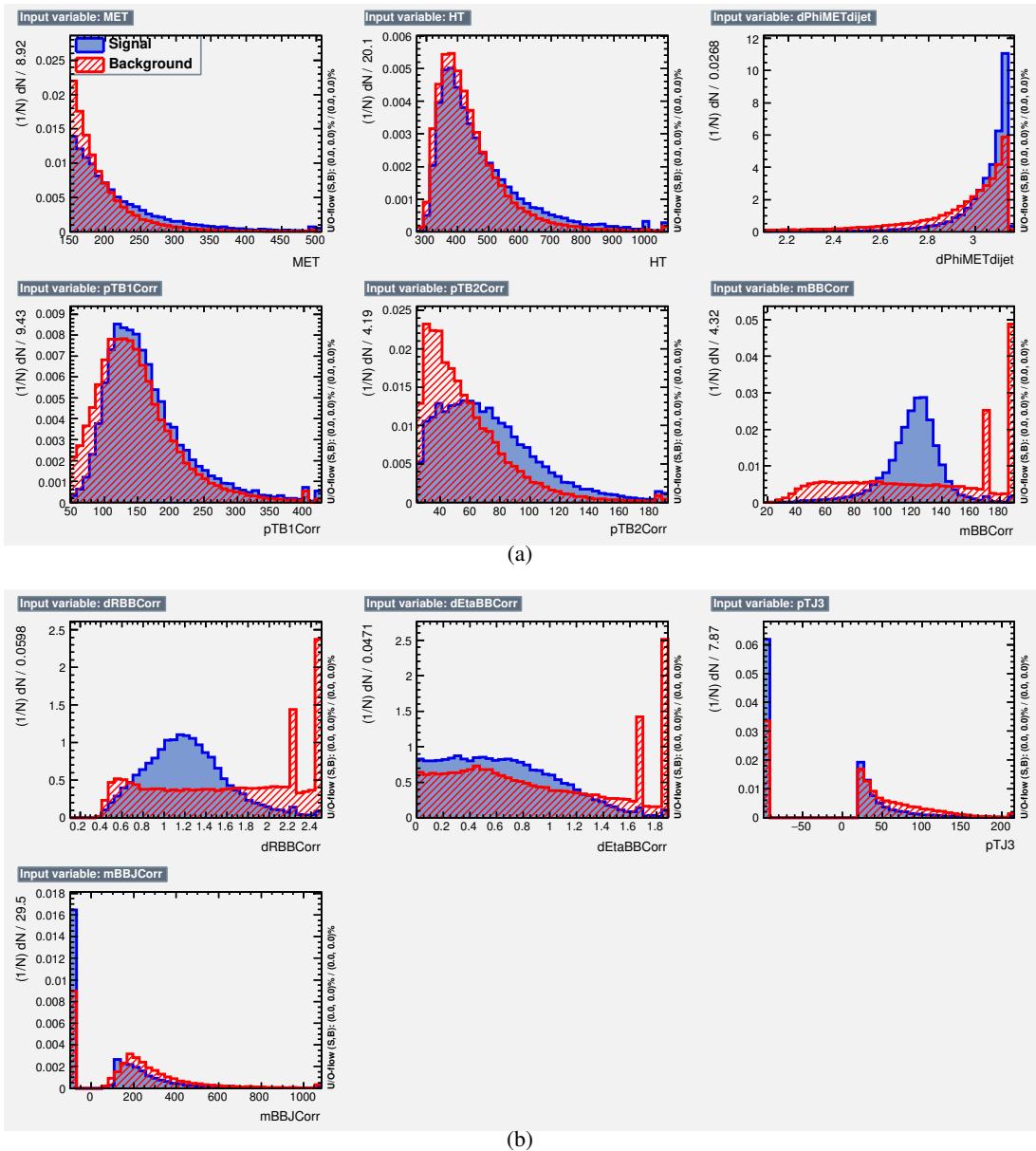


Figure 26: Input variables passed to the BDT for signal (blue) and background (red) samples in the 0-lepton channel are shown.

811 In Figures 28-29 (a) the expected background rejection as a function of signal efficiency, obtained after
 812 training the BDT in the 0,2-lepton channel, is shown, whereas in (b), the BDT distributions obtained while
 813 training (dots) and testing (histogram), for both signal and background, are found to be in reasonable
 814 agreement, indicating that no overtraining is affecting the procedure. Correlations between the input
 815 variables are shown in Figures 30-31 for the 0,2-lepton channel, for background and signal samples. For
 816 2-lepton channel, BDT inputs for three or more than three jet, $highp_T^V$ category are shown. Similar plots
 817 for other categories are shown in Appendix I.2.1.

Not reviewed, for internal circulation only

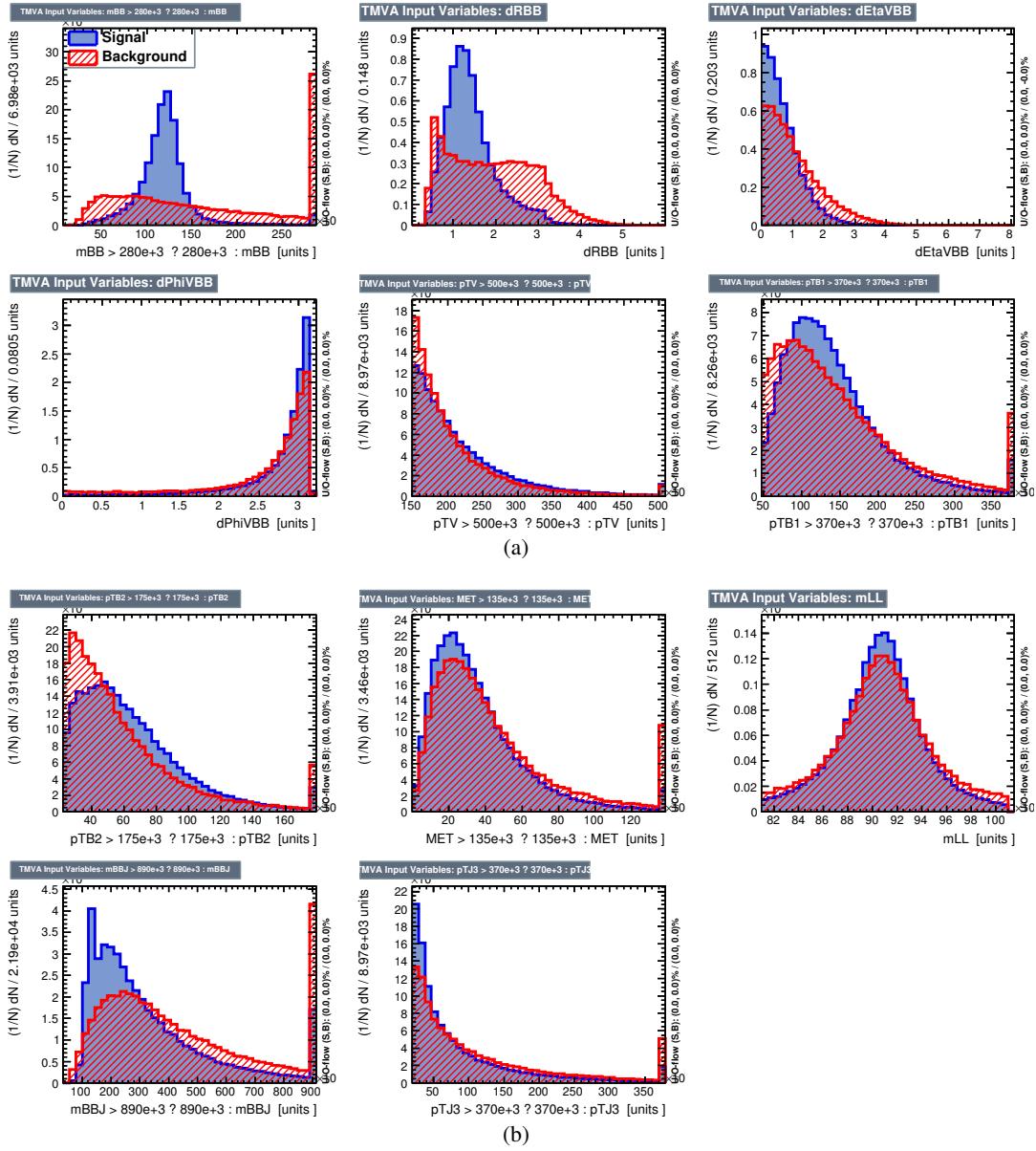


Figure 27: Input variables passed to the BDT for signal (blue) and background (red) samples in the 2-lepton channel are shown.

Not reviewed, for internal circulation only

TMVA Setting	Value	Definition
BoostType	AdaBoost	Boost procedure
AdaBoostBeta	0.15	Learning rate
SeparationType	GiniIndex	Node separation gain
PruneMethod	NoPruning	Pruning method
NTrees	200	Number of trees
MaxDepth	4	Maximum tree depth
nCuts	100	Number of equally spaced cuts tested per variable per node
nEventsMin	5%	Minimum number of events in a node

Table 29: BDT configuration parameters.

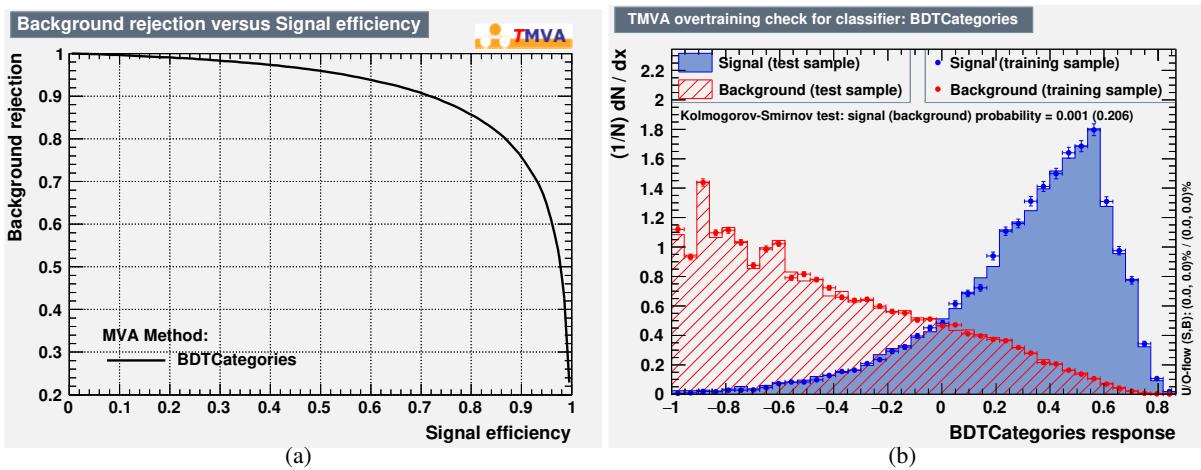


Figure 28: Distributions related to the BDT training procedure, in the 0-lepton channel, are shown. In (a) the expected background rejection as a function of signal efficiency is provided, and in (b), the BDT distributions obtained while training (dots) and testing (histogram), for both signal and background, are found to be in reasonable agreement.

Not reviewed, for internal circulation only

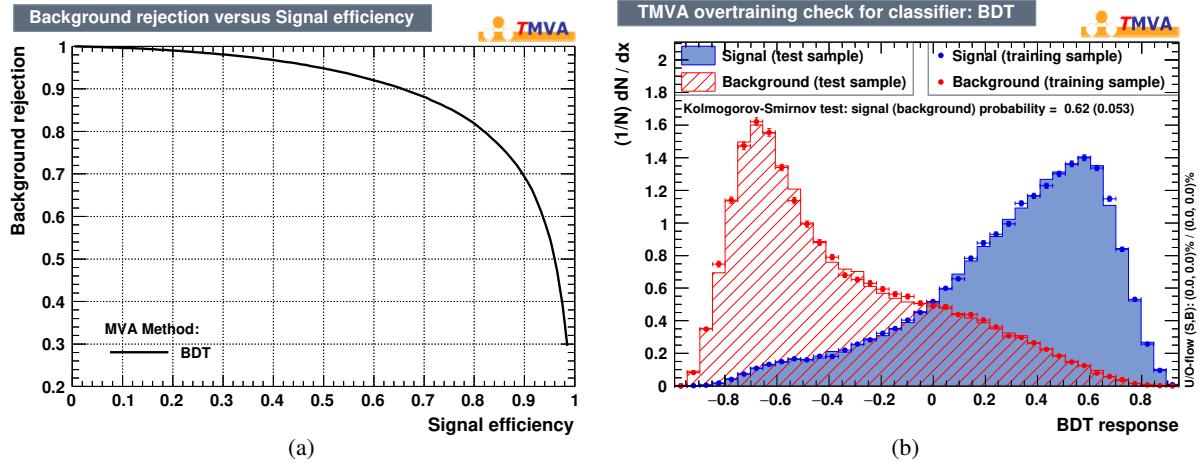


Figure 29: Distributions related to the BDT training procedure, in the 2-lepton channel, are shown. In (a) the expected background rejection as a function of signal efficiency is provided, and in (b), the BDT distributions obtained while training (dots) and testing (histogram), for both signal and background, are found to be in reasonable agreement.

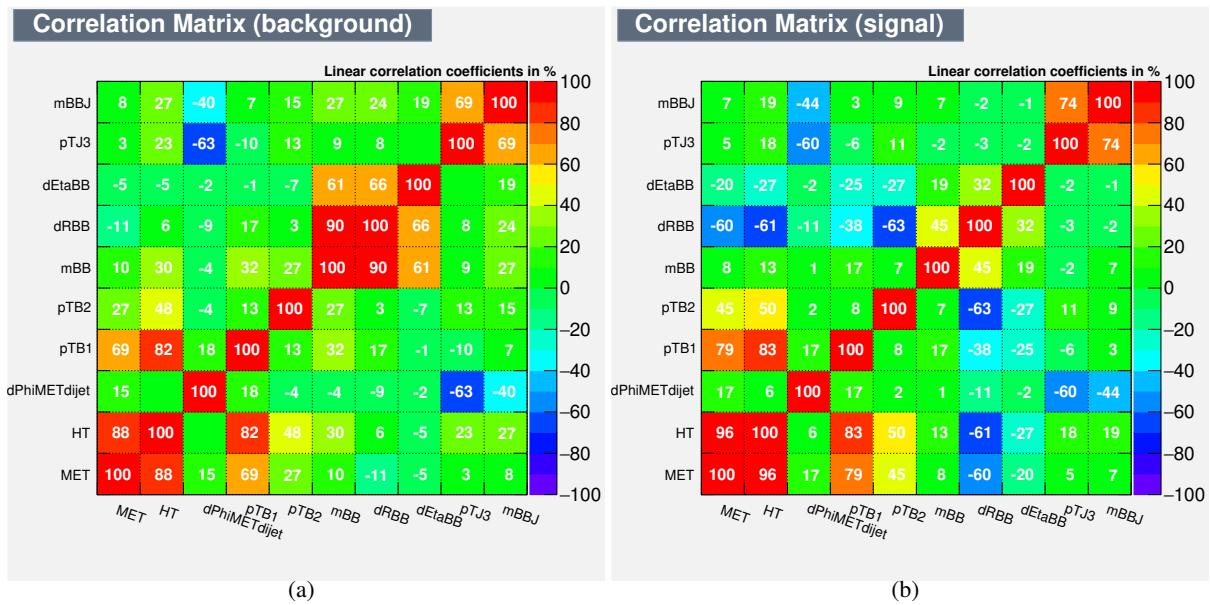


Figure 30: Correlation matrices of the background (a) and signal (b) input variables obtained in the BDT training procedure, in the 0-lepton channel, are shown.

Not reviewed, for internal circulation only

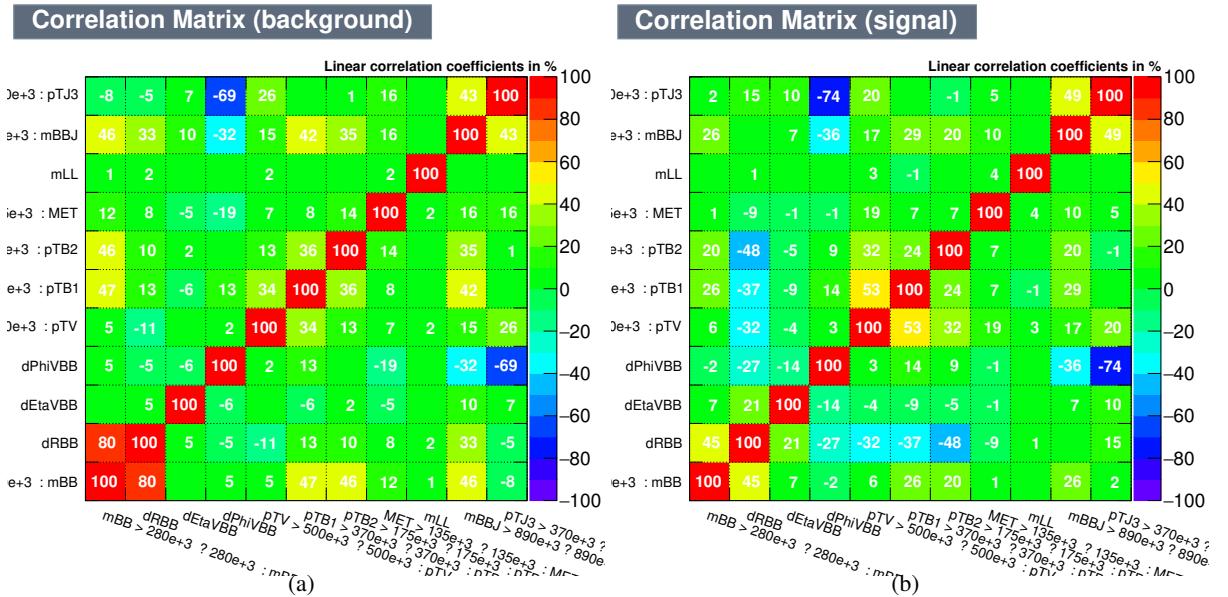


Figure 31: Correlation matrices of the background (a) and signal (b) input variables obtained in the BDT training procedure, in the 2-lepton channel, are shown.

818 5.3. BDT Transformation

819 Since the output of the BDT is designed to maximise the separation of the signal and background
 820 populations, the optimal performance is not necessarily achieved with the default binning. For example,
 821 in the dijet mass tails, wider bins are required to reduce statistical uncertainty, but this is at a cost to the
 822 BDT sensitivity. Therefore, a transformation of the BDT output, designed to optimise the effect of the
 823 final sensitivity and reduction in the number of bins, is implemented.

824 As a general description, to remap the histograms entering the final fit, consider the function:

$$Z(I[k, l]) = Z(z_s, n_s(I[k, l]), N_s, z_b, n_b(I[k, l]), N_b), \quad (6)$$

825 where

- 826 • $I[k, l]$ is an interval of the histograms, containing the bins between bin k and bin l ;
- 827 • N_s is the total number of signal events in the histogram;
- 828 • N_b is the total number of background events in the histogram;
- 829 • $n_s(I[k, l])$ is the total number of signal events in the interval $I[k, l]$;
- 830 • $n_b(I[k, l])$ is the total number of background events in the interval $I[k, l]$;
- 831 • z_s and z_b are parameters used to tune the algorithm.

832 Whilst several different possible Z functions exist to transform the BDT output, in the results below,
 833 only **Transformation D** is considered:

$$Z = z_s n_s / N_s + z_b n_b / N_b. \quad (7)$$

834 The re-binning is then conducted using the following algorithm:

- 835 1. Starting from the last bin on the right of the original histogram, increase the range of the interval
 $I(k, \text{last})$ by adding one after the other, the bins from the right to the left;
- 836 2. Calculate the value of Z at each step;
- 837 3. Once $Z(I[k_0, \text{last}]) > 1$, rebin all the bins in the interval $I(k_0, \text{last})$ into a single bin;
- 838 4. Repeat steps 1-3, starting this time from the last bin on the right, not included in the previous remap
 (the new last is $k_0 - 1$), until k_0 in the first bin.

841 In Run-1, the implementation of **Transformation D** was found to offer a significant decrease in the
 842 number of bins, whilst comparatively increasing the expected sensitivity [27]. For the current analysis
 843 these sensitivity studies have been repeated and find $z_s = 10, z_b = 5$ as optimal parameters. Additionally
 844 we require that the MC statistical uncertainty in each bin is $< 20\%$ to minimise possible biases in the fit,
 845 see appendix F.

846 5.4. Optimization studies and performances

847 In this section, the optimization studies and performances of the BDTs in each lepton channel are shown.

848 5.4.1. 0-lepton MVA performance

849 In the 0-lepton channel, the MVA is trained using VH signal MC samples, and background MC samples
 850 ($V+jets$, $t\bar{t}$, single top, and diboson) for 2-tag events only and in each n-jet category considered, 2 jets and
 851 3 jets, separately.

852 Figure 32 shows the BDT output distributions in the 2-tag, 2 and 3 jet categories, where transformation D
 853 ($z_s = 10$, $z_b = 5$), described in Section 5.3, is used to rebin. Signal-like events correspond to BDT output
 854 values close to 1, and more background-like events are found when going to lower values.

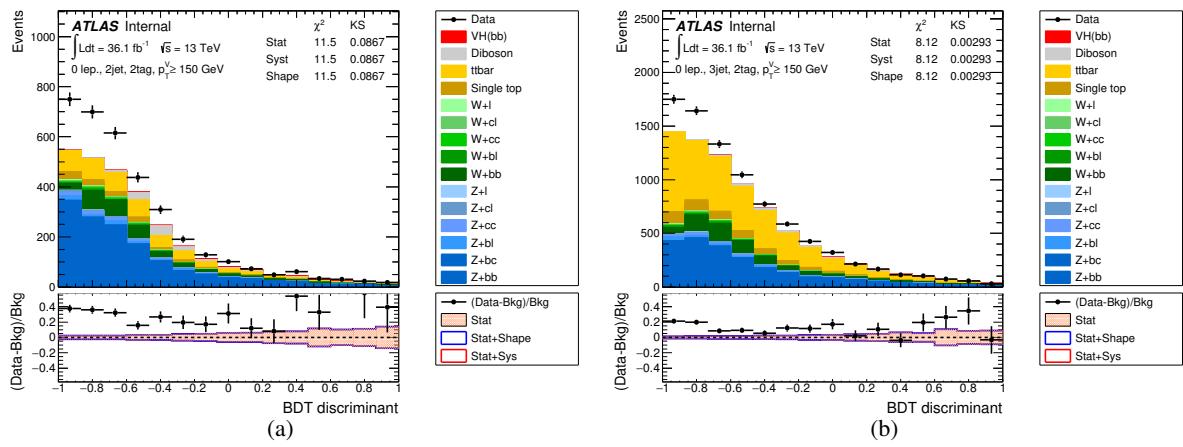


Figure 32: The BDT distributions in the 0-lepton channel for 2-tag events, in the (a) 2-jet and (b) 3-jet categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

855 Table 30 shows the comparison of significance between a BDT-based and a simple m_{bb} cut-based analysis.
 856 For the cut-based analysis, a slightly tighter signal selection is applied, as described in Section 3.2.5. The
 857 significance is estimated using a binned likelihood ratio, no systematic uncertainties are considered in
 858 the computation. The BDT-based analysis is seen to offer an improvement of about 20% in significance
 859 compared to the cut-based using pre-fit yields.

Category	binned log-likelihood ratio		
	MVA	$m_{bb} \ 150 < E_T^{\text{miss}} < 200 \text{ GeV}$	$m_{bb} \ 200 < E_T^{\text{miss}} \text{ GeV}$
2 jet	2.43 ± 0.06	1.16 ± 0.04	1.80 ± 0.06
3 jet	1.62 ± 0.04	0.71 ± 0.01	1.11 ± 0.03
combination	2.92 ± 0.05	2.51 ± 0.04	

Table 30: Expected significance of BDT-based and cut-based analyses in the 0-lepton channel, estimated using a binned log-likelihood ratio using prefit distributions with backgrounds normalized to the SM crossections (no fit scale factors). No systematic uncertainties are included here.

Not reviewed, for internal circulation only

860 5.4.2. 1-lepton MVA performance

861 In the 1-lepton channel, the MVA is trained on VH signal samples and background samples ($V+jets$, $t\bar{t}$,
862 single top, ZZ , WZ and WW) for each signal category, presently 2 jets and 3 jets, for $p_T^W > 150$ GeV.

863 Figure 33 shows the BDT output distributions in the 2-tag, 2 and 3 jet categories, where the BDT output is
864 rebinned using the transformation D ($z_s = 10$, $z_b = 5$), as described in Section 5.3. These figures clearly
865 indicate the discriminating power of the BDT, where an enhanced signal purity is seen for BDT output
866 values close to 1. The background prediction of BDT output is in good agreement with the observed
867 data.

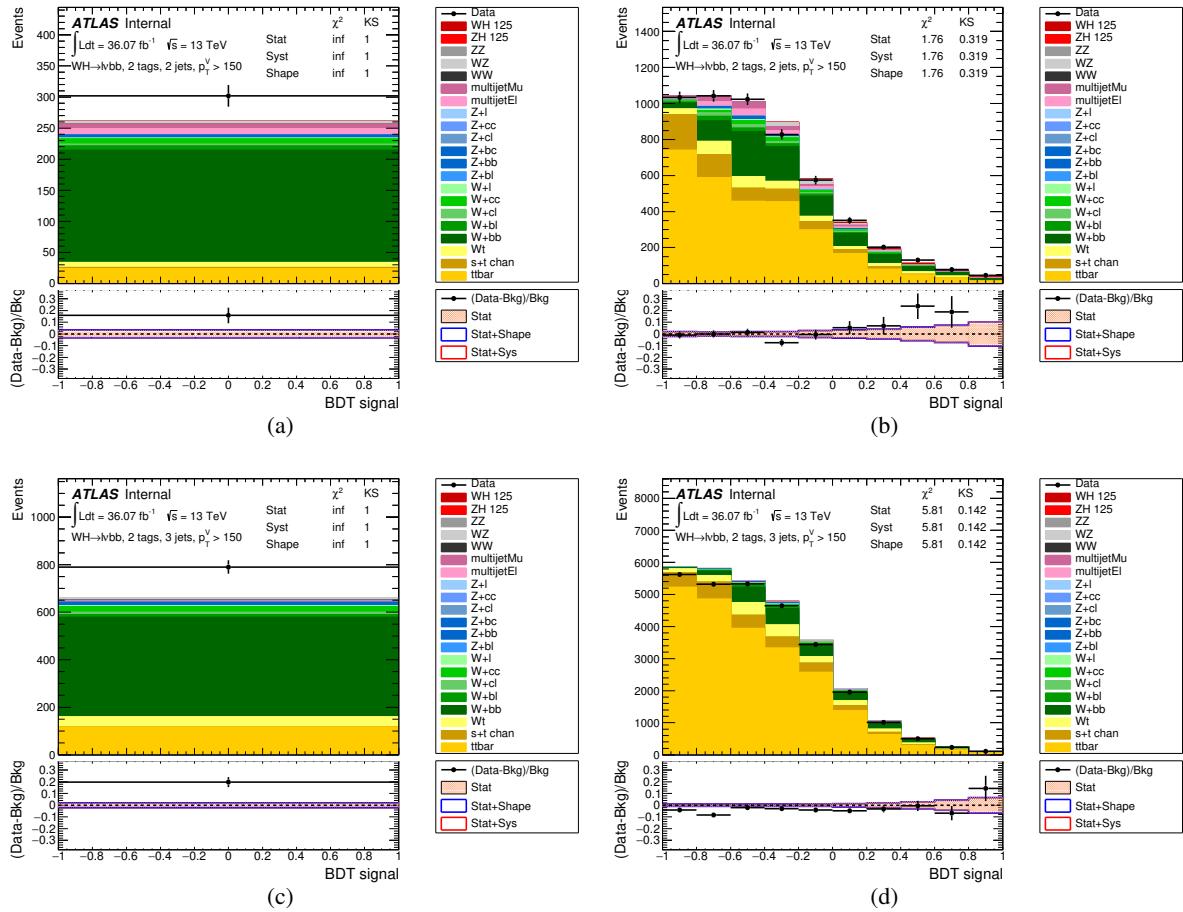


Figure 33: The $p_T^W > 150$ GeV BDT distributions in the 1-lepton channel for 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

868 Table 31 shows the comparison of significance between BDT-based and m_{bb} -based analysis. The signifi-
869 ficance calculations, performed using a simple binned likelihood ratio calculation where no systematic
870 uncertainties are included, for the BDT analysis use truth-tagging for the input MC $V+jets$ samples, whilst
871 all other samples use direct tagging.

Category	binned log-likelihood ratio		
	MVA	$m_{bb} \ 150 < p_T^V < 200 \text{ GeV}$	$m_{bb} \ 200 < p_T^V \text{ GeV}$
2 jet	2.42 ± 0.07	0.93 ± 0.01	1.79 ± 0.06
3 jet	1.40 ± 0.03	0.41 ± 0.01	0.86 ± 0.02
combination	2.79 ± 0.06		2.27 ± 0.05

Table 31: Expected significance of BDT-based and cut-based analyses in the 1-lepton channel, estimated using a binned log-likelihood ratio. Simulated samples are normalised to SM cross-sections. No systematic uncertainties are included in this computation.

Not reviewed, for internal circulation only

872 5.4.3. 2-lepton MVA performance

873 In the 2 lepton channel, the MVA(BDT) is trained on signal samples ($qq \rightarrow ZH \rightarrow \ell\ell b\bar{b}$, $gg \rightarrow ZH \rightarrow \ell\ell b\bar{b}$, and background samples ($Z+jets$, $t\bar{t}$, single top, ZZ and WZ). The weighted sum of all background 874 samples forms background distributions for BDT training. The training has been separately performed 875 for each category (2-jets and 3+jets, middle p_T^Z and high p_T^Z). 876

877 Figures 34–35 show the BDT output distributions in the 2-tag category for each p_T^Z category, respectively. 878 The BDT output is transformed with the transformation D ($z_s = 10$, $z_b = 5$), as described in Section 5.3. 879 These figures clearly indicate the discriminating power of the BDT, where an enhanced signal purity is 880 seen for BDT output values close to 1, while the low BDT output values are dominated by background 881 processes. BDT output distributions in $e\mu$ control region are shown in Figure 36–37. In the $e\mu$ control 882 region, The background prediction of the BDT output is in good agreement with the observed data.

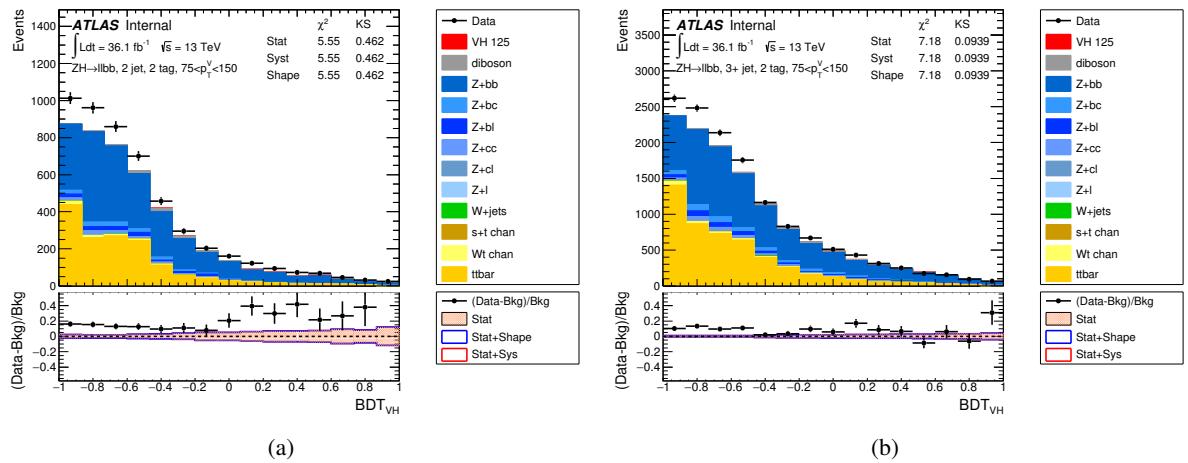


Figure 34: Transformed BDT output distribution in the category of 2 b -tagged(2 jets), $75 < p_T^Z < 150$ GeV (a), transformed BDT output distribution in the category of 2 b -tagged(3+ jets), $75 < p_T^Z < 150$ GeV (b).

883 Table 32 shows the comparison of significance between BDT-based and m_{bb} -based analysis. For m_{bb} - 884 based analysis, tighter signal selections have been applied (see Section 3.2). The significance is evaluated 885 in quadrature over each bin to take into account shape information. The BDT-based analysis improves 886 significance by about 23% compared to m_{bb} -based analysis.

Not reviewed, for internal circulation only

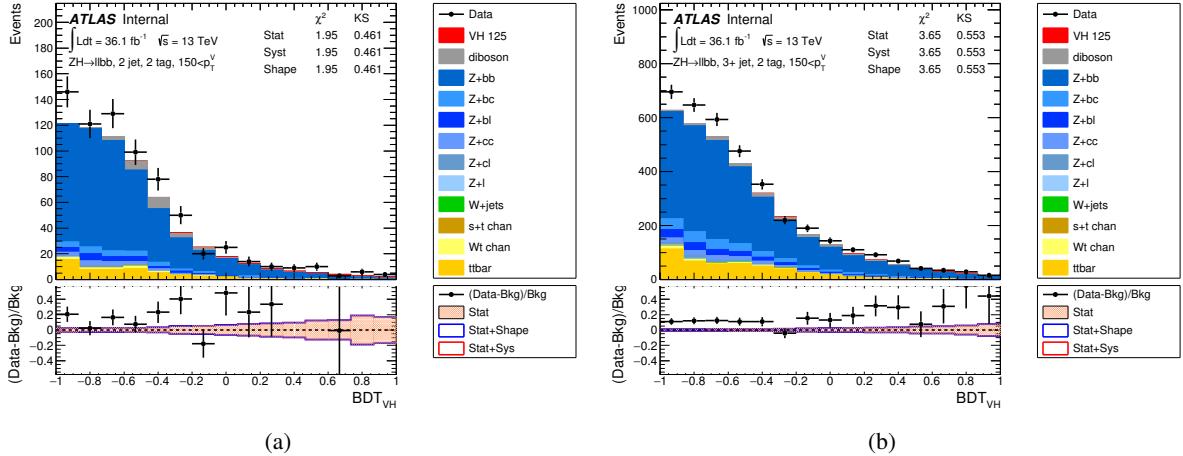


Figure 35: Transformed BDT output distribution in the category of 2 b -tagged(2 jets), $p_T^Z < 150$ GeV (a), transformed BDT output distribution in the category of 2 b -tagged(3+ jets), $p_T^Z > 150$ GeV (b).

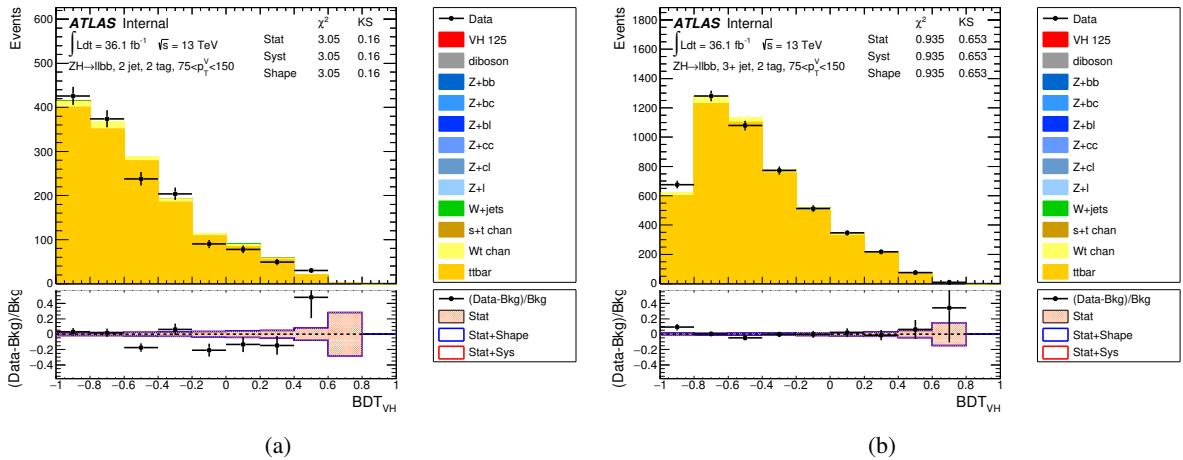


Figure 36: Raw BDT output distribution in the $e\mu$ control region of (a) : 2 b -tagged(2 jets), $75 < p_T^Z < 150$ GeV, (b) 2 b -tagged(3+ jets) . Note : transformation does not work in current configuration. The distribution is before transformation.

Not reviewed, for internal circulation only

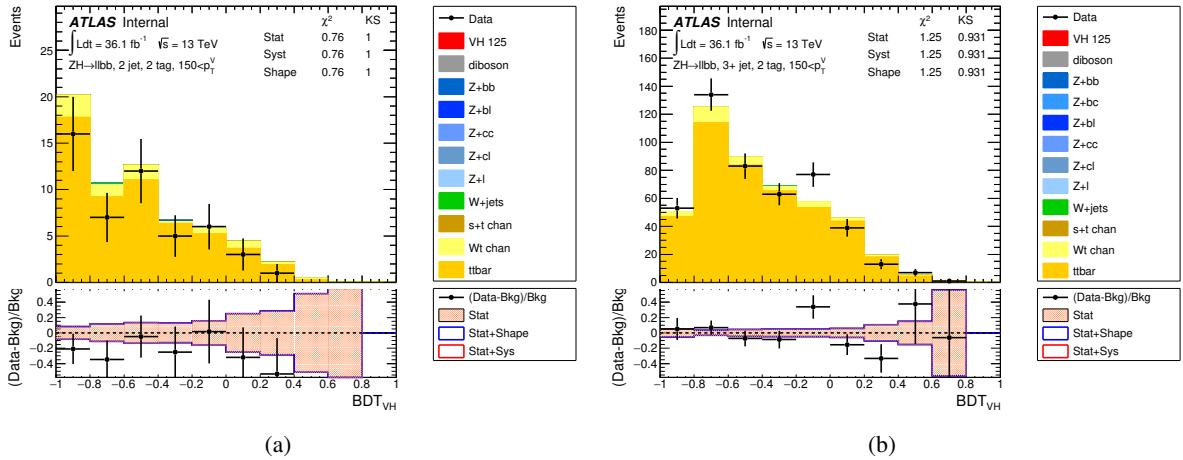


Figure 37: Raw BDT output distribution in the $e\mu$ control region of 2 b -tagged(2 jets), $p_T^Z < 150$ GeV (a), in the category of 2 b -tagged(3+ jets), $p_T^Z > 150$ GeV (b). Note : transformation does not work in current configuration. The distribution is before transformation.

Category	binned log-likelihood ratio	
Category	MVA	m_{bb}
2jet, (p_T^Z 0-75 GeV), mva(m_{bb})	0.753 ± 0.016	0.551 ± 0.011
3+jet, (p_T^Z 0-75 GeV), mva(m_{bb})	0.570 ± 0.008	0.411 ± 0.005
2jet, (p_T^Z 75-150 GeV), mva(m_{bb})	1.039 ± 0.03	0.774 ± 0.015
3+jet, (p_T^Z 75-150 GeV), mva(m_{bb})	0.953 ± 0.014	0.689 ± 0.007
2jet, (p_T^Z 150- GeV), mva	1.634 ± 0.051	
3+jet, (p_T^Z 150- GeV), mva	1.523 ± 0.028	
2jet, (p_T^Z 150-200 GeV), m_{bb}		0.835 ± 0.028
3+jet, (p_T^Z 150-200 GeV), m_{bb}		0.722 ± 0.014
2jet, (p_T^Z 200- GeV), m_{bb}		1.271 ± 0.097
3+jet, (p_T^Z 200- GeV), m_{bb}		0.929 ± 0.024
Combined mva(m_{bb})	2.81 ± 0.035	2.29 ± 0.056

Table 32: Significance comparison between BDT-based analysis and m_{bb} -based analysis in the 2-lepton channel. Transformation D with $z_s = 10$, $z_b = 5$ is used for BDT output transformation. No systematic uncertainties are included and signal yield is normalized to the SM prediction ($\mu = 1$) in these computations.

6. Experimental Systematic Uncertainties

In this section the experimental systematic uncertainties that are taken into account in this analysis are summarized. Detailed information on the systematic uncertainties listed in Table 33 can be found in Ref [15]. Modelling systematic uncertainties are discussed in detail in Ref [16] and are summarized once more in the next section.

Not reviewed, for internal circulation only

¹⁰ Due to bug in the MCP tool in AnalysisBase, the uncertainty for muons with $p_T < 10$ GeV is incorrect. A fixed 0.2% uncertainty is applied to these muons.

Not reviewed, for internal circulation only

Systematic uncertainty	Short description	Reference
	Event	
Luminosity	uncertainty on total integrated luminosity	Section 11.1 in Ref. [15]
Pileup Reweighting	uncertainty on pileup reweighting	Section 11.1 in Ref. [15]
	Electrons	
EL_EFF_Trigger_Total_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty	Section 11.2.2. in Ref. [15]
EL_EFF_Reco_Total_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty	Section 11.3.1. in Ref. [15]
EL_EFF_ID_Total_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty	Section 11.3.1. in Ref. [15]
EL_EFF_Iso_Total_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty	Section 11.3.1. in Ref. [15]
EG_SCALE_ALL	energy scale uncertainty	Section 11.3.2. in Ref. [15]
EG_RESOLUTION_ALL	energy resolution uncertainty	Section 11.3.2. in Ref. [15]
	Muons	
MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainty	Section 11.2.2. in Ref. [15]
MUON_EFF_TrigSystUncertainty	reconstruction and ID efficiency uncertainty for muons with $p_T > 15$ GeV	Section 11.4.1. in Ref. [15]
MUON_EFF_STAT	reconstruction and ID efficiency uncertainty for muons with $p_T < 15$ GeV	Section 11.4.1. in Ref. [15]
MUON_EFF_SYS	isolation efficiency uncertainty	Section 11.4.1. in Ref. [15]
MUON_EFF_STAT_LOWPT	track-to-vertex association efficiency uncertainty	Section 11.4.1. in Ref. [15]
MUON_EFF_SYST_LOWPT	momentum resolution uncertainty from inner detector	Section 11.4.2. in Ref. [15]
MUON_ISO_STAT	momentum resolution uncertainty from muon system	Section 11.4.2. in Ref. [15]
MUON_ISO_SYS	momentum scale uncertainty	Section 11.4.2. in Ref. [15]
MUON_TTVA_STAT ¹⁰	charge dependent momentum scale uncertainty	Section 11.4.2 in Ref. [15]
MUON_TTVA_SYS ¹⁰		
MUON_ID		
MUON_MS		
MUON_SCALE		
MUON_SAGITTA_RHO		
MUON_SAGITTA_RESBIAS		
	Jets	
JET_21NP_JET_EffectiveNP_1	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EffectiveNP_2	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EffectiveNP_3	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EffectiveNP_4	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EffectiveNP_5	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EffectiveNP_6	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EffectiveNP_7	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EffectiveNP_8restTerm	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EtaIntercalibration_Modeling	energy scale uncertainty on eta-intercalibration (modeling)	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EtaIntercalibration_TotalStat	energy scale uncertainty on eta-intercalibrations (statistics/method)	Section 11.5.1. in Ref. [15]
JET_21NP_JET_EtaIntercalibration_NonClosure	energy scale uncertainty on eta-intercalibrations (non-closure)	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Pileup_OffsetMu	energy scale uncertainty on pile-up (mu dependent)	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Pileup_OffsetNPV	energy scale uncertainty on pile-up (NPV dependent)	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Pileup_PtTerm	energy scale uncertainty on pile-up (pt term)	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Pileup_RhoTopology	energy scale uncertainty on pile-up (density ρ)	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Flavor_Composition_Zjets	energy scale uncertainty on Z+jets sample's flavour composition	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Flavor_Composition_Wjets	energy scale uncertainty on W+jets sample's flavour composition	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Flavor_Composition_top	energy scale uncertainty on top sample's flavour composition	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Flavor_Composition	energy scale uncertainty on VV and VH sample's flavour composition	Section 11.5.1. in Ref. [15]
JET_21NP_JET_Flavor_Response	energy scale uncertainty on samples' flavour response	Section 11.5.1. in Ref. [15]
JET_21NP_JET_BJES_Response	energy scale uncertainty on b-jets	Section 11.5.1. in Ref. [15]
JET_21NP_JET_PunchThrough_MC15	energy scale uncertainty for punch-through jets	Section 11.5.1. in Ref. [15]
JET_21NP_JET_SingleParticle_HighPt	energy scale uncertainty from the behaviour of high-pT jets	Section 11.5.1. in Ref. [15]
JET_JER_SINGLE_NP	energy resolution uncertainty	Section 11.5.1. in Ref. [15]
JET_JvtEfficiency	JVT efficiency uncertainty	Section 11.5.1 in Ref. [15]
FT_EFF_Eigen_B	b-tagging efficiency uncertainties ("BTAG_MEDIUM"): 3 components for b jets, 3 for c jets and 5 for light jets	Section 11.7. in Ref. [15]
FT_EFF_Eigen_C		
FT_EFF_Eigen_L		
FT_EFF_Eigen_extrapolation	b-tagging efficiency uncertainty on the extrapolation to high- p_T jets	Section 11.7. in Ref. [15]
FT_EFF_Eigen_extrapolation_from_charm	b-tagging efficiency uncertainty on tau jets	Section 11.7. in Ref. [15]
	MET	
METTrigStat	trigger efficiency uncertainty	Section 11.2.1. in Ref. [15]
METTrigTop/Z	track-based soft term related longitudinal resolution uncertainty	Section 11.6. in Ref. [15]
MET_SoftTrk_ResoPara	track-based soft term related transverse resolution uncertainty	Section 11.6. in Ref. [15]
MET_SoftTrk_ResoPerp	track-based soft term related longitudinal scale uncertainty	Section 11.6. in Ref. [15]
MET_SoftTrk_Scale	track MET scale uncertainty due to tracks in jets	Section 11.6. in Ref. [15]

Table 33: Summary of the experimental systematic uncertainties considered. Details on the individual systematic uncertainties can be found in the given Sections of Ref. [15].

892 7. Statistical Treatment

893 The signal yield is estimated using a binned likelihood fit to the BDT discriminant distribution, obtained
 894 as described in Section 5. A similar fit is performed on the m_{bb} distribution, and provides a simple
 895 cross-check to the nominal result. Both fits are performed in all the analysis categories simultaneously,
 896 combining the 0-, 1- and 2-lepton channels. The definition of the likelihood and the configuration of the
 897 fit in terms of inputs and parameters are given in the following sections.

898 7.1. Likelihood function definition

899 The statistical analysis of the data uses a binned likelihood function built as the product of Poisson
 900 probability terms,

$$\text{Pois}(n|\mu S + B) \left[\prod_{b \in \text{bins}} \frac{\mu v_b^{\text{sig}} + v_b^{\text{bkg}}}{\mu S + B} \right], \quad (8)$$

901 where n is the total number of data events, and S and B represent the total number of expected signal and
 902 background events, respectively. The parameter of interest μ , common to all channels and categories, is
 903 the signal strength, defined as the ratio between the measured and the expected signal cross-sections. It is
 904 multiplied by the expected signal and background yields, v_b^{sig} and v_b^{bkg} respectively, in each histogram bin
 905 b .

906 The dependence of the signal and background predictions on the systematic uncertainties is described by
 907 a set of nuisance parameters (NP) θ , which are parametrized by Gaussian priors; the latter are used for
 908 maintaining the normalisation positive and thus the likelihood positive. The expected numbers of signal
 909 and background events in each bin are functions of θ and parametrized such that the rates in each category
 910 are log-normally distributed for a normally distributed θ .

911 The priors act to constrain the NPs to their nominal values within their assigned uncertainties. They are
 912 implemented via so-called penalty or auxiliary measurements added to the likelihood which will always
 913 increase when any nuisance parameter is shifted from the nominal value. The likelihood function, $\mathcal{L}(\mu, \theta)$,
 914 is therefore a function of μ and θ .

915 The nominal fit result in terms of μ and σ_μ is obtained by maximising the likelihood function with respect
 916 to all parameters. This is referred to as the maximised log-likelihood value, MLL. The profile likelihood
 917 ratio test statistic, q_μ , is then constructed as follows:

$$q_\mu = -2 \ln(\mathcal{L}(\mu, \hat{\theta}_\mu) / \mathcal{L}(\hat{\mu}, \hat{\theta})), \quad (9)$$

918 where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximise the likelihood (with the constraint $0 \leq \hat{\mu} \leq \mu$), and $\hat{\theta}_\mu$
 919 are the nuisance parameter values that maximise the likelihood for a given μ . This test statistic is used to
 920 measure the compatibility of the background-only model with the observed data, extracting the local p_0
 921 value, and, if no hint of a signal is found in this procedure, for the derivation of exclusion intervals using
 922 the CL_s method [28, 29].

923 7.2. Fit inputs

924 The observable used in the fit to separate signal from background is the BDT discriminant, optimised as
 925 described in Section 5. The following processes are considered in the fit:

- 926 • signal: $Vh \rightarrow Vb\bar{b}$, with $V = Z, W$ decaying to $\ell\ell, \ell\nu$, or $\nu\nu$ final states;
- 927 • $Z/W +$ jets, split into flavour components as described in Section 3.2;
- 928 • $t\bar{t}$
- 929 • single-top: s -, t -, Wt -channel;
- 930 • diboson: WW, WZ, ZZ ;
- 931 • multijet background (for the 1-lepton channel only).

932 Truth-tagging is used for the WW , $V + cl$, cc , and l samples, in the 2 tag regions, to reduce the MC
 933 statistical uncertainties. Direct tagging is employed for all other samples and regions. The fit is performed
 934 simultaneously in the 0-, 1- and 2-lepton channels, and the BDT discriminant defined in Section 5 is used.
 935 The event categories used in the fit are summarized in Table 34, and were optimized to maximize the
 936 expected significance, while prioritizing the model simplicity and the limited MC statistical samples.

Channel Regions	0-lep SR	SR	1-lep $W + hf$ (Sec.3.2) CR	SR	2-lep $(e - \mu)$ (Sec.3.2) CR
p_T^V [GeV]	≥ 150		≥ 150		75-150, >150
# of jets	2,3		2,3		2, 3+
# of b-tag	2		2		2
Discriminant	BDT output	BDT output	one bin	BDT output	m_{bb} ¹

Table 34: Categories of events used in the fit for the 0-, 1- and 2-lepton channels.

937 7.3. Nuisance parameters: normalisation and systematic uncertainties

938 As stated above, systematic uncertainties in the analysis are modelled with nuisance parameters in the
 939 fit. Two different types of nuisance parameters are considered: unconstrained normalisation factors
 940 and parameters associated to systematics that have external constraints (implemented as priors on the
 941 NP's distribution function). The priors are identified with experimental or modelling uncertainties. All
 942 experimental systematic uncertainties listed in Table 33 are included in the fit; the considered modelling
 943 systematic uncertainties are briefly summarized in the following section. A detailed description of their
 944 motivation and derivation can be found in Ref. [16].

945 The fit model contains eight freely-floating normalisation parameters, for the signal as well as the main
 946 backgrounds, that are fitted in the signal and control regions described above:

947 μ : SM VH signal strength;

¹ still use 1-bin m_{bb} shape in 2tag2jet high- p_T^Z region due to lack of statistics.

Nuisance Parameter	Description	Samples/Categories	Value	Effect
norm_Zbb_J2	Zhf	Zhf, 2jet, all regions in combined fit, 0- and 2-lepton channels for per-channel fits	Float	Normalisation
norm_Zbb_J3	Zhf	Zhf, 3jet, all regions in combined fit, 0- and 2-lepton channels for per-channel fits	Float	Normalisation
SysZlNorm	Zl normalisation	Zl, all regions	18%	Normalisation
SysZclNorm	Zcl normalisation	Zcl, all regions	23%	Normalisation
SysZbbNorm_L0	Zhf normalisation	Zhf normalisation in 0-lepton	7%	Normalisation
SysZbcZbbRatio	Zbc/Zbb ratio	Zbc, all regions	40% (0-l), 40% (2-l, 2-j), 30% (2-j, 3+j)	Normalisation
SysZblZbbRatio	Zbl/Zbb ratio	Zbl, all regions	25% (0-l), 28% (2-l, 2-j), 20% (2-j, 3+j)	Normalisation
SysZccZbbRatio	Zcc/Zbb ratio	Zcc, all regions	15% (0-l), 16% (2-l, 2-j), 13% (2-j, 3+j)	Normalisation
SysZPtV	p_T^V shape	Z + jets samples, all regions	-	Normalisation + Shape
SysZmbb	m_{bb} shape	Z + jets samples, all regions	-	Shape

Table 35: Summary of the $Z +$ jets systematic uncertainties: the first column quotes the name of the nuisance parameter implemented in the fit referring to a specific systematic uncertainty, the second column the source of the uncertainty, the third the categories and sample on which it is applied, the fourth column the value of the Gaussian prior on the NP (if applicable) and the fifth column the type of systematic uncertainty. The listed systematic uncertainties are separated in normalisation effects (first block), acceptance effects (second block), and shape systematic uncertainties (third block).

948 **norm_Zbb_J2, norm_Zbb_J3:** normalisation of $Z+{\rm jets}$ (bb, bc, bl, cc) background (decorrelated
949 between 2 and 3(+) jet);

950 **norm_Wbb_J2, norm_Wbb_J3:** normalisation of $W+{\rm jets}$ (bb, bc, bl, cc) background;

951 **norm_ttbar, norm_ttbar_J2_L2, norm_ttbar_J3_L2:** ttbar normalisation (decorrelated between
952 0+1 and 2-lepton (2- and 3+-jet)).

953 Here and in the following, the expression “heavy-flavour jets” (abbreviated “hf”) is used in case of $Z +$
954 jets and $W +$ jets events for the sum of the bb, bc, cc and bl components.¹ On top of these eight floating
955 parameters, additional, constrained, nuisance parameters are introduced for the signal and background
956 components to provide control over the relative acceptance of the different categories and samples in the
957 analysis (across lepton channels, jet multiplicity categories, different flavour components of $Z/W+{\rm jets}$).
958 Priors on these nuisance parameters are obtained from data-MC comparison or MC studies, and reflect
959 our knowledge on these relative acceptances variations from MC comparisons.

960 Table 35, 36, 37, and 38 summarize all modelling nuisance parameters considered for the $Z +$ jets, $W +$
961 jets, $t\bar{t}$, as well as the signal and the smaller backgrounds, respectively. Besides the names of the nuisance
962 parameters (corresponding to the names used in pull plots etc.) a short description of to which samples
963 / regions they apply, whether they affect the normalisation and/or shape and the value of the Gaussian
964 prior (if applicable) is given. In Section 6, reconstruction systematic uncertainties considered in the fit
965 are discussed. For details on the evaluation of these systematic uncertainties, please read Ref. [16].

966 7.3.1. Smoothing of the Systematic Uncertainties

967 The uncertainties on reconstructed objects are propagated in the analysis in two different ways: by shifting
968 weights or by re-selecting events. For flavour tagging, where a scale factor is used to correct the simulation
969 efficiency to data, this weight is shifted up (down) and the change in the final distribution is noted as the
970 $+1$ (-1) σ shift. For jet energy scale (JES) uncertainties, the jet energies are shifted and therefore events
971 can migrate in and out of the acceptance. Again the difference in the final variable is noted as the 1σ error
972 but if the variations are small and/or the sample statistics are small, the MC statistical uncertainty can

¹ Due to historical reasons some nuisance parameter names contain bb instead of “hf”, but meaning the latter, e.g. `norm_Zbb`; each such case is indicated when the respective nuisance parameter is introduced.

Not reviewed, for internal circulation only

Nuisance Parameter	Description	Samples/Categories	Value	Effect
<code>norm_Wbb_J2</code>	Whf , 2jet normalisation	Whf	Float	Normalisation
<code>norm_Wbb_J3</code>	Whf , 3jet normalisation	Whf	Float	Normalisation
<code>WlNorm</code>	Wl normalisation	Wl , all regions	32%	Normalisation
<code>WclNorm</code>	Wcl normalisation	Wcl , all regions	37%	Normalisation
<code>SysWbbNorm_DWhfCR_L1</code>	Whf normalisation	Whf CR to SR	10%	Normalisation
<code>SysWbbNorm_L0</code>	Whf normalisation	Whf , 0-lepton channel	5%	Normalisation
<code>SysWbcWbbRatio</code>	Wbc/Wbb ratio	Wbc , all regions	15% (0-l), 30% (1-l)	Normalisation
<code>SysWb1WbbRatio</code>	Wcl/Wbb ratio	Wcl , all regions	26% (0-l), 23% (1-l)	Normalisation
<code>SysWccWbbRatio</code>	Wcc/Wbb ratio	Wcc , all regions	10% (0-l), 30% (1-l)	Normalisation
<code>SysWPtV</code>	p_T^V shape	$W + \text{jets}$ samples, all regions	-	Shape
<code>SysWMbb</code>	m_{bb} shape	$W + \text{jets}$ samples, all regions	-	Shape

Table 36: Summary of the $W+\text{jets}$ systematic uncertainties: the first column quotes the name of the nuisance parameter implemented in the fit referring to a specific systematic uncertainty, the second column the source of the uncertainty, the third the categories and sample on which it is applied, the fourth column the value of the Gaussian prior on the NP (if applicable) and the fifth column the effect of the systematic uncertainty. The listed systematic uncertainties are separated in normalisation effects (first block), acceptance effects (second block), and shape systematic uncertainties (third block).

Nuisance Parameter	Description	Categories	Value	Effect
<code>norm_ttbar</code>	$t\bar{t}$ normalisation	0- and 1-lepton channels	Float	Normalisation
<code>norm_ttbar_L0</code>	$t\bar{t}$ normalisation	0-lepton	8%	Normalisation
<code>norm_ttbar_J2</code>	$t\bar{t}$ normalisation	3 jet regions, 0- and 1-lepton channels	9%	Normalisation
<code>SystTtbarNorm_DWhfCR_L1</code>	$t\bar{t}$ Whf CR to SR		25%	Normalisation
<code>norm_ttbar_J2_L2</code>	$t\bar{t}$ normalisation	2-jet 2-lepton	Float	Normalisation
<code>norm_ttbar_J3_L2</code>	$t\bar{t}$ normalisation	3+ jet regions, 2-lepton channel	Float	Normalisation
<code>SysTTbarPtV</code>	p_T^V shape	0- and 1- lepton channels	-	Shape
<code>SysTTbarMBB</code>	m_{bb} shape	0- and 1- lepton channels	-	Shape
<code>SysTTbarPTV_L2</code>	p_T^V shape	2-lepton channels, low to high p_T^V	-	Normalisation+Shape
<code>SysTTbarMBB_L2</code>	m_{bb} shape	2-lepton channels	-	Shape

Table 37: Summary of the $t\bar{t}$ systematic uncertainties: the first column quotes the name of the nuisance parameter implemented in the fit referring to a specific systematic uncertainty, the second column the source of the uncertainty, the third the categories and sample on which it is applied, the fourth column the value of the Gaussian prior on the NP (if applicable) and the fifth column the effect of the systematic uncertainty. The listed systematic uncertainties are separated in normalisation effects in 0- and 1-lepton (first block), 2-lepton (second block), and shape systematic uncertainties (third block).

Not reviewed, for internal circulation only

Nuisance Parameter	Description	Categories	Value	Effect
Multi-jet				
SysMNorm_2J_E1	multi-jet normalisation	1-lepton channel, 2-jet, e	+63% - 59% (+190% - 100% cut-based)	Normalisation
SysMNorm_3J_E1	multi-jet normalisation	1-lepton channel, 3-jet, e	+416% - 100% (+195% - 100% cut-based)	Normalisation
SysMNorm_2J_Mu	multi-jet normalisation	1-lepton channel, 2-jet, μ	+56% - 100% (+72% - 100% cut-based)	Normalisation
SysMNorm_3J_Mu	multi-jet normalisation	1-lepton channel, 3-jet, μ	+198% - 100% (+275% - 100% cut-based)	Normalisation
SysMJTrigger	multi-jet shape	1-lepton, electron channel	-	Shape
SysMReduced	multi-jet shape	1-lepton, electron and muon channel	-	Shape
SysMWCORRECTION	multi-jet shape	1-lepton, electron and muon channel	-	Shape
SysM2Tag	multi-jet shape	1-lepton, electron and muon channel	-	Shape
SysMJSFsCR	multi-jet shape	1-lepton, electron and muon channel	-	Shape
Single top				
stop _s	single-top (s -channel) normalisation	all regions	4.6%	Normalisation
stop _t	single-top (t -channel) normalisation	all regions	4.4%	Normalisation
stop _{tW}	single-top (Wt -channel) normalisation	all regions	6.2%	Normalisation
stop _{tAcc}	single-top (t -channel) acceptance	2,3jet regions	17% (2jets), 20% (3jets)	Normalisation
stop _{tWtAcc}	single-top (Wt channel) acceptance	2,3jet regions	35% (2jets), 41% (3jets)	Normalisation
SysSstop _t PTV	single-top (t -channel) p_T^V shape	all regions	-	Shape
SysSstop _t MBB	single-top (t -channel) m_{bb} shape	all regions	-	Shape
SysSstop _{tW} PTV	single-top (Wt -channel) p_T^V shape	all regions	-	Shape
SysSstop _{tW} MBB	single-top (Wt -channel) m_{bb} shape	all regions	-	Shape
Diboson				
ZZNorm	ZZ normalisation	all regions	20%	Normalisation
WZNorm	WZ normalisation	all regions	26%	Normalisation
WWNorm	WW normalisation	all regions	25%	Normalisation
SysVZUEPSACC	acceptance variation	2jet regions	WZ ℓvbb : 3.9%, ZZ $\ell \ell bb$: 5.8%, ZZ $v vbb$: 5.6%	Normalisation
SysVZUEPSACC	acceptance variation	3(+jet regions	WZ ℓvbb : 3.9%, ZZ $\ell \ell bb$: 5.8%, ZZ $v vbb$: 5.6%	Normalisation
SysVZUEPS_32JR	acceptance variation	3(+jet regions	WZ ℓvbb : 10.8%, ZZ $\ell \ell bb$: 3.1%, ZZ $v vbb$: 7.3%	Normalisation
SysZZUEPSResid_0L	ZZ extrapolation from 2 to 0	0-lepton channel	6%	Normalisation
SysWZUEPSResid_0L	WZ extrapolation from 1 to 0	0-lepton channel	11%	Normalisation
QCDscaleAcc_J2_V2	acceptance variation	2 jet regions	WZ ℓvbb : 12.7%, ZZ $\ell \ell bb$: 11.9%, ZZ $v vbb$: 10.3%	Normalisation
QCDscaleAcc_J3_V2	acceptance variation	2 jet regions	WZ ℓvbb : -17.7%, ZZ $\ell \ell bb$: -16.4%, ZZ $v vbb$: -15.2%	Normalisation
QCDscaleAcc_J3_V2	acceptance variation	3(+jet regions	WZ ℓvbb : 21.2%, ZZ $\ell \ell bb$: 10.1%, ZZ $v vbb$: 17.4%	Normalisation
QCDscaleAcc_PTVAzeto_V2	acceptance variation	3(+jet regions	WZ ℓvbb : 19.0%, ZZ $v vbb$: 18.2%	Normalisation
SysVVPTVME	p_T^V shape	all regions	-	Shape(+normalisation for 2lep)
SysVVMbbME	m_{bb} shape	all regions	-	Shape
SysVZUEPSVPT	p_T^V shape variation	all regions	-	Shape(+normalisation for 2lep)
SysVZUEPSMbb	m_{bb} shape variation	all regions	-	Shape
SysVZQCDscaleVPT	p_T^V shape variation	all regions	-	Shape(+normalisation for 2lep)
SysVZQCDscaleMbb	m_{bb} shape variation	all regions	-	Shape

Table 38: Summary of the systematic uncertainties for single-top, diboson : the first column quotes the name of the nuisance parameter implemented in the fit referring to a specific systematic uncertainty, the second column the source of the uncertainty, the third the categories and sample on which it is applied, the fourth column the value of the Gaussian prior on the NP (if applicable) and the fifth column the effect of the systematic uncertainty. The listed systematic uncertainties are separated in normalisation effects, acceptance variations and shape systematic uncertainties.

make up a substantial part of this supposed systematic difference. If there are multiple JES uncertainties, as in this analysis, then this MC uncertainty should not be included in each one.

To mitigate these effects, two so-called “smoothing” algorithms (developed for the Run 1 analysis) are used to merge consecutive bins in the MC templates. First, bins from one extremum to the next are merged until no local extrema remain in the BDT distribution (or up to on in the m_{bb} distribution for the cut-based analysis). Merging is performed at each step of this iterative process where the difference between merged and unmerged templates is the smallest. Second, the bins resulting from this first algorithm are sequentially merged, starting from the upper end of the distribution, until the statistical uncertainty in each of the merged bins, calculated in the nominal template, is smaller than 5%. In each of these sets of bins, the integrals of the nominal and systematically shifted distributions are compared to give the $\pm 1\sigma$ variation. This value is then used as the associated uncertainty for all the nominal bins in the set.

The result of smoothing systematic uncertainties is checked to ensure that this is behaving as expected.

Not reviewed, for internal circulation only

Nuisance Parameter	Description	Categories	Value	Effect
SM Vh signal				
TheoryQCDscale	QCD scale variation	all regions	$qqVH: 0.7\%$, $ggZH: 27\%$	Normalisation
TheoryPDF_qqVH	PDF variation	all regions	$WH: 1.9\%$, $qqZH: 1.6\%$	Normalisation
TheoryPDF_ggZH	PDF variation	all regions	$ggZH: 5.0\%$	Normalisation
TheoryBRbb	BR variation	all regions	1.7%	Normalisation
TheoryAcc_J2_qqZh	acceptance variation	2 jet regions	$WH: 8.8\%$, $qqZ\ell\ell H: 3.3\%$, $qqZvvH: 6.9\%$	Normalisation
TheoryAcc_J2_ggZH	acceptance variation	2 jet regions	$ggZ\ell\ell H: 3.3\%$, $ggZvvH: 6.9\%$	Normalisation
TheoryAcc_J3_qqZH	acceptance variation	2 jet regions	$WH: -8.6\%$, $qqZ\ell\ell H: -3.2\%$, $qqZvvH: -7.0\%$	Normalisation
TheoryAcc_J3_ggZH	acceptance variation	2 jet regions	$ggZ\ell\ell H: -3.2\%$, $ggZvvH: -7.0\%$	Normalisation
theoryAcc_J3_qqZh	acceptance variation	3(+jet) regions	$WH: 6.8\%$, $qqZ\ell\ell H: 3.9\%$, $qqZvvH: 5.0\%$	Normalisation
TheoryAcc_J3_ggVH	acceptance variation	3(+jet) regions	$ggZ\ell\ell H: 3.9\%$, $ggZvvH: 5.0\%$	Normalisation
SysTheoryAcc_JVeto_qqZH	acceptance variation	3(+jet) regions	$WH: 3.8\%$, $qqZvvH: -2.5\%$	Normalisation
SysTheoryAcc_JVeto_ggZH	acceptance variation	3(+jet) regions	$ggZvvH: -2.5\%$	Normalisation
SysTheoryPDFAcc_qqVH	acceptance variation	all regions	$WH: 1.3\%$, $qqZ\ell\ell H: 0.5\%$, $qqZvvH: 1.1\%$	Normalisation
SysTheoryPDFAcc_ggZH	acceptance variation	all regions	$ggZ\ell\ell H: 0.5\%$, $ggZvvH: 1.1\%$	Normalisation
SysTheoryUEPSAcc	acceptance variation	2jet regions	$WH: 12.1\%$, $Z\ell\ell H: 13.9\%$, $ZvvH: 10.0\%$	Normalisation
SysTheoryUEPSAcc	acceptance variation	3(+jet) regions	$WH: 12.1\%$, $Z\ell\ell H: 13.9\%$, $ZvvH: 10.0\%$	Normalisation
SysTheoryUEPSAcc_32JR	acceptance variation	3(+jet) regions	$WH: 12.9\%$, $Z\ell\ell H: 13.4\%$, $ZvvH: 13.0\%$	Normalisation
SysVHQCDscalePTV	p_T^V shape variation	all regions	-	Shape(+normalisation for only 2lep)
SysVHQCDscaleMbb	m_{bb} shape variation	all regions	-	Shape
SysVHQCDscalePTV_ggZH	p_T^V shape variation	all regions	-	Shape(+normalization for only 2lep)
SysVHQCDscaleMbb_ggZH	m_{bb} shape variation	all regions	-	Shape
SysVHPDFPTV	p_T^V shape variation	all regions	-	Shape
SysVHPDFPTV_ggZH	p_T^V shape variation	all regions	-	Shape
SysVHUEPSPTV	p_T^V shape variation	all regions	-	Shape(+normalisation for only 2lep)
SysVHUEPSMbb	m_{bb} shape variation	all regions	-	Shape
SysVHNLOEWK	EWK weight shape variation	all regions	-	Shape

Table 39: Summary of the systematic uncertainties for SM Vh processes: the first column quotes the name of the nuisance parameter implemented in the fit referring to a specific systematic uncertainty, the second column the source of the uncertainty, the third the categories and sample on which it is applied, the fourth column the value of the Gaussian prior on the NP (if applicable) and the fifth column the effect of the systematic uncertainty. The listed systematic uncertainties are separated in normalisation effects, acceptance variations and shape systematic uncertainties.

7.3.2. Pruning of the Systematic Uncertainties

Several of the uncertainties described in Section 6 and 7.3 have a negligible effect on the distributions entering in the fit. In addition, limited statistics in the MC nominal distributions can produce systematic templates with large fluctuations, introducing artificial variations in the fit. Therefore, following the Run 1 strategy, uncertainties are removed following a “pruning” procedure, which is carried out for each category/sub-channel in each region.

Pruning is performed as follows:

- Neglect the normalisation uncertainty for a given sample in a region if either of the following is true:
 - the variation is less than 0.5%;
 - both up and down variations have the same sign.
- Neglect the shape uncertainty for a given sample in a given region if the following is true:
 - not one single bin has a deviation over 0.5% after the overall normalisation is removed;
 - if only the up or the down variation is non-zero and passed the previous pruning steps.
- Neglect the shape and normalisation uncertainties for a given sample in a given region if the sample is less than 2% of the total background:

- 1001 – if the signal < 2% of the total background in all bins and the shape and normalisation error
 1002 are each < 0.5% of the total background;
- 1003 – if at least one bin has a signal contribution > 2% of the total background, only in those bins
 1004 where the shape and normalisation error are each < 2% of the signal yield.

1005 The list of pruned uncertainties is regularly checked to ensure that this is behaving as expected.

1006 **7.4. Fit configuration validation**

1007 The likelihood can be maximized with respect to the observed data, or to the pseudo data derived from
 1008 the nominal MC expectations (or any systematic variation) by creating an Asimov dataset. When fitting
 1009 to the nominal MC expectation, the constrained NPs should be found at 0, and the floating normalisations
 1010 at 1. It's possible than when performing the fit to data some of the NPs are pulled away from the expected
 1011 value: in this case we want to be confident that we understand the cause of this pull, and that it isn't biasing
 1012 the fitted signal.

1013 The post-fit uncertainties on a given parameter extracted from a fit on the Asimov dataset, are expected
 1014 to be equal to 1, but might be smaller or larger if the parameter is over or under constrained, respectively.
 1015 The post-fit uncertainties from a data fit should be compared to the result of the Asimov fit, to check that
 1016 the constraints obtained with data match the expectations, given the size of the uncertainty, the samples,
 1017 and the region(s) on which it acts.

1018 Unless otherwise stated, the nominal MC expectation is used when the Asimov data set is mentioned.
 1019 An important tool to understand the impact of the NPs on the signal is the so called NP ranking which
 1020 evaluates the impact of each NP on the best fit of the signal strength.

1021 **7.4.1. Nuisance parameter pulls and constraints**

1022 In Figure 38, the nuisance parameter pulls and constraints are shown for the unconditional combined fit,
 1023 performed on an Asimov dataset and 2015+2016 data. In Figures 40 to 41, the same pulls and constraints
 1024 are shown in more detail. MC statistical uncertainties, normalisation and shape modelling systematic
 1025 uncertainties, and experimental systematic uncertainties are considered in the fit.

Not reviewed, for internal circulation only

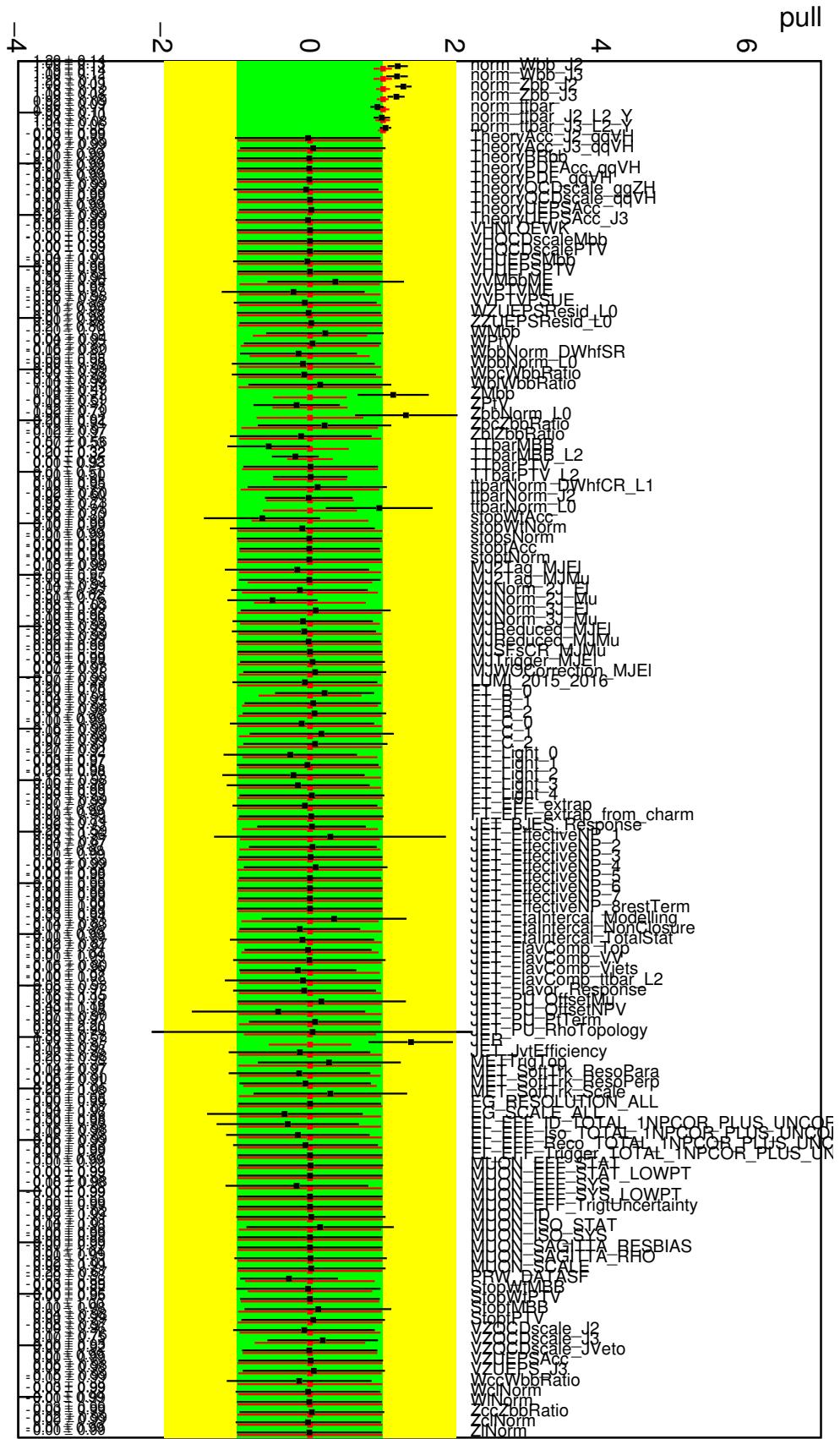


Figure 38: The nuisance parameter pulls and the free parameter scale factors corresponding to the combined unconditional fits performed on Asimov datasets (red) and 2015+2016 data (black) are shown.

23rd June 2017 – 14:35

Not reviewed, for internal circulation only

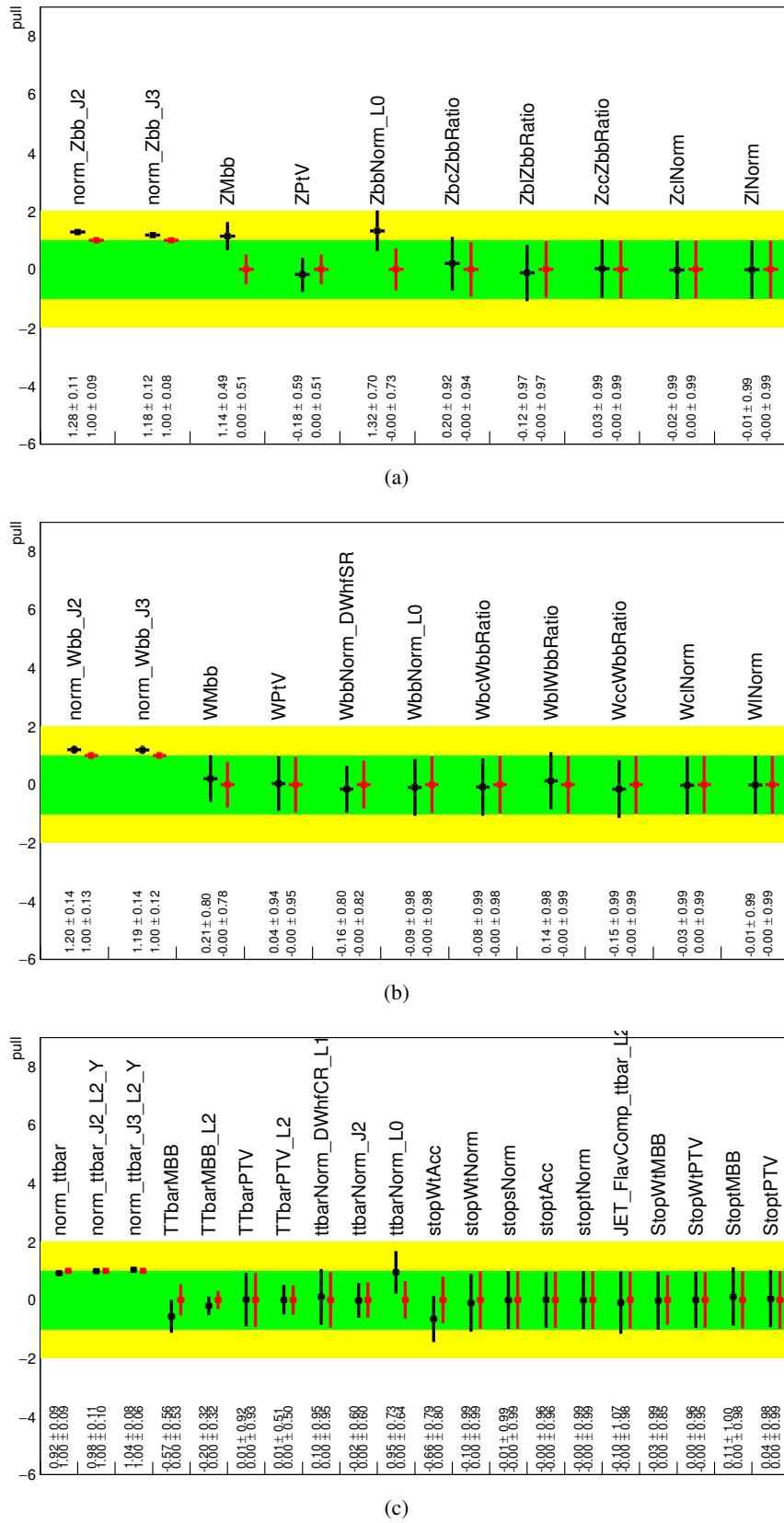


Figure 39: Pull distribution of modeling NPs in the combined unconditional VH fit. Z jets (a), W jets (b), top (c) for the combined unconditional fits performed on Asimov datasets (red) and 2015+2016 data (black).

Not reviewed, for internal circulation only

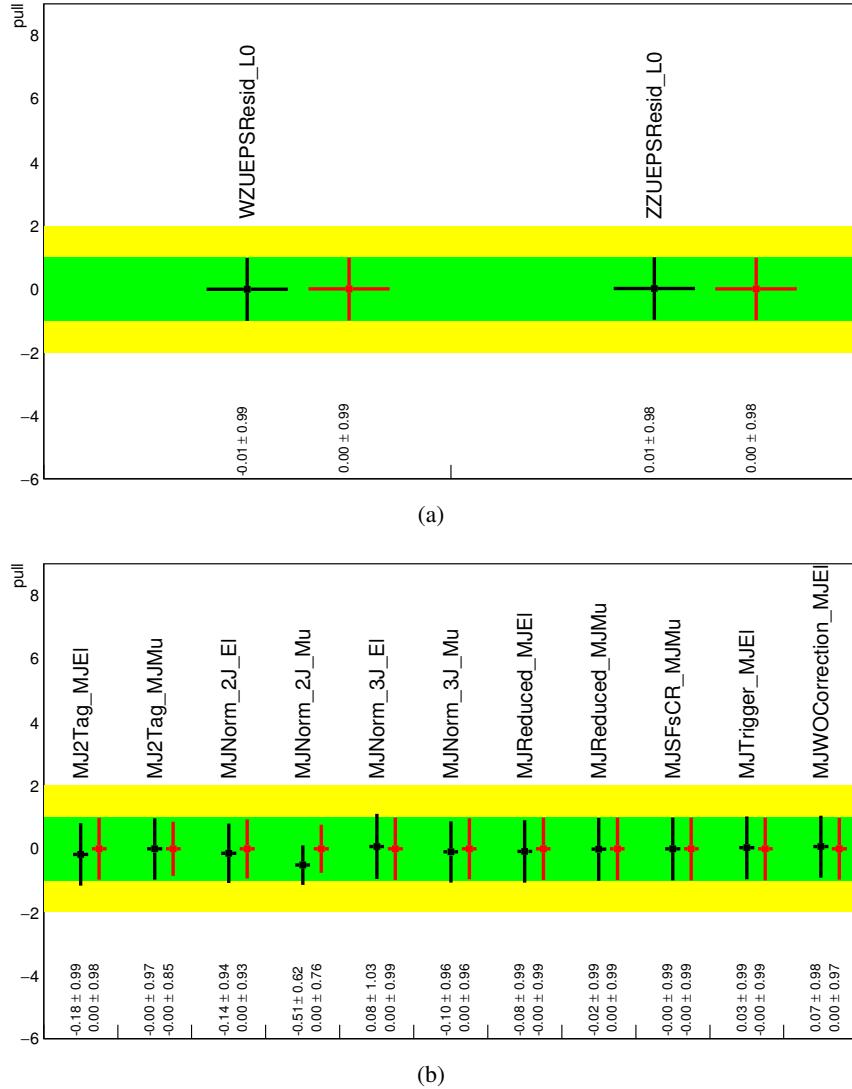


Figure 40: Pull distribution of modeling NPs in the combined unconditional VH fit. diboson (a) and MJ (b) for the combined unconditional fits performed on Asimov datasets (red) and 2015+2016 data (black)..

Not reviewed, for internal circulation only

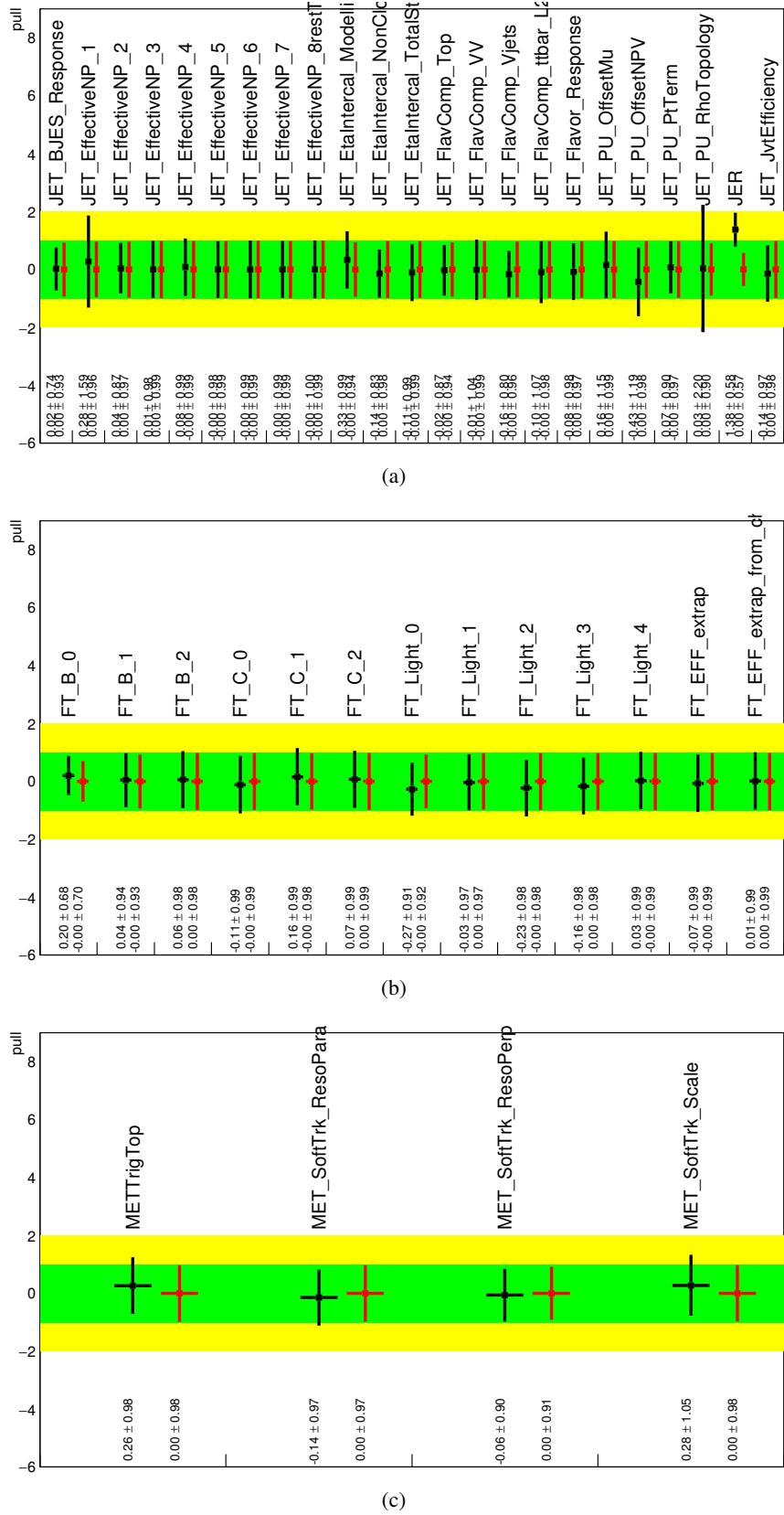


Figure 41: Pull distribution of experimental NPs in the combined unconditional VH fit. Jet systematics (a), b-tagging (b) and E_T^{miss} (c) for the combined unconditional fits performed on Asimov datasets (red) and 2015+2016 data (black)..

1026 **7.4.2. Correlation matrices**

1027 The correlation matrices corresponding to the combined fit when fitting an Asimov dataset and data can be
 1028 found in Figure 42. Only correlations among parameters with a magnitude greater than 0.25 are shown.

Not reviewed, for internal circulation only

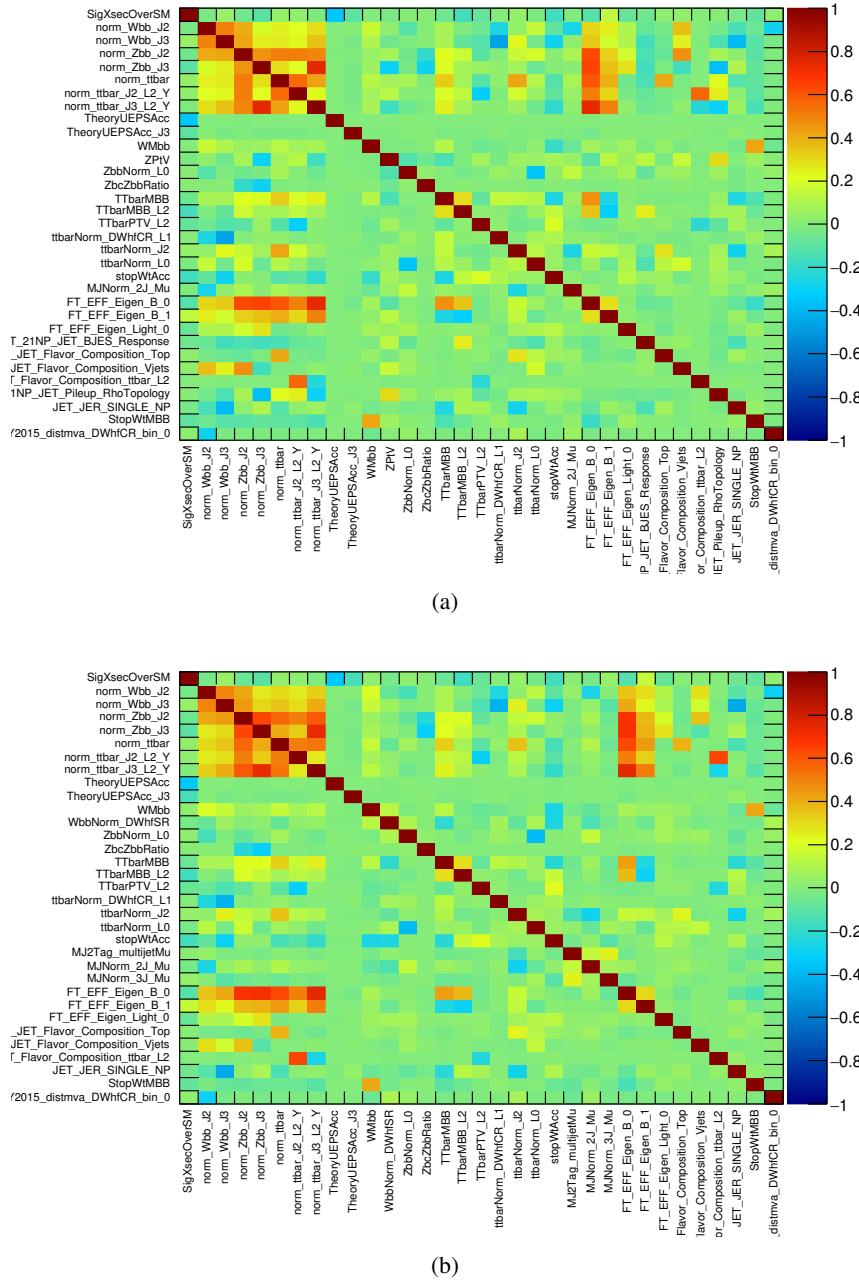


Figure 42: Correlation matrices from the unconditional combined VH fit to an Asimov dataset (a) and to 2015 and 2016 data (b).

1029 **7.4.3. Nuisance parameter ranking**

1030 The pre-fit and post-fit ranking plots for the combined fit on data are shown in Figure 43. For the pre-fit
 1031 ranking, the value of the prior is used, which is equivalent for the nuisance parameter to $\sigma = 1$. For the
 1032 post-fit ranking, the error on the profiled nuisance parameter at the maximum of the likelihood is used.

1033 The blue boxes show the postfit impact of the nuisance parameters, which is evaluated by changing them by
 1034 their profiled error at the maximum of the likelihood. The change in the best fit μ value gives the inclusive
 1035 sensitivity of the measured value to the given NP. The yellow bands correspond to the so-called prefit
 1036 impact of the nuisance parameters, which is evaluated by changing the values of the nuisance parameters
 1037 by 1 (which is the prefit uncertainty on the NP). By construction, floating parameters (without any prior)
 1038 do not have any prefit impact. Only the systematics with the largest impact are shown.

Not reviewed, for internal circulation only

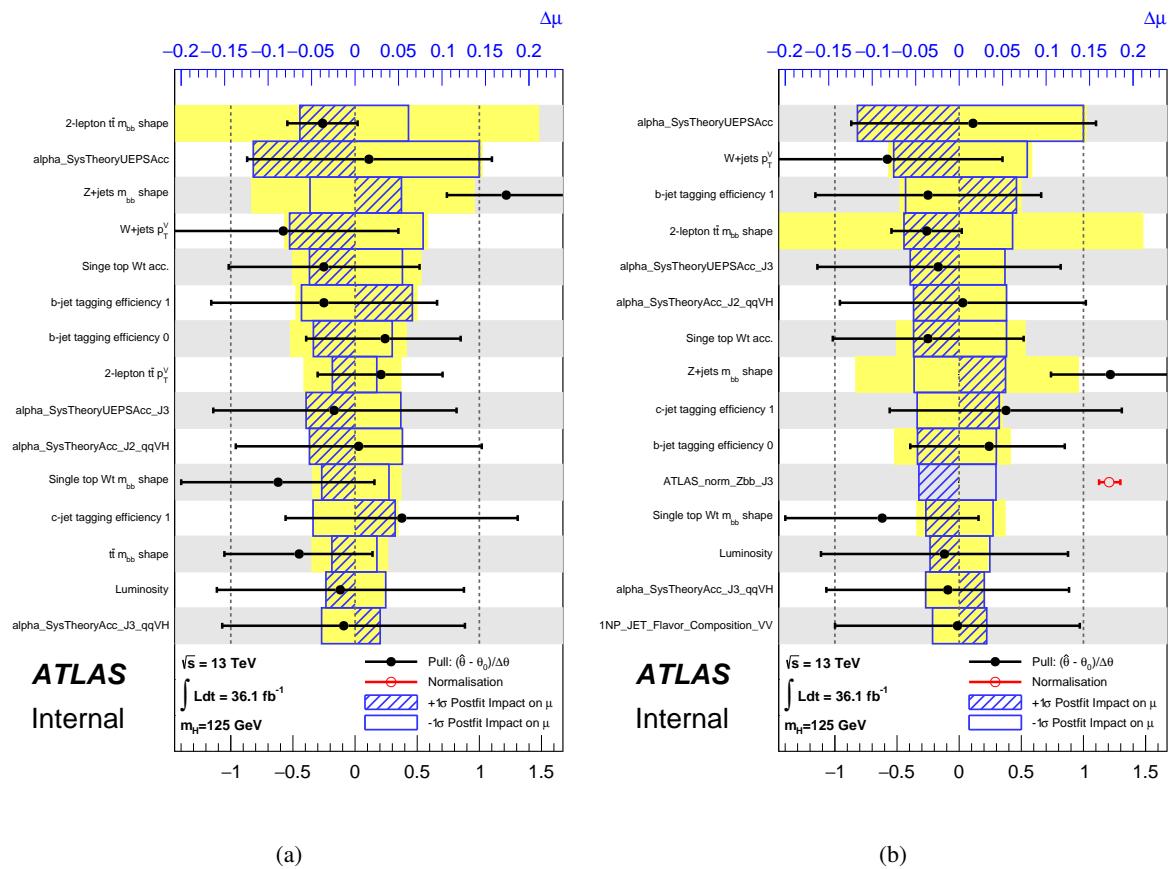


Figure 43: The post-fit impact nuisance parameter ranking for the combined fit based on prefit (a) and postfit (b) impact on $\hat{\mu}$.

