

# CMS Draft Analysis Note

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

2020/01/09

Archive Hash: c4444da-D

Archive Date: 2020/01/09

## Search for VBF Higgs bosons decaying to invisible particles at 13 TeV with 2017 and 2018 data

Alp Akpinar, Andreas Albert, Zeynep Demiragli, and Siqi Yuan  
Boston University (US)

### Abstract

This note describes the search for invisible decays of Higgs boson produced with vector boson fusion (VBF). The search is performed using a shape-based analysis, using the data collected by CMS at  $\sqrt{s} = 13 \text{ TeV}$  in 2017 and 2018, corresponding to integrated luminosities of  $41.3 \text{ fb}^{-1}$  and  $59.7 \text{ fb}^{-1}$ , respectively.

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

PDFAuthor: Alp Akpinar, Andreas Albert, Zeynep Demiragli, Siqi Yuan  
PDFTitle: Run-II VBFHinv BU  
PDFSubject: CMS  
PDFKeywords: CMS, physics, your topics

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



## 1 Contents

2	1	Introduction . . . . .	2
3	2	Samples . . . . .	2
4	2.1	Data Samples . . . . .	2
5	2.2	Background Samples . . . . .	2
6	2.3	Signal simulation samples . . . . .	4
7	3	Trigger . . . . .	4
8	3.1	Efficiency measurement . . . . .	5
9	4	Physics objects . . . . .	11
10	4.1	Jets . . . . .	11
11	4.2	Missing transverse momentum and recoil . . . . .	12
12	4.3	Leptons . . . . .	13
13	4.4	Photons . . . . .	15
14	5	Reweighting of simulated events . . . . .	19
15	5.1	Trigger efficiency reweighting . . . . .	19
16	5.2	Pileup reweighting . . . . .	19
17	5.3	Lepton and photon identification/reconstruction efficiency reweighting . . . . .	19
18	5.4	Higher-order reweighting . . . . .	26
19	6	Event selection . . . . .	30
20	6.1	Signal region selection . . . . .	30
21	6.2	Single muon control region selection . . . . .	33
22	6.3	Single electron control region selection . . . . .	38
23	6.4	Double muon control region selection . . . . .	43
24	6.5	Double electron control region selection . . . . .	48
25	6.6	Photon control region . . . . .	53
26	7	Background estimation . . . . .	57
27	7.1	Signal extraction strategy . . . . .	57
28	7.2	Transfer factors . . . . .	57
29	7.3	Systematic uncertainties . . . . .	60
30	7.4	Control sample validation . . . . .	62
31	7.5	QCD multijet estimation . . . . .	62
32	8	Results . . . . .	65
33	A	Comparison of 1D and 2D QCD k-factors . . . . .	66
34	B	Trigger efficiencies in double muon CR . . . . .	72

## 35 1 Introduction

36 This analysis note describes a search for events in which the Higgs boson is produced by vector  
 37 boson fusion (VBF), decaying into invisible particles in the final state. The signature of such a  
 38 final state will be two separated jets and an imbalance in  $\vec{p}_T$  due to the undetected particles.  
 39 Two separated jets are the result of the hadronization of two final state quarks emerging from  
 40 the VBF process.

41 This analysis makes use of data collected with the CMS detector in proton-proton (pp) colli-  
 42 sions in 2017 and 2018 at  $\sqrt{s} = 13$  TeV, corresponding to an integrated luminosity of  $41.3 \text{ fb}^{-1}$   
 43 and  $59.7 \text{ fb}^{-1}$ , respectively.

44 The analysis strategy is similar to that of 2016 VBF analysis, with several improvements. One  
 45 of the major improvements in this analysis is the addition of the photon control region as the  
 46 fifth control region. The addition of this new control region provides additional constraint and  
 47 sensitivity on the final fit. In addition to this, the theory scale factors are re-derived in this  
 48 analysis for VBF topology.

## 49 2 Samples

50 The analysis described in this note is performed using data collected in 2017 by CMS at 13 TeV  
 51 and corresponds to an integrated luminosity of  $41.5 \text{ fb}^{-1}$ . The MC simulation samples for the  
 52 background processes have been produced in the Fall17 and Autumn18 campaigns. Further  
 53 details are given in the following sub-sections.

### 54 2.1 Data Samples

55 The datasets listed in Tab. 1 are used to select events in the signal and the control regions, and  
 56 to perform measurements on physics objects used in the analysis (e.g. trigger turn-ons).

Table 1: List of datasets used to select events in the signal and control regions. Datasets for both years correspond to the Nano1June2019 campaign, otherwise known as v5. For the 2017 data, the 31Mar2018 reconstruction is used for all periods. For 2018 data, the 17Sep2018 reconstruction is used for runs A to C, and the 22Jan2019 reconstruction is used for run D.

Year	Dataset name	Events selected for
both	/MET/Run201*/NANO AOD	Signal, single muon, double muon control regions
2017	/SingleElectron/Run2017*/NANO AOD	Single electron, double electron control regions
	/SinglePhoton/Run2017*/NANO AOD	Single photon control region
2018	/EGamma/Run2018*/NANO AOD	Single electron, double electron control, single photon control regions

### 57 2.2 Background Samples

58 Simulation datasets for the background processes are listed in Table 2 and 3. There are several  
 59 Standard Model processes that pose as backgrounds to the VBF  $H_{inv}$  signal, experimental sig-  
 60 nature of the final state being two jets with large rapidity separation and invariant mass, along  
 61 with  $p_T^{\text{miss}}$ . These processes are as follows:

62 **Z( $\nu\nu$ ) + jets** : This process yields the largest irreducible background in the analysis, consist-  
 63 ing of a Z boson and 2 or more jets coming from either QCD or EWK vertices. Simulated

samples for this background have been produced at leading order (LO) in QCD using the aMC@NLO generator in several bins of  $H_T$  for the QCD case, and in one inclusive sample for the EWK case.

**W + jets** : This process is the second largest source of background in this analysis, consisting of a W boson and 2 or more jets coming from either QCD or EWK vertices. The contribution of this background can be reduced by rejecting events with charged lepton candidates (electron/muon/tau). However, this process becomes irreducible in the case where the charged leptons are outside of the detector acceptance. Simulation samples for this background have been generated at LO in QCD using the aMC@NLO generator in several bins  $H_T$  for the QCD case, and in once inclusive sample for the EWK case.