1039 **7.4.4. Breakdown of statistical and systematic uncertainties**

1040 Statistical uncertainties due to the limited sizes of the simulated MC samples are taken into account
 1041 in the fit by means of nuisance parameters, which are parametrized by Poisson priors. Based on the
 1042 study described in Appendix F the binning on the final discriminant is chosen so that the MC statistical
 1043 uncertainty for all bins is less than 20%.

1044 The fitted parameter of interest $\hat{\mu}$ and the breakdown of the uncertainties coming from data statistics (“Data
 1045 Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only
 1046 (“MC Stat.”), and each category of systematic uncertainty, is shown for the combined fit on 36.1 fb^{-1} of
 1047 Asimov data in Table 40.

Set of nuisance parameters	Impact on error
Total	+0.418 / -0.363
DataStat	+0.239 / -0.234
FullSyst	+0.343 / -0.279
Floating normalizations	+0.072 / -0.071
All normalizations	+0.083 / -0.082
All but normalizations	+0.328 / -0.260
Jets MET	+0.072 / -0.053
BTag	+0.117 / -0.105
Leptons	+0.012 / -0.009
Luminosity	+0.046 / -0.026
Diboson	+0.055 / -0.034
Zjets	+0.076 / -0.081
Wjets	+0.080 / -0.074
Model ttbar	+0.081 / -0.071
Model Single Top	+0.080 / -0.073
Model Multi Jet	+0.023 / -0.022
Signal Systematics	+0.219 / -0.117
MC stat	+0.135 / -0.129

Table 40: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and different sub-groups of systematic uncertainties, is shown, as well as the fitted $\hat{\mu}$ extracted from the combined unconditional fit on 36.1 fb^{-1} of data.

1048 **7.4.5. Post-fit BDT distributions**

1049 In Figures 44, 45, 46 and 47, the postfit BDT distributions in the 0-, 1-, and 2-lepton channels are shown,
 1050 as obtained from the combined fit on 2015 and 2016 data. The overall agreement between data and MC
 1051 distributions is very good.

1052 Additional post-fit distributions for all BDT input variables are presented in Appendix K.

Not reviewed, for internal circulation only

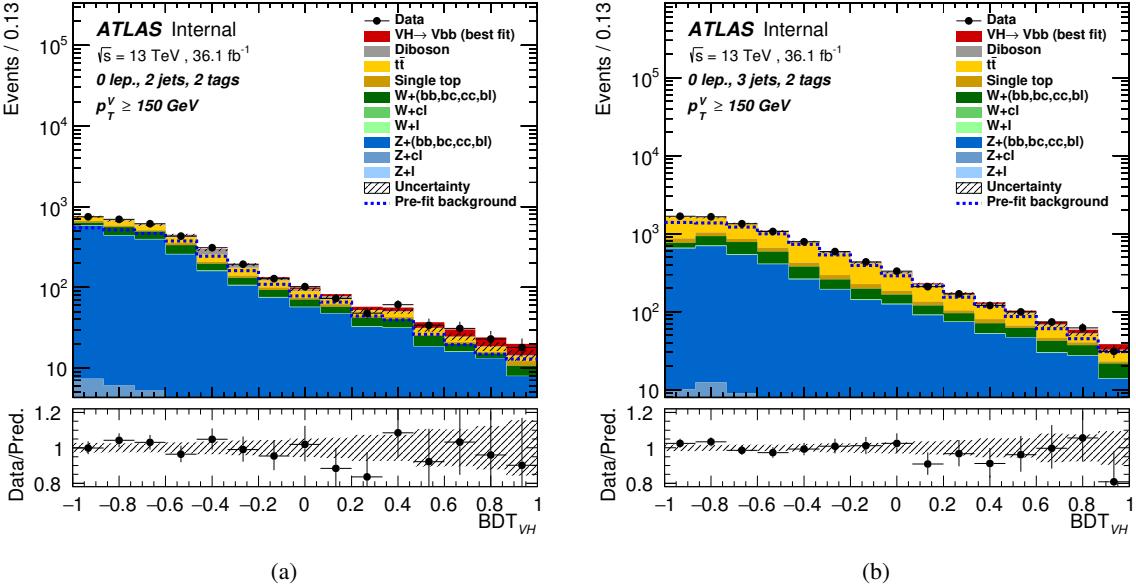


Figure 44: BDT postfit distributions obtained in the 0-lepton channel for the combined unconditional VH fit on 2015 and 2016 data in the 2-tag, 2-jet region (a), and 2-tag 3-jet region (b). The best-fit signal value is used.

Not reviewed, for internal circulation only

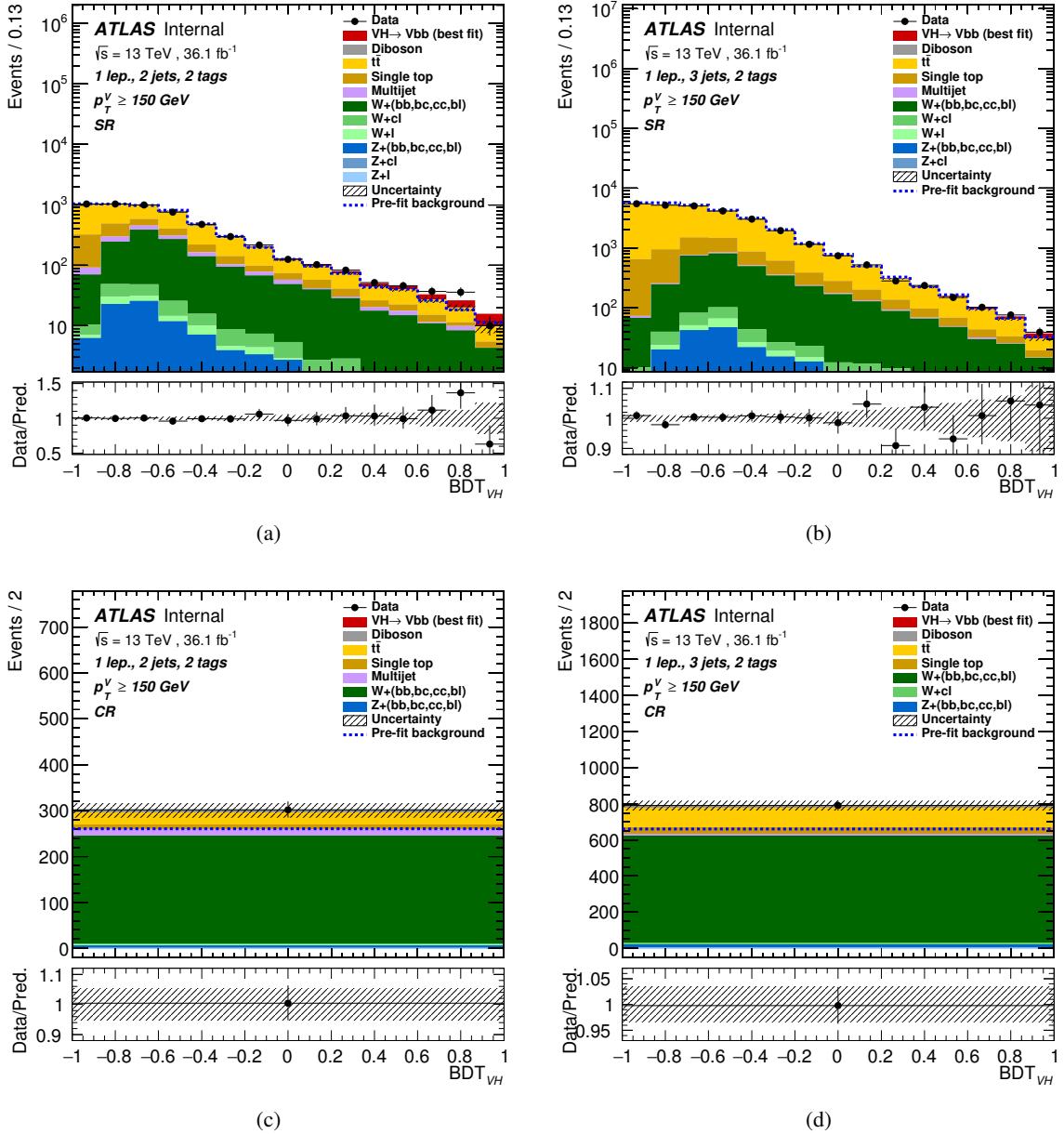


Figure 45: BDT postfit distributions obtained in the 1-lepton channel for the combined unconditional VH fit on 2015 and 2016 data in the 2-tag, 2-jet CR (a), SR (b), and 2-tag 3-jet CR (c) and SR (d). The best-fit signal value is used.

Not reviewed, for internal circulation only

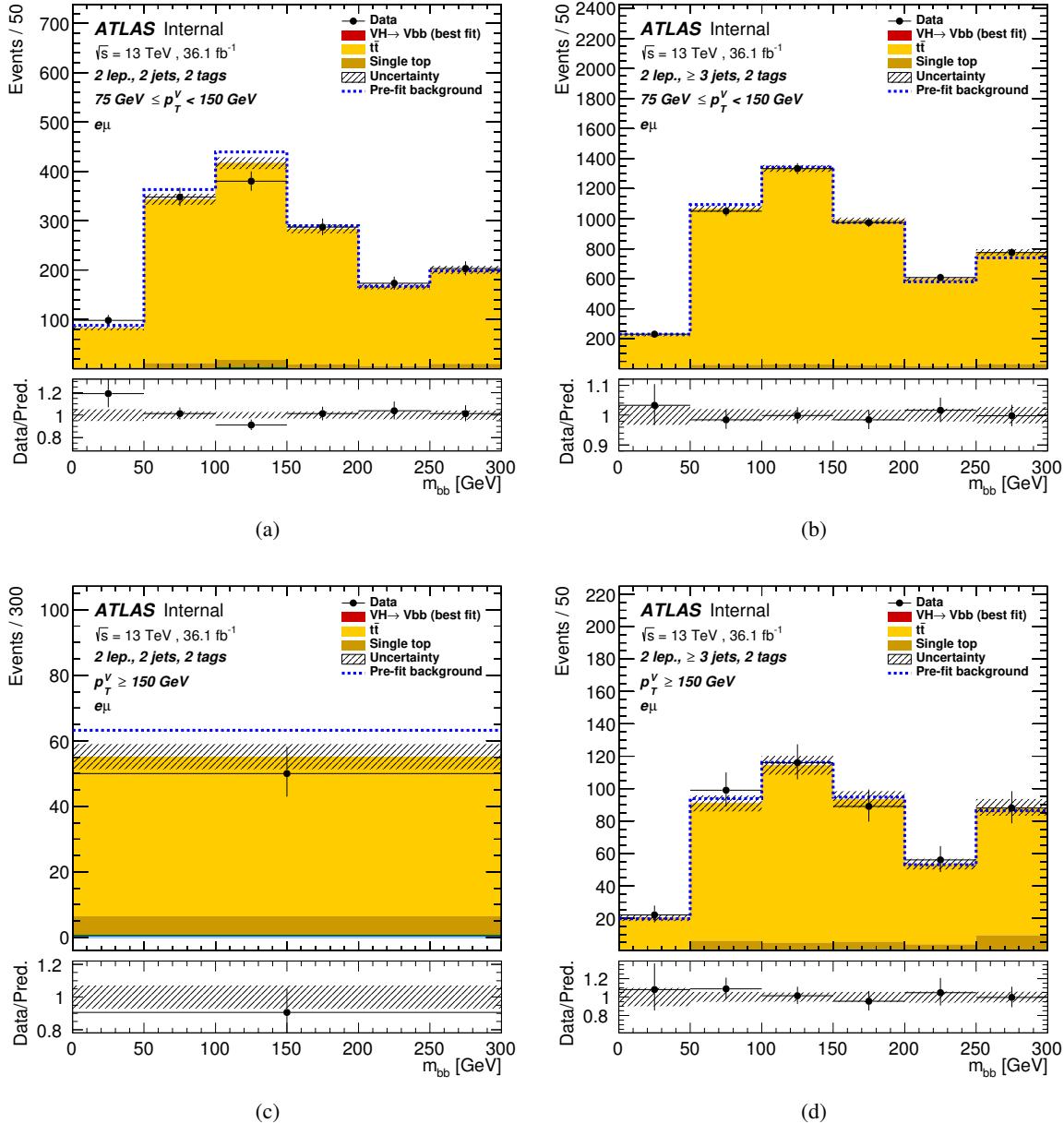


Figure 46: m_{bb} postfit distributions obtained in the 2-lepton channel $e - \mu$ control regions for the combined unconditional VH fit on 2015 and 2016 data. The best-fit signal value is used.

Not reviewed, for internal circulation only

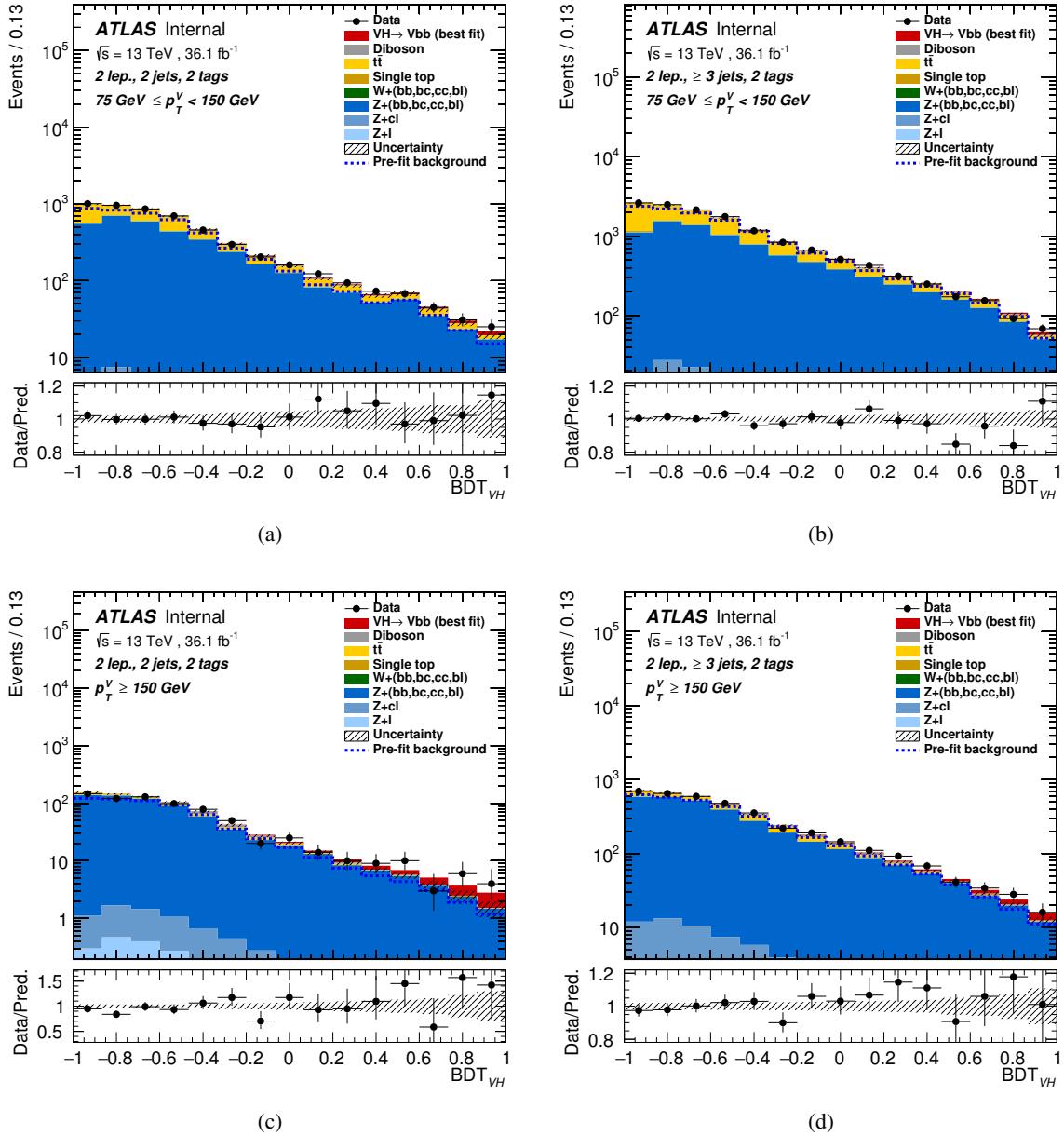


Figure 47: BDT postfit distributions obtained in the 2-lepton channel for the combined unconditional VH fit on 2015 and 2016 data. The best-fit signal value is used.

1053 **7.4.6. Background-Subtracted Post-fit BDT distributions**

1054 In this section, background subtracted data/MC comparison plots of the post-fit VH BDT distribution are
 1055 presented. Figure 48 presents the background subtracted MVA distribution summed for all fitted signal
 1056 regions, additional with each region weighted by the Higgs S/B, when running the unconditional fit to
 1057 data. Figures 49 - 51 then present the background subtracted MVA distributions in the 0-, 1- and 2-lepton
 1058 regions, respectively, after running the unconditional fit to data on all regions.

1059 *Note: due to a plotting error, the error bands on these plots are incorrect, and will be fixed for the next
 1060 circulation.*

Not reviewed, for internal circulation only

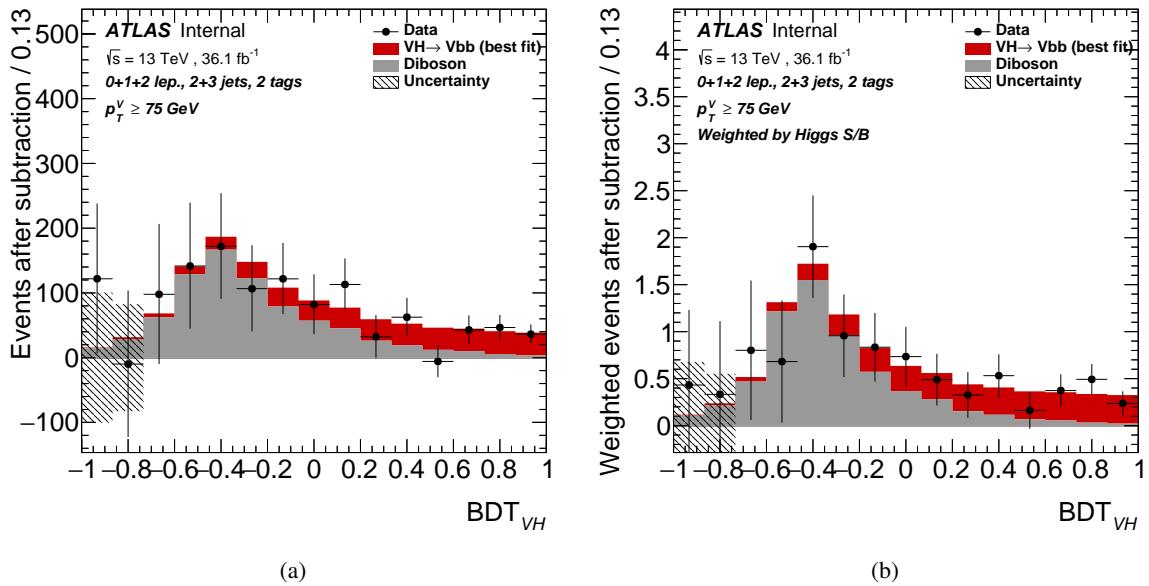


Figure 48: Background subtracted MVA distributions summed for all fitted signal regions and channels (a), with each region weighted by the Higgs S/B (b), when running the unconditional fit to data.

Not reviewed, for internal circulation only

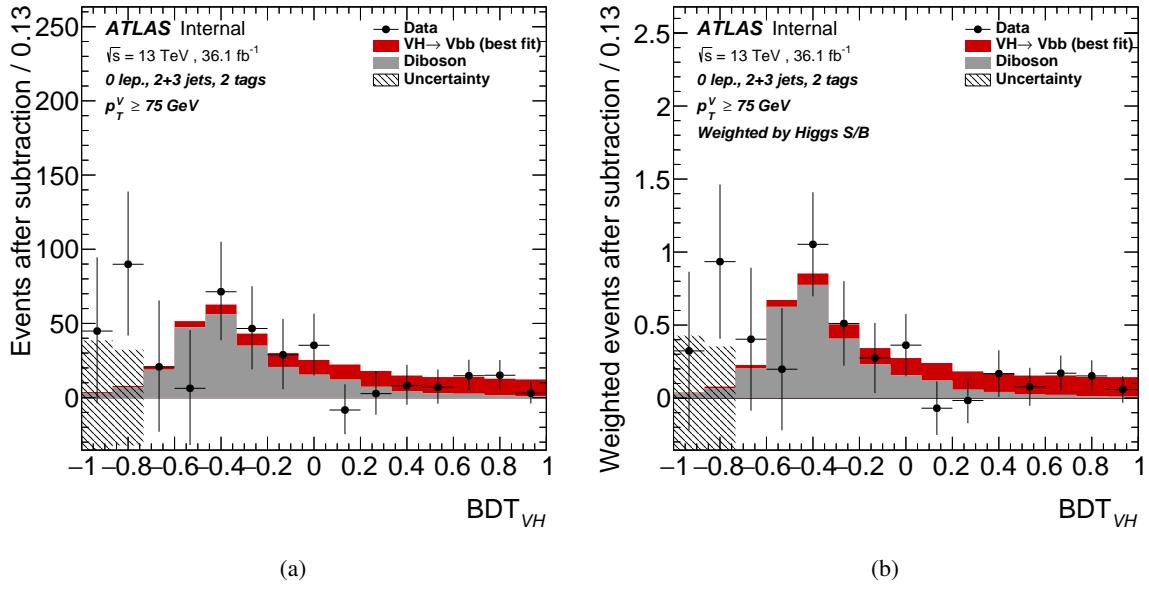


Figure 49: Background subtracted MVA distributions summed for all fitted signal regions for the 0-lepton channel (a), with each region weighted by the Higgs S/B (b), when running the unconditional fit to data.

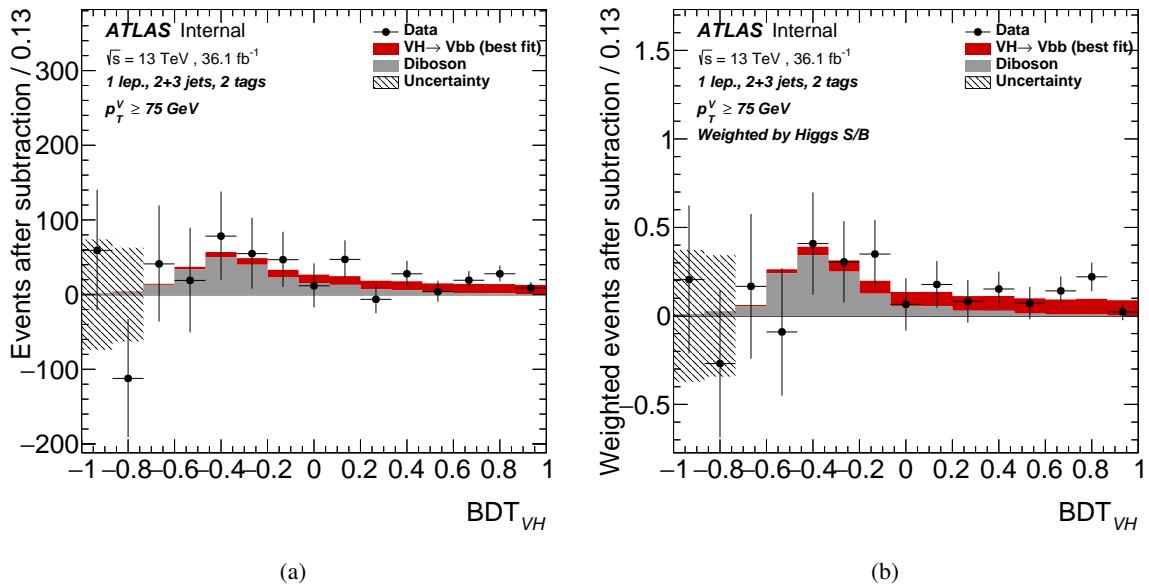


Figure 50: Background subtracted MVA distributions summed for all fitted signal regions for the 1-lepton channel (a), with each region weighted by the Higgs S/B (b), when running the unconditional fit to data.

Not reviewed, for internal circulation only

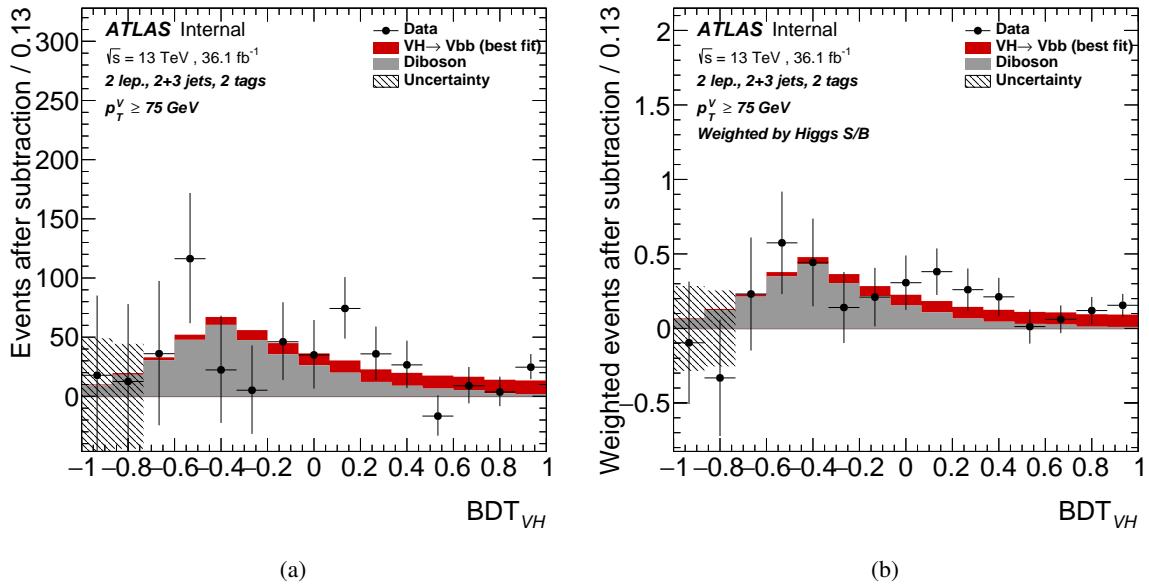


Figure 51: Background subtracted MVA distributions summed for all fitted signal regions for the 2-lepton channel (a), with each region weighted by the Higgs S/B (b), when running the unconditional fit to data.

1061 7.4.7. Signal over Background Plots

1062 In Figure 52 all BDT bins are summed up across all channel and regions used in the combined fit, and
1063 ordered by their signal to background ratio is shown, with the signal strength taken from the best-fit value.
1064 The same distribution is shown three times, associated to ratios computed as (a) data over expected back-
1065 ground ($data/background$), (b) background-subtracted data over signal ($(data - background)/signal$),
1066 and (c) background subtracted data over square-root of background ($(data - background)/\sqrt{background}$).
1067

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

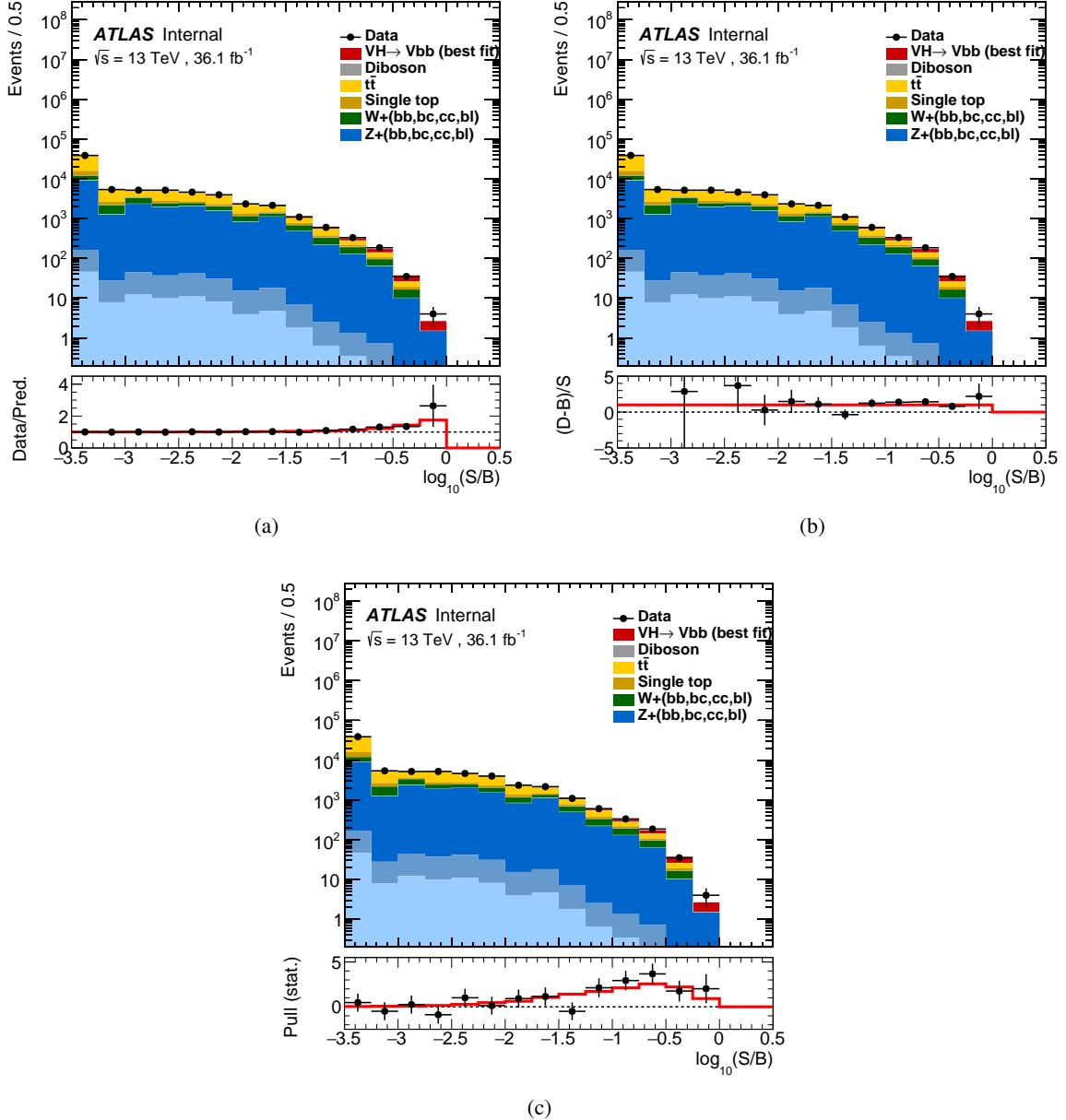


Figure 52: All BDT bins summed up across all regions and channels used in the combined unconditional VH fit, and ordered by their signal to background ratio is shown. Signal strength is taken from the best-fit value. The same distribution is shown three times, associated to ratios computed as (a) data over expected background ($\text{data}/\text{background}$), (b) background-subtracted data over signal ($(\text{data} - \text{background})/\text{signal}$), and (c) background subtracted data over square-root of background ($(\text{data} - \text{background})/\sqrt{\text{background}}$).

1068 **7.5. Fit to diboson VZ resonance**

1069 Studies on the diboson ($VZ \rightarrow b\bar{b}$) fit offers a good cross-check of our full analysis procedure for the
1070 SM VH search. The final state is identical to the signal, and the kinematics are similar to the SM VH ,
1071 except for the substitution of Z boson for a H boson. The diboson signal has a softer p_T^V distribution
1072 than the Higgs. This has the important consequence to increase the importance of the 2-lepton channel
1073 (which has a low p_T^V regions), relatively to the 0- and 1-lepton channels, when comparing with the Higgs
1074 fit. For the diboson measurement, the BDT for each channel is retrained with the diboson $VZ \rightarrow b\bar{b}$ as
1075 the signal. The input variables and BDT configurations are same as for the SM VH BDT. VZ diboson
1076 fit and measurement of $VZ \rightarrow b\bar{b}$ signal strength are then performed. Results from the unconditional fit
1077 performed to all fitting regions is presented, including comparisons of the pulls, rankings and correlations
1078 of nuisance parameters to the VH fit. Results from the diboson fit carried out in each individual lepton
1079 channel are presented in Appendix L.

1080 There is, however, one difference between the VZ fit and the VH fit. Due to limited MC statistics in the
1081 diboson samples, the parameters of the BDT transformation, Section 5.3, are varied. Instead of using
1082 $z_s = 10, z_b = 5$, this is altered to $z_s = 5, z_b = 5$. This increases the diboson signal statistics in each bin,
1083 reducing the impact of MC statistical uncertainties on the fit result.

1084 **7.5.1. Nuisance parameter pulls and constraints**

1085 In Figure 53, the nuisance parameter pulls and constraints are shown for the combined unconditional fit,
1086 performed on 2015+2016 data for the VH fit (black) and the diboson VZ fit (red). In Figures 55 to 56, the
1087 same pulls and constraints are shown in more detail.

1088 MC statistical uncertainties, normalisation and shape modelling systematic uncertainties, and experimental
1089 systematic uncertainties are considered in the fit.

Not reviewed, for internal circulation only

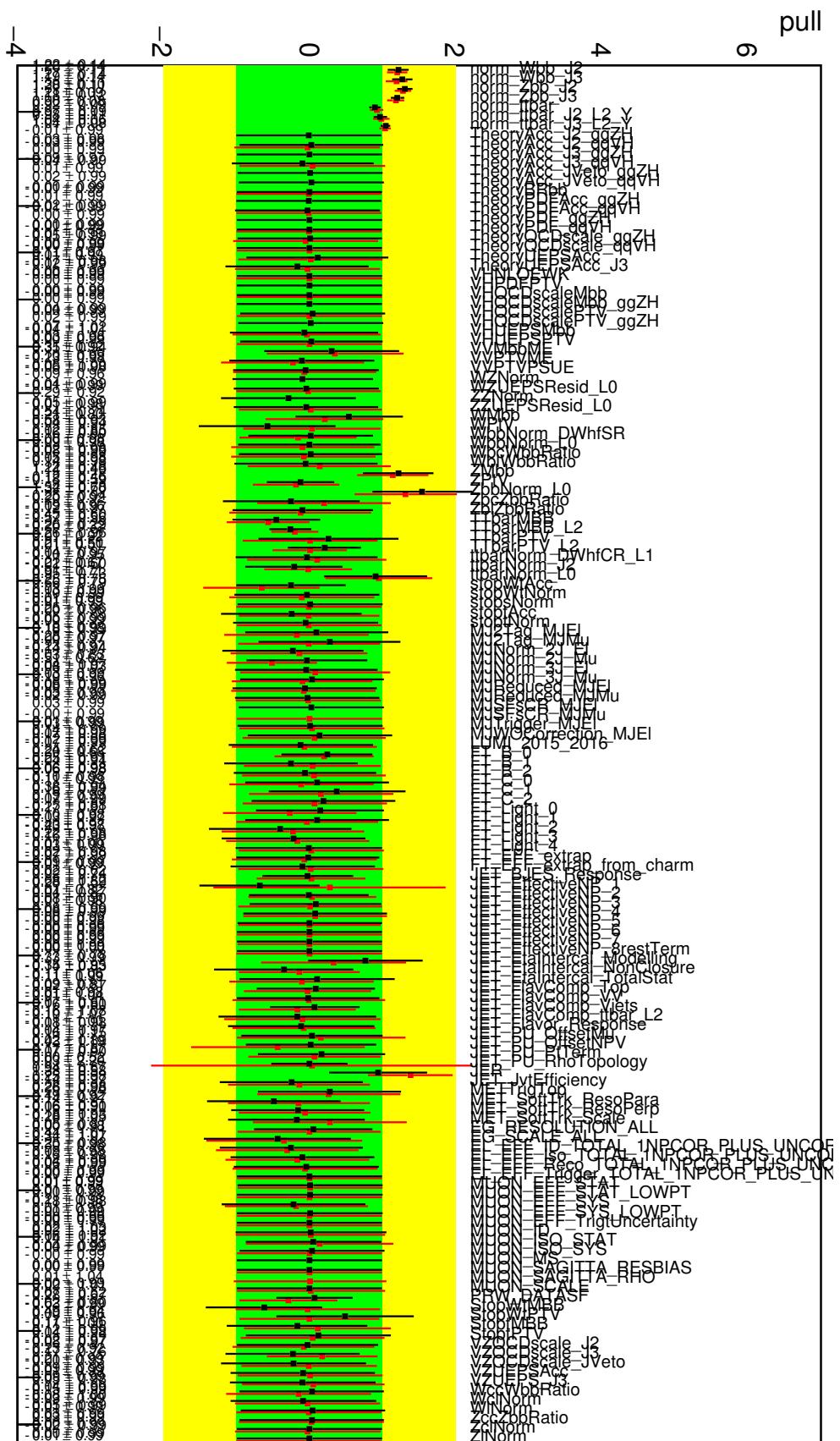


Figure 53: The nuisance parameter pulls and the free parameter scale factors corresponding to the combined unconditional fit to data for the VH fit (black) and the VZ fit (red).

Not reviewed, for internal circulation only

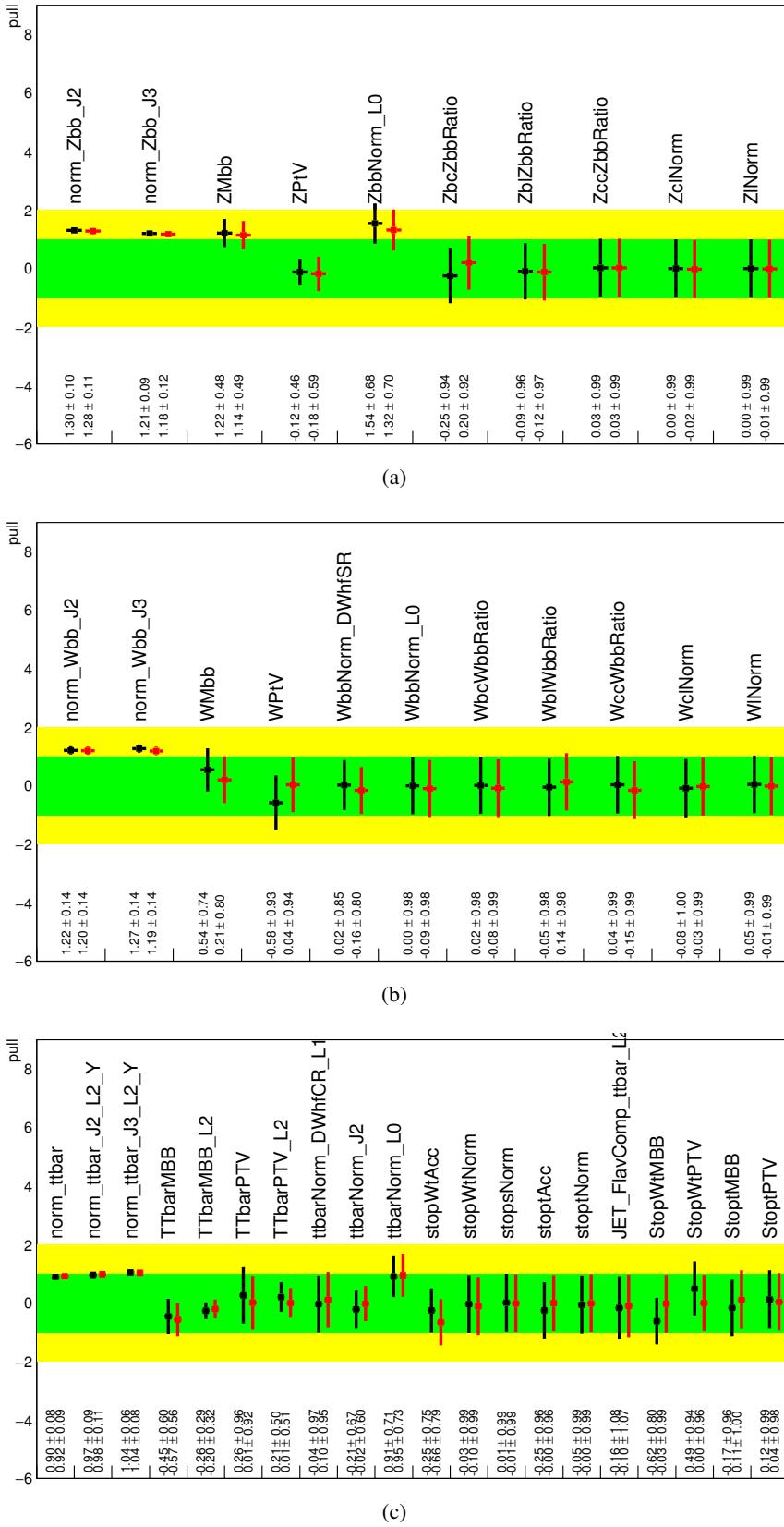


Figure 54: Pull distribution of modeling NPs in the combined unconditional VZ fit (red) compared to the VH fit (black). Z jets (a), W jets (b), top (c).

Not reviewed, for internal circulation only

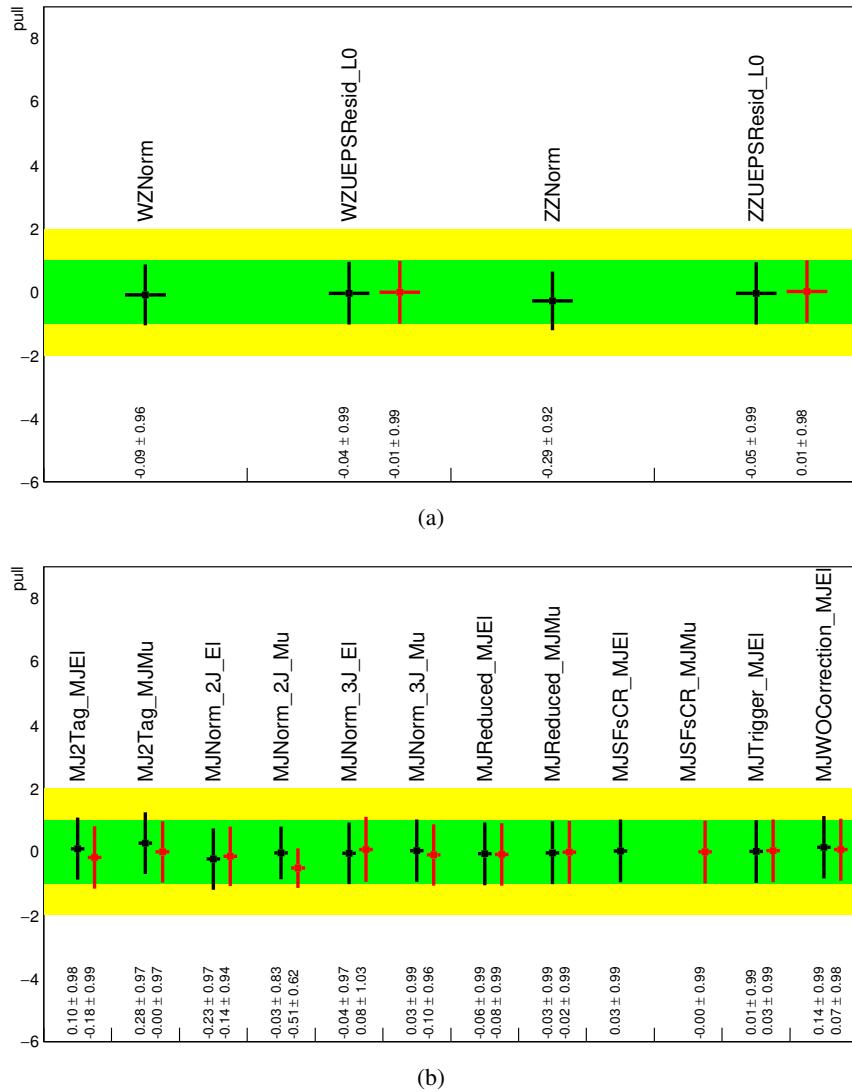


Figure 55: Pull distribution of modeling NPs in the combined unconditional VZ fit (red) compared to the VH fit (black). (a) and (b).

Not reviewed, for internal circulation only

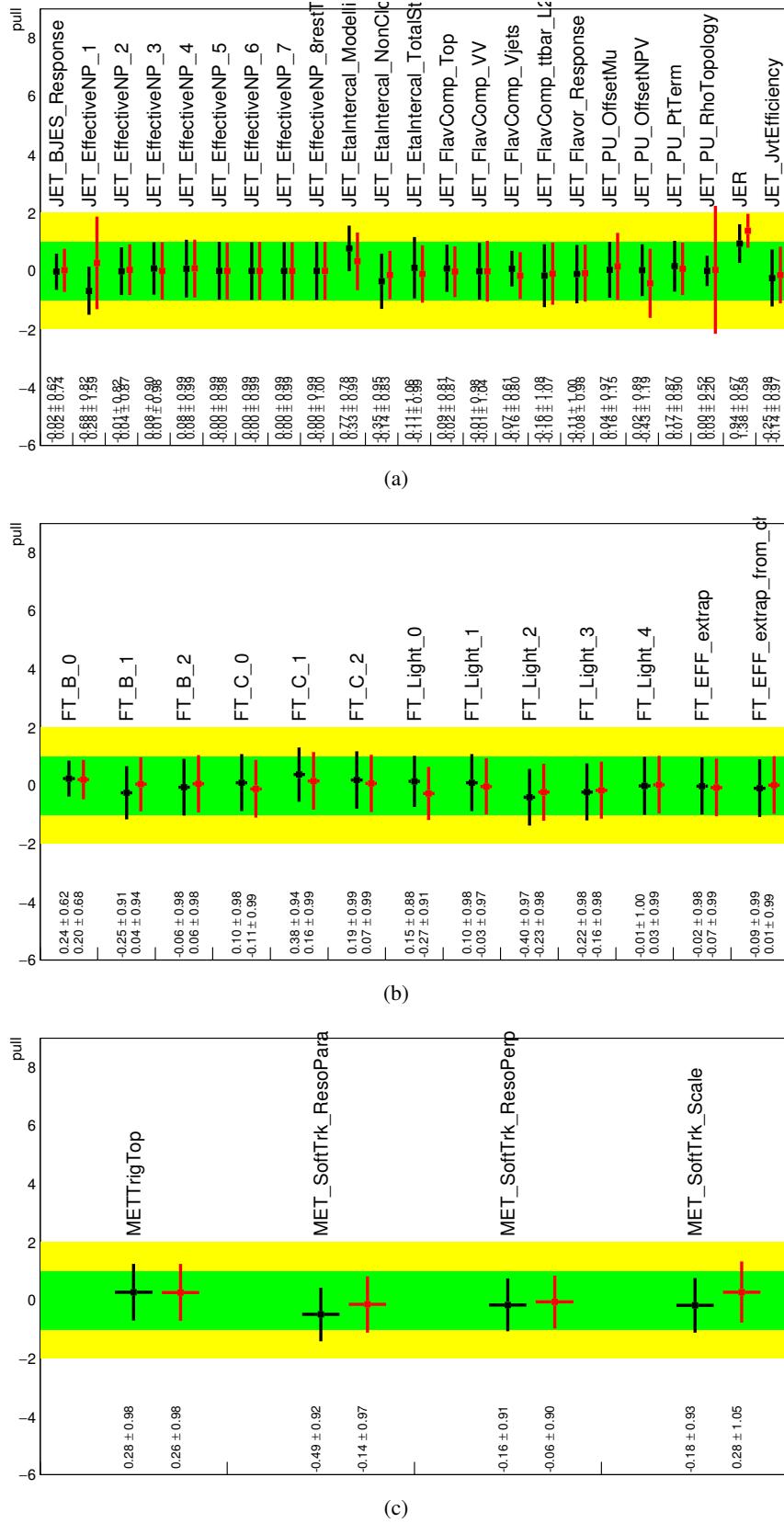


Figure 56: Pull distribution of experimental NPs in the combined unconditional VZ fit (red) compared to the VH fit (black). Jet systematics (a), b-tagging (b) and E_T^{miss} (c).

1090 **7.5.2. Correlation matrices**

1091 The correlation matrices corresponding to the combined VZ fit when fitting an Asimov dataset and data
 1092 can be found in Figure 57. Only correlations among parameters with a magnitude greater than 0.25 are
 1093 shown.

Not reviewed, for internal circulation only

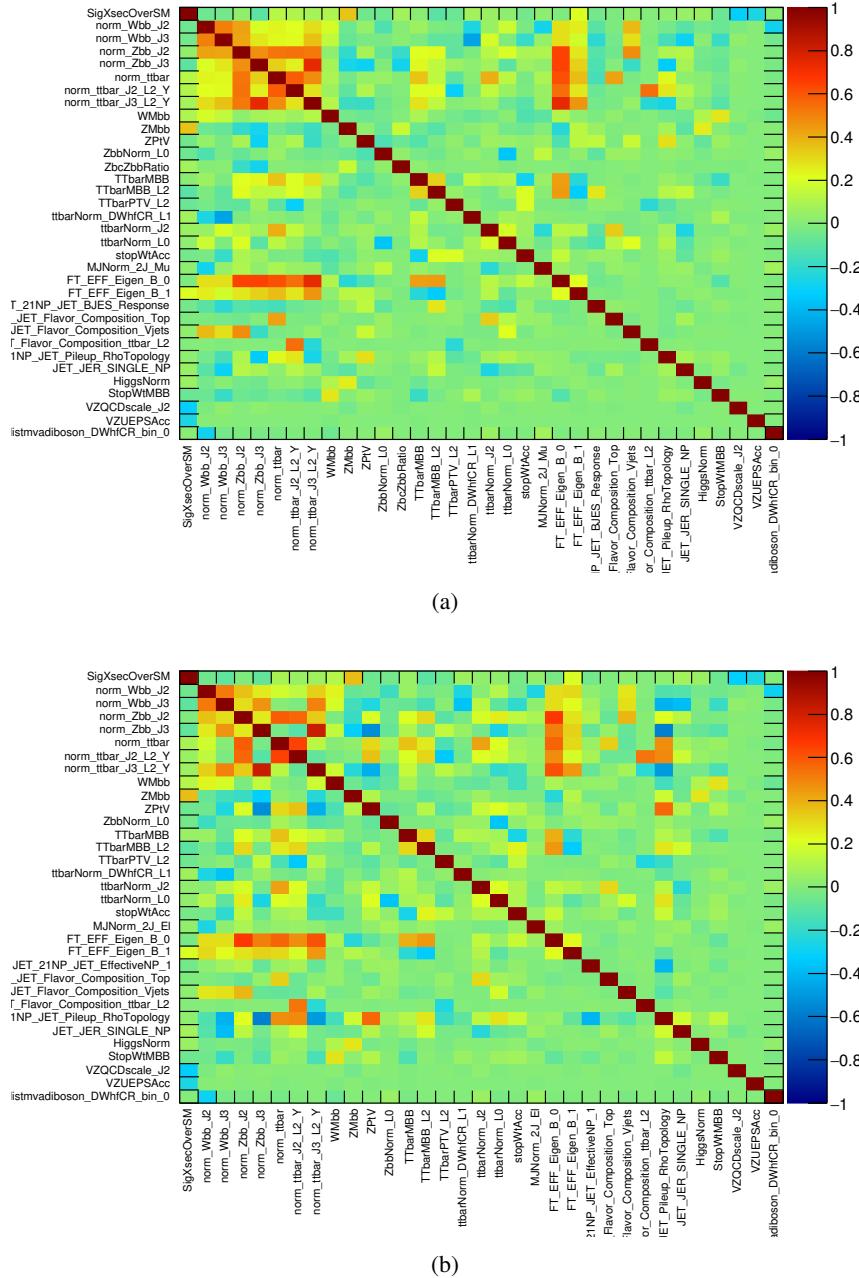


Figure 57: Correlation matrices from the unconditional combined VZ fit to an Asimov dataset (a) and to 2015 and 2016 data (b).

1094 **7.5.3. Nuisance parameter ranking**

1095 The pre-fit and post-fit ranking plots for the combined fit on data are shown in Figure 58. For the pre-fit
 1096 ranking, the value of the prior is used, which is equivalent for the nuisance parameter to $\sigma = 1$. For the
 1097 post-fit ranking, the error on the profiled nuisance parameter at the maximum of the likelihood is used.

1098 The blue boxes show the postfit impact of the nuisance parameters, which is evaluated by changing them by
 1099 their profiled error at the maximum of the likelihood. The change in the best fit μ value gives the inclusive
 1100 sensitivity of the measured value to the given NP. The yellow bands correspond to the so-called prefit
 1101 impact of the nuisance parameters, which is evaluated by changing the values of the nuisance parameters
 1102 by 1 (which is the prefit uncertainty on the NP). By construction, floating parameters (without any prior)
 1103 do not have any prefit impact. Only the systematics with the largest impact are shown.

Not reviewed, for internal circulation only

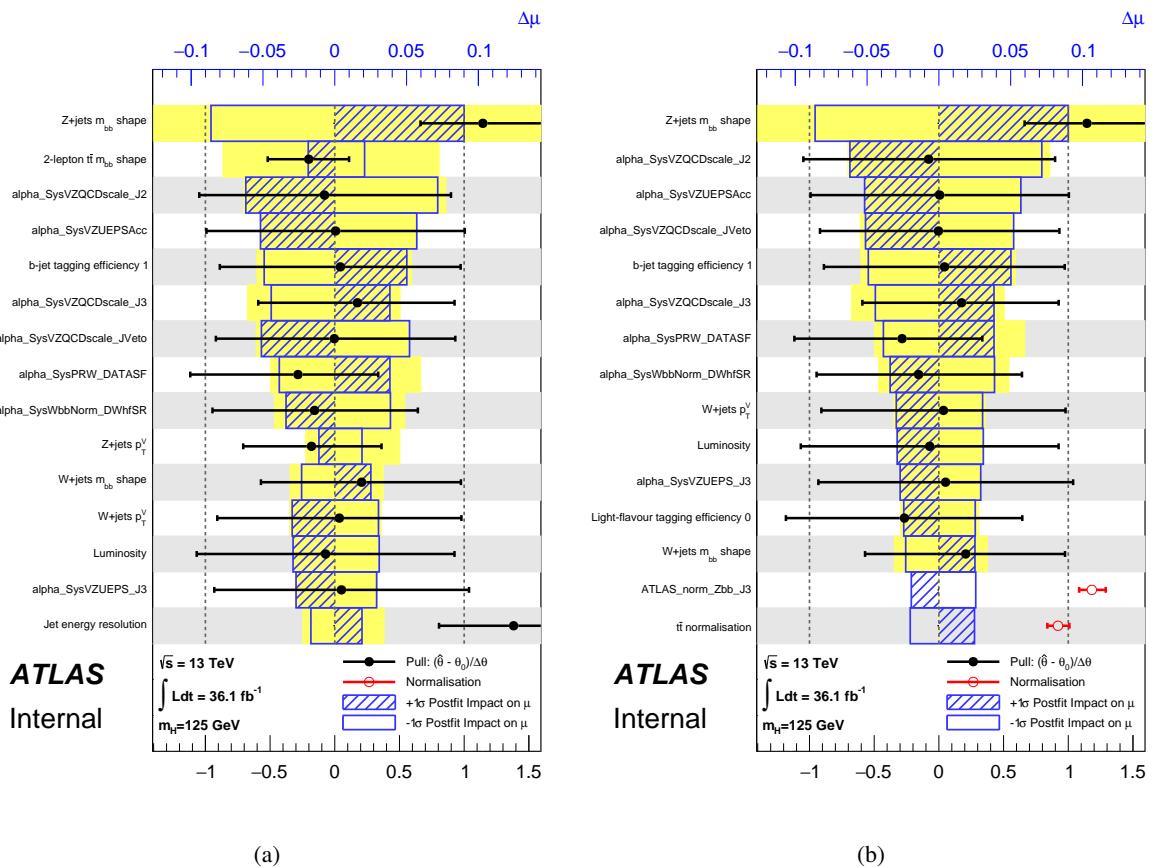


Figure 58: The nuisance parameter ranking from the unconditional VZ fit to data in the combined fit, ordered by pre-fit impact (a), when ordered by post-fit impact (b).

1104 **7.5.4. Breakdown of statistical and systematic uncertainties**

1105 Statistical uncertainties due to the limited sizes of the simulated MC samples are taken into account in
1106 the fit by means of nuisance parameters, which are parametrized by Poisson priors Based on the study
1107 described in Appendix F the binning on the final discriminant is chosen so that the MC stat uncertainties
1108 for all bins is less than 20%.

1109 The fitted parameter of interest $\hat{\mu}$ and the breakdown of the uncertainties coming from data statistics (“Data
1110 Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only
1111 (“MC Stat.”), and each category of systematic uncertainty, is shown for the combined unconditional fit on
1112 36.1 fb^{-1} of data in Table 41.

Set of nuisance parameters	Impact on error
Total	+0.251 / -0.222
DataStat	+0.116 / -0.114
FullSyst	+0.223 / -0.191
Floating normalizations	+0.069 / -0.059
All normalizations	+0.082 / -0.071
All but normalizations	+0.207 / -0.173
Jets MET	+0.045 / -0.034
BTag	+0.067 / -0.057
Leptons	+0.007 / -0.004
Luminosity	+0.041 / -0.029
Diboson	+0.056 / -0.038
Zjets	+0.098 / -0.089
Wjets	+0.066 / -0.062
Model ttbar	+0.048 / -0.038
Model Single Top	+0.016 / -0.015
Model Multi Jet	+0.016 / -0.012
Signal Systematics	+0.004 / -0.004
MC stat	+0.072 / -0.068

Table 41: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and different sub-groups of systematic uncertainties, is shown, as well as the fitted $\hat{\mu}$ extracted from the combined unconditional VZ fit on 36.1 fb^{-1} of data.

1113 **7.5.5. Post-fit BDT distributions**

1114 In Figures 59, 60, and 62, the postfit BDT distributions in the 0-, 1-, and 2-lepton channels are shown,
 1115 as obtained from the combined fit on 2015 and 2016 data. The overall agreement between data and MC
 1116 distributions is very good.

Not reviewed, for internal circulation only

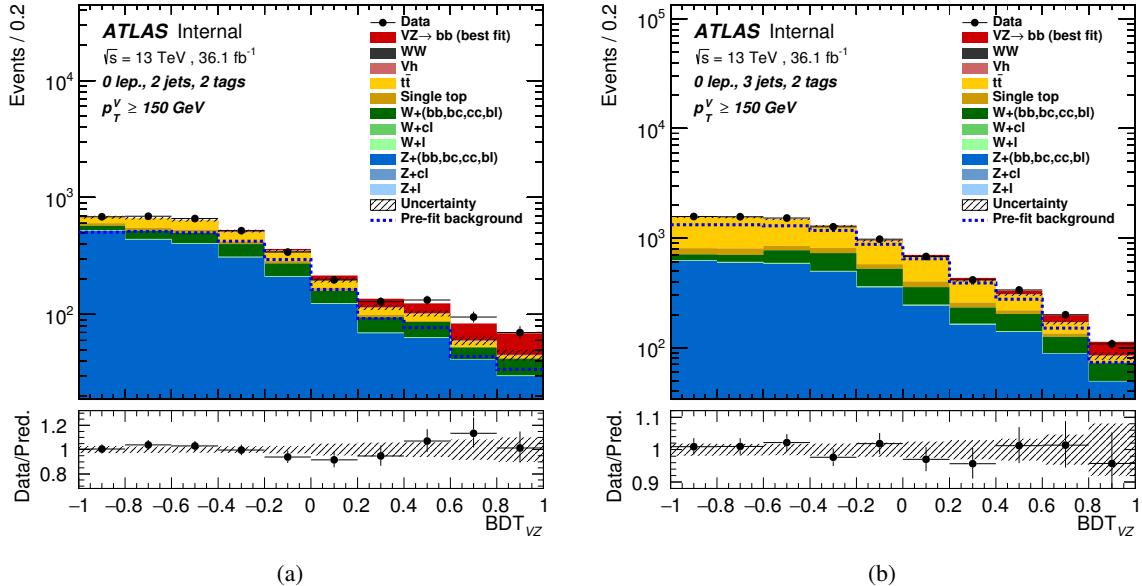


Figure 59: BDT postfit distributions obtained in the 0-lepton channel for the combined unconditional VZ fit on 2015 and 2016 data. The best-fit signal value is used.

Not reviewed, for internal circulation only

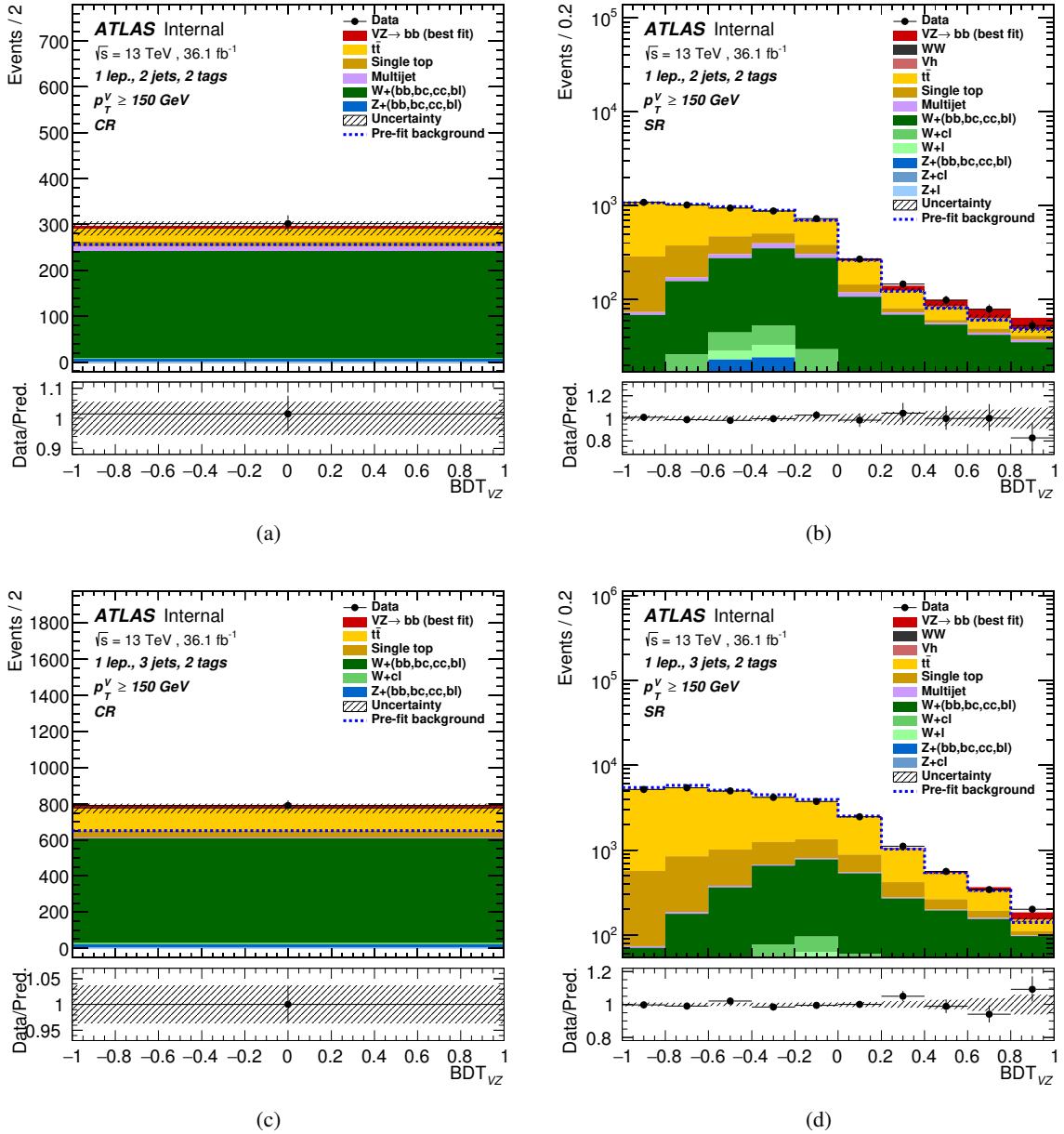


Figure 60: BDT postfit distributions obtained in the 1-lepton channel for the combined unconditional VZ fit on 2015 and 2016 data. The best-fit signal value is used.

Not reviewed, for internal circulation only

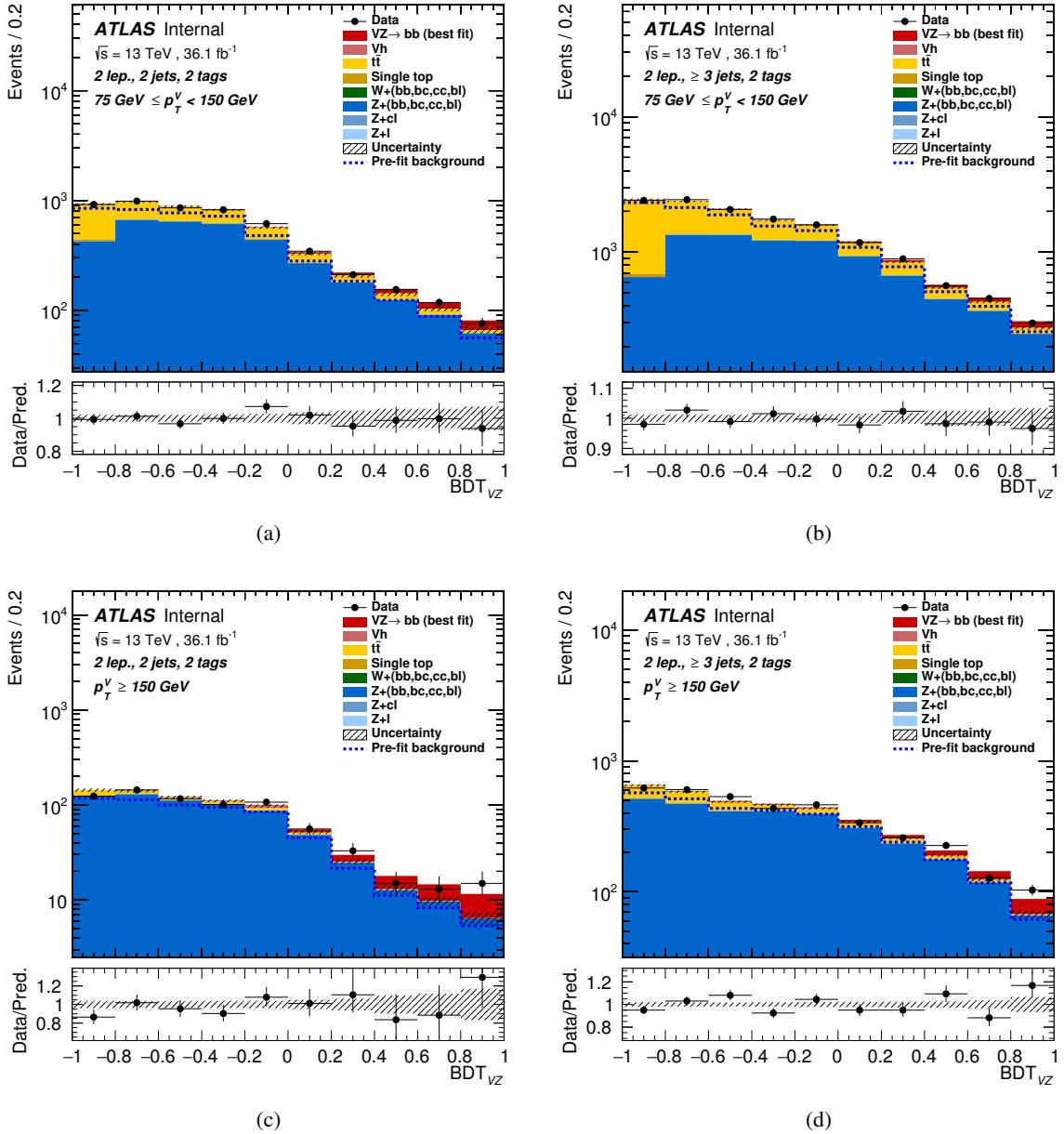


Figure 61: BDT postfit distributions obtained in the 2-lepton channel for the combined unconditional VZ fit on 2015 and 2016 data. The best-fit signal value is used.

Not reviewed, for internal circulation only

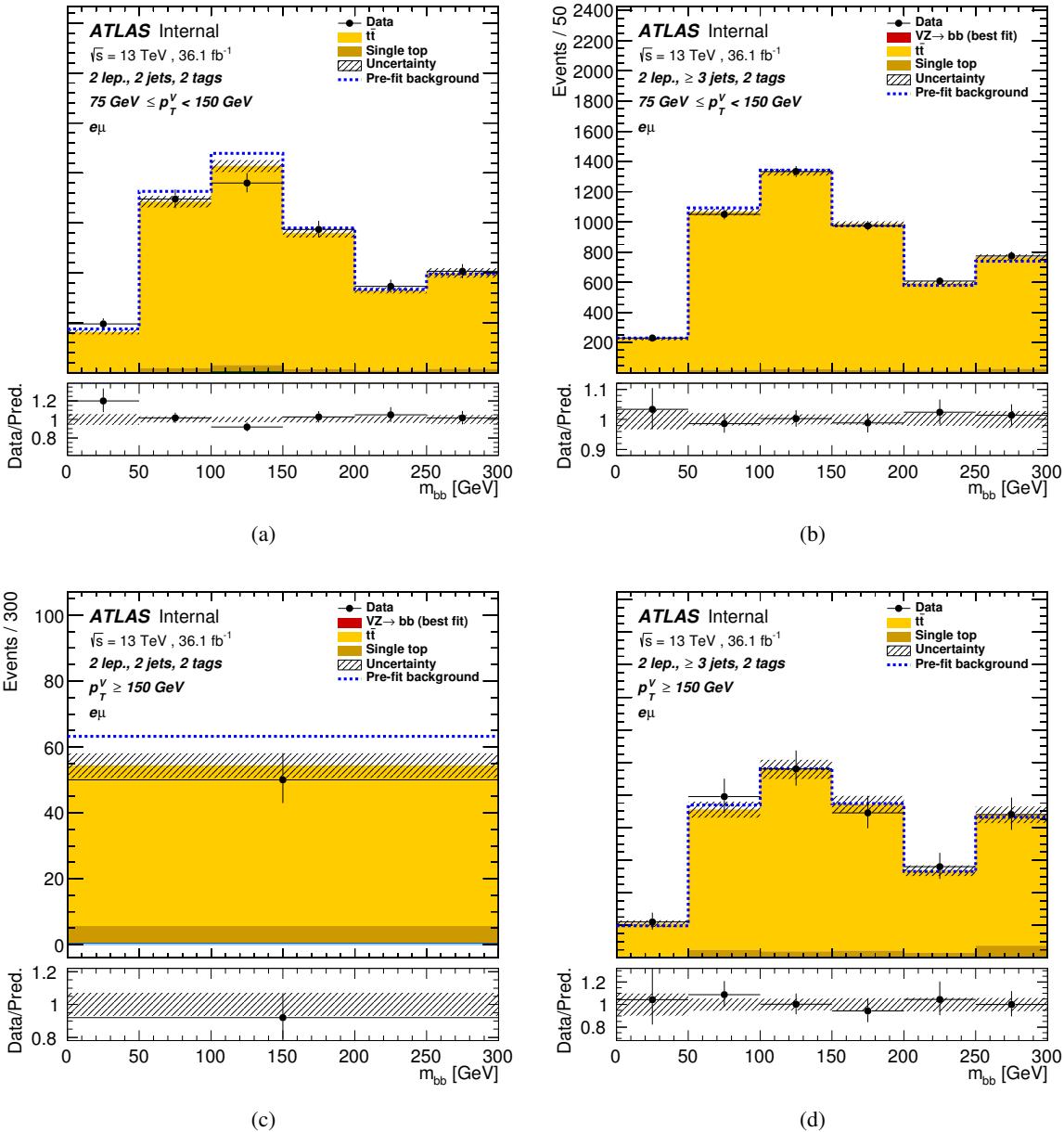


Figure 62: m_{bb} postfit distributions obtained in the 2-lepton channel $e - \mu$ control regions for the combined unconditional VZ fit on 2015 and 2016 data. The best-fit signal value is used.

1117 7.5.6. Signal over Background Plots

1118 In Figure 63 all BDT bins are summed up across all channel and regions used in the combined fit, and
 1119 ordered by their signal to background ratio is shown, with the signal strength taken from the best-fit value.
 1120 The same distribution is shown three times, associated to ratios computed as (a) data over expected back-
 1121 ground ($data/background$), (b) background-subtracted data over signal ($(data - background)/signal$),
 1122 and (c) background subtracted data over square-root of background ($(data - background)/\sqrt{background}$).

Not reviewed, for internal circulation only

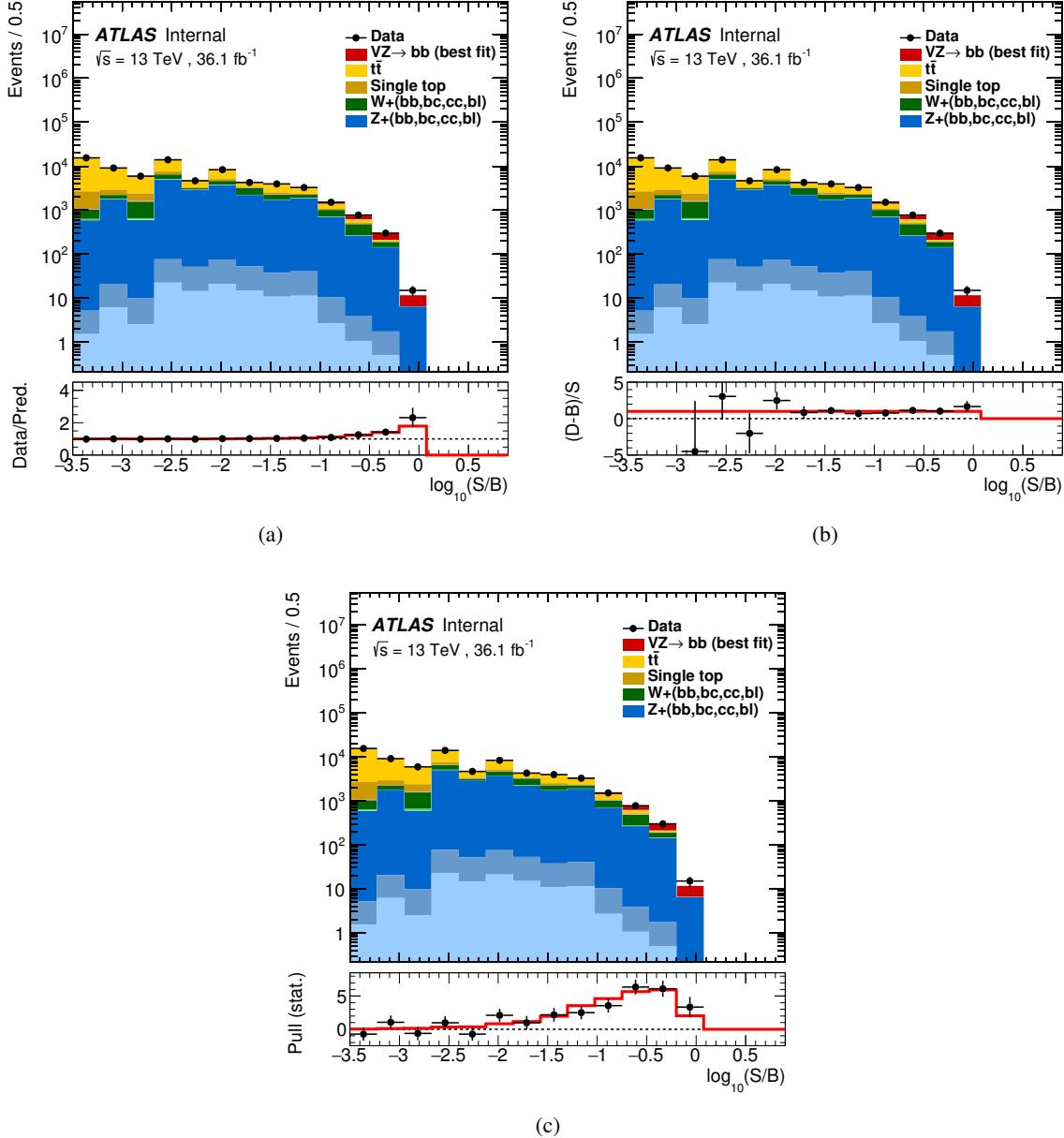


Figure 63: All BDT bins summed up across all regions and channels used in the combined unconditional VZ fit, and ordered by their signal to background ratio is shown. The signal strength taken from the best-fit value. The same distribution is shown three times, associated to ratios computed as (a) data over expected background ($\text{data}/\text{background}$), (b) background-subtracted data over signal ($(\text{data} - \text{background})/\text{signal}$), and (c) background subtracted data over square-root of background ($(\text{data} - \text{background})/\sqrt{\text{background}}$).

1124 8. Run 2 Results

1125 8.1. VH results

1126 Table 42 shows the expected significance estimated from an Asimov dataset corresponding to 36.1 fb^{-1}
 1127 and from the fit to data using the signal regions described in the previous section. A detailed version of
 1128 this table, outlining the VH signal significance of each fitting region is provided in Section J.5.

Channel	Exp. significance (Asimov)	Exp. significance (data)	Obs. Significance
0-lepton (SR)	1.99	1.73	0.53
1-lepton (SR+CR)	1.81	1.81	2.30
2-lepton (SR+CR)	1.95	1.86	3.55
0,1,2-lepton (SR+CR)	3.19	3.03	3.54

Table 42: Observed significance from an unconditional fit to the data corresponding to 36.1 fb^{-1} and expected significances from a fit to an Asimov dataset and from a fit to the data. Expected significances from individual regions are estimated separately.

1129 In Figure 64, the μ obtained in the combined VH fit to data is compared to the ones obtained when running
 1130 the fit in each of the three leptonic channels and from a 3 POI fit.

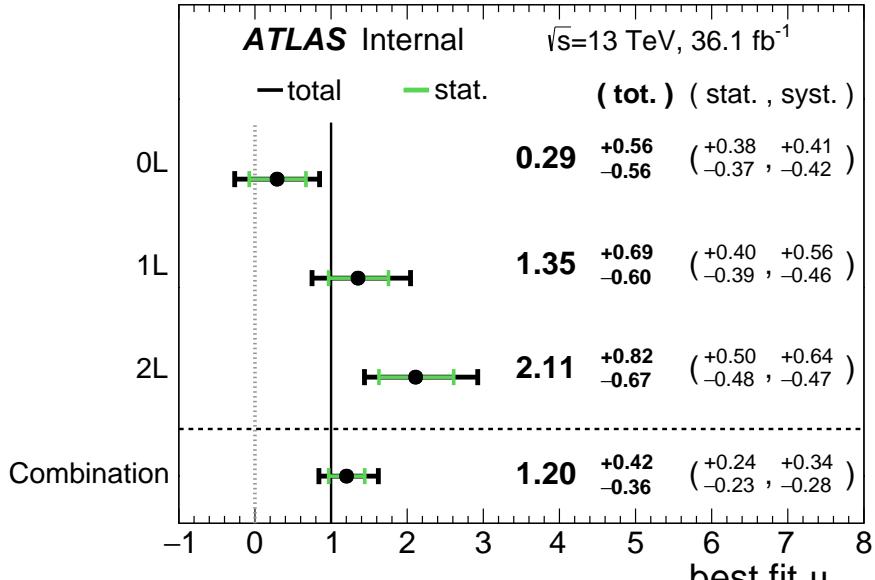
1131 In Figure 65, the μ obtained in the combined fit for the WH and ZH processes is presented. In this fit,
 1132 two independent floating μ parameters (PoIs) for each of WH and ZH processes are obtained from a
 1133 simultaneous fit.

1134 In Figure 66, the μ obtained in the combined fit with separate POIs for each leptonic channel are compared
 1135 to the combined fit μ obtained when running the fit in each of three leptonic channels. Good compatibility
 1136 between the 3 POI fit and individual channel fits is observed.

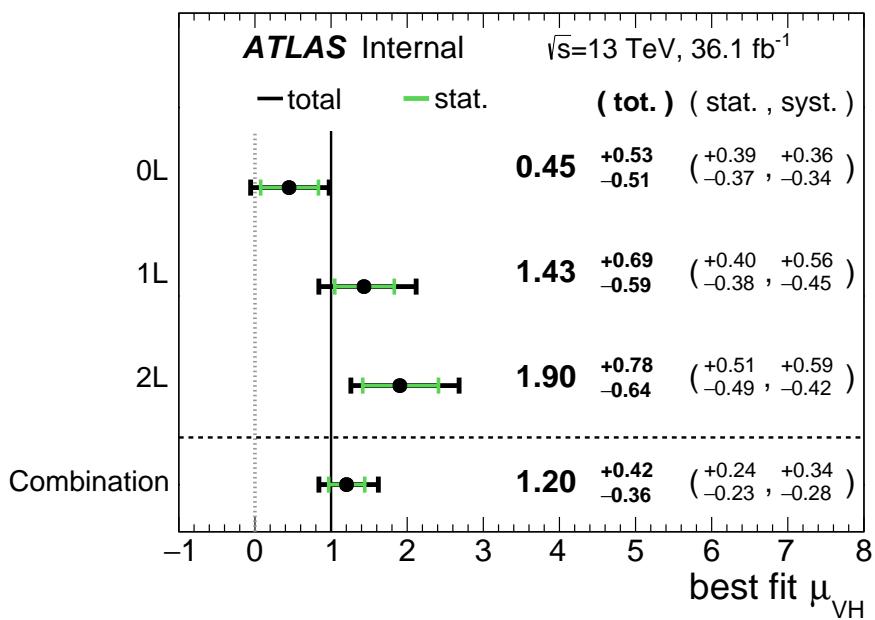
1137 The compatibility between the signal strength obtained in the lepton channels inside the combined fit can
 1138 be evaluated. Given the same fit model is fitted to the data, the negative log likelihood (NLL) values of a
 1139 combined fit with 1 common parameter of interest or 3 parameters of interest (1 per lepton channel) are
 1140 related to each other. The distribution of $2 \times \Delta\text{NLL}$ should follow a χ^2 distribution with $3-1=2$ degrees
 1141 of freedom. This allows to evaluate a compatibility probability. For the VH combined fit, it has been
 1142 estimated to be 10.3%.

1143 Compatibility with the cut-based analysis results is presented in Appendix M.

Not reviewed, for internal circulation only



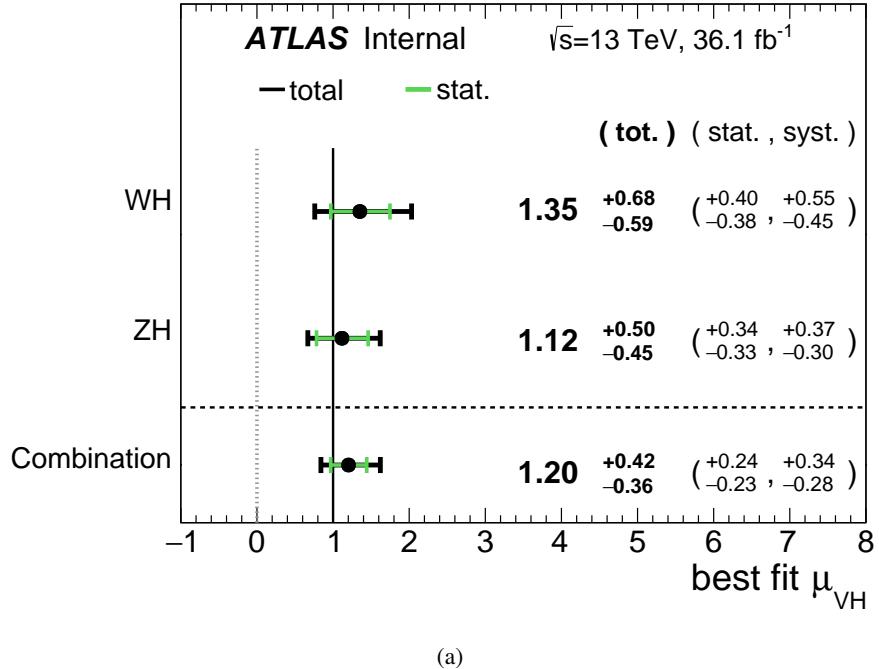
(a)



(b)

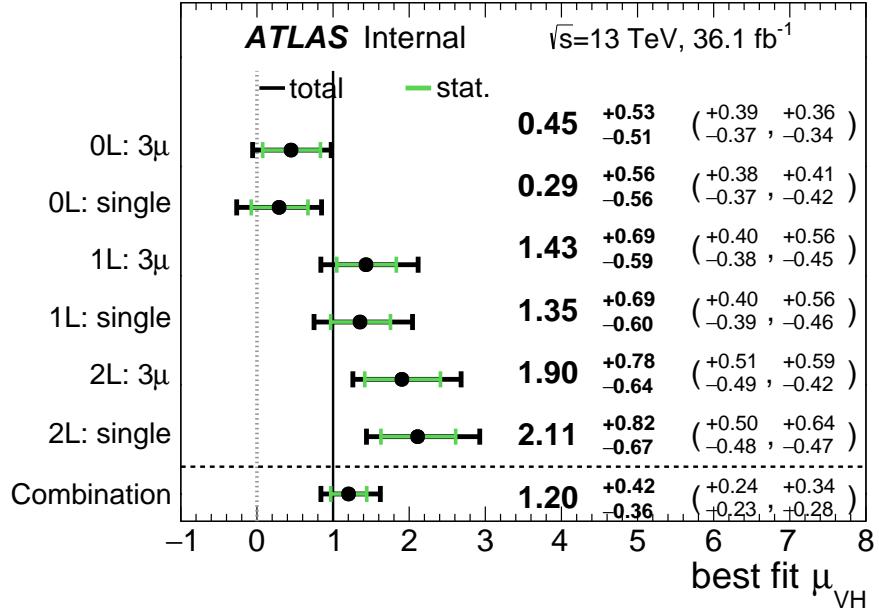
Figure 64: The μ obtained in the combined VH fit to data is compared to the values obtained when running separate fits in each of the three leptonic channels (a) and when running the combined fit with the signal strength floated separately for each of the leptonic channels (a 3 POI fit), (b).

Not reviewed, for internal circulation only



(a)

Figure 65: The μ obtained from combined fit for the WH and ZH processes and the combination of them. The individual μ value for WH/ZH processes are obtained with separate PoIs.



(a)

Figure 66: The μ obtained from combined fit for the 0-, 1- and 2-lepton processes when running a 3 POI fit, fitting to each individual channel, and the combination of them.

¹¹⁴⁴ **8.2. VZ Diboson results**

¹¹⁴⁵ Table 43 shows the expected significances estimated from an Asimov dataset corresponding to 36.1 fb^{-1}
¹¹⁴⁶ and from the fit to data using the signal regions described in the previous section for the VZ diboson fit.

¹¹⁴⁷ In 1-lepton, the VZ significance is generally poorer than the other channels, while it's not the case for
¹¹⁴⁸ VH. Given results the ranking and breakdown, the main reason is that there is a larger contamination of
¹¹⁴⁹ W + hf background under WZ, while it is simultaneously the least constrained background (especially
¹¹⁵⁰ with large p_T^V and m_{bb} systematics). Under the Higgs signal, there is a lower W + hf contamination and
¹¹⁵¹ more $t\bar{t}$.

¹¹⁵² A detailed version of this table, outlining the VZ signal significance of each fitting region is provided in
¹¹⁵³ Section L.4.

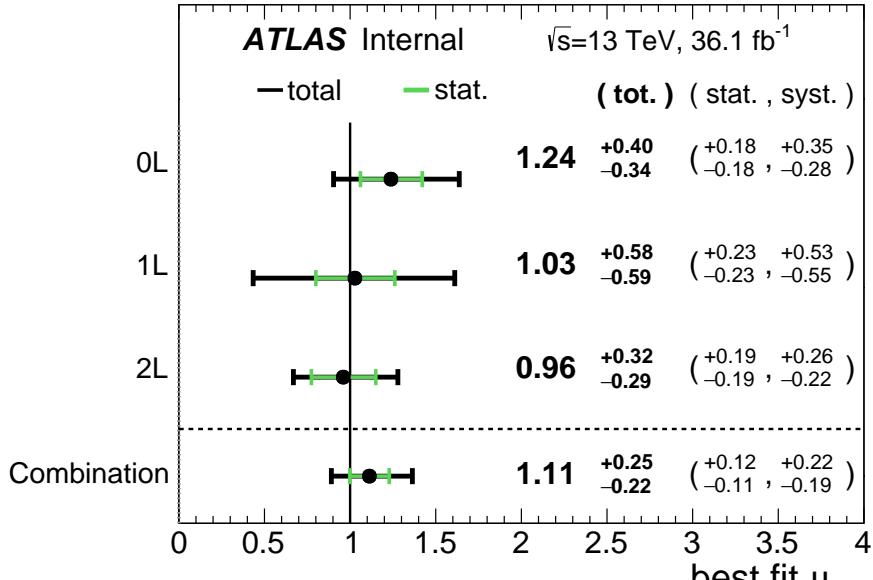
Channel	Exp. significance (Asimov)	Exp. significance (data)	Obs. Significance
0-lepton (SR)	3.77	3.61	4.16
1-lepton (SR+CR)	1.89	1.79	1.69
2-lepton (SR+CR)	3.79	3.48	3.45
0,1,2-lepton (SR+CR)	5.86	5.25	5.76

Table 43: Expected significances from a fit to an Asimov dataset corresponding to 36.1 fb^{-1} and from a fit to the data, treating VZ as the signal. Expected significances from individual regions are estimated separately.

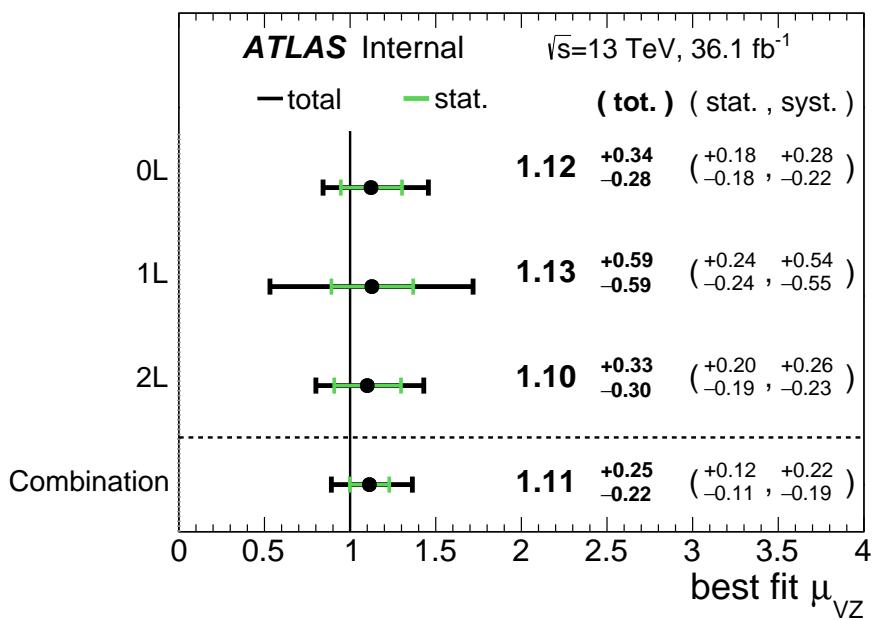
¹¹⁵⁴ In Figure 67, the μ obtained in the combined VZ fit to data is compared to the ones obtained when running
¹¹⁵⁵ the fit in each of the three leptonic channels and from a 3 POI fit.

¹¹⁵⁶ In Figure 68, the μ obtained in the combined fit for the WZ and ZZ processes is presented. In this fit,
¹¹⁵⁷ two independent floating μ parameters (PoIs) for each of WZ and ZZ processes are obtained from a
¹¹⁵⁸ simultaneous fit.

Not reviewed, for internal circulation only



(a)



(b)

Figure 67: The μ obtained in the combined VH fit to data is compared to the values obtained when running separate fits in each of the three leptonic channels (a) and when running the combined fit with the signal strength floated separately for each of the leptonic channels (a 3 POI fit), (b).

Not reviewed, for internal circulation only

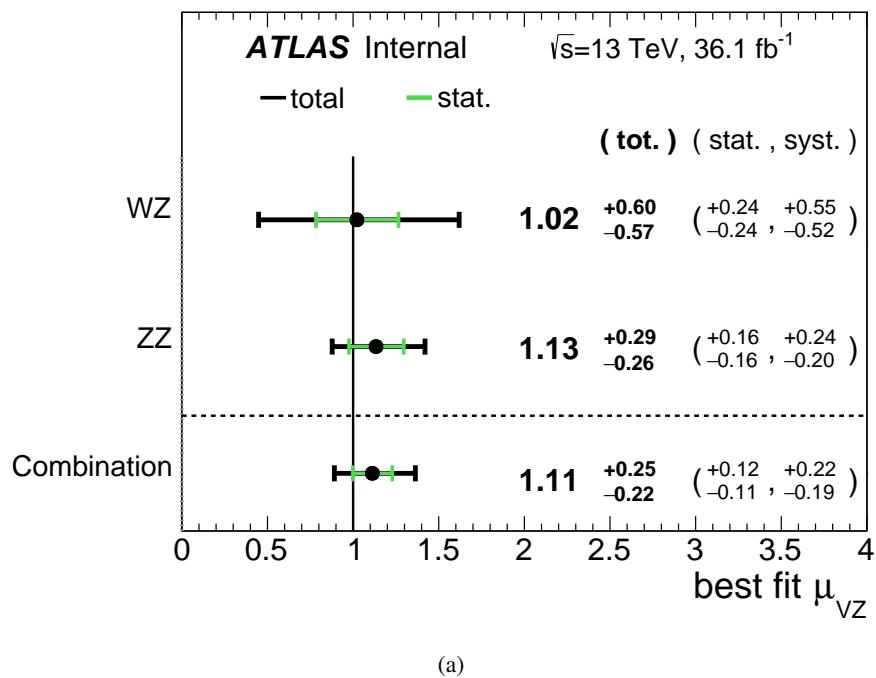


Figure 68: The μ obtained from combined fit for the WZ and ZZ processes and the combination of them. The individual μ value for WZ/ZZ processes are obtained with separate PoIs (to-be-updated).