**Z( $\ell\ell$ ) + jets** : This process mimics signal-like events in the case where the leptons coming from the Z boson decay are not reconstructed. As in the case of  $W \rightarrow \ell\nu$ , the contribution of this background is reduced by rejecting events with charged leptons. Simulation samples for this process have been generated at LO in QCD using the aMC@NLO generator in bins of  $H_T$  for the QCD case, and in once inclusive sample for the EWK case.

**Top:** Top-quark decays (both  $t\bar{t}$  and single top) also contribute background events to this analysis. In these processes, the W boson produced in a top-quark decay further decays leptonically, which produces genuine  $p_T^{\text{miss}}$  in the event. Next-to-leading order (NLO)  $t\bar{t}$  simulation samples have been produced with the aMC@NLO generator with two additional partons in the matrix element. Single-top events have been generated with the Powheg generator at NLO in QCD with one additional matrix element parton.

**Dibosons:** Decays of diboson pairs (WW, WZ, ZZ) also constitute background processes. Typically, one of the bosons decays leptonically ( $W \rightarrow \ell\nu, Z \rightarrow \nu\nu$ ) while the other boson decays hadronically, thus producing jets and  $p_T^{\text{miss}}$  in the final state. Simulated samples for WW, WZ and ZZ production have been generated at LO with Pythia 8.

**QCD Multijet:** QCD multijet events typically do not have large genuine  $p_T^{\text{miss}}$ . However, given the large cross section with which these events are produced, even a small fraction of events with jet mismeasurement results in a non-zero contribution of this process as background in the analysis. Simulated QCD samples have been generated at LO in QCD using the MadGraph generator in several bins of  $H_T$ .

The MC samples produced using MADGRAPH5\_aMC@NLO, and POWHEG generators are interfaced with PYTHIA using the CP5 tune [1] for the fragmentation, hadronization, and underlying event description. In the case of the MADGRAPH5\_aMC@NLO samples, jets from the matrix element calculations are matched to the parton shower description following the MLM [2] (FxFx [3]) prescription to match jets from matrix element calculations and parton shower description for LO (NLO) samples. The NNPDF 3.1 NNLO [4] parton distribution functions (PDFs) are used in all generated samples. The propagation of all final-state particles through the CMS detector is simulated with the GEANT 4 software [5]. The simulated events include the effects of pileup, with the multiplicity of reconstructed primary vertices matching that in data. The average number of pileup interactions per proton bunch crossing is found to be 32 for both the 2017 and 2018 data samples used in this analysis (assuming a total inelastic proton-proton cross-section).

Dataset name	Cross section (pb)	Order in QCD
WJetsToLNu_HT-70To100_TuneCP5_13TeV-madgraphMLM-pythia8	1296	LO
WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1392	LO
WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	410.2	LO
WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	57.95	LO
WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	12.98	LO
WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	5.39	LO
WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.08	LO
WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.008098	LO
ZJetsToNuNu_HT-100To200_13TeV-madgraph	305.3	LO
ZJetsToNuNu_HT-200To400_13TeV-madgraph	91.86	LO
ZJetsToNuNu_HT-400To600_13TeV-madgraph	13.13	LO
ZJetsToNuNu_HT-600To800_13TeV-madgraph	3.242	LO
ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.501	LO
ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.3431	LO
ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.005146	LO
DYJetsToLL_M-50_HT-70to100_TuneCP5_13TeV-madgraphMLM-pythia8	146.9	LO
DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	160.9	LO
DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	48.68	LO
DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	6.998	LO
DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.745	LO
DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.8077	LO
DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.1923	LO
DYJetsToLL_M-50_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.003477	LO
GJets_HT-40To100_TuneCP5_13TeV-madgraphMLM-pythia8	18640	LO
GJets_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	8641	LO
GJets_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	2196	LO
GJets_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	258.4	LO
GJets_HT-600ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	85.23	LO

Table 2: Simulated datasets for the modelling of single electroweak boson backgrounds. All datasets are accessed at the NanoAOD data tier from the v5 campaign, also known as 1Jun19.

### 2.3 Signal simulation samples

Simulated signal samples for invisible decays of the Higgs boson are generated using the Powheg generator at NLO in QCD. The samples are normalized to the standard model cross section for Higgs production in the given mode [6]. They are listed in Tab. 4.

## 3 Trigger

Events for the signal region are collected using a set of dedicated triggers designed to select events with large  $p_T^{\text{miss}}$  and large  $H_T^{\text{miss}}$  based on the online particle flow (PF) algorithm. In these dedicated trigger algorithms, identified PF muons are removed from the event before the  $p_T^{\text{miss}}$  and the  $H_T^{\text{miss}}$  objects are calculated. With this definition, the signal trigger paths can also be used to select single and double muon events for the W and Z control regions, respectively.

Electron events for the W and Z regions are selected using a single electron trigger. To ensure the trigger efficiency also for high- $p_T$  electrons, the single electron trigger is used in combination with a single photon trigger [7]. The same photon trigger is used to select events for the photon control region.

The full list of triggers used, along with the L1 seeds and the associated primary datasets are