9. Combination of Run 1 and Run 2 results

The Run1 ($\sqrt{s} = 7$ TeV with 4.7 fb^{-1} of data, and $\sqrt{s} = 8$ TeV with 20.3 fb^{-1} of data) and Run2 workspaces are combined according to the correlation scheme in Table 44. The impact of correlating experimental and modelling nuisance parameters (NP's) was checked and found to have negligible impact on fit results (details may be found in Appendix N). Hence, only signal NP's and the b -jet energy scale are correlated.

7 TeV NP	8 TeV NP	13 TeV NP
ATLAS_BR_bb		SysTheoryBRbb
SysTheoryQCDscale_ggZH		SysTheoryQCDscale_ggZH
SysTheoryQCDscale_qqVH		SysTheoryQCDscale_qqVH
—	SysTheoryPDF_ggZH_8TeV	SysTheoryPDF_ggZH
—	SysTheoryPDF_qqVH_8TeV	SysTheoryPDF_qqVH
—	SysTheoryVHPt_8TeV	SysVHNLOEWK
SysJetFlavB_7TeV	SysJetFlavB_8TeV	SysJET_21NP_JET_BJES_Response

Table 44: A summary of correlated nuisance parameters among the 7, 8, and 13 TeV datasets.

The combined results yields an observed (expected) significance of 3.57 (4.00) and an observed (expected) limit of 1.37 ($0.510^{+0.200}_{-0.143}$), with a signal strength of $\hat{\mu} = 0.898^{+0.278}_{-0.261}$.

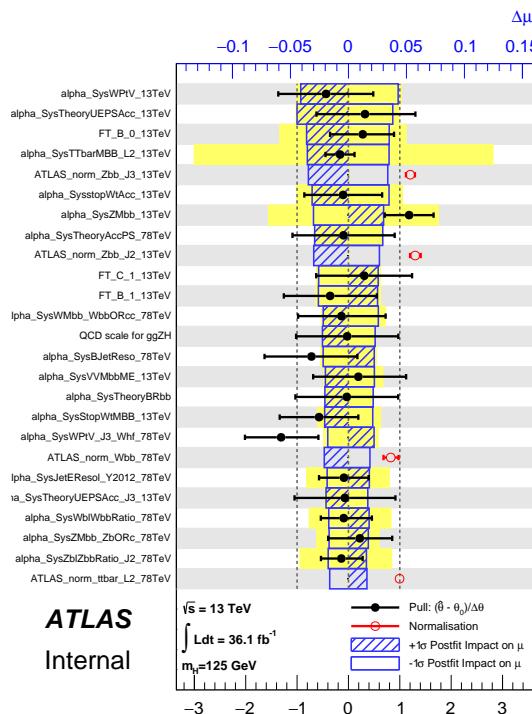


Figure 69: Ranked nuisance parameters for the Run1+Run2 combination.

Two (34.2% compatibility with single parameter of interest fit) and three (1.49% compatibility with single parameter of interest fit) parameter of interest fits were also conducted, and summaries of $\hat{\mu}$ for these fits

¹¹⁶⁸ and for 7, 8, and 13 TeV may be found in Figures 70-72. The main results for Run 1, Run 2, and the
¹¹⁶⁹ combination may be found in Table 45.

Dataset	$\hat{\mu}$	Total Error in $\hat{\mu}$	Obs. (Exp.)	Significance
Run 1	0.51	+0.40 / -0.37		1.4 (2.6)
Run 2	1.20	+0.42 / -0.36		3.54 (3.03)
Combined	0.90	+0.28 / -0.26		3.57 (4.00)

Table 45: A summary of main results for the Run 1, Run 2, and combined fits.

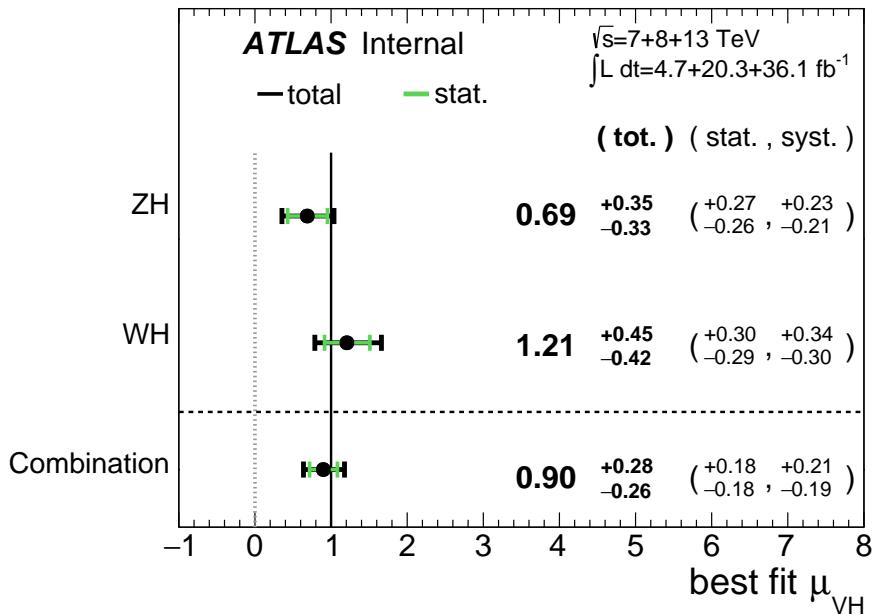


Figure 70: $\hat{\mu}$ summary plot for a two parameter of interest fit.

Not reviewed, for internal circulation only

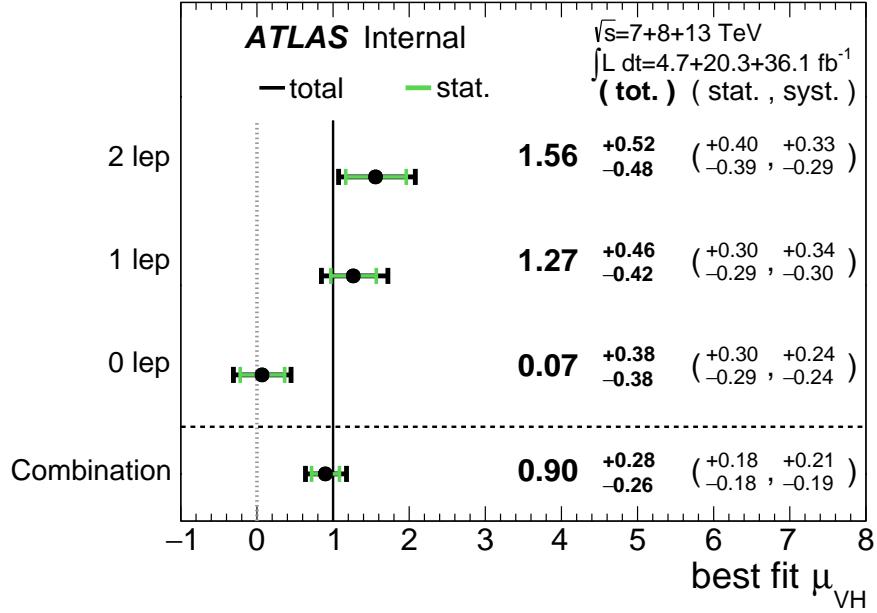


Figure 71: $\hat{\mu}$ summary plot for a three parameter of interest fit.

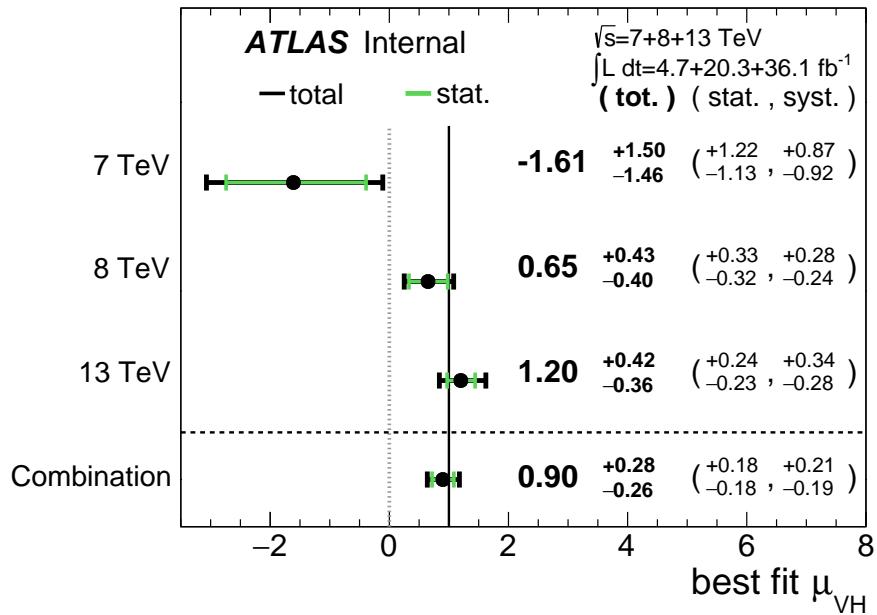


Figure 72: $\hat{\mu}$ summary plot for different \sqrt{s} values.

1170 10. Conclusions

1171 A search for the decay of a Standard Model Higgs boson into $b\bar{b}$ pair when produced in association with a
1172 W and Z boson has been performed with Run1 and Run2 full data. In Run2 dataset, the measured signal
1173 strength with respect to the SM expectation is found to be $1.20^{+0.24}_{-0.23}(\text{stat.})^{+0.34}_{-0.28}(\text{syst.})$ at $m_H = 125$ GeV. The
1174 observed (expected) significance is 3.5(3.0) standard deviations. The analysis has been validated by meas-
1175 uring the signal strength of $V(W/Z)Z(\rightarrow b\bar{b})$. The measured signal strength is $1.11^{+0.12}_{-0.11}(\text{stat.})^{+0.22}_{-0.19}(\text{syst.})$,
1176 corresponding to observed(expected) significance of 5.8(5.3) standard deviations.

1177 This result based on Run2 data has been also combined with previous results on the Run1 dataset. The
1178 observed(expected) significance combined Run1 and Run2 is 3.6(4.0) standard deviation. The measured
1179 signal strength is $0.90^{+0.18}_{-0.18}(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$.

¹¹⁸⁰ **Appendices**

[Not reviewed, for internal circulation only]

1181 A. 0 Lepton Multijet distributions

1182 In this section the PYTHIA 8 MC samples (A14 tune with NNPDF2.3LO PDFs) (listed in Table 46) are
 1183 included in the MC background estimation in the 0 lepton channel without the anti-QCD cuts. The
 1184 samples are split in 13 orthogonal slices according to the true p_T of the leading jet, going from 20 GeV to
 1185 infinity, and merged with the corresponding crosssections. In Figures 73, 74, and 75 the distributions of
 1186 (a) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \mathbf{E}_{T,\text{trk}}^{\text{miss}})$, (b) $\Delta\Phi(jet1, jet2)$, (c) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, h)$, and (d) $\min[\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \text{pre-sel. jets})]$ are shown
 1187 for the 0, 1, and 2 tag categories for inclusive number of jets. Despite some spikes in the multijet simulated
 1188 sample, due to few events with very large normalization weights, the agreement in shape between data
 1189 and MC is rather good, and MC tends to overestimate the multijet yield.

Table 46: List of MC QCD samples used to study the multi-jet distributions in 0 lepton

Sample	Events
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ0W	1515
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ1W	7689
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2W	162441
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3W	2826351
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4W	6918784
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ5W	7976971
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ6W	1893399
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ7W	1770198
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ8W	1743197
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ9W	1813197
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ10W	1995998
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ11W	1993199
Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ12W	1974598

Not reviewed, for internal circulation only

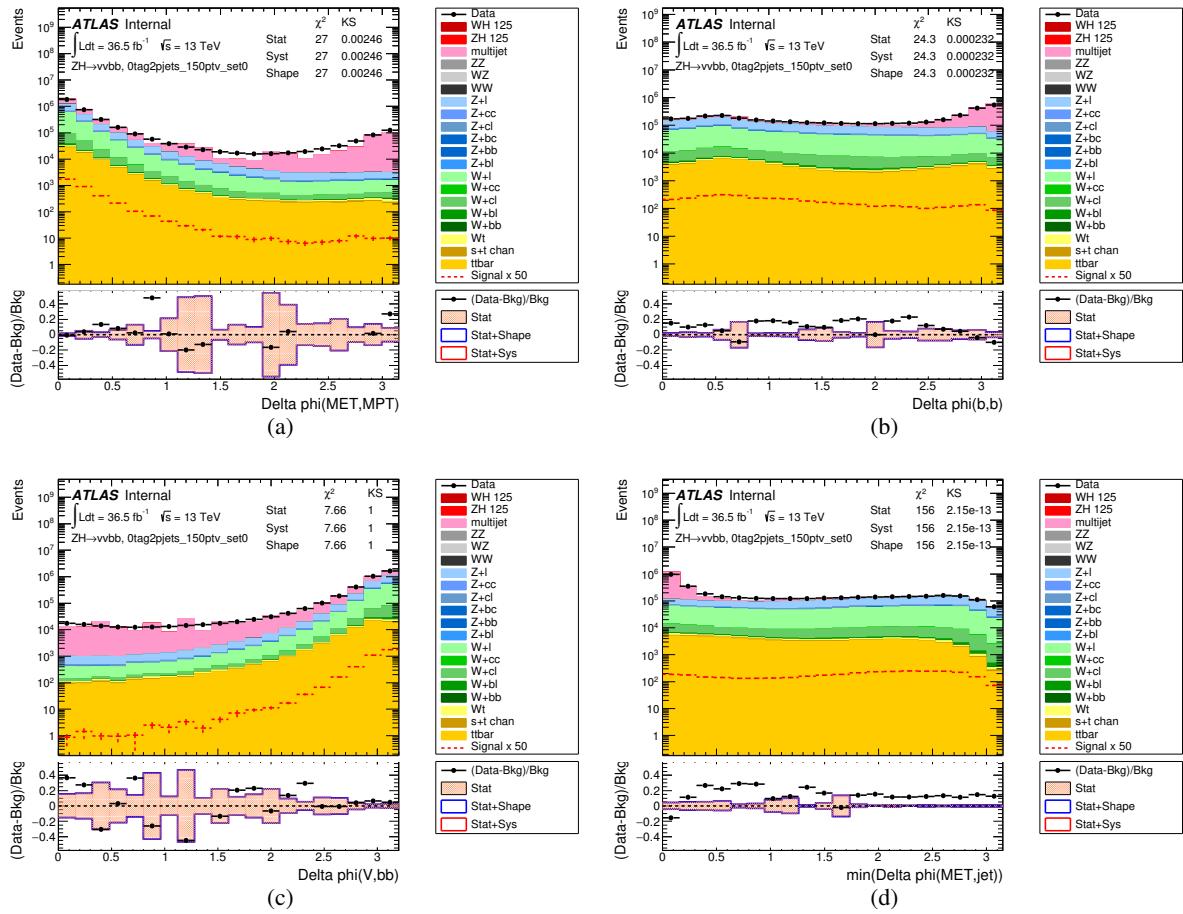


Figure 73: The distributions, for the 0-tag category, and inclusive in number of jets, of (a) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \mathbf{E}_{T,\text{trk}}^{\text{miss}})$, (b) $\Delta\Phi(jet_1, jet_2)$, (c) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, h)$, (d) and $\min[\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \text{pre-sel. jets})]$ are shown after the full 0-lepton channel selection, except for the anti-QCD cuts, is applied.

Not reviewed, for internal circulation only

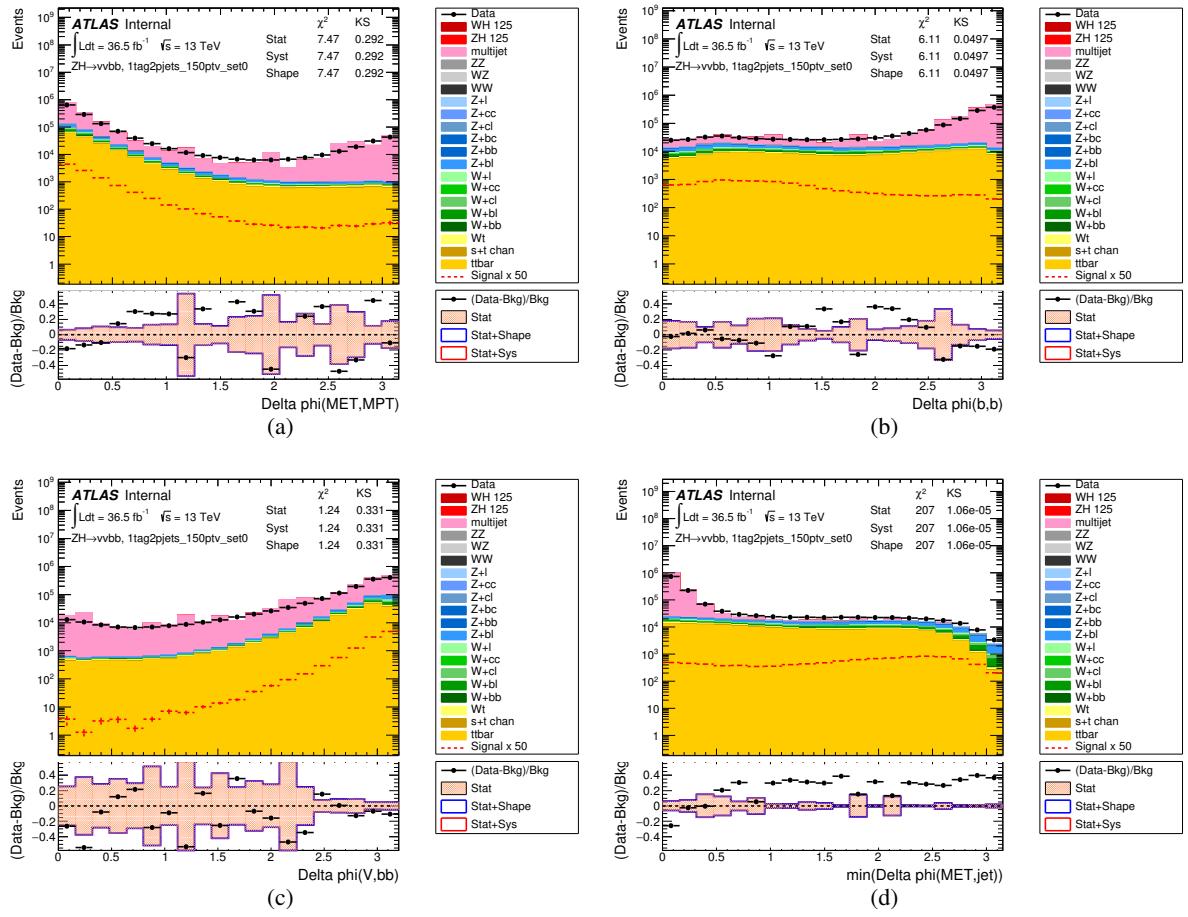


Figure 74: The distributions, for the 1-tag category, and inclusive in number of jets, of (a) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \mathbf{E}_{T,\text{trk}}^{\text{miss}})$, (b) $\Delta\Phi(jet1, jet2)$, (c) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, h)$, (d) and $\min[\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \text{pre-sel. jets})]$ are shown after the full 0-lepton channel selection, except for the anti-QCD cuts, is applied.

Not reviewed, for internal circulation only

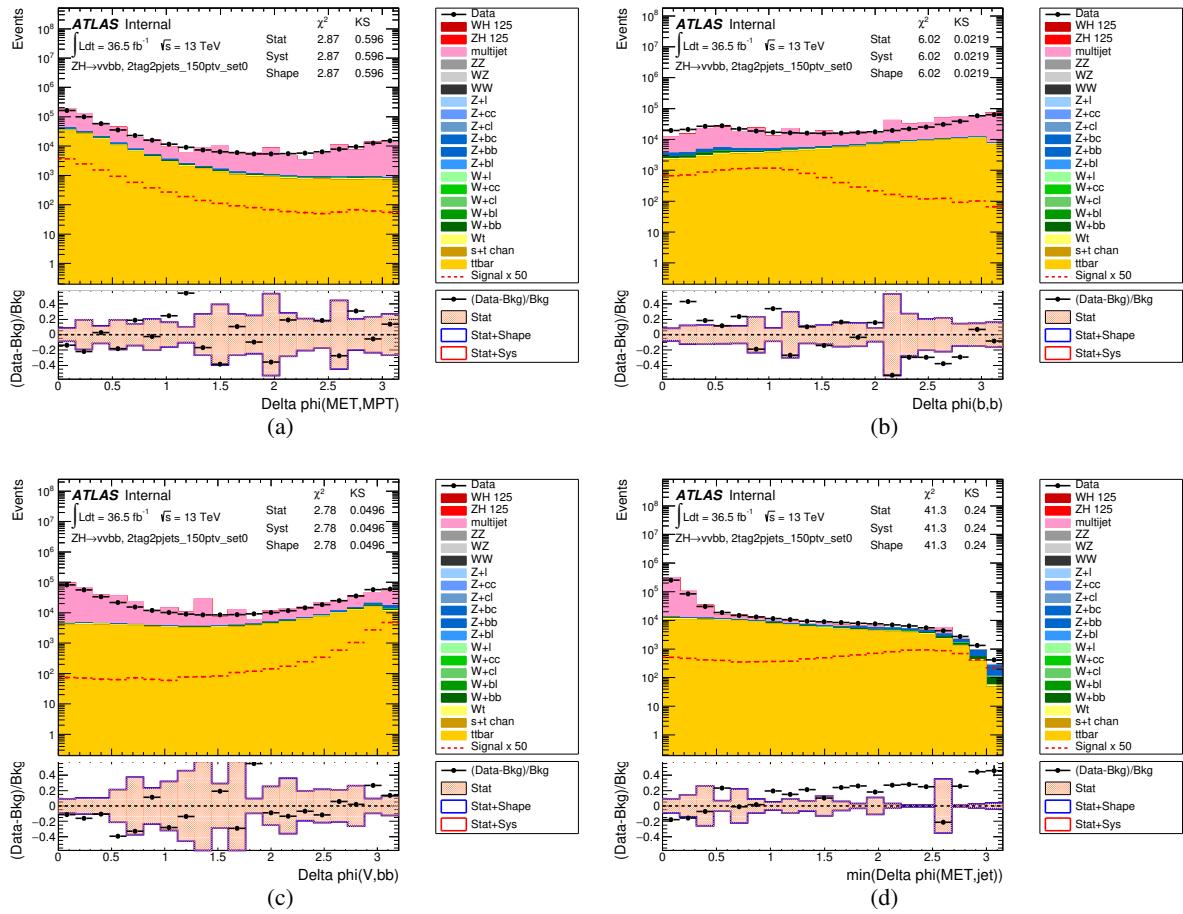


Figure 75: The distributions, for the 2-tag category, and inclusive in number of jets, of (a) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \mathbf{E}_{T,\text{trk}}^{\text{miss}})$, (b) $\Delta\Phi(jet1, jet2)$, (c) $\Delta\Phi(\mathbf{E}_T^{\text{miss}}, h)$, (d) and $\min[\Delta\Phi(\mathbf{E}_T^{\text{miss}}, \text{pre-sel. jets})]$ are shown after the full 0-lepton channel selection, except for the anti-QCD cuts, is applied.

1190 B. 2 Lepton Multijet studies

1191 B.1. p_T^V dependence

1192 Figure 76 shows the comparison of several kinematic distribution between Multi-jet and VH signal.
 1193 Table 47 shows p_T^V efficiency for each category. Extracted multi-jet p_T^V distribution is softer than signal.
 1194 Multi-jet contamination can be suppressed in high p_T^V region which is most sensitive category.

Not reviewed, for internal circulation only

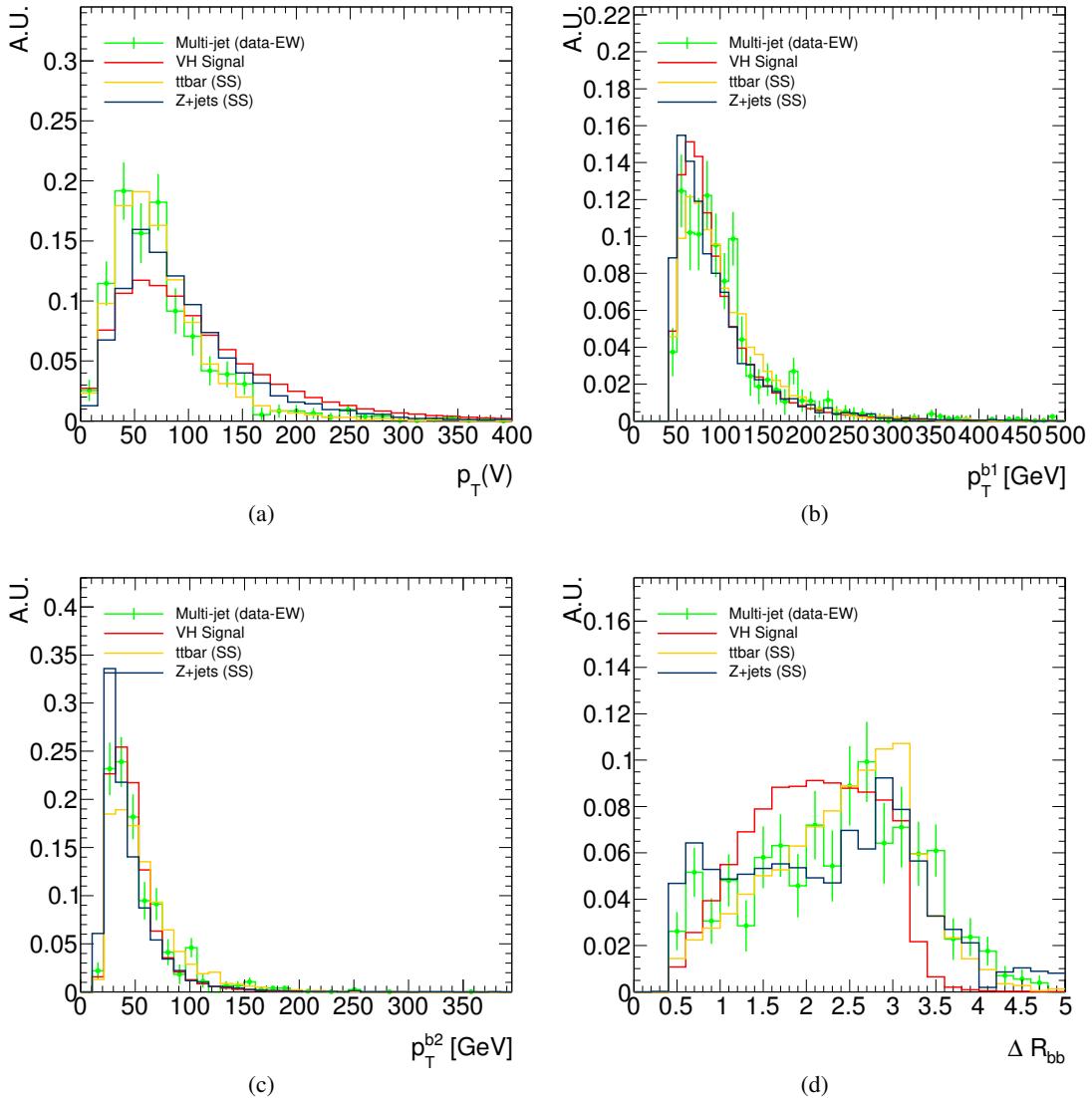


Figure 76: Shape comparison among signal, ttbar/Z+jets and multi-jet which is extracted from SS data. a) p_T^V , b) p_T of first leading b -jet, c), p_T of second leading b -jet and d) ΔR_{bb}

p_T^V efficiency	0-75 GeV	75-150 GeV	150-GeV
Signal	0.38	0.37	0.25
Multi-jet	0.67	0.26	0.06

Table 47: Estimated efficiency in each p_T^V category

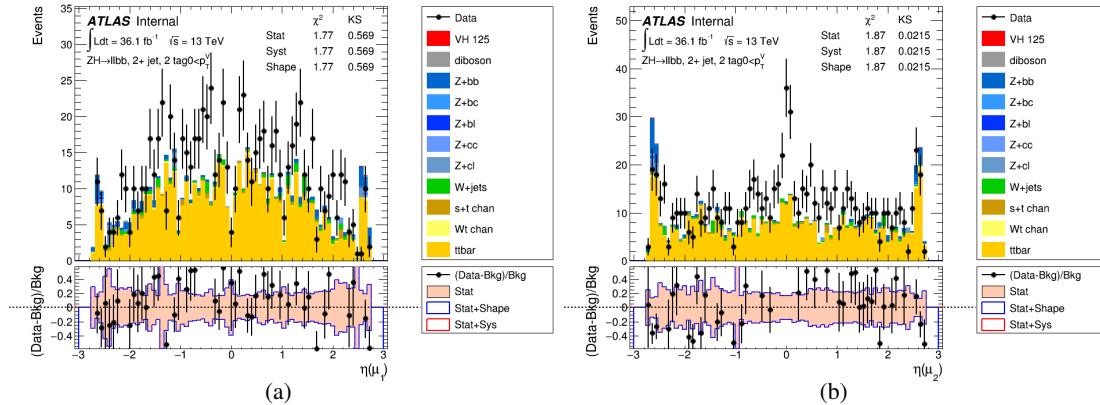
1195 B.2. Impact on buggy muon reconstruction and charge mis-assignment

1196 Strange dimuon peak in SS region results from charge mis-assignment of SiliconAssociatedForwardMuon.
 1197 In this reconstruction, silicon tracklet has no correct charge information. However, charge in CxAOD28
 1198 is taken from silicon tracklet due to reconstruction bug. Table 48 shows the comparison of signal yield
 1199 between opposite sign (default) and opposite+same sign to recover events misassigning muon charge. The
 1200 impact on signal yield is just 0.2% level. In this stage, we concluded no impact on the final results.

Signal Yield ($\mu\mu$ -ch)	Only opposite-sign	opposite+same-sign	Loss
2tag2jet (inclusive p_T^V)	26.57	26.63	0.2%
2tag3+jet (inclusive p_T^V)	38.61	38.69	0.2%

Table 48: Signal yield in opposite-sign and opposite plus same-sign events

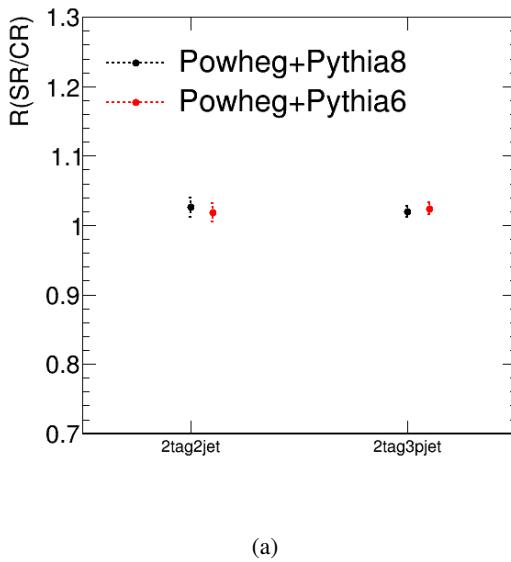
1201 Figure 77 show the eta distribution including buggy muon. Z mass events are seen in only high eta region.
 1202 The modeling of such muons are not so good. Given small amount of this muon contribution, it is no
 1203 impact on the analysis.

Figure 77: Muon η distribution of same-sign events in 2tag, 2+jets region. Multi-jet contribution is not added in this distributions.

1204 C. Data-driven top background estimation

1205 As described in Eq.3, top background shape and yield can be directly extracted from data with only
 1206 acceptance correction between same-flavor signal region and different-flavor $e\mu$ control region. This
 1207 correction factor which is defined the ratio of the yield of control region to the yield of signal region is
 1208 evaluated from MC. Dominant systematic uncertainty comes from the stability of this correction factor.

1209 Figure 78 shows the correction factor for 2jet and 3+jet category. The correction factor is stable around
 1210 1.03 and no MC dependence thanks to canceling out systematic sources.



(a)

Figure 78: The correction factor estimated from $t\bar{t}$ MC for each jet bin. Black points(red points) are correction factor for Powheg+Pythia8(Powheg+Pythia6) $t\bar{t}$ MC.

1211 Figure 79 shows the variation of $t\bar{t}$ yield for lepton-related systematic uncertainties in inclusive signal
 1212 region and control region, respectively. The yield variations in the signal region change at most a few %
 1213 that is relatively small. Furthermore it is found that the yields change coherently between signal region
 1214 and control region. Thanks to this cancellation effect, the correction factor which is ratio of signal region
 1215 to control region is quite stable for all systematic variations as shown in Figure 80. For this reason, it is
 1216 not necessary to consider lepton systematic uncertainty on this data-driven method.

1217 As discussed in Section 4.3.2, figures 4-5 show no big dependence on lepton p_T and η between signal
 1218 region and control region. m_{bb} and BDT distribution is in good agreement between SR and CR. It
 1219 is sufficient enough to use single correction factor for each region.

1220 Figure 81, 83 and 85 shows the $t\bar{t}$ yield variation with regard to nominal for the flavor tagging, MET and
 1221 JET related systematic uncertainties in signal region and control region. Some of jet and flavor tagging
 1222 systematic uncertainties are relatively large. However, as shown in the lepton-related systematic, all other
 1223 experimental systematic source can be canceled out by taking the ratio of signal region and control region.
 1224 Therefore the correction factor is stable for all systematic uncertainties. No systematic uncertainty is taken
 1225 into account in this method.

Not reviewed, for internal circulation only

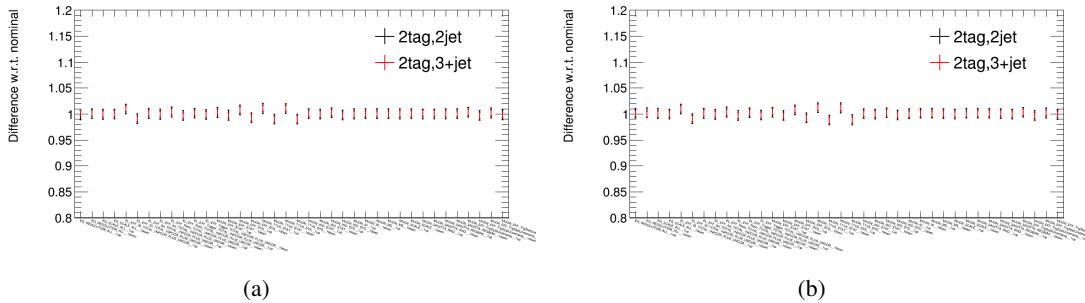


Figure 79: The difference of yield with regard to nominal for each lepton-related systematic variation in signal region (a) and control region (b).

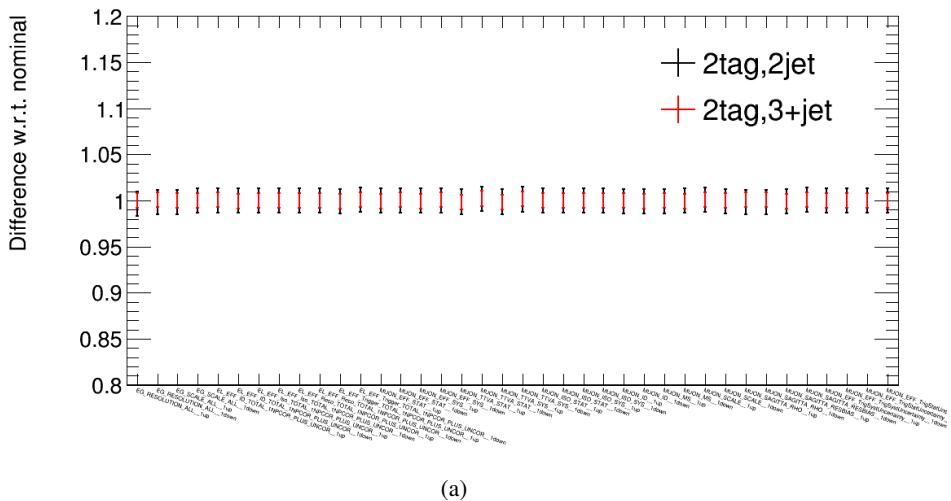


Figure 80: The difference of correction factor with regard to nominal for each lepton-related systematic variation. The deviation observed in each signal region and control region is cancel out.

Figures 87-88 shows data-driven top background estimation results. In these figures, flavor symmetric backgrounds (mainly $t\bar{t}$ and Wt) are directly estimated from $e\mu$ control region data with correction factor evaluated from MC. Other backgrounds has been scaled by the normalization factor from fit. The background estimation using control region are in good agreement with data. No experimental and theoretical systematic uncertainties are assigned since the shape and normalization are directly derived from data.

1232 This data-driven method is under implement in full fit analysis.

Not reviewed, for internal circulation only

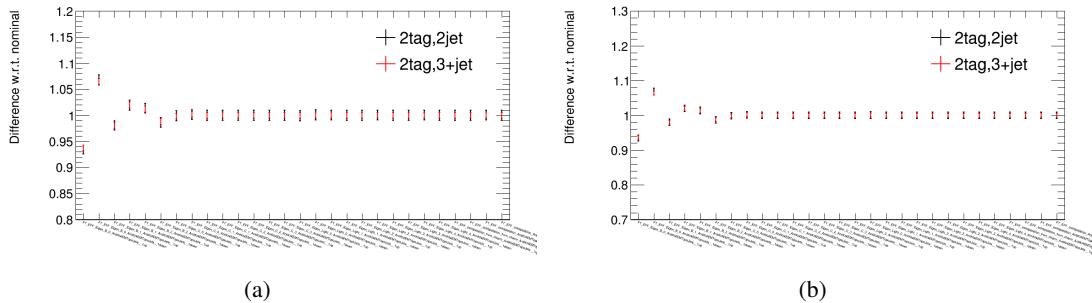


Figure 81: The difference of yield with regard to nominal for each flavor-tagging-related systematic variation in signal region (a) and control region (b).

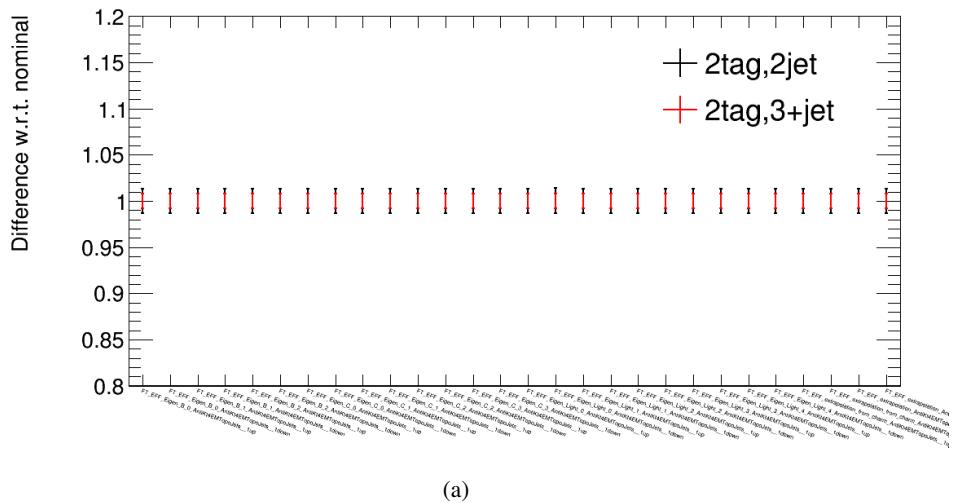


Figure 82: The difference of correction factor with regard to nominal for each flavor-tagging-related systematic variation. The deviation observed in each signal region and control region is cancel out.

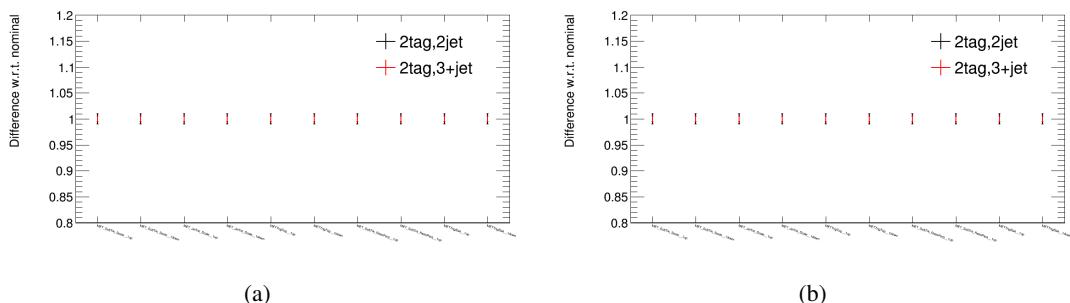


Figure 83: The difference of yield with regard to nominal for each met-related systematic variation in signal region (a) and control region (b).

Not reviewed, for internal circulation only

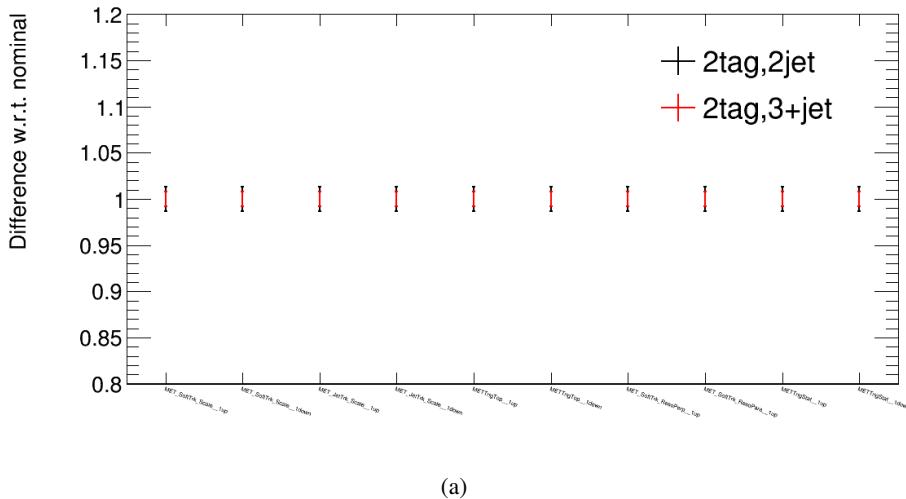


Figure 84: The difference of correction factor with regard to nominal for each met-related systematic variation. The deviation observed in each signal region and control region is cancel out.

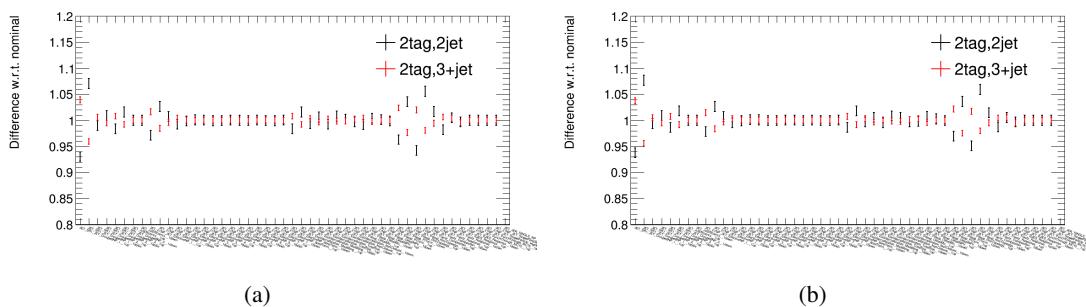


Figure 85: The difference of yield with regard to nominal for each jet-related systematic variation in signal region (a) and control region (b).

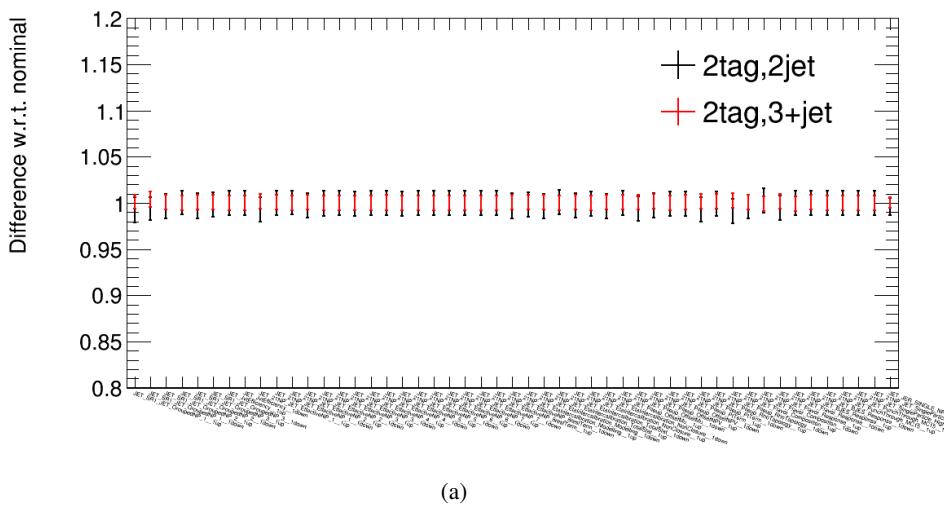


Figure 86: The difference of correction factor with regard to nominal for each jet-related systematic variation. The deviation observed in each signal region and control region is cancel out.

Not reviewed, for internal circulation only

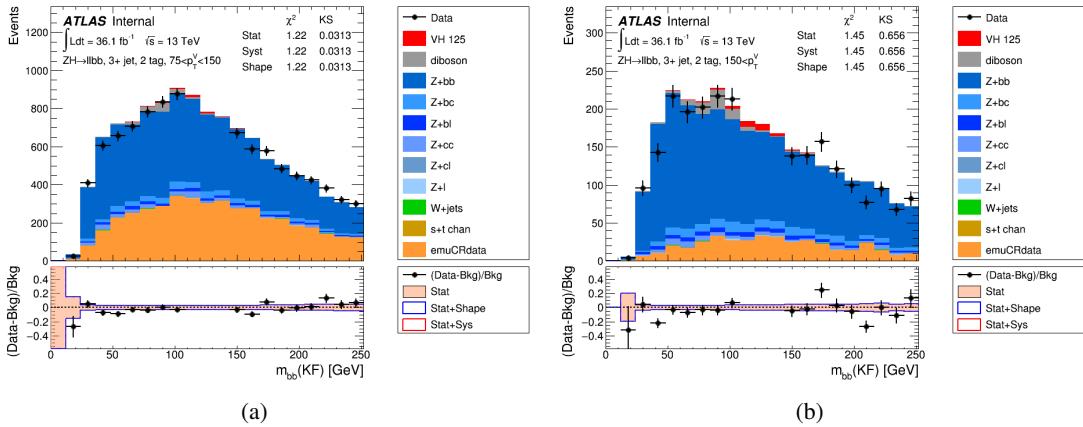


Figure 87: m_{bb} distributions for two p_T^V categories. $e\mu$ control region data applied correction factor is used to model flavor symmetric background. Background except for $t\bar{t}$ and Wt are scaled by normalization factor from fit.

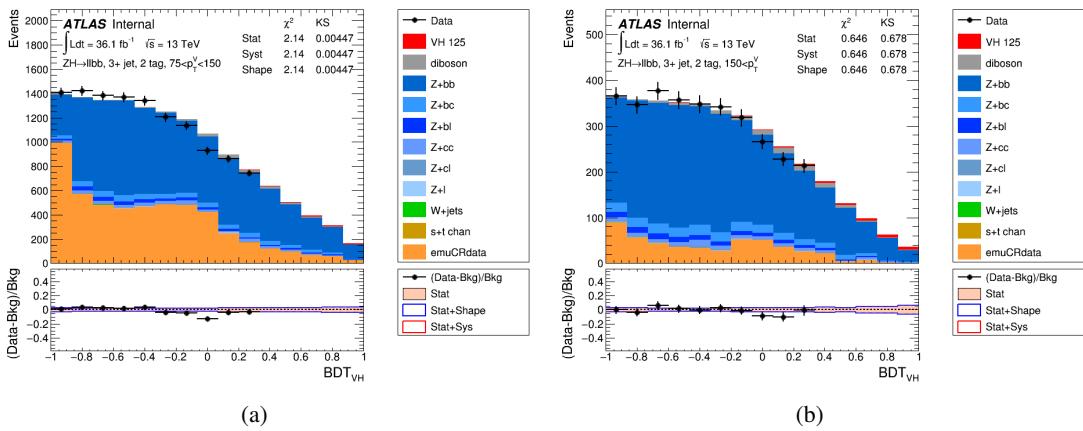


Figure 88: BDT output distributions for two p_T^V categories. $e\mu$ control region data applied correction factor is used to model flavor symmetric background. Background except for $t\bar{t}$ and Wt are scaled by normalization factor from fit.

1233 D. Kinematic Fitter Performance Study and Bias check

1234 The Higgs mass (m_{bb}) is one of the most important discriminants in $H \rightarrow b\bar{b}$ analysis. An improvement
 1235 of Higgs mass resolution gives better separation power and search sensitivity. Various b -jet corrections
 1236 are applied to improve the mass resolution. In the 2-lepton channel, one powerful technique to improve
 1237 mass resolution can be applied because $ZH \rightarrow b\bar{b}$ system can be fully reconstructed. It is possible to
 1238 utilize an event-level Kinematic Fit (KF) which constrains $\ell\ell b\bar{b}$ event topology with the uncertainty of its
 1239 reconstructed objects. A KF based on Run-1 has been developed for Run-2 lepton and jet objects. The
 1240 same likelihood model is used as in Run-1. The details of likelihood model can be found [25, 26]. The
 1241 likelihood is comprised by 12 fit parameters,

- 1242 • energies of 2 electron or p_T of 2 muons
- 1243 • energies of 2 b -jets
- 1244 • η, ϕ of 2 leptons and 2 jets
- 1245 • p_X and p_Y of $\ell\ell b\bar{b}$ system.
- 1246 • $m_{\ell\ell}$

1247 and 3 constraints for the variation of these parameters,

- 1248 • parameters : Gaussian (b -jet energy : Transfer Functions (TF))
- 1249 • p_X and p_Y of $\ell\ell b\bar{b}$ system : zero with a width of 9 GeV obtained from ZH signal MC.
- 1250 • m_{ll} : Breit-Wigner (BW) distribution of Z boson

1251 Figure 89 shows the reconstructed m_{bb} distribution using various b -jet corrections. When compared to
 1252 the m_{bb} mass resolution with muon-in-jet correction, that with the KF improves by 20-30%.

1253 Figures 90 and 91 show the difference of m_{bb} shape for dominant backgrounds, $t\bar{t}$ and $Z + bb$ in the
 1254 2-lepton analysis. It is found that the KF does not sculpt the background distribution and does not enhance
 1255 background around 125GeV.

1256 Figure 92 shows p_X^{llbb} and p_Y^{llbb} shape distributions in high- p_T^V category which is one of constraint para-
 1257 meters in kinematic fitter. Comparing $p_{X,Y}^{llbb}$ distribution before/after applying kinematic fitter, observed
 1258 data distributions are in reasonable agreement with expected MC distributions. No significant kinematic
 1259 fit bias is observed. Figure 93 show m_{ll} distribution that is also constraint parameter in kinematic fitter.
 1260 No mismodeling is observed after applying kinematic fitter.

Not reviewed, for internal circulation only

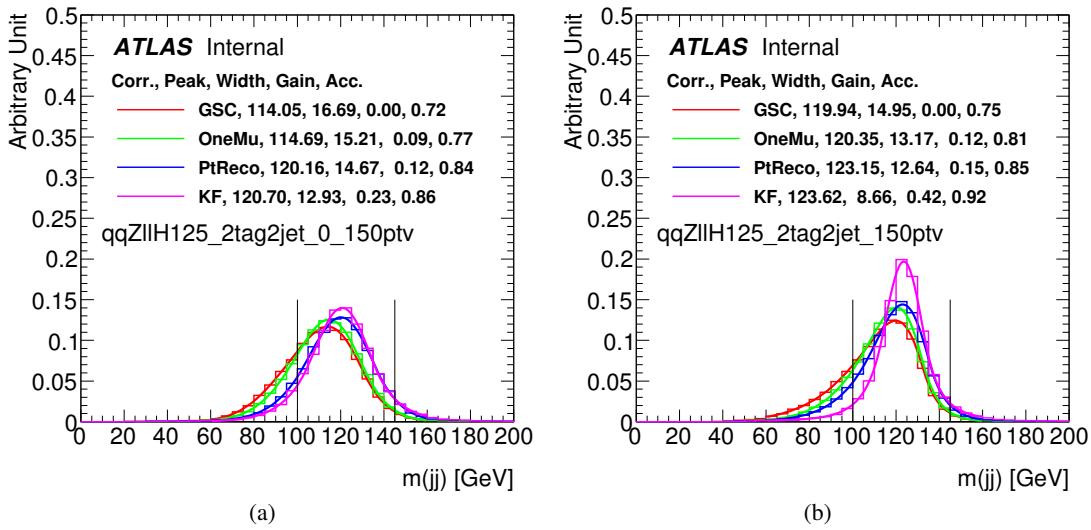


Figure 89: m_{bb} distribution of $qq \rightarrow Z h \rightarrow \ell\ell bb$ signal in 2 b -tagged jets events, (a) $0 < p_T^Z < 150$ GeV , (b) $p_T^Z > 150$ GeV. Each color shows each b -jet energy correction level. GSC : no specific b -jet correction (same as general jet), OneMu : with muon-in-jet correction, PtReco : PtReco correction after OneMu. KF : Kinematic fitter correction after OneMu.

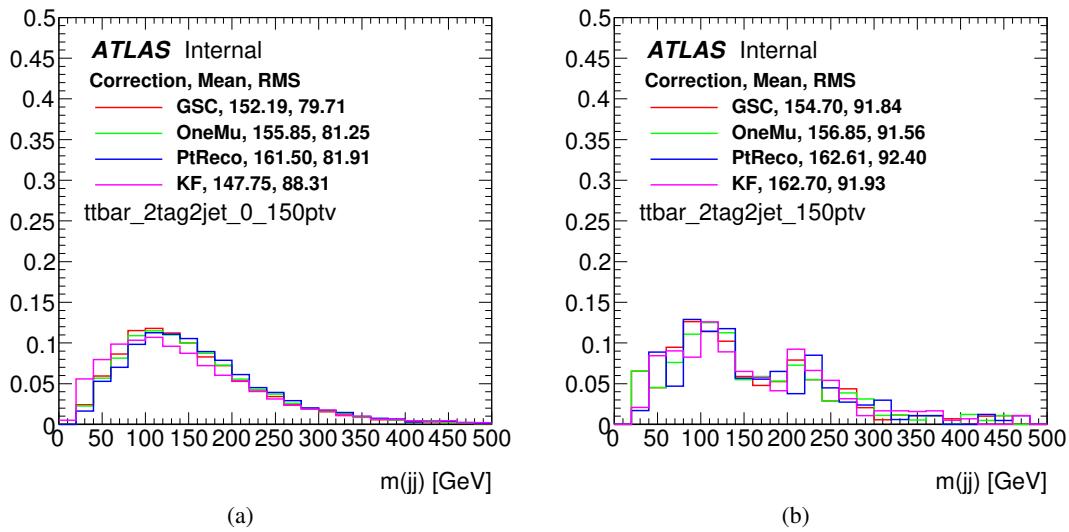


Figure 90: m_{bb} distribution of $t\bar{t}$ background in 2 b -tagged jets events, (a) $0 < p_T^Z < 150$ GeV , (b) $p_T^Z > 150$ GeV. Each color shows each b -jet energy correction level. GSC : no specific b -jet correction (same as general jet), OneMu : with muon-in-jet correction, PtReco : PtReco correction after OneMu. KF : Kinematic fitter correction after OneMu.

Not reviewed, for internal circulation only

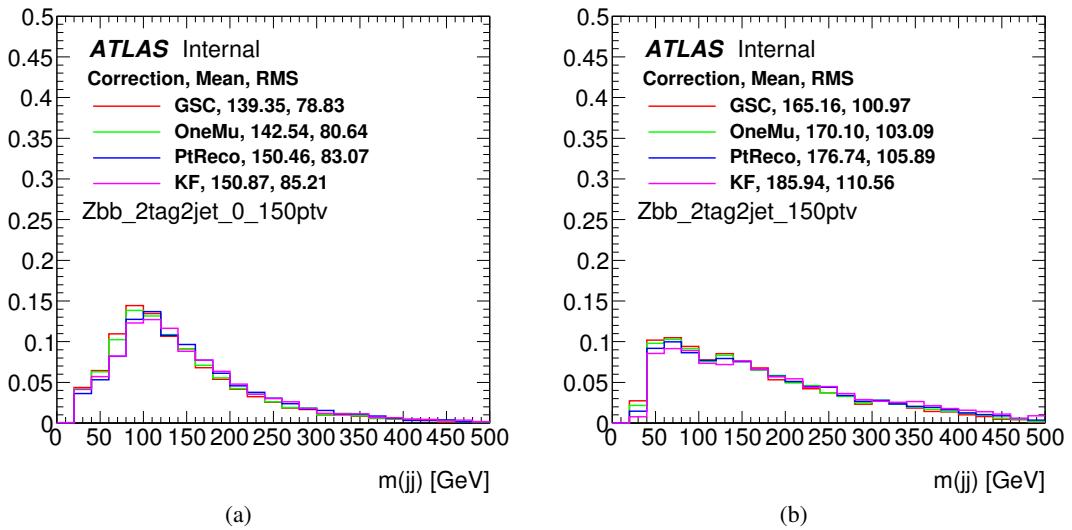


Figure 91: m_{bb} distribution of $t\bar{t}$ background in 2 b -tagged jets events, (a) $0 < p_T^Z < 150$ GeV , (b) $p_T^Z > 150$ GeV. Each color shows each b -jet energy correction level. GSC : no specific b -jet correction (same as general jet), OneMu : with muon-in-jet correction, PtReco : PtReco correction after OneMu. KF : Kinematic fitter correction after OneMu.

Not reviewed, for internal circulation only

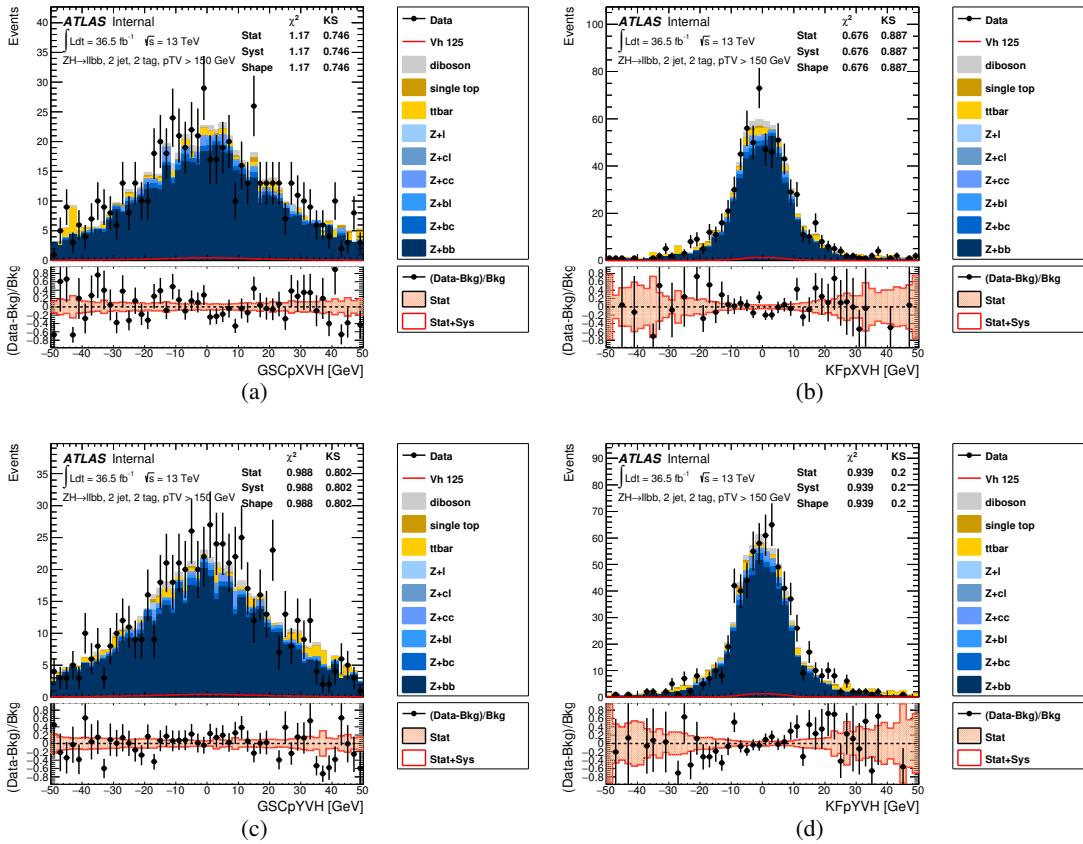


Figure 92: p_T^{VH} distribution in 2 tag 2 jet, $p_T^Z > 150$ GeV category. Left plots show before applying kinematic fitter (no constraint), Right plots show after constraint.

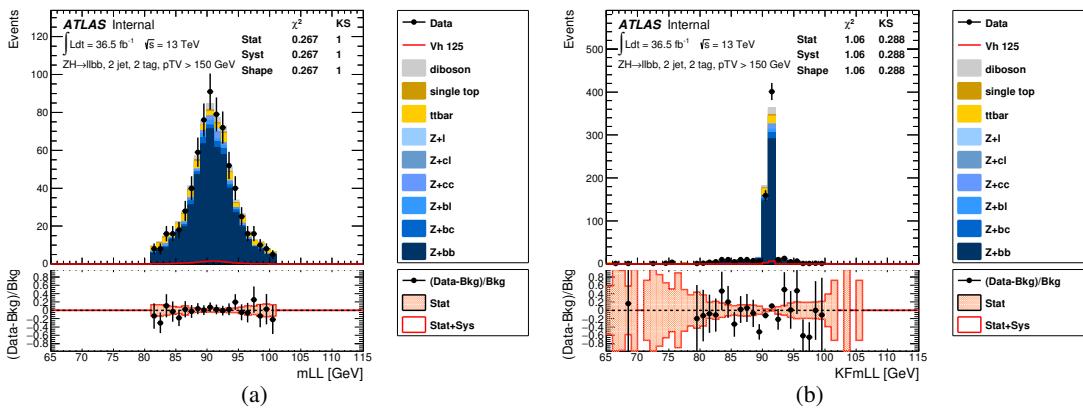


Figure 93: m_{ll} distribution in 2 tag 2 jet, $p_T^Z > 150$ GeV category. Left plots show before applying kinematic fitter (no constraint), Right plots show after constraint.

1261 E. Prefit data-MC comparisons

1262 Additional prefit, data-MC distributions are shown here, corresponding to an integrated luminosity of 36
 1263 fb^{-1} of 13 TeV data.

1264 E.0.1. 0-lepton channel

1265 In this section, prefit data-MC comparisons are shown for a number of distributions in the 0 lepton
 1266 channel.

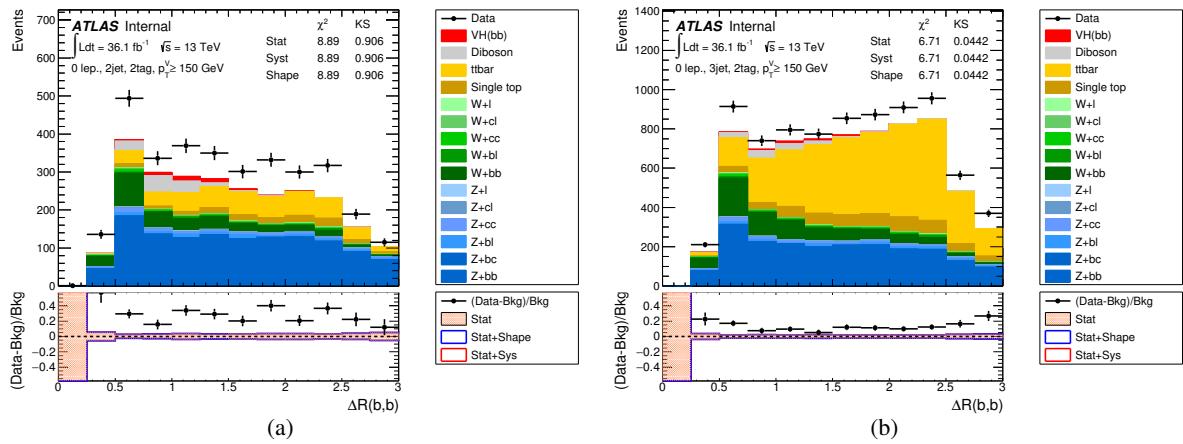


Figure 94: The pre-fit $\Delta R(B_1, B_2)$ distributions in the 0-lepton channel for 2-tag 2-jet (a) and 2-tag 3-jet (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

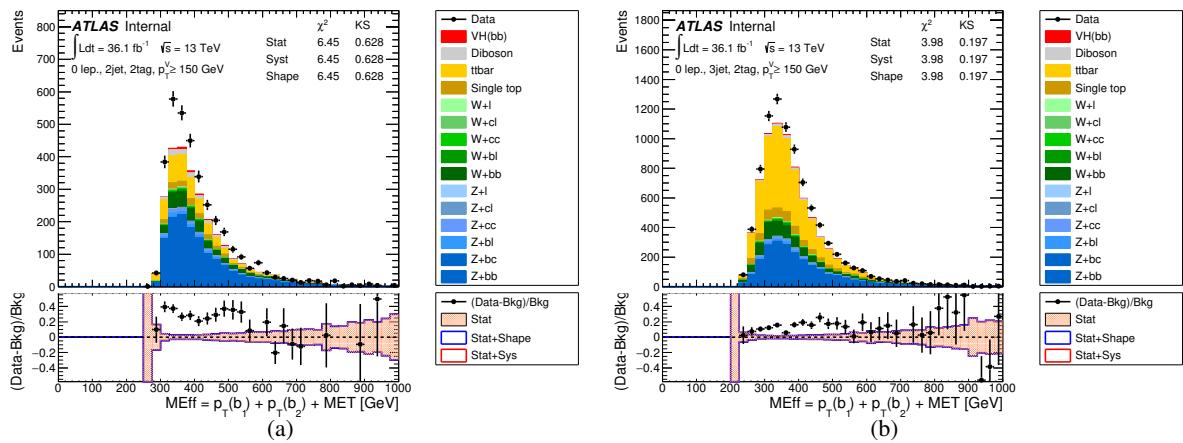


Figure 95: The pre-fit m_{eff} distributions in the 0-lepton channel for 2-jet events (a) and 3-jet events (b) in the 2-tag categories. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

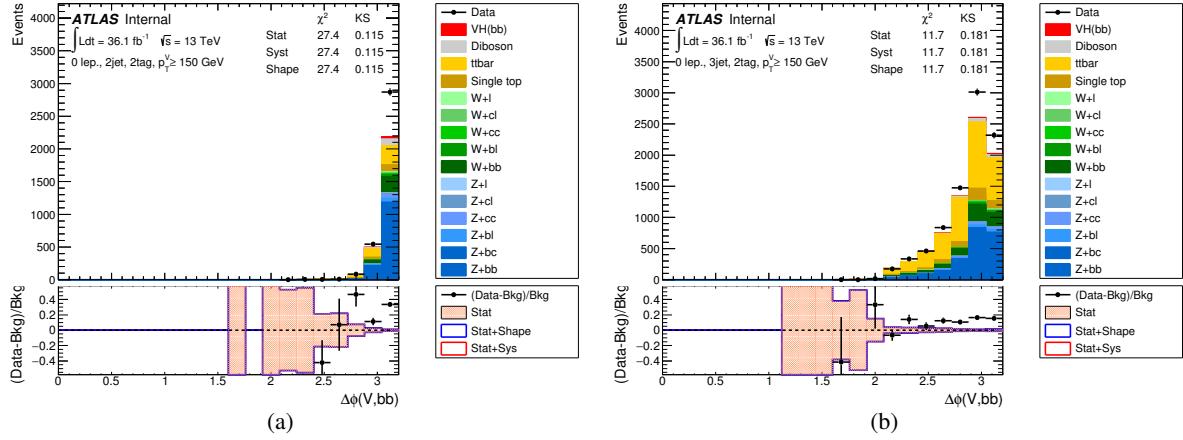


Figure 96: The pre-fit $\Delta\phi(E_T^{\text{miss}} \text{ dijet})$ distributions in the 0-lepton channel for 2-jet events (a) and 3-jet events (b) in the 2-tag categories. The background and signal samples are normalized to the expected cross-section predictions.

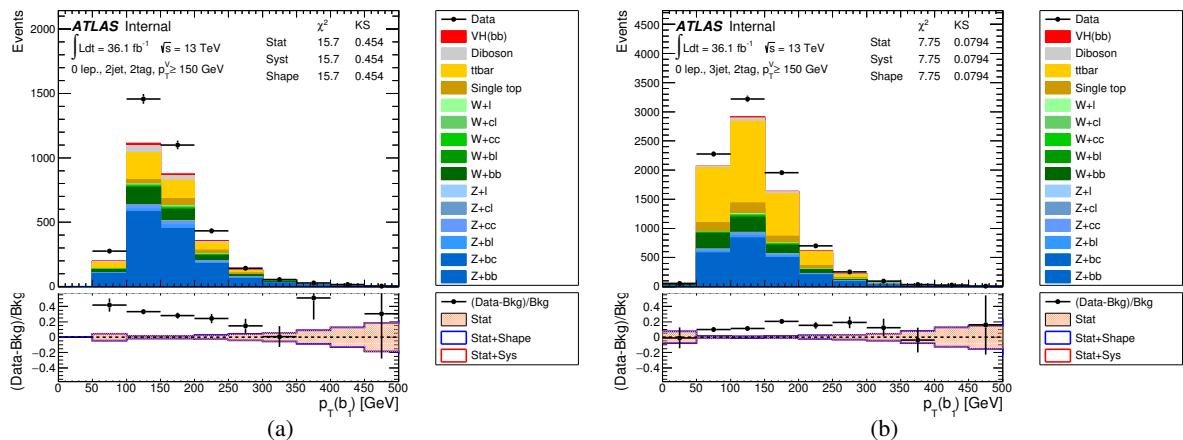


Figure 97: The pre-fit distributions in the 0-lepton channel for 2-tag 2-jet (a) and 2-tag 3-jet (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

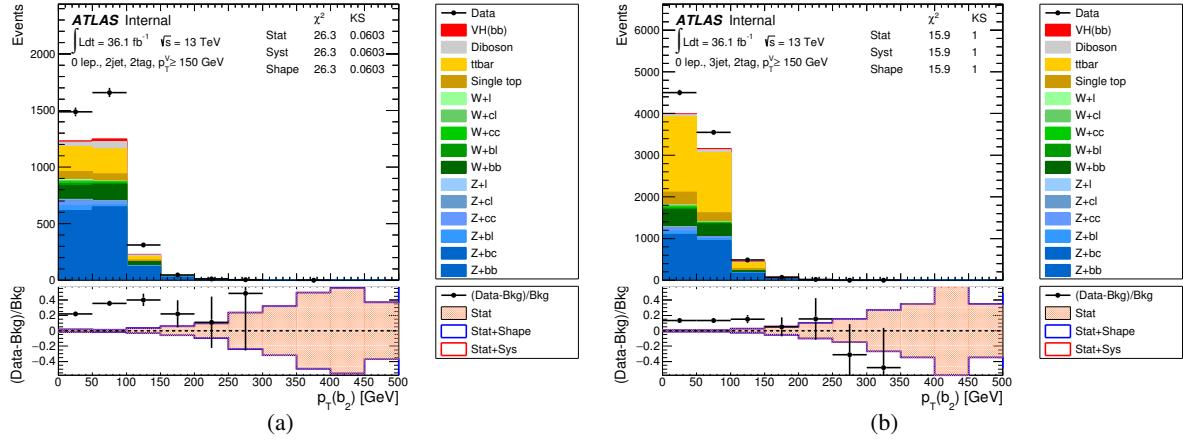


Figure 98: The pre-fit $p_T^{\text{B}2}$ distributions in the 0-lepton channel for 2-jet events (a) and 3-jet events (b) in the 2-tag categories. The background and signal samples are normalized to the expected cross-section predictions.

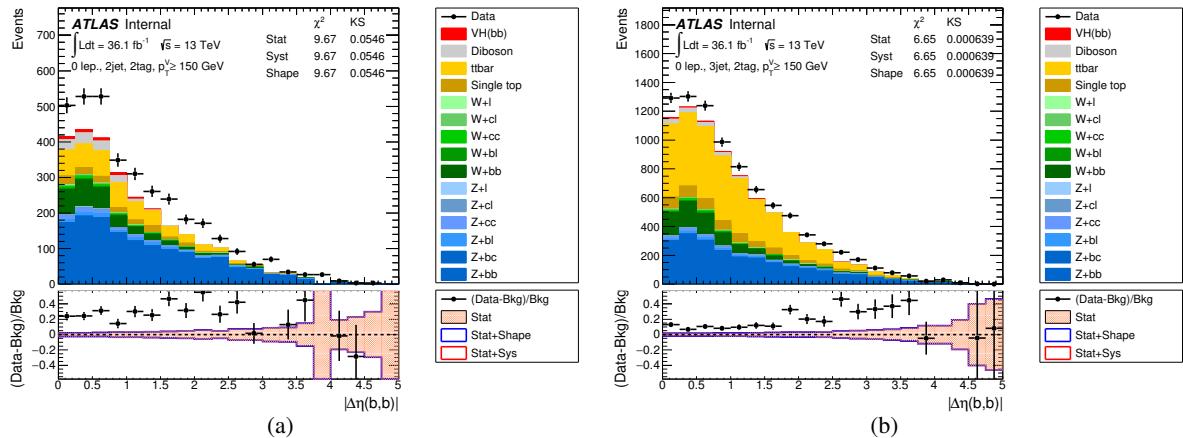


Figure 99: The pre-fit $\Delta\eta(\text{B}1,\text{B}2)$ distributions in the 0-lepton channel for 2-jet events (a) and 3-jet events (b) in the 2-tag categories. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

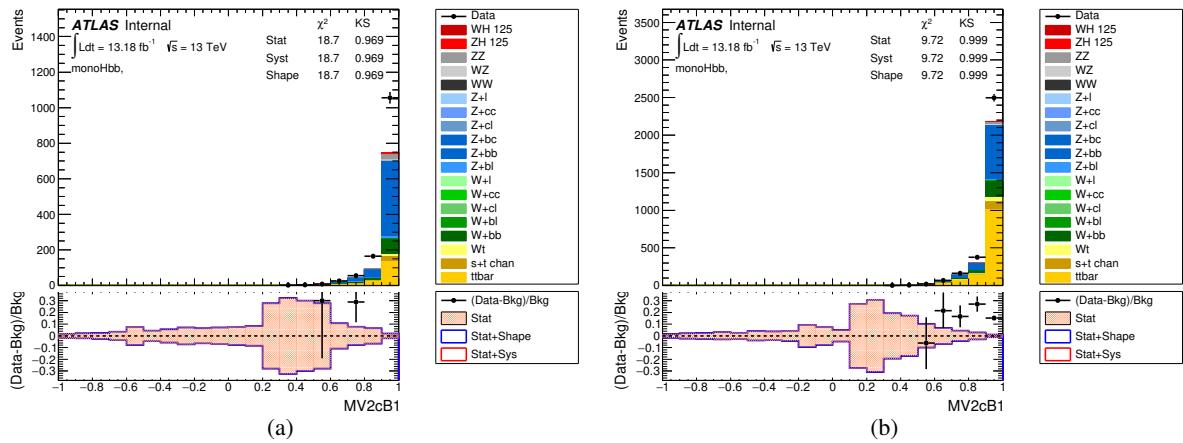


Figure 100: The pre-fit MV2C10 distributions of the leading jet in the 0-lepton channel for 2-jet events (a) and 3-jet events (b) in the 2-tag categories. The background and signal samples are normalized to the expected cross-section predictions.

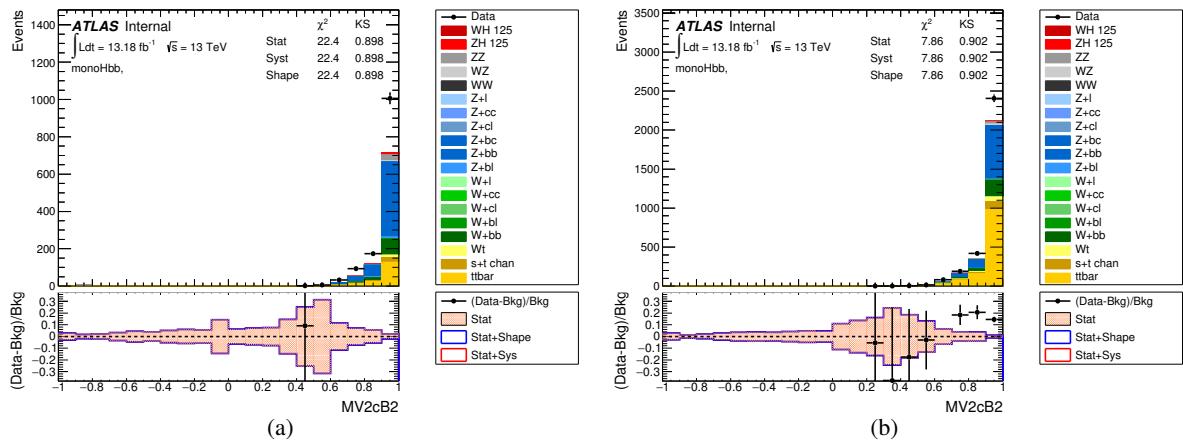


Figure 101: The pre-fit MV2C10 distributions of the sub-leading jet in the 0-lepton channel for 2-jet events (a) and 3-jet events (b) in the 2-tag categories. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

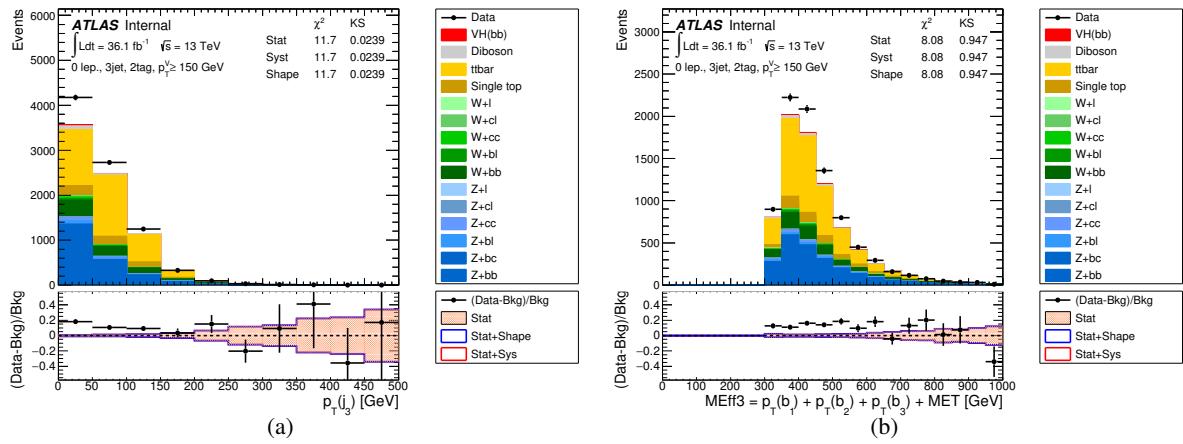


Figure 102: The pre-fit p_T^{J3} (a) and m_{eff}^{J3} (b) distribution in the 0-lepton channel for 3-jet 2-tag events are shown. The background and signal samples are normalized to the expected cross-section predictions.

1267 **E.0.2. 1-lepton channel**

1268 In this section, prefit data-MC comparisons are shown for a number of distributions in the 1-lepton
 1269 channel.

Not reviewed, for internal circulation only

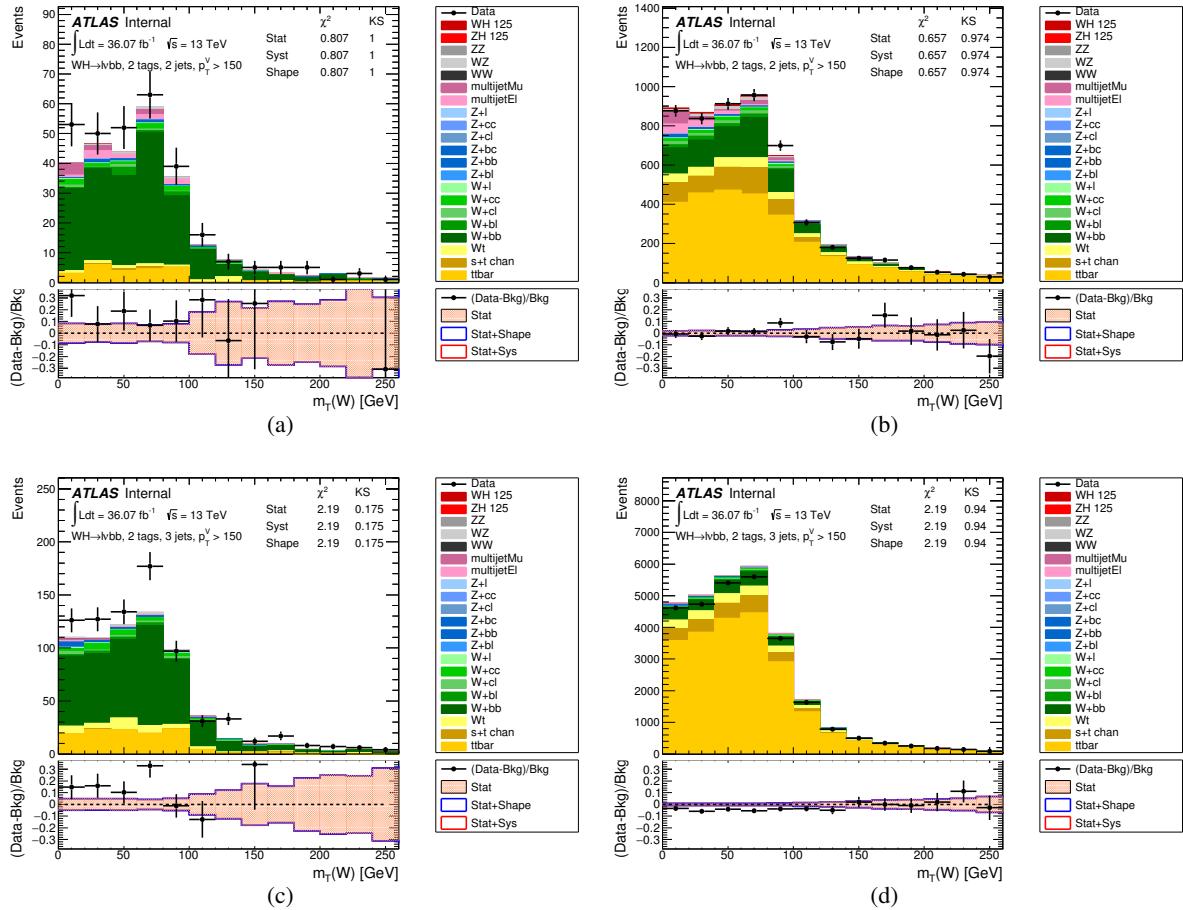


Figure 103: The pre-fit $p_T^W > 150$ GeV m_T^W distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

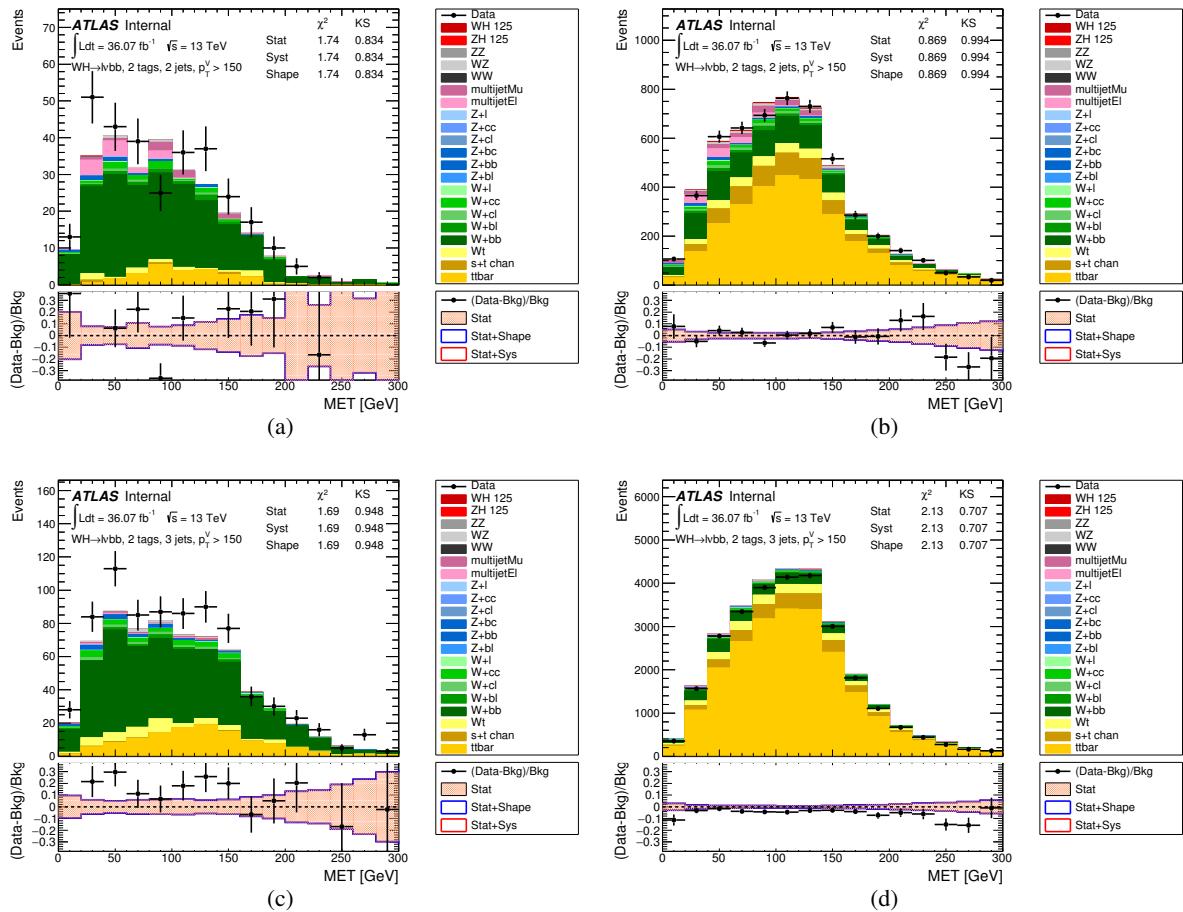


Figure 104: The pre-fit $p_T^W > 150 \text{ GeV } E_T^{\text{miss}}$ distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

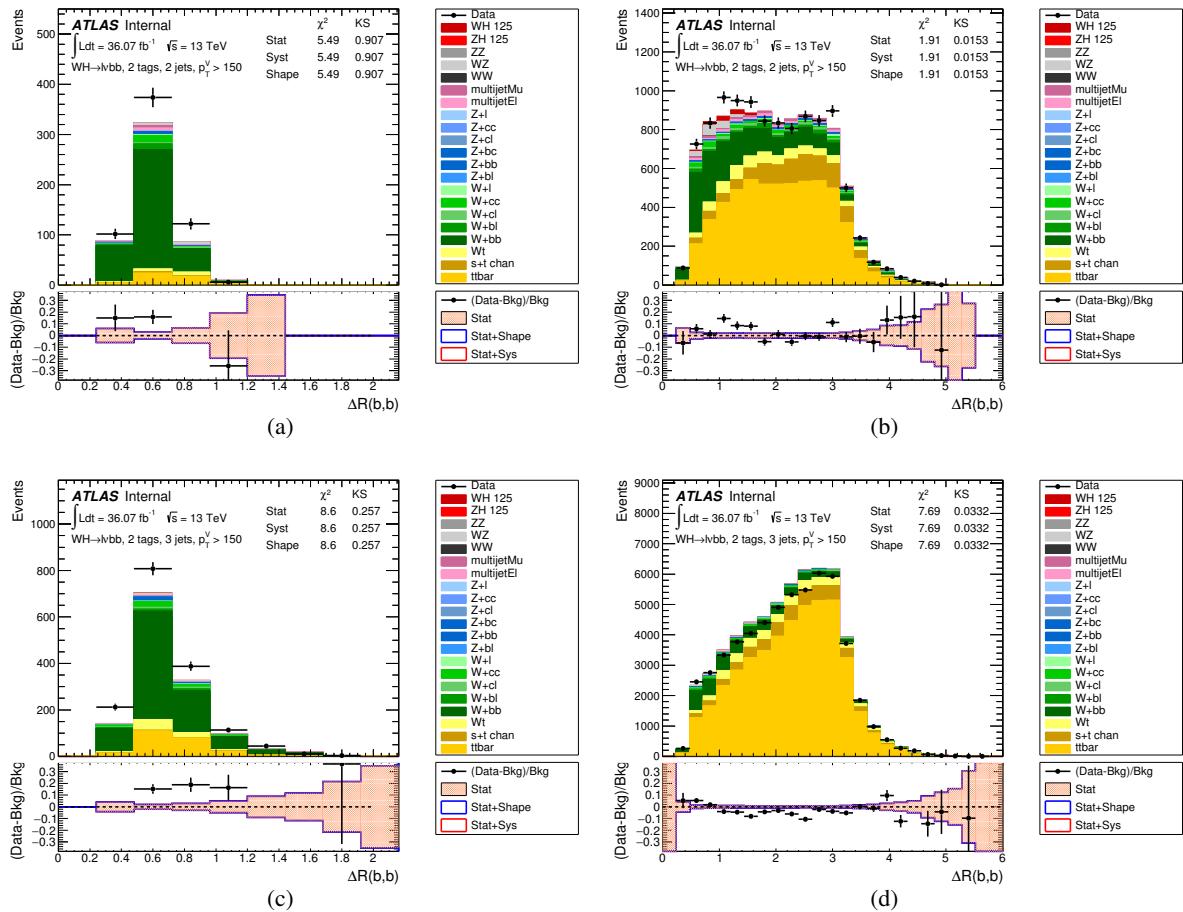


Figure 105: The pre-fit $p_T^W > 150$ GeV $\Delta R(b,b)$ distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

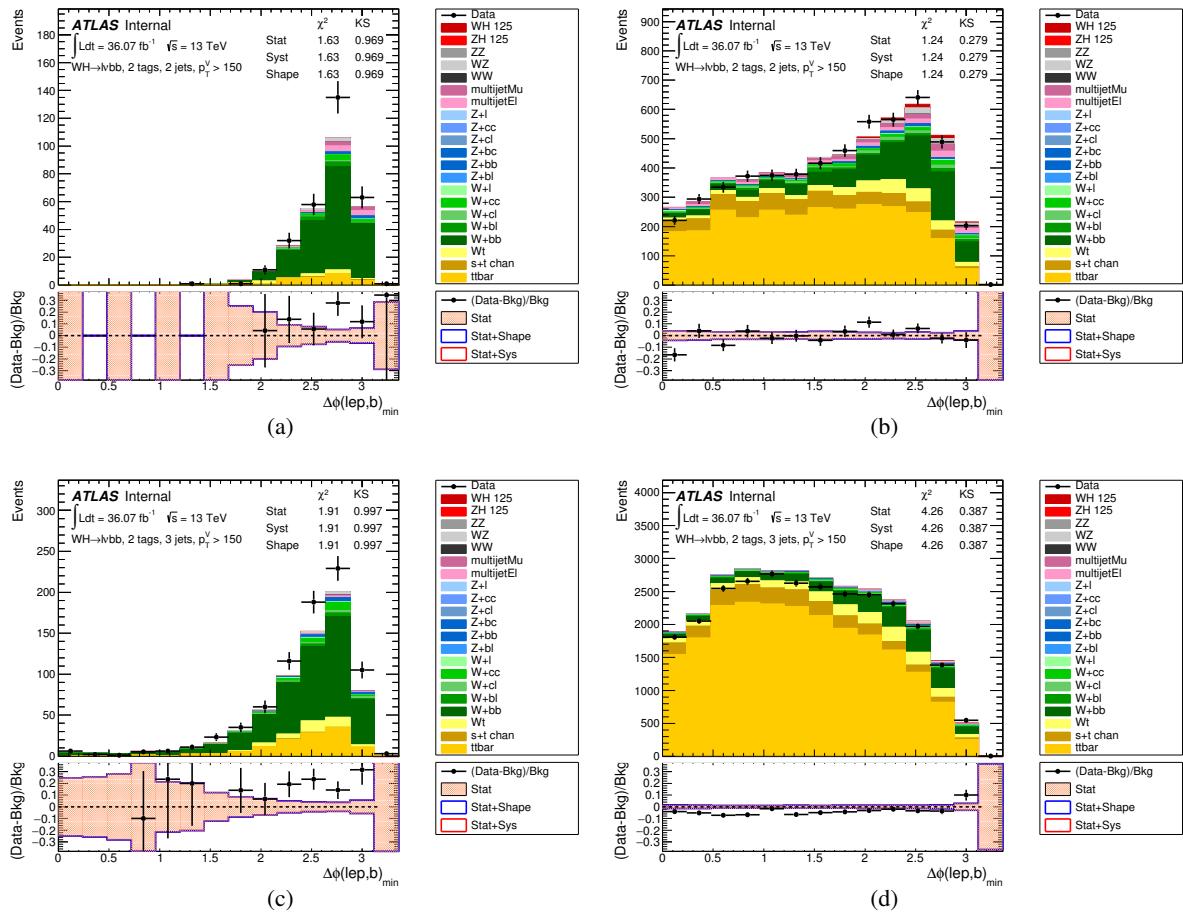


Figure 106: The pre-fit $p_T^W > 150$ GeV $\Delta\phi(l, b)$ distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

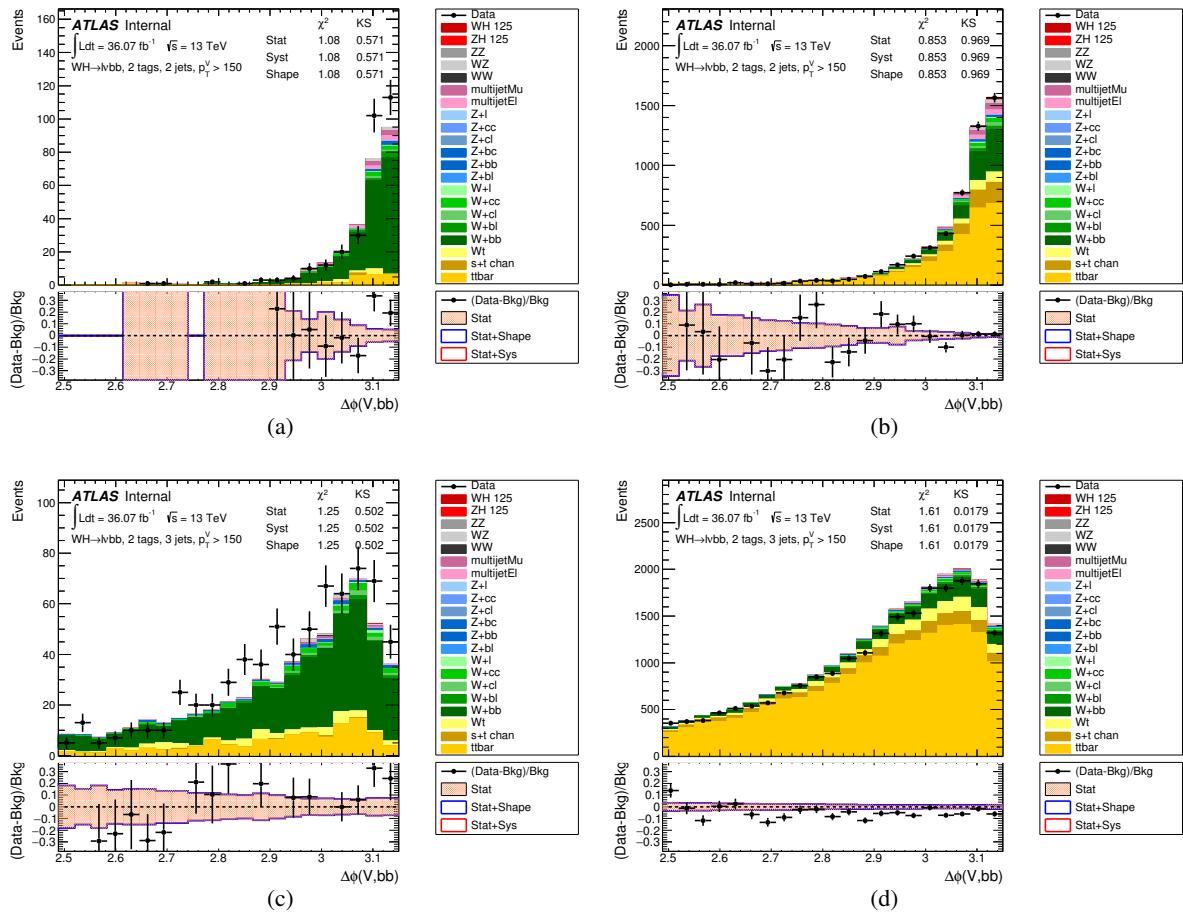


Figure 107: The pre-fit $p_T^W > 150$ GeV $\Delta\phi(V, H)$ distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

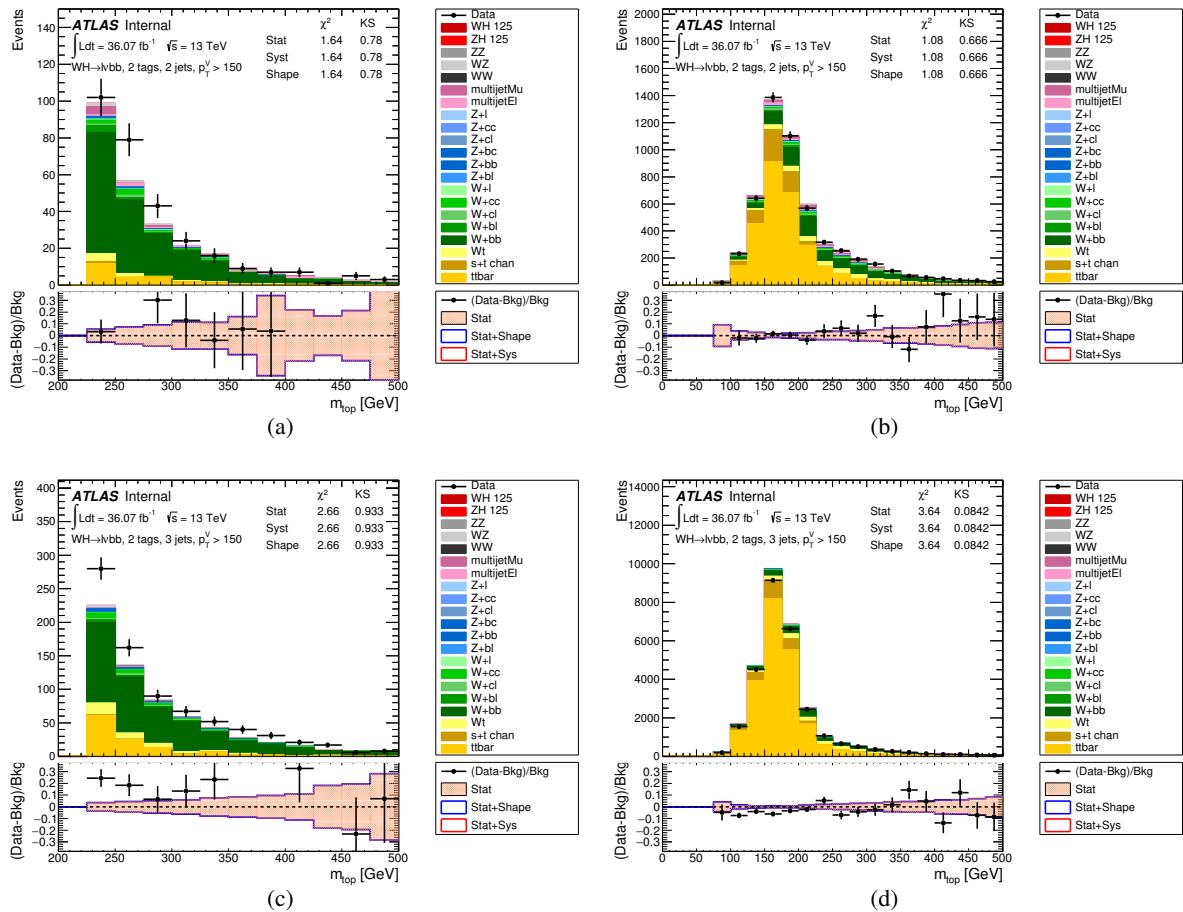


Figure 108: The pre-fit $p_T^W > 150$ GeV m_{top} distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

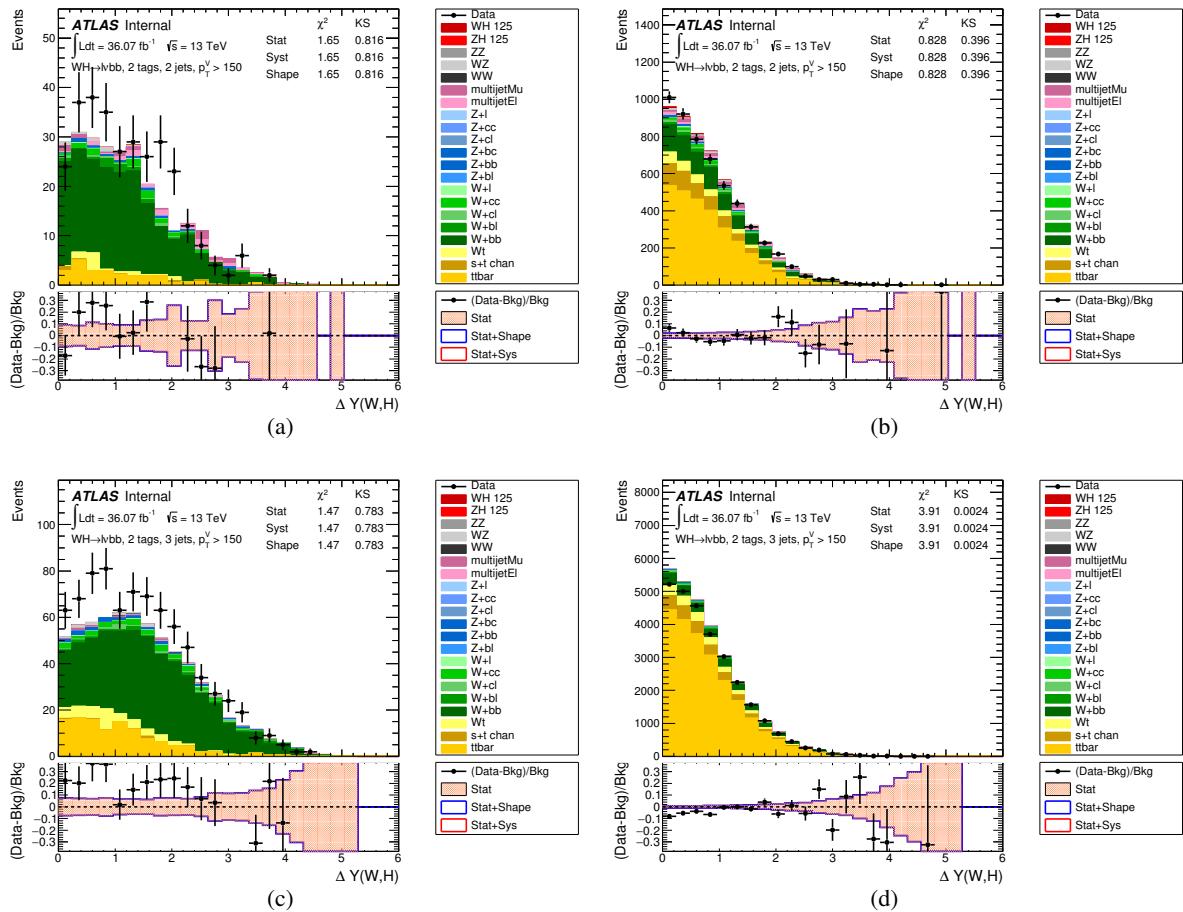


Figure 109: The pre-fit $p_T^W > 150$ GeV $\Delta Y(W, H)$ distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

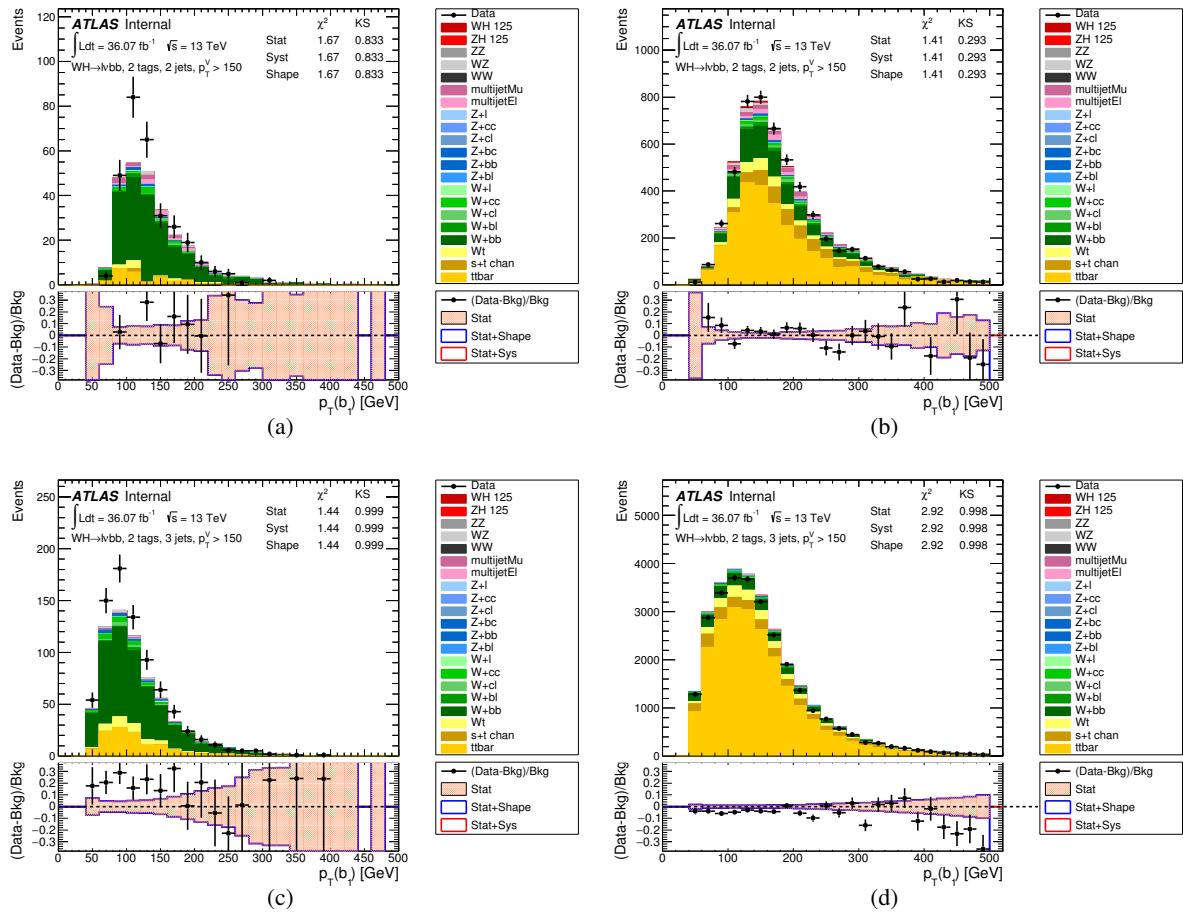


Figure 110: The pre-fit $p_T^W > 150$ GeV p_T^{B1} distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

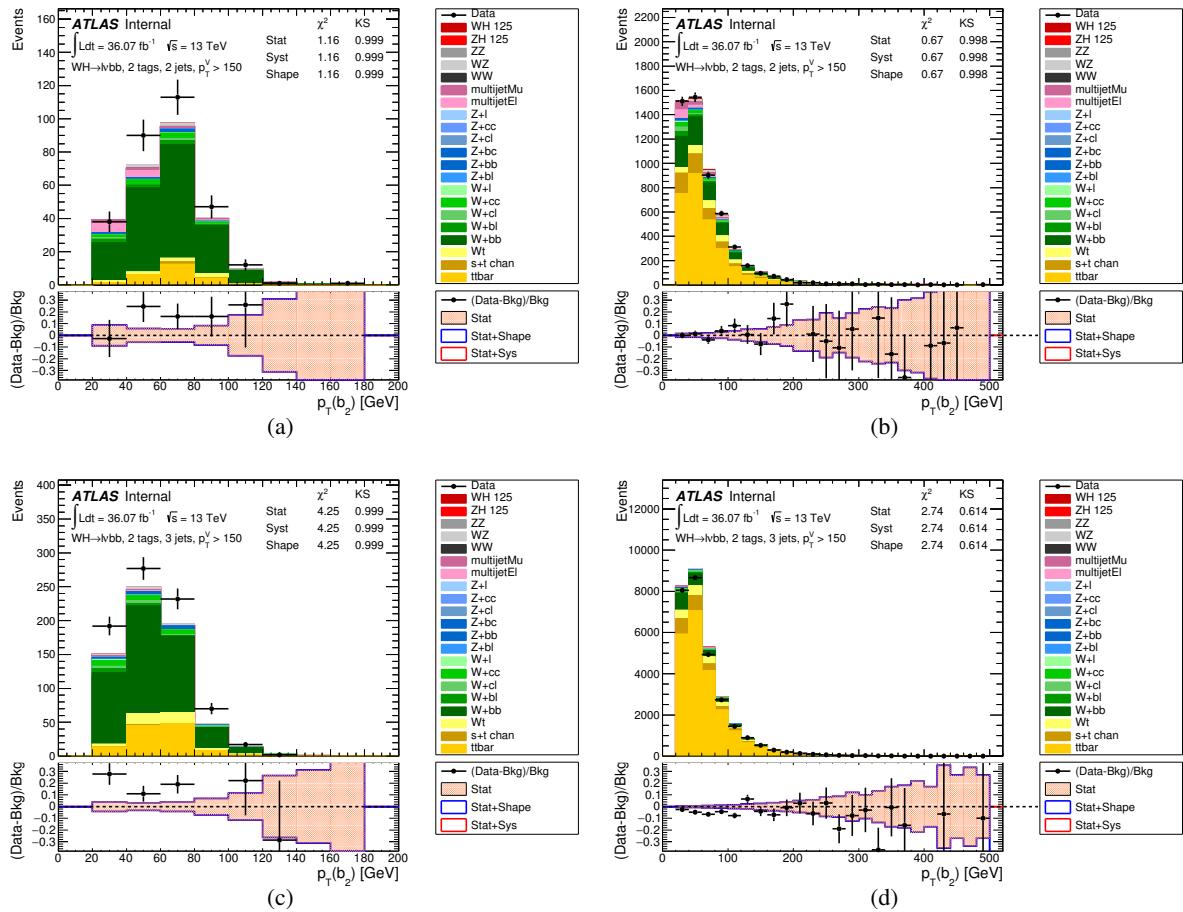


Figure 111: The pre-fit $p_T^W > 150 \text{ GeV}$ p_T^{B2} distributions in the 1-lepton channel for the 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

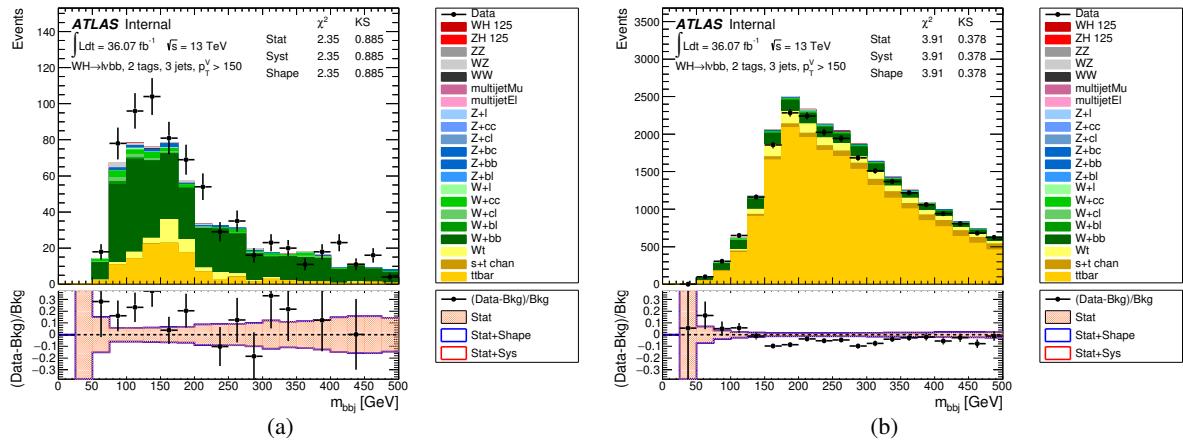


Figure 112: The pre-fit $p_T^W > 150$ GeV m_{bbj} distributions in the 1-lepton channel for the 2-tag 3-jet $W + hf$ CR (a), SR (b) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

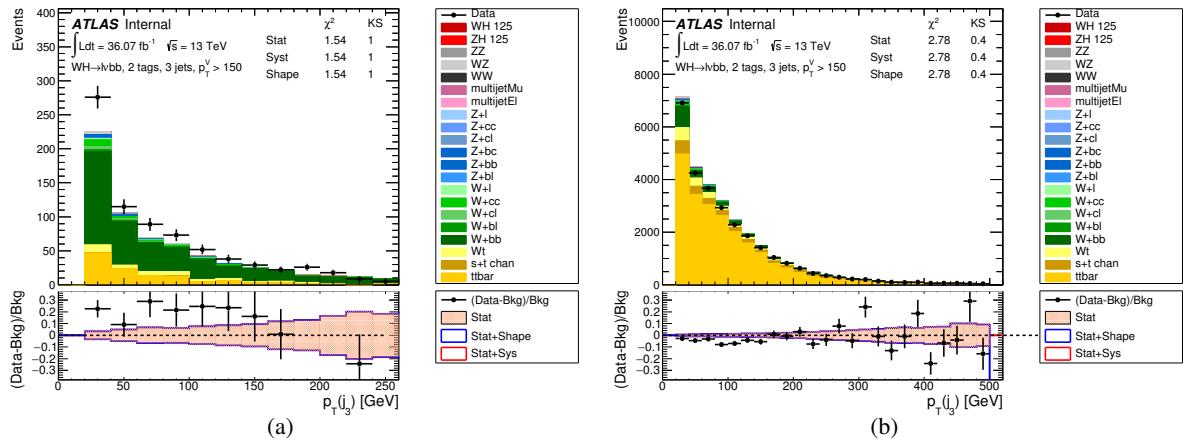


Figure 113: The pre-fit $p_T^W > 150$ GeV $p_T^{j_3}$ distributions in the 1-lepton channel for the 2-tag 3-jet $W + hf$ CR (a), SR (b) categories, are shown. The background and signal samples are normalized to the expected cross-section predictions.

1270 **E.0.3. 2-lepton channel**

1271 Figure 114 shows the jet multiplicity distribution after pre-selection for 2-tag events in the two p_T^Z
 1272 categories.

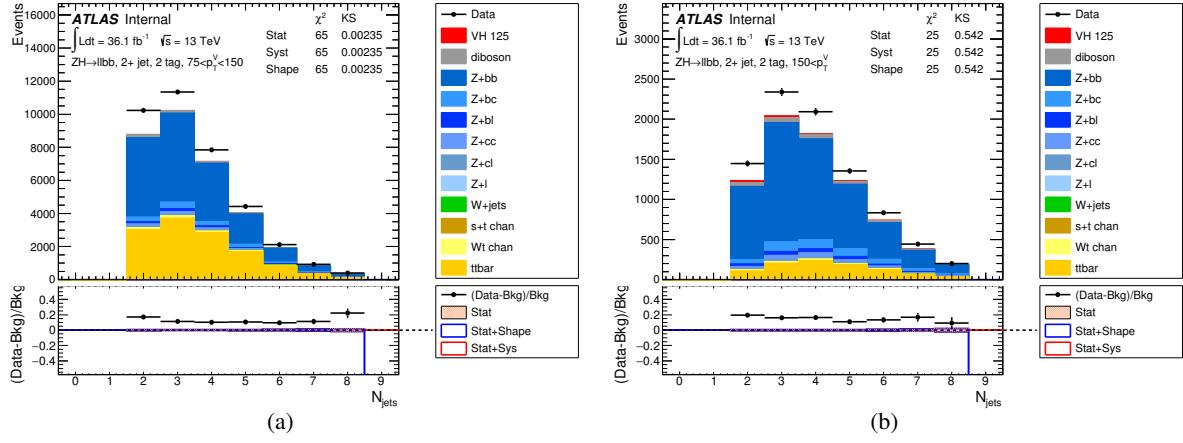


Figure 114: The pre-fit jet multiplicity distributions in the 2-leptons channel for 2-tag events, (a) $75 < p_T^V < 150$ GeV and (b) $p_T^V > 150$ categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

1273 Figures 115-119 and 120-124 show other input variables used in the MVA analysis.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

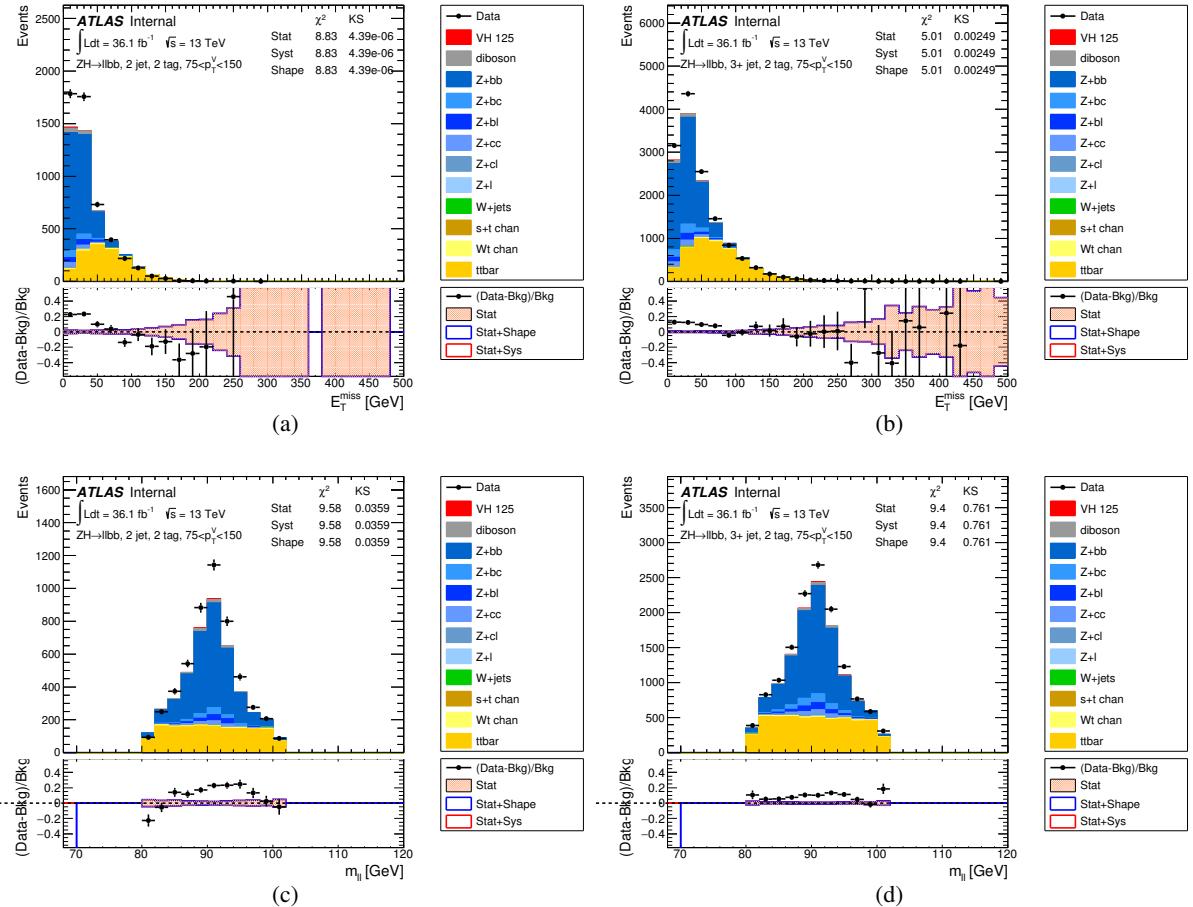


Figure 115: The pre-fit missing transverse energy distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $75 < p_T^V < 150$ GeV and (b) 3+jets, $75 < p_T^V < 150$ GeV categories are shown. The pre-fit m_{ll} distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $75 < p_T^V < 150$ GeV and (d) 3+jets, $75 < p_T^V < 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

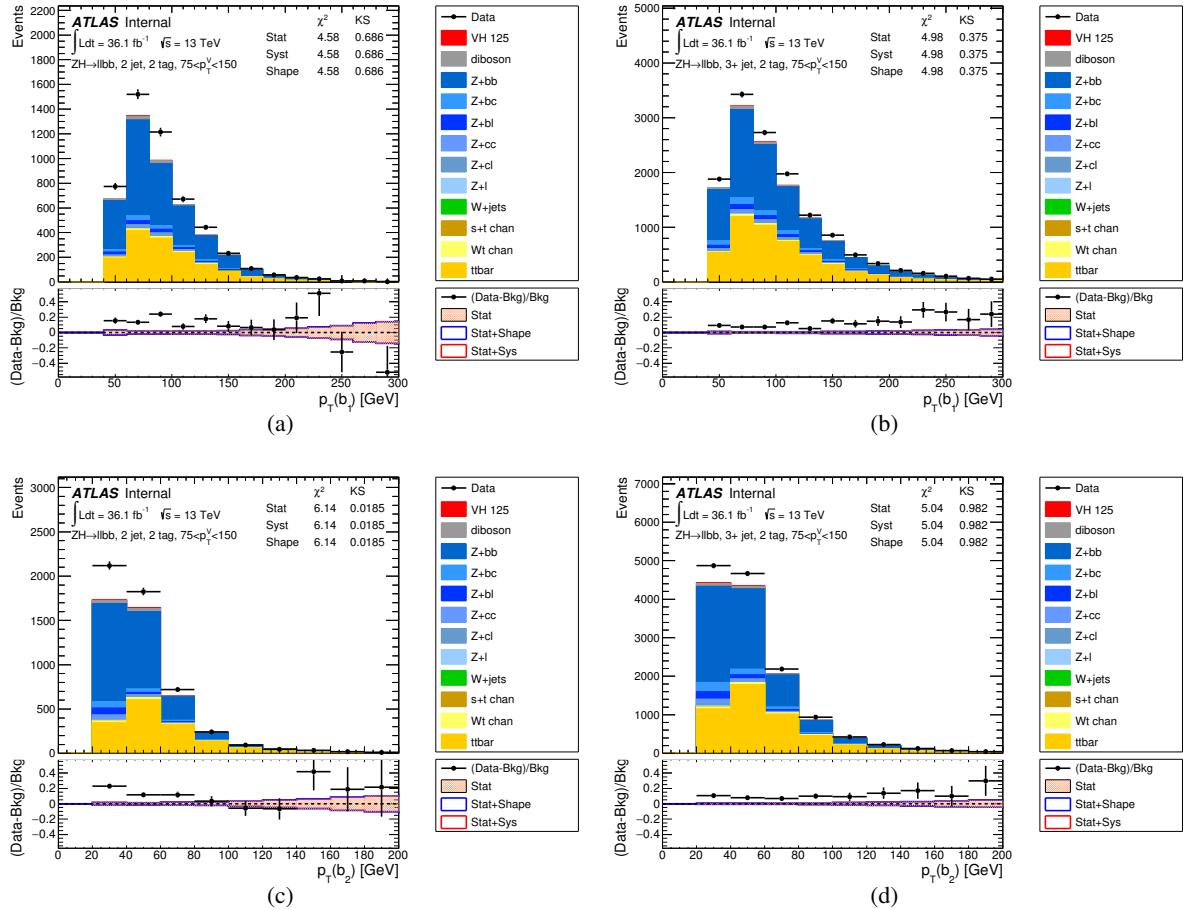


Figure 116: The pre-fit p_T of leading b-tagged jet distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $75 < p_T^V < 150$ GeV and (b) 3+jets, $75 < p_T^V < 150$ GeV categories are shown. The pre-fit p_T of second leading b-tagged jet distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $75 < p_T^V < 150$ GeV and (d) 3+jets, $75 < p_T^V < 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

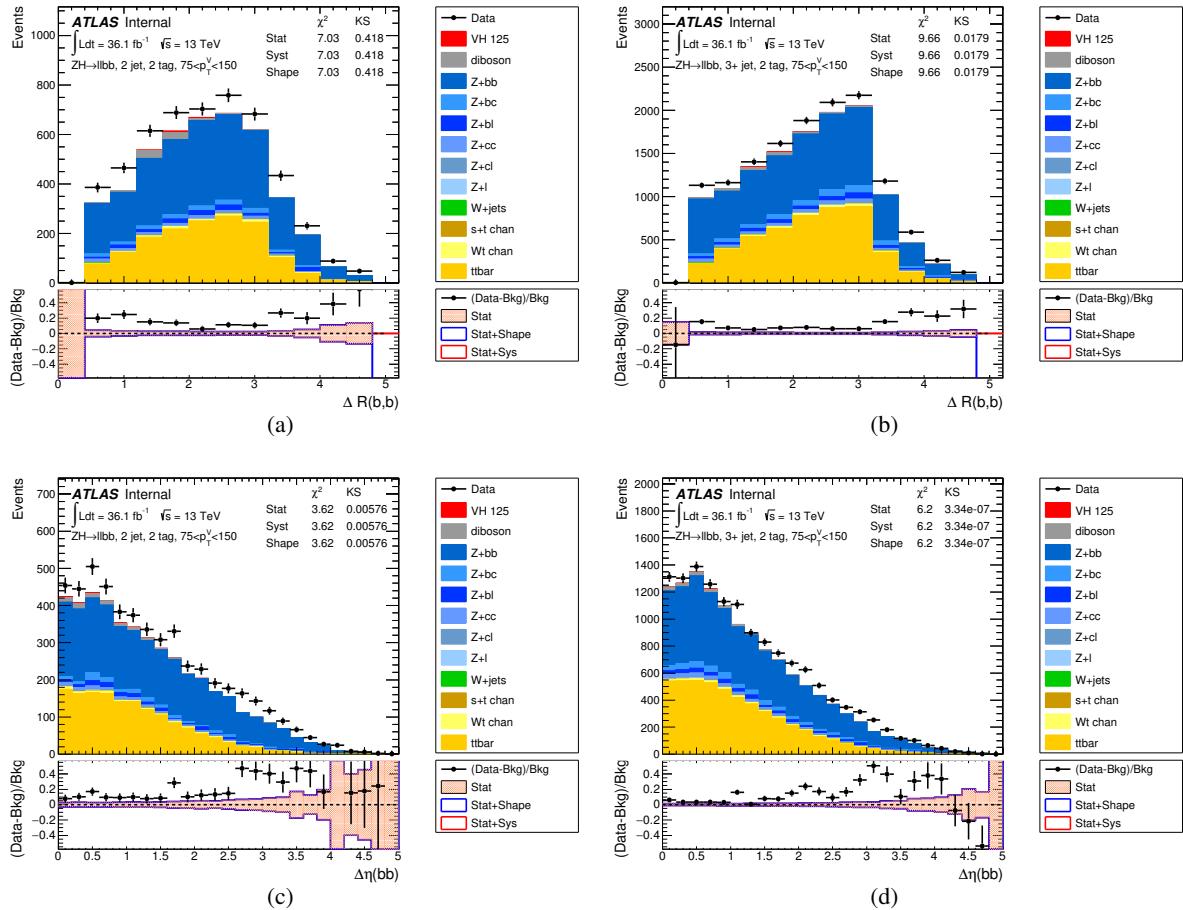


Figure 117: The pre-fit ΔR_{bb} distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $75 < p_T^V < 150$ GeV and (b) 3+jets, $75 < p_T^V < 150$ GeV categories are shown. The pre-fit $\Delta\eta_{bb}$ distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $75 < p_T^V < 150$ GeV and (d) 3+jets, $p_T^V > 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

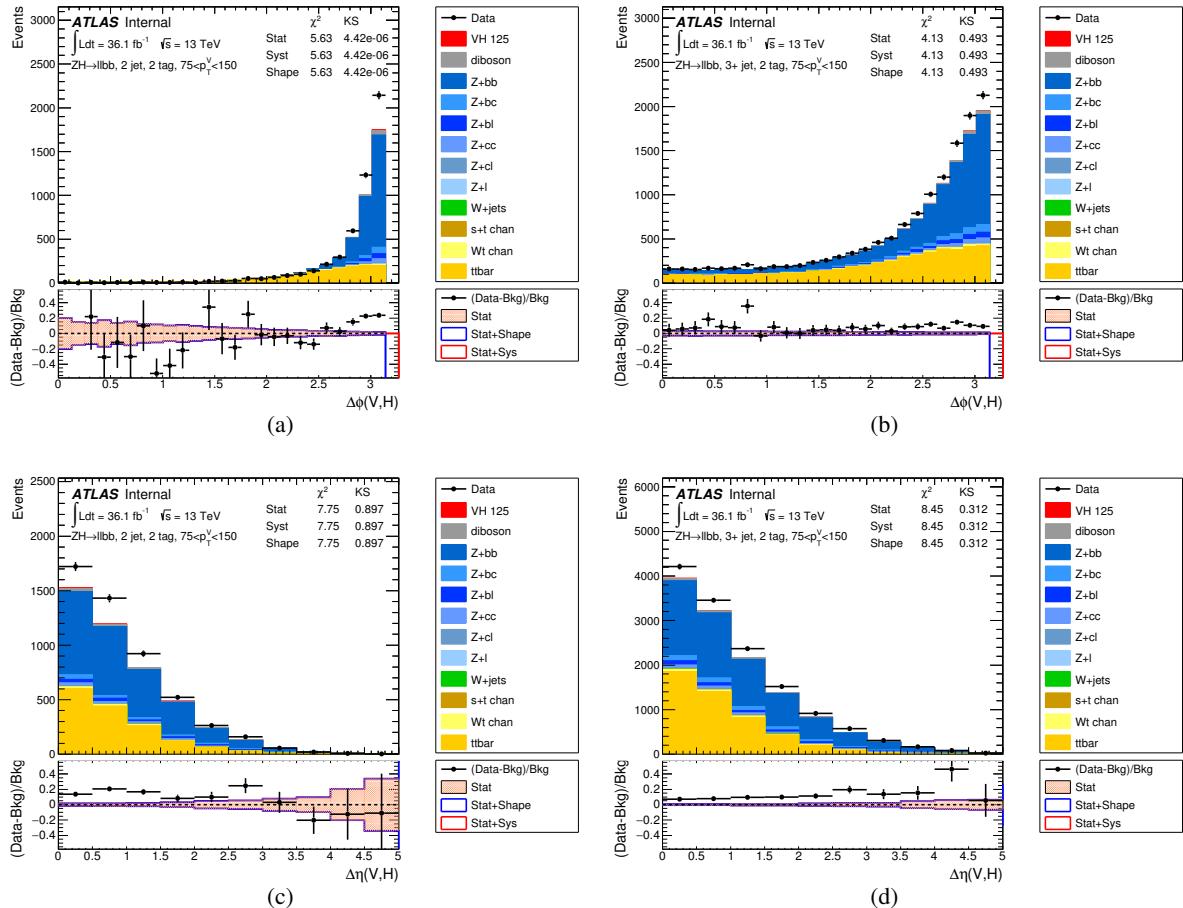


Figure 118: The pre-fit $\Delta\phi_{Vbb}$ distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $p_T^V > 150$ GeV and (b) 3+jets, $p_T^V > 150$ GeV categories are shown. The pre-fit $\Delta\eta_{Vbb}$ distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $p_T^V > 150$ GeV and (d) 3+jets, $p_T^V > 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

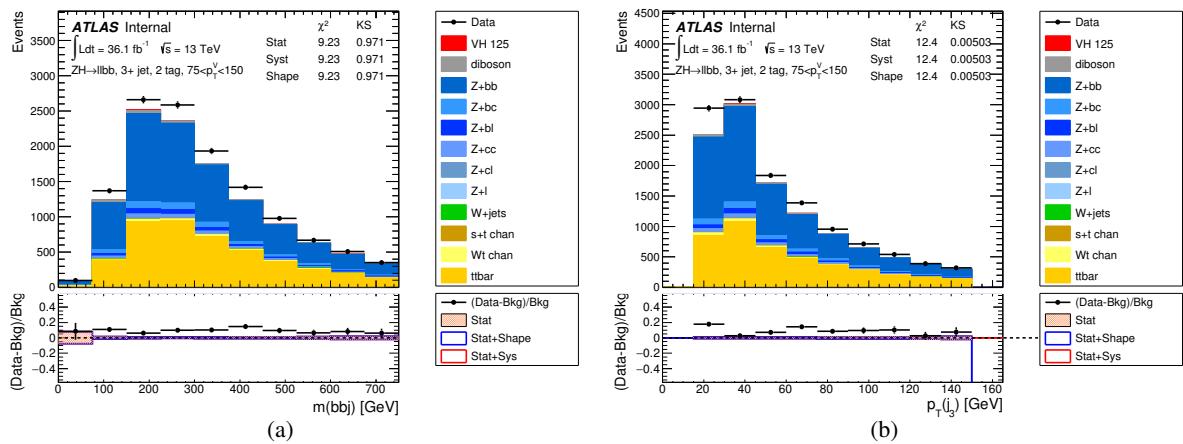


Figure 119: The pre-fit m_{BBj} and third leading jet p_T distributions in the 2-leptons channel for 2-tag events, in the 3+jets, $75 < p_T^V < 150$ GeV are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

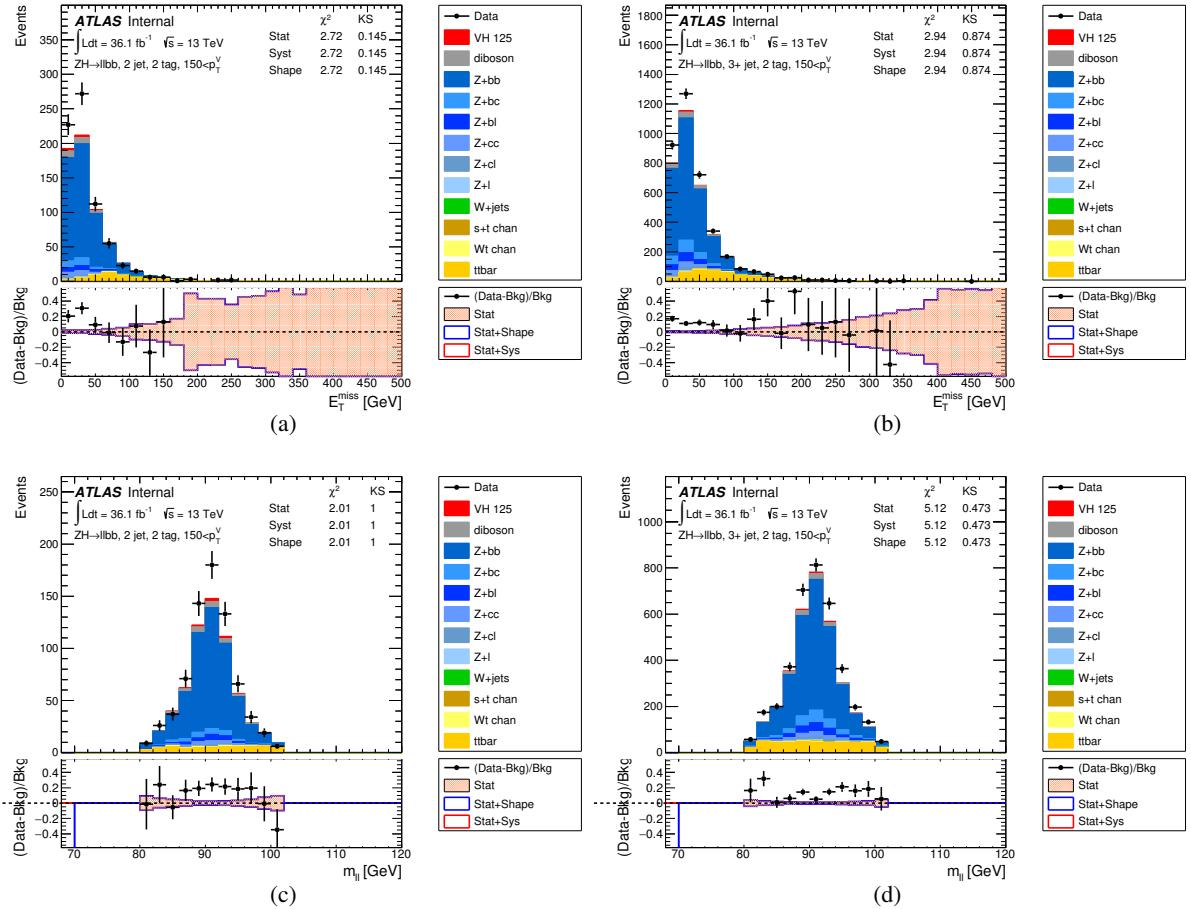


Figure 120: The pre-fit missing transverse energy distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $p_T^V > 150$ GeV and (b) 3+jets, $p_T^V > 150$ GeV categories are shown. The pre-fit m_{ll} distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $p_T^V > 150$ GeV and (d) 3+jets, $p_T^V > 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

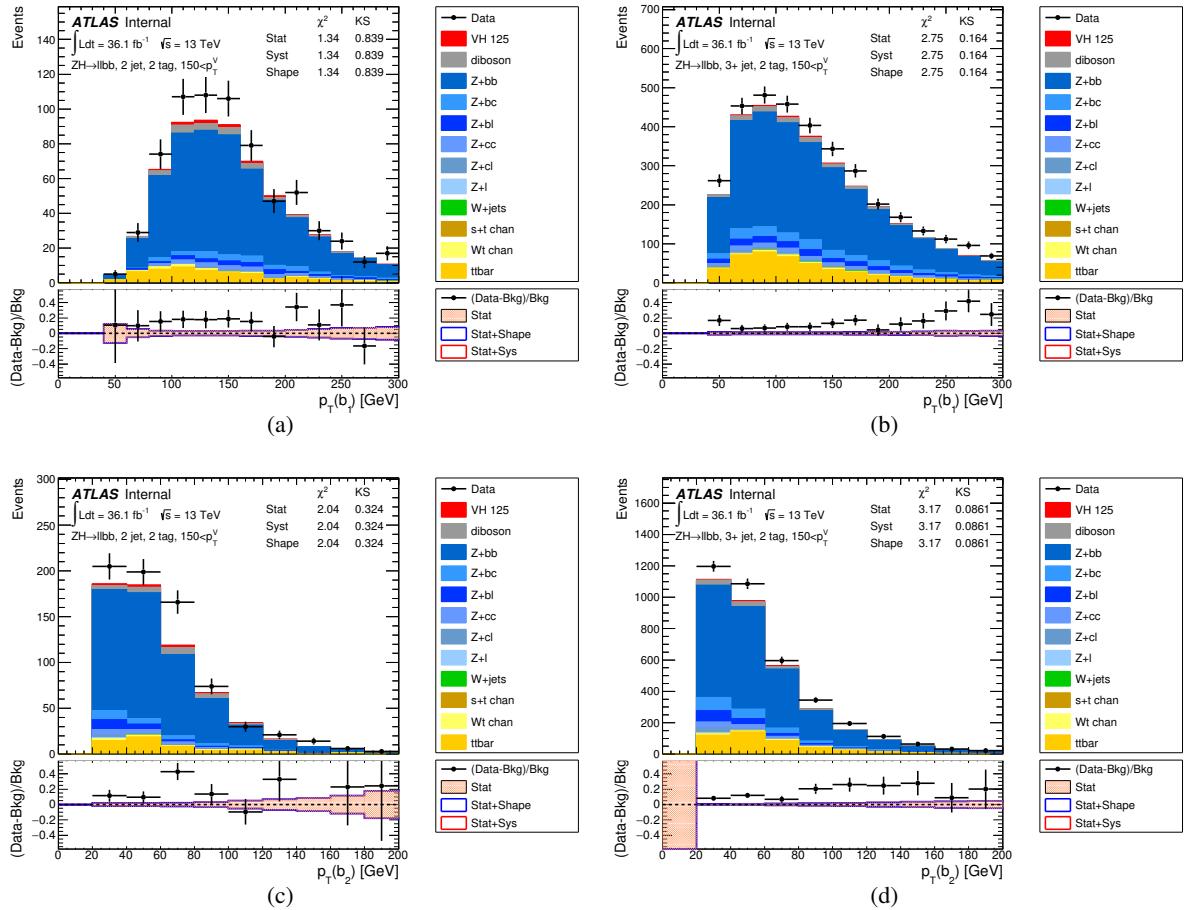


Figure 121: The pre-fit p_T of leading b-tagged jet distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $p_T^V > 150$ GeV and (b) 3+jets, $p_T^V > 150$ GeV categories are shown. The pre-fit p_T of second leading b-tagged jet distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $p_T^V > 150$ GeV and (d) 3+jets, $p_T^V > 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

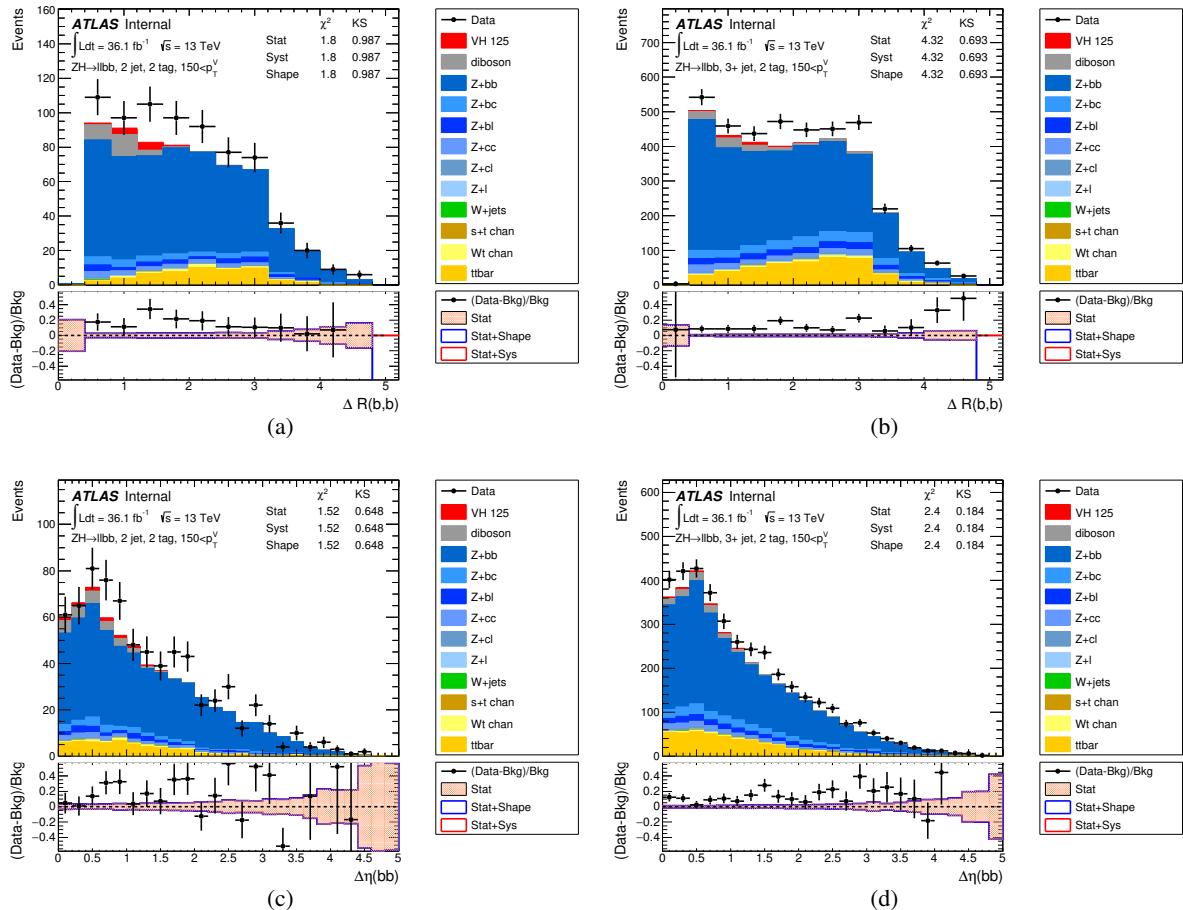


Figure 122: The pre-fit ΔR_{bb} distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $p_T^V > 150$ GeV and (b) 3+jets, $p_T^V > 150$ GeV categories are shown. The pre-fit $\Delta\eta_{bb}$ distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $p_T^V > 150$ GeV and (d) 3+jets, $p_T^V > 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

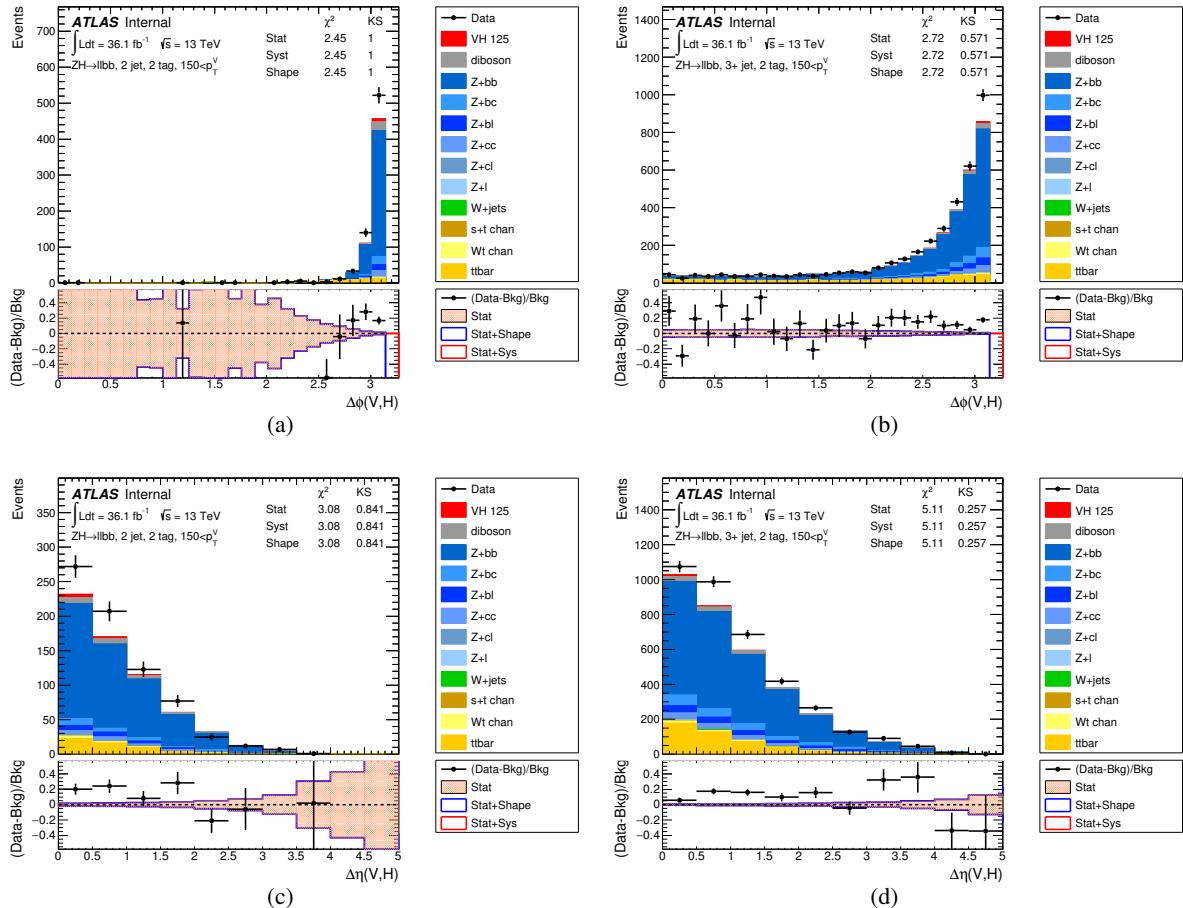


Figure 123: The pre-fit $\Delta\phi_{Vbb}$ distributions in the 2-leptons channel for 2-tag events, in the (a) 2jets, $p_T^V > 150$ GeV and (b) 3+jets, $p_T^V > 150$ GeV categories are shown. The pre-fit $\Delta\eta_{Vbb}$ distributions in the 2-leptons channel for 2-tag events, in the (c) 2jets, $p_T^V > 150$ GeV and (d) 3+jets, $p_T^V > 150$ GeV categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

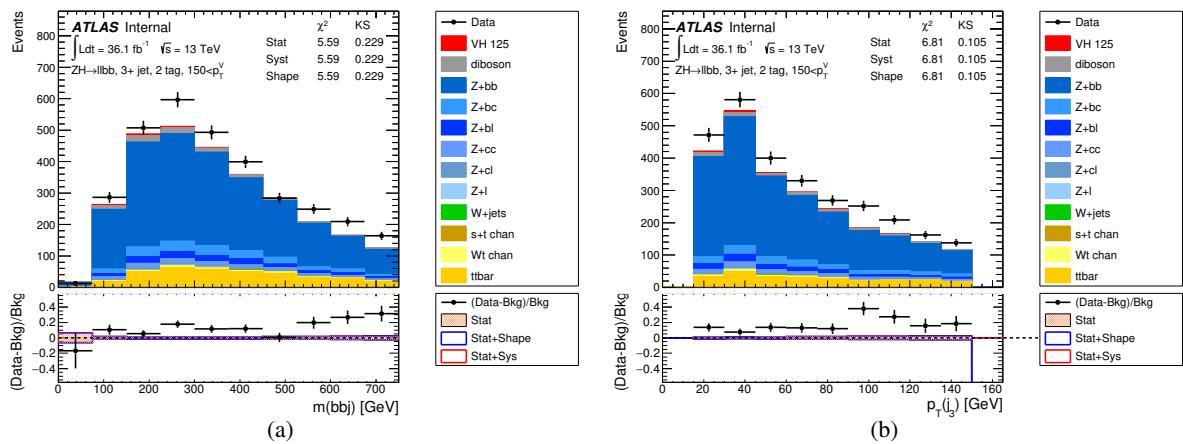


Figure 124: The pre-fit m_{BBj} and third leading jet p_T distributions in the 2-leptons channel for 2-tag events, in the 3+jets, $p_T^V > 150$ GeV are shown. The background and signal samples are normalized to the expected cross-section predictions.

Figure 125 shows MET significance defined in Sec 3.2.5. Signal and Z+jets background which does not have intrinsic missing transverse momentum shows lower MET significance. In case that high- p_T objects exist in the events tend to be high MET due to resolution. However MET significance still has good separation power in such a case. Low MET significance $E_T^{\text{miss}}/\sqrt{H_T} < 3.5\sqrt{\text{GeV}}$ has been required in the m_{bb} cut-base analysis in 2lepton.

Not reviewed, for internal circulation only

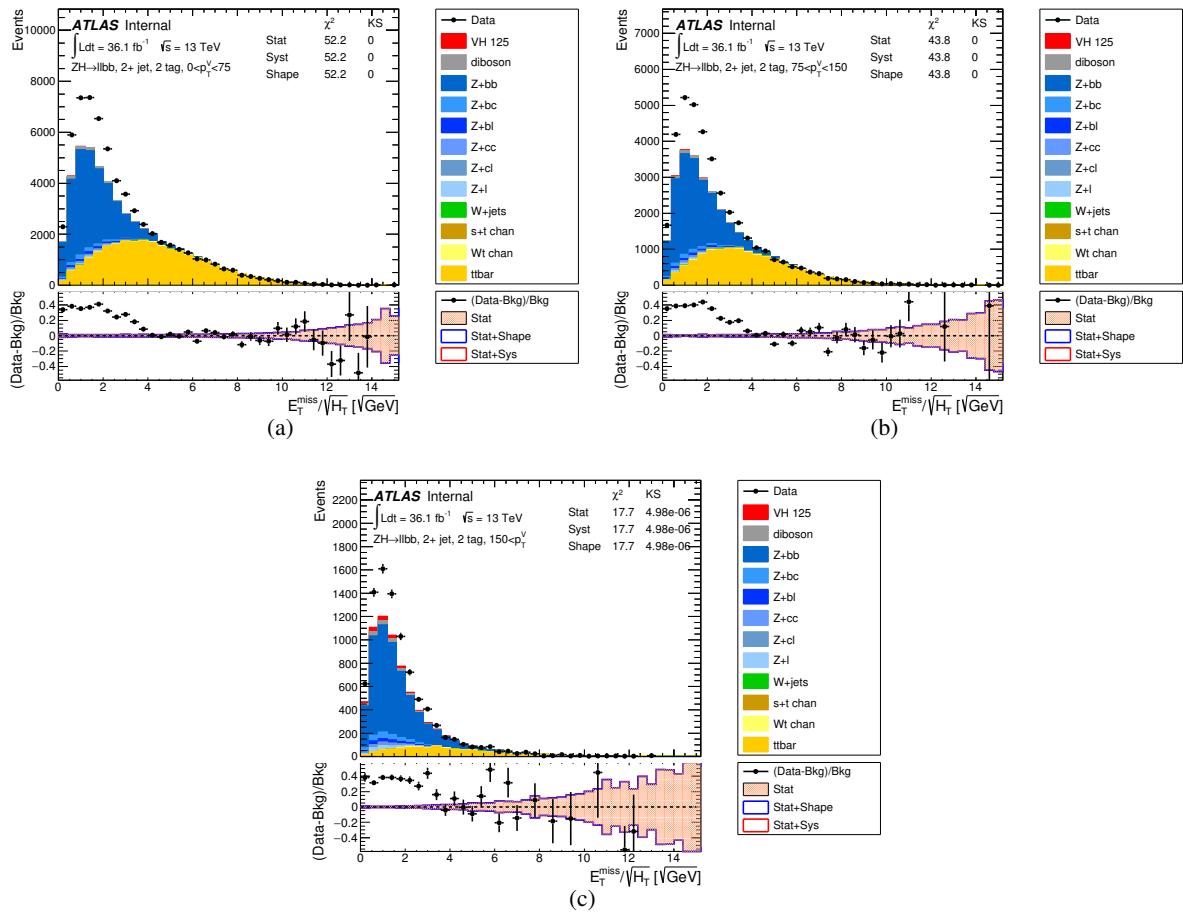


Figure 125: The pre-fit MET significance distributions in the 2-leptons channel for 2-tag events. (a) $p_T^V < 75 \text{ GeV}$, (b) $75 < p_T^V < 150 \text{ GeV}$ (c) $150 < p_T^V$ categories are shown. Two or more than two jets categories are merged and additional ΔR_{bb} selection for cut-based analysis not applied here. The background and signal samples are normalized to the expected cross-section predictions.

1279 F. Effect of MC stat uncertainties on the fit

1280 A possible bias on the mean of the fitted signal strength, $\langle \hat{\mu} \rangle$, from nuisance parameters (NPs) with large
 1281 priors is investigated in this section. A minimal example likelihood is constructed, in which the NPs for
 1282 the MC statistical uncertainties (γ -parameters, uncorrelated among bins) are employed to demonstrate
 1283 the effect. The example likelihood consists of two bins, each containing an expected background of
 1284 $b = 0.5 \pm 0.25$ (stat.) and an expected signal of $s = 0.5 \pm 0$ (stat.). An illustration is given in Figure 126.

Two alternatives to parameterise the γ -parameters in the binned likelihood are investigated:

$$\mathcal{L} = \prod_{i \in \text{bins}} P(N_i | \gamma_i(\mu s + b)) \cdot P(n_i | \gamma_i n_0) \quad n_0 = (s + b)^2 / \sigma_{s+b}^2 \quad (10)$$

$$\mathcal{L} = \prod_{i \in \text{bins}} P(N_i | \mu s + \gamma_i b) \cdot P(n'_i | \gamma_i n'_0) \quad n'_0 = b^2 / \sigma_b^2. \quad (11)$$

1285 Here, P is the Poisson distribution, μ is the signal strength parameter and n_0 (n'_0) is the effective number
 1286 of generated MC events, calculated from the corresponding statistical uncertainty, σ_{s+b} (σ_b). The second
 1287 parametrization is used in the actual analysis as the uncertainties in the signal are negligible. Fits to toy
 1288 data are performed to obtain the corresponding $\hat{\mu}_{\text{SB}}$ ($\hat{\mu}_B$). Toy data are drawn in two different ways: either
 1289 only the expected number of data events in each bin, N_i , are sampled (“conditional ensemble”) or the n_i
 1290 are sampled as well (“unconditional ensemble”). The $\hat{\mu}$ -distributions obtained from this setup are shown
 1291 in Figure 127. Significant pulls with respect to unity are observed and considered to be a possible bias for
 1292 a fit result obtained from data.

1293 Biases are investigated by modifying parameters of the model: the relative MC statistical uncertainty of
 1294 the background, σ_b/b , the expected number of signal and background events, s and b , and the number
 1295 of bins, N_{bins} . The results are shown in Figure 128. A strong dependence of the bias on σ_b , a mild
 1296 dependence on N_{bins} and a weak dependence on $s = b$ is observed.

1297 The median and 68 % confidence intervals of the $\hat{\mu}$ -distributions are evaluated as well. However, the
 1298 distributions of $\hat{\mu}$ for the previous examples with $s = b = 0.5$, shown in Figure 127, are largely discrete,
 1299 which prohibits reliable estimates of these quantities. We therefore perform this investigation using
 1300 $s = b = 20$ which removes the discreteness in the distributions. As before, a scan of σ_b/b is performed to
 1301 investigate the behaviour of a possible bias. Additionally, the median and 68 % confidence intervals of $\hat{\mu}$
 1302 are shown. The results for $\langle \hat{\mu} \rangle$ are shown in Figure 129 showing significant biases when the background
 1303 uncertainty is larger than 20%.

1304 In conclusion, a possible bias on the mean of $\hat{\mu}$ originating from large MC statistical uncertainties was
 1305 studied. For $\sigma_b/b \leq 20\%$ the bias is at the percent level and is deemed to be small enough to obtain
 1306 reliable results for the present analysis.

Not reviewed, for internal circulation only

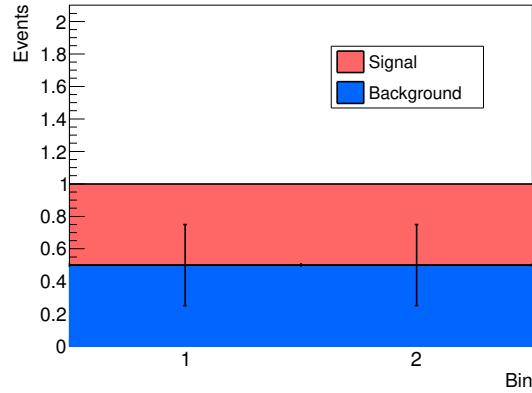


Figure 126: Illustration of example likelihood, consisting of two bins: stacked histograms of the expected signal and background.

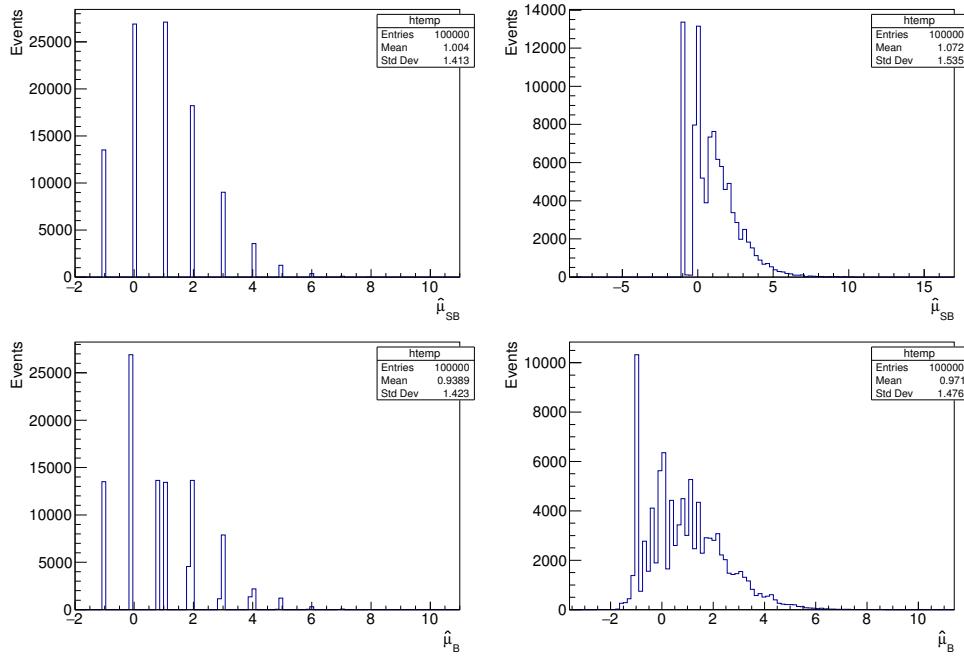


Figure 127: Distributions of $\hat{\mu}$ obtained from fits with $\gamma_i(\mu s + b)$ -parameterization (top) and $(\mu s + \gamma_i b)$ -parameterization (bottom) for fits to the conditional ensembles (left) and to the unconditional ensembles (right). Each distribution corresponds to fits to 100,000 toy datasets. A likelihood with two bins is employed, each containing expected yields of $s = 0.5 \pm 0$ and $b = 0.5 \pm 0.25$.

Not reviewed, for internal circulation only

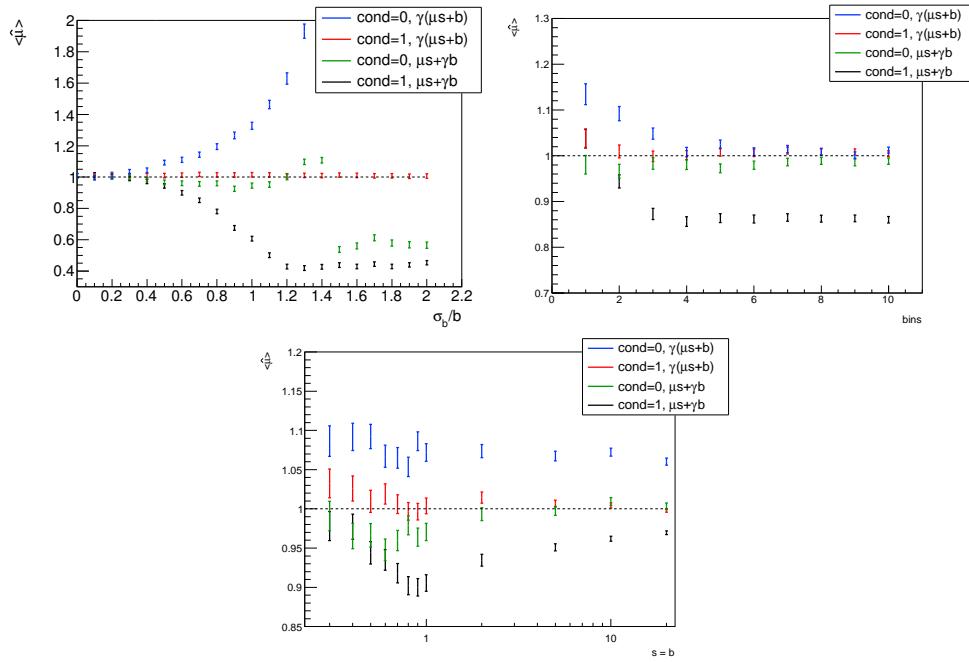


Figure 128: The mean of $\hat{\mu}$ as a function of σ_b/b , N_{bins} , and $s = b$ obtained from fits to toy data in the various setups listed in the legends. Each result is obtained from fits to 10,000 toy datasets. For the scan over σ_b/b the parameters are kept at $s = 0.5 \pm 0$ and $b = 0.5$, for the scan over N_{bins} : $s = 0.5 \pm 0$ and $b = 0.5 \pm 0.25$, finally in the scan over $s=b$ the uncertainty on b is always kept at 50%.

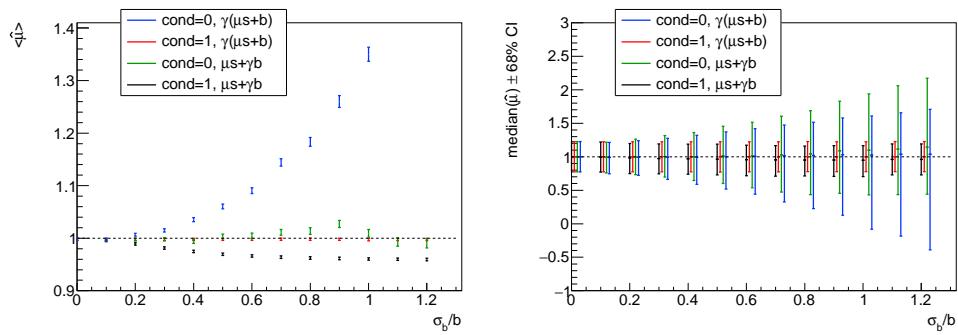


Figure 129: The mean (left), and median and 68 % confidence intervals (right) of $\hat{\mu}$ as a function of σ_b/b . Each result is obtained from fits to 10,000 toy datasets. A likelihood with two bins is employed, each containing expected yields of $s = 20 \pm 0$ and $b = 20 \pm \sigma_b$.

G. Truth Tagging Closure

1307 Studies have been carried out to determine the impact of possible biases which may be introduced through
1308 the use of truth tagging. Figure 130 presents comparisons of the m_{bb} distribution when using direct
1309 tagging (black) and truth tagging (red) for $W + cc, cl, l$ processes in the 1-lepton 2-jet and 3-jet signal
1310 regions. Truth tagging is used be default for these processes.

1312 From these comparisons, a number of features can be seen. Firstly, the use of truth tagging can be seen to
1313 be highly beneficial all of these distributions; the statistical uncertainty is greatly reduced compared to the
1314 case with direct tagging. Whilst difficult to conclude due to the large statistical fluctuations in the direct
1315 tagging case, in the $W + cc$ distribution, it is possible that the yield with truth tagging is approximately
1316 10 – 20% lower. The $W + cl$ and $W + l$ shapes appear to be in good agreement.

1317 Given the small contribution of these background processes, and the large normalisation uncertainties,
1318 the impact of any bias introduced through the use of truth tagging on the final result will be minimal. The
1319 benefits of the extra statistics from truth tagging will likely outweigh any negatives from a small bias.

Not reviewed, for internal circulation only

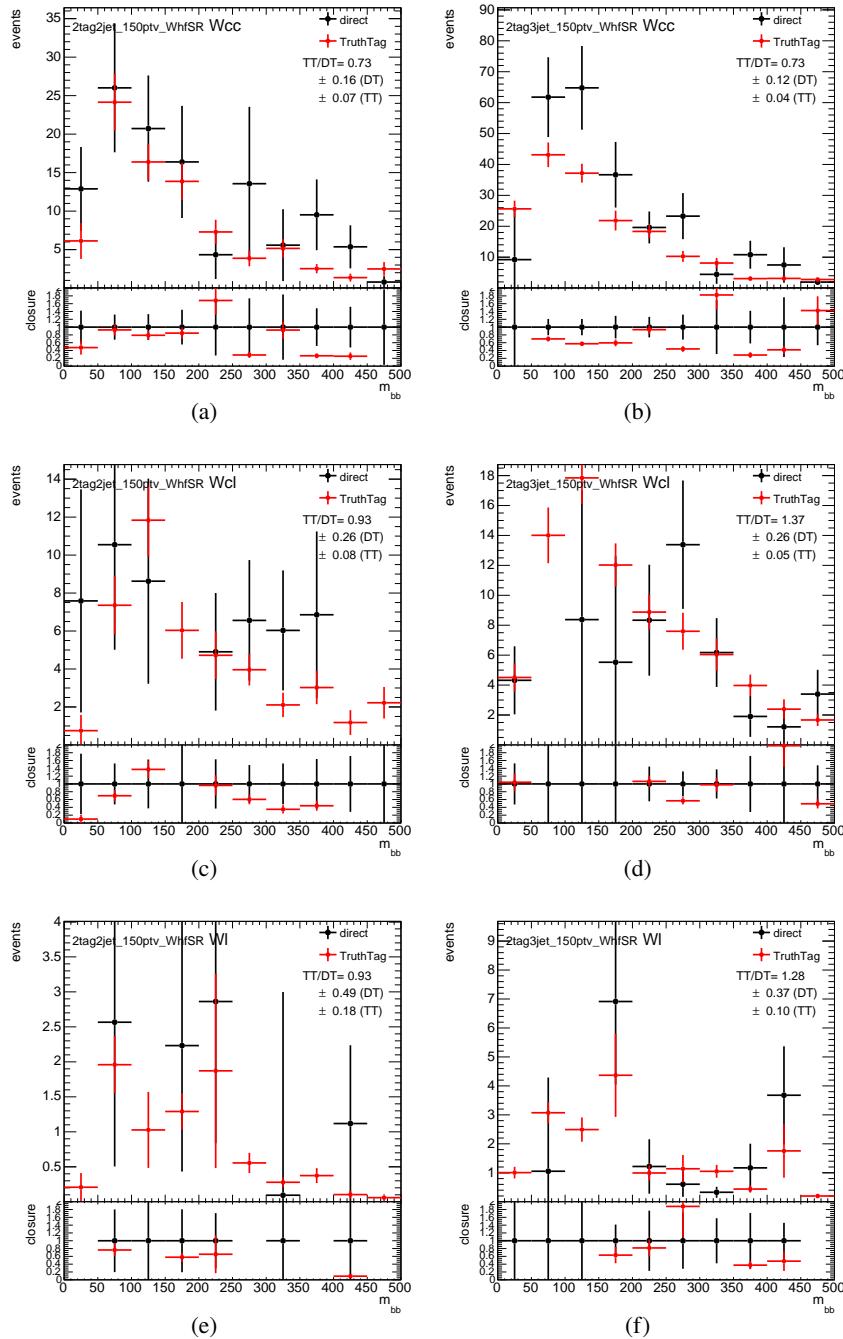


Figure 130: Comparison of m_{bb} shape distributions for $W + cc$ (top row), $W + cl$ (middle row) and $W + l$ (bottom row), for the 2-jet (left column) and 3-jet (right column) signal regions with direct tagging (black) and truth tagging (red).

1320 H. Continuous Tagging

1321 In Run1 analysis the pseudo-continuous tagging algorithm were used [13]. We have investigated similar
 1322 method in Run2 b -tagging algorithm, MV2c10.

1323 The MV2c10 weight variables are included in the BDT, and sensitivity improvement has been checked
 1324 in simple LLR using BDT output distribution. The systematic uncertainties are not taken into account.
 1325 Table 49-51 shows the sensitivity comparison between default b -tagging strategy and continuous tagging.
 1326 Sensitivity has been checked in some b -tagging operating point because continuous tagging would be
 1327 more powerful in looser b -tagging operating point. Potential improvement, 10-20% has been obtained
 1328 from continuous b -tagging. However the systematic is not taken into account in this studies. This alters the
 1329 background composition significantly and brings additional backgrounds. We decided not to implement
 1330 the continuous tagging to simplify the analysis in this stage. This potential improvement should be tested
 1331 for future.

b -tagging WP	Njet=2		Njet=3	
	nominal	+MV2c10	nominal	+MV2c10
70%(default)	1.31	1.39	0.97	1.07
77%	1.33	1.52	0.99	1.20
85%	1.21	1.59	0.88	1.20

Table 49: LLR significant calculated on BDT output distribution. Systematic uncertainties are not taken into account.

b -tagging WP	2tag,2jet	2tag,3+jet
70%(default)	0.621	0.426
77%+MV2c10	0.645	0.465
85%+MV2c10	0.659	0.456

Table 50: LLR significant calculated on BDT output distribution. Systematic uncertainties are not taken into account. In 3 jet categories, MV2c10 information up to third jet is used.

b -tagging WP	2tag,2jet ($low p_T^Z$)	2tag,3+jet ($low p_T^Z$)	2tag,3+jet ($high p_T^Z$)	2tag,3+jet ($high p_T^Z$)
70%(default)	0.753	0.596	0.968	0.84
77%+MV2c10	0.781	0.64	1.03	0.922
85%+MV2c10	0.821	0.67	1.10	0.922

Table 51: LLR significant calculated on BDT output distribution. Systematic uncertainties are not taken into account. In 3+jet categories, MV2c10 information up to third jet is used.

I. BDT optimization studies

1332 Details of the MVA training are discussed in this appendix.

1334 **I.1. 1-lepton**

1335 **I.1.1. MVA performance in 2-jet**

1336 Figure 131 shows the BDT input variables after the event selection in the 2-jet category. Figure 132 shows
1337 expected background rejection vs signal obtained from BDT training procedure and the BDT distribution
1338 for both test and training samples. No significant overtraining can be seen from this comparison. Figure 133
1339 shows the correlation among input variables.

Not reviewed, for internal circulation only

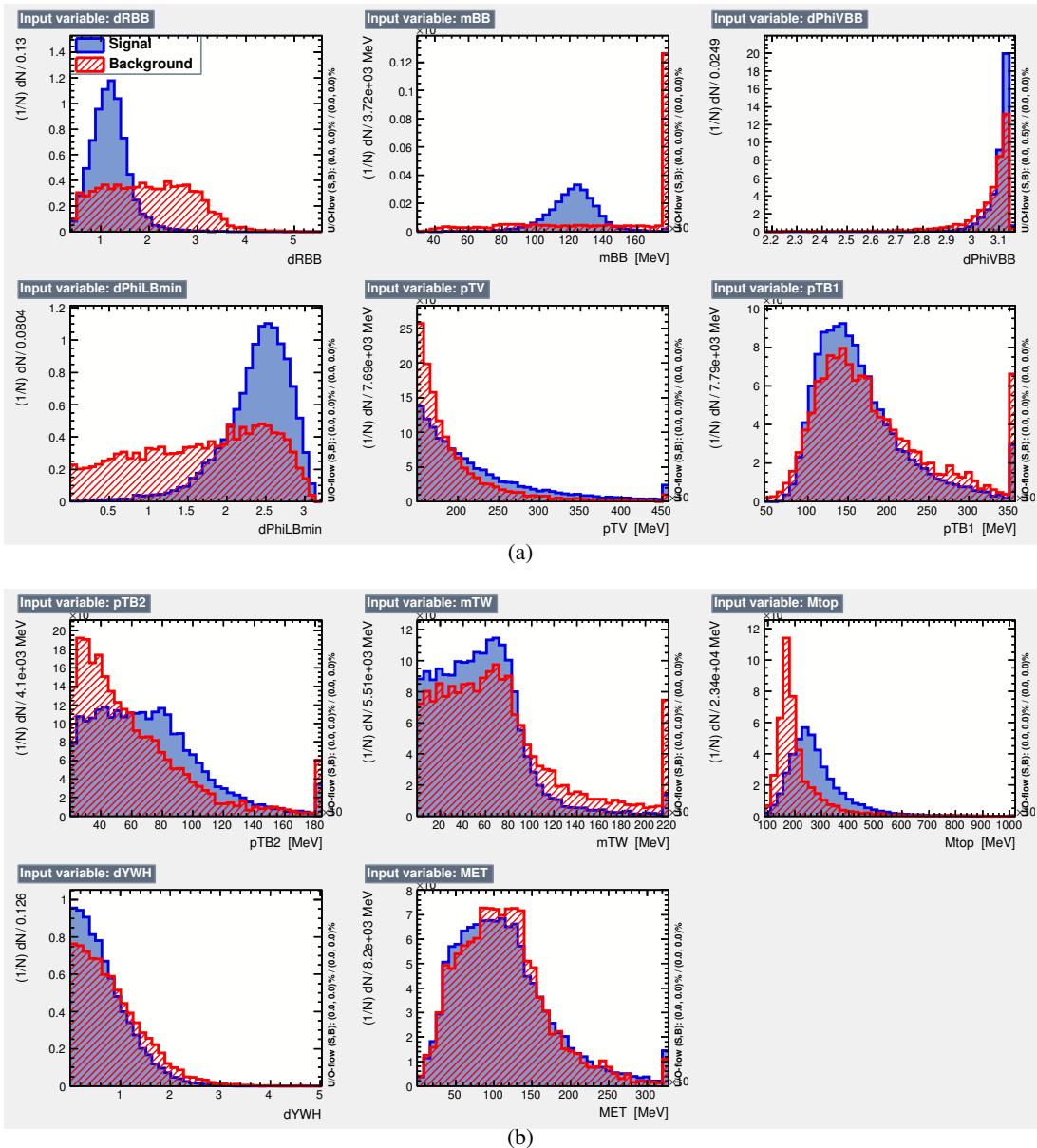


Figure 131: Input variables passed to the BDT for signal (blue) and background (red) samples in the 1-lepton 2-tag 2-jet region are shown.

Not reviewed, for internal circulation only

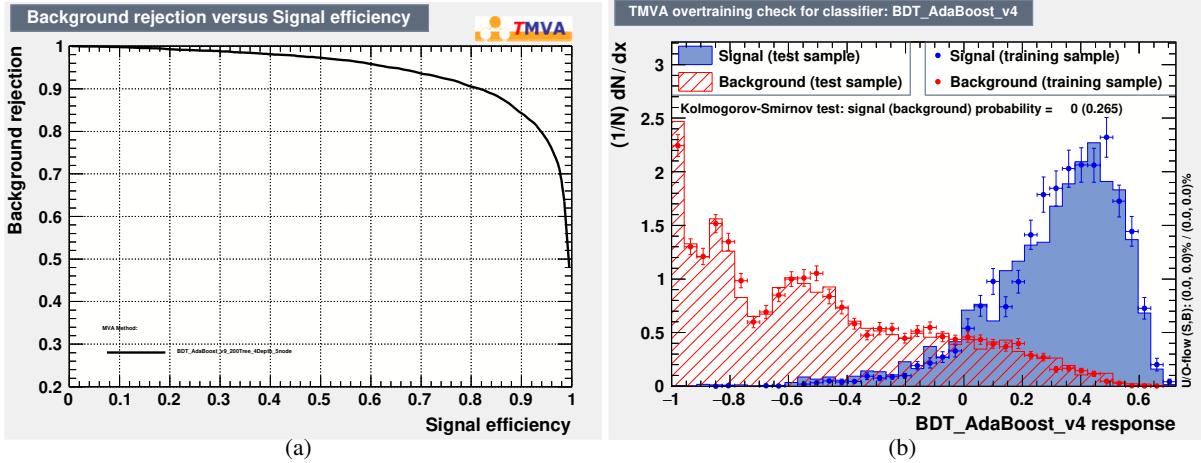


Figure 132: Distributions related to the BDT training procedure, in the 1-lepton 2-tag 2-jet region, are shown. In (a) the expected background rejection as a function of signal efficiency is provided, and in (b), the BDT distributions obtained while training (dots) and testing (histogram), for both signal and background, are found to be in reasonable agreement.

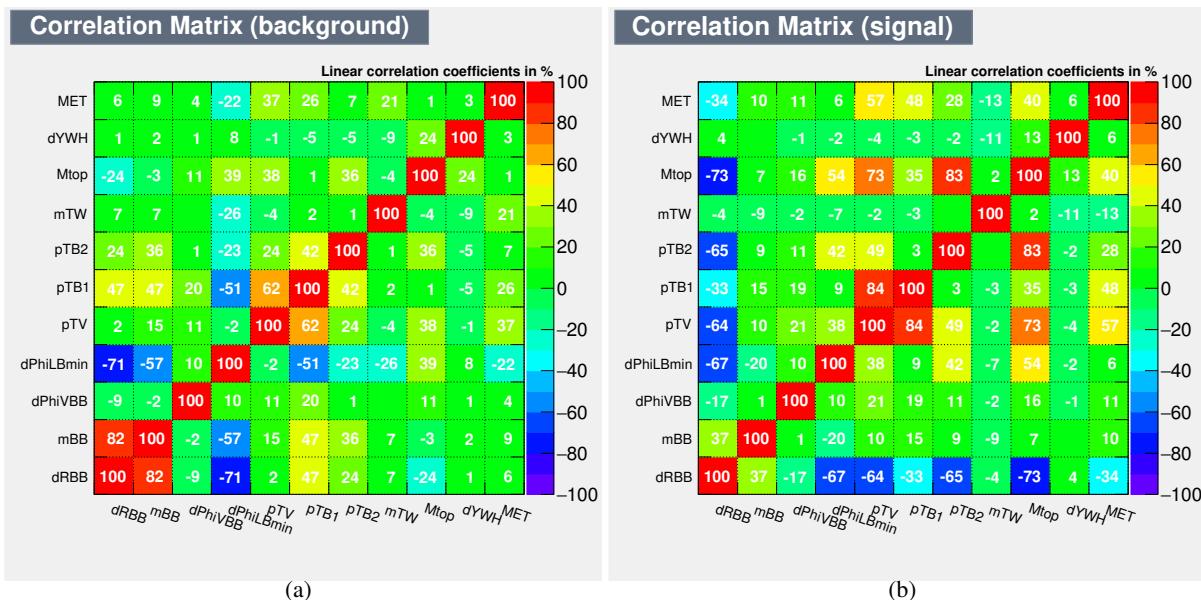


Figure 133: Correlation matrices of the background (a) and signal (b) input variables obtained in the BDT training procedure, in the 1-lepton 2-tag 2-jet region, are shown.

I.1.2. MVA performance in 3-jet

1340 **I.1.2. MVA performance in 3-jet**

1341 Figure 134 shows the BDT input variables after the event selection for the 3-jet category. Figure 135 shows
1342 expected background rejection vs signal obtained from BDT training procedure and the BDT distribution
1343 for both test and training samples. No significant overtraining can be seen from this comparison. Figure 136
1344 shows the correlation among input variables.

[Not reviewed, for internal circulation only]

Not reviewed, for internal circulation only

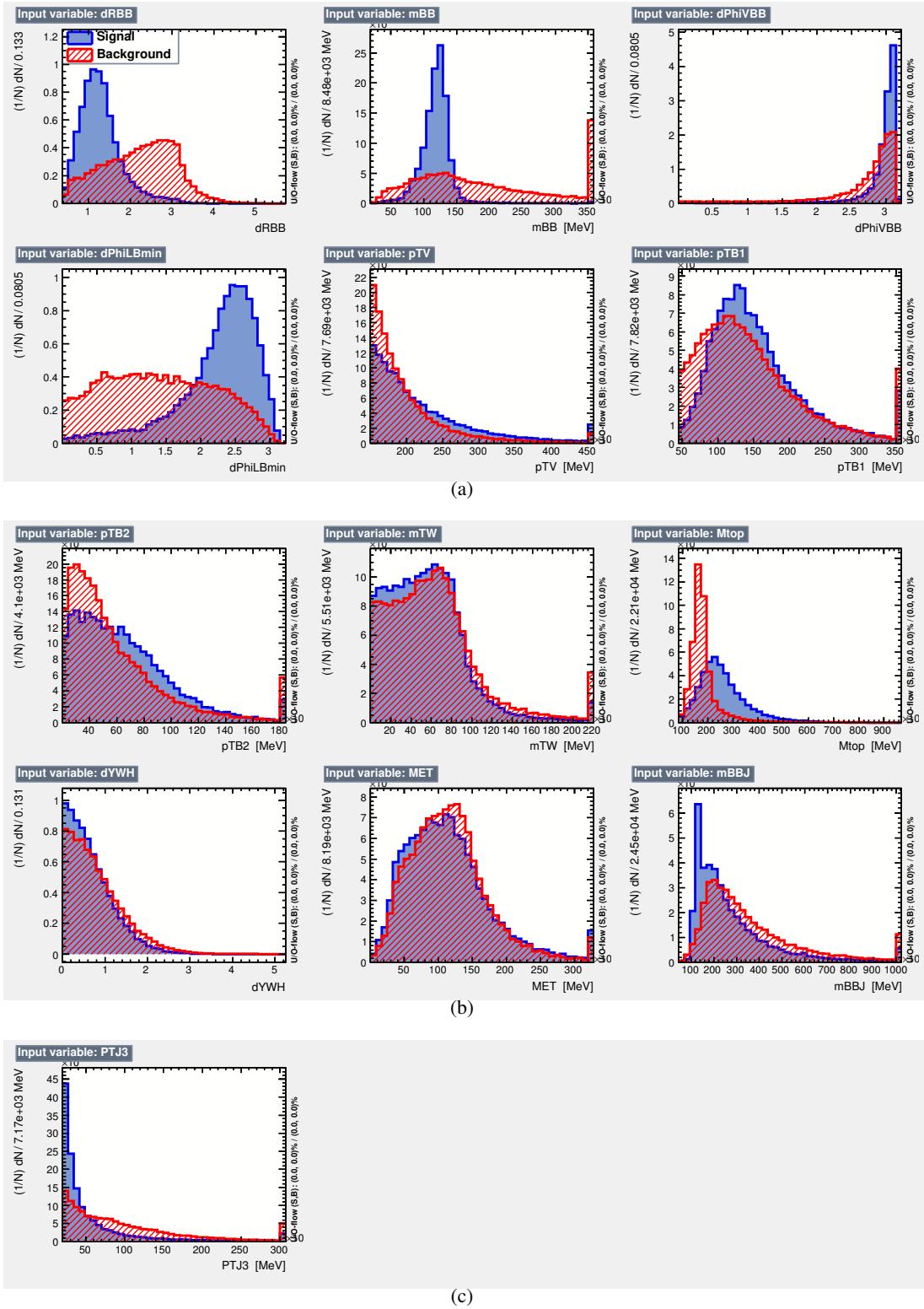


Figure 134: Input variables passed to the BDT for signal (blue) and background (red) samples in the 1-lepton channel are shown, in the 2-tag 3-jet region.

Not reviewed, for internal circulation only

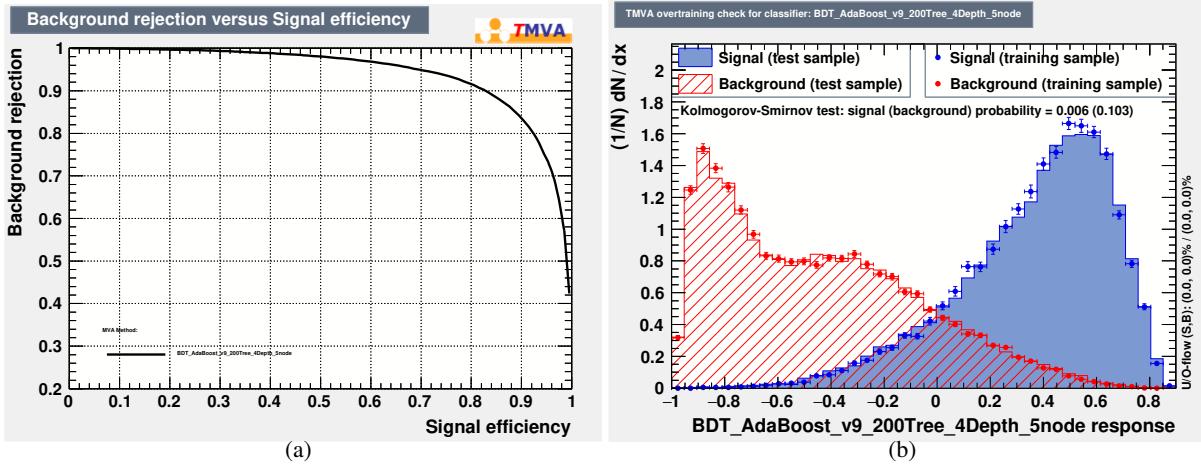


Figure 135: Distributions related to the BDT training procedure, in the 1-lepton channel, are shown for the 2-tag 3-jet region. In (a) the expected background rejection as a function of signal efficiency is provided, and in (b), the BDT distributions obtained while training (dots) and testing (histogram), for both signal and background, are found to be in reasonable agreement.