Data set name	Cross section (pb)	Order in QCD
EWKWMINUS2JETS_WTOLNU_M-50_TUNECP5_13TeV-madgraph-pythia8	20.35	LO
EWKWPPLUS2JETS_WTOLNU_M-50_TUNECP5_13TeV-madgraph-pythia8	25.81	LO
EWKZ2JETS_ZTOLL_M-50_TUNECP5_13TeV-madgraph-pythia8	4.321	LO
EWKZ2JETS_ZTONUNU_TUNECP5_13TeV-madgraph-pythia8	10.04	LO
AJJ_EWK_TUNECP5_13TeV_amcatnlo-pythia8	6.096	NLO
GJETS_MJJ-500_SM_5f_TUNECP5_EWK_13TeV-madgraph-pythia8	4.937	LO
GJETS_SM_5f_TUNECP5_EWK_13TeV-madgraph-pythia8	32.91	LO
TTJETS_TUNECP5_13TeV-amcatnloFXFX-pythia8	831.76	NLO
ST_t-channel_top_4f_inclusiveDecays_TUNECP5_13TeV-powhegV2-madspin-pythia8	137.458	NLO
ST_t-channel_antitop_4f_inclusiveDecays_TUNECP5_13TeV-powhegV2-madspin-pythia8	83.0066	NLO
ST_tW_top_5f_inclusiveDecays_TUNECP5_*_13TeV-powheg-pythia8	35.85	NLO
ST_tW_antitop_5f_inclusiveDecays_TUNECP5_*_13TeV-powheg-pythia8	35.85	NLO
WW_TUNECP5_13TeV-pythia8	75.91	LO
WZ_TUNECP5_13TeV-pythia8	27.56	LO
ZZ_TUNECP5_13TeV-pythia8	12.14	LO
QCD_HT1000to1500_TUNECP5_13TeV-madgraph-pythia8	1095	LO
QCD_HT1000to1500_TUNECP5_13TeV-madgraph-pythia8	1095	LO
QCD_HT100to200_TUNECP5_13TeV-madgraph-pythia8	2.369e+07	LO
QCD_HT1500to2000_TUNECP5_13TeV-madgraph-pythia8	99.27	LO
QCD_HT2000toInf_TUNECP5_13TeV-madgraph-pythia8	20.25	LO
QCD_HT200to300_TUNECP5_13TeV-madgraph-pythia8	1.554e+06	LO
QCD_HT300to500_TUNECP5_13TeV-madgraph-pythia8	324300	LO
QCD_HT300to500_TUNECP5_13TeV-madgraph-pythia8	324300	LO
QCD_HT500to700_TUNECP5_13TeV-madgraph-pythia8	29990	LO
QCD_HT50to100_TUNECP5_13TeV-madgraphMLM-pythia8	1.85e+08	LO
QCD_HT700to1000_TUNECP5_13TeV-madgraph-pythia8	6374	LO
QCD_HT700to1000_TUNECP5_13TeV-madgraph-pythia8	6374	LO

Table 3: Simulated datasets for the modelling of EWK V production and other processes. All datasets are accessed at the NanoAOD data tier from the v5 campaign, also known as 1Jun19.

Data set name	Cross section (pb)	Order in QCD
Gluglu_HToInvisible_M125_TUNECP5_13TeV_powheg_pythia8	48.58	NLO
VBF_HToInvisible_M125_13TeV_TUNECP5_powheg_pythia8	3.782	NLO

Table 4: Simulated signal datasets for the modelling of invisible Higgs boson decays. All datasets are accessed at the NanoAOD data tier from the v5 campaign, also known as 1Jun19.

121 shown in Table 5.

### 122 3.1 Efficiency measurement

#### 123 3.1.1 $p_T^{\text{miss}} + H_T^{\text{miss}}$ triggers

124 The performance of the  $p_T^{\text{miss}} + H_T^{\text{miss}}$  triggers is measured using single muon events. The events  
 125 are selected from the SingleMuon using the HLT\_IsoMu27 (HLT\_IsoMu24) trigger for 2017  
 126 (2018), and the offline muon is required to be well-identified and have  $p_T$  larger than 40 GeV.  
 127 The same selection is required as for the single-muon control region used in the final fit (cf. sec.  
 128 6.2):

- 129 1. Veto on additional leptons, photons, b jets,  $\tau_{had}$  candidates.  
 130 2.  $\Delta\phi(jet, \vec{p}_T^{\text{miss}}) > 0.5$  for the four leading jets with  $p_T > 30$  GeV.

Table 5: HLT paths and the associated L1 seeds used in the analysis for the 2017 and 2018 datasets. The HLT paths ending in “\_HT60” are backup triggers introduced to mitigate noise rate problems in 2017. Their inclusion is not strictly necessary for 2018, but is done for consistency.

Year	HLT path	L1 seed	Primary dataset
2017	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight	L1_ETMHF70	MET
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60	L1_ETMHF80_HTT60er	MET
	HLT_Ele35_WPTight_Gsf	L1_SingleEG24	SingleElectron
		L1_SingleEG30	
	HLT_Photon200	L1_SingleJet170	SinglePhoton
2018	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight	L1_SingleTau100er2p1	
		L1_ETM150	MET
		L1_ETMHF90_HTT60er	
	HLT_Ele32_WPTight_Gsf	L1_SingleIsoEG24er2p1	
		L1_SingleEG26er2p5	EGamma
		L1_SingleEG60	
	HLT_Photon200	L1_SingleEG34er2p5	
		L1_SingleJet160er2p5	
		L1_SingleJet180	
		L1_SingleTau120er2p1	
		L1_SingleEG60	EGamma

131 3. (Calo  $p_T^{\text{miss}}$ - PF  $p_T^{\text{miss}}$ ) / recoil  $\geq 0.5$

132 4.  $M_T(\ell, p_T^{\text{miss}}) < 160 \text{ GeV}$ .

133 5. Leading AK4 jet with  $p_T > 80 \text{ GeV}$ , passing the tight jet ID.

134 6. Subleading AK4 jet with  $p_T > 40 \text{ GeV}$ .

135 7.  $\Delta\eta_{jj} > 1.0$

136 8.  $\Delta\phi_{jj} < 1.5$

137 To understand the dependence of the efficiencies on the jet kinematics, we measured the efficiency in three different categories:

139 • Two central VBF jets: Leading two jets both satisfying  $|\eta| \leq 2.4$

140 • Two forward VBF jets: Leading two jets both satisfying  $|\eta| > 2.4$

141 • One central and one forward VBF jet: One of the leading jets is central with  $|\eta| \leq 2.4$  and the other jet in the pair is forward with  $|\eta| > 2.4$

143 In these three categories, efficiencies are measured as a function of  $M_{jj}$  and recoil, using both data and MC samples for 2017 and 2018, separately. Figs. 60, 61, 1 and 2 show the results.

145 Fig. 1 shows the efficiencies as a function of recoil in data and MC for the three categories, for 146 2017 samples, whereas Fig. 2 shows the efficiencies as a function of recoil in data and MC for the

<sup>147</sup> three categories, for 2018 samples. Efficiencies in double muon control region, and efficiencies  
<sup>148</sup> in both single and double muon regions as a function of  $M_{jj}$  can be found in B.