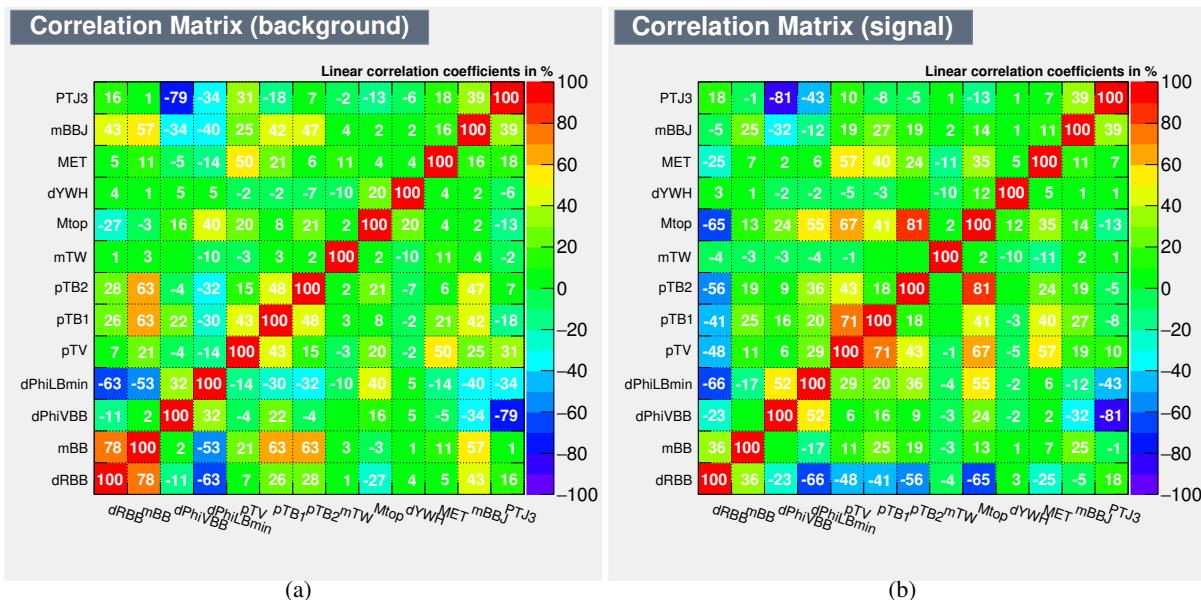
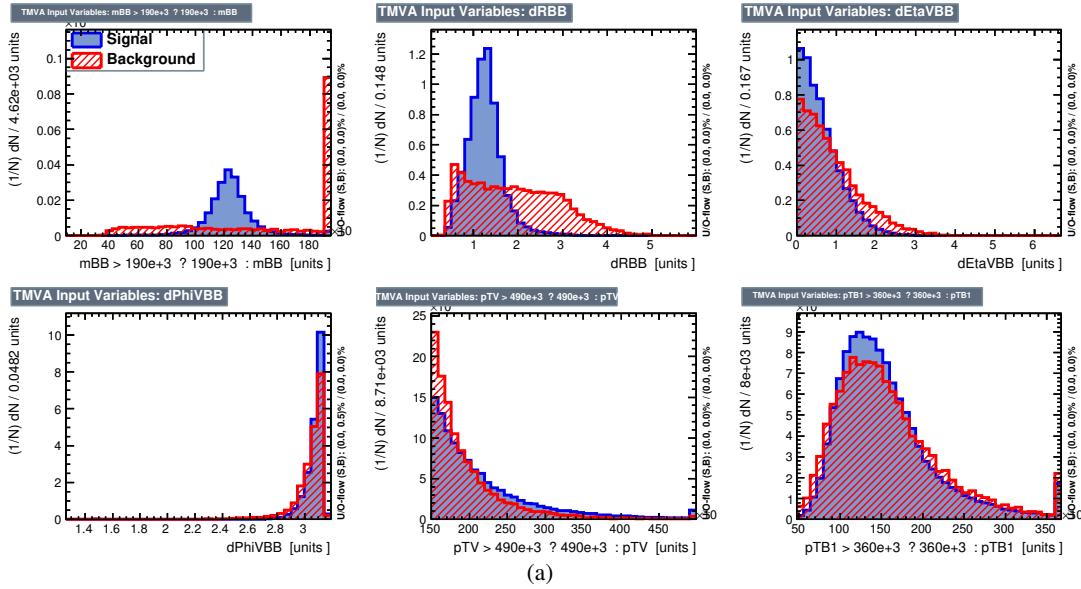


Figure 136: Correlation matrices of the background (a) and signal (b) input variables obtained in the BDT training procedure, in the 1-lepton 2-tag 3-jet region, are shown.

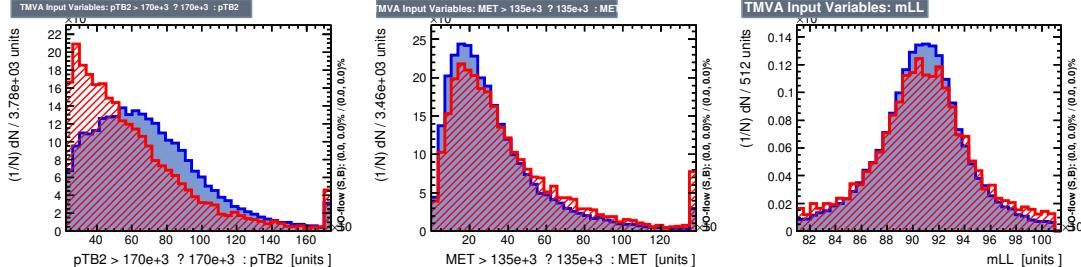
1345 **I.2. 2-lepton**

1346 **I.2.1. MVA performance in 2jet, $150 \text{ GeV} < p_T^V$**

1347 Figure 137 shows input variables after event selection in 2jet, $\text{high}p_T^V$ category. Figure 138 shows expected
 1348 background rejection vs signal obtained from BDT training procedure and the BDT distribution for both
 1349 test and training samples. No significant overtraining can be seen from this comparison. Figure 139 shows
 1350 the correlation among input variables.



(a)



(b)

Figure 137: Input variables passed to the BDT for signal (blue) and background (red) samples in the 2-lepton channel are shown.

Not reviewed, for internal circulation only

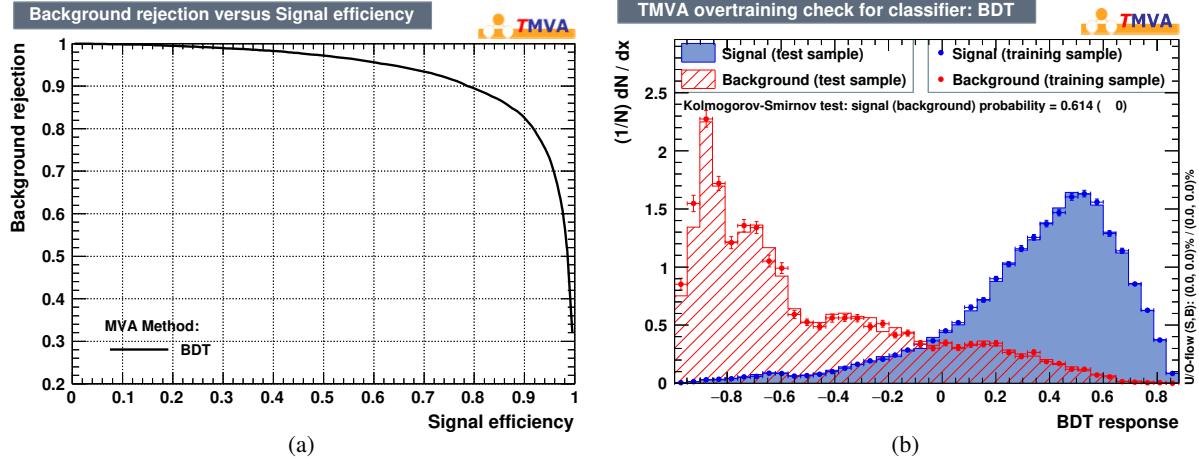


Figure 138: Distributions related to the BDT training procedure, in the 2-lepton channel, are shown. In (a) the expected background rejection as a function of signal efficiency is provided, and in (b), the BDT distributions obtained while training (dots) and testing (histogram), for both signal and background, are found to be in reasonable agreement.

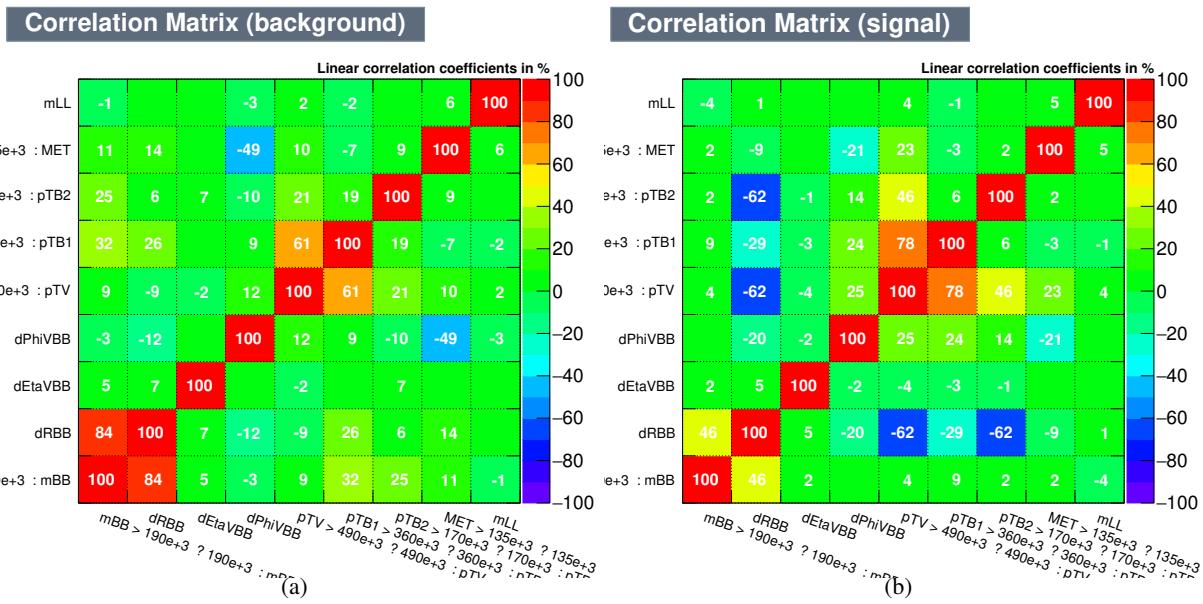


Figure 139: Correlation matrices of the background (a) and signal (b) input variables obtained in the BDT training procedure, in the 2-lepton channel, are shown.

1351 I.2.2. MVA performance in 2jet, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$

1352 Figure 140 shows input variables after event selection in 2jet, low p_T^V category. Figure 141 shows expected
 1353 background rejection vs signal obtained from BDT training procedure and the BDT distribution for both
 1354 test and training samples. No significant overtraining can be seen from this comparison. Figure 142 shows
 1355 the correlation among input variables.

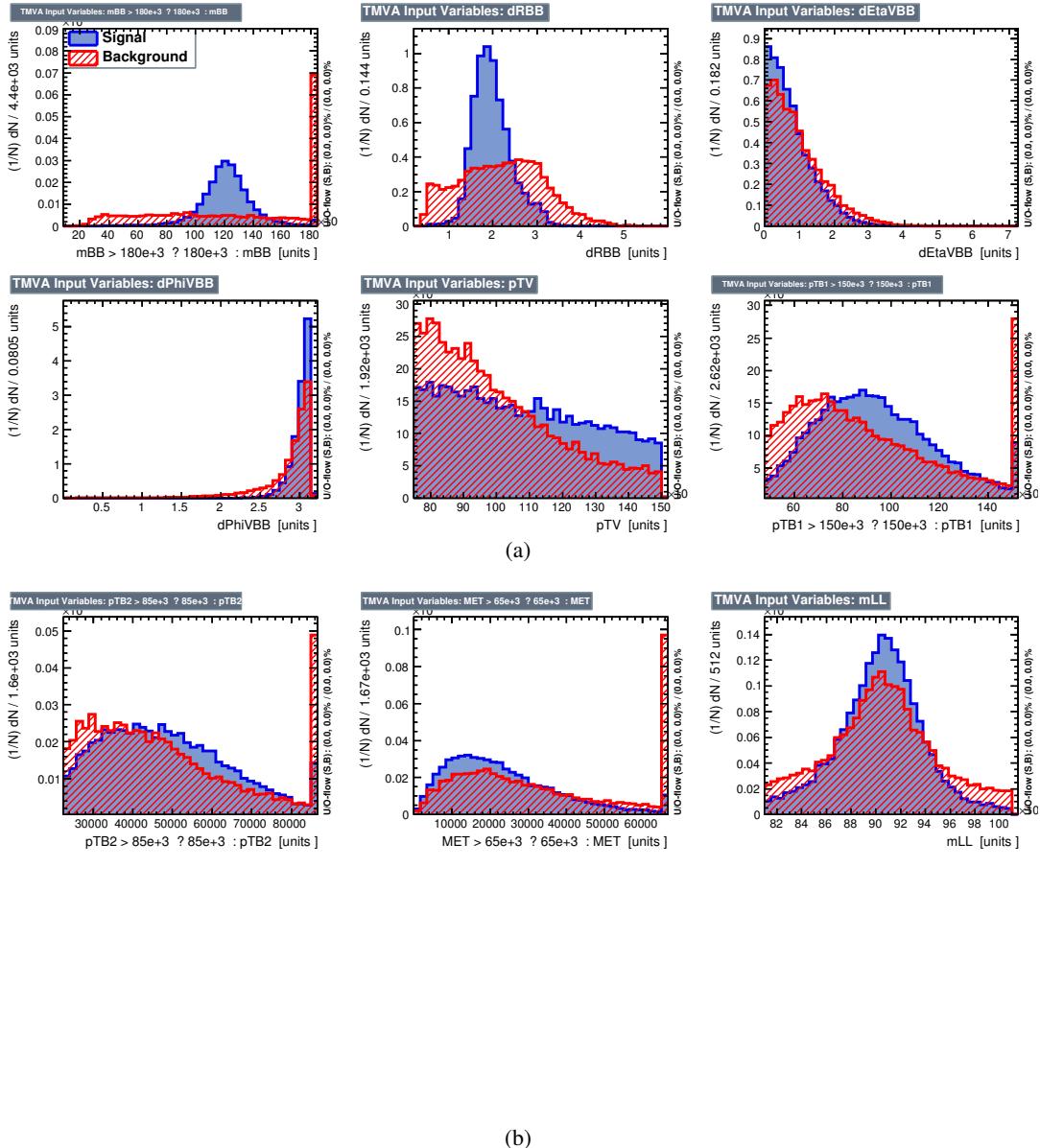


Figure 140: Input variables passed to the BDT for signal (blue) and background (red) samples in the 2-lepton channel are shown.

Not reviewed, for internal circulation only

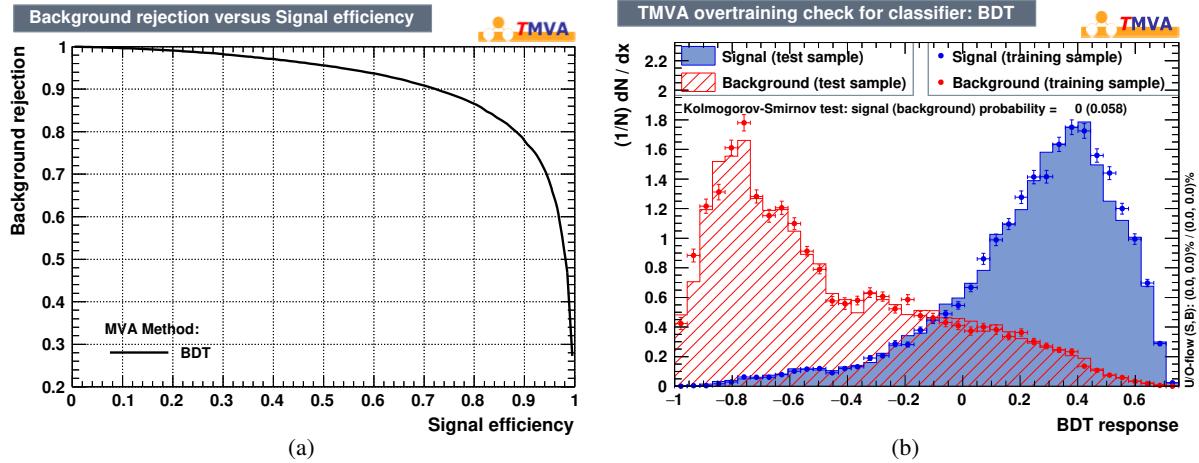


Figure 141: Distributions related to the BDT training procedure, in the 2-lepton channel, are shown. In (a) the expected background rejection as a function of signal efficiency is provided, and in (b), the BDT distributions obtained while training (dots) and testing (histogram), for both signal and background, are found to be in reasonable agreement.

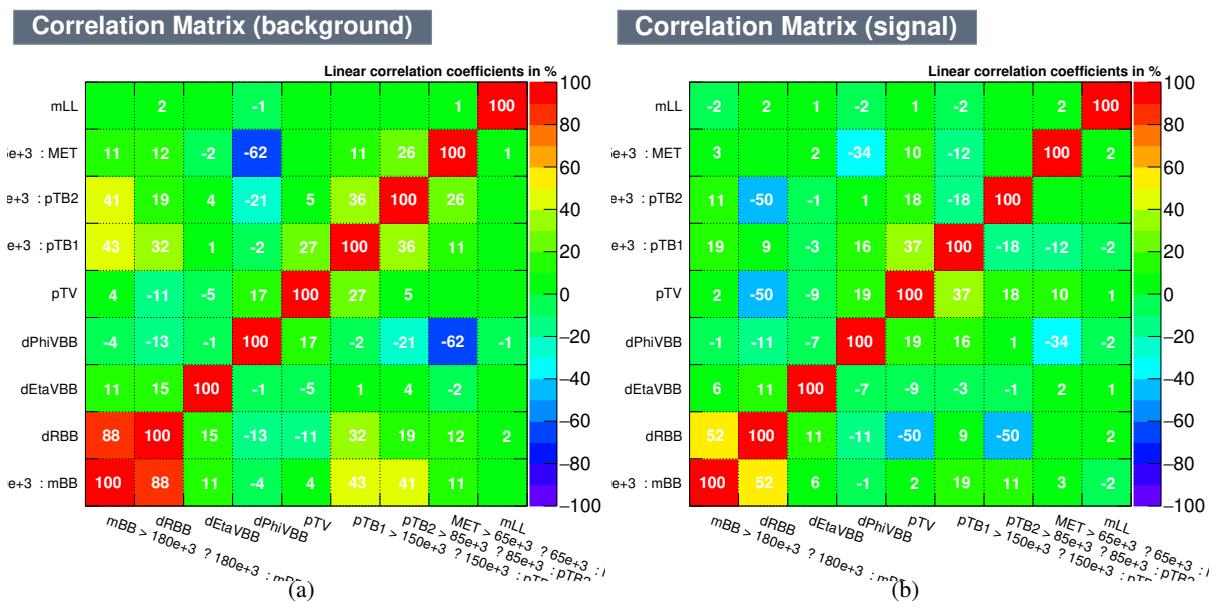


Figure 142: Correlation matrices of the background (a) and signal (b) input variables obtained in the BDT training procedure, in the 2-lepton channel, are shown.

1356 I.2.3. MVA performance in 3+jet, $75 \text{ GeV} < p_T^V < 150 \text{ GeV}$

1357 Figure 143 shows input variables after event selection in three or more than three jets, low p_T^V category.
 1358 Figure 144 shows expected background rejection vs signal obtained from BDT training procedure and the
 1359 BDT distribution for both test and training samples. No significant overtraining can be seen from this
 1360 comparison. Figure 145 shows the correlation among input variables.

Not reviewed, for internal circulation only

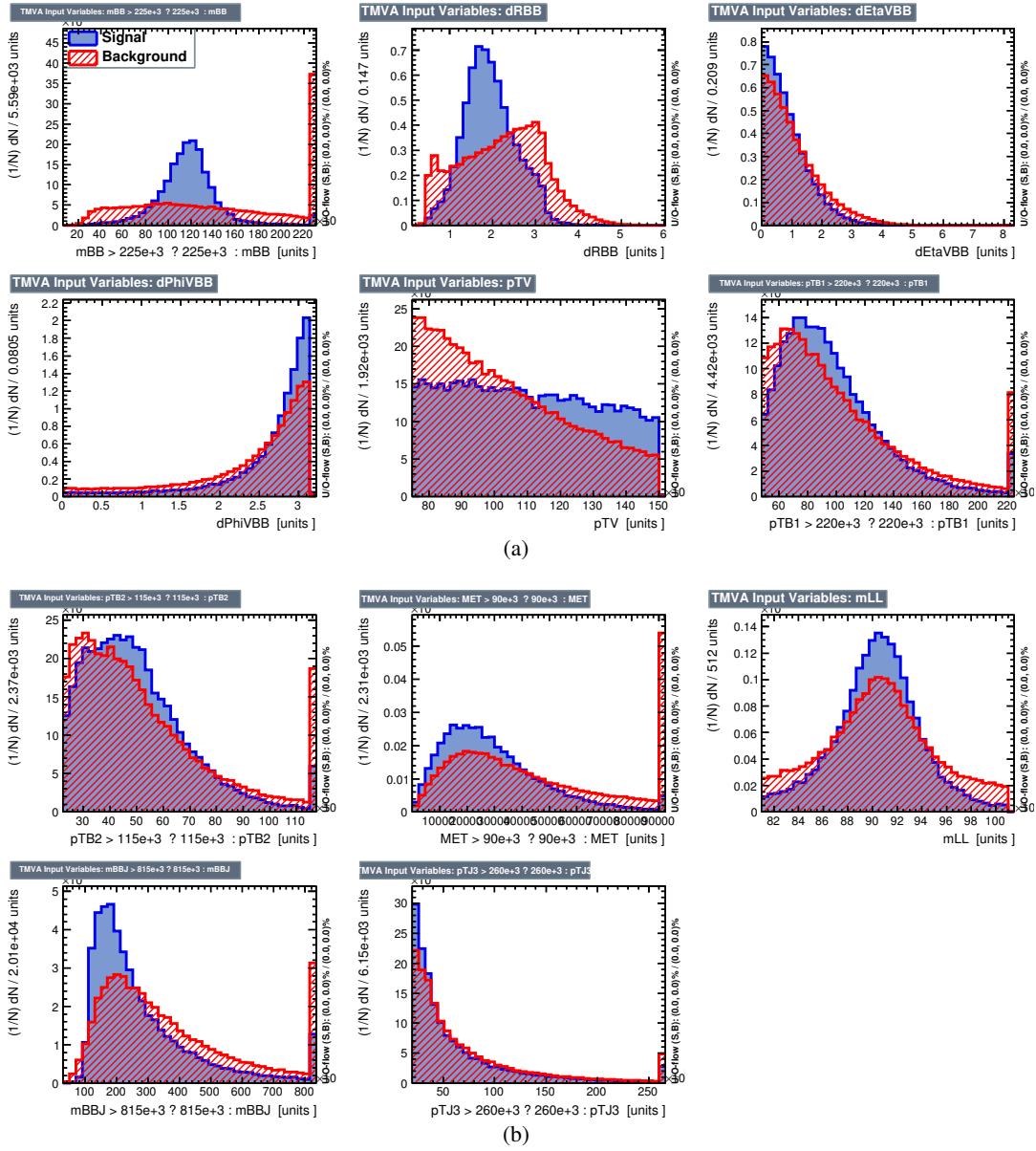


Figure 143: Input variables passed to the BDT for signal (blue) and background (red) samples in the 2-lepton channel are shown.

Not reviewed, for internal circulation only

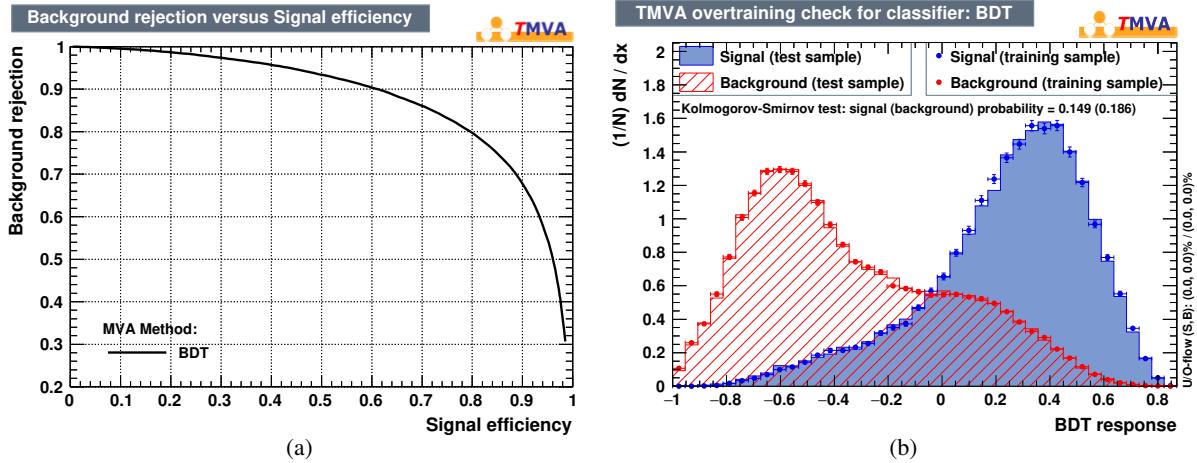


Figure 144: Distributions related to the BDT training procedure, in the 2-lepton channel, are shown. In (a) the expected background rejection as a function of signal efficiency is provided, and in (b), the BDT distributions obtained while training (dots) and testing (histogram), for both signal and background, are found to be in reasonable agreement.

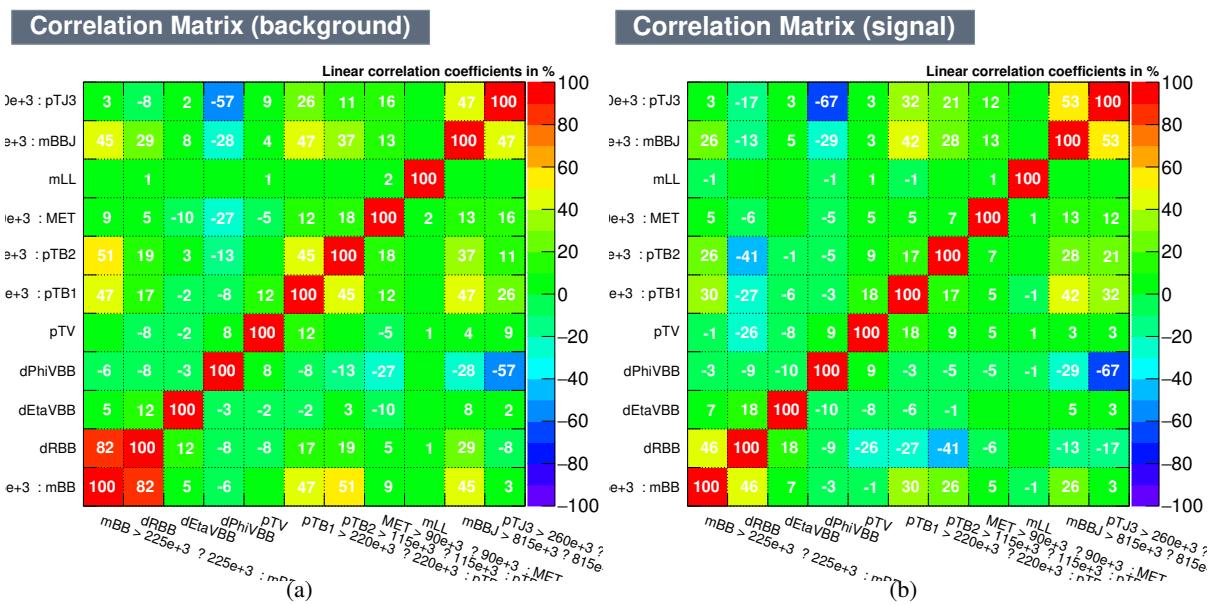


Figure 145: Correlation matrices of the background (a) and signal (b) input variables obtained in the BDT training procedure, in the 2-lepton channel, are shown.

1361 **J. Fit performances in each leptonic channel**

1362 In this section, the results of the fit to the Vh BDT, performed separately in each channel, are shown.

1363 **J.1. 0-lepton channel**

1364 In this section are shown the results of the fit performed in the 0-lepton channel Vh BDT.

1365 **J.1.1. Nuisance parameter pulls and constraints**

1366 In Figure 146, the nuisance parameter pulls and constraints are shown for the 0-lepton channel, when
1367 performing a unconditional fit on an Asimov dataset and the 2015 + 2016 data.

1368 **J.1.2. Correlation matrices**

1369 The correlation matrices from the unconditional fit in the 0-lepton channel are shown when fitting an
1370 Asimov dataset, and data, in Figure 147.

1371 **J.1.3. Nuisance parameter ranking**

1372 The ranking of the most important systematics is shown, for the 0-lepton unconditional fit, in Figure 148.

1373 **J.1.4. Post-fit distributions**

1374 Figure 149 shows the postfit BDT distributions in the 0-lepton channel after the unconditional fit.

1375 **J.1.5. Breakdown of statistical and systematic uncertainties**

1376 Breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together
1377 with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and each category
1378 of systematic uncertainty, is shown for the 0-lepton unconditional fit on 36.1 fb^{-1} of Asimov data in
1379 Table 52.

Total	+0.558 / -0.557
DataStat	+0.379 / -0.366
FullSyst	+0.409 / -0.421
Floating normalizations	+0.063 / -0.115
All normalizations	+0.066 / -0.125
All but normalizations	+0.376 / -0.390
Jets MET	+0.075 / -0.097
BTag	+0.193 / -0.228
Leptons	+0.000 / -0.000
Luminosity	+0.030 / -0.005
Diboson	+0.033 / -0.026
Zjets	+0.144 / -0.184
Wjets	+0.062 / -0.068
Model ttbar	+0.027 / -0.036
Model Single Top	+0.110 / -0.114
Model Multi Jet	+0.000 / -0.000
Signal Systematics	+0.115 / -0.061
MC stat	+0.234 / -0.254

Table 52: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and different sub-groups of systematic uncertainties, is shown in the 0-lepton unconditional fit.

Not reviewed, for internal circulation only

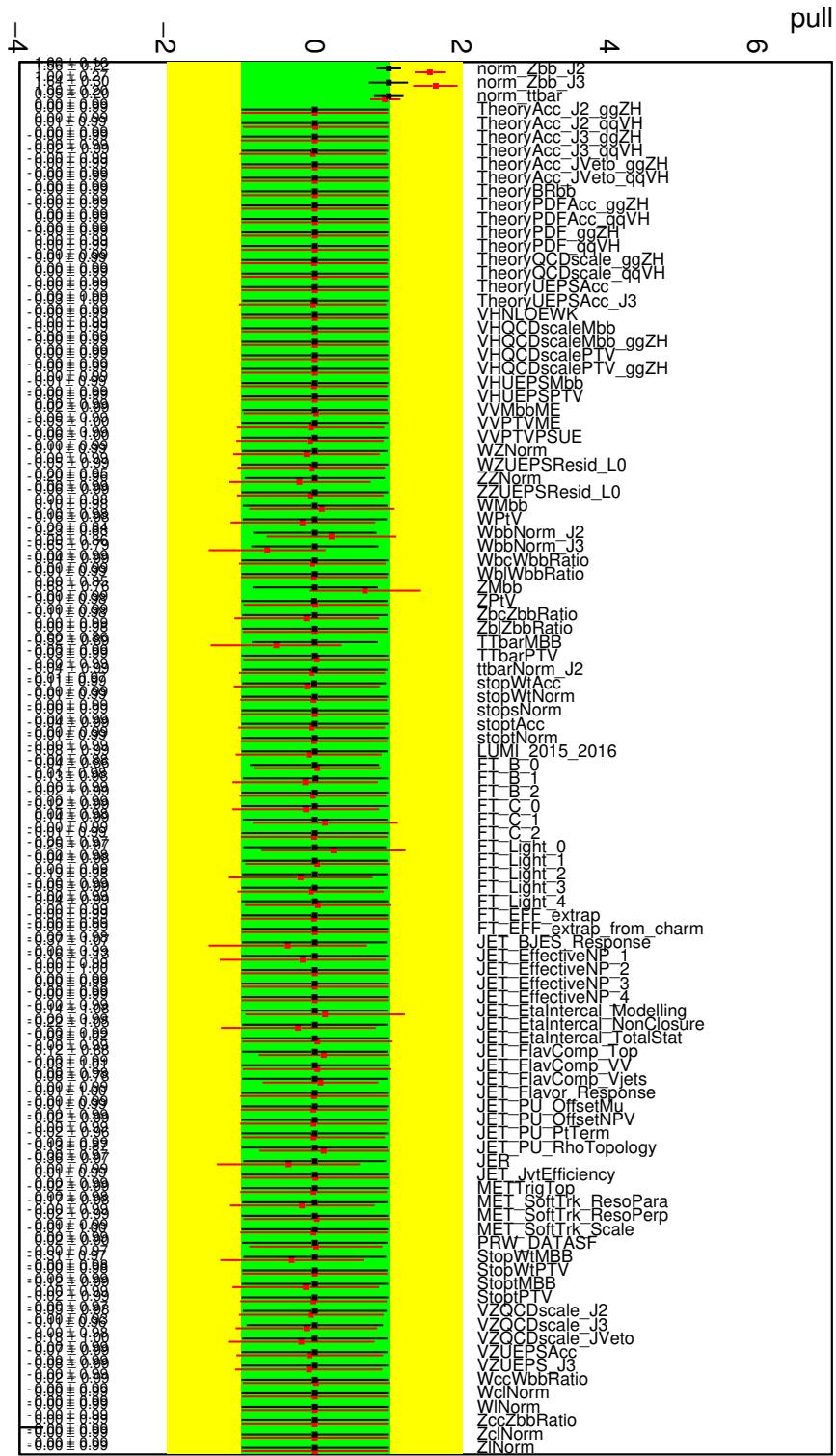


Figure 146: Nuisance parameter pulls and the free parameter scale factors corresponding to an unconditional fit performed in 0-lepton Asimov dataset (black) and 2015 and 2016 data (red)

Not reviewed, for internal circulation only

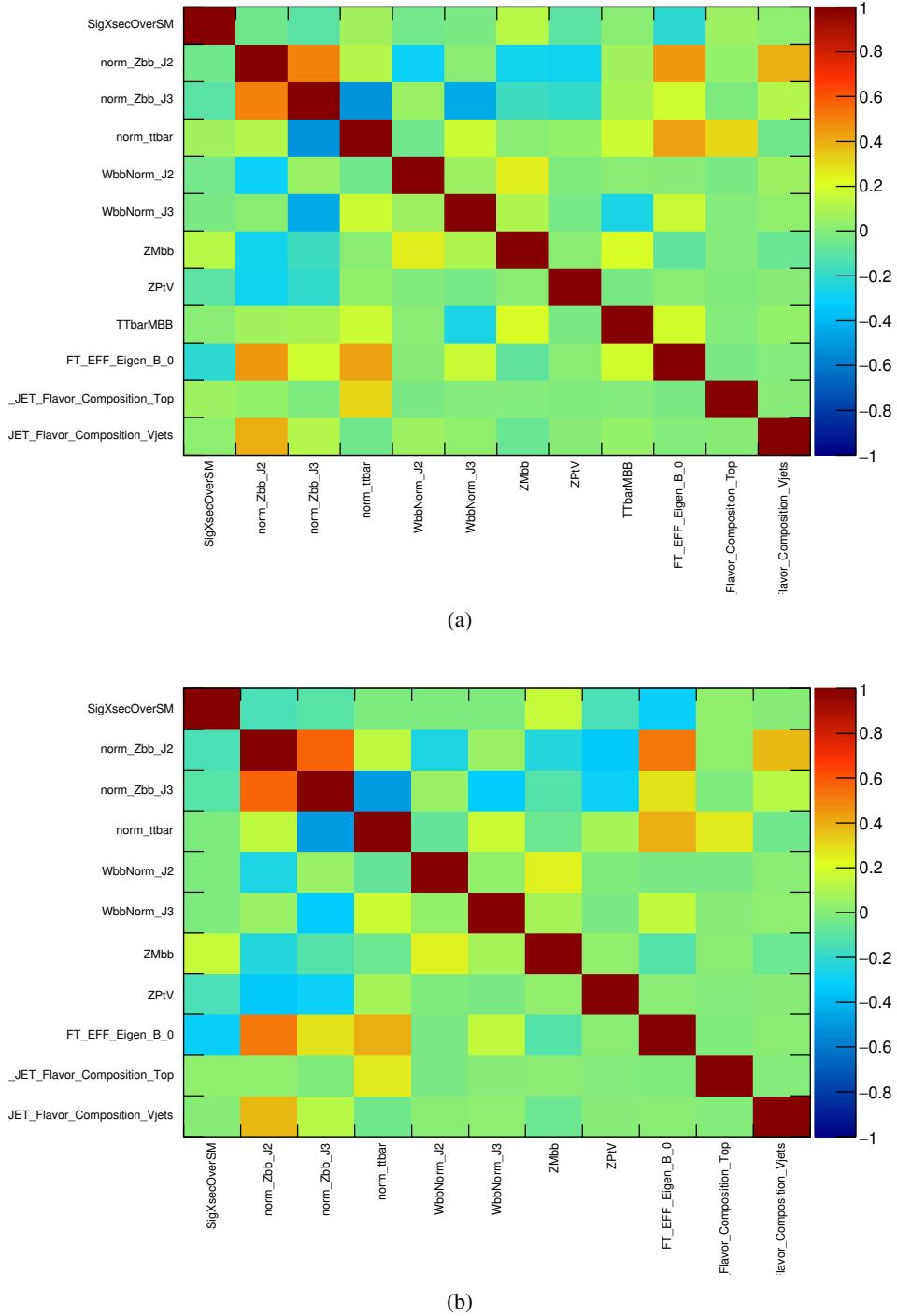


Figure 147: Correlation matrices from the unconditional fit in the 0-lepton channel to an Asimov dataset (a) and to 2015 and 2016 data (b). Only correlations with a magnitude greater than 0.25 are shown.

Not reviewed, for internal circulation only

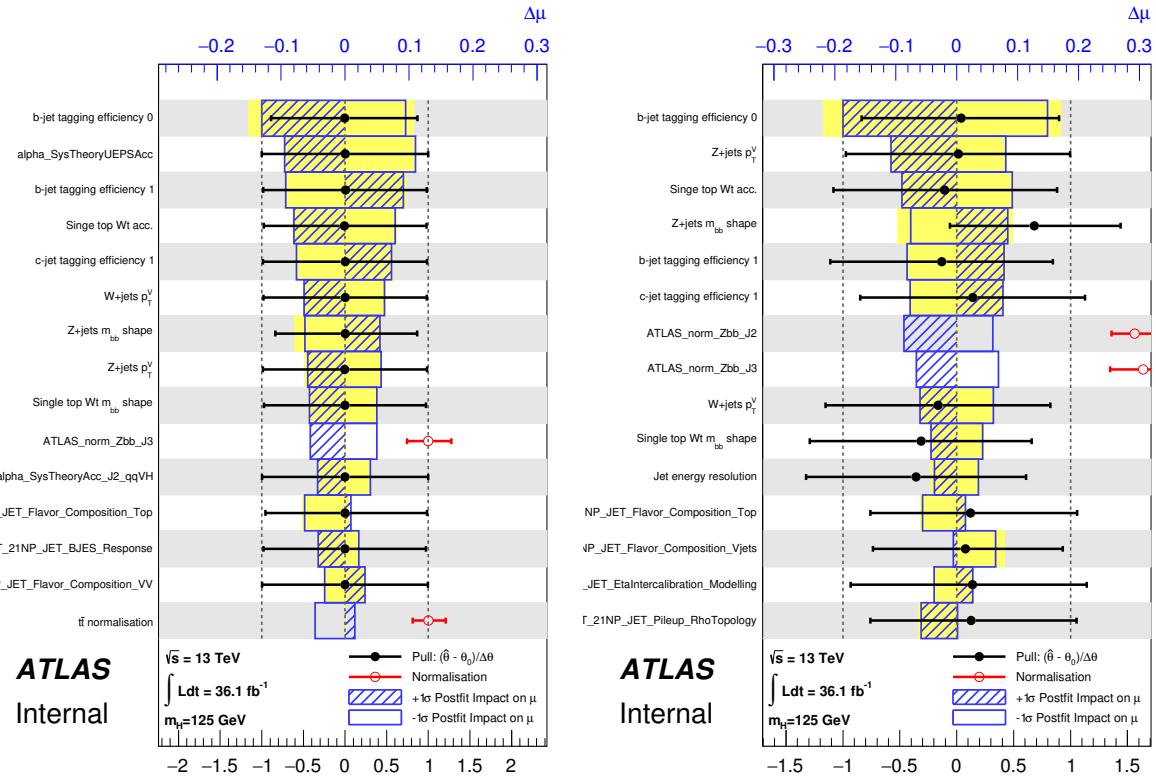


Figure 148: The nuisance parameter ranking based on the impact on μ for the 0-lepton based on prefit (a) and postfit unconditional (b).

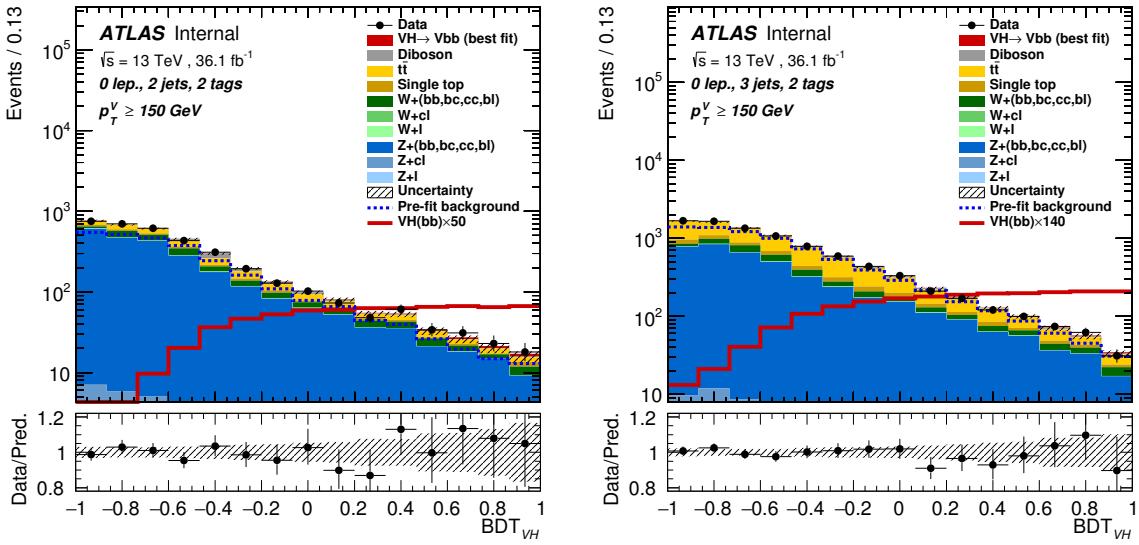


Figure 149: Post-fit distributions obtained in the 0-lepton channel for 2-tag 2-jet SR (a), and for the 2-tag 3-jet SR (b) categories after the unconditional fit on 2015 and 2016 data.

1380 J.2. 1-lepton channel

1381 In this section, the results of the fit, performed in the 1-lepton channel, are shown.

1382 J.2.1. Nuisance parameter pulls and constraints

1383 In Figure 150, the nuisance parameter pulls and constraints are shown for the 1-lepton channels , when
1384 performing an unconditional fit on an Asimov dataset (red) and the 2015 and 2016 data (black). Nuisance
1385 parameters associated to the MC statistical uncertainties in each bin are not shown.

1386 J.2.2. Correlation matrices

1387 The correlation matrices from the unconditional fit in the 1-lepton channels are shown when fitting an
1388 Asimov dataset, and data, in Figure 151. Only the fit parameters with at least a correlation coefficient
1389 with a magnitude greater than 0.25 are shown.

Not reviewed, for internal circulation only

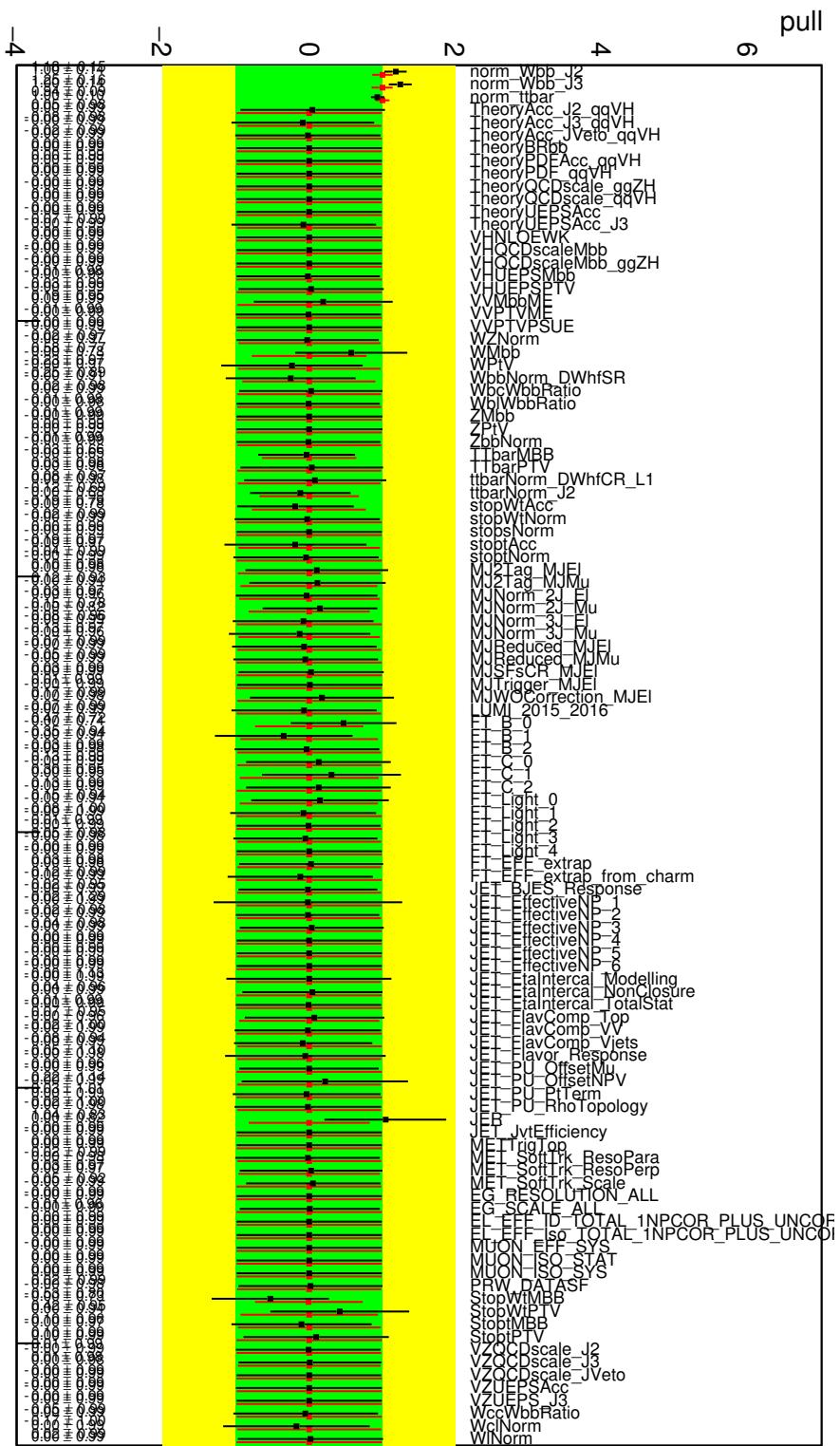


Figure 150: The nuisance parameter pulls and the free parameter scale factors corresponding to the unconditional fits performed on Asimov datasets (red) and 2015 and 2016 data (black), in the 1-lepton channel, are shown.

Not reviewed, for internal circulation only

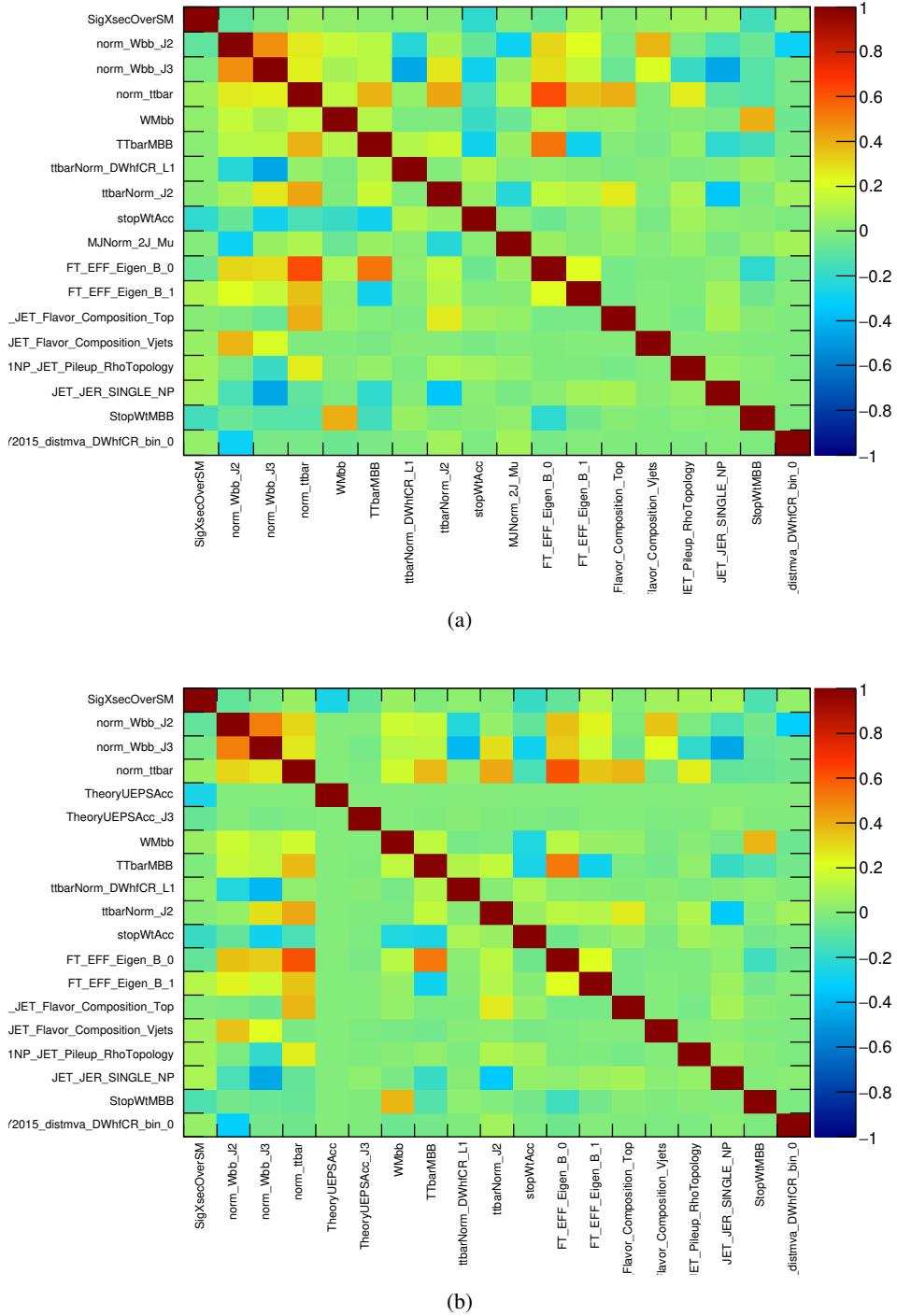


Figure 151: Correlation matrices from the unconditional fit in the 1-lepton channel to an Asimov dataset (a) and to 2015 and 2016 data (b).

1390 **J.2.3. Nuisance parameter ranking**

1391 The ranking of the most important systematics is shown, for the 1-lepton fit, in Figure 152. For the pre-fit
 1392 ranking, the value of the prior is used, which is equivalent for the nuisance parameter to $\sigma = 1$. For the
 1393 post-fit ranking, the error on the profiled nuisance parameter at the maximum of the likelihood is used.

Not reviewed, for internal circulation only

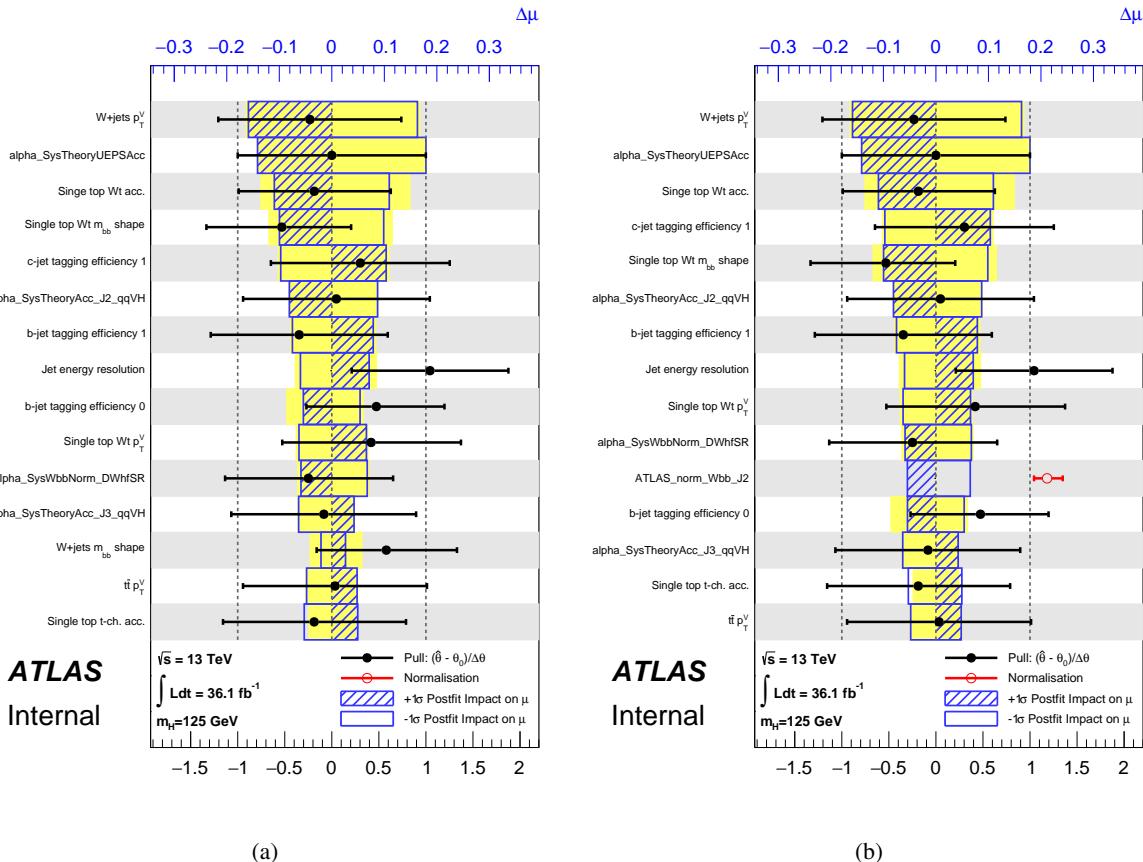


Figure 152: The post-fit impact nuisance parameter ranking for the 1-lepton unconditional fit based on prefit (a) and postfit (b) impact on $\hat{\mu}$.

1394 **J.2.4. Post-fit distributions**

1395 In Figure 153, the postfit BDT distributions in the 1-lepton channel is shown. These are obtained after an
 1396 unconditional fit is performed in each channel separately, on 2015 and 2016 data.

1397 **J.2.5. Breakdown of statistical and systematic uncertainties**

1398 The fitted parameter of interest $\hat{\mu}$ and the breakdown of the uncertainties coming from data statistics (“Data
 1399 Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only
 1400 (“MC Stat.”), and each category of systematic uncertainty, is shown for the 1-lepton unconditional fit on
 1401 36.1 fb^{-1} of data in Table 53.

Not reviewed, for internal circulation only

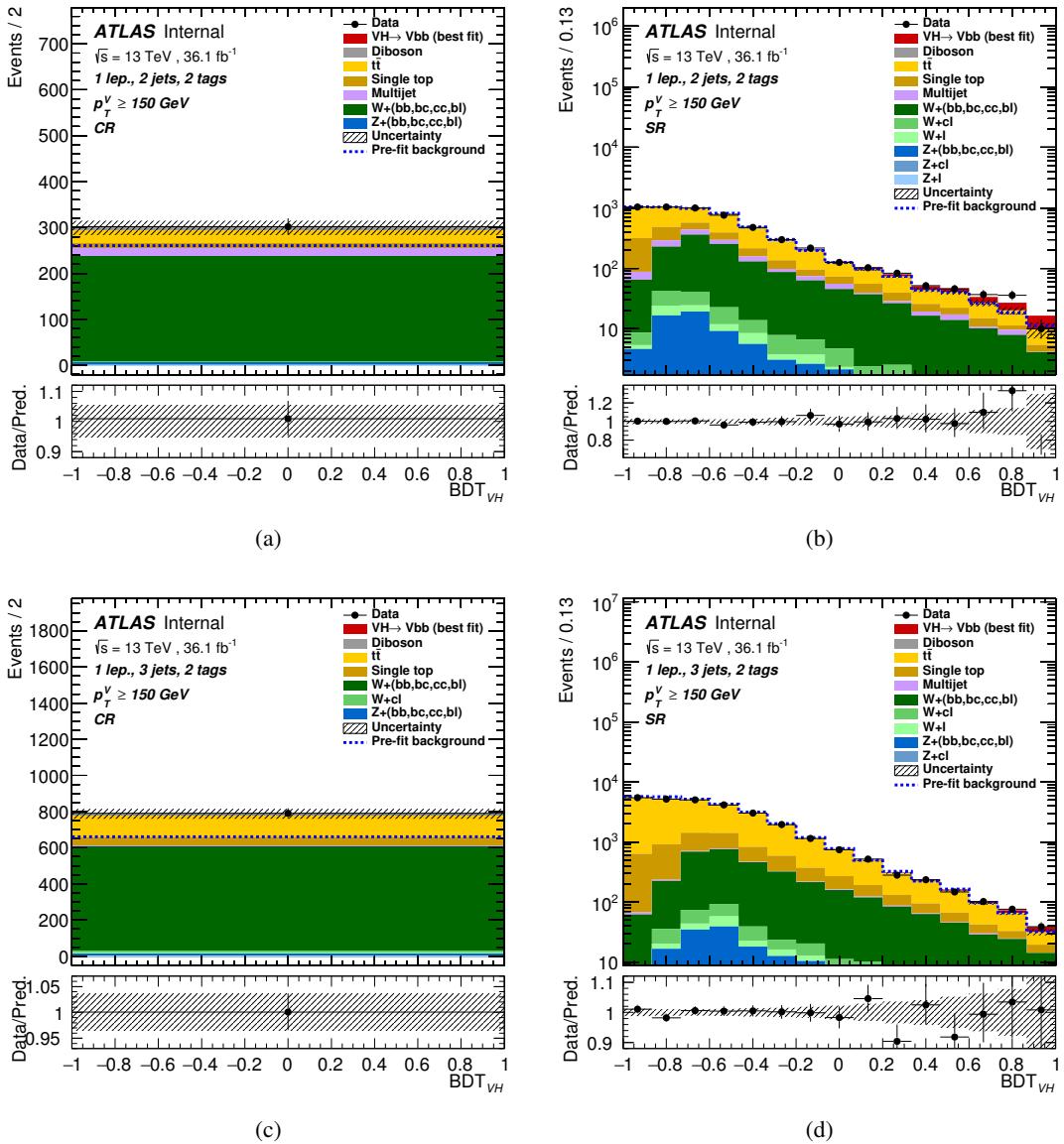


Figure 153: Post-fit distributions obtained in the 1-lepton channel for 2-tag 2-jet, $W + hf$ CR (a), SR (b), and for the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories with an unconditional fit on 2015 and 2016 data.

Set of nuisance parameters	Impact on error
Total	+0.690 / -0.604
DataStat	+0.400 / -0.386
FullSyst	+0.562 / -0.465
Floating normalizations	+0.087 / -0.077
All normalizations	+0.126 / -0.110
All but normalizations	+0.531 / -0.429
Jets MET	+0.152 / -0.103
BTag	+0.162 / -0.138
Leptons	+0.010 / -0.005
Luminosity	+0.057 / -0.021
Diboson	+0.090 / -0.051
Zjets	+0.004 / -0.005
Wjets	+0.178 / -0.170
Model ttbar	+0.074 / -0.048
Model Single Top	+0.179 / -0.152
Model Multi Jet	+0.052 / -0.051
Signal Systematics	+0.311 / -0.117
MC stat	+0.255 / -0.253

Table 53: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and different sub-groups of systematic uncertainties, is shown, as well as the fitted $\hat{\mu}$ extracted from the 1-lepton unconditional fit on 36.07 fb^{-1} of data.

J.3. 2-lepton channel

1402 In this section, the results of the fit, performed in the 2-lepton channel, are shown.

J.3.1. Nuisance parameter pulls and constraints

1403 In Figure 154, the nuisance parameter pulls and constraints are shown for the 2-lepton channels , when
1404 performing an unconditional fit on an Asimov dataset (red) and the 2015 and 2016 data (black). Nuisance
1405 parameters associated to the MC statistical uncertainties in each bin are not shown.

J.3.2. Correlation matrices

1406 The correlation matrices from the unconditional fit ($\mu = \hat{\mu}$) in the 2-lepton channels are shown when the
1407 unconditional fit to an Asimov dataset, and data, in Figure 155. Only the fit parameters with at least a
1408 correlation coefficient with a magnitude greater than 0.25 are shown.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

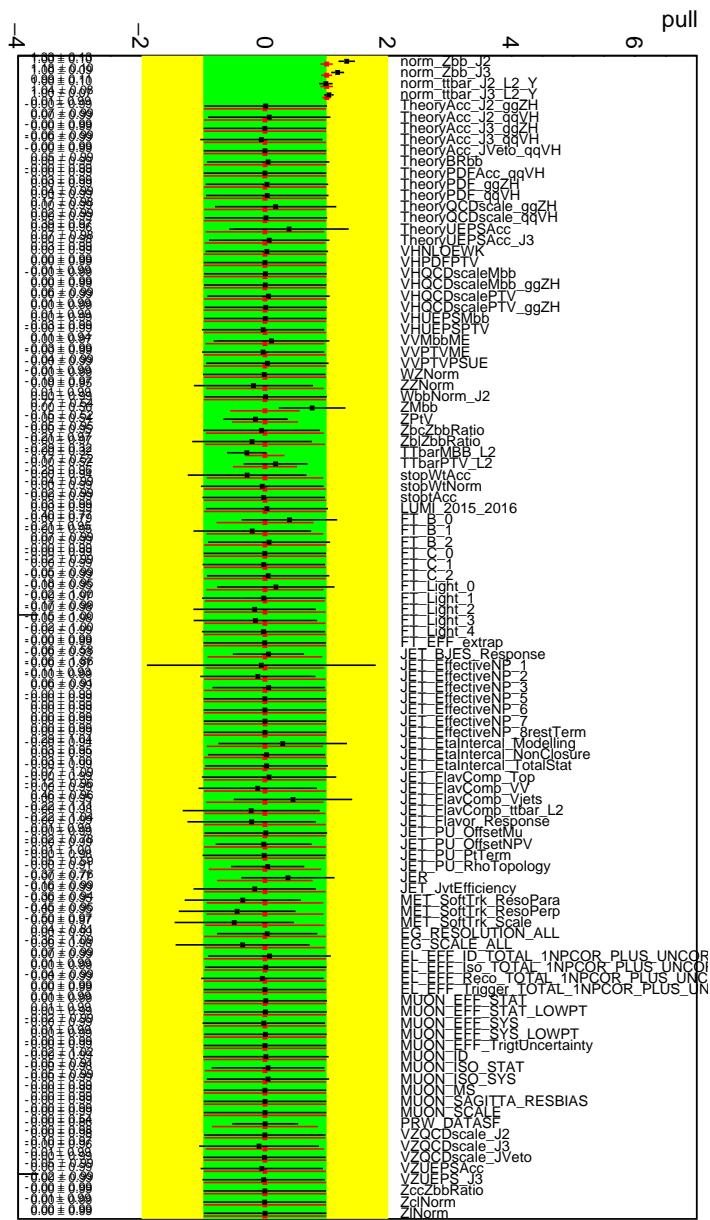


Figure 154: The nuisance parameter pulls and the free parameter scale factors corresponding to the unconditional fits performed on Asimov datasets generated with $\hat{\mu}$ (red) and 2015 and 2016 data (black), in the 2-lepton channel, are shown.

Not reviewed, for internal circulation only

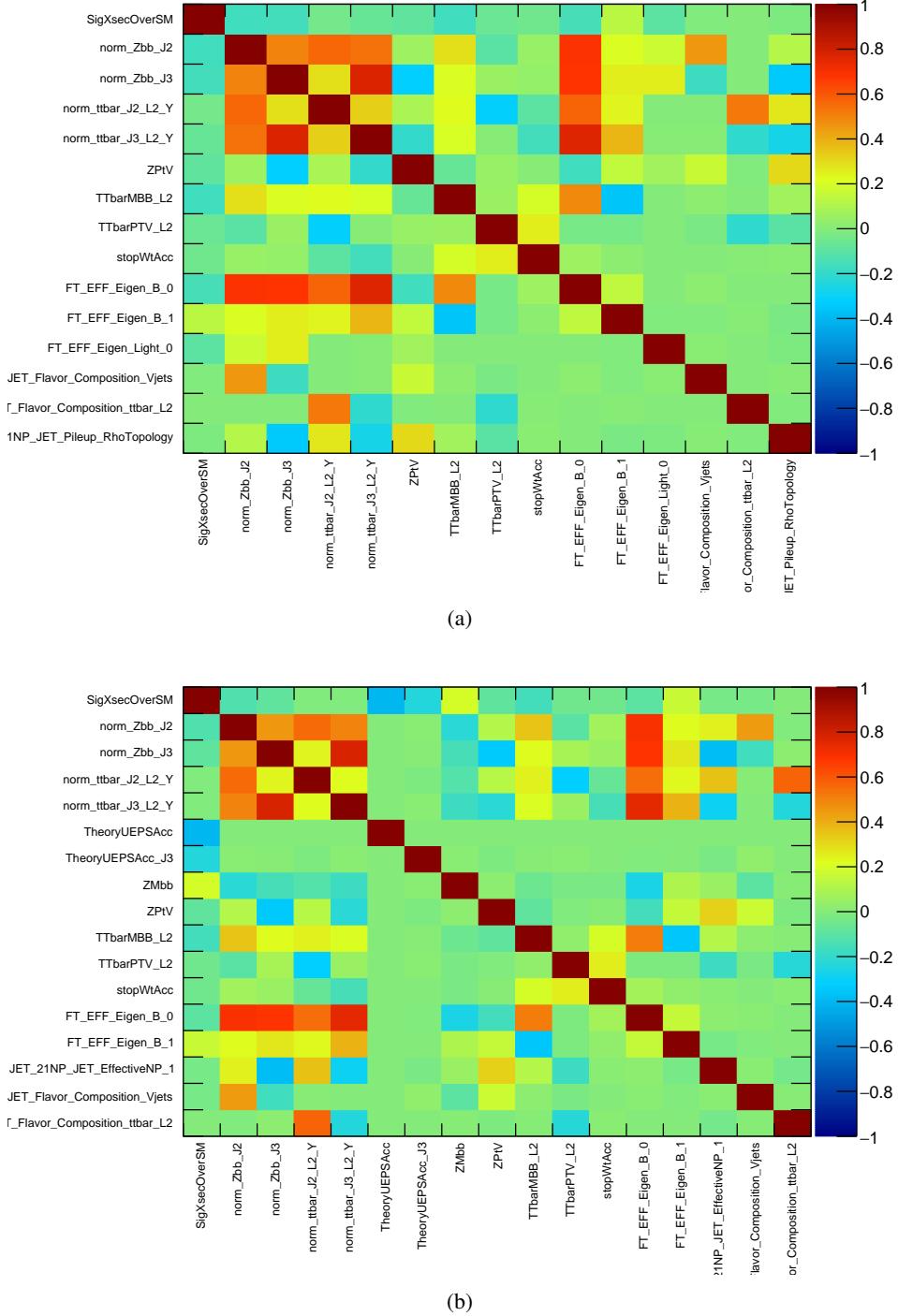


Figure 155: Correlation matrices from the unconditional fit in the 2-lepton channel to an Asimov dataset (a) and to 2015 and 2016 data (b). Matrices show only high correlation with a magnitude greater than 0.25.

1412 **J.3.3. Nuisance parameter ranking**

1413 The ranking of the most important systematics is shown, for the 2-lepton fit, in Figure 156.

Not reviewed, for internal circulation only

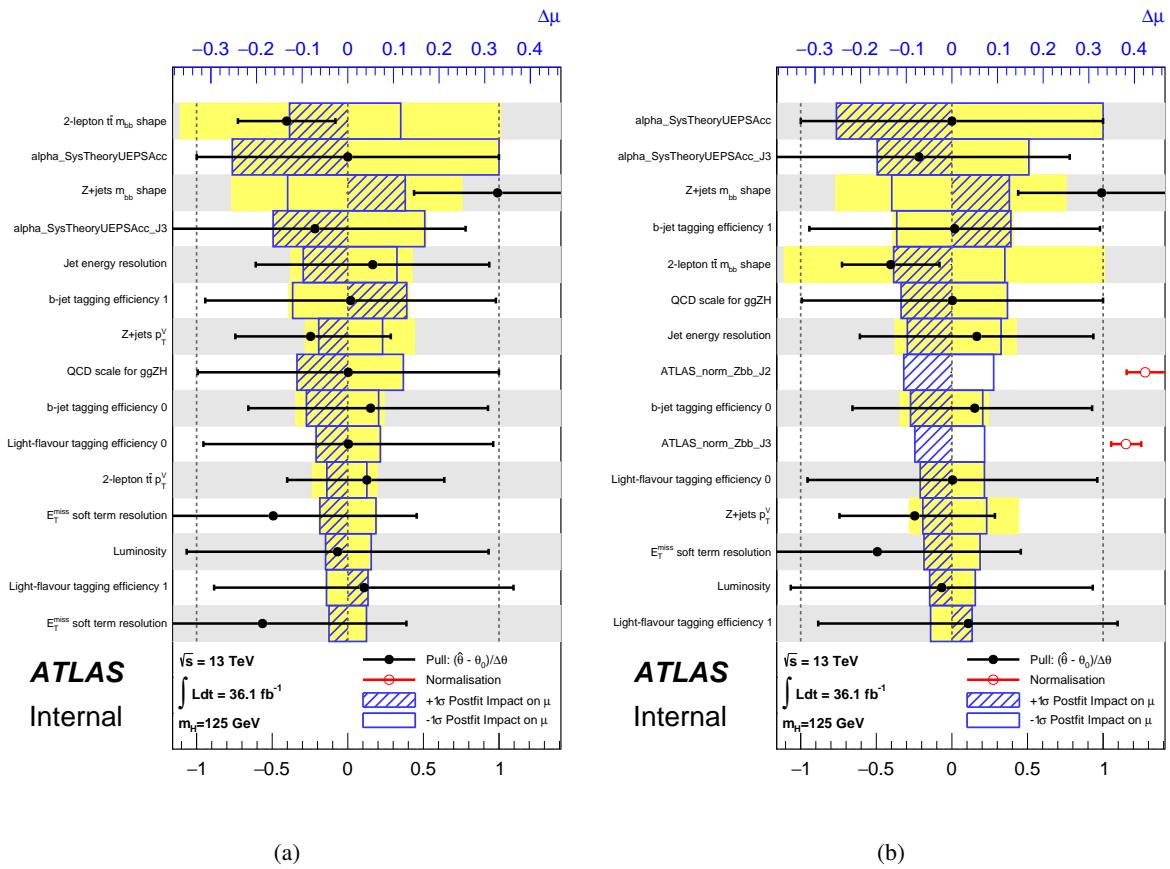


Figure 156: The nuisance parameter ranking for the 2-lepton channel unconditional fit (a) ordered by prefit impact (b) ordered by postfit impact.

1414 **J.3.4. Post-fit distributions**

1415 In Figure 157-158 the postfit BDT distributions in the 2-lepton channel, are shown. Figures 159-160 show
 1416 the postfit m_{bb} distributions in the 2-lepton $e\mu$ control region. These are obtained after an unconditional
 1417 fit is performed in each signal region separately, on 2015 and 2016 data.

1418 **J.3.5. Breakdown of statistical and systematic uncertainties**

1419 The fitted parameter of interest $\hat{\mu}$ and the breakdown of the uncertainties coming from data statistics (“Data
 1420 Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only
 1421 (“MC Stat.”), and each category of systematic uncertainty, is shown for the 2-lepton unconditional $\mu = \hat{\mu}$
 1422 fit on 36.1 fb^{-1} of data in Table 54.

Not reviewed, for internal circulation only

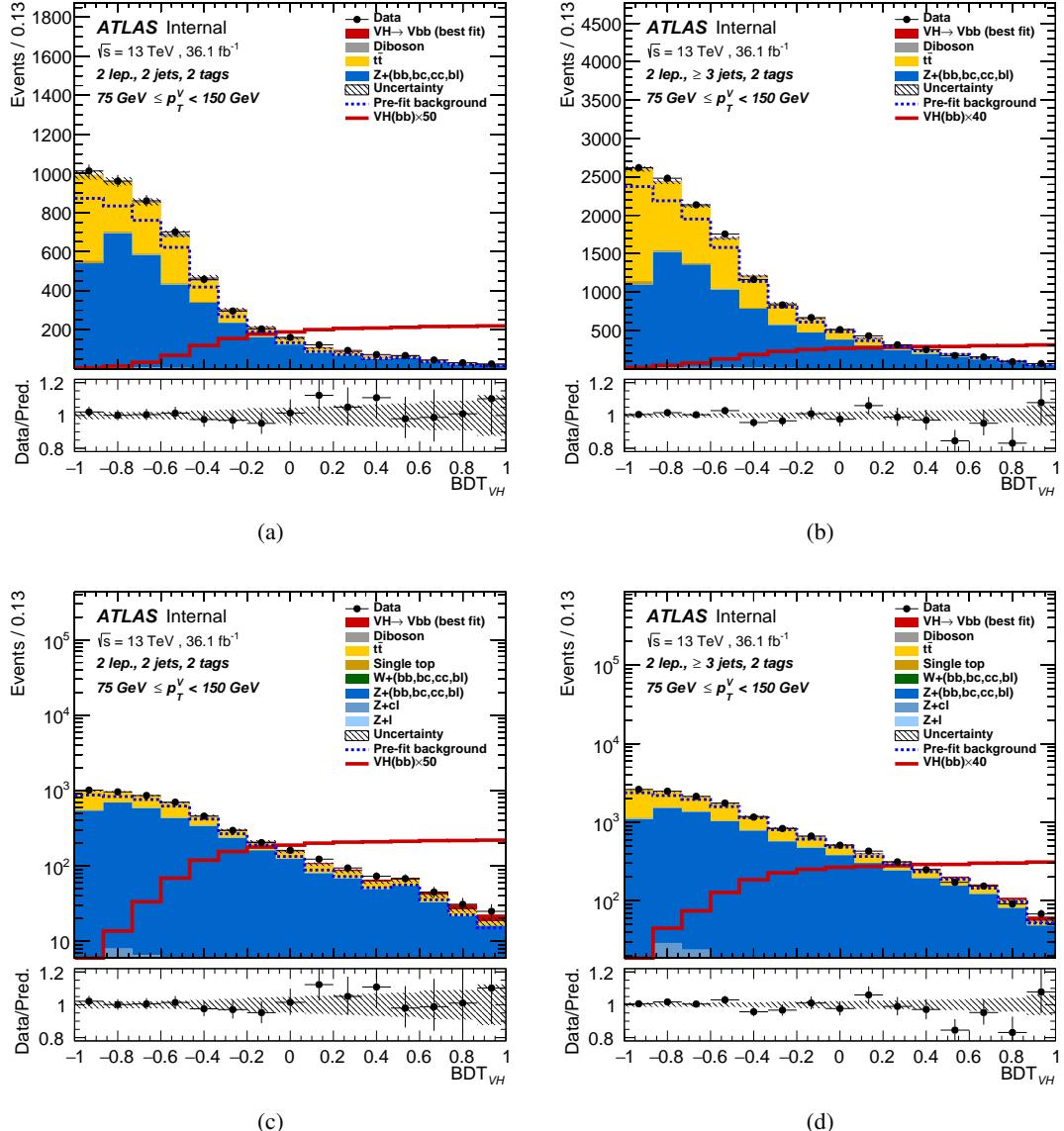


Figure 157: Post-fit BDT output distributions obtained in the $75 < p_T^Z < 150$ 2-lepton channel, for 2-tag 2-jet SR (a), 2-tag 3+jet SR (b) with an unconditional fit on 2015 and 2016 data. Each log scaled plot is also shown in (c) and (d), respectively.

Not reviewed, for internal circulation only

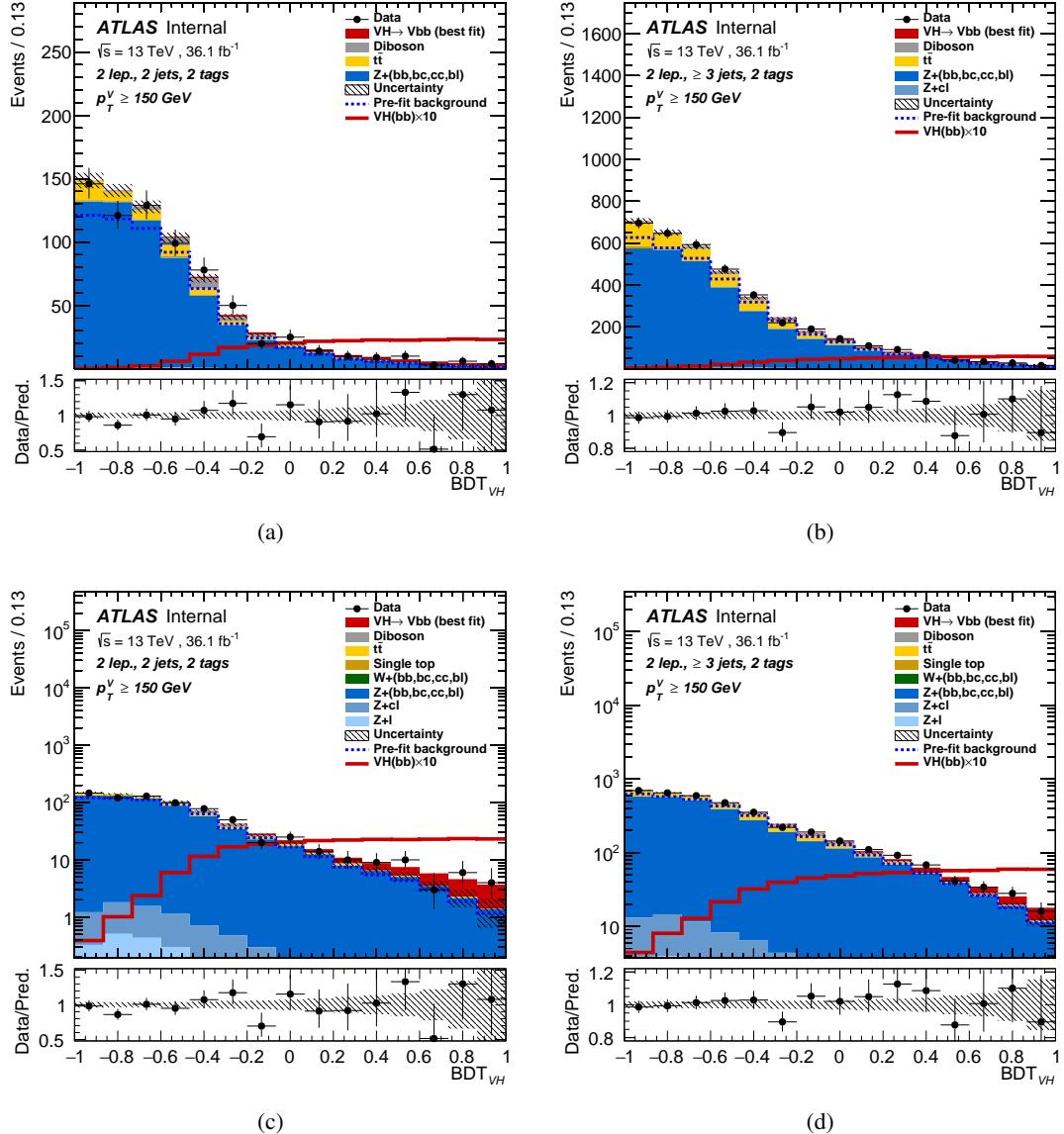


Figure 158: Post-fit BDT output distributions obtained in the $150 < p_T^Z$ 2-lepton channel, for 2-tag 2-jet SR (a), 2-tag 3+jet SR (b) with an unconditional fit on 2015 and 2016 data. Each log scaled plot is also shown in (c) and (d), respectively.

Not reviewed, for internal circulation only

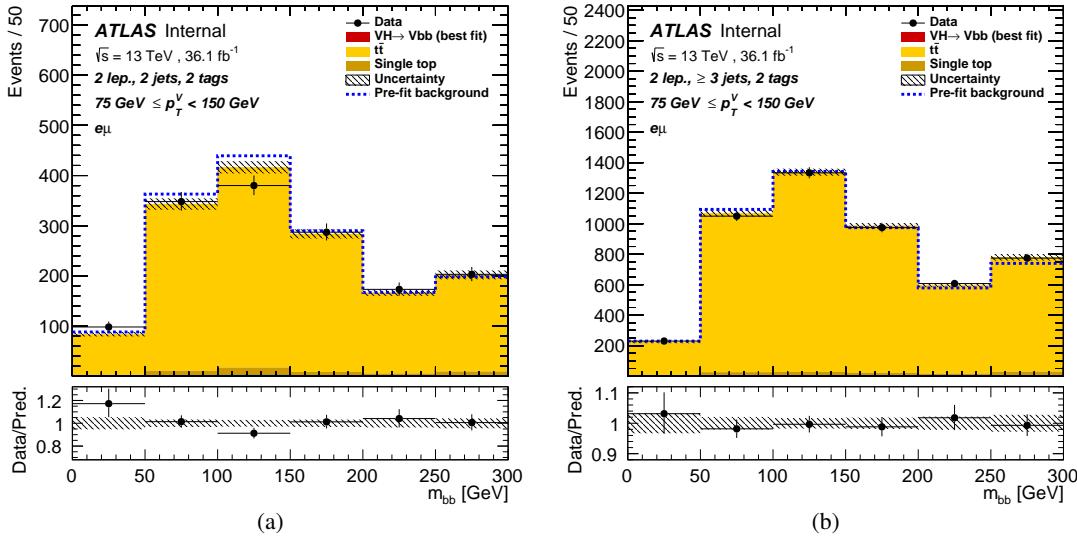


Figure 159: The post-fit $75 \text{ GeV} < p_T^Z < 150 \text{ GeV}$ m_{bb} distributions for an unconditional fit in the 2-lepton channel for the 2-tag 2-jet $e\mu$ CR (a), and for the 2-tag 3+jet $e\mu$ CR (b) categories, are shown.

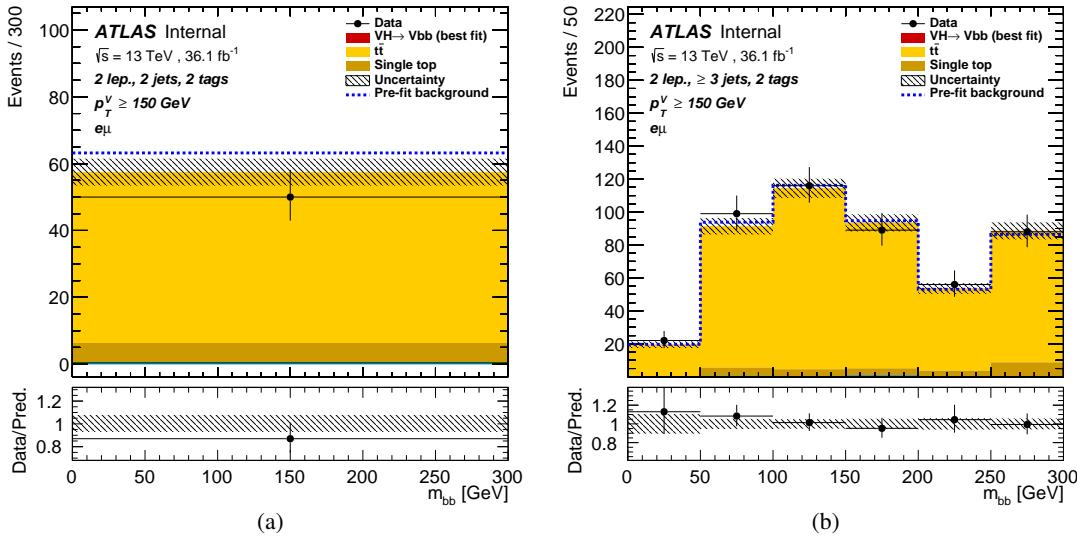


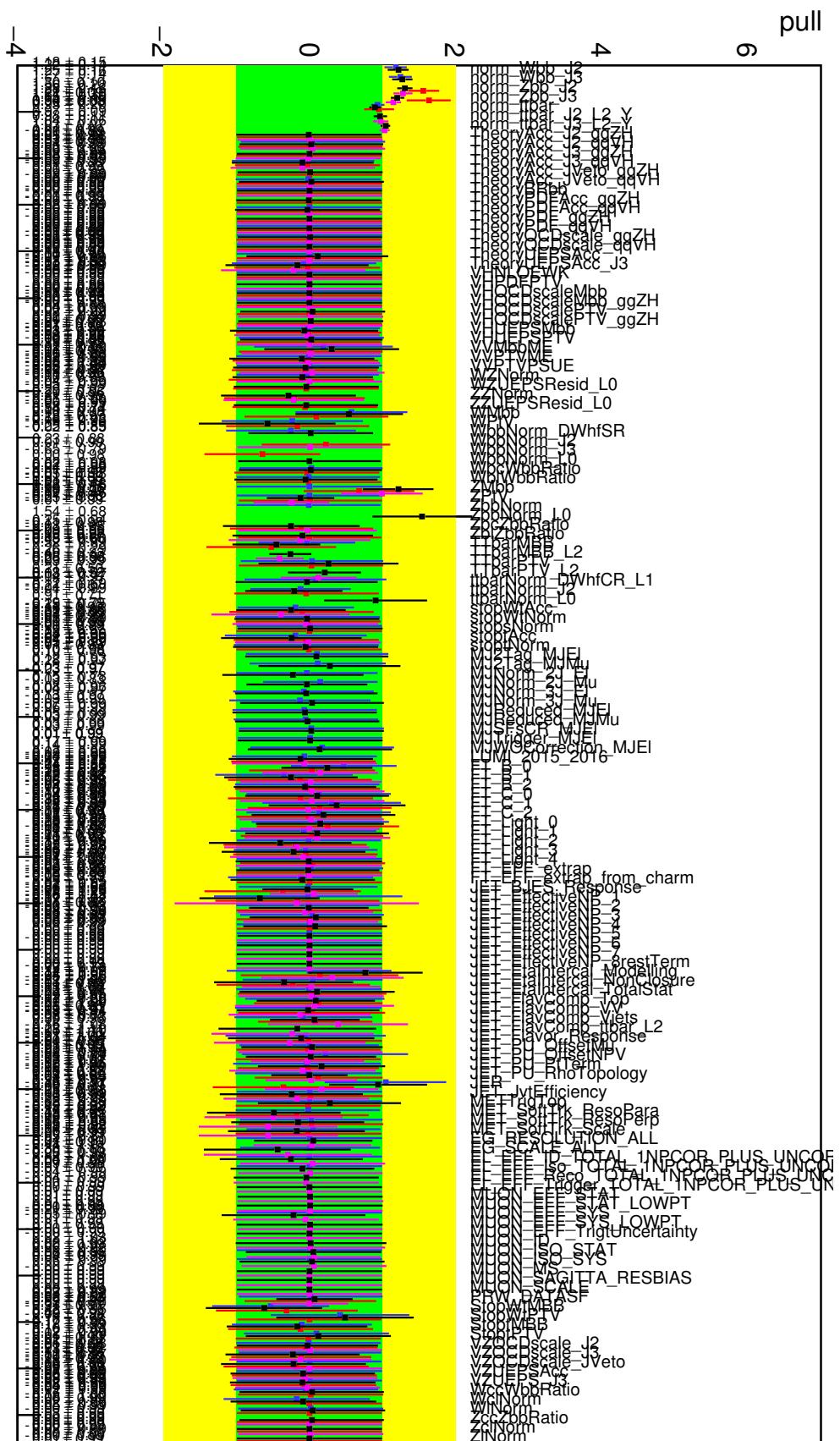
Figure 160: The post-fit $p_T^Z > 150 \text{ GeV}$ m_{bb} distributions for an unconditional $\mu = 1$ fit in the 2-lepton channel for the 2-tag 2-jet $e\mu$ CR (a), and for the 2-tag 3+jet $e\mu$ CR (b) categories, are shown.

Set of nuisance parameters	Impact on error
Total	+0.816 / -0.672
DataStat	+0.500 / -0.482
FullSyst	+0.645 / -0.468
Floating normalizations	+0.156 / -0.151
All normalizations	+0.157 / -0.151
All but normalizations	+0.612 / -0.427
Jets MET	+0.157 / -0.135
BTag	+0.199 / -0.171
Leptons	+0.039 / -0.026
Luminosity	+0.088 / -0.049
Diboson	+0.049 / -0.035
Zjets	+0.175 / -0.182
Wjets	+0.000 / -0.000
Model ttbar	+0.132 / -0.120
Model Single Top	+0.032 / -0.036
Signal Systematics	+0.498 / -0.262
MC stat	+0.174 / -0.157

Table 54: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and different sub-groups of systematic uncertainties, is shown, as well as the fitted $\hat{\mu}$ extracted from the 2-lepton unconditional fit on the data.

J.4. Nuisance parameter pulls and constraints in all channels

1423 In Figure 161, the nuisance parameter pulls and constraints are shown for performed on the unconditional
1424 fit to 2015+2016 data for the combined fit on all regions (black), 0-lepton fit (red), 1-lepton fit (blue) and
1425 2-lepton fit (pink). In Figures 162 to 163, the same pulls and constraints are shown in more detail. MC
1426 statistical uncertainties, normalisation and shape modelling systematic uncertainties, and experimental
1427 systematic uncertainties are considered in the fit.
1428



Not reviewed, for internal circulation only

Figure 161: The nuisance parameter pulls and the free parameter scale factors corresponding to the combined fits performed on 2015+2016 data generated for the combined unconditional fit on all regions (black), 0-lepton fit (red), 23rd June 2017 and 135 lepton fit (pink). 195

Not reviewed, for internal circulation only

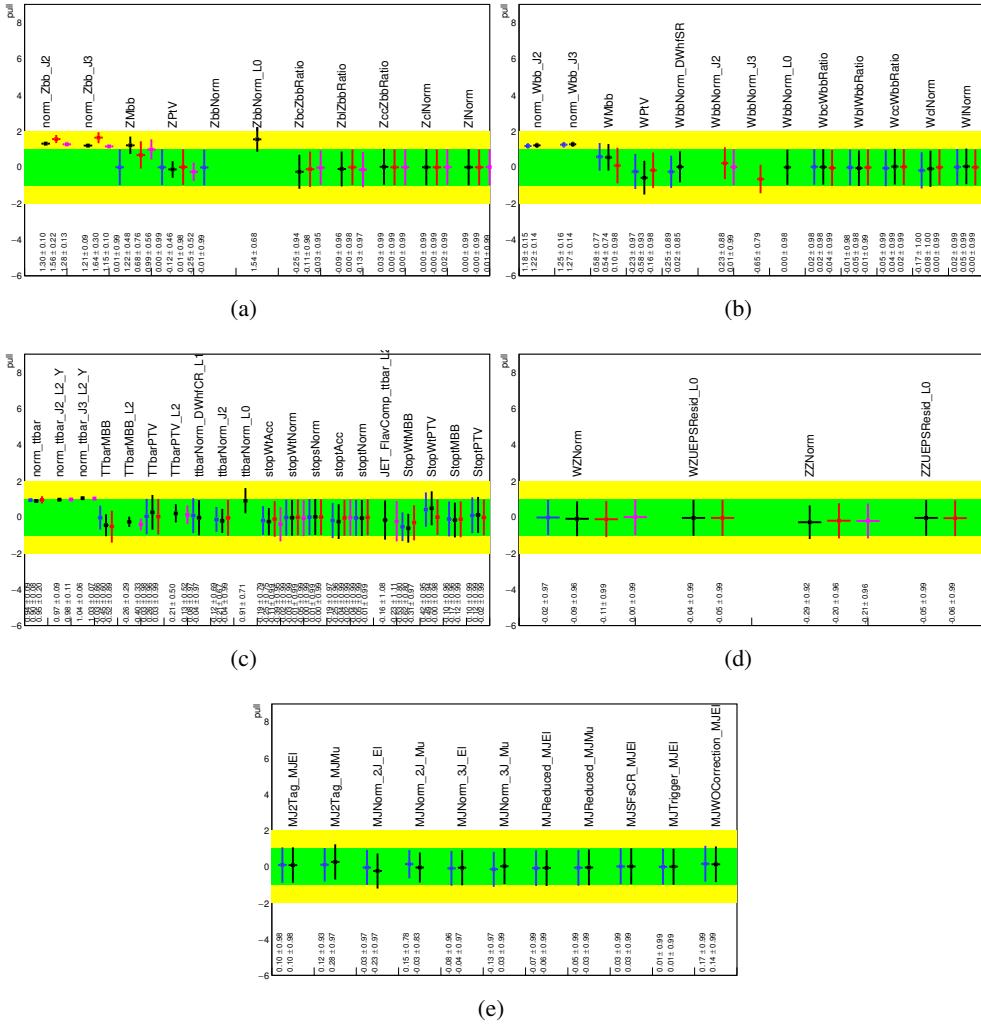


Figure 162: Pull distributions of modeling NPs in the conditional $\mu = \text{fit to } 2015+2016$ data, for the combined unconditional fit on all regions (black), 0-lepton fit (red), 1-lepton fit (blue) and 2-lepton fit (pink). Z jets (a), W jets (b), top (c), diboson (d) and MJ (e).

Not reviewed, for internal circulation only

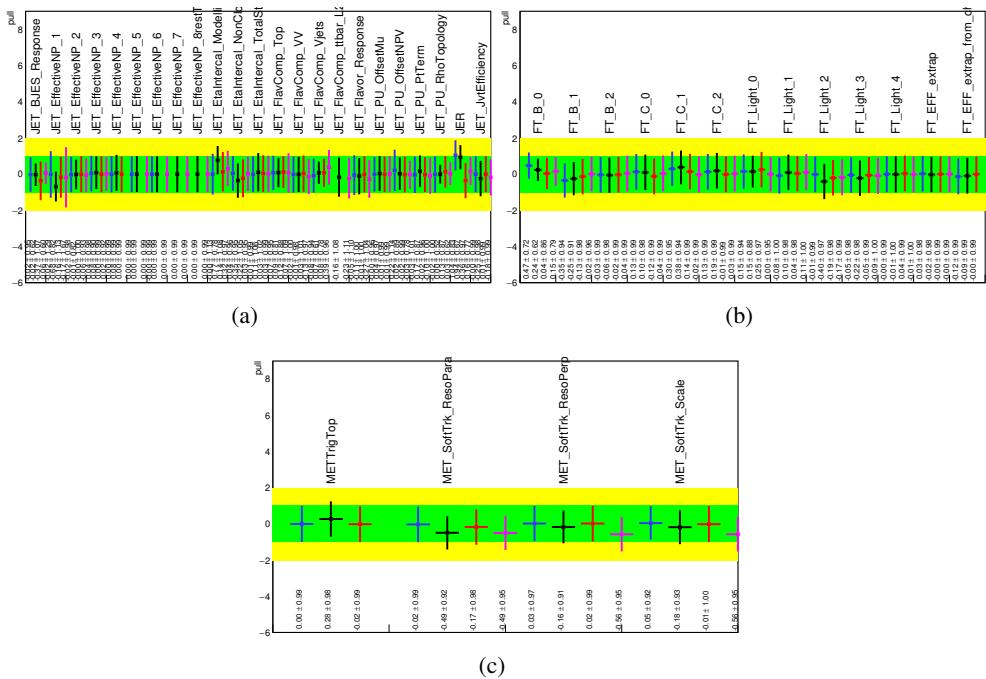


Figure 163: Pull distributions of experimental NPs in the conditional $\mu = \text{fit to 2015+2016 data}$, for the combined unconditional fit on all regions (black), 0-lepton fit (red), 1-lepton fit (blue) and 2-lepton fit (pink). Jet systematics (a), b-tagging (b) and E_T^{miss} (c).

¹⁴²⁹ **J.5. Detailed fit significances**

¹⁴³⁰ Detailed fit significances are provided in Table 55, detailing the VH significance in each fitting region.

Channel	n_{jet}	p_T^V [GeV]	Expected sig. (Asimov)	Expected sig. (data)	Observed sig. (data)
0-lepton(SR)	2	> 150	1.81	1.55	0.83
	3	> 150	0.94	0.89	-0.37
1-lepton(SR)	2	> 150	1.65	1.64	2.00
	3	> 150	0.83	0.78	0.57
2-lepton(SR+CR)	2	75 – 150	0.66	0.58	1.94
		> 150	1.40	1.31	2.77
	3p	75 – 150	0.59	0.52	-0.65
		> 150	1.20	1.16	2.03
0-lepton (SR)			1.99	1.73	0.53
1-lepton (SR)			1.81	1.81	2.30
2-lepton (SR+CR)			1.95	1.86	3.55
0,1,2-lepton (SR+CR)			3.19	3.03	3.54

Table 55: Expected significances from a fit to an Asimov dataset corresponding to 36.1 fb^{-1} and from a fit to the data. Expected significances from individual regions are estimated separately.

1431 K. Post-fit BDT input Distributions

1432 In this section the nominal results obtained using the BDT analysis combining 2015 and 2016 data for
 1433 each leptonic channel are presented. The post-fit BDT outputs and some key kinematic distributions are
 1434 shown for each channel from the combined unconditional fit to data. BDT input variables are broken
 1435 down by channel. Sections K.1, K.2 and K.3, present BDT input variables from the 0-, 1-, and 2-lepton
 1436 channels, respectively.

1437 **K.1. 0-lepton**

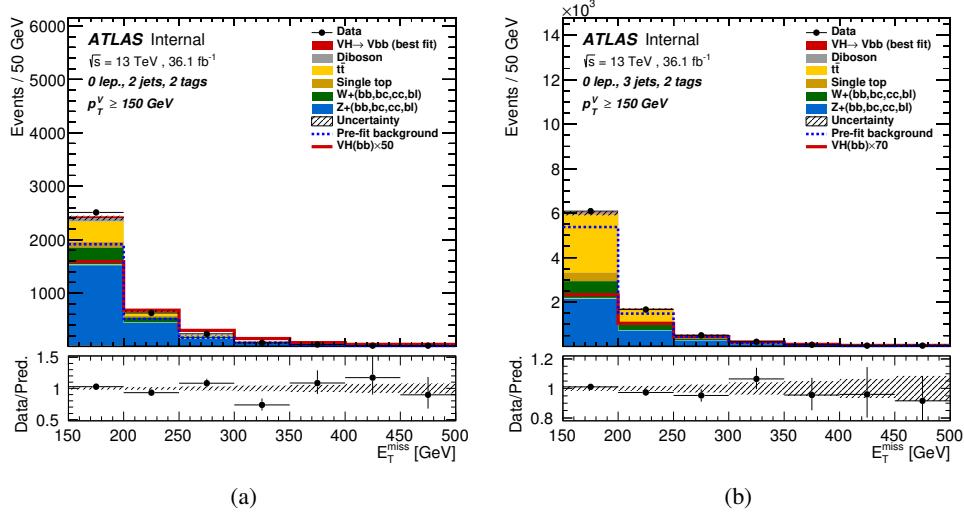


Figure 164: E_T^{miss} postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

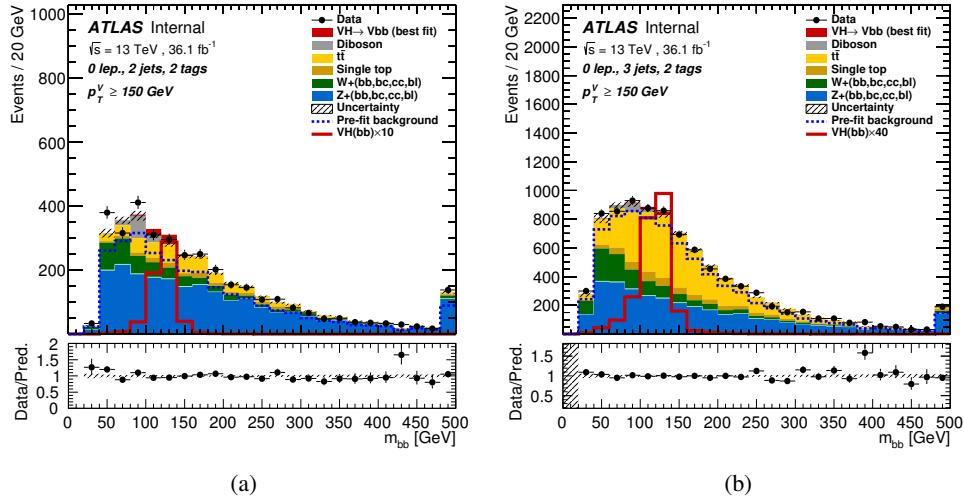


Figure 165: m_{bb} postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

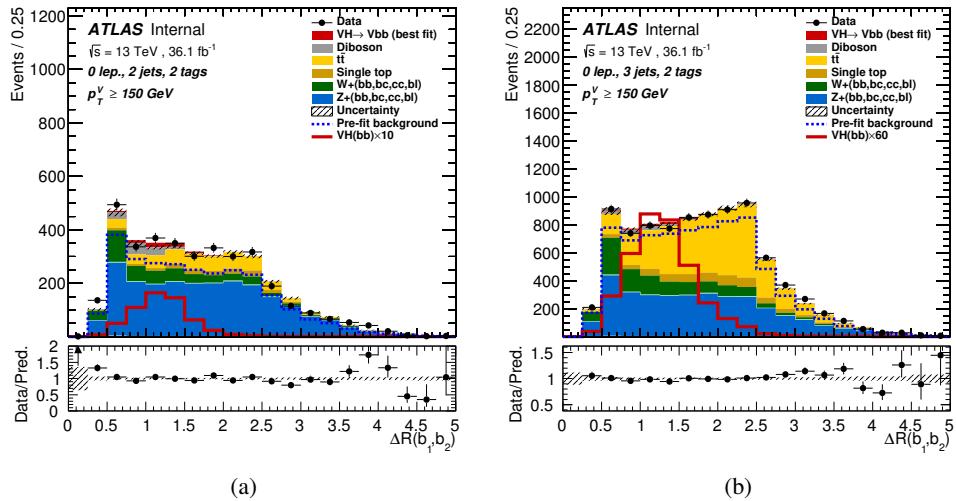


Figure 166: $\Delta R(b, b)$ postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

Not reviewed, for internal circulation only

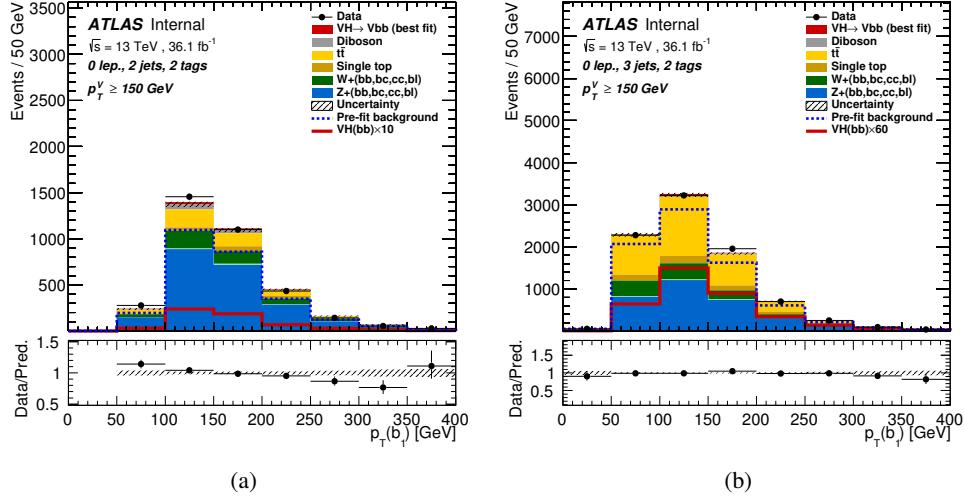


Figure 167: p_T (jet 1) postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

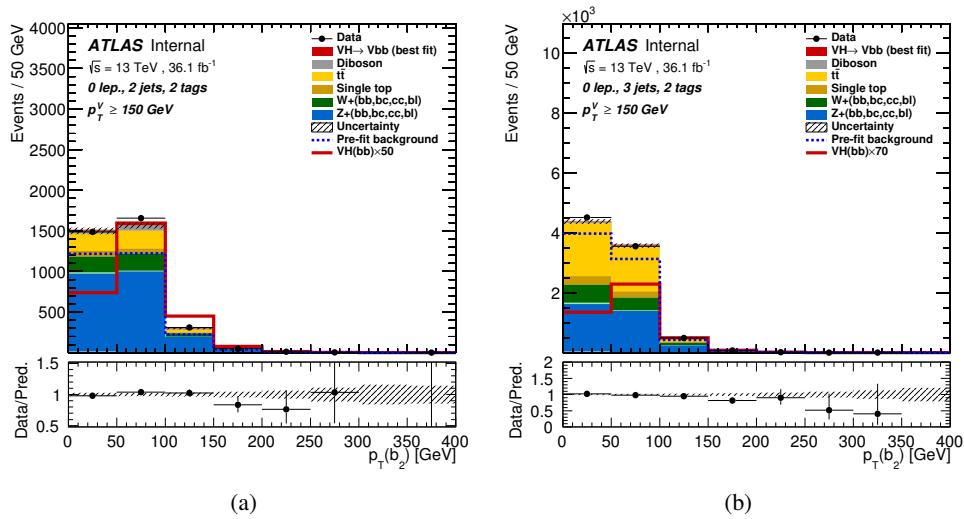


Figure 168: p_T (jet 2) postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

Not reviewed, for internal circulation only

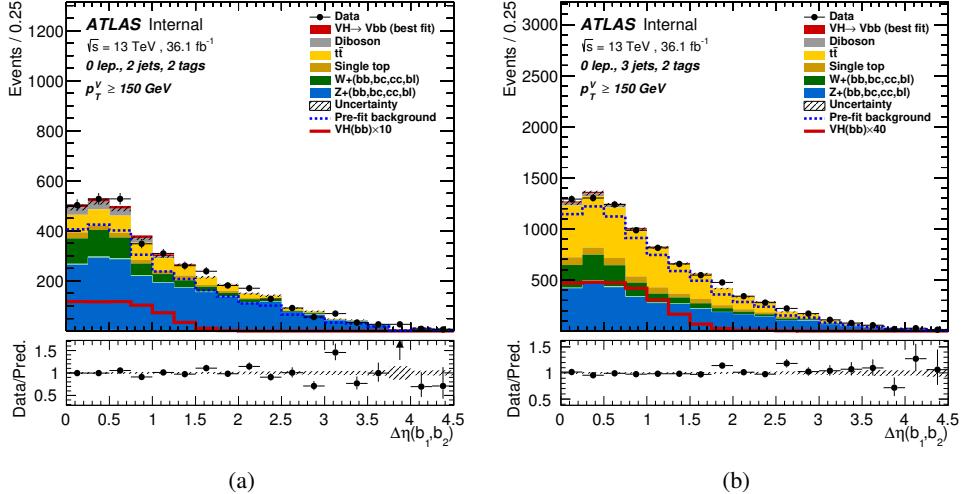


Figure 169: $\Delta\eta(b, b)$ postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

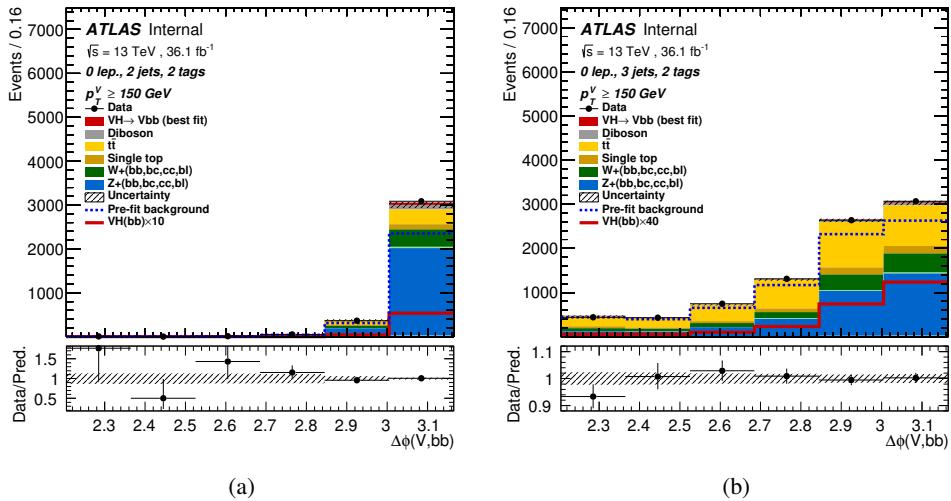


Figure 170: $\Delta\phi(V, H)$ postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

Not reviewed, for internal circulation only

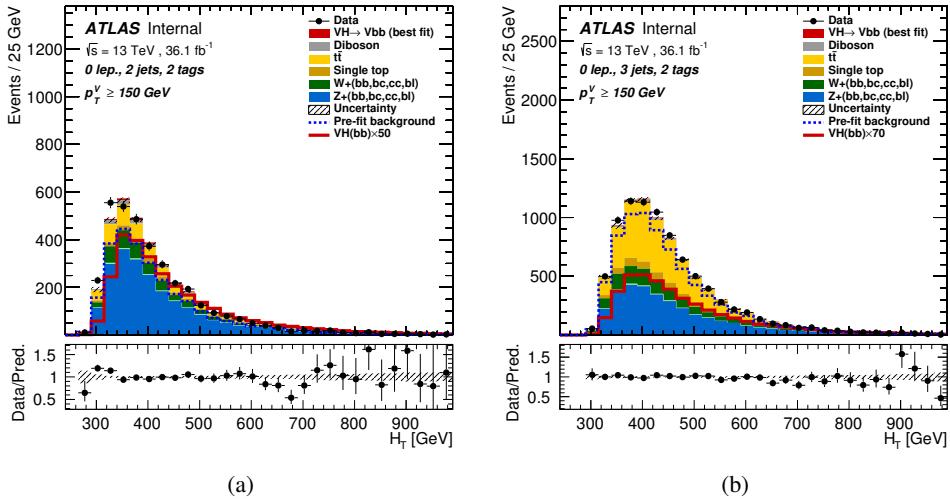


Figure 171: M_{eff} postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a) and 2-tag 3-jet (b) regions.

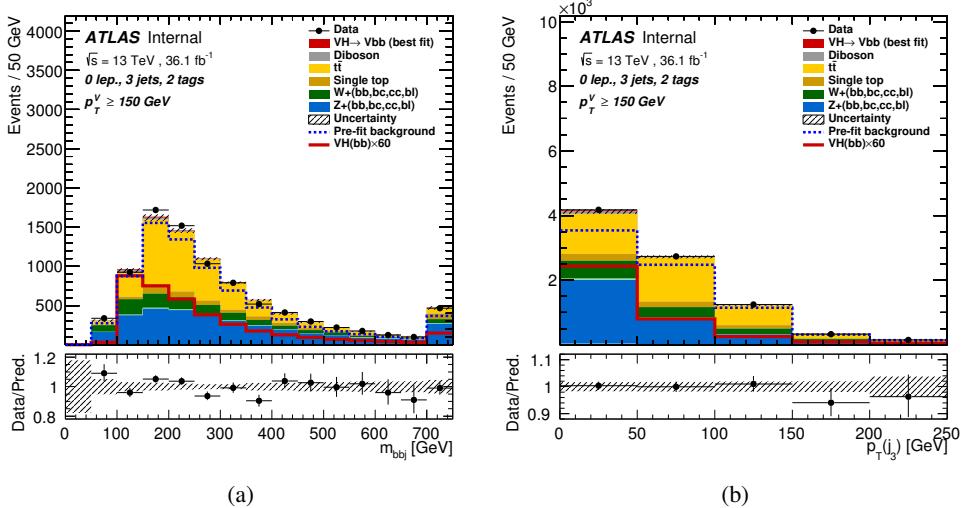


Figure 172: m_{bbj} postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 3-jet region. p_T (jet 3) postfit distributions obtained in the 0-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 3-jet region.

1438 K.2. 1-lepton

Not reviewed, for internal circulation only

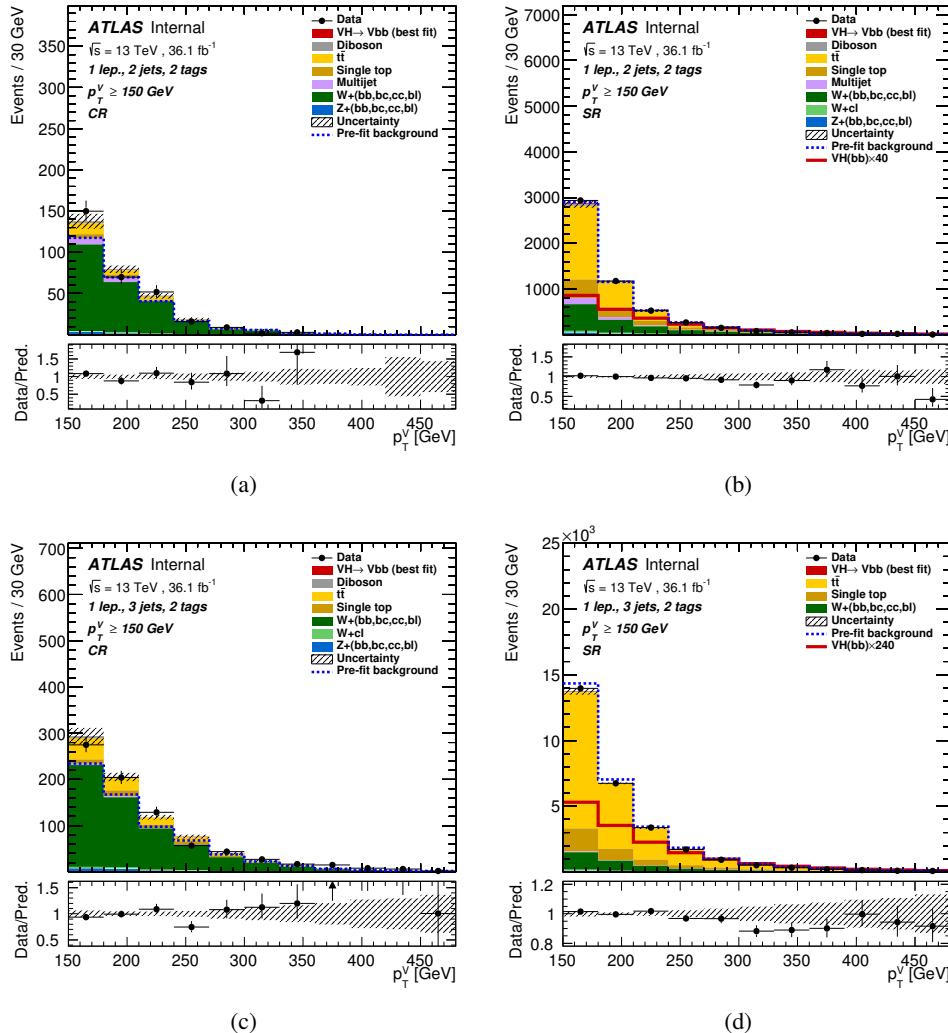


Figure 173: p_T^V postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

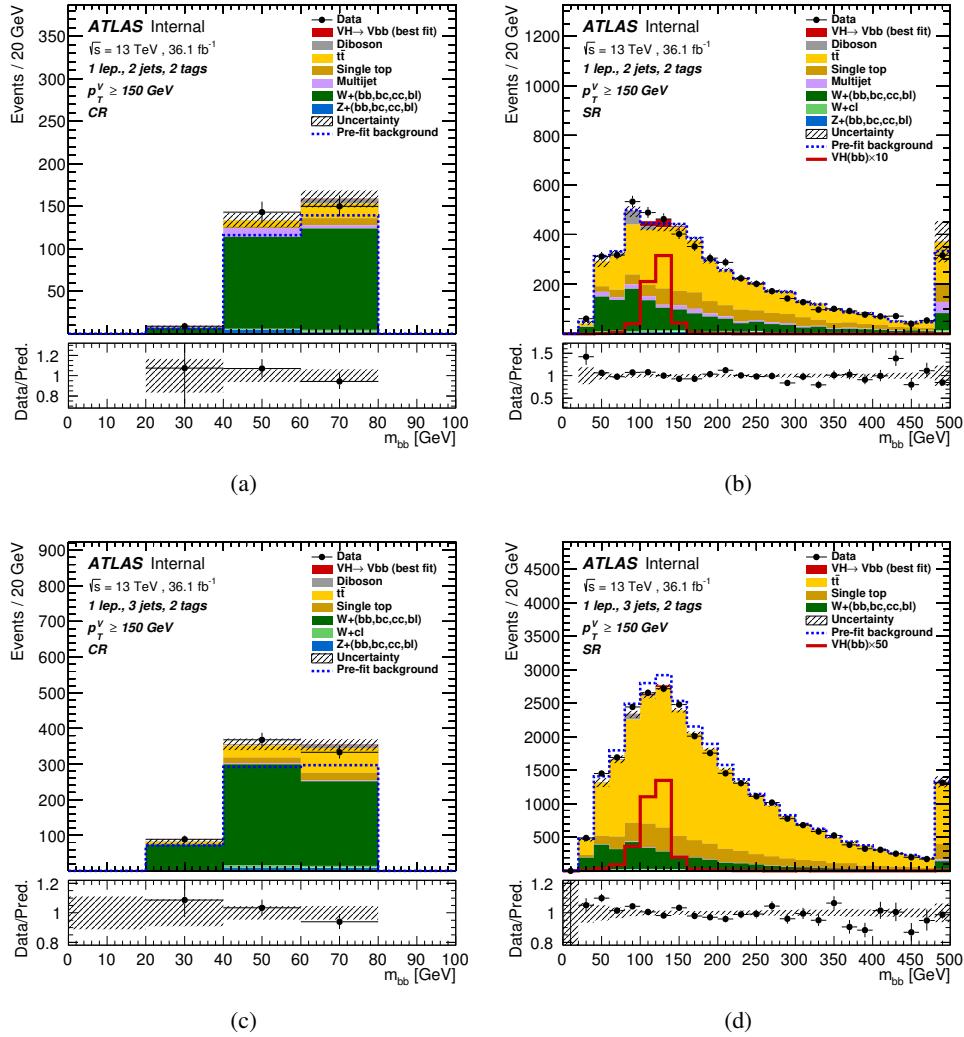


Figure 174: m_{bb} postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

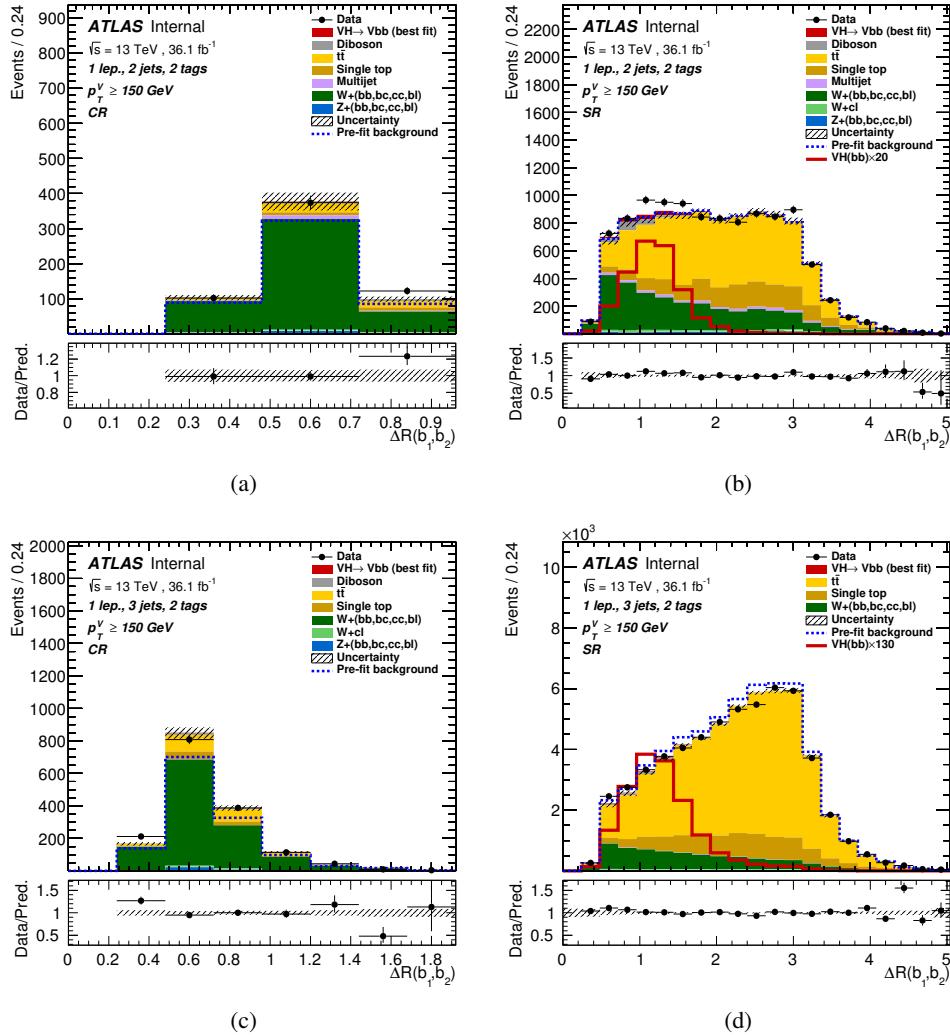


Figure 175: $\Delta R(b, b)$ postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

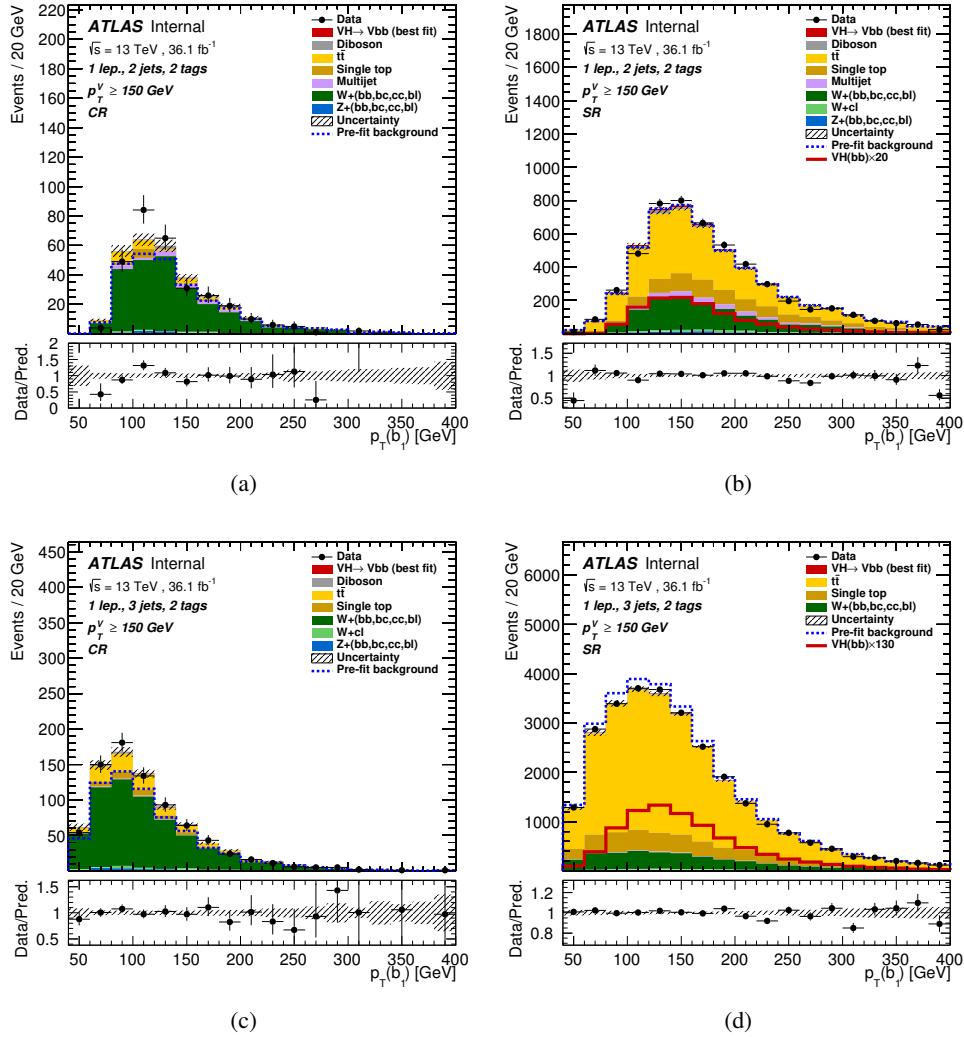


Figure 176: $p_T(\text{jet 1})$ postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

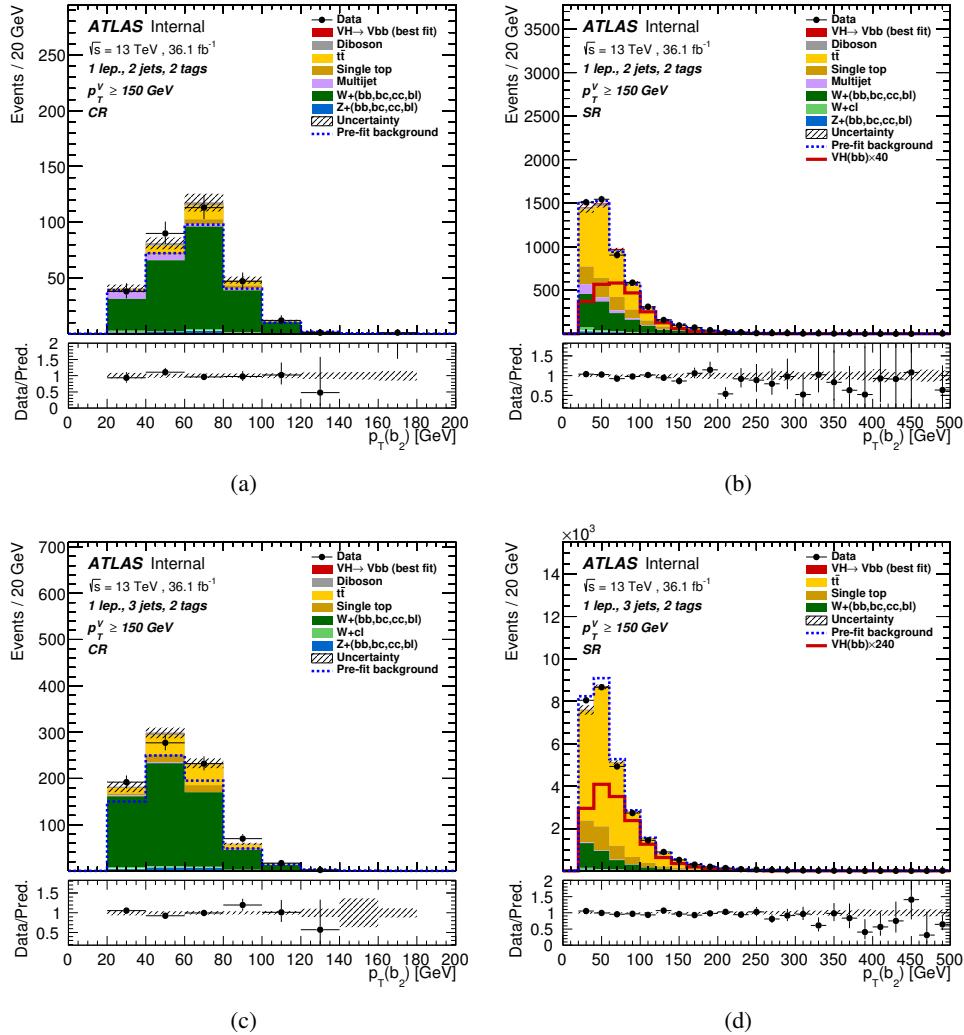


Figure 177: $p_T(\text{jet 2})$ postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

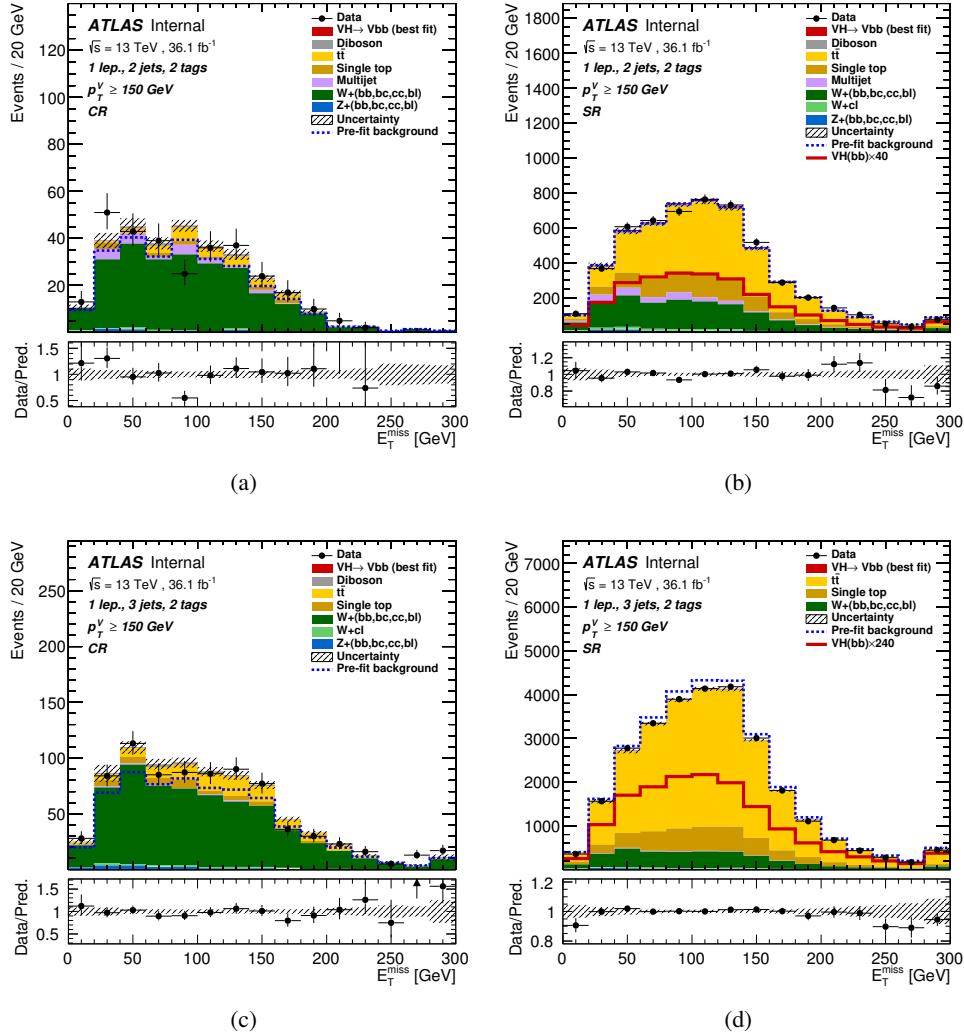


Figure 178: E_T^{miss} postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

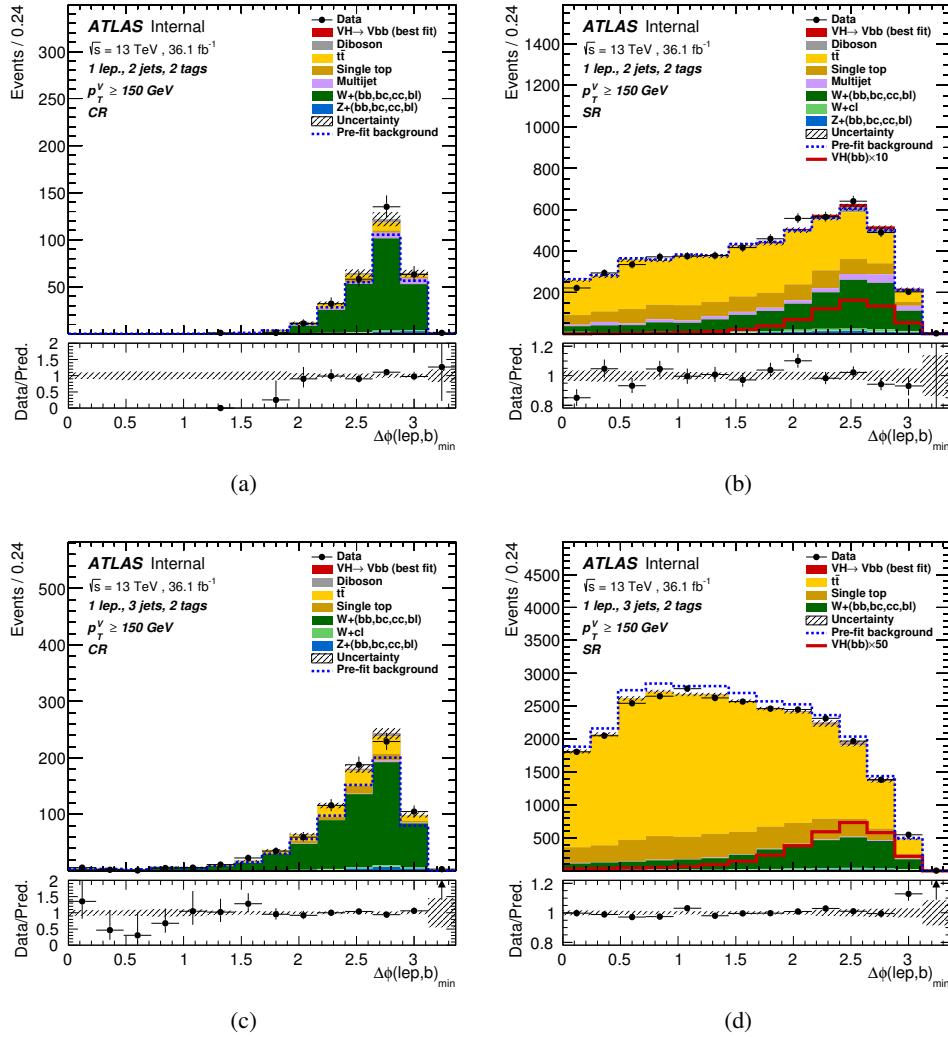


Figure 179: $\Delta\phi \min(l, b)$ postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

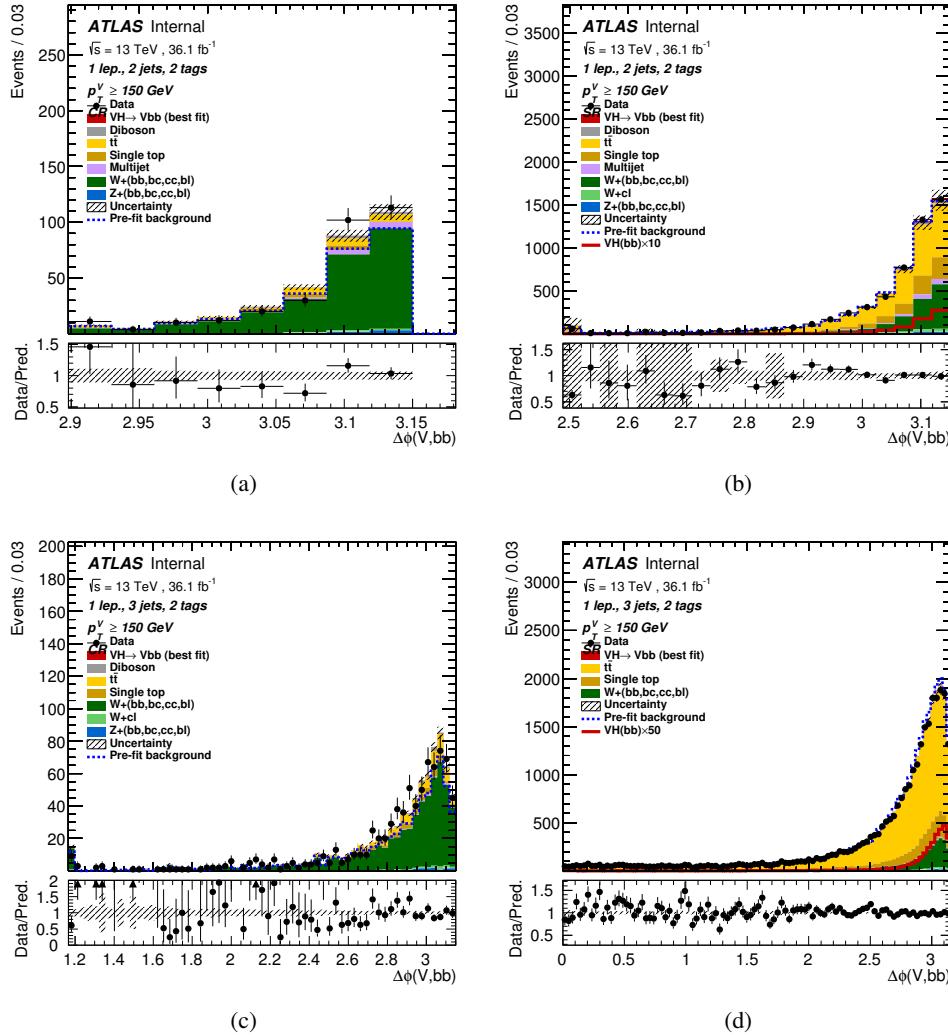


Figure 180: $\Delta\phi(V, H)$ postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

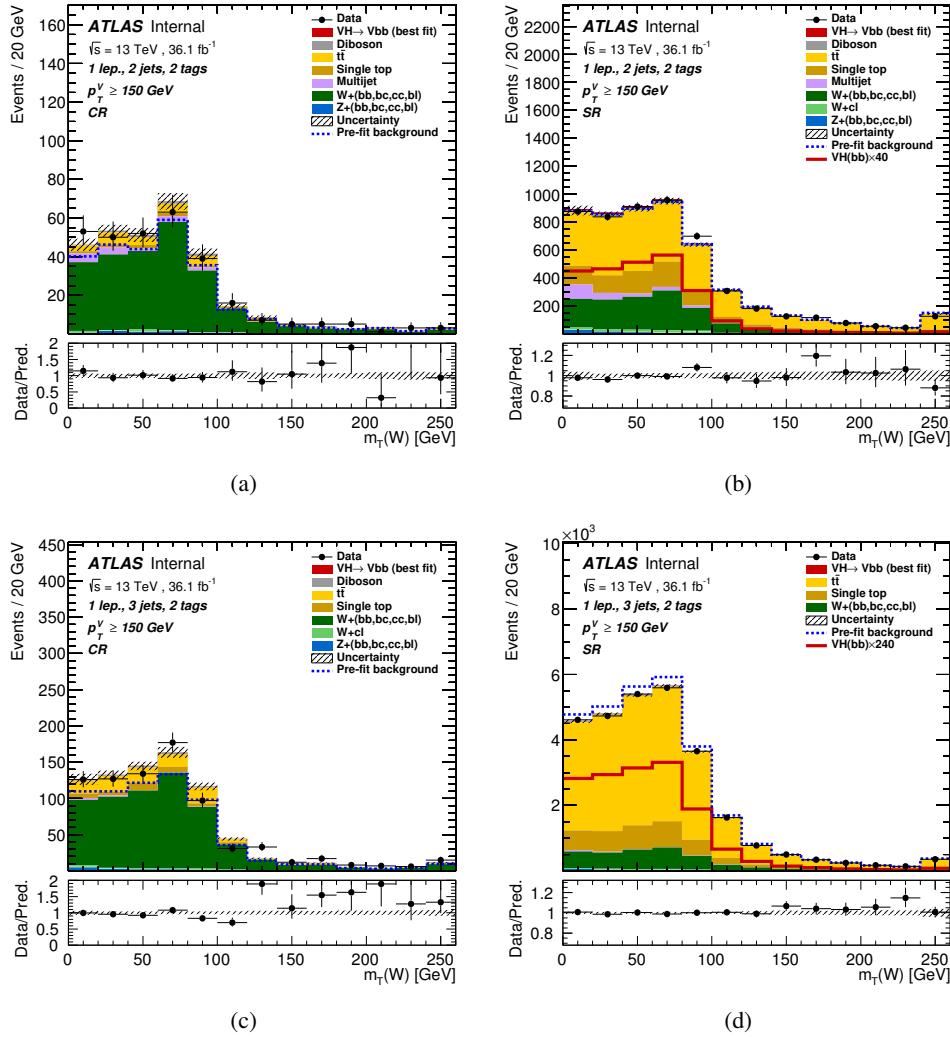


Figure 181: m_T^W postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

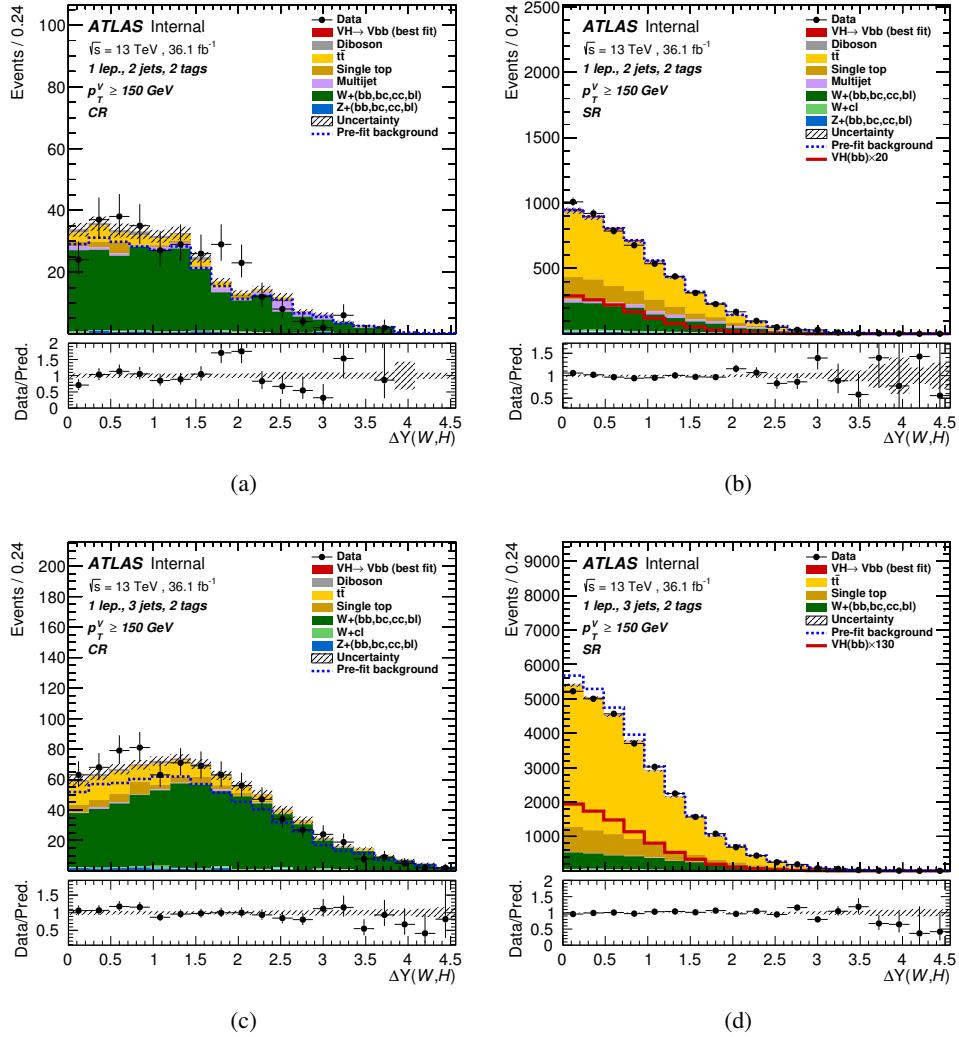


Figure 182: $\Delta Y(W, H)$ postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

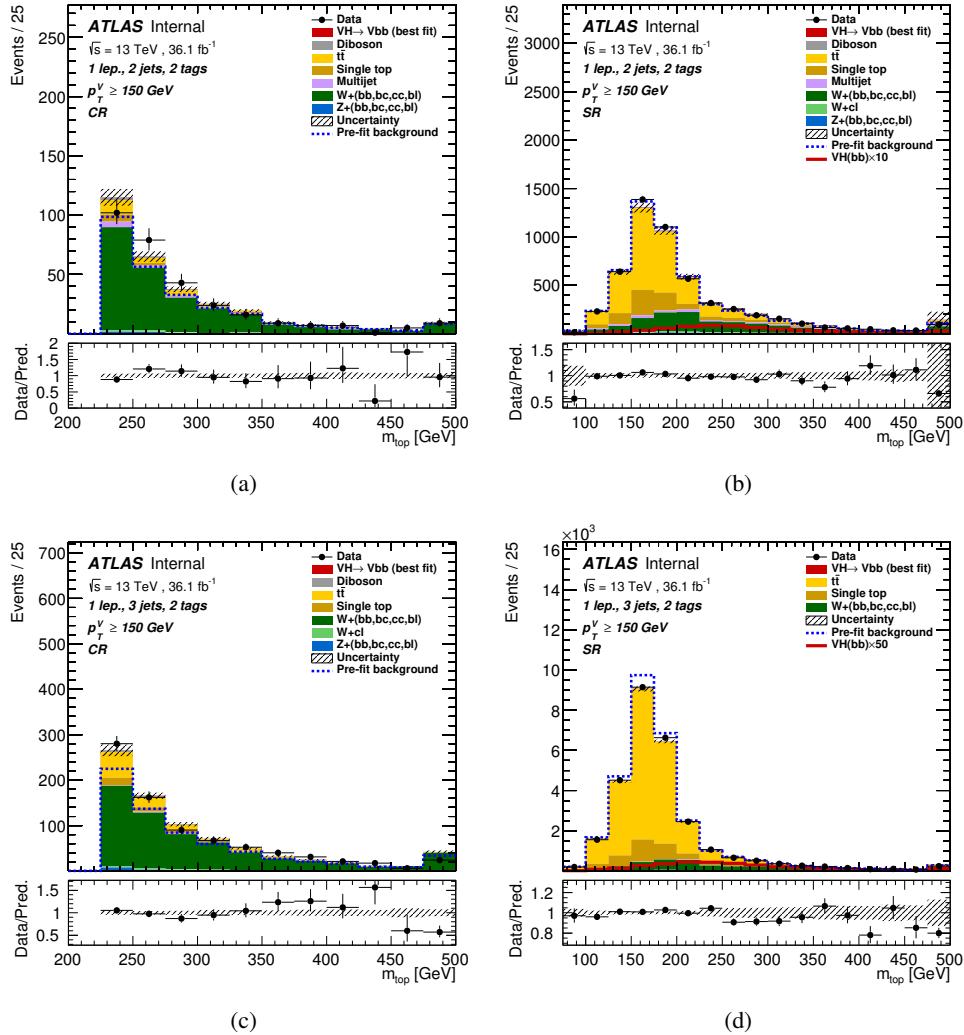


Figure 183: m_{top} postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 2-jet $W + hf$ CR (a), SR (b), 2-tag 3-jet $W + hf$ CR (c) and SR (d).

Not reviewed, for internal circulation only

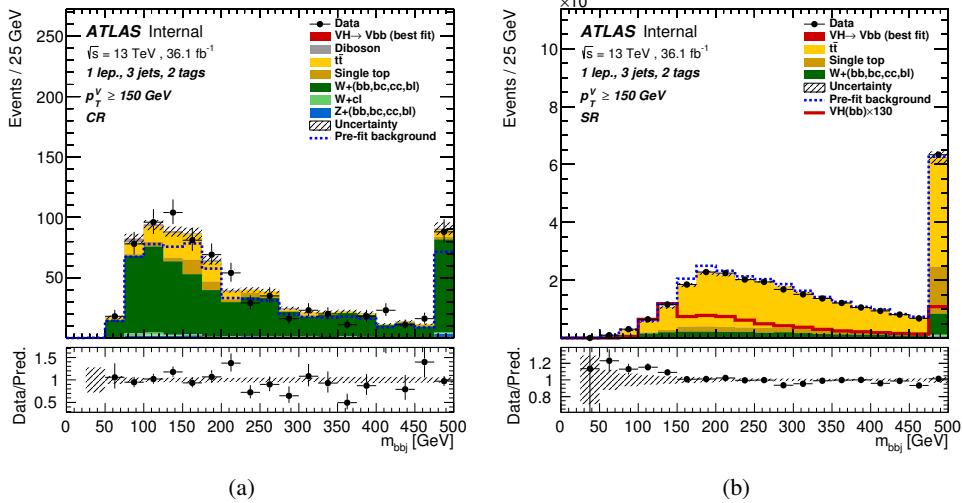


Figure 184: m_{bbj} postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 3-jet $W + hf$ CR (a) and SR (b).

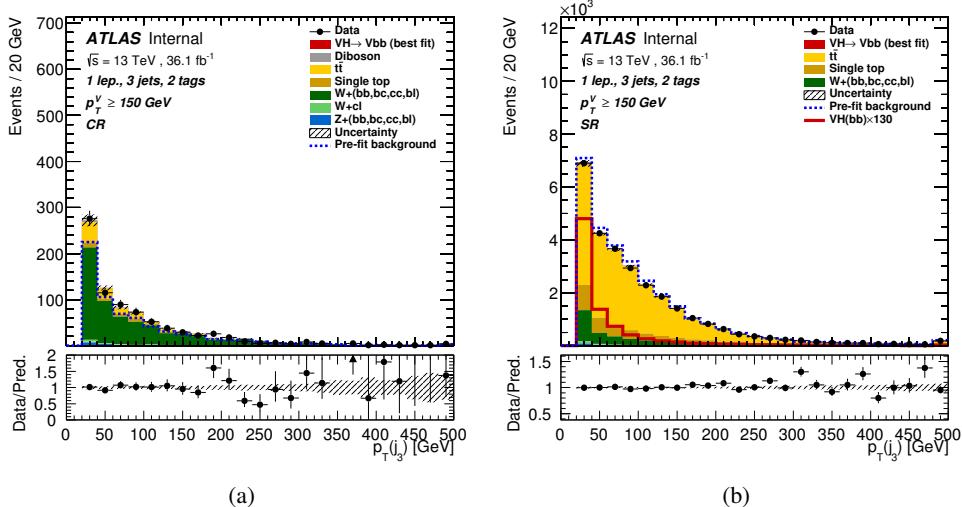


Figure 185: p_T (jet 3) postfit distributions obtained in the 1-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 3-jet $W + hf$ CR (a) and SR (b).

1439 K.3. 2-lepton

Not reviewed, for internal circulation only

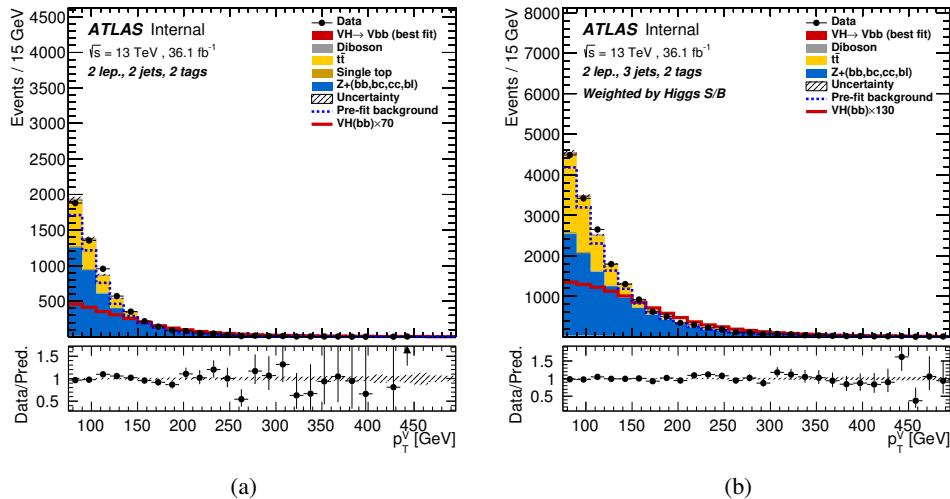


Figure 186: p_T^V postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet (a), 2-tag 3-jet (b) regions

Not reviewed, for internal circulation only

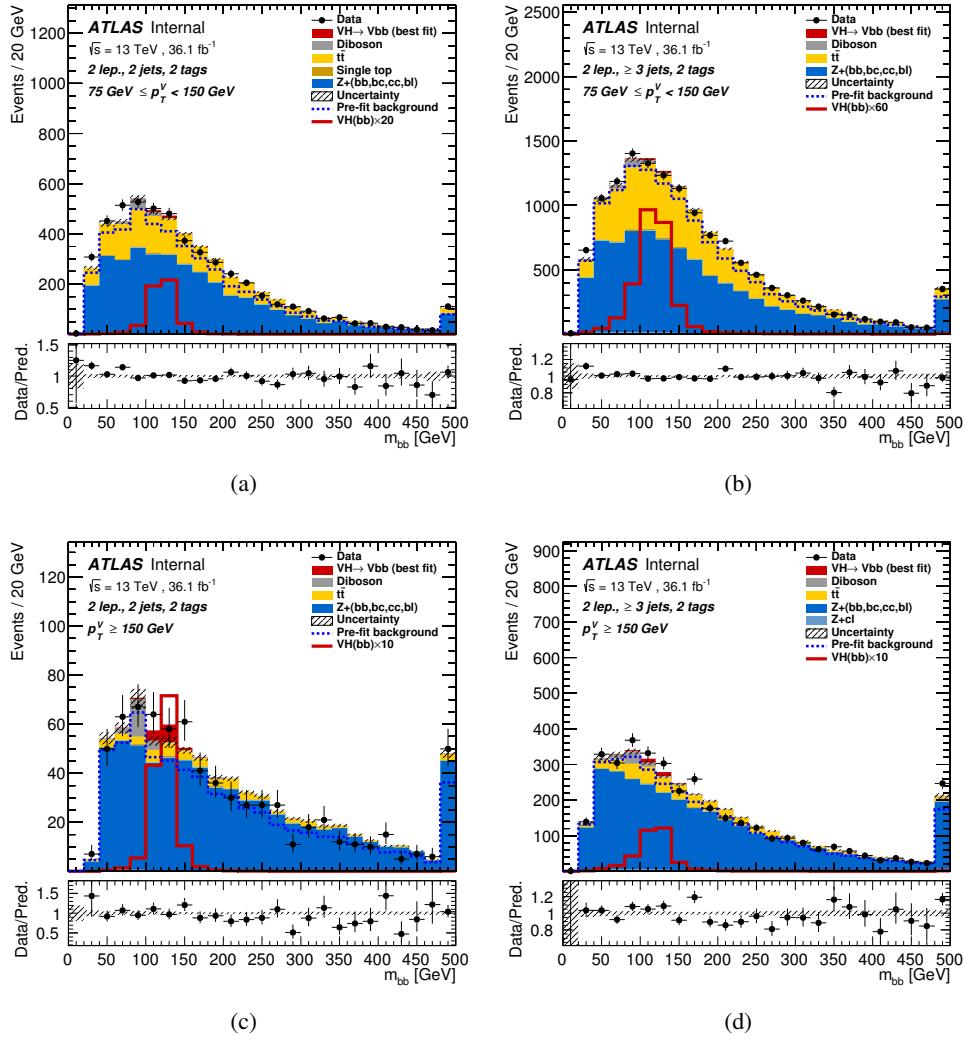


Figure 187: m_{bb} postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet $75 < p_T^V < 150$ GeV (a), 2-tag 3-jet $75 < p_T^V < 150$ GeV (b), 2-tag 2-jet $p_T^V \geq 150$ GeV (c) and 2-tag 3-jet $p_T^V \geq 150$ GeV (d)

Not reviewed, for internal circulation only

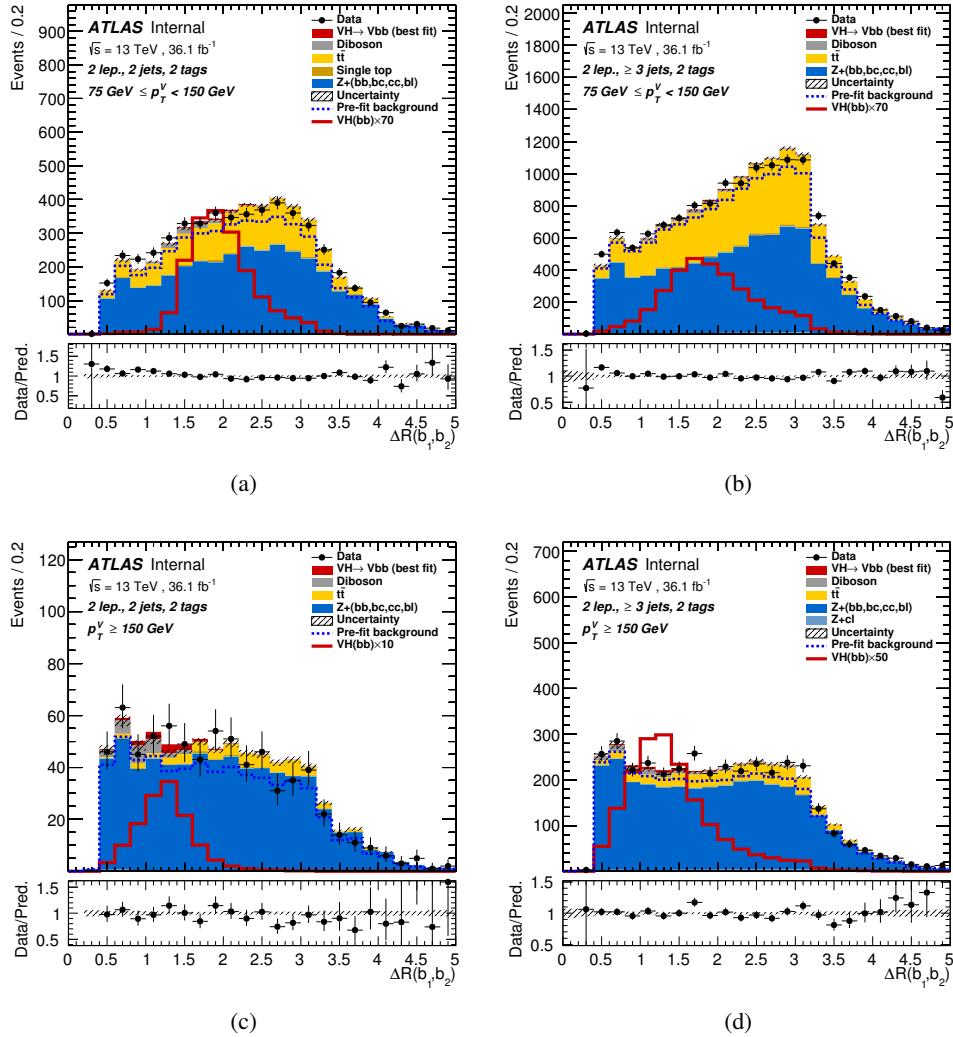


Figure 188: $\Delta R(b, b)$ postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet $75 < p_T^V < 150 \text{ GeV}$ (a), 2-tag 3-jet $75 < p_T^V < 150 \text{ GeV}$ (b), 2-tag 2-jet $150 < p_T^V \text{ GeV}$ (c) and 2-tag 3-jet $150 < p_T^V \text{ GeV}$ (d)

Not reviewed, for internal circulation only

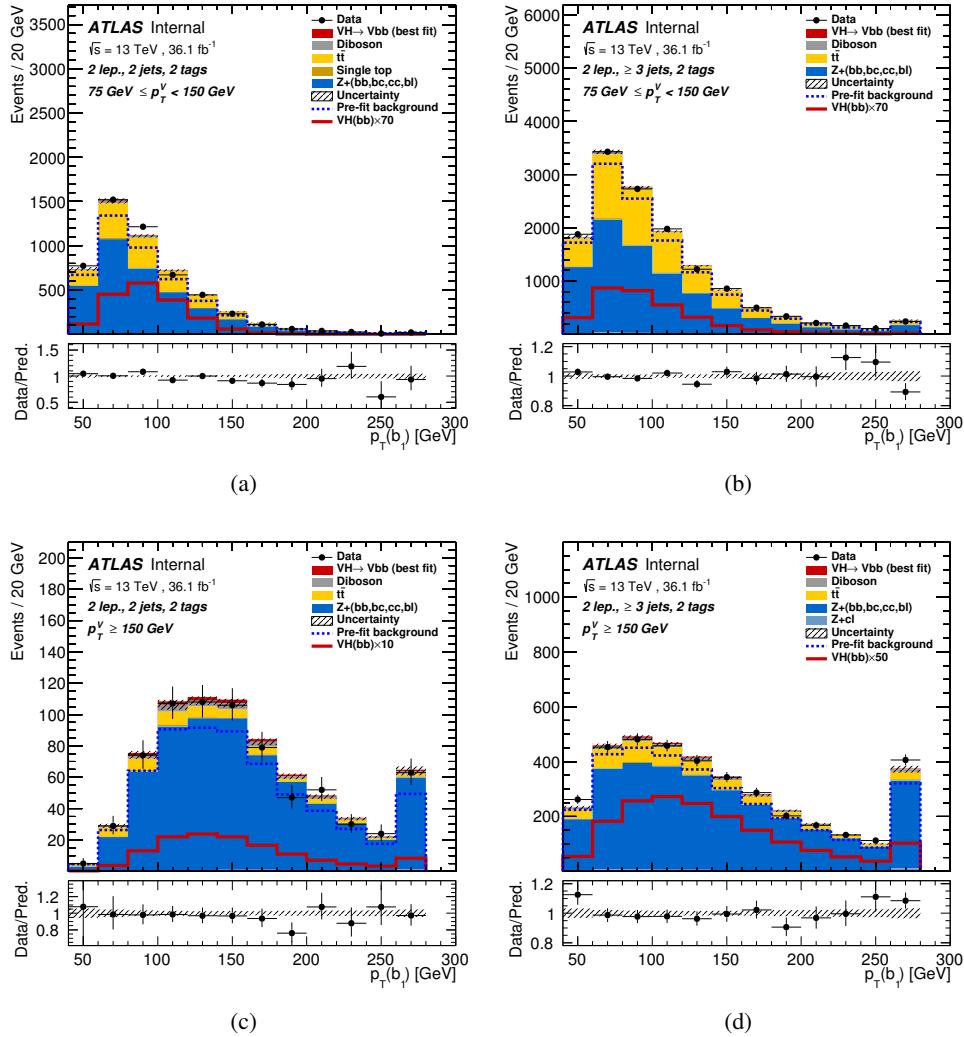


Figure 189: p_T (jet 1) postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet $75 < p_T^V < 150$ GeV (a), 2-tag 3-jet $75 < p_T^V < 150$ GeV (b), 2-tag 2-jet $150 < p_T^V$ GeV (c) and 2-tag 3-jet $150 < p_T^V$ GeV (d)

Not reviewed, for internal circulation only

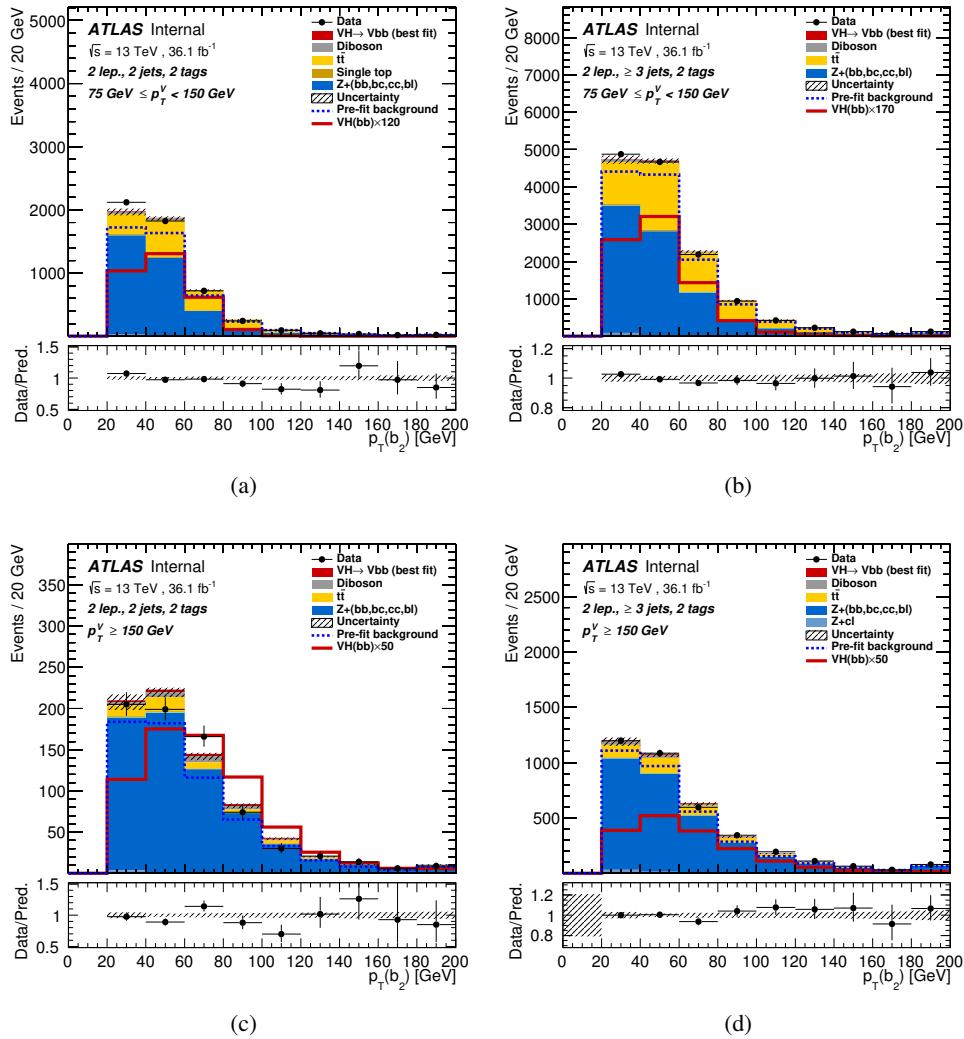


Figure 190: p_T (jet 2) postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet $75 < p_T^V < 150$ GeV (a), 2-tag 3-jet $75 < p_T^V < 150$ GeV (b), 2-tag 2-jet $150 < p_T^V$ GeV (c) and 2-tag 3-jet $150 < p_T^V$ GeV (d)

Not reviewed, for internal circulation only

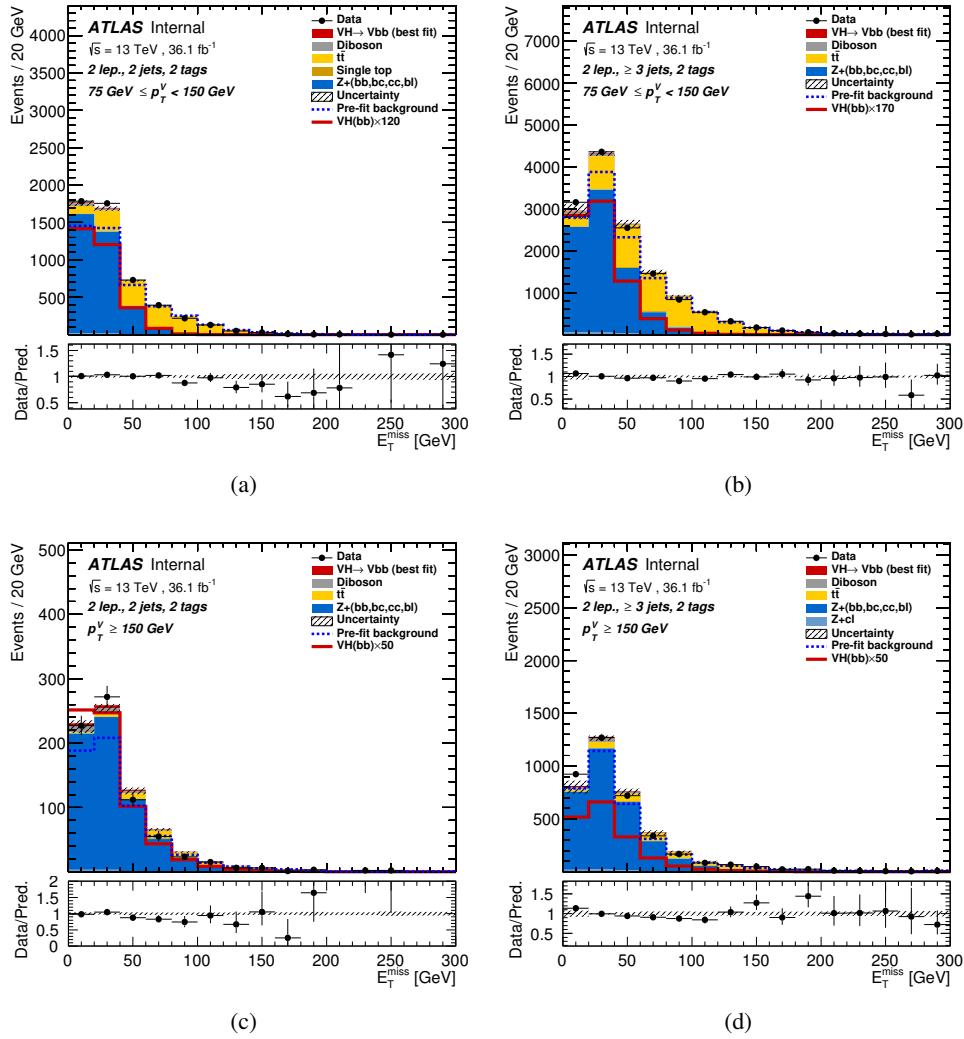


Figure 191: E_T^{miss} postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet $75 < p_T^V < 150 \text{ GeV}$ (a), 2-tag 3-jet $75 < p_T^V < 150 \text{ GeV}$ (b), 2-tag 2-jet $150 < p_T^V \text{ GeV}$ (c) and 2-tag 3-jet $150 < p_T^V \text{ GeV}$ (d)

Not reviewed, for internal circulation only

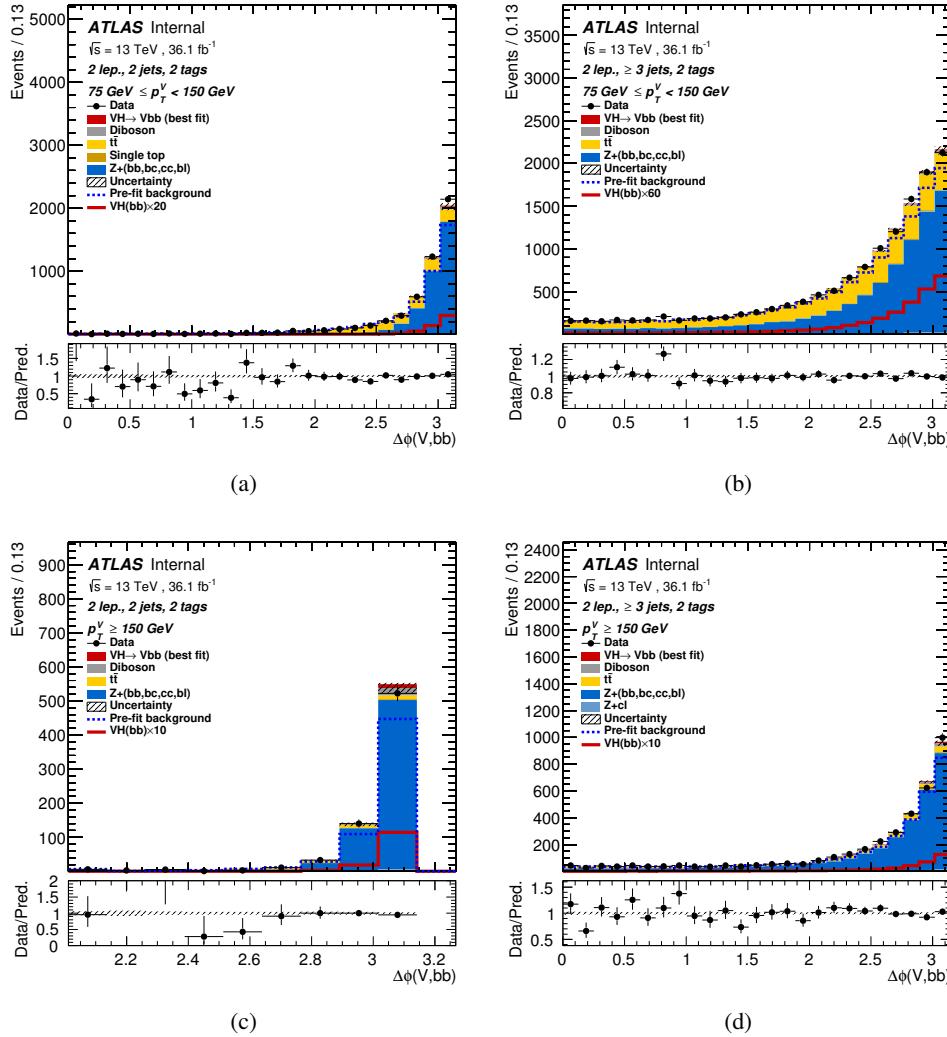


Figure 192: $\Delta\phi(V, H)$ postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet $75 < p_T^V < 150 \text{ GeV}$ (a), 2-tag 3-jet $75 < p_T^V < 150 \text{ GeV}$ (b), 2-tag 2-jet $150 < p_T^V \text{ GeV}$ (c) and 2-tag 3-jet $150 < p_T^V \text{ GeV}$ (d)

Not reviewed, for internal circulation only

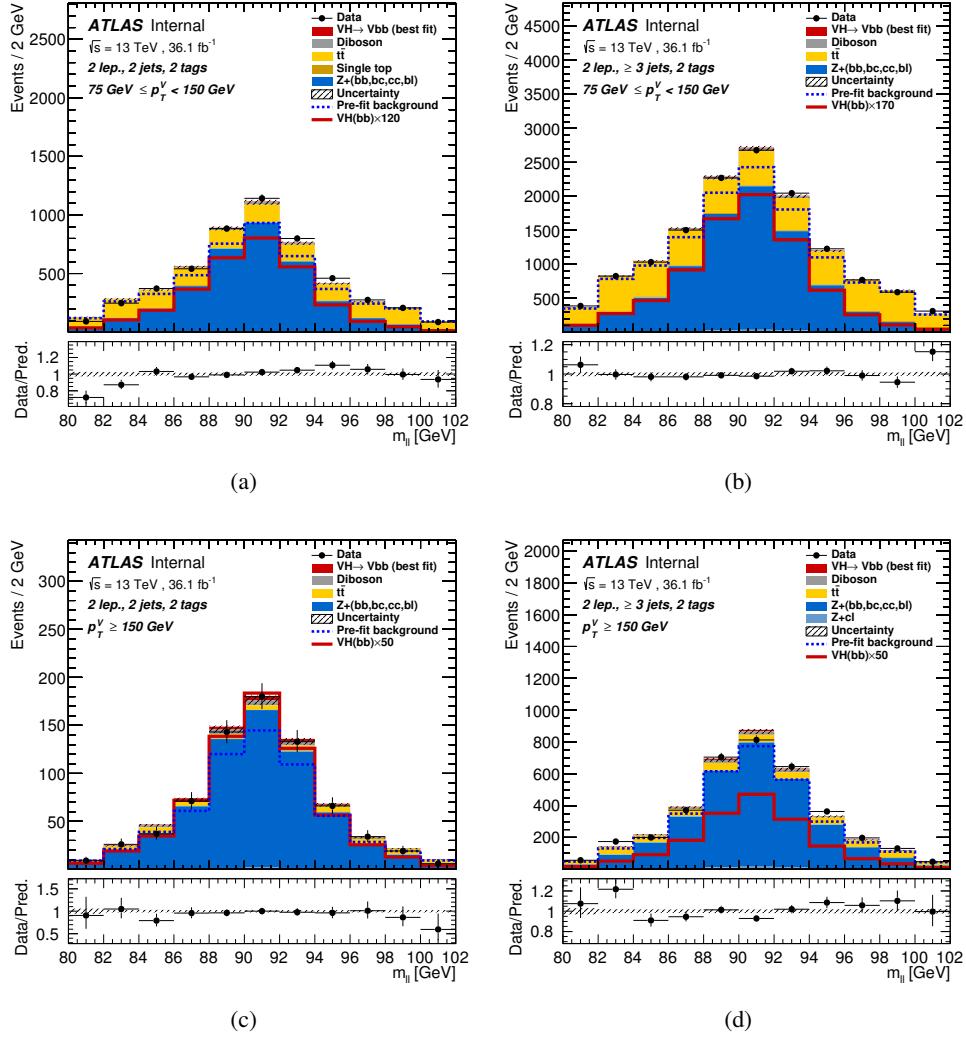


Figure 193: m_{ll} postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data, for the 2-tag 2-jet $75 < p_T^V < 150$ GeV (a), 2-tag 3-jet $75 < p_T^V < 150$ GeV (b), 2-tag 2-jet $150 < p_T^V$ GeV (c) and 2-tag 3-jet $150 < p_T^V$ GeV (d)

Not reviewed, for internal circulation only

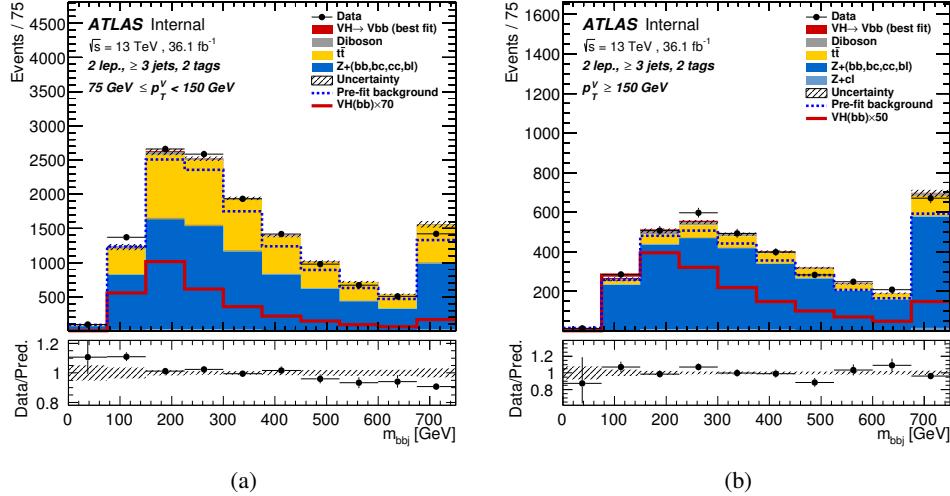


Figure 194: m_{bbj} postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 3-jet $75 < p_T^V < 150 \text{ GeV}$ (a), 2-tag 2-jet $150 < p_T^V \text{ GeV}$ (b) regions.

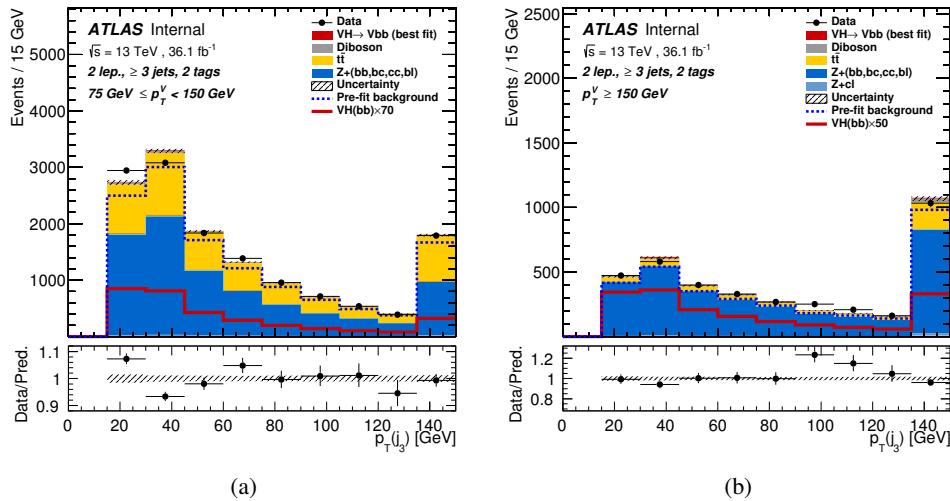


Figure 195: $p_T(\text{jet 3})$ postfit distributions obtained in the 2-lepton channel from the unconditional fit on 2015 and 2016 data for the 2-tag 3-jet $75 < p_T^V < 150 \text{ GeV}$ (a), 2-tag 2-jet $150 < p_T^V \text{ GeV}$ (b) regions.

Not reviewed, for internal circulation only

1440 L. Cross-check: Standard Model Diboson ($VZ \rightarrow b\bar{b}$) measurement

1441 Studies on the diboson ($VZ \rightarrow b\bar{b}$) fit offers a good cross-check of our full analysis procedure for the
 1442 SM VH search. The final state is identical to the signal, and the kinematics are similar to the SM VH ,
 1443 except for the substitution of Z boson for a H boson. The diboson signal has a softer p_T^V distribution
 1444 than the Higgs. This has the important consequence to increase the importance of the 2-lepton channel
 1445 (which has a low p_T^V regions), relatively to the 0- and 1-lepton channels, when comparing with the Higgs
 1446 fit. For the diboson measurement, the BDT for each channel is retrained with the diboson $VZ \rightarrow b\bar{b}$ as
 1447 the signal. The input variables and BDT configurations are same as for the SM VH BDT. VZ diboson
 1448 fit and measurement of $VZ \rightarrow b\bar{b}$ signal strength are then performed. Results from the unconditional fit
 1449 performed to all fitting regions is presented, including comparisons of the pulls, rankings and correlations
 1450 of nuisance parameters to the VH fit.

1451 There is, however, one difference between the VZ fit and the VH fit. Due to limited MC statistics in the
 1452 diboson samples, the parameters of the BDT transformation, Section 5.3, are varied. Instead of using
 1453 $z_s = 10, z_b = 5$, this is altered to $z_s = 5, z_b = 5$. This increases the diboson signal statistics, reducing the
 1454 impact of MC statistical uncertainties on the fit result.

1455 L.1. 0-lepton

1456 L.1.1. Prefit BDT distribution

1457 Figure 196 presents prefit distributions of the VZ MVA in the 0-lepton signal. Transformation D has been
 1458 applied, with the highest BDT output corresponding to the region most sensitive to the VZ signal.

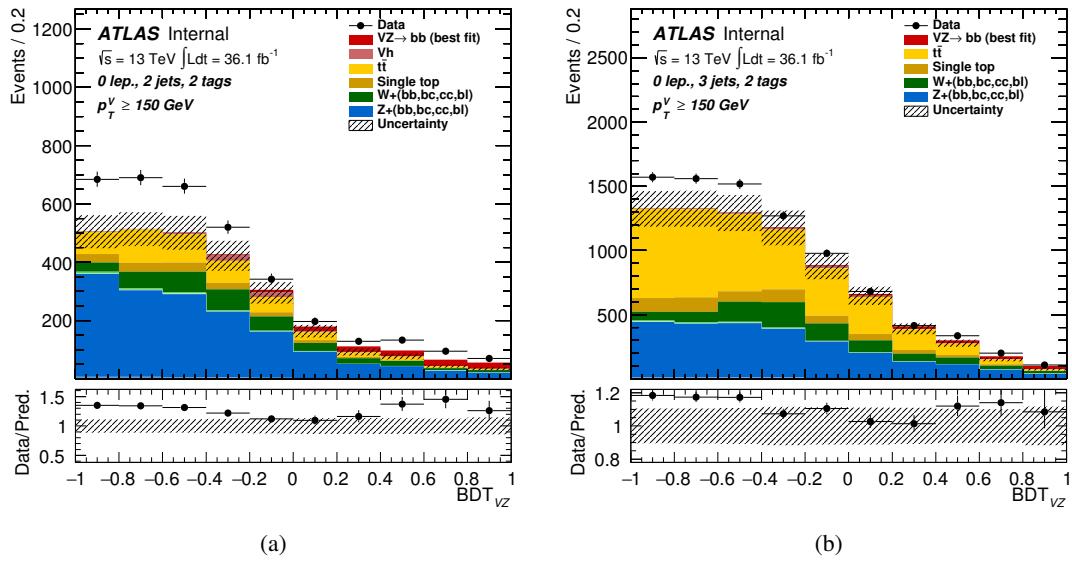


Figure 196: Unblinded prefit BDT distribution in the 0-lepton channel with BDT transformation D applied. The 2-tag 2-jet SR, and the 2-tag 3-jet SR categories are shown. The background and signal samples are normalized to the cross-section predictions.

1459 L.1.2. Comparison to VH Fit

In this section, the results from the fit to data in the VH and ZH fits are compared.

Figure 197 shows pull distributions for all NPs in the 0-lepton fit to data in the VH fit, and the VZ fit.

Not reviewed, for internal circulation only

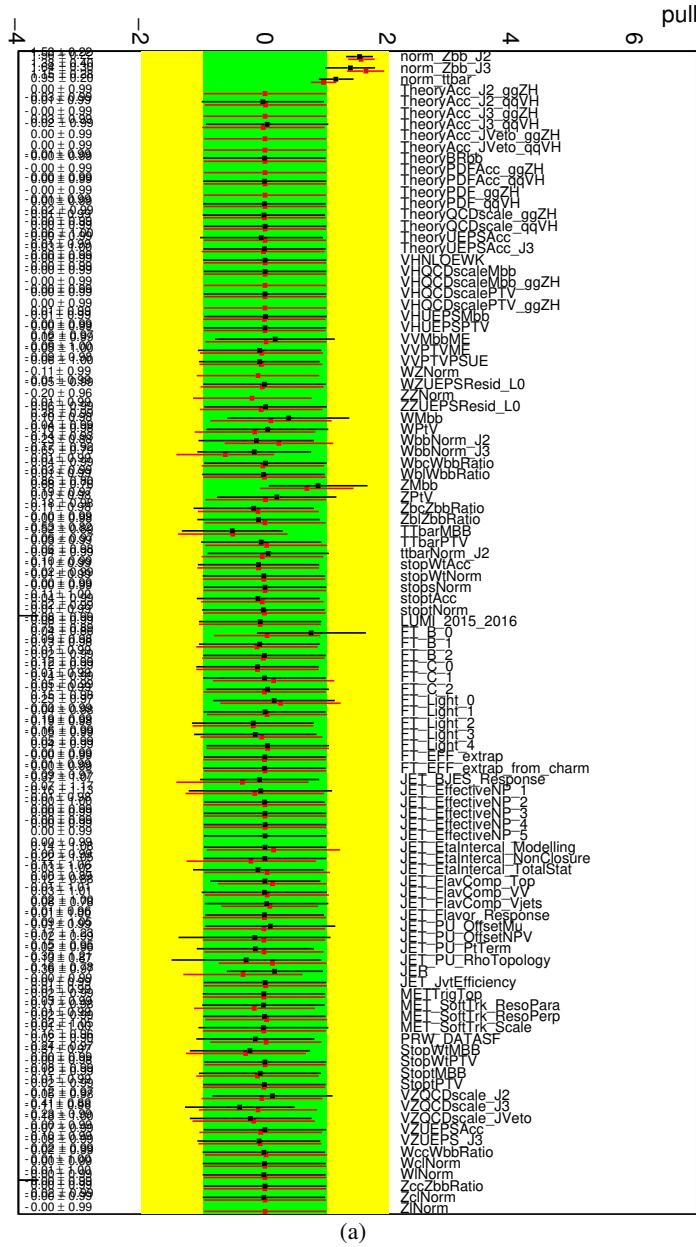


Figure 197: Comparison of the pulls in the 0-lepton unconditional Vh (red) and diboson (black) fits to data.

L.1.3. Results

1462 Figure 200 shows postfit BDT distribution from the VZ diboson fit for the 0-lepton channel after the
1463 unconditional fit.

1464 Figure 199 shows the ranking of the systematic uncertainties based on the impact on μ after the
1465 unconditional fit.

1466 Table 56 shows the breakdown of the uncertainties from data statistics and systematic uncertainties after
1467 the unconditional fit.

Table 56: The breakdown of the uncertainties in the 0 lepton diboson fit, coming from data statistics (“Data Stat.”),
 systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”),
 and experimental and modelling systematic uncertainties only (“All Syst.”) is shown in terms of the total impact on
 the uncertainty on μ after the unconditional fit.