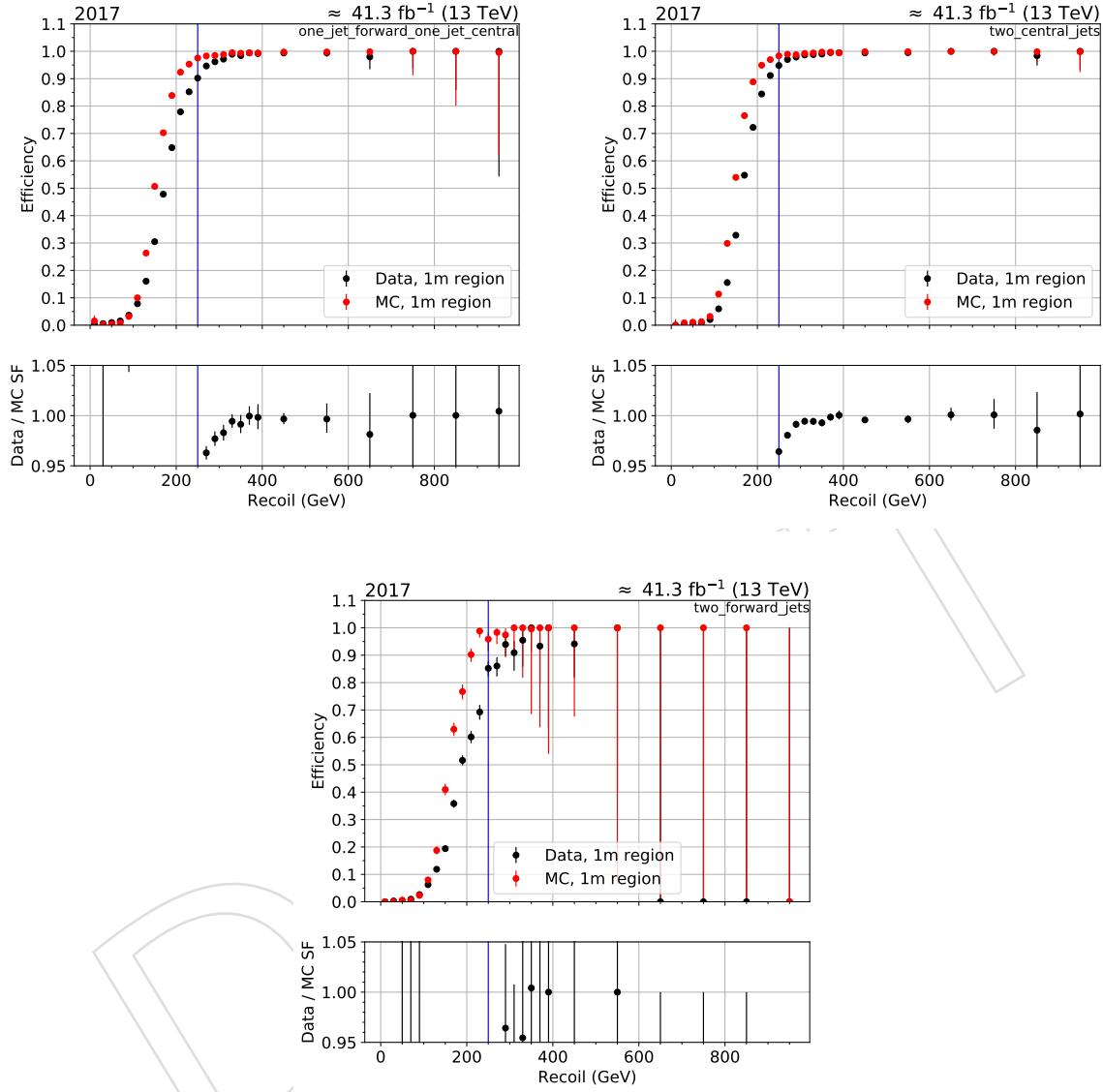


Figure 1: MET trigger efficiency as a function of recoil in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2017 data and MC samples with the selection of single muon events.

<sup>149</sup> Using the measured efficiencies in data and MC, data-MC scale factors are measured for the  
<sup>150</sup> three jet kinematic cases. Fig. 3 shows the scale factors as a function of recoil in single muon  
<sup>151</sup> control region, using 2017 (left) and 2018 (right) samples. We have low statistics for the two  
<sup>152</sup> forward jets category, but overall the scale factors are close to 1. The scale factors measured in  
<sup>153</sup> the double muon region can be found in B.