Set of nuisance parameters	Impact on error
Total	+0.399 / -0.335
DataStat	+0.183 / -0.178
FullSyst	+0.354 / -0.283
Floating normalizations	+0.189 / -0.154
All normalizations	+0.195 / -0.161
All but normalizations	+0.313 / -0.234
Jets MET	+0.092 / -0.079
BTag	+0.074 / -0.053
Leptons	+0.000 / -0.000
Luminosity	+0.048 / -0.030
Diboson	+0.112 / -0.068
Zjets	+0.188 / -0.157
Wjets	+0.085 / -0.073
Model ttbar	+0.183 / -0.143
Model Single Top	+0.022 / -0.018
Model Multi Jet	+0.000 / -0.000
Signal Systematics	+0.002 / -0.002
MC stat	+0.121 / -0.111

Not reviewed, for internal circulation only

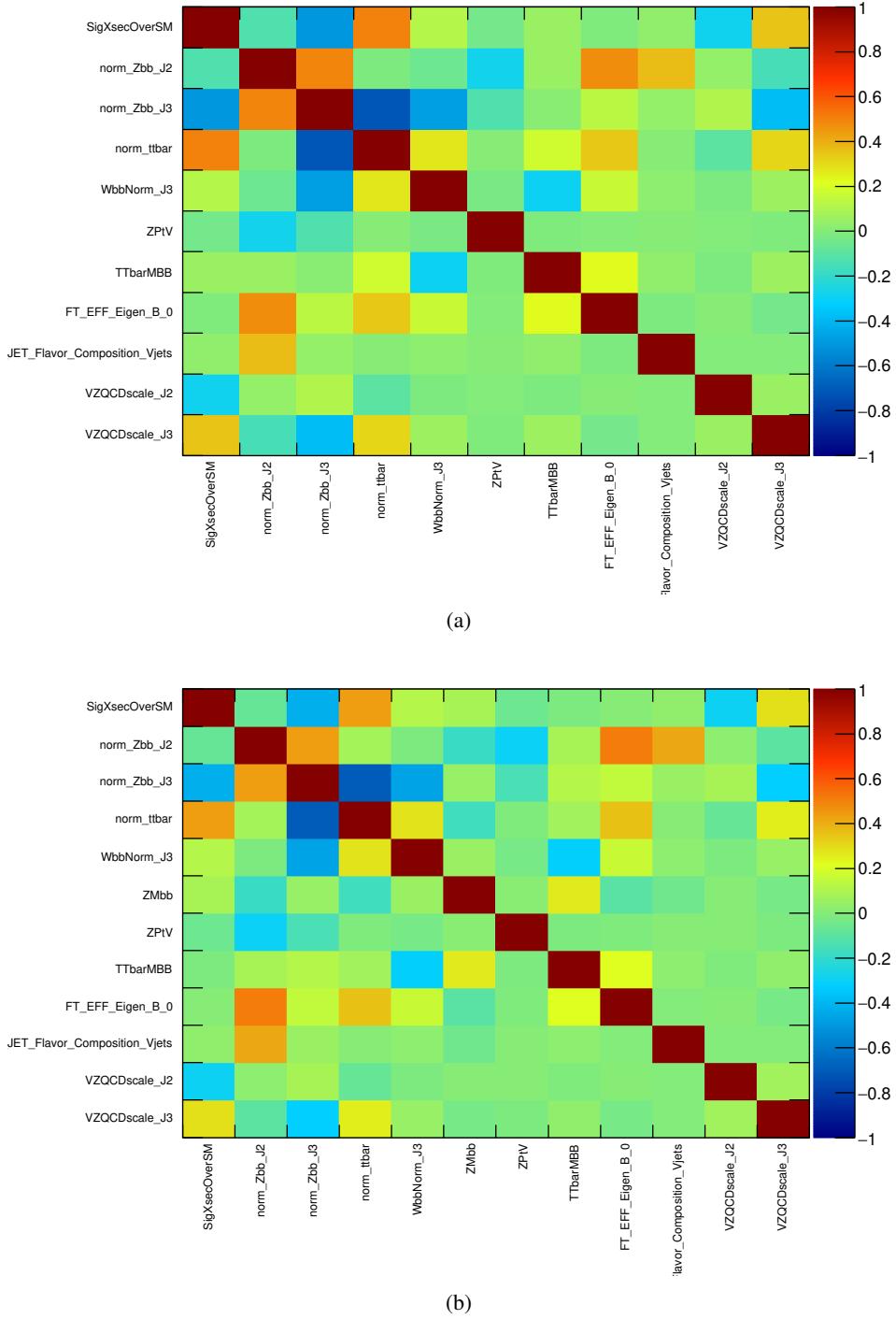


Figure 198: Highest correlation plots in the 0-lepton unconditional fits (a) Asimov (b) Data

Not reviewed, for internal circulation only

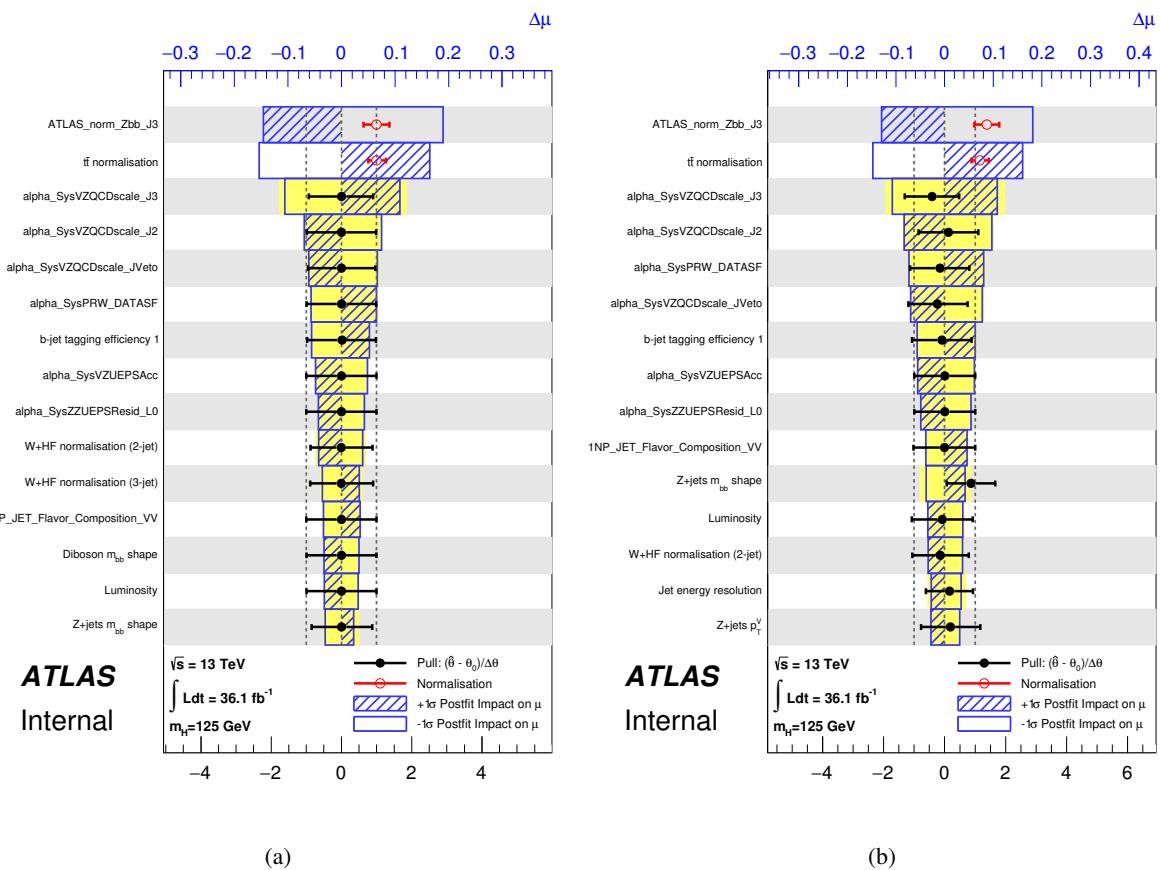


Figure 199: The nuisance parameter ranking for the 0-lepton channel fit (a) based Asimov built from MC prediction prefit and (b) based on the unconditional postfit to data.

Not reviewed, for internal circulation only

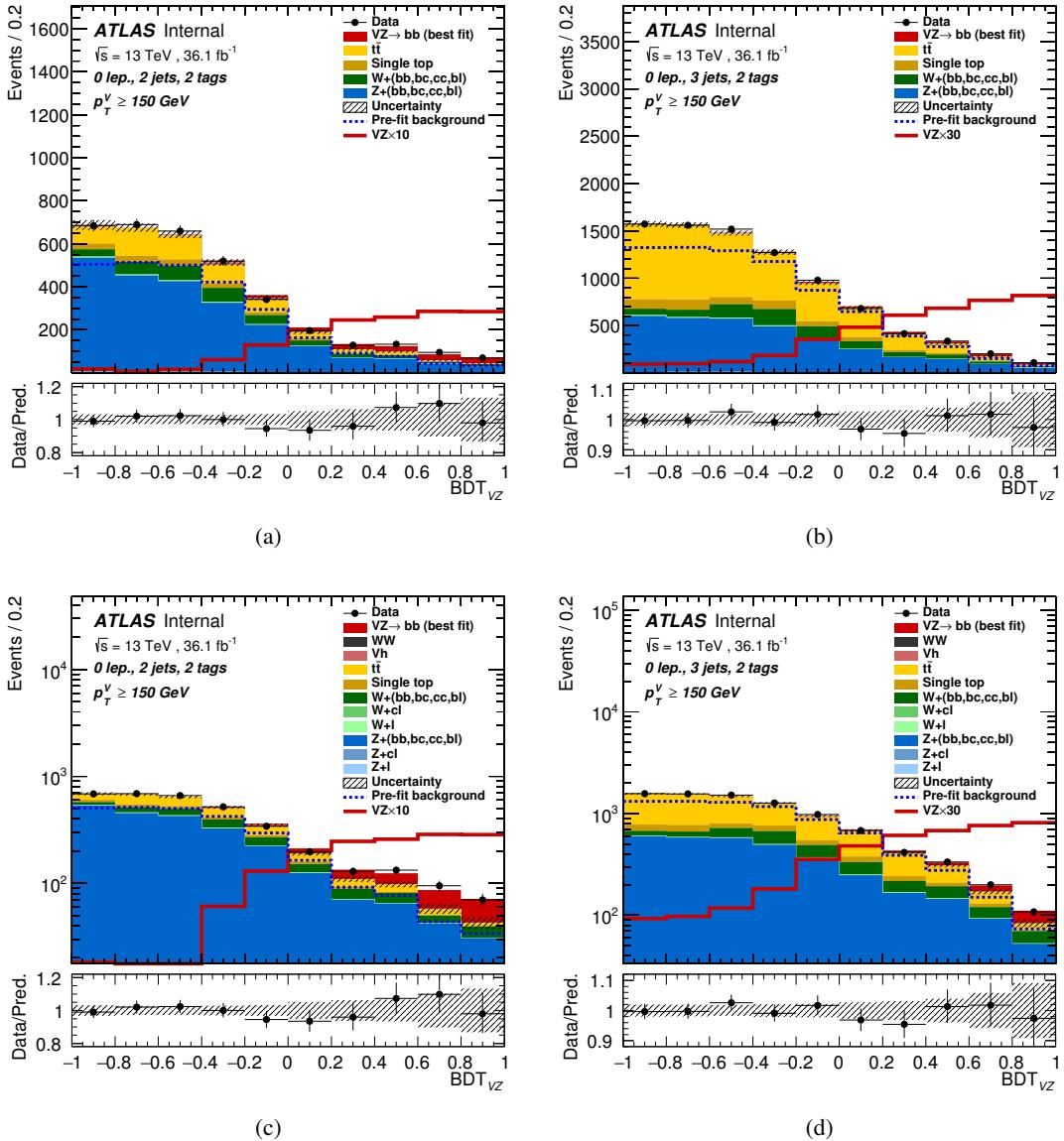


Figure 200: Post-fit diboson BDT distribution in the 0-lepton channel with BDT transformation D applied, after running the unconditional fit to data. The 2-tag 2-jet SR (left), and the 2-tag 3-jet SR (right) categories are shown.

1469 **L.2. 1-lepton**1470 **L.2.1. Prefit BDT distribution**

1471 Figure 201 presents prefit distributions of the VZ MVA in the 1-lepton signal and control regions.
 1472 Transformation D has been applied, with the highest BDT output corresponding to the region most
 1473 sensitive to the VZ signal.

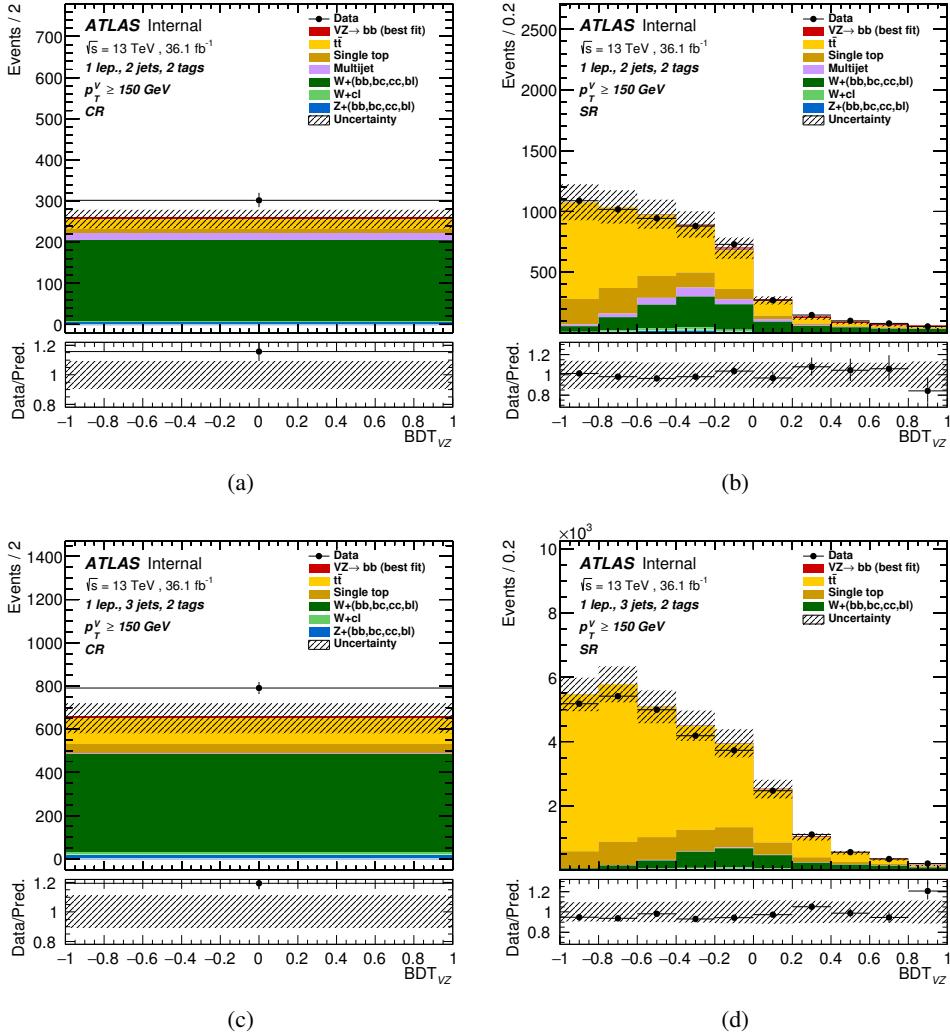


Figure 201: Unblinded prefit BDT distribution in the 1-lepton channel with BDT transformation D applied. The 2-tag 2-jet, $W + hf$ CR (a), SR (b), and the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

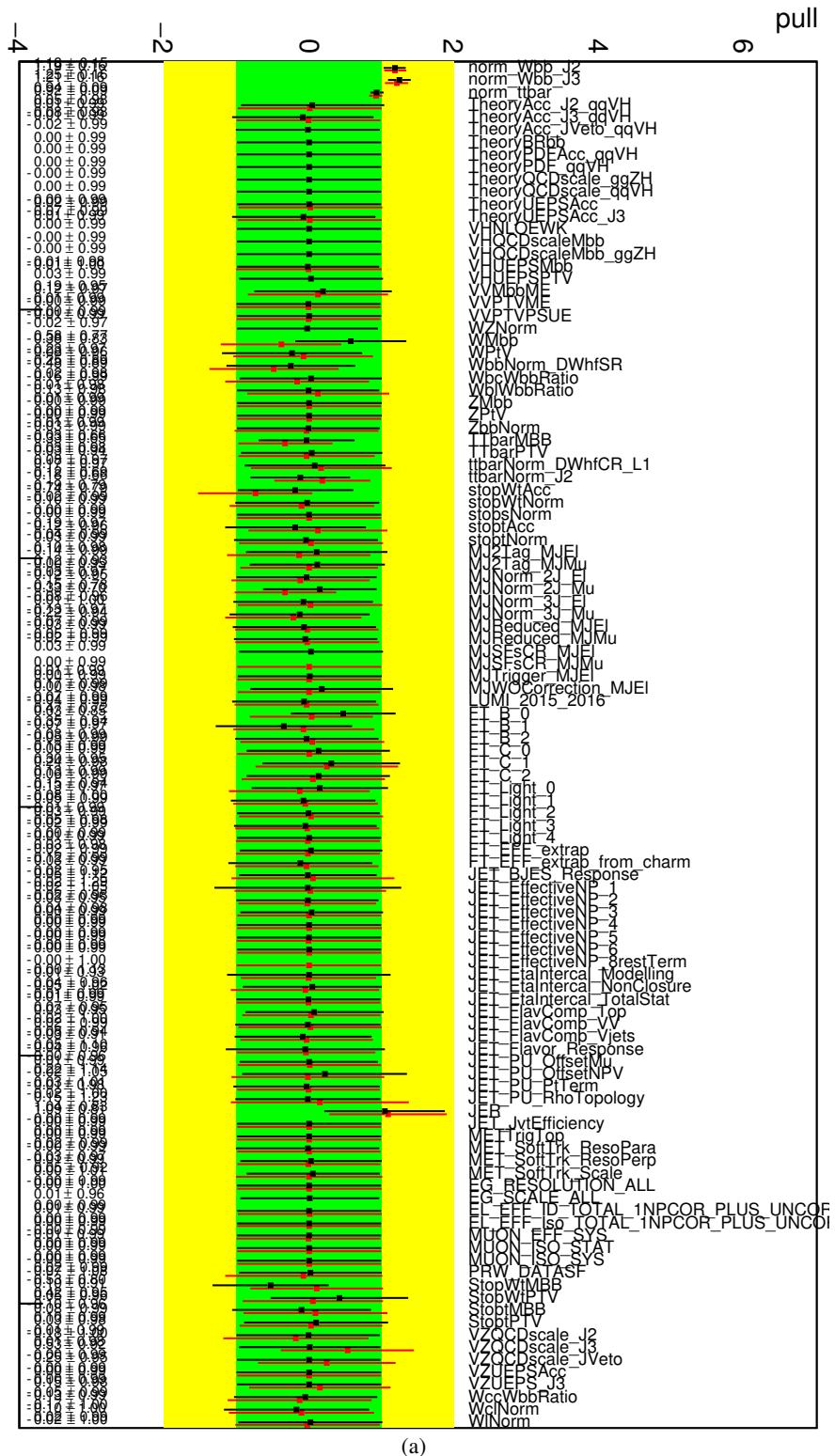
1474 L.2.2. Comparison to VH Fit

1475 In this section, the results from the unconditional fit to data in the VH and ZH fits are compared.

1476 Figure 202 shows pull distributions for all NPs in the 1-lepton unconditional fit to data in the VZ fit (red),
1477 and the unconditional VZ fit to data (black).

1478 Figure 203 shows correlation between nuisance parameters in the VH and ZH fits. The plots shows the
1479 NPs having high correlation (more than 20%). Figure 204 presents the ranking of nuisance parameters,
1480 based on their impact on μ , again in the VH and ZH fits.

Not reviewed, for internal circulation only



(a)

Figure 202: Pull plot comparing all NPs in the 1-lepton unconditional fit to the VH signal (red), and the unconditional VZ fit to data (black).

Not reviewed, for internal circulation only

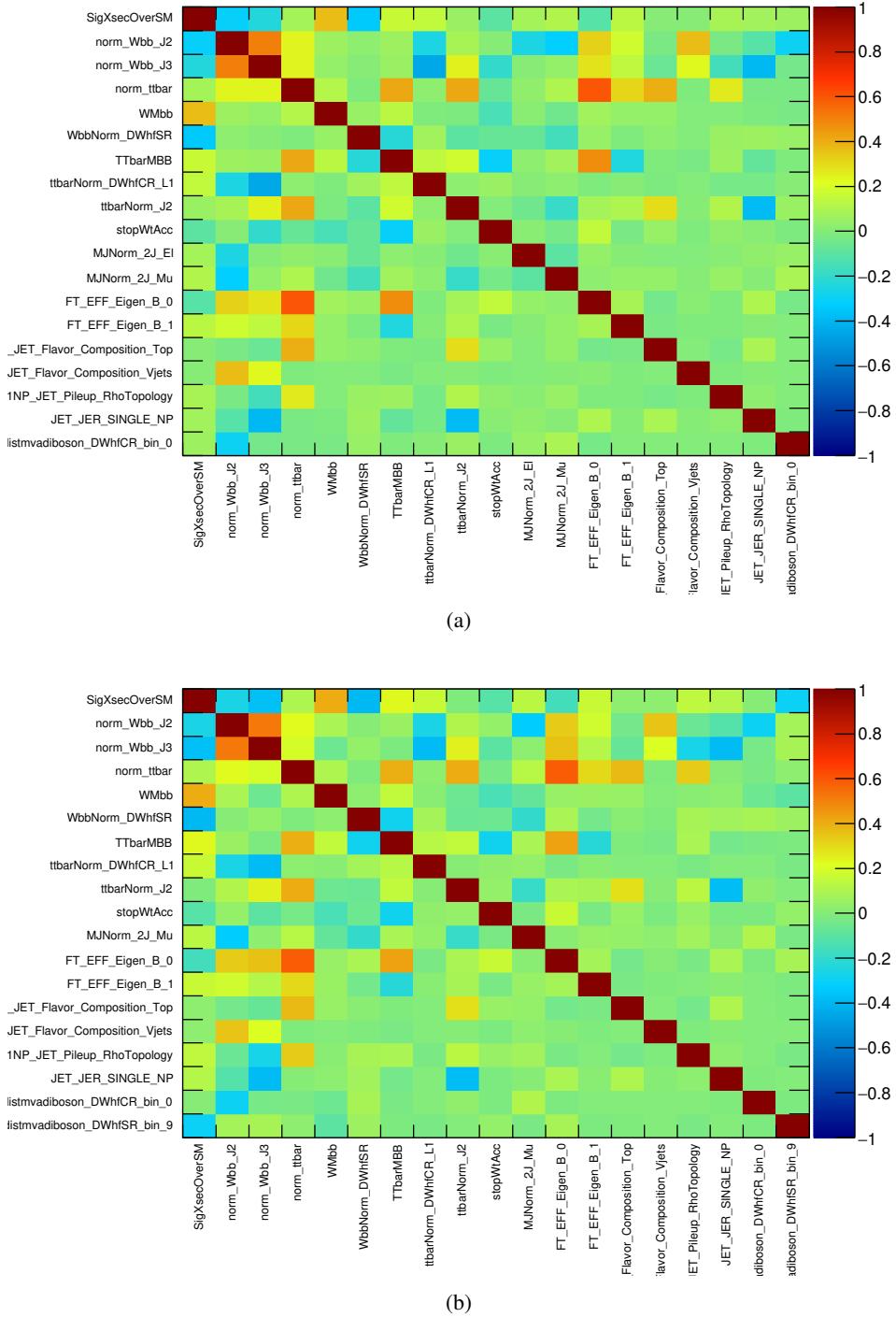


Figure 203: Correlation matrices from the unconditional VZ fit in the 1-lepton channel to an Asimov dataset (a) and to 2015 and 2016 data (b).

Not reviewed, for internal circulation only

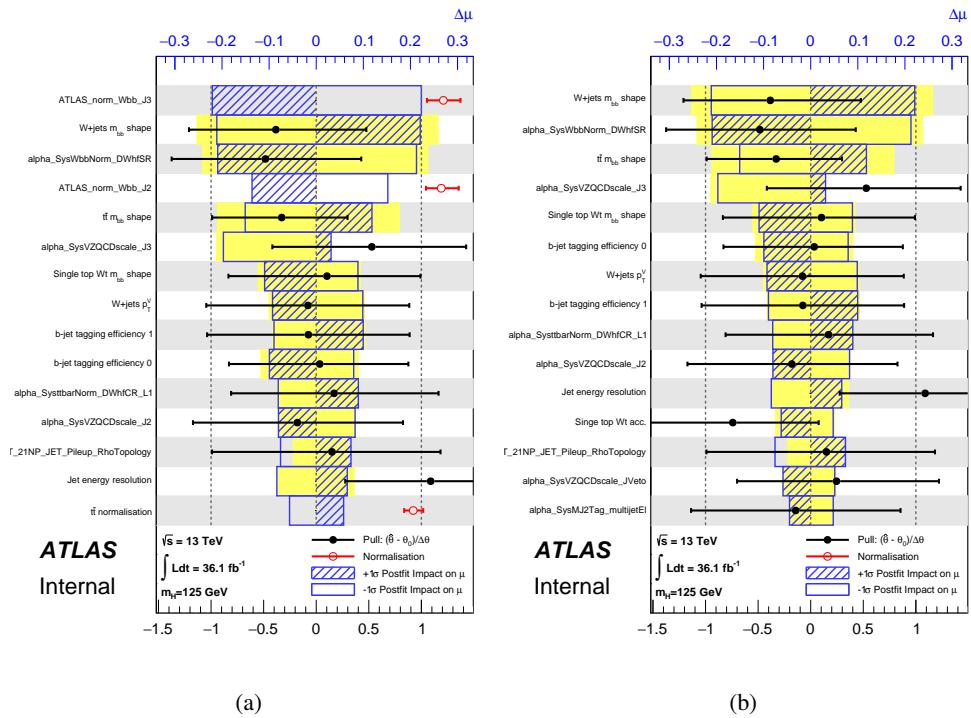


Figure 204: The nuisance parameter ranking from the unconditional VZ fit to data in the 1-lepton channel, ordered by pre-fit impact (a), when ordered by post-fit impact (b).

1481 **L.2.3. Results**

1482 Table 57 shows the breakdown of the uncertainties from data statistics and systematic uncertainties for the
1483 1-lepton unconditional fit to data.

Set of nuisance parameters	Impact on error
Total	+0.583 / -0.594
DataStat	+0.233 / -0.228
FullSyst	+0.535 / -0.549
Floating normalizations	+0.221 / -0.284
All normalizations	+0.310 / -0.372
All but normalizations	+0.476 / -0.491
Jets MET	+0.118 / -0.141
BTag	+0.134 / -0.157
Leptons	+0.003 / -0.010
Luminosity	+0.053 / -0.021
Diboson	+0.091 / -0.057
Zjets	+0.002 / -0.005
Wjets	+0.383 / -0.435
Model ttbar	+0.156 / -0.175
Model Single Top	+0.124 / -0.126
Model Multi Jet	+0.125 / -0.143
Signal Systematics	+0.012 / -0.014
MC stat	+0.198 / -0.250

Table 57: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and experimental and modelling systematic uncertainties only (“All Syst.”) is shown in terms of the total impact on the uncertainty on μ .

1484 Figure 205 presents postfit diboson BDT distribution from the VZ diboson fit for the 1-lepton channel.

Not reviewed, for internal circulation only

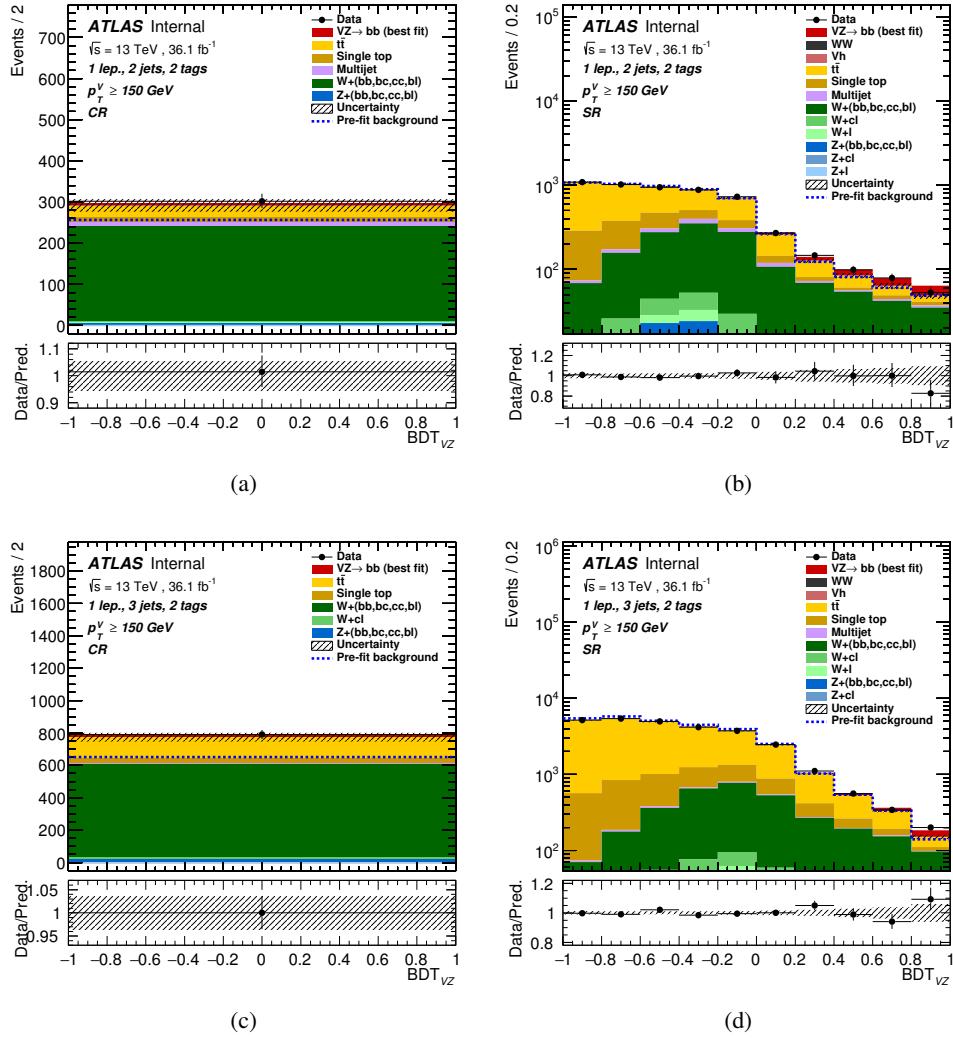


Figure 205: Unblinded VZ BDT distribution in the 1-lepton channel with BDT transformation D applied, after running the unconditional fit to data. The 2-tag 2-jet, $W + hf$ CR (a), SR (b), and the 2-tag 3-jet $W + hf$ CR (c), SR (d) categories are shown.

1485 **L.3. 2-lepton**

1486 **L.3.1. Prefit BDT distribution**

1487 Figure 206 presents prefit distributions of the diboson MVA.

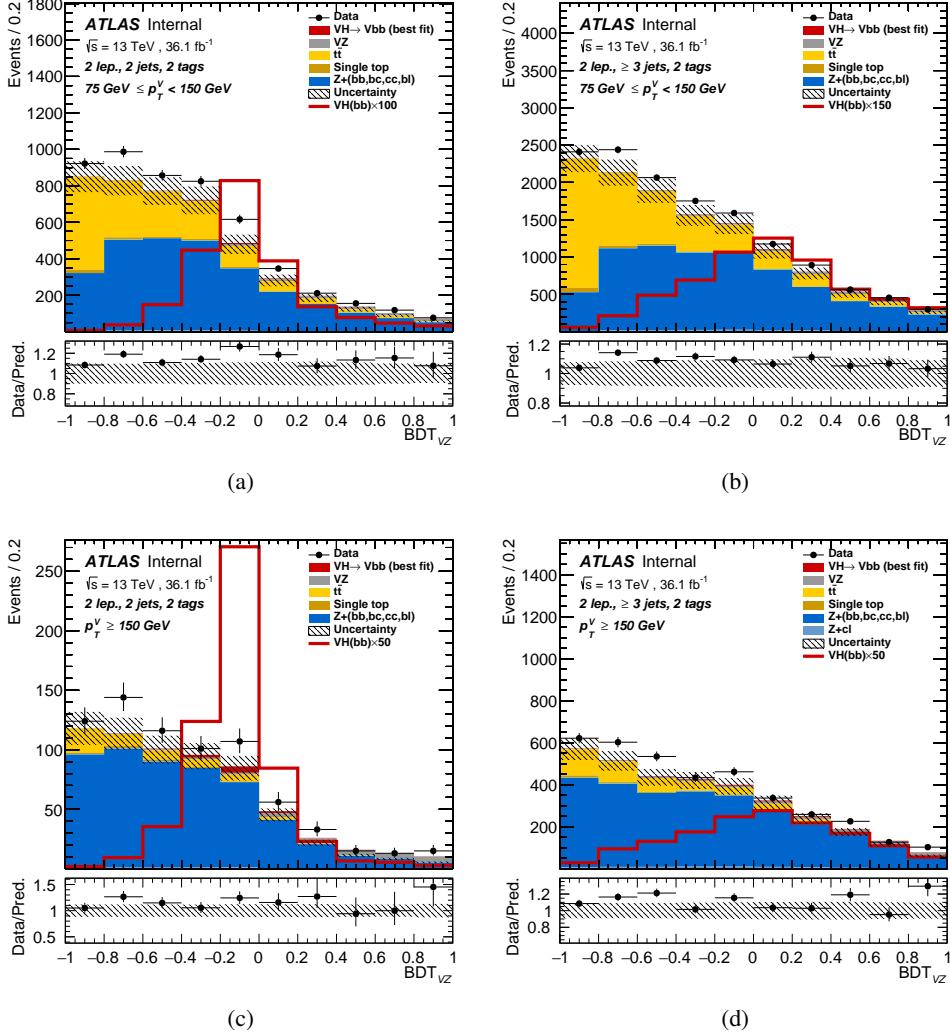
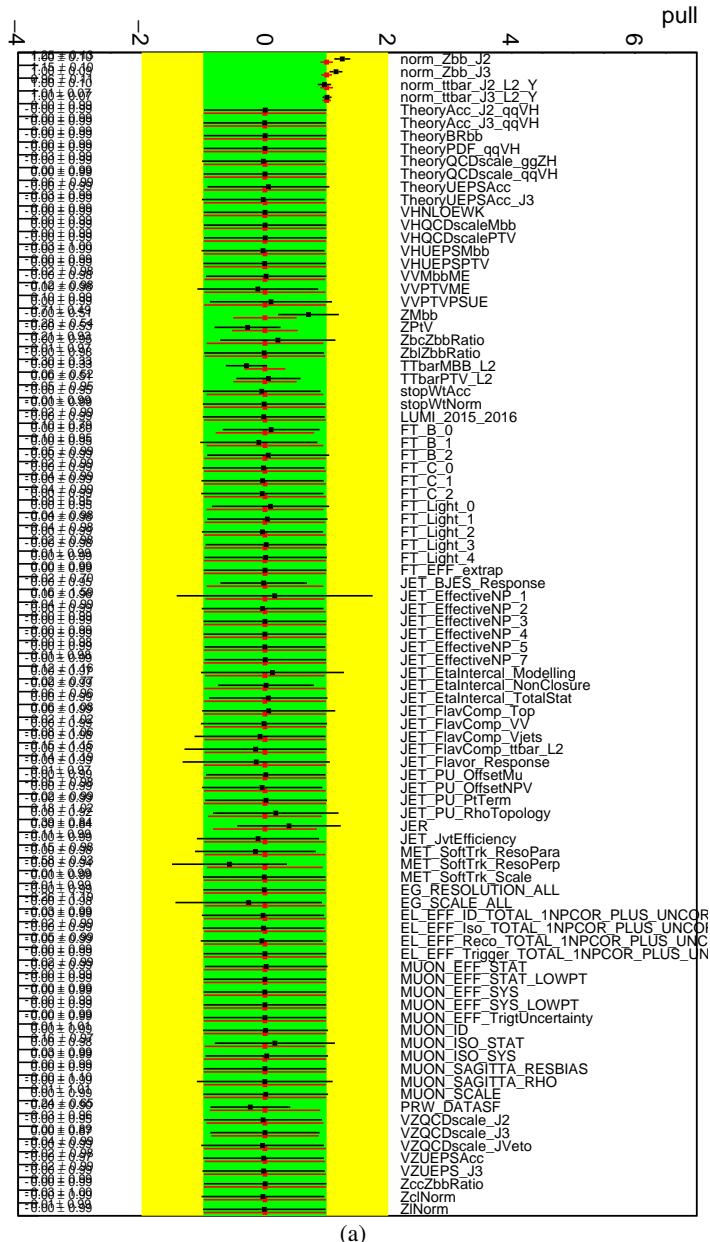


Figure 206: prefit BDT distribution in the 2-lepton channel with BDT transformation D applied. The $75 < p_T^Z < 150$, 2-tag 2-jet, (a), 2-tag 3+jet (b), and $150 < p_T^Z$, 2-tag 2jet (c), 2-tag 3+jet (d) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

1488 **L.3.2. Pull distribution, correlation and ranking check**

1489 Figure 207 shows pull distributions for all NPs in the 2-lepton unconditional fit to data in the VH fit (red),
1490 and the VZ fit (black).

Not reviewed, for internal circulation only



(a)

Figure 207: Pull plot comparing all NPs in the 2-lepton unconditional fit to the Asimov data (red), and unconditional fit to data (black).

1491 Figure 208 shows correlation between nuisance parameters. The plots shows the NPs having high
 1492 correlation (more than 20%). Figure 209 presents the ranking of nuisance parameters, based on their
 1493 impact on μ .

Not reviewed, for internal circulation only

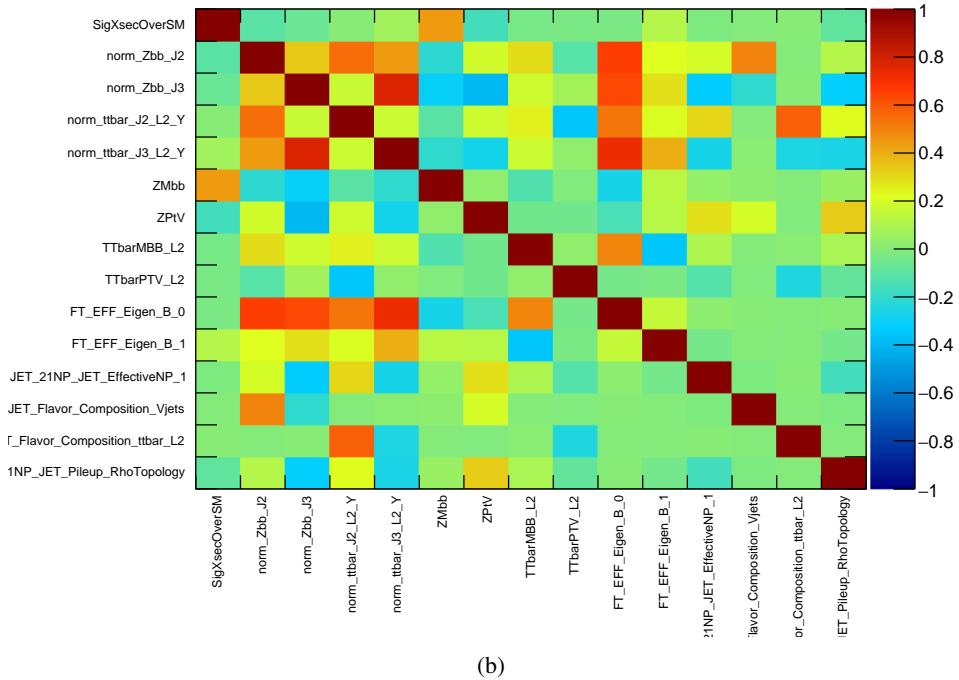
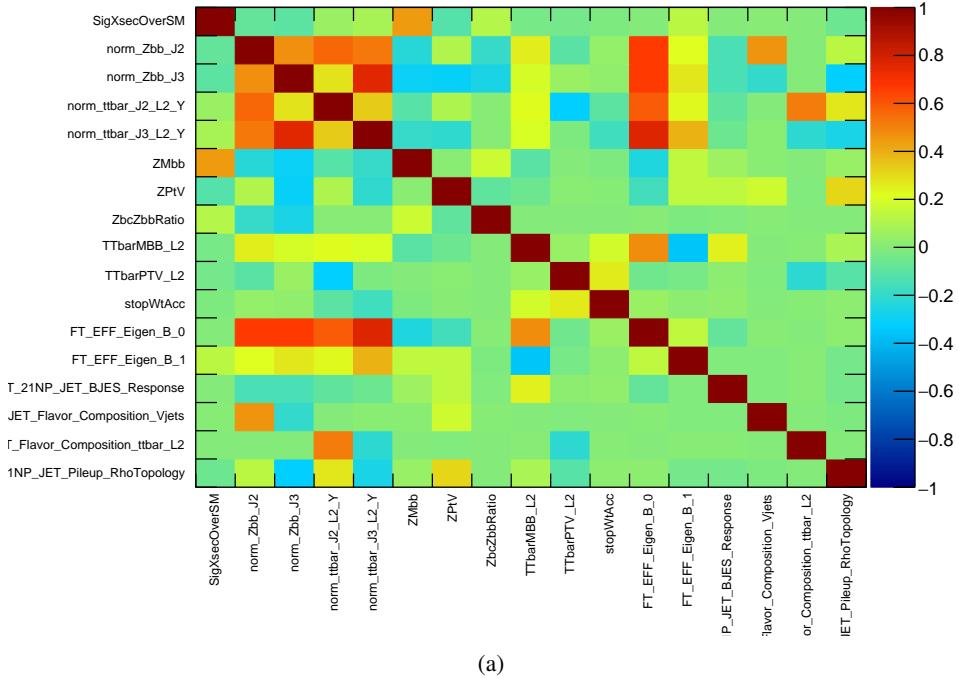


Figure 208: Highest correlation plots in the 2-lepton (a) Asimov unconditional fit. (b) Data unconditional fit

Not reviewed, for internal circulation only

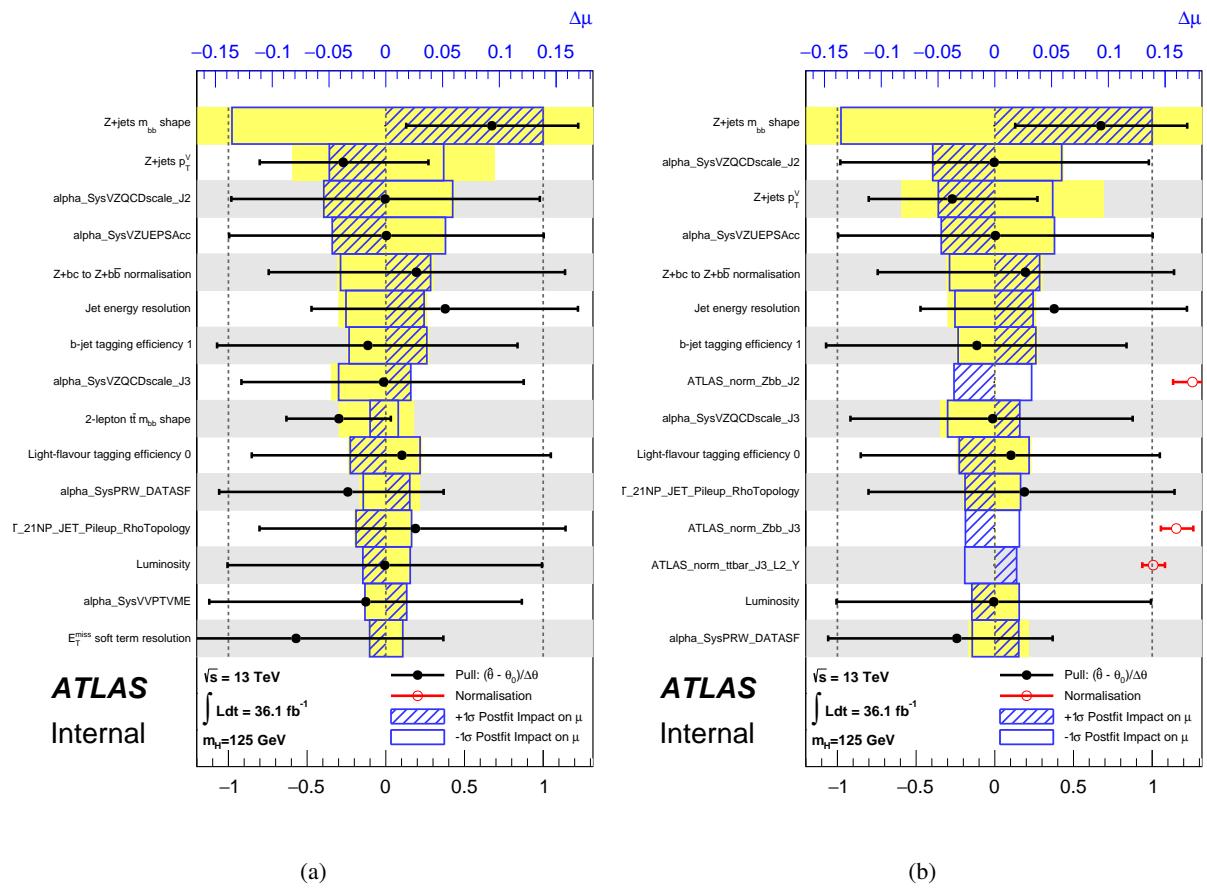


Figure 209: The nuisance parameter ranking for the 2-lepton channel (a) prefit and (b) unconditional fit to data.

1494 **L.3.3. Results**

1495 Figure 210 and 211 presents postfit BDT distribution from the VZ diboson fit for the 2-lepton channel.

Not reviewed, for internal circulation only

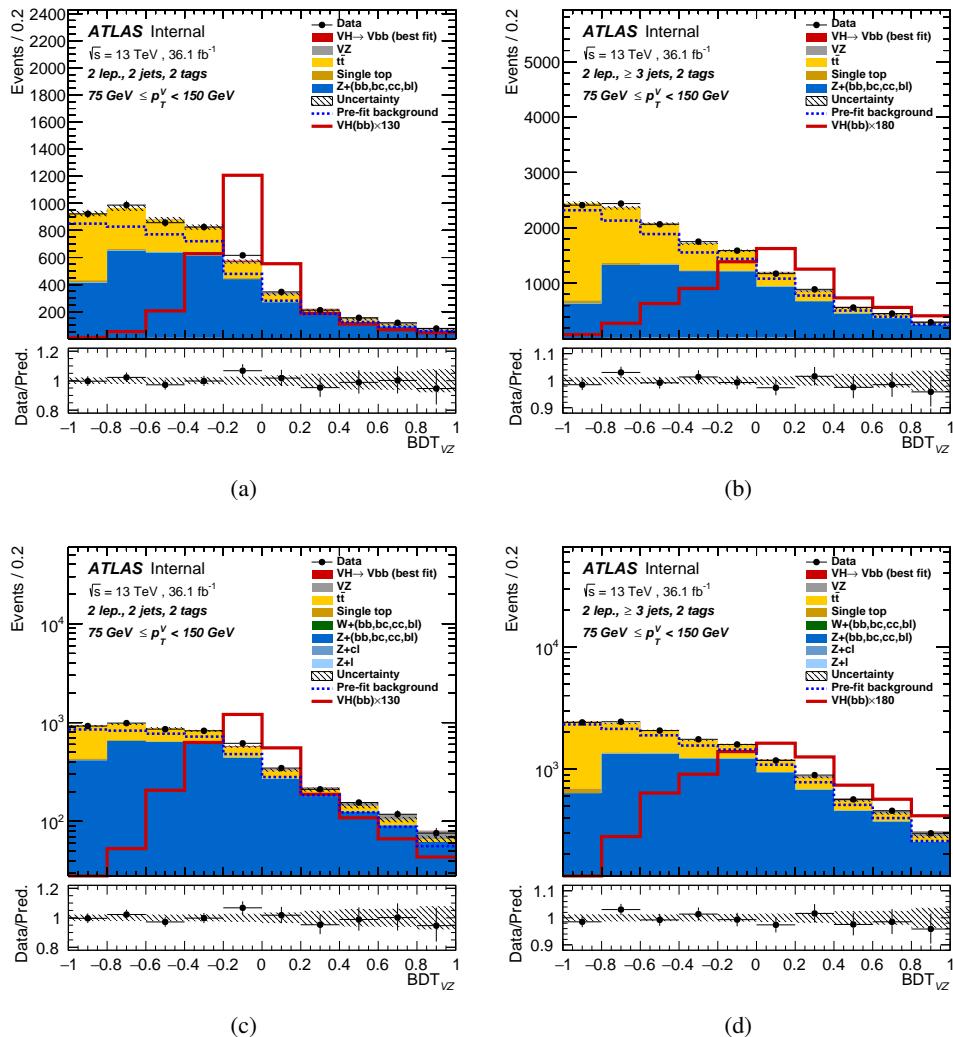


Figure 210: postfit BDT distribution in the 2-lepton channel with BDT transformation D applied, after running the $\mu = 1$ conditional fit to data. The $75 < p_T^Z < 150$, 2-tag 2-jet, (a), 2-tag 3+jet (b). Postfit BDT distributions in log-scale are shown in (c) and (d).

1496 Table 58 shows the breakdown of the uncertainties from data statistics and systematic uncertainties
1497 performed the unconditional fit to data .

Not reviewed, for internal circulation only

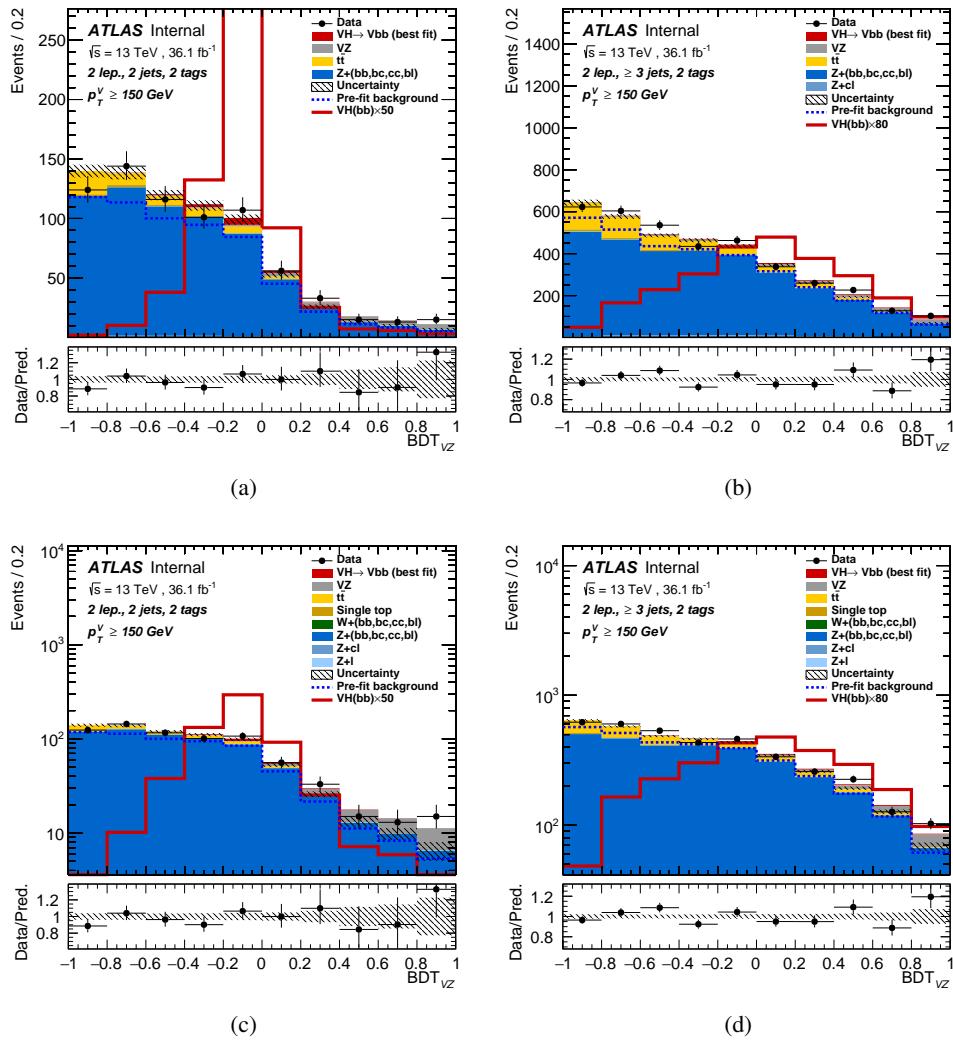


Figure 211: postfit BDT distribution in the 2-lepton channel with BDT transformation D applied, after running the $\mu = 1$ conditional fit to data. The $150 < p_T^V \leq 170$ GeV, 2-tag 2-jet, (a), 2-tag 3+jet (b). Postfit BDT distributions in log-scale are shown in (c) and (d).

1498 L.3.4. Comparison to VH Fit

1499 Figure 212 shows pull distributions for all NPs in the 2-lepton conditional $\mu = 1$ fit to data in the VH fit
1500 (black), and the unconditional VH fit (red).

Set of nuisance parameters	Impact on error
Total	+0.319 / -0.291
DataStat	+0.190 / -0.186
FullSyst	+0.256 / -0.224
Floating normalizations	+0.090 / -0.080
All normalizations	+0.090 / -0.080
All but normalizations	+0.226 / -0.191
Jets MET	+0.060 / -0.044
BTag	+0.063 / -0.054
Leptons	+0.016 / -0.009
Luminosity	+0.038 / -0.020
Diboson	+0.042 / -0.019
Zjets	+0.152 / -0.143
Wjets	+0.000 / -0.000
Model ttbar	+0.043 / -0.023
Model Single Top	+0.000 / -0.004
Signal Systematics	+0.003 / -0.003
MC stat	+0.092 / -0.091

Table 58: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and experimental and modelling systematic uncertainties only (“All Syst.”) is shown, as well as the fitted $\hat{\mu}$ extracted from the 2-lepton unconditional fit on the data.

Not reviewed, for internal circulation only

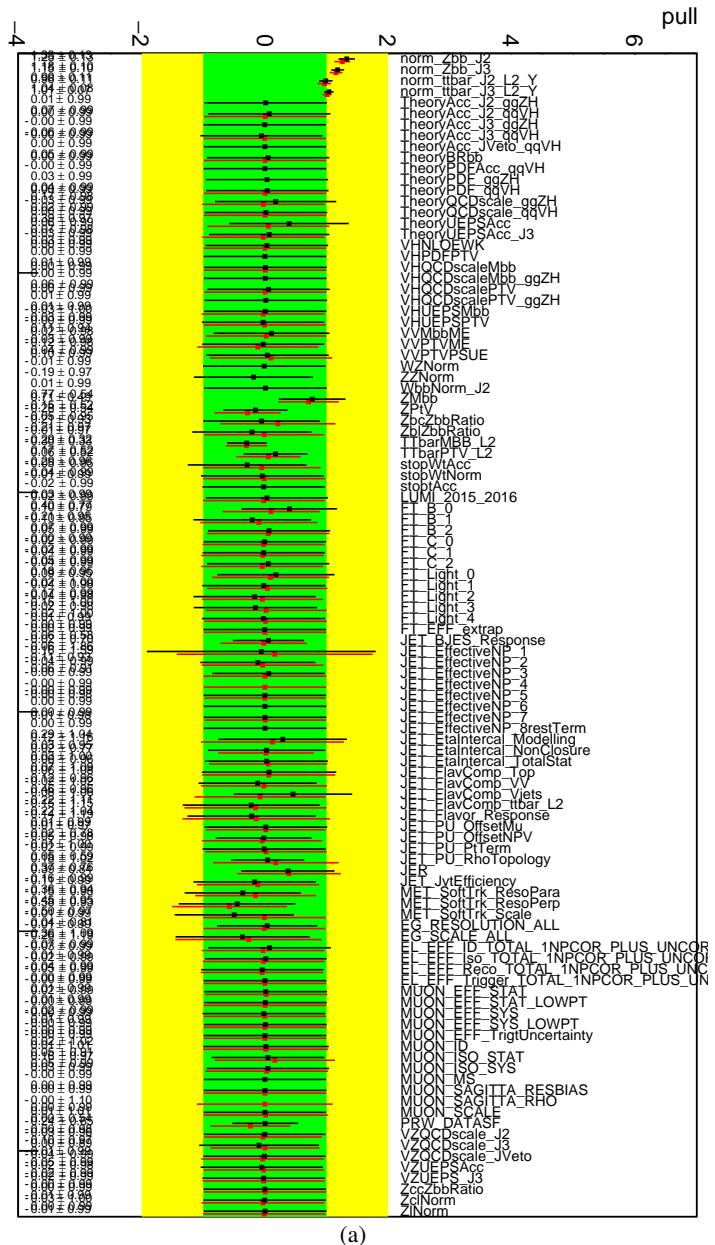


Figure 212: Pull distribution of all NPs in the 2-lepton conditional fit to data. red : VZ unconditional fit, black : VH conditional $\mu = 1$ fit

1501 **L.4. Detailed fit significances**

1502 Detailed fit significances are provided in Table 59, detailing the VZ significance in each fitting region.

Channel	n_{jet}	p_T^V [GeV]	Expected significance (Asimov)	Expected significance (data)	Measured Significance
0-lepton (SR)			3.77	3.72	4.17
1-lepton (SR+CR)			1.89	1.77	1.56
2-lepton (SR+CR)			3.79	3.47	3.41
0,1,2-lepton (SR+CR)			5.86	5.34	5.74

Table 59: Expected significances from a fit to an Asimov dataset corresponding to 36.1 fb^{-1} and from a fit to the data, treating VZ as the signal, expected significances from a conditional fit to data, and observed significances from an unconditional fit to data. Expected significances from individual regions are estimated separately.

1503 As a further cross-check, the diboson VZ fit has also been conducted using the cut-based selection, with a
 1504 fit to m_{bb} . Fit significances are provided in Table 60, detailing the VZ significance in each fitting region,
 1505 good compatibility with the MVA fit is observed.

Channel	n_{jet}	p_T^V [GeV]	Expected significance (Asimov)	Expected significance (data)	Measured Significance
0-lepton (SR)			4.20	3.80	3.73
1-lepton (SR+CR)			2.26	2.28	2.02
2-lepton (SR+CR)			3.74	3.65	3.45
0,1,2-lepton (SR+CR)			5.93	5.98	5.87

Table 60: Expected significances from a fit to an Asimov dataset corresponding to 36.1 fb^{-1} and from a fit to the data, treating VZ as the signal in the cut-based analysis. Expected significances from a conditional fit to data, and observed significances from an unconditional fit to data. Expected significances from individual regions are estimated separately.

1506 M. Cross-check: dijet mass analysis

1507 As a cross-check of the fit to the MVA output, a fit is conducted to the m_{bb} distribution. As described
 1508 in Section 3.2.5, a number of modifications are made to the event selection to optimise the sensitivity.
 1509 Furthermore, p_T^V is split into regions of [150, 200] GeV and $[200, \infty]$ GeV in all lepton channels.

1510 M.1. 0-lepton

1511 M.1.1. Prefit m_{bb} distribution

1512 Figures 213 presents prefit distributions of the cut-based m_{bb} distributions in the 0-lepton signal.

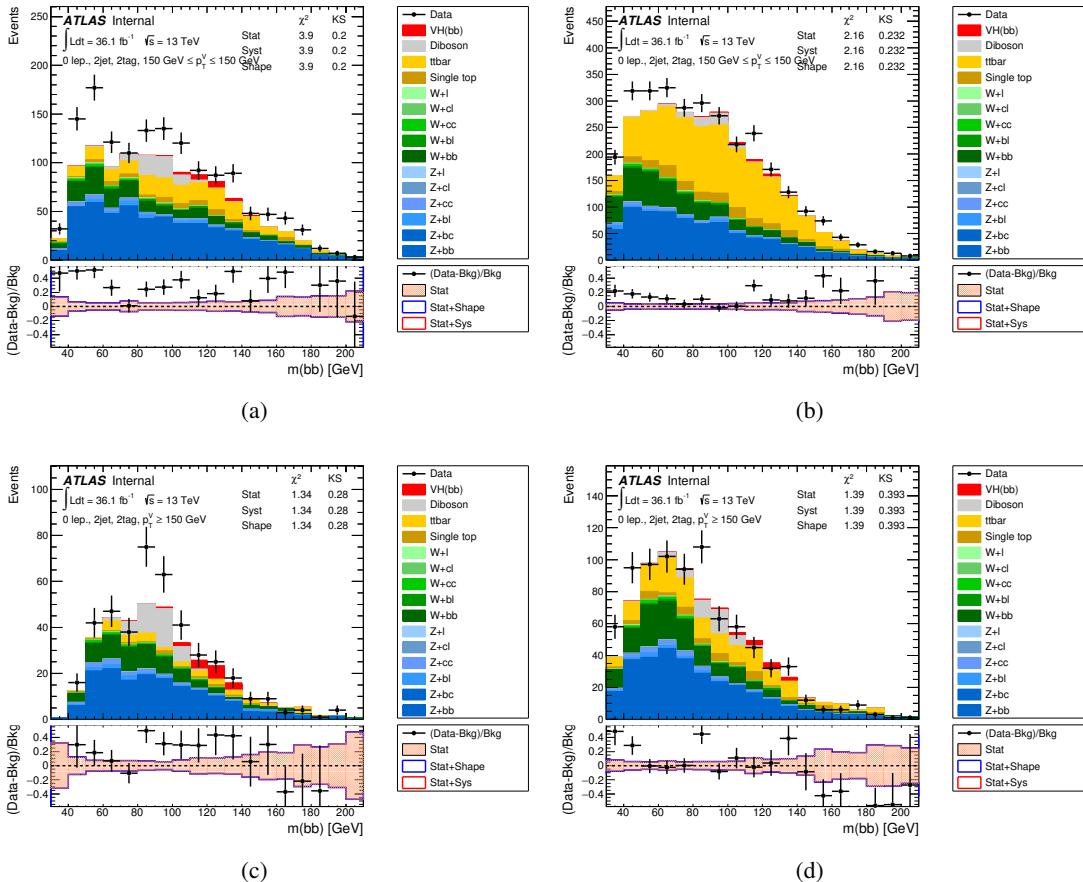


Figure 213: Pre-fit m_{bb} distributions in the 0-lepton channel. The 2-tag 2-jet SR (a), and the 2-tag 3-jet SR (b) categories are shown for the $150 < p_T^V < 200$ GeV region. The 2-tag 2-jet SR (c), and the 2-tag 3-jet SR (d) categories are shown for the $200 < p_T^V$ GeV region.

1513 **M.1.2. Comparison to VH Fit**

1514 Figure 214 shows pull distributions for all NPs in the 0-lepton unconditional fit to Vh BDT and the
 1515 unconditional fit to m_{bb} in the cut-based analysis.

Not reviewed, for internal circulation only

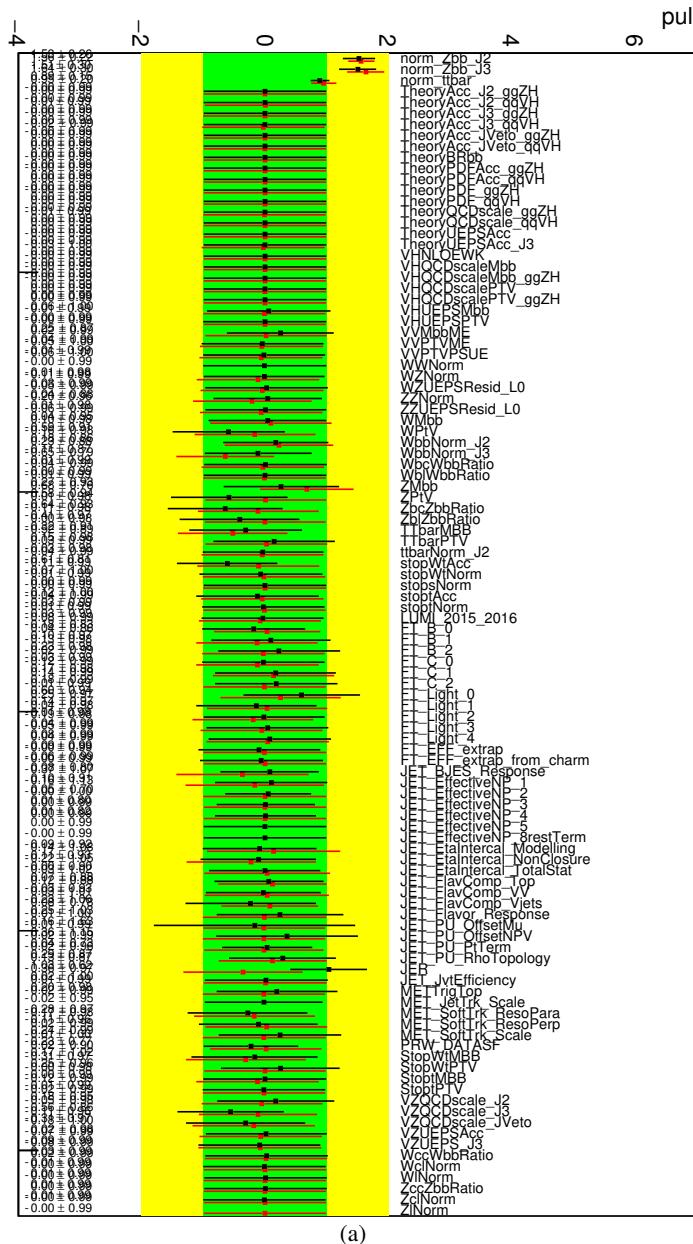


Figure 214: Comparison of the pulls in the 0-lepton unconditional VH BDT (red) and cut-based m_{bb} fits (black) to data.

1516 **M.1.3. Results**

1517 Table 61 shows the breakdown of the uncertainties from data statistics and systematic uncertainties in the
 1518 unconditional fit.

1519 Figure 215 shows correlation between nuisance parameters. Figure 216 presents the ranking of nuisance
 1520 parameters, based on their impact on μ . Figures 217 show postfit m_{bb} distribution from the cut-based
 1521 analysis fit for the 0-lepton channel.

Table 61: The breakdown of the uncertainties in the 0 lepton cut-based analysis m_{bb} unconditional fit, coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and experimental and modelling systematic uncertainties only (“All Syst.”) is shown in terms of the total impact on the uncertainty on μ .

Set of nuisance parameters	Impact on error
Total	+0.640 / -0.636
DataStat	+0.425 / -0.409
FullSyst	+0.478 / -0.487
Floating normalizations	+0.179 / -0.227
All normalizations	+0.197 / -0.237
All but normalizations	+0.389 / -0.405
Jets MET	+0.090 / -0.100
BTag	+0.122 / -0.152
Leptons	+0.000 / -0.000
Luminosity	+0.010 / -0.026
Diboson	+0.081 / -0.069
Zjets	+0.203 / -0.253
Wjets	+0.122 / -0.121
Model ttbar	+0.138 / -0.162
Model Single Top	+0.055 / -0.060
Model Multi Jet	+0.000 / -0.000
Signal Systematics	+0.148 / -0.120
MC stat	+0.285 / -0.310

Not reviewed, for internal circulation only

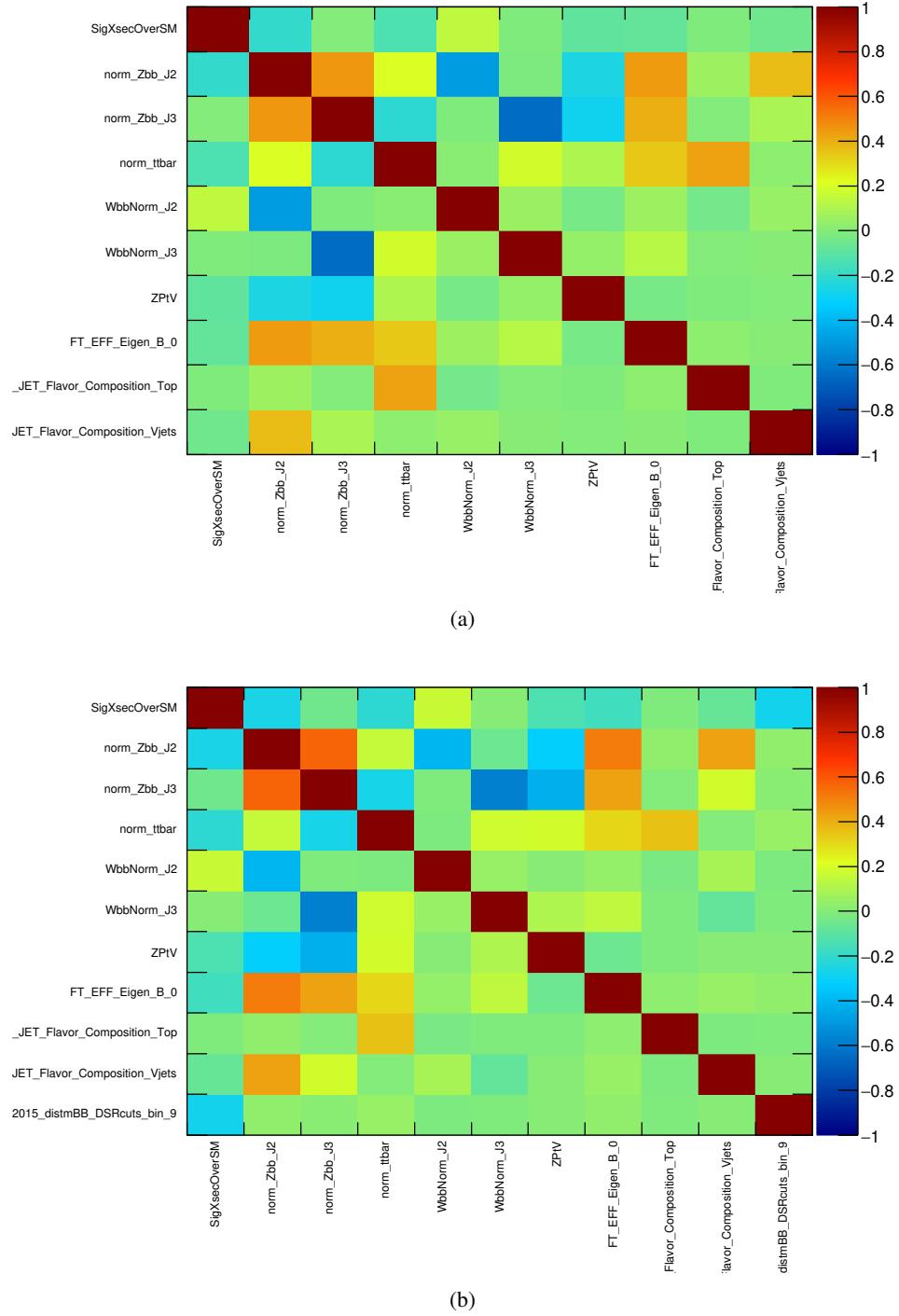


Figure 215: Highest correlations in the 0-lepton unconditional fits (a) Asimov and (b) Data

Not reviewed, for internal circulation only

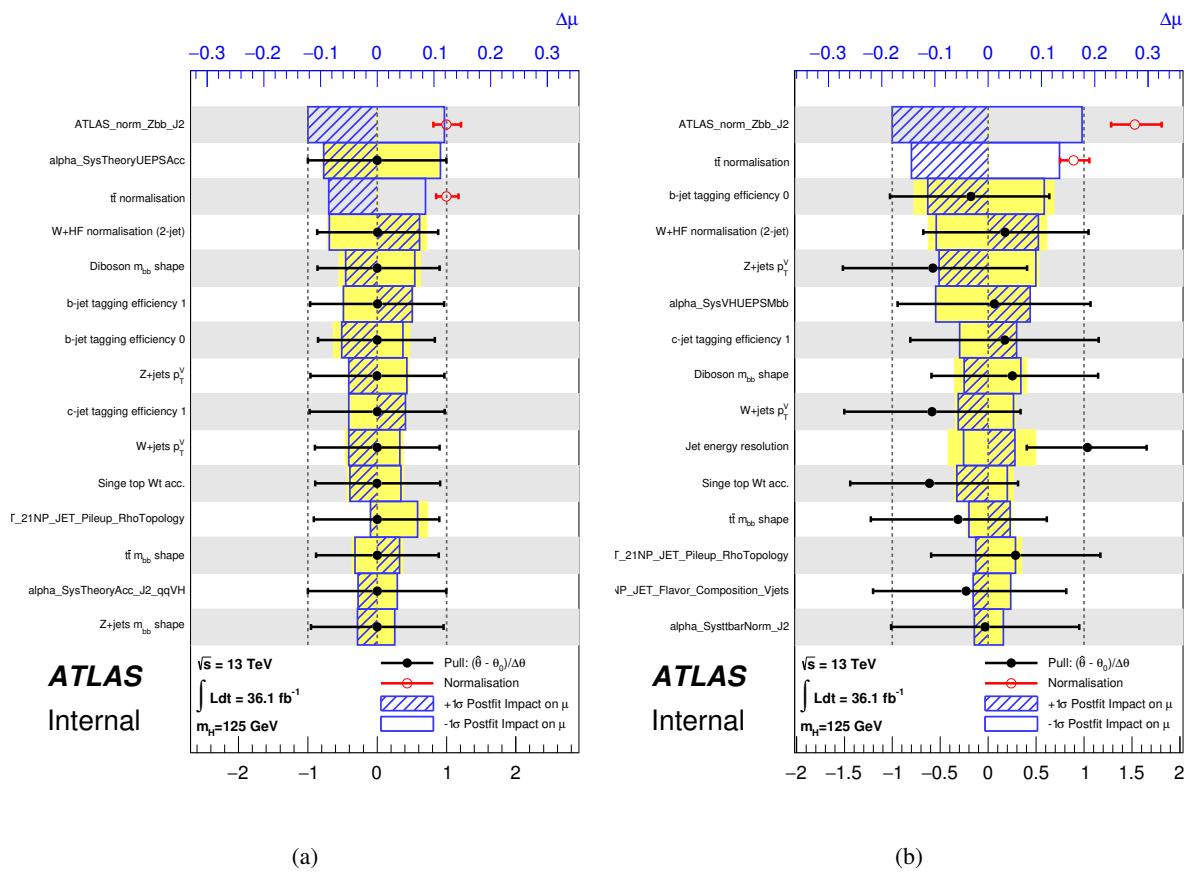


Figure 216: The nuisance parameter ranking for the 0-lepton channel (a) prefit and (b) data postfit unconditional.

Not reviewed, for internal circulation only

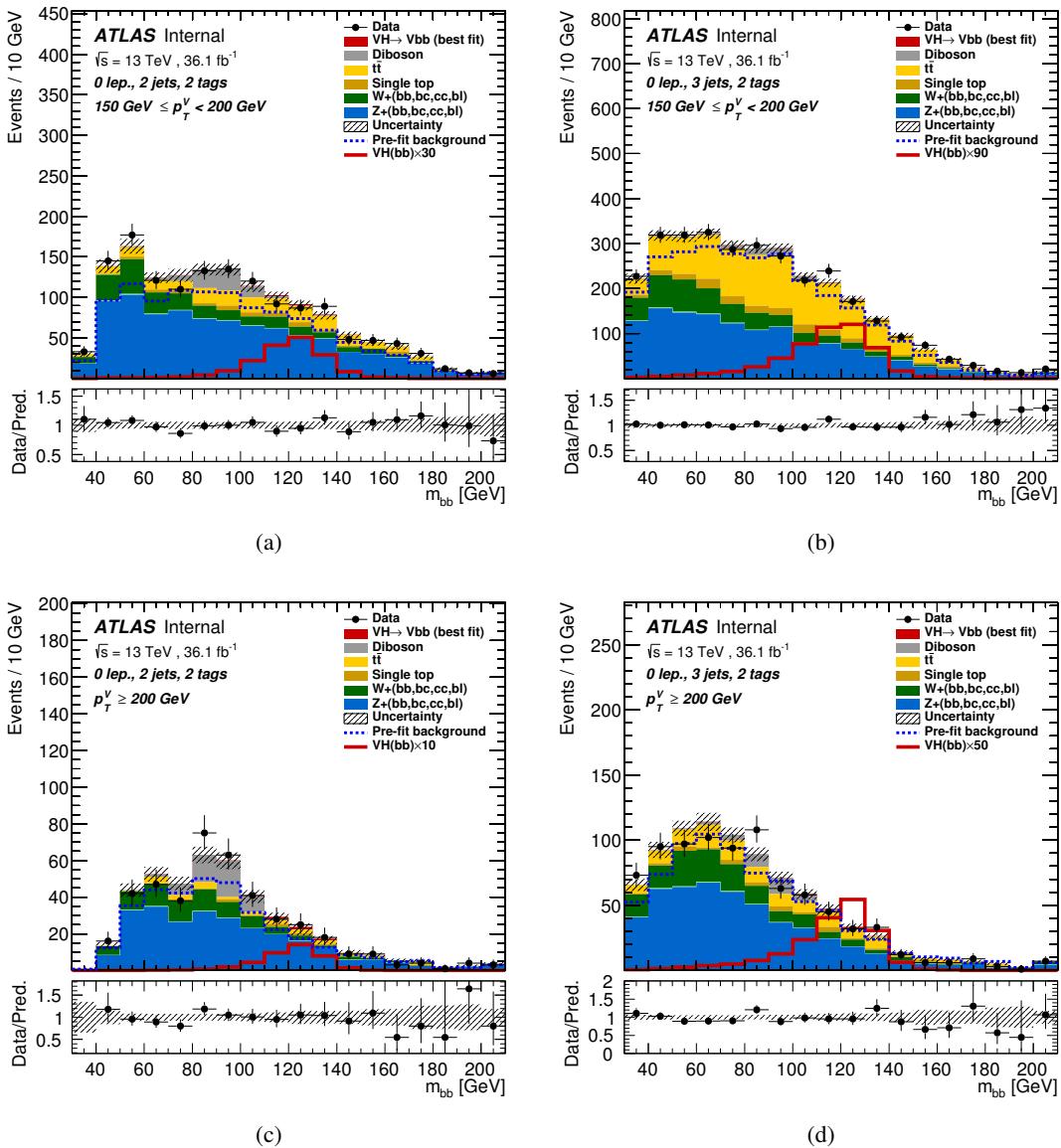


Figure 217: Post-fit m_{bb} distribution in the 0-lepton channel, after running the unconditional fit to data. The 2-tag 2-jet SR (a), and the 2-tag 3-jet SR (b) categories are shown for the $150 < p_T^V < 200 \text{ GeV}$ region. The 2-tag 2-jet SR (c), and the 2-tag 3-jet SR (d) categories are shown for the $p_T^V \geq 200 \text{ GeV}$ region.

1522 **M.2. 1-lepton**

1523 **M.2.1. Prefit m_{bb} distribution**

1524 Figures 218 and 219 presents prefit distributions of the cut-based m_{bb} distributions in the 1-lepton signal
 1525 and control regions. Bins are set to have 10 GeV width. For the purposes of the cut-based analysis, the
 1526 $W + hf$ CR and signal region are merged together.

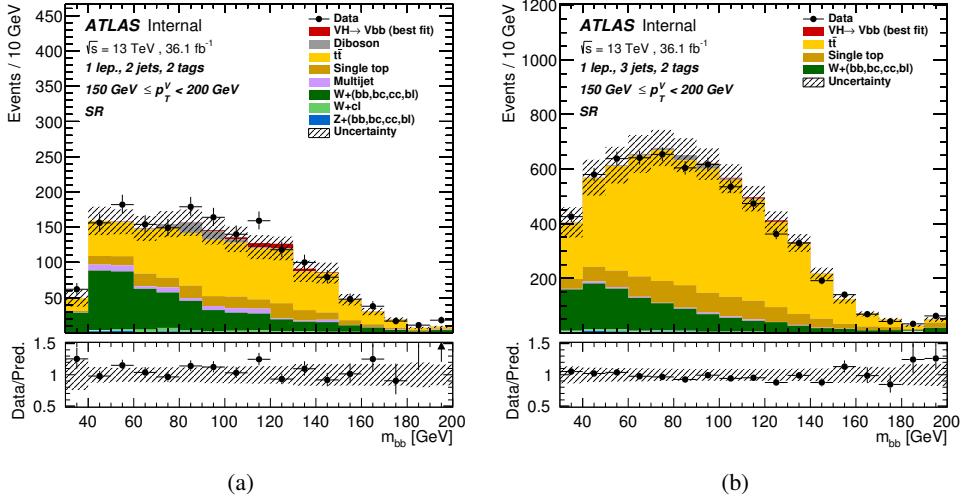


Figure 218: Unblinded prefit m_{bb} distribution in the 1-lepton channel, for the $150 < p_T^V < 200$ GeV region. The 2-tag 2-jet SR (a), and the 2-tag 3-jet SR (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

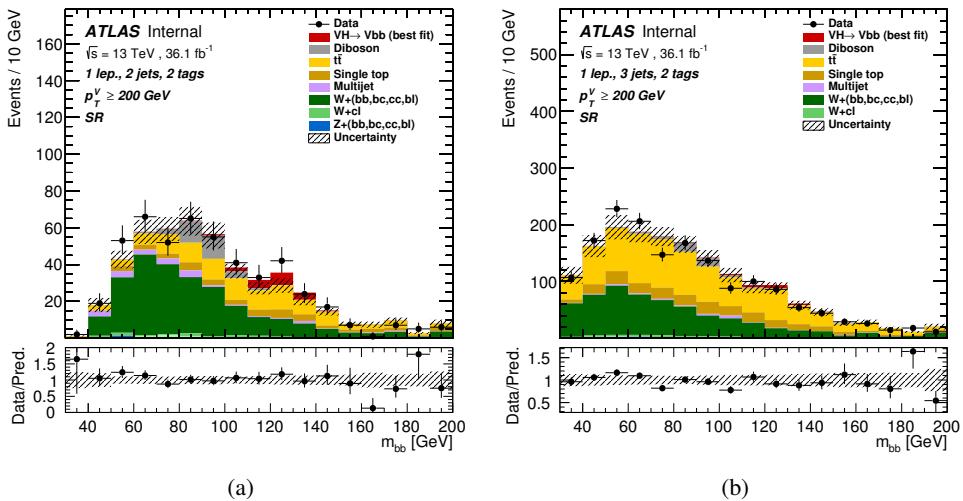
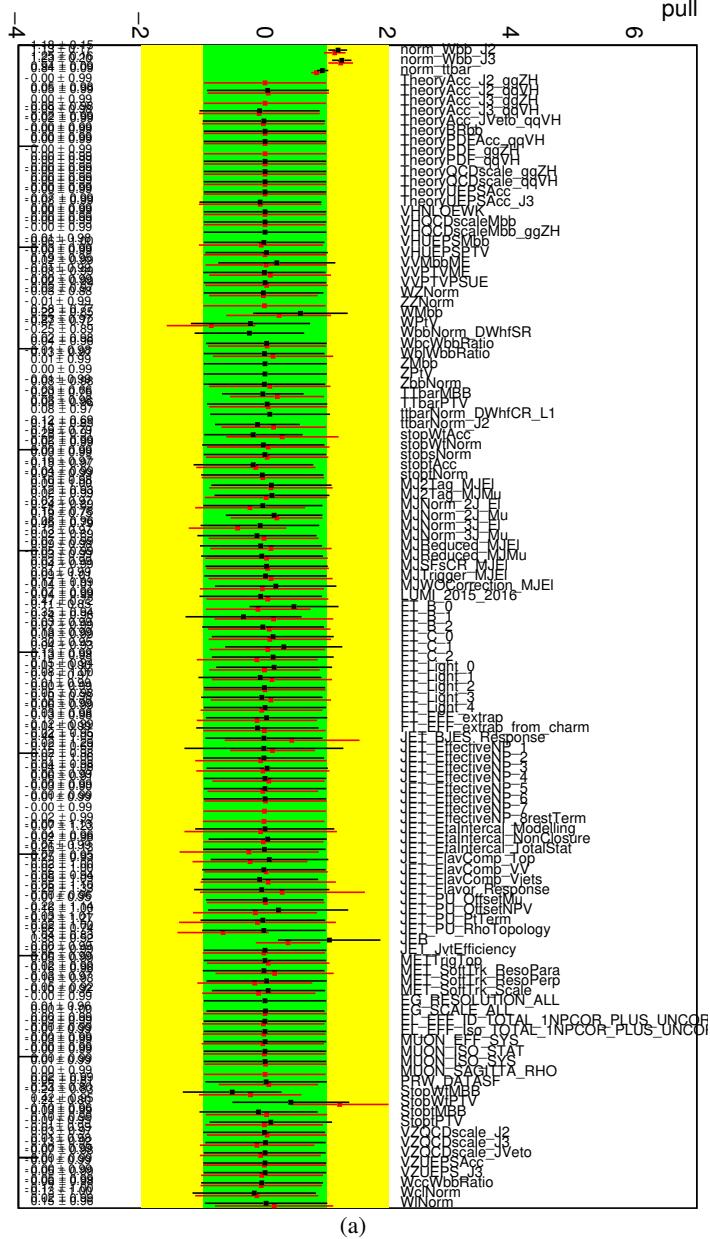


Figure 219: Unblinded prefit m_{bb} distribution in the 1-lepton channel, for the $200 < p_T^V$ GeV region. The 2-tag 2-jet SR (a), and the 2-tag 3-jet SR (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

1527 **M.2.2. Comparison to VH Fit**

1528 In this section, the results from the unconditional fit to data in the VH and m_{bb} fits are compared.

1529 Figure 220 shows pull distributions for all NPs in the 1-lepton unconditional fit to data in the MVA VH fit
 1530 (black), and the m_{bb} fit (red).



(a)

Figure 220: Pull plot comparing all NPs in the 1-lepton unconditional fit to the VH signal (black), and m_{bb} fit (red).

1531 Figure 221 shows correlation between nuisance parameters in the m_{bb} fit. The plots shows the NPs having
 1532 high correlation (more than 20%). Figure 222 presents the ranking of nuisance parameters, based on their
 1533 impact on μ in the m_{bb} fits.

1534 **M.2.3. Results**

1535 Figures 223 and 224 presents post-fit cut-based m_{bb} distribution from the unconditional fit for the 1-lepton
1536 channel.

1537 Table 62 shows the breakdown of the uncertainties from data statistics and systematic uncertainties from
1538 the 1-lepton unconditional fit to data.

Set of nuisance parameters	Impact on error
Total	+0.904 / -0.787
DataStat	+0.548 / -0.522
FullSyst	+0.719 / -0.588
Floating normalizations	+0.168 / -0.172
All normalizations	+0.197 / -0.198
All but normalizations	+0.643 / -0.502
Jets MET	+0.211 / -0.148
BTag	+0.164 / -0.145
Leptons	+0.027 / -0.018
Luminosity	+0.071 / -0.023
Diboson	+0.123 / -0.063
Zjets	+0.007 / -0.007
Wjets	+0.189 / -0.160
Model ttbar	+0.153 / -0.193
Model Single Top	+0.125 / -0.118
Model Multi Jet	+0.065 / -0.073
Signal Systematics	+0.412 / -0.127
MC stat	+0.346 / -0.350

Table 62: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and experimental and modelling systematic uncertainties only (“All Syst.”) is shown in terms of the total impact on the uncertainty on μ from the 1-lepton unconditional fit to data.

Not reviewed, for internal circulation only

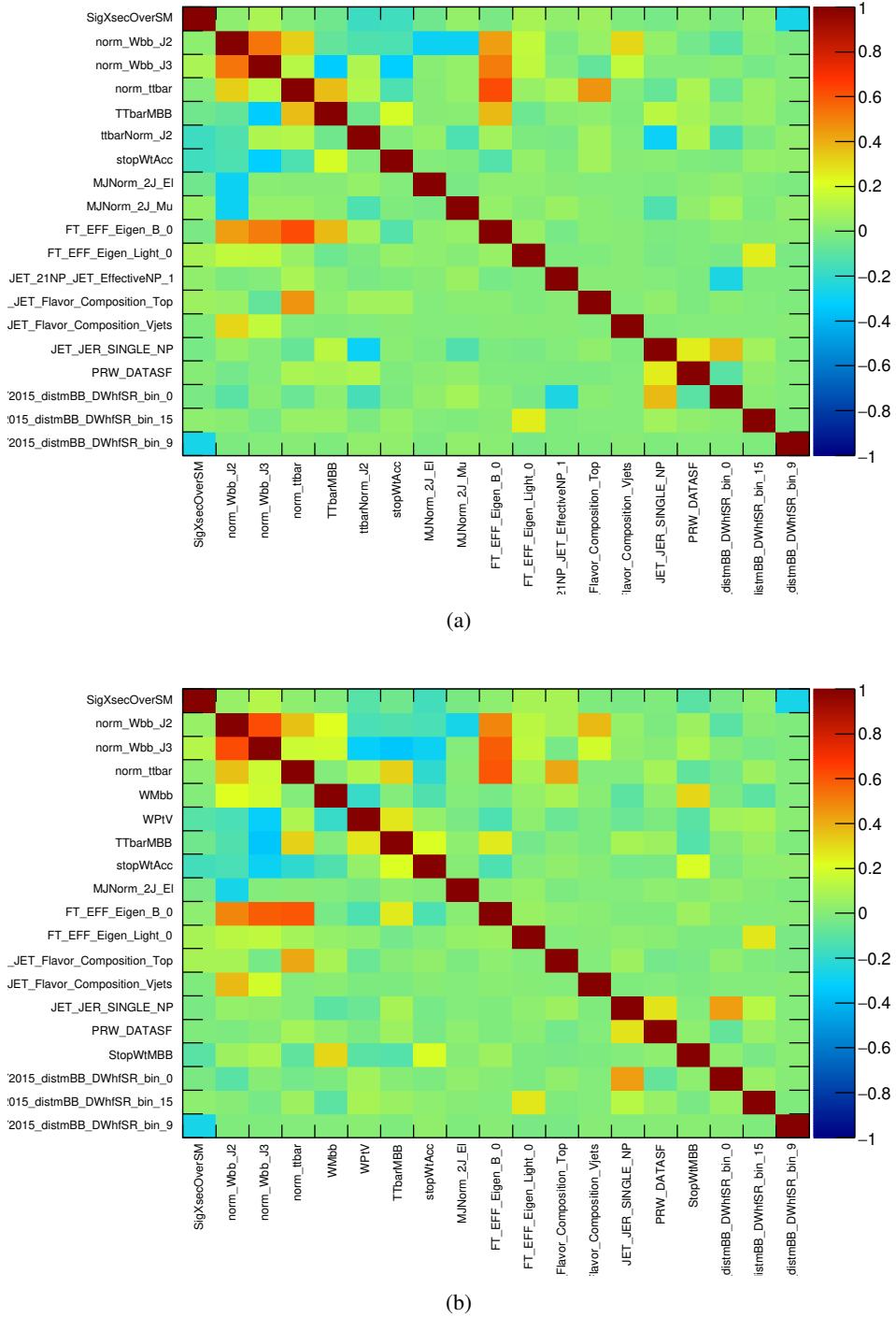


Figure 221: Correlation matrices from the unconditional 1-lepton m_{bb} fit to an Asimov dataset (a) and to 2015 and 2016 data (b).

Not reviewed, for internal circulation only

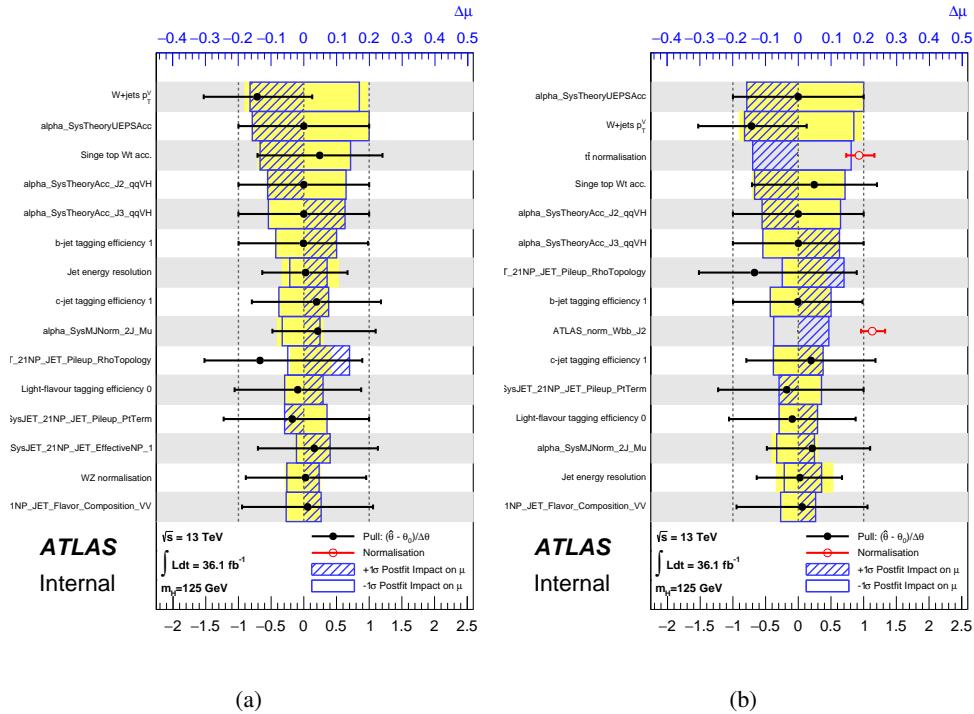


Figure 222: The post-fit impact nuisance parameter ranking for the 1-lepton channel fit based on prefit (a) and postfit (b) impact on $\hat{\mu}$ from the unconditional fit to data.

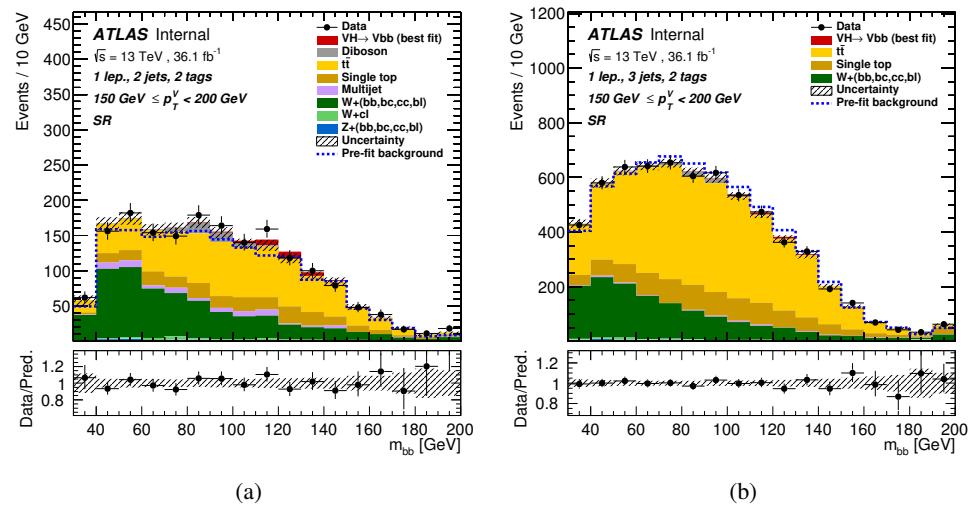


Figure 223: Unblinded m_{bb} distribution in the 1-lepton channel for the $150 < p_T^V < 200 \text{ GeV}$ region, after running the unconditional fit to data. The 2-tag 2-jet SR (a), and the 2-tag 3-jet SR (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

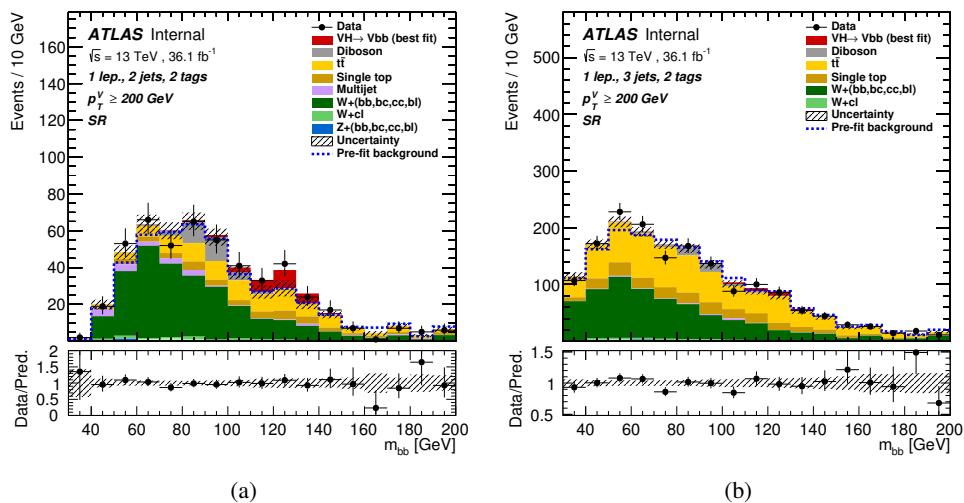


Figure 224: Unblinded m_{bb} distribution in the 1-lepton channel for the $200 < p_T^V$ GeV region, after running the unconditional fit to data. The 2-tag 2-jet SR (a), and the 2-tag 3-jet SR (b) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

1539 M.3. 2-lepton**1540 M.3.1. Prefit m_{bb} distribution**

1541 Figure 225 presents prefit distributions of dijet mass distribution in cut-based analysis.

1542 M.3.2. Pull distribution, correlation and ranking check

1543 Figure 226 shows pull distributions for all NPs in the unconditional fit to data in the MVA VH fit (black)
1544 and m_{bb} cut-based analysis (red).