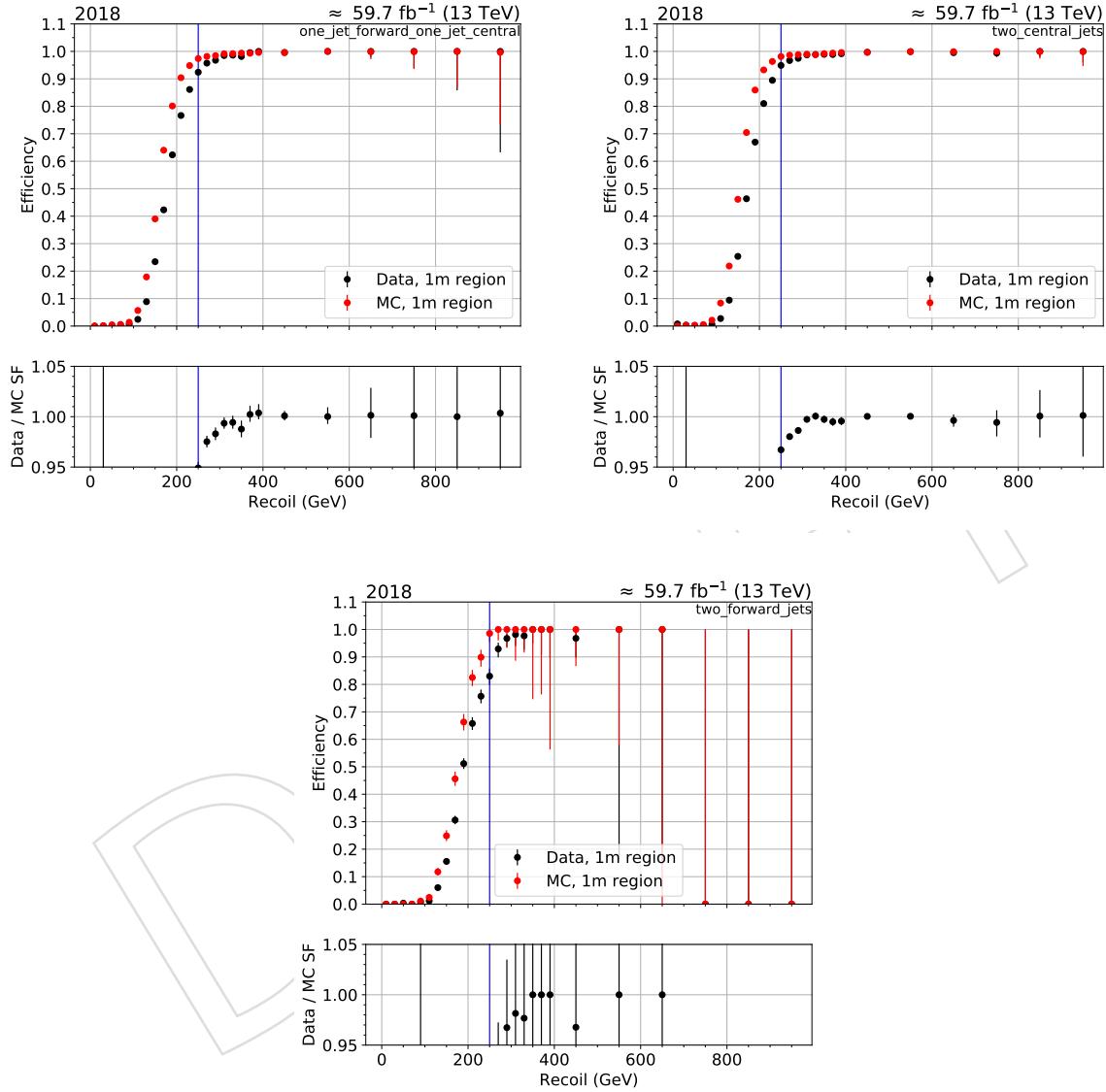


Figure 2: MET trigger efficiency as a function of recoil in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2018 data and MC samples with the selection of single muon events.

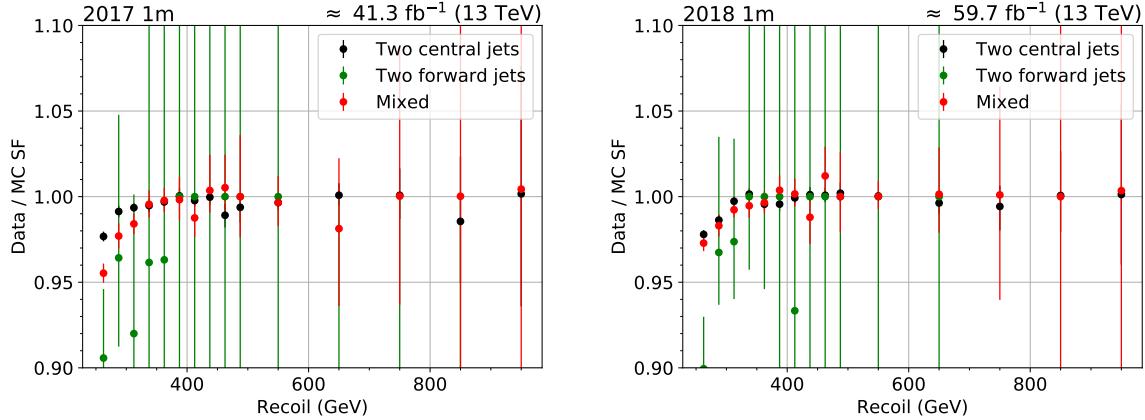


Figure 3: Data-MC scale factors for the three jet kinematic cases in the single muon control region, using 2017 (left) and 2018 (right) samples.

### 154 3.1.2 Photon trigger

155 The photon trigger efficiency is measured using events from the JetHT dataset collected with  
 156 the HLT\_PFHT1050 trigger, which was fully unprescaled in 2017 and 2018<sup>1</sup>. Events are selected  
 157 in the same way as for the photon analysis control region (cf. sec. 6.6), except for the photon  
 158  $p_T$ , recoil and trigger requirements. The trigger efficiency  $\epsilon$  is then determined as:

$$\epsilon(\text{HLT\_Photon200}) = \frac{\text{Offline selection \&& HLT\_PFHT1050 \&& HLT\_Photon200}}{\text{Offline selection \&& HLT\_PFHT1050}}$$

159 The resulting efficiency in data and GJets  $H_T$ -binned simulation is shown in Fig. 4. The trigger  
 160 efficiency in data is larger than 95% for a photon  $p_T$  of larger than 215 GeV, and larger than  
 161 99% for photon  $p_T$  larger than 400 GeV. Between 250 and 400 GeV, there is a slight inefficiency  
 162 amounting to approximately 1% at the most, with a larger amplitude in 2017 than in 2018. In  
 163 both years, the turn-on behavior is almost immediate in simulated events, resulting in an MC-  
 164 to-data scale factor almost entirely driven by the efficiency in data. The scale factor is within  
 165 1% of unity for all bins except the lowest 2017 bin at 215 GeV, where it deviates from unity by  
 166 about 4%.

167 Based on these results, the offline  $p_T$  cut for the photon in the photon control region is chosen  
 168 to be 215 GeV.

### 169 3.1.3 Electron trigger

170 We also need to add information about the single electron trigger efficiencies

<sup>1</sup>The other, prescaled HLT\_PFHTXXX paths yield lower statistical precision.