1545 Figure 227 shows correlation between nuisance parameters. The plots shows the NPs having high
1546 correlation (more than 20%). Figure 228 presents the ranking of nuisance parameters, based on their
1547 impact on μ .

Not reviewed, for internal circulation only

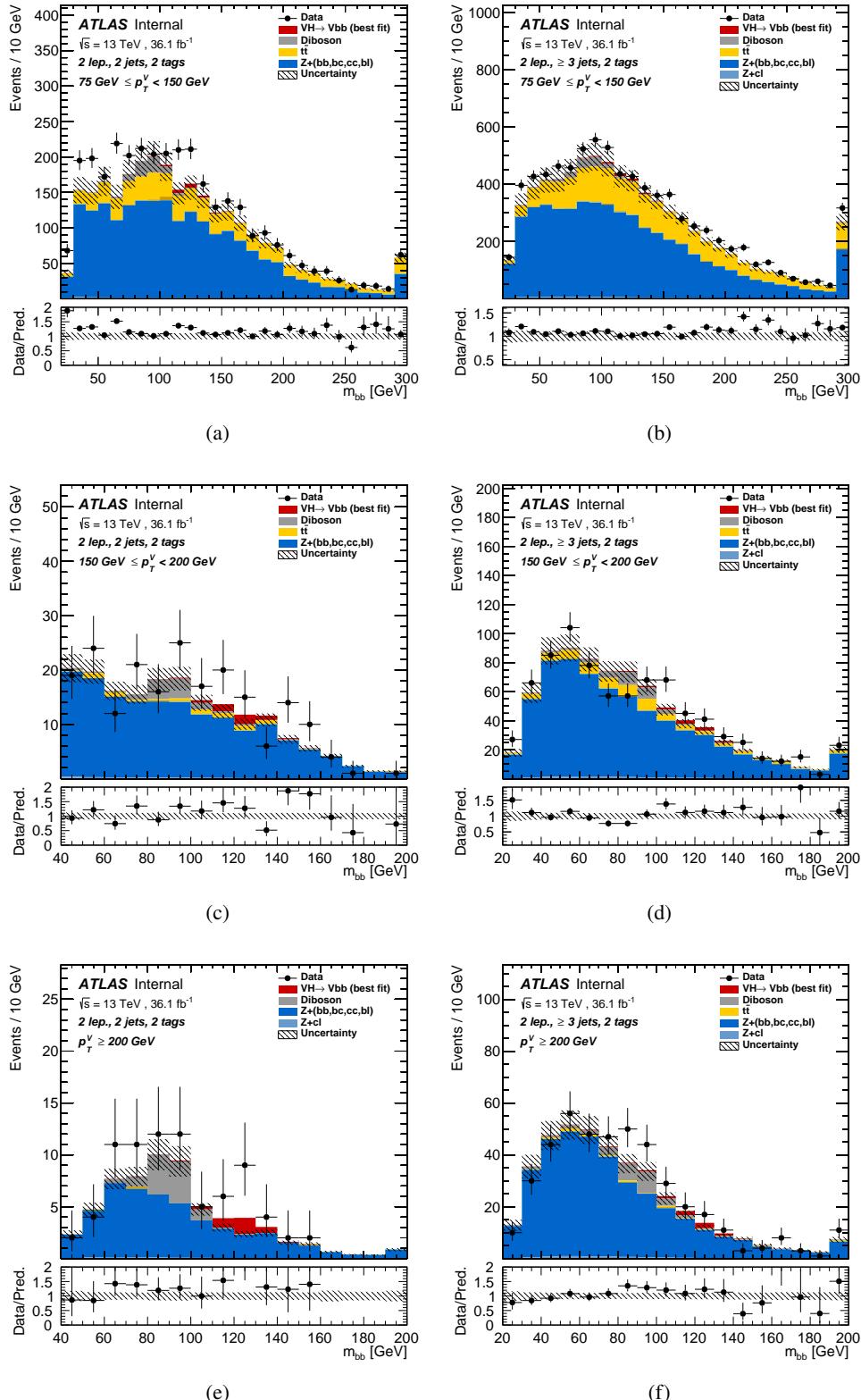


Figure 225: prefit m_{BB} distribution in the 2-lepton channel. The $75 < p_T^Z < 150$ GeV, 2-tag 2-jet, (a), 2-tag 3+jet (b), and $150 < p_T^Z < 200$ GeV, 2-tag 2jet (c), 2-tag 3+jet (d) and $200 \text{ GeV} < p_T^Z$, 2-tag 2jet (e), 2-tag 3+jet (f) categories are shown. The background and signal samples are normalized to the expected cross-section predictions.

Not reviewed, for internal circulation only

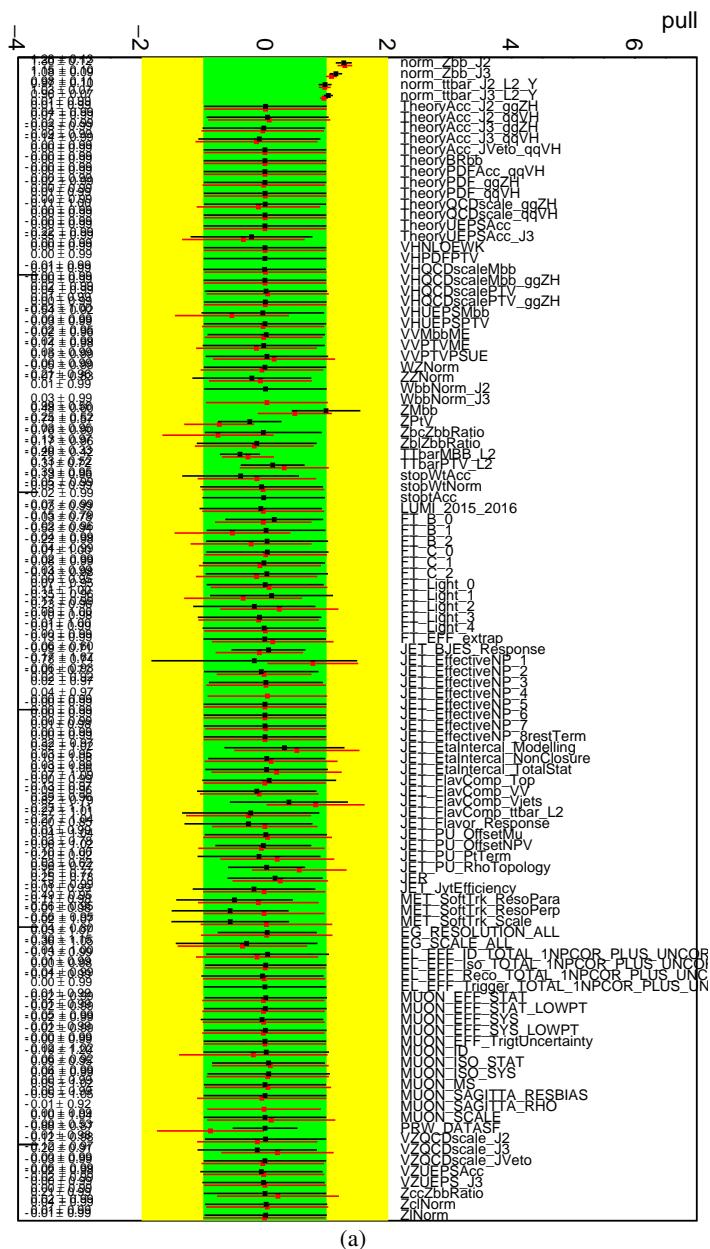


Figure 226: The nuisance parameter pulls and the free parameter scale factors corresponding to the unconditional fits performed on MVA (black) and cut-based m_{bb} analysis (red).

Not reviewed, for internal circulation only

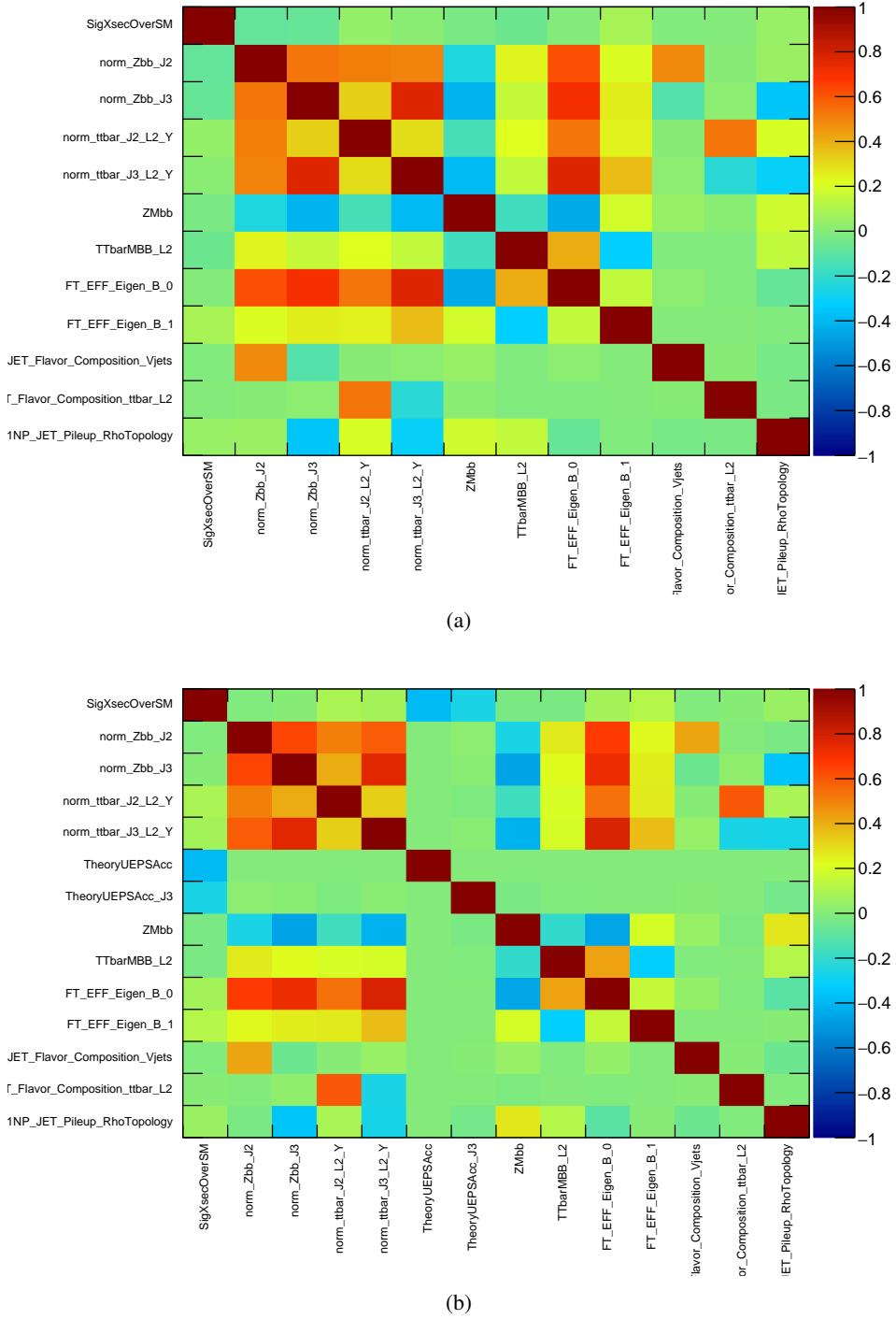


Figure 227: Highest correlation plots in the 2-lepton fit. (a) Asimov unconditional fit. (b) Data unconditional fit.

Not reviewed, for internal circulation only

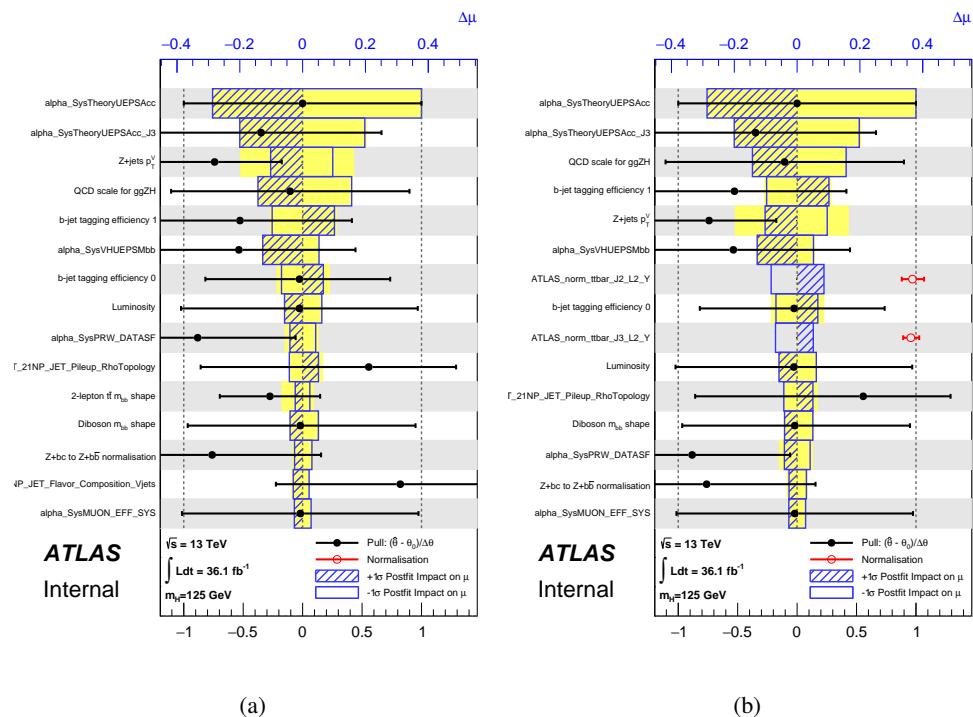


Figure 228: The nuisance parameter ranking for the 2-lepton channel unconditional fit (a) ordered by prefit impact (b) ordered by postfit impact.

1548 **M.3.3. Results**

1549 Figure 229 presents postfit m_{bb} distribution from the cut-based VH fit for the 2-lepton channel.

1550 Table 63 shows the breakdown of the uncertainties from data statistics and systematic uncertainties.

Set of nuisance parameters	Impact on error
Total	+0.972 / -0.793
DataStat	+0.616 / -0.591
FullSyst	+0.752 / -0.528
Floating normalizations	+0.175 / -0.122
All normalizations	+0.175 / -0.123
All but normalizations	+0.692 / -0.452
Jets MET	+0.079 / -0.049
BTag	+0.158 / -0.086
Leptons	+0.048 / -0.036
Luminosity	+0.099 / -0.052
Diboson	+0.074 / -0.058
Zjets	+0.120 / -0.127
Wjets	+0.003 / -0.003
Model ttbar	+0.146 / -0.072
Model Single Top	+0.002 / -0.004
Signal Systematics	+0.594 / -0.325
MC stat	+0.264 / -0.226

Table 63: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and experimental and modelling systematic uncertainties only (“All Syst.”) is shown, as well as the fitted $\hat{\mu}$ extracted from the 2-lepton unconditional fit on the data.

Not reviewed, for internal circulation only

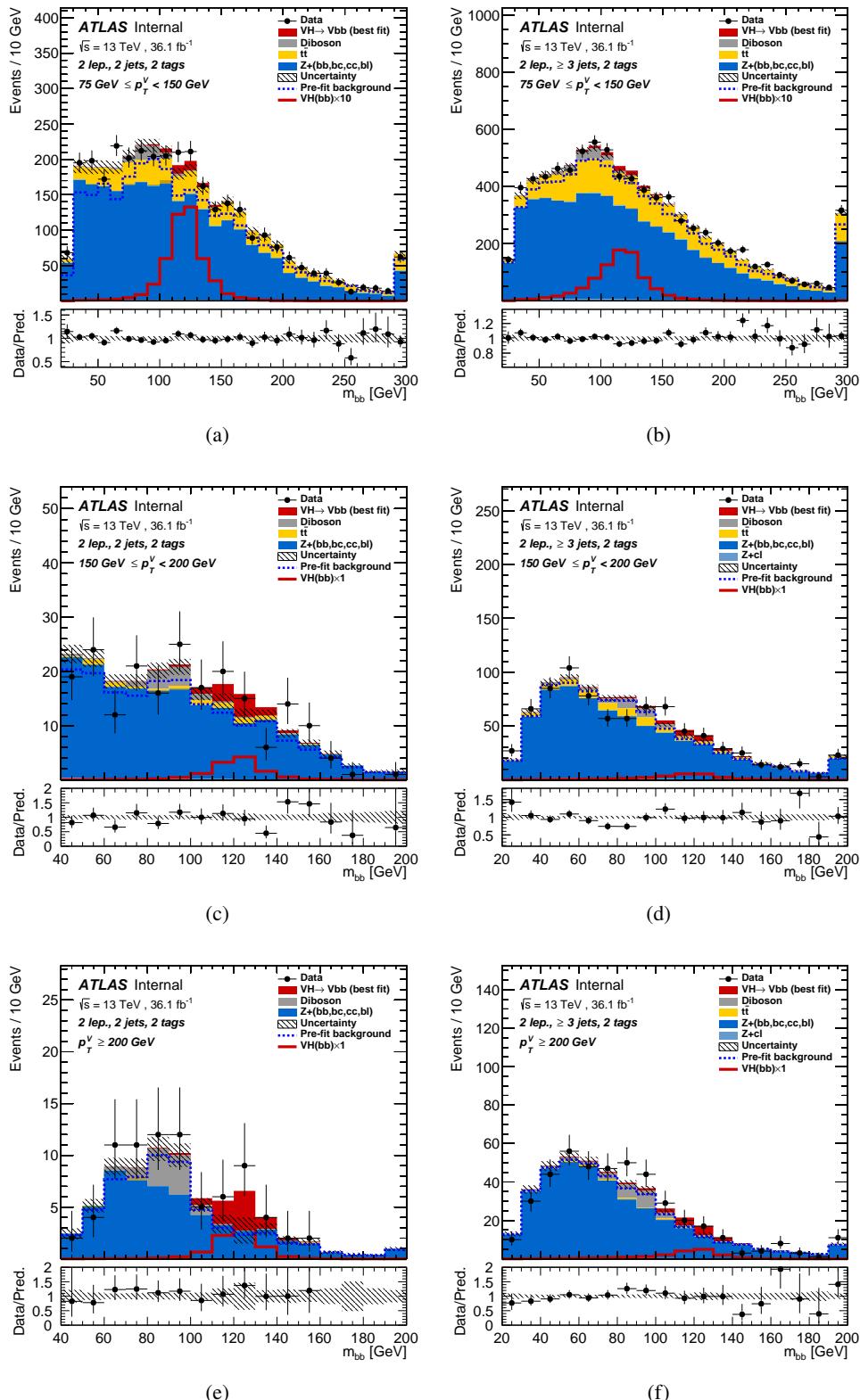


Figure 229: postfit m_{bb} distribution in the 2-lepton channel, after running an unconditional fit to data. The $75 < p_T^Z < 150$, 2-tag 2-jet, (a), 2-tag 3+jet (b), $150 < p_T^Z < 200$, 2-tag 2jet (c), 2-tag 3+jet (d), and $200 < p_T^Z$, 2-tag 2jet (e), 2-tag 3+jet (f) categories are shown.

1551 M.4. Combined Fit

M.4.1. Pull Comparison with the unconditional fit to the data in VH Cut-based analysis

Figure 230 shows pull distributions for all NPs in the combined unconditional fit to data in the MVA VH fit (black), and the unconditional fit to the data in the cut-based VH fit (red).

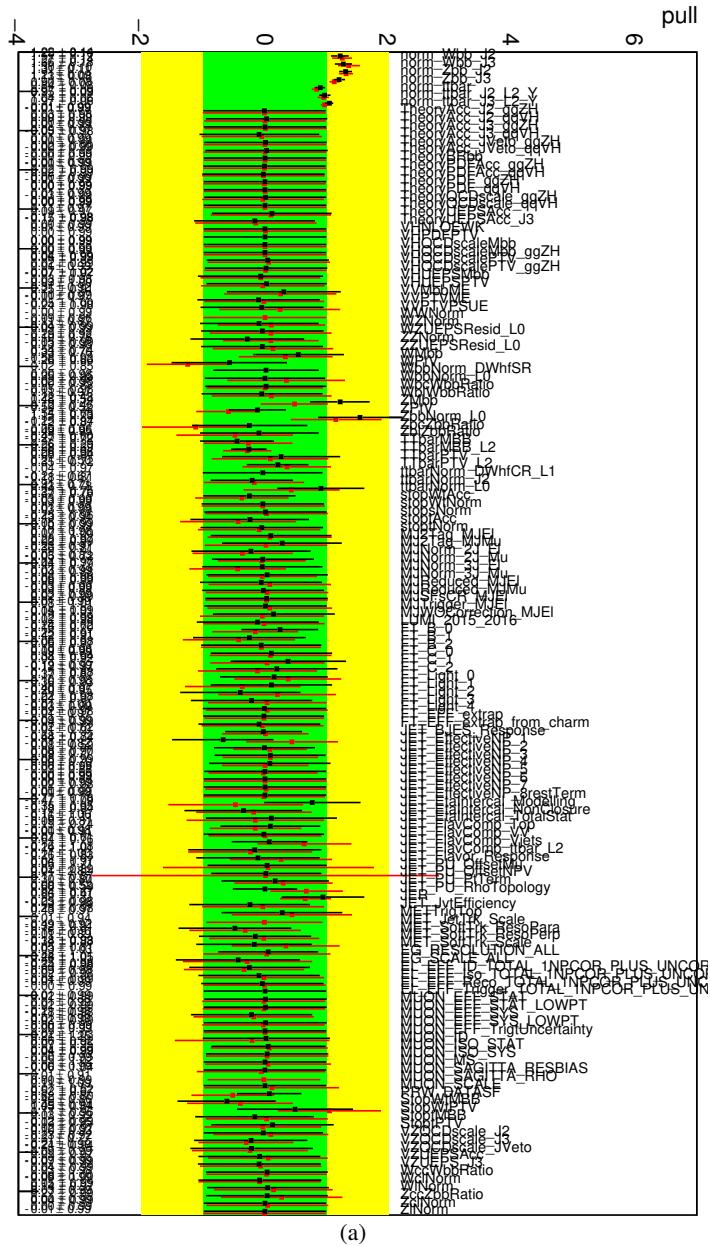


Figure 230: Pull plot comparing all NPs in the combined unconditional fit to the VH signal (black), and unconditional fit to the VH signal (red).

Figure 233 shows correlation between nuisance parameters in the m_{bb} fit. The plots shows the NPs having high correlation (more than 20%). Figure 234 presents the ranking of nuisance parameters, based on their

Not reviewed, for internal circulation only

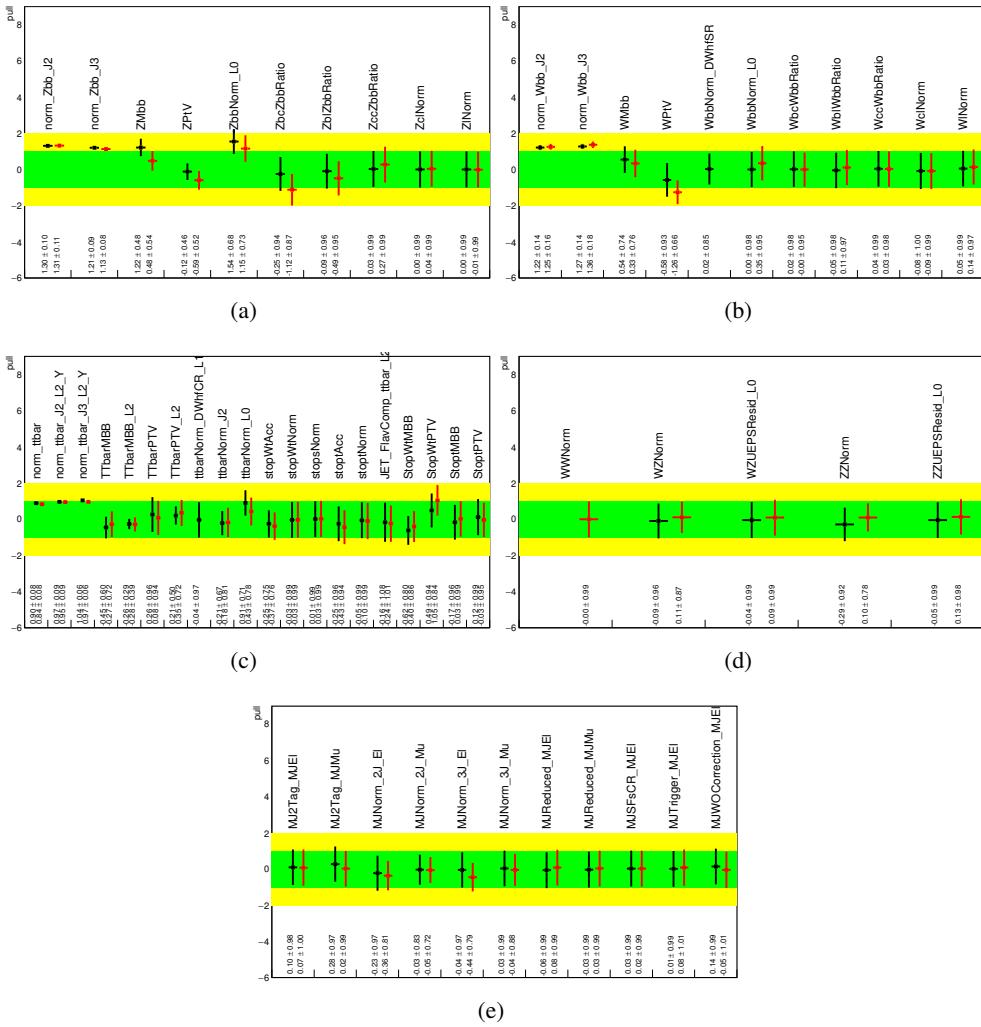


Figure 231: Pull distribution of modeling NPs in the combined unconditional fit. Black: the combined unconditional MVA fit to the VH signal, red: unconditional cut-based fit to the VH signal.

1557 impact on μ in the m_{bb} fits.

1558 M.4.2. Results

1559 Figures 235 to 237 presents post-fit cut-based m_{bb} distribution from the unconditional fit.

Table 64 shows the breakdown of the uncertainties from data statistics and systematic uncertainties from the combined, unconditional cut-based fit to data.

Not reviewed, for internal circulation only

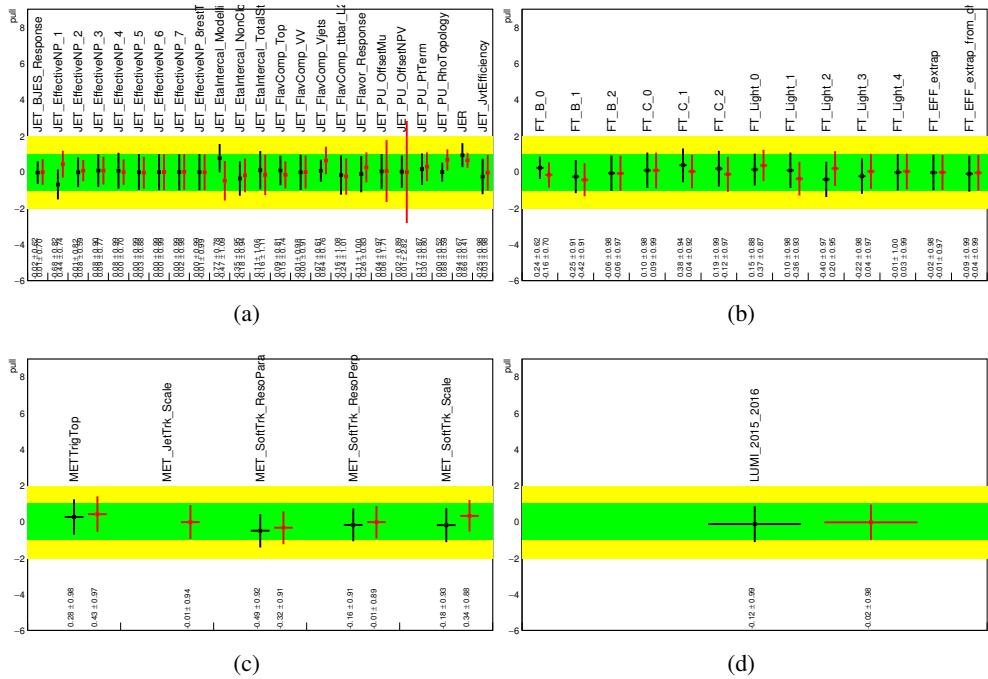


Figure 232: Pull distribution of experimental NPs in the combined unconditional fit. Black: the combined unconditional MVA fit to the VH signal, red: unconditional cut-based fit to the VH signal.

Not reviewed, for internal circulation only

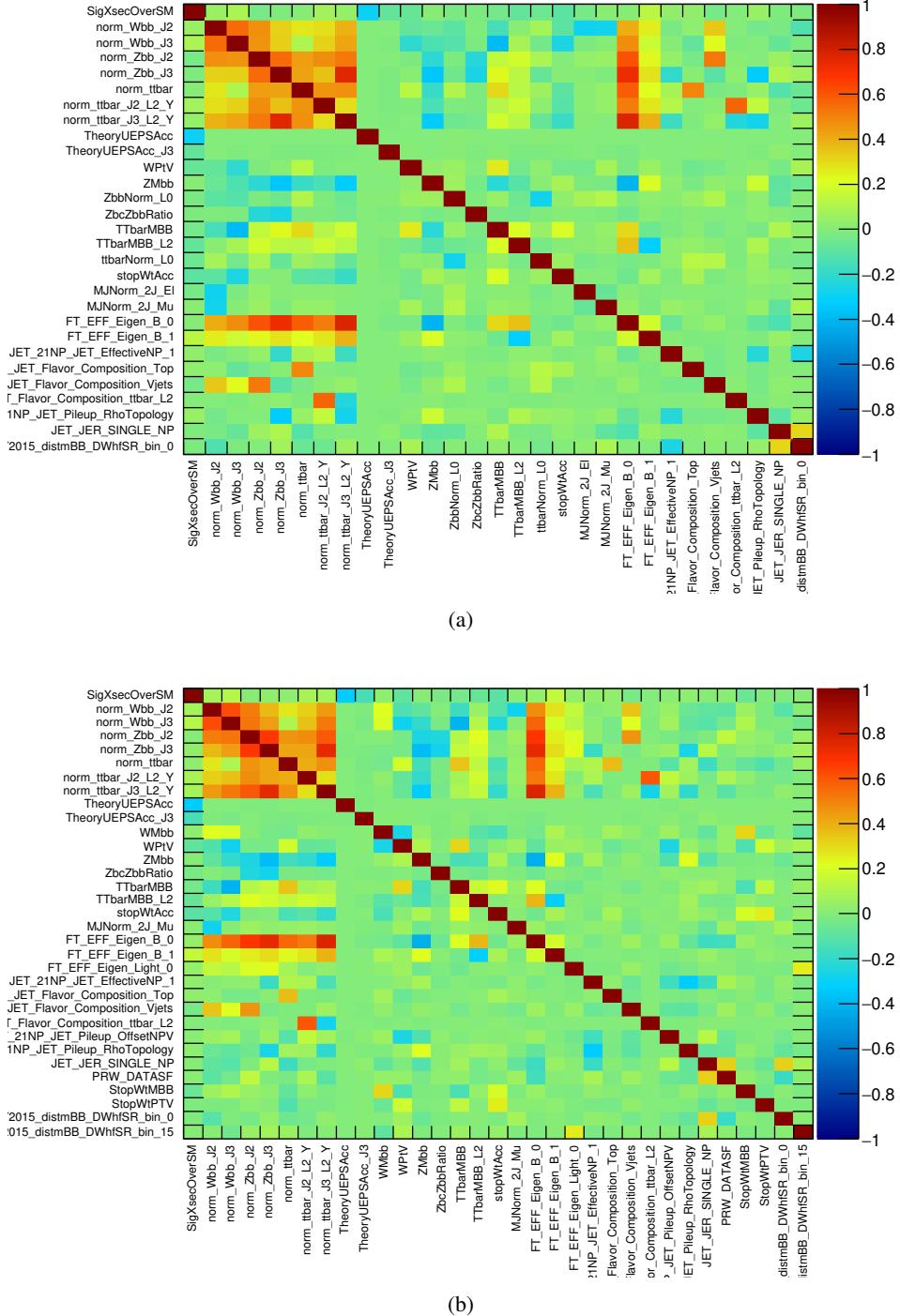


Figure 233: Correlation matrices from the unconditional combined m_{bb} fit to an Asimov dataset (a) and to 2015 and 2016 data (b).

Not reviewed, for internal circulation only

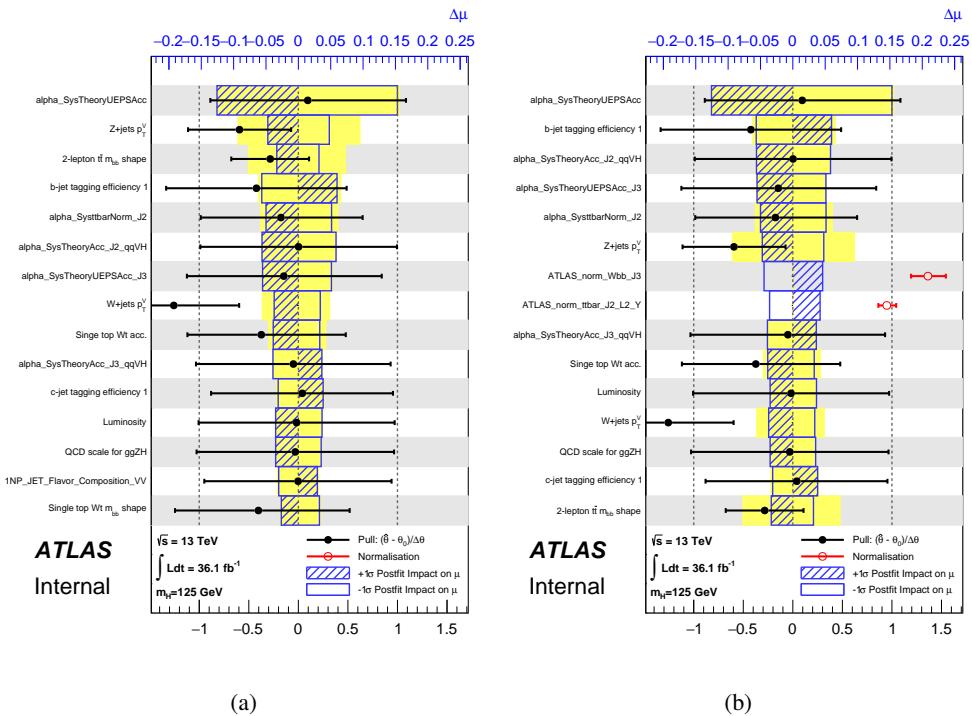


Figure 234: The post-fit impact nuisance parameter ranking for the combined fit based on prefit (a) and postfit (b) impact on $\hat{\mu}$ from the combined, unconditional cut-based fit to data.

Not reviewed, for internal circulation only

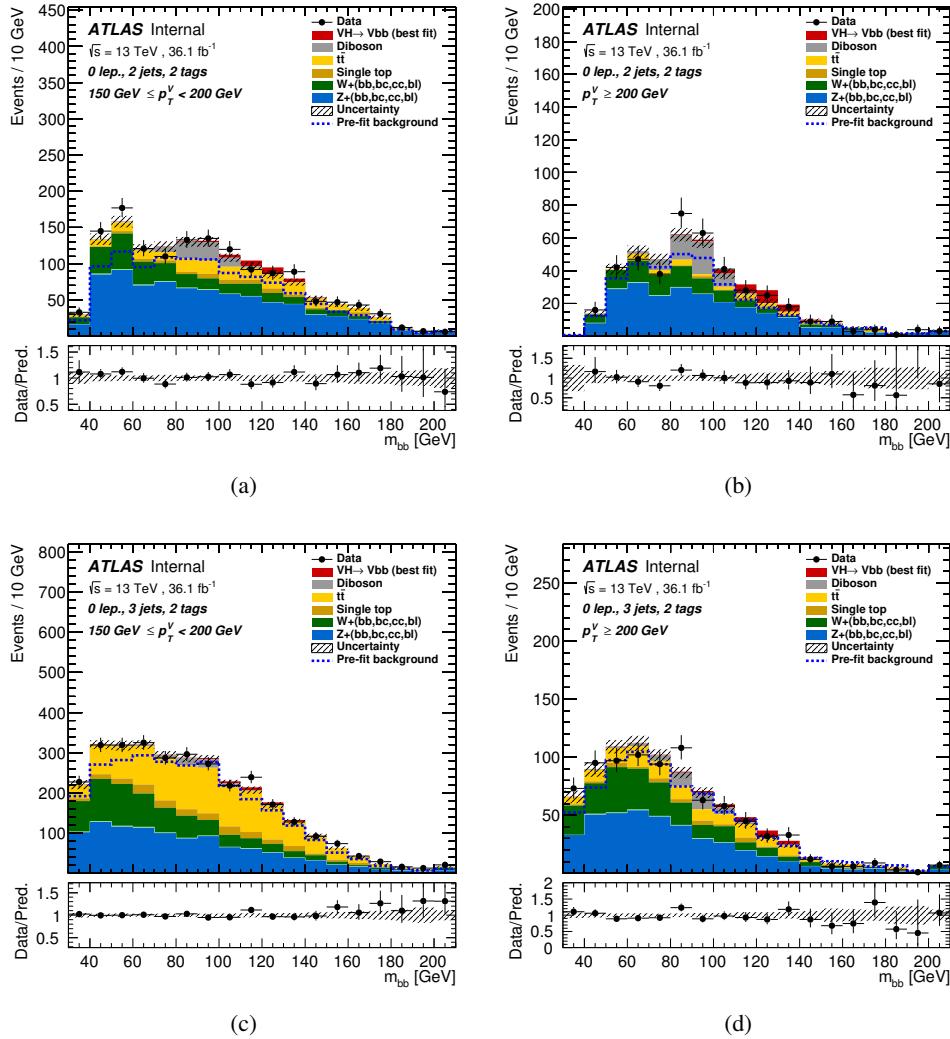


Figure 235: Blinded m_{bb} distribution in the 0-lepton channel after running the unconditional fit to data. The 2-tag 2-jet SR, $150 < p_T^V < 200$ GeV (a), 2-tag 2-jet SR, $200 < p_T^V$ GeV (b), and the 2-tag 3-jet SR, $150 < p_T^V < 200$ GeV (c) and 2-tag 3-jet SR, $200 < p_T^V$ GeV (d) categories are shown.

Not reviewed, for internal circulation only

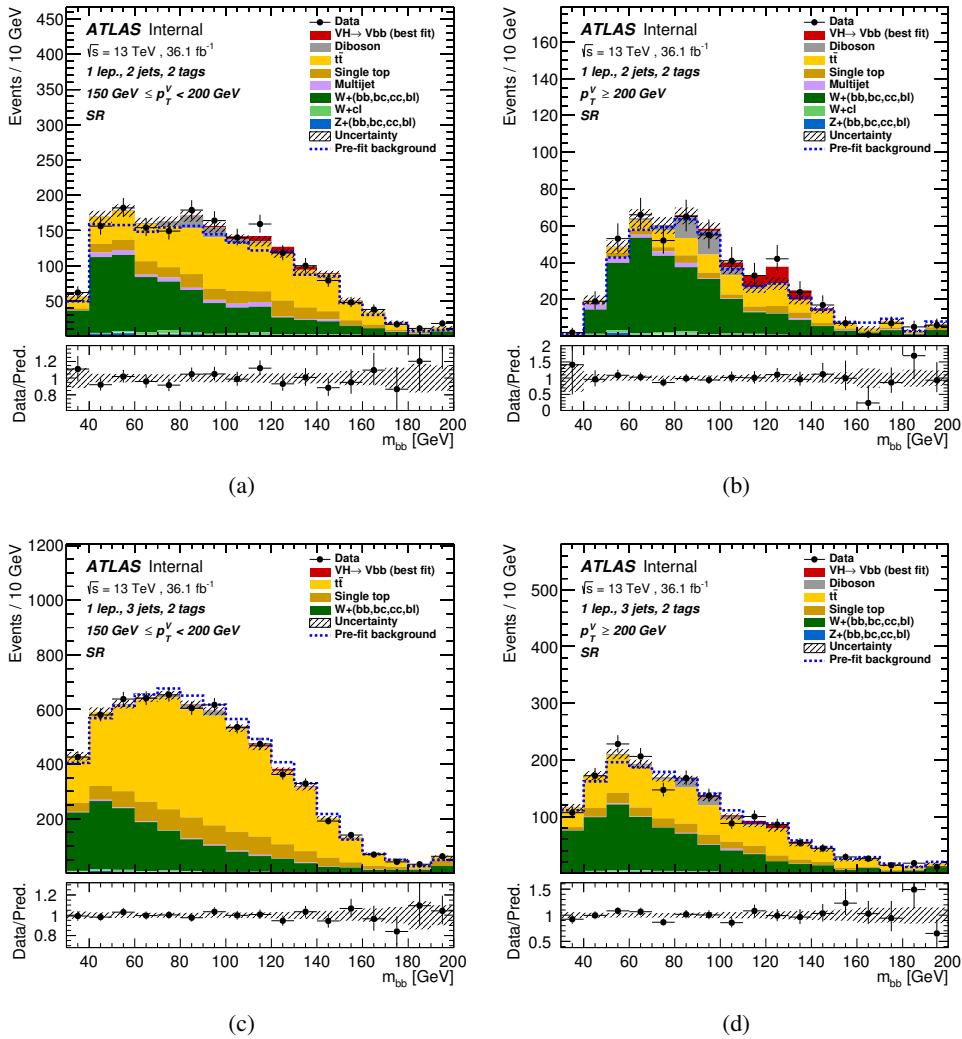


Figure 236: Blinded m_{bb} distribution in the 1-lepton channel after running the unconditional fit to data. The 2-tag 2-jet SR, $150 < p_T^V < 200$ GeV (a), 2-tag 2-jet SR, $p_T^V \geq 200$ GeV (b), and the 2-tag 3-jet SR, $150 < p_T^V < 200$ GeV (c) and 2-tag 3-jet SR, $p_T^V \geq 200$ GeV (d) categories are shown.

Not reviewed, for internal circulation only

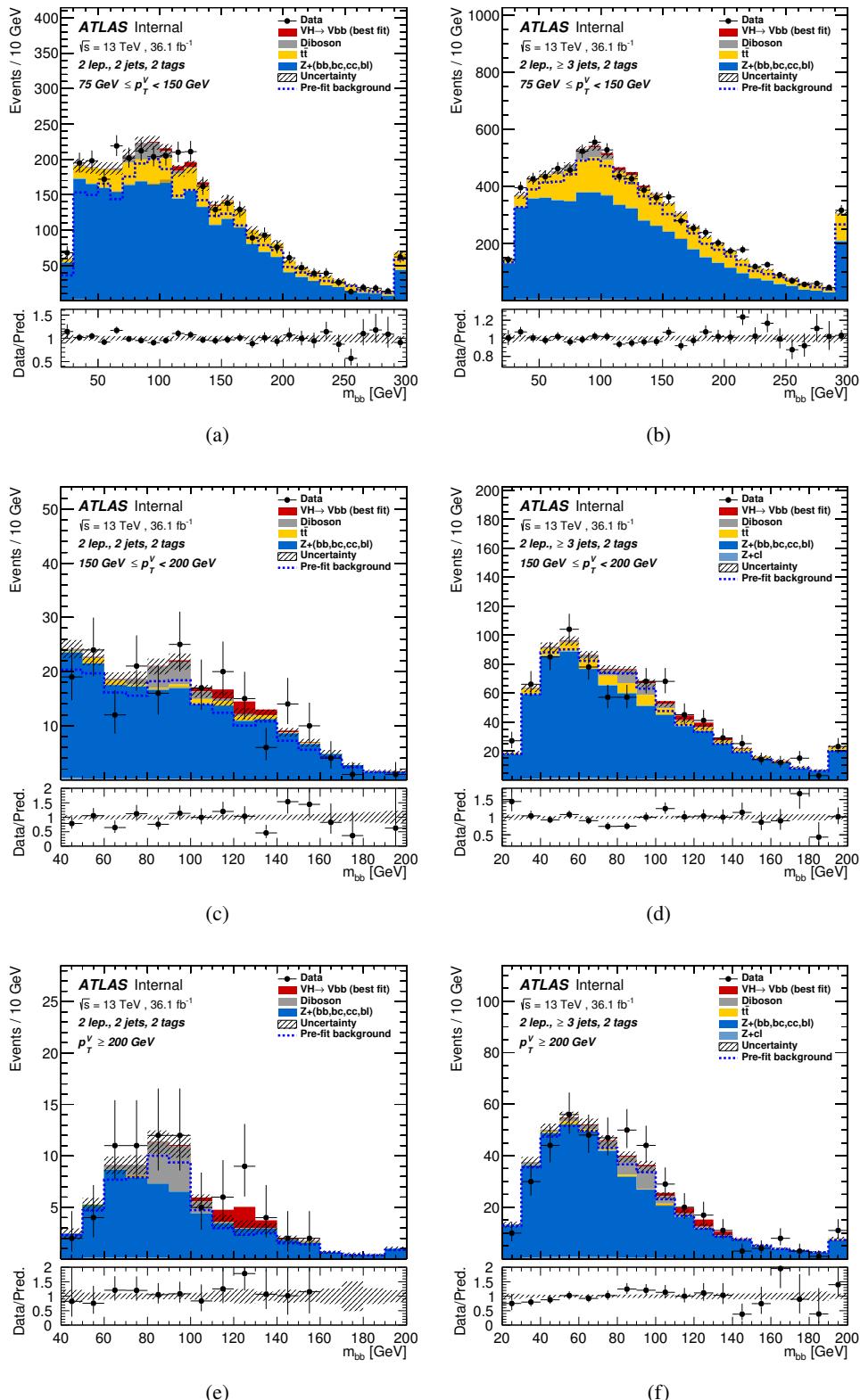


Figure 237: Blinded m_{bb} distribution in the 2-lepton channel after running the unconditional fit to data. The 2-tag 2-jet SR, $75 < p_T^V < 200 \text{ GeV}$ (a), 2-tag 3-jet SR, $75 < p_T^V < 150 \text{ GeV}$ (b), 2-tag 2-jet SR, $150 < p_T^V < 200 \text{ GeV}$ (c), 2-tag 3-jet SR, $150 < p_T^V < 200 \text{ GeV}$ (d), and the 2-tag 2-jet SR, $200 < p_T^V \text{ GeV}$ (e) and 2-tag 3-jet SR, $200 < p_T^V \text{ GeV}$ (f) categories are shown.

Set of nuisance parameters	Impact on error
Total	+0.459 / -0.398
DataStat	+0.277 / -0.270
FullSyst	+0.366 / -0.292
Floating normalizations	+0.079 / -0.063
All normalizations	+0.105 / -0.091
All but normalizations	+0.335 / -0.253
Jets MET	+0.068 / -0.047
BTag	+0.090 / -0.066
Leptons	+0.010 / -0.009
Luminosity	+0.048 / -0.026
Diboson	+0.058 / -0.038
Zjets	+0.059 / -0.059
Wjets	+0.069 / -0.052
Model ttbar	+0.098 / -0.077
Model Single Top	+0.051 / -0.050
Model Multi Jet	+0.019 / -0.021
Signal Systematics	+0.239 / -0.127
MC stat	+0.158 / -0.149

Table 64: The breakdown of the uncertainties coming from data statistics (“Data Stat.”), systematic uncertainties together with MC statistical uncertainties (“Full Syst.”), MC statistics only (“MC Stat.”), and experimental and modelling systematic uncertainties only (“All Syst.”) is shown in terms of the total impact on the uncertainty on μ for the combined, unconditional cut-based fit to data.

1562 **M.4.3. Background subtracted post-fit m_{bb} distributions**

1563 In this section, background subtracted data/MC comparison plots of the post-fit VH m_{bb} distribution are
 1564 presented for the cut-based fit. Figure 238 presents the background subtracted m_{bb} distribution summed
 1565 for all fitted signal regions, additional with each region weighted by the Higgs S/B, when running the
 1566 unconditional fit to data. Figures 239 - 241 then present the background subtracted MVA distributions in
 1567 the 0-, 1- and 2-lepton regions, respectively, after running the unconditional fit to data on all regions.

Not reviewed, for internal circulation only

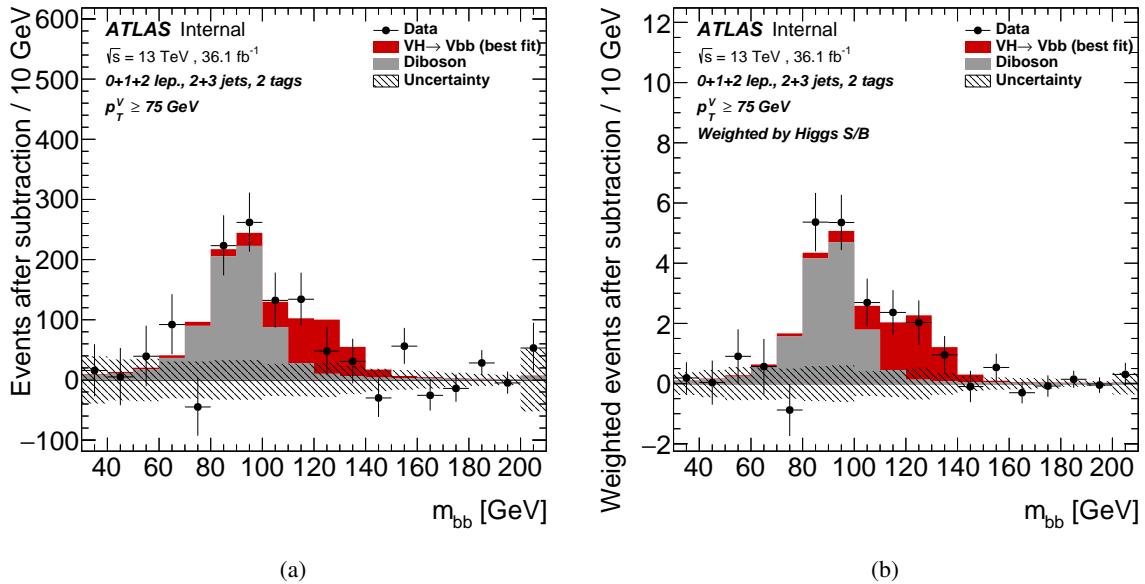


Figure 238: Background subtracted m_{bb} distributions summed for all fitted signal regions and channels (a), with each region weighted by the Higgs S/B (b), when running the unconditional cut-based fit to data.

Not reviewed, for internal circulation only

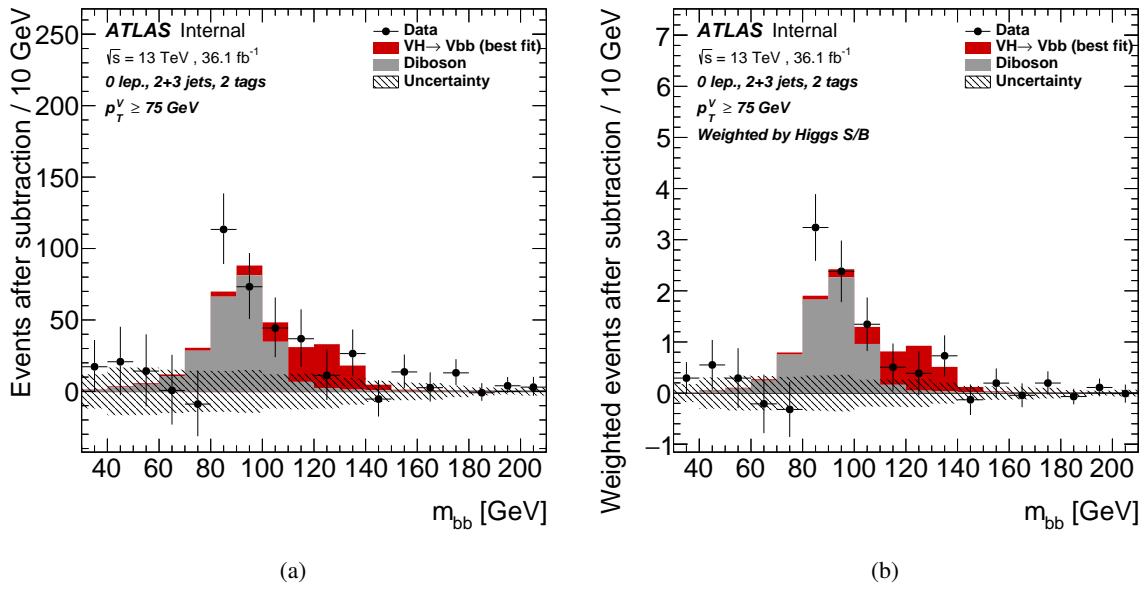


Figure 239: Background subtracted m_{bb} distributions summed for all fitted signal regions for the 0-lepton channel (a), with each region weighted by the Higgs S/B (b), when running the unconditional cut-based fit to data.

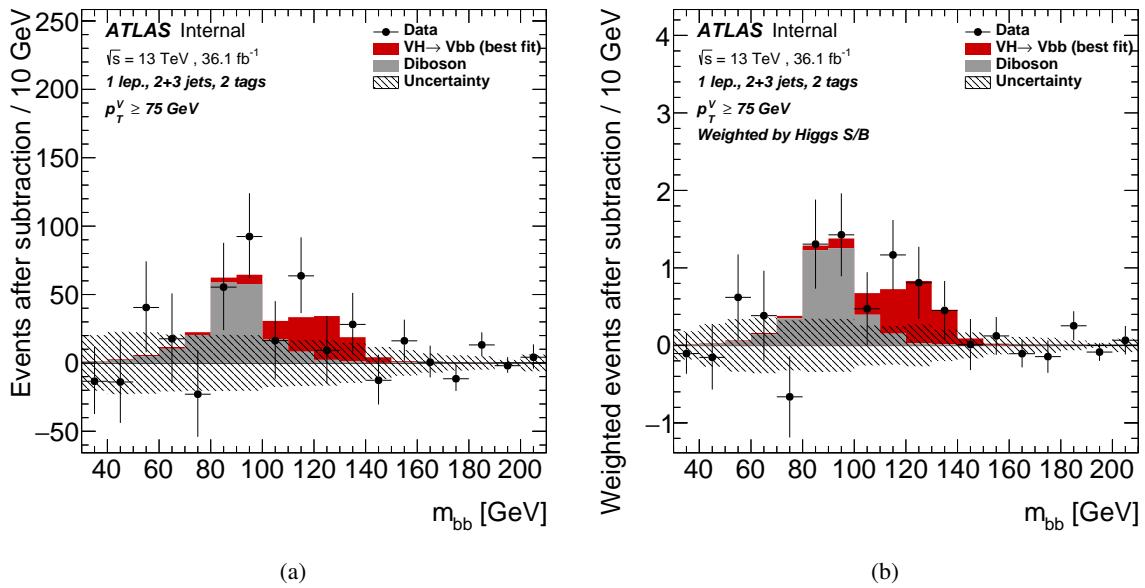


Figure 240: Background subtracted m_{bb} distributions summed for all fitted signal regions for the 1-lepton channel (a), with each region weighted by the Higgs S/B (b), when running the unconditional cut-based fit to data.

Not reviewed, for internal circulation only

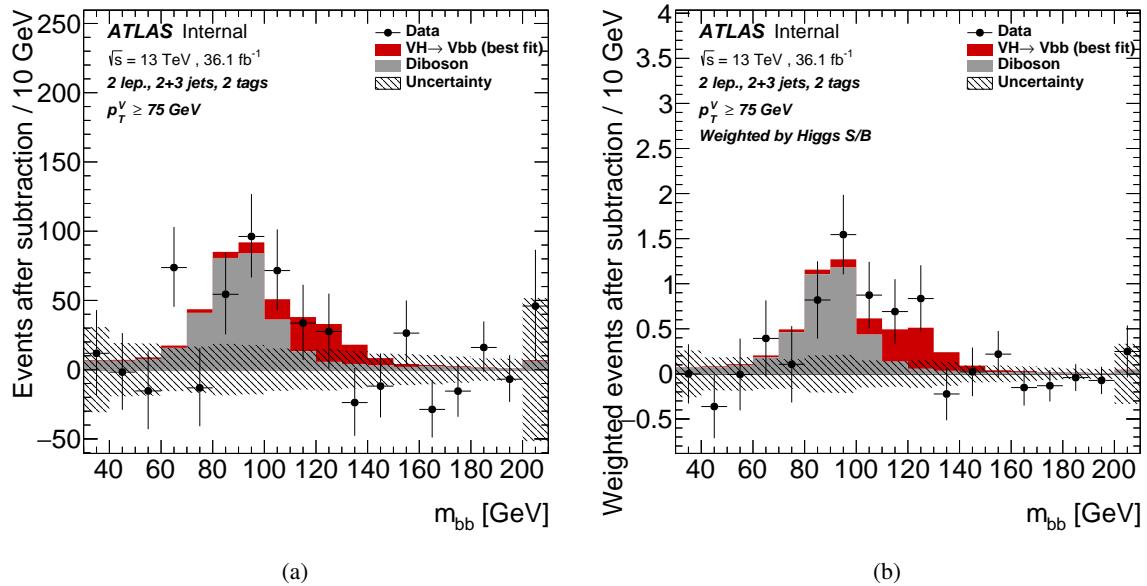


Figure 241: Background subtracted m_{bb} distributions summed for all fitted signal regions for the 2-lepton channel (a), with each region weighted by the Higgs S/B (b), when running the unconditional cut-based fit to data.

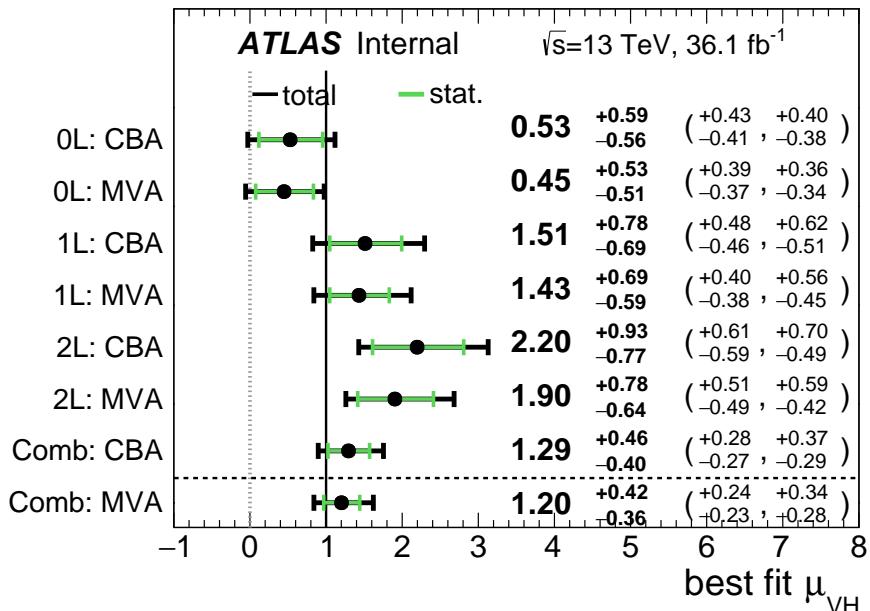
1568 **M.5. Fit significances**

1569 Detailed fit significances are provided in Table 65, detailing the m_{bb} significance in each fitting region.

Channel	Expected significance (Asimov)	Expected significance (data)	Observed significance (data)
0-lepton (SR)	1.75	1.55	0.38
1-lepton (SR+CR)	1.43	1.49	2.02
2-lepton (SR+CR)	1.60	1.60	3.45
0,1,2-lepton (SR+CR)	2.78	2.78	3.51

Table 65: Expected significances from a fit to an Asimov dataset corresponding to 36.1 fb^{-1} and observed significances from a fit to the data in the m_{bb} cut-based analysis. Significances from individual regions are estimated separately.

1570 In Figure 242, the $\hat{\mu}$ obtained in the combined MVA VH fit to data is compared to the ones obtained when
 1571 running the cut-based VH fit to data. Results are also presented when running the fit in each leptonic
 1572 channel. Very good agreement between the cut-based and MVA $\hat{\mu}$ values is observed.



(a)

Figure 242: The $\hat{\mu}$ obtained in the combined MVA VH fit to data is compared to the ones obtained when running the cut-based VH fit to data. Results are also presented when running the fit in each leptonic channel.

1573 **N. Detailed Studies on the combination of Run 1 and Run 2**

1574 **N.1. Breakdown of statistical and systematic uncertainties**

1575 A breakdown of the uncertainties from data statistics (“DataStat”), systematic uncertainties together with
 1576 MC statistical uncertainties (“FullSyst”), and other systematic uncertainty categories on the total error on
 1577 $\hat{\mu}$ may be found in Table N.1, and a ranking of nuisance parameters is shown in Figure 69.

Total	+0.278 / -0.261
DataStat	+0.185 / -0.181
FullSyst	+0.208 / -0.188
Floating normalizations	+0.055 / -0.056
All normalizations	+0.068 / -0.069
All but normalizations	+0.192 / -0.172
Jets, MET	+0.046 / -0.040
Jets	+0.041 / -0.036
MET	+0.023 / -0.018
BTag	+0.077 / -0.076
BTag b	+0.062 / -0.059
BTag c	+0.033 / -0.032
BTag light	+0.028 / -0.028
Leptons	+0.008 / -0.008
Luminosity	+0.026 / -0.014
Diboson	+0.030 / -0.027
Model Zjets	+0.049 / -0.050
Zjets flt. norm.	+0.032 / -0.040
Model Wjets	+0.082 / -0.083
Wjets flt. norm.	+0.031 / -0.027
Model ttbar	+0.047 / -0.046
ttbar flt. norm.	+0.025 / -0.026
Model Single Top	+0.047 / -0.045
Model Multi Jet	+0.027 / -0.038
Signal Systematics	+0.098 / -0.052
MC stat	+0.080 / -0.084

Table 66: Summary of the impact of different nuisance parameter categories on the total error on $\hat{\mu}$ for the combined Run1+Run2 fit.

1578 Figures N.1–N.1 show $\hat{\mu}$ summary plots for multiple parameter of interest fits. A fit with two parameters
 1579 of interest as the $\hat{\mu}$ ’s of Run 1 and Run 2 show has a 20.1% compatibility with the single parameter of
 1580 interest fit, and a six parameter interest fit (one each for each lepton channel in Run 1 and Run2) has a
 1581 7.10% compatibility with the single parameter of interest fit, while a four parameter of interest fit has a
 1582 34.6% compatibility.

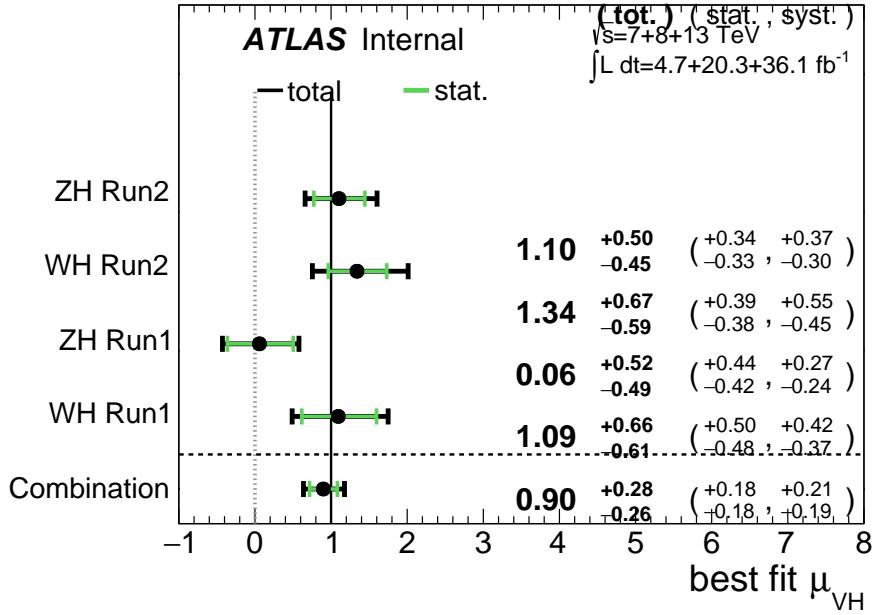


Figure 243: $\hat{\mu}$ summary plot for a four parameter fit.

Not reviewed, for internal circulation only

1583 N.2. Correlation Studies

1584 In order to decide which NP's would and would not be correlated between Run 1 and Run 2, a number of
 1585 studies were conducted concerning jet energy scale (JES) and modeling systematics. Each is described
 1586 in turn below. It should be noted that correlation studies were conducted using postfit Asimov data with
 1587 $\mu = 0.51$ (1.0) for Run 1 (2), which is why the results differ from those in the final, unblinded combination
 1588 with actual data. Updated results using the unblinded data are in progress.

1589 N.2.1. Jet Energy Scale

1590 The JetEtMiss recommended “strong” and “weak” correlation schemes between Run 1 and Run 2 were
 1591 used as a point of departure for the combination correlation scheme regarding JES NP's. However, the JES
 1592 NP's in both the Run 1 and Run2 workspaces are a reduced set of NP's, with some 56 (75) NP's reduced
 1593 to 6 (8) for Run 1 (2). In order to restore the full set of JES NP's, the effective NP's in each workspace are
 1594 unfolded using maps detailing the linear combinations of unfolded NP's that form the effective NP's.

1595 The linear combinations used to unfold the effective JES NP's were calculated as follows:

$$NP_{i,eff} = \frac{\sum_j A_{ij} |NP_{j,unf}| NP_{j,unf}}{\sqrt{\sum_j A_{ij}^2 |NP_{j,unf}|^2}} \quad (12)$$

1596 where *eff* and *unf* are for effective and unfolded NP's, respectively, the A_{ij} 's are scalar coefficients taken
 1597 from raw maps, and $|NP_{j,unf}|$ are the amplitudes of the unfolded NP's. The raw A_{ij} and scaled maps for
 1598 Run 1 and Run2 may be found in Figure 246

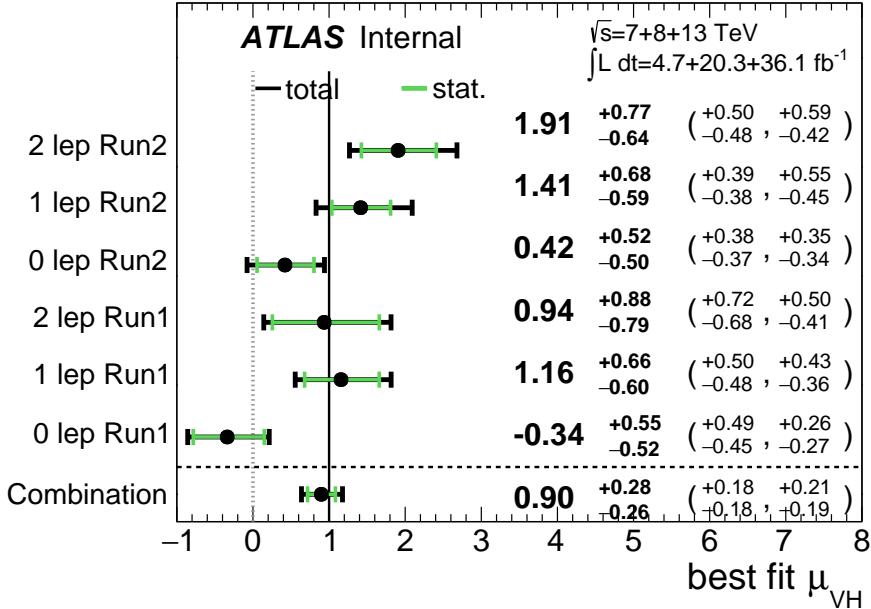


Figure 244: $\hat{\mu}$ summary plot for a six parameter of interest fit.

1599 Unfolding was found to have very little effect on both expected sensitivities and errors, as can be seen in
Tables N.2.1 and 68.

	R1 Unfold	R1 Eff	R2 Unfold	R2 Eff	Comb Unfold	Comb Eff
Exp. Sig.	2.604	2.606	2.952	2.952	3.96	3.951
Obs. Sig.	1.369	1.374	3.539	3.539	3.58	3.572
Exp. Limit	$0.755^{+0.296}_{-0.211}$	$0.755^{+0.296}_{-0.211}$	$0.707^{+0.277}_{-0.198}$	$0.707^{+0.277}_{-0.198}$	$0.506^{+0.198}_{-0.141}$	$0.505^{+0.198}_{-0.141}$
Obs. Limit	1.21	1.21	1.87	1.87	1.37	1.38

Table 67: Expected and observed sensitivities for Run 1, Run 2, and combined workspaces with effective and unfolded JES NP's.

1600

1601 As a result of these studies, the weak JES correlation scheme with uncorrelated effective JES NP's has
1602 been chosen as the treatment of JES in the Run 1 + Run 2 combined fit.

1603 N.2.2. Modeling Systematics

1604 Another principal systematic category is modeling uncertainties. The effect of correlating groups of
1605 systematics was estimated using the same strategy employed by the ATLAS/CMS SHVHbb combination
1606 for Run 1. This extrapolation can be used to estimate the impact of correlations on the estimated signal
1607 strength, the total error on the signal strength, and the χ^2 of the result. The impact of such correlations is
1608 no more than a few percent effect, as the following tables demonstrate, beginning with the category with
1609 the greatest shift, W+jets modeling, in Table N.2.2

Not reviewed, for internal circulation only

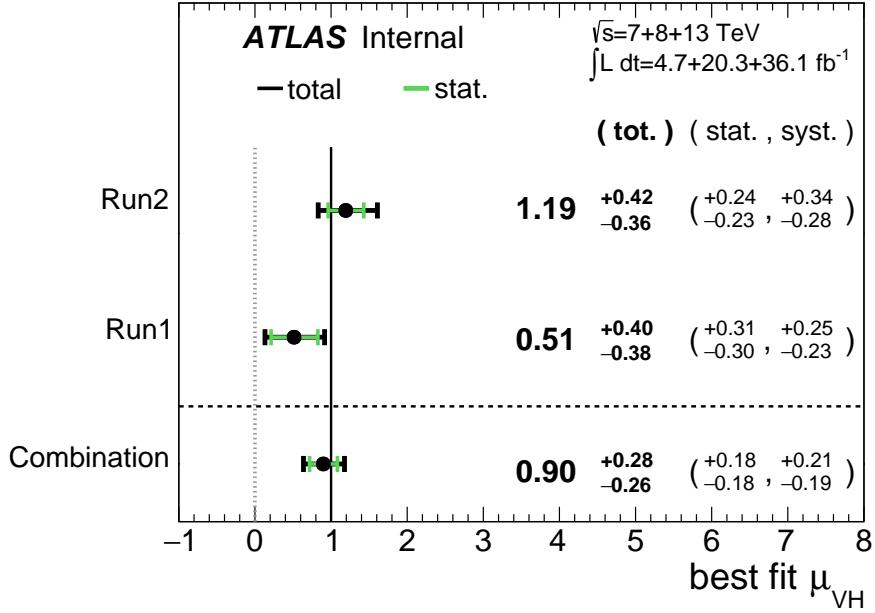


Figure 245: $\hat{\mu}$ summary plot for a two parameter of interest (Run 1 and Run 2) values.

	R1 Unfold	R1 Eff	R2 Unfold	R2 Eff	Comb Unfold	Comb Eff
$\Delta\hat{\mu}$	0.0018		0.0		0.0002	
muhat	0.5064	0.5082	1.2135	1.2135	0.9113	0.9115
Total	+0.400 -0.373	+0.401 -0.373	+0.386 -0.356	+0.386 -0.356	+0.277 -0.264	+0.277 -0.265
DataStat	+0.312 -0.301	+0.312 -0.301	+0.241 -0.236	+0.241 -0.236	+0.189 -0.185	+0.189 -0.185
FullSyst	+0.250 -0.220	+0.251 -0.220	+0.301 -0.267	+0.301 -0.267	+0.203 -0.189	+0.203 -0.190
Jets	+0.060 -0.051	+0.060 -0.052	+0.066 -0.048	+0.066 -0.048	+0.039 -0.036	+0.040 -0.037
BTag	+0.094 -0.079	+0.095 -0.079	+0.119 -0.110	+0.119 -0.110	+0.079 -0.080	+0.080 -0.080

Table 68: Error on signal strength breakdowns for Run 1, Run 2, and combined workspaces with effective and unfolded JES NP's.

1610 N.2.3. Flavour Tagging

1611 Given that c -tagging changed significantly between Run 1 and Run2 and that light tagging NP's are very
 1612 lowly ranked, these sets of NP's are left uncorrelated. Moreover, the change in the physical meaning of
 1613 the effective b -tagging NP's means a full correlation of such NP's (insomuch as they exist in each result)
 1614 is one of limited utility. Hence, it was decided to leave flavour tagging NP's uncorrelated. However,
 1615 since the meaning of the leading b -tagging NP's is approximately constant across years and since Run 2
 1616 b -tagging NP's are very highly ranked in both the Run 2 only and combined fits, tests correlating these
 1617 NP's were conducted, the results of which can be seen below. It should be noted that the leading B NP
 1618 at 8 TeV, SysBTagB0Effic_Y2012_8TeV, has an opposite effect on $t\bar{t}$ normalization than the 7 and 13
 1619 TeV NP's, and so must be flipped using a similar strategy as for JES unfolding. Initial studies of flavour
 1620 tagging correlations did not flip this NP, and so results for this scheme (labeled "B0 8TeV Not Flipped")

Not reviewed, for internal circulation only

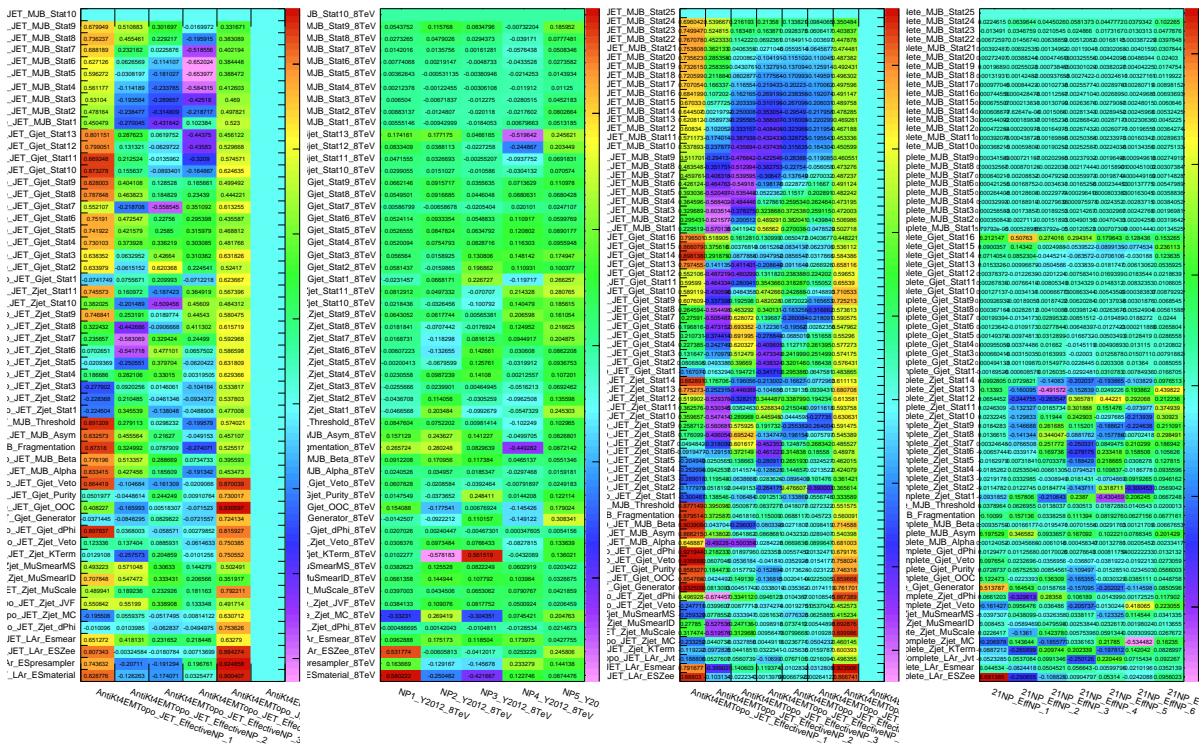


Figure 246: The raw and scaled coefficients for unfolding Run 1 (a and b) and Run 2 (c and d), respectively

1621 have also been included for comparison.

1622 It is clear from these results that correlating the leading effective Eigen NP associated with b 's can have
 1623 a noticeable effect on final fit results and that the 8 TeV B0 NP is the most important component of a
 1624 combined B0 NP. What is less clear is whether there are sufficient grounds for doing so (i.e. does the
 1625 correspondence of these NP's across years warrant a full correlation).

Not reviewed, for internal circulation only

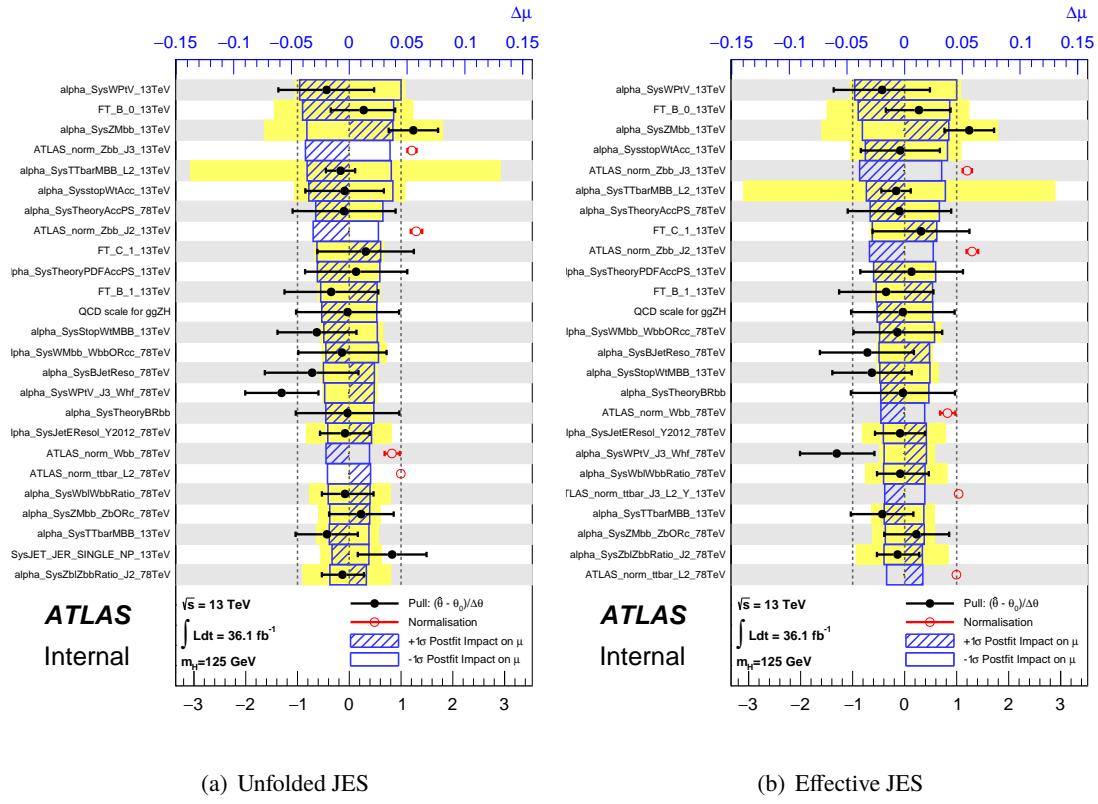


Figure 247: Ranks for the effective and unfolded JES NP Run1+Run2 combined workspaces.

	$ \Delta\mu $	σ	$ \Delta\sigma $	χ^2
$\rho=-1$	0.0024	0.2448	0.011 (4.3%)	0.95
$\rho=-0.6$	0.0015	0.2493	0.00654 (2.55%)	0.9804
$\rho=-0.3$	0.0008	0.2526	0.00325 (1.27%)	1.0045
$\rho=0$	—	0.2558	—	1.0298
$\rho=0.3$	0.0008	0.259	0.0032 (1.25%)	1.0564
$\rho=0.6$	0.0017	0.2622	0.00636 (2.49%)	1.0844
$\rho=1$	0.0029	0.2664	0.0105 (4.11%)	1.1242

Table 69: Run 1 + Run 2 W+jets modeling correlation projections

	Comb Eff	BTag B0	B0 8TeV Not Flipped
Exp. Sig.	3.951	4.071	3.868
Obs. Sig.	3.572	3.866	3.4
Exp. Limit	$0.505^{+0.198}_{-0.141}$	$0.493^{+0.193}_{-0.138}$	$0.513^{+0.201}_{-0.143}$
Obs. Limit	1.38	1.42	1.36

Table 70: Expected and observed sensitivities for a combination featuring the weak JES scheme, combination with the weak JES scheme + leading b NP's correlated, and the b correlation with the 8 TeV NP with sign unflipped.

Not reviewed, for internal circulation only

	Comb Eff	BTag B0	B0 8TeV Not Flipped
$\hat{\mu}$	0.9115	0.9598	0.8806
Total	+0.277 / -0.265	+0.273 / -0.260	+0.282 / -0.268
DataStat	+0.189 / -0.185	+0.185 / -0.181	+0.193 / -0.189
FullSyst	+0.203 / -0.190	+0.201 / -0.186	+0.206 / -0.189
BTag	+0.080 / -0.080	+0.073 / -0.072	+0.083 / -0.079
BTag b	+0.064 / -0.062	+0.056 / -0.051	+0.066 / -0.063

Table 71: Breakdowns of the impact of different NP sets on total error on $\hat{\mu}$ for a combination featuring the weak JES scheme and a combination with the weak JES scheme + leading b NP's correlated.

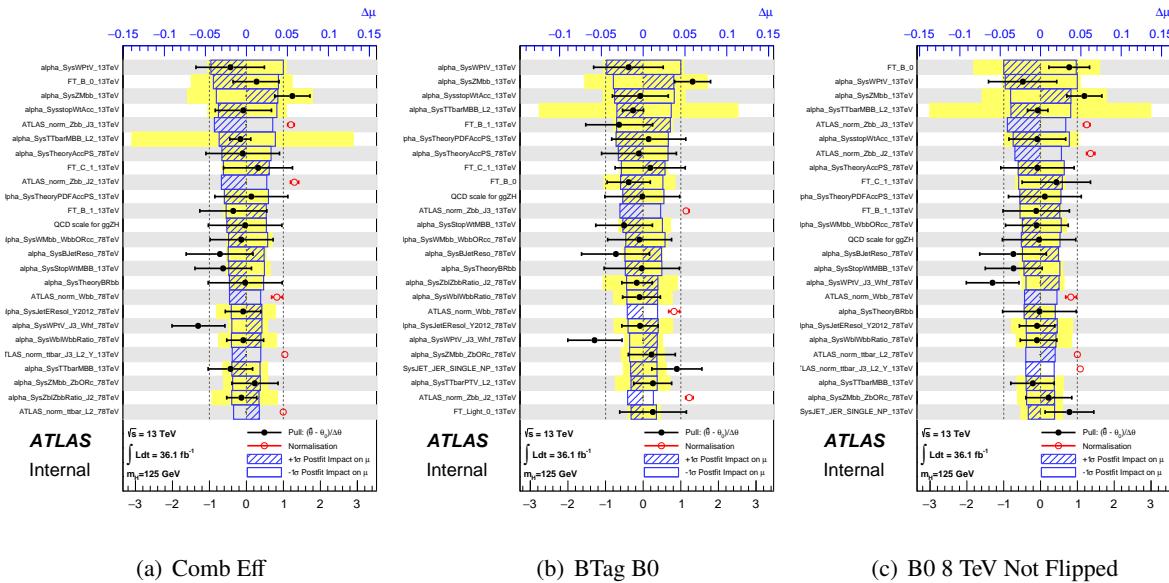


Figure 248: NP rankings for a combination featuring the weak JES scheme and a combination with the weak JES scheme + leading b NP's correlated.

1626 References

- 1627 [1] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, **13** (1964) 321.
- 1628 [2] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, **13** (1964) 508.
- 1629 [3] P. W. Higgs, *Broken Symmetries, Massless Particles and Gauge Fields*, **12** (1964) 132.
- 1630 [4] G. Guralnik, C. Hagen and T. Kibble, *Global Conservation Laws and Massless Particles*,
1631 **13** (1964) 585.
- 1632 [5] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson*
1633 *with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1–29,
1634 arXiv: [1207.7214 \[hep-ex\]](https://arxiv.org/abs/1207.7214).
- 1635 [6] S. Chatrchyan et al.,
1636 *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*,
1637 *Phys. Lett. B* **716** (2012) 30–61, arXiv: [1207.7235 \[hep-ex\]](https://arxiv.org/abs/1207.7235).
- 1638 [7] G. Aad et al., *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8*
1639 *TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803,
1640 arXiv: [1503.07589 \[hep-ex\]](https://arxiv.org/abs/1503.07589).
- 1641 [8] G. Aad et al.,
1642 *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*,
1643 *Eur. Phys. J. C* **75**.10 (2015) 476, [Erratum: *Eur. Phys. J.C*76,no.3,152(2016)],
1644 arXiv: [1506.05669 \[hep-ex\]](https://arxiv.org/abs/1506.05669).
- 1645 [9] V. Khachatryan et al., *Constraints on the spin-parity and anomalous HVV couplings of the Higgs*
1646 *boson in proton collisions at 7 and 8 TeV*, *Phys. Rev. D* **92**.1 (2015) 012004,
1647 arXiv: [1411.3441 \[hep-ex\]](https://arxiv.org/abs/1411.3441).
- 1648 [10] ‘Measurements of the Higgs boson production and decay rates and constraints on its couplings
1649 from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV’,
1650 tech. rep. ATLAS-CONF-2015-044, CERN, 2015,
1651 URL: <http://cds.cern.ch/record/2052552>.
- 1652 [11] LHC Higgs Cross Section Working Group, *LHC Higgs Cross Section Working Group*,
1653 <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>, 2016,
1654 URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG> (visited on
1655 18/03/2016).
- 1656 [12] S. Chatrchyan et al., *Search for the standard model Higgs boson produced in association with a W*
1657 *or a Z boson and decaying to bottom quarks*, *Phys. Rev. D* **89**.1 (2014) 012003,
1658 arXiv: [1310.3687 \[hep-ex\]](https://arxiv.org/abs/1310.3687).
- 1659 [13] G. Aad et al., *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated (W/Z)H*
1660 *production with the ATLAS detector*, *JHEP* **01** (2015) 069, arXiv: [1409.6212 \[hep-ex\]](https://arxiv.org/abs/1409.6212).
- 1661 [14] ‘Search for the Standard Model Higgs boson produced in association with a vector boson and
1662 decaying to a $b\bar{b}$ pair in pp collisions at 13 TeV using the ATLAS detector’,
1663 tech. rep. ATLAS-CONF-2016-091, CERN, 2016,
1664 URL: <http://cds.cern.ch/record/2206813>.

- [15] A. Buzatu and W. Wang, *Object selections for SM Higgs boson produced in association with a vector boson in which $H \rightarrow b\bar{b}$ and V decays leptonically with Run-2 data: Object support note for $VH(b\bar{b})$ 2015+2016 dataset publication*, ATL-COM-PHYS-2016-1674 (2016), URL: <https://cds.cern.ch/record/2233686>.
- [16] A. Robson, G. Piacquadio and E. Schopf, *Signal and Background Modelling Studies for the Standard Model VH , $H \rightarrow b\bar{b}$ and Related Searches: Modelling support note for $VH(b\bar{b})$ 2015+2016 dataset publication*, ATL-COM-PHYS-2016-1747 (2016), This is a support note for the $VH(b\bar{b})$ SM publication using the 2015+2016 datasets., URL: <https://cds.cern.ch/record/2235887>.
- [17] LHC Higgs Cross Section Working Group et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, CERN-2013-004 (CERN, Geneva, 2013), arXiv: [1307.1347 \[hep-ph\]](https://arxiv.org/abs/1307.1347).
- [18] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $O(\alpha(s)^{**2})$* , Phys.Rev. **D74** (2006) 114017, arXiv: [hep-ph/0609070 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0609070).
- [19] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\alpha(\frac{4}{S})$* , Phys. Rev. Lett. **110** (2013) 252004, arXiv: [1303.6254 \[hep-ph\]](https://arxiv.org/abs/1303.6254).
- [20] M. Aliev et al., *HATHOR: Hadronic Top and Heavy quarks cross section calculator*, Comput.Phys.Commun. **182** (2011) 1034–1046, arXiv: [1007.1327 \[hep-ph\]](https://arxiv.org/abs/1007.1327).
- [21] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Computer Physics Communications **191** (2015) 74–89, ISSN: 0010-4655, URL: <http://www.sciencedirect.com/science/article/pii/S0010465515000454>.
- [22] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, Nucl. Phys. Proc. Suppl. **205-206** (2010) 10–15, arXiv: [1007.3492 \[hep-ph\]](https://arxiv.org/abs/1007.3492).
- [23] *Supporting Document for the Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated (W/Z) H production with the ATLAS detector*, ATL-COM-2014-051 (2014), URL: <https://cds.cern.ch/record/1645654>.
- [24] ‘Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run’, tech. rep. ATL-PHYS-PUB-2016-012, CERN, 2016, URL: <https://cds.cern.ch/record/2160731>.
- [25] F Ahmadov et al., ‘Supporting Document for the Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated (W/Z) H production with the ATLAS detector’, tech. rep. ATL-COM-PHYS-2014-051, CERN, 2014, URL: <https://cds.cern.ch/record/1645654>.
- [26] HSG5, ‘Invariant Mass Studies for the $H \rightarrow b\bar{b}$ Measurements for LHCP’, tech. rep. ATL-COM-PHYS-2013-449, CERN, 2013.
- [27] *Supporting Document for the Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated (W/Z) H production with the ATLAS detector*, ATL-COM-PHYS-2014-051 (2014), eprint: <https://cds.cern.ch/record/1645654>.
- [28] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv: [1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727).

- ¹⁷⁰⁸ [29] A. L. Read, *Presentation of search results: The CL(s) technique*, J.Phys. **G28** (2002) 2693–2704.

[Not reviewed, for internal circulation only]

1709 List of contributions

Not reviewed, for internal circulation only

	Ahmadov, Faig	1-lepton analyzer, MVA and statistical analysis
	Argyropoulos, Spyridon	MC contact, CxAOD framework, Study of systematics smoothing
	Arnold, Hannah	CxAOD framework
	Amaral Coutinho, Yara	jet resolution studies
	Sanchez Pineda, Arturo	CxAOD framework
	Rodolfo	
	Bell, Andrew Stuart	1-lepton analyzer, primary input producer, MVA studies, main support note editor, BG modeling, Statistical analysis, Flavour tagging contact
	Benitez, Jose	0-lepton analyzer, Main support note editor
	Buzatu, Adrian	dijet mass reconstruction, CxAOD framework, Objects support note editor
	Calderini, Giovanni	Supervision, BG modeling
	Chan, Stephen Kam-wah	Statistical analysis
	Chen, Chunhui	Supervision
	Buckley, Andrew	Supervision
	Buscher, Daniel	BG modelling, CxAOD framework, statistical analysis
	Delporte, Charles	0-lepton analyzer, b-tagging studies
1710	De Lorenzi, Francesco	VH boosted studies, statistical analysis
	De Vivie De Regie, Jean-Baptiste	Supervision
	Dao, Valerio	Coordination of modelling studies, statistical analysis
	Enari, Yuji	Supervision
	Francavilla, Paolo	BG modeling, Statistical analysis, paper editor
	Gargiulo, Simona	W+jets modelling uncertainties
	Gaycken, Goetz	Supervision
	Gray, Chloe	Signal modelling uncertainties
	Grivaz, Jean-Francois	Supervision, analysis review
	Hays, Jonathan Michael	CxAOD framework coordinator
	Hesketh, Gavin Grant	Supervision
	Hobbs, John	Supervision
	Huth, John	Supervision
	Ishijima, Naoki	1-lepton analyzer, multijet studies
	Kado, Marumi	Supervision
	Kato, Chikuma	dijet mass reconstruction, CxAOD production
	Li, Changqiao	0-lepton analyzer, primary input producer, CxAOD production, statistical analysis

	Li, Haifeng	Diboson modelling uncertainties, statistical analysis
	Liu, Yanwen	Supervision
	Marchiori, Giovanni	Egamma contact, Supervision
	Ma, Lianliang	Supervison
	Ma, Yanhui	1-lepton analyzer, CxAOD framework, multijet studies
	Mario Jose da Cunha Sargedas de Sousa	Statistical analysis
1711	Masubuchi, Tatsuya	2-lepton analyzer, BG modelling, Main supporting note editor
	Mallik, Usha	Supervision
	Mehta, Andrew	2-lepton analyzer, primary input producer, BG modelling, statistical analysis
	Montalbano, Alyssa	1-lepton analyzer, multijet studies, statistical analysis
	Morange, Nicolas	Coordination
	Moser, Brian	BG modelling
	Pandini, Carlo Enrico	0-lepton analyzer, BG and signal modelling
	Piacquadio, Giacinto	Paper editor, coordination of multijet studies
	Prell, Soeren	Supervision
	Robson, Aidan	Single-top modelling uncertainties
	Scanlon, Tim	Coordination
	Schopf, Elisabeth	Modelling support note editor, background modelling studies, statistical analysis
	Sekula, Stephen	CxAOD framework, supervision
	Thompson, Paul	2-lepton analyzer, primary input producer, BG modelling
	Tsybychev, Dmitri	Supervision
	Vergel Infante, Carlos Miguel	MET trigger studies
	Wang, Song-Ming	MET trigger study, Trigger contact, supervision
	Wang, Wei	MET trigger study, Objects support note editor
	Weiser, Christian	Supervision
	Yamaguchi, Yohei	Derivations contact, multijet studies, 1-lepton analyzer
	Yang, Siqi	MVA studies
1712	Yu, Jie	VH boosted studies, statistical analysis