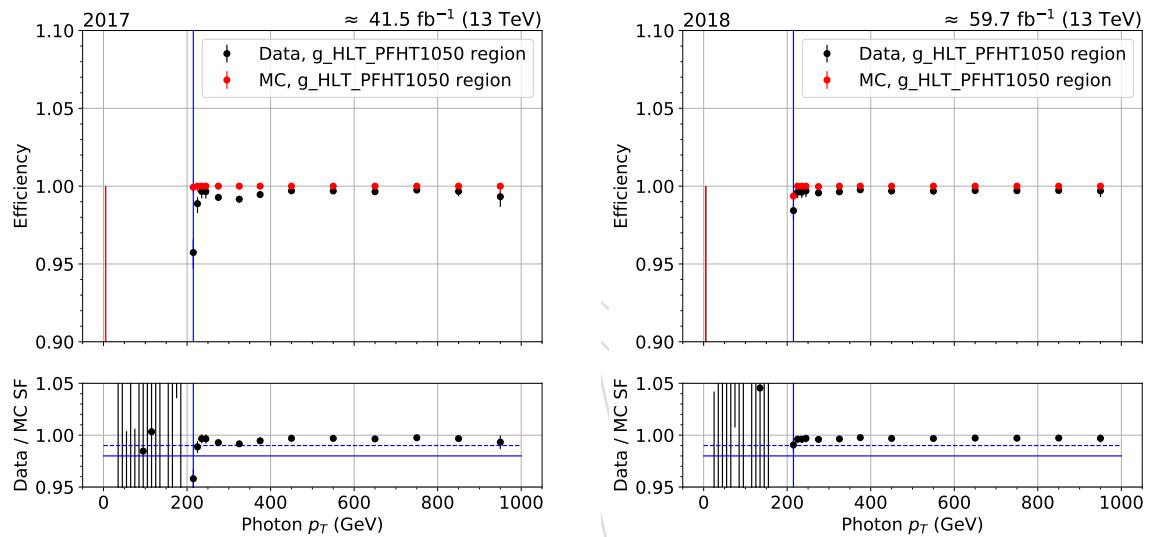


Figure 4: Efficiency of the `HLT_Photon200` trigger in data (black) and  $H_T$ -binned GJets simulation (red) for 2017 (left) and 2018 (right) as a function of photon  $p_T$ . The bottom panel shows the MC-to-data efficiency scale factor. The blue vertical line indicates a photon  $p_T$  of 215 GeV, which is the requirement used in the analysis selection.

## 171 4 Physics objects

172 All physics objects are used to identify signal-like events, to suppress backgrounds and to de-  
 173 fine control regions for background estimation. For the object definitions, we mostly follow the  
 174 CMS POG endorsed recommendations. The physics objects and the selection requirements are  
 175 described below.

### 176 4.1 Jets

177 In this analysis, jets are reconstructed by clustering PF candidates using the infrared and collinear  
 178 safe anti- $k_T$  algorithm [8]. Jets are clustered with a distance parameter of 0.4 and are referred  
 179 to as AK4 jets. The reconstructed vertex with the largest value of summed physics-object  $p_T^2$  is  
 180 taken to be the primary pp interaction vertex. The physics objects are those returned by a jet  
 181 finding algorithm [8, 9] applied to all charged PF candidates associated with the vertex, plus  
 182 the corresponding associated missing transverse momentum.

183 Jet momentum is determined as the vector sum of all particle momenta in the jet, and is found  
 184 from simulation to be within 5 to 10% of the true momentum over the full  $p_T$  spectrum and de-  
 185 tector acceptance. An offset correction is applied to jet energies to take into account the contri-  
 186 bution from additional proton-proton interactions within the same or nearby bunch crossings  
 187 (pileup). Jet energy corrections are derived from simulation, and are confirmed with *in situ*  
 188 measurements of the energy balance in dijet, multijet,  $\gamma$ +jet, and leptonic Z+jet events [10]. The  
 189 Fall17\_17Nov2017\_V32 and Autumn18\_V8 versions of the jet energy corrections are used for  
 190 the 2017 and 2018 data sets, respectively.

191 The AK4 jets used in this analysis are required to pass loose jet identification criteria. In ad-  
 192 dition, all the jets with  $p_T$  smaller than 50GeV must pass the medium pileup ID criteria. This  
 193 additional constraint on all AK4 jets is found to improve the modeling of jet distributions, es-  
 194 specially in the horn regions near  $|\eta| = 2.9$ . The effect of this requirement is demonstrated in  
 195 the sub-leading jet  $\eta$  distribution in Fig. 5.

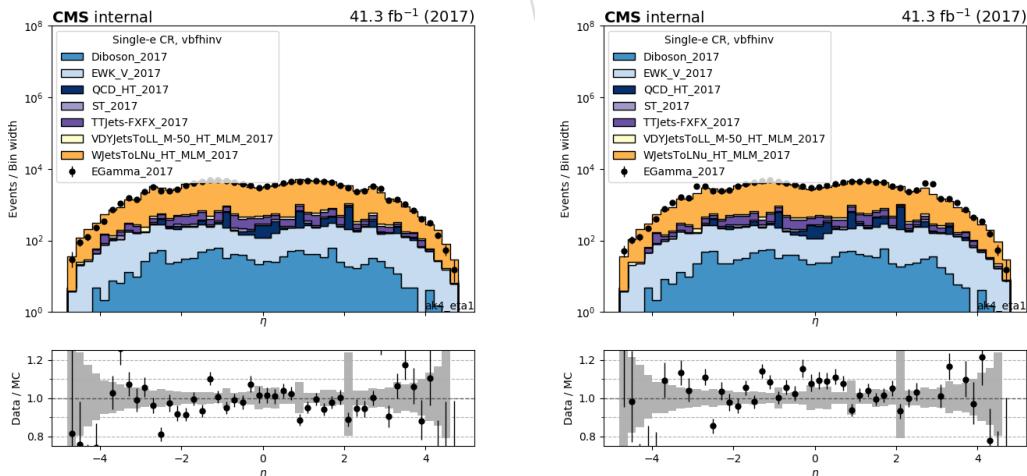


Figure 5: Subleading AK4 jet  $\eta$  distribution in single electron control region with pileup ID requirement (left) and without pileup ID requirement (right).

196 In addition to the pileup ID requirement, JER corrections are applied to all AK4 jets. Appli-  
 197 cation of this correction on all jets is observed to improve the modeling of the jets, especially

198 near the horn regions near  $|\eta| = 2.9$ . The effect of JER corrections on the subleading jet eta  
 199 distribution is shown in Fig. 6.

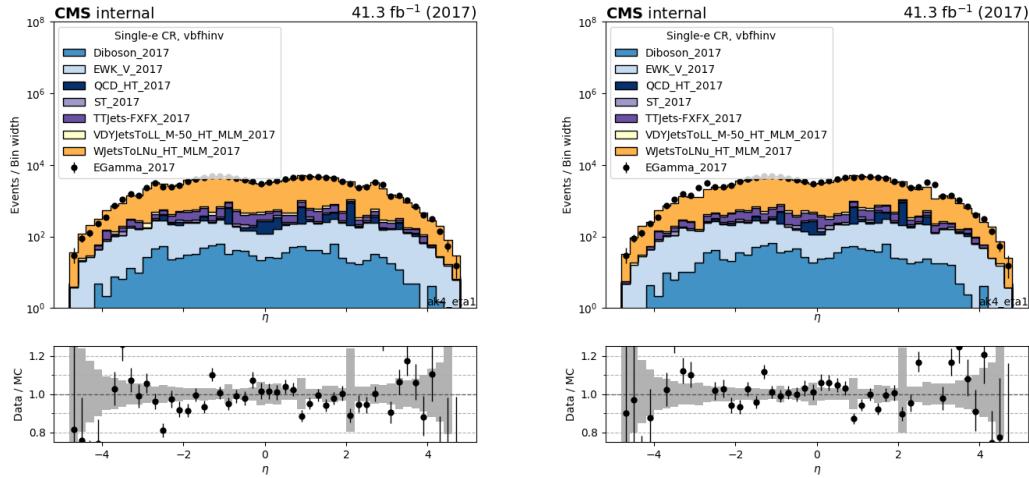


Figure 6: Subleading AK4 jet  $\eta$  distribution in single electron control region with JER corrections applied (left) and without JER corrections applied (right).

#### 200 4.1.1 b-tagged jets

201 In this analysis b jets with  $p_T > 20$  GeV and  $|\eta| < 2.4$  are identified using the “DeepCSV”  
 202 algorithm [11], adopting a working point (medium) corresponding to correctly identifying a  
 203 b-jet with a probability of 80%, and misidentifying a light-flavor jet with a probability of 10%.  
 204 This working point corresponds to the value of DeepCSV tagger to be greater than 0.4941.  
 205 Events with identified b jets are rejected to reduce the contamination from top quark processes.

## 206 4.2 Missing transverse momentum and recoil

207 The vector  $\vec{p}_T^{\text{miss}}$  is defined as the imbalance in the transverse momentum of all particles that  
 208 interact with the detectors. Due to momentum conservation in the plane transverse to the beam  
 209 axis,  $\vec{p}_T^{\text{miss}}$  corresponds to the transverse momentum that is carried by undetected particles such  
 210 as neutrinos. Practically,  $\vec{p}_T^{\text{miss}}$  is computed as the negative of the vectorial sum of transverse  
 211 momenta of all PF candidates and is therefore also referred to as PF  $\vec{p}_T^{\text{miss}}$ . The magnitude of  
 212 the  $\vec{p}_T^{\text{miss}}$  is referred to as  $p_T^{\text{miss}}$ .

213 Minimum energy thresholds in the calorimeters, inefficiencies in the tracker, nonlinearity of  
 214 the response of the calorimeter for hadronic particles can lead to an over- or underestimation  
 215 of  $p_T^{\text{miss}}$ . The bias on the  $p_T^{\text{miss}}$  measurement is reduced by propagating the effect of the jet  
 216 energy corrections introduced in section 4.1 according to

$$\vec{p}_T^{\text{miss}}(\text{corr}) = \vec{p}_T^{\text{miss}} - \sum_{\text{jets}} (\vec{p}_{T,\text{jet}}(\text{corr}) - \vec{p}_{T,\text{jet}}), \quad (1)$$

217 where the “corr” refers to the scale energy corrected measurements of the related objects.

218 This “type-I” correction for  $\vec{p}_T^{\text{miss}}$  uses jet energy scale corrections for all corrected jets with  
 219  $p_T > 15$  GeV that have less than 0.9 of their energy deposited in the ECAL. Furthermore, if a  
 220 muon is found in a jet, its 4-momentum is subtracted from the 4-momentum of the jet when  
 221 performing the correction and is added back to a corrected object.

Since signal events in this analysis contain only jets and no other reconstructed candidates,  $p_T^{\text{miss}}$  is equivalent to the total hadronic momentum in the event. For the leading backgrounds, this also corresponds to the transverse momentum of the W or Z boson. To mimic this behavior in the control regions of this analysis, the transverse momentum of the hadronic recoil  $\vec{U}$ , defined as the vectorial sum of the transverse momenta of all particles except the vector boson (or its decay products), is used. The variable is computed as

$$\vec{U} = \vec{p}_T^{\text{miss}} + \sum_{i \in \text{leptons, photons}} \vec{p}_T^i \quad (2)$$

where the sum takes into account the leptons and photons used to define the respective control region. The uncertainty of  $p_T^{\text{miss}}$  has a strong dependence on the event topology. Therefore, the uncertainty on  $p_T^{\text{miss}}$  is often factorized into its components of jets, leptons and unclustered energy. Each sub-component is then varied within its scale and resolution uncertainty.

Anomalous high- $p_T^{\text{miss}}$  events can appear due to various phenomena. In the ECAL, spurious deposits may appear due to particles striking sensors in the ECAL photodetectors, or from real showers with non-collision origins such as those caused by beam halo particles. ECAL dead cells can cause real energy to be missed, again leading to a spurious imbalance. In the HCAL, spurious energy can arise due to noise in the hybrid photodiode and readout box electronics, as well as direct particle interactions with the light guides and photomultiplier tubes of the forward calorimeter. A number of filters has been developed by the POG/DPG groups to identify and suppress anomalous high  $p_T^{\text{miss}}$  events [12]. The recommended filters are listed in Tab. 6 and are applied in the analysis.

Table 6: The  $p_T^{\text{miss}}$  filters recommended by the JME POG [12]. The recommendations apply to both 2017 and 2018. Except for the bad super cluster filter (“ee badSC”), all filters are applied both in data and simulation.

Filter	Name in NanoAOD	Applied in data (MC)
primary vertex filter	Flag_goodVertices	✓(✓)
beam halo filter	Flag_globalSuperTightHalo2016Filter	✓(✓)
HBHE noise filter	Flag_HBHENoiseFilter	✓(✓)
HBHEiso noise filter	Flag_HBHENoiseIsoFilter	✓(✓)
ECAL TP filter	Flag_EcalDeadCellTriggerPrimitiveFilter	✓(✓)
Bad PF Muon Filter	Flag_BadPFMuonFilter	✓(✓)
ee badSC noise filter	Flag_eeBadScFilter	✓(✗)
ECAL bad calibration filter update	Flag_ecalBadCalibFilterV2	✓(✓)

### 4.3 Leptons

#### 4.3.1 Electrons

Electrons within the geometrical acceptance of  $|\eta| < 2.5$  are reconstructed by associating tracks reconstructed in the silicon detector with clusters of energy in the ECAL [13]. Well-identified electron candidates are required to satisfy additional identification criteria based on the shower shape of the energy deposit in the ECAL and the consistency of the electron track with the

primary vertex [14]. Electron candidates that are identified as coming from photon conversions in the detector material are removed. An isolation variable is calculated based on the sum of the energies of the PF candidates within a cone of  $\Delta R < 0.3$  around the electron. The mean energy deposit in the isolation cone of the electron coming from pileup is estimated following the method described in Ref. [13] and subtracted from the isolation sum. In this note, ‘veto’ [15] electrons with a minimum  $p_T$  of 10 GeV are selected with an average efficiency of 95% and their presence is used as a condition to reject events, whereas ‘tight’ [15] electrons with a minimum  $p_T$  of 40 GeV and an average efficiency of 70% are used to select the events in the control regions. Full selection criteria are shown in Table 7.

Table 7: Tight and veto electron identification criteria.

Variable	Selection Tight	Selection Veto
	Barrel (Endcaps)	Barrel (Endcap)
Full 5x5 $\sigma_{i\eta i\eta}$	$< 0.0104$ ( $< 0.0353$ )	$< 0.0126$ ( $< 0.0457$ )
$ \Delta\eta_{in} $	$< 0.00255$ ( $< 0.00501$ )	$< 0.00463$ ( $< 0.00814$ )
$ \Delta\phi_{in} $	$< 0.022$ ( $< 0.0236$ )	$< 0.148$ ( $< 0.19$ )
H/E	$< 0.026 + 1.15/E_{SC} + 0.0324\rho/E_{SC}$ ( $< 0.0188 + 2.06/E_{SC} + 0.183 * \rho/E_{SC}$ )	$< 0.05 + 1.16/E_{SC} + 0.0324\rho/E_{SC}$ ( $< 0.05 + 2.54/E_{SC} + 0.183\rho/E_{SC}$ )
Relative isolation ( $\rho$ correction)	$< 0.0287 + 0.506/p_T$ ( $< 0.0445 + 0.963/p_T$ )	$< 0.198 + 0.506/p_T$ ( $< 0.203 + 0.963/p_T$ )
$1/E - 1/p$	$< 0.159$ ( $< 0.0197$ )	$< 0.209$ ( $< 0.132$ )
$ d_{xy}(\text{vtx}) $	$< 0.050$ ( $< 0.100$ )	$< 0.050$ ( $< 0.100$ )
$ d_z(\text{vtx}) $	$< 0.100$ ( $< 0.200$ )	$< 0.100$ ( $< 0.200$ )
Expected Inner Missing Hits	$\leq 1$ ( $\leq 1$ )	$\leq 2$ ( $\leq 3$ )
Pass conversion veto	Yes (Yes)	Yes (Yes)

### 4.3.2 Muons

Muons within the geometrical acceptance of  $|\eta| < 2.4$  are reconstructed by combining information from the silicon tracker and the muon system [16]. The muons are required to pass set of quality criteria based on the number of spatial points measured in the tracker and in the muon system, the fit quality of the muon track and its consistency with the primary vertex of the event. Similar to electron case, the isolation requirements for muons are also based on the sum of the energies of the PF candidates, but a different cone size of a  $\Delta R < 0.4$  is used.

257 The muon isolation variable is corrected for pileup effects by subtracting half of the sum of the  
 258 transverse momenta of charged particles that are inside the isolation cone and not associated  
 259 with the primary vertex. In this note, “loose” [17] muons with  $p_T > 10$  GeV are selected with  
 260 an average efficiency of 98% and are used as a condition to reject events, whereas “tight” [18]  
 261 muons with  $p_T > 20$  GeV are selected with an average efficiency of 95% and are used to select  
 262 events in the control samples. A full list of tight identification criteria is given here:

- 263 • Muon reconstructed as a global muon
- 264 • Muon reconstructed as a particle flow muon
- 265 • Normalized  $\chi^2$  of the global track less than 10
- 266 • At least one muon chamber hit included in the global track fit
- 267 • Muon segments in at least two muon stations
- 268 • Transverse impact parameter w.r.t. the primary vertex less than 2 mm.
- 269 • Longitudinal impact parameter w.r.t. the primary vertex less than 5 mm.
- 270 • At least one pixel hit
- 271 • Hits on at least 5 tracker layers
- 272 •  $\Delta\beta$  relative isolation less than 0.15

### 273 4.3.3 Taus

274 Hadronically decaying  $\tau$  leptons are required to pass identification criteria using the hadron-  
 275 plus-strips algorithm [19]. The algorithm identifies a jet as an hadronically decaying tau lepton  
 276 candidate if a subset of the particles assigned to the jet is consistent with the decay products  
 277 of a  $\tau$  candidate. Candidate  $\tau$  jets are required to pass both the “DecayModeNewDMs” and  
 278 “DecayMode” identifiers.

279 In addition,  $\tau$  candidates are required to be isolated from other activity in the event. The isolat-  
 280 ion requirement is computed by summing the  $p_T$  of the charged PF candidates and PF photon  
 281 candidates within an isolation cone of  $\Delta R = 0.5(0.3)$ , around the tau candidate direction. The  
 282 charged and photon candidates associated with the tau candidate are removed from this sum  
 283 and further described in Ref. [19]. The “VLoose\_IsolationMVArun2v1DBoldDMwLT” isolation  
 284 working point [20] is employed in this analysis for tau candidates with  $p_T$  larger than 20 GeV  
 285 within  $|\eta| < 2.3$ .

## 286 4.4 Photons

287 Photon candidates are reconstructed from energy deposits in the ECAL using algorithms that  
 288 constrain the clusters to the size and shape expected from a photon [21]. The identification of  
 289 the candidates is based on shower-shape and isolation variables. For isolated photons, scalar  
 290 sums of the  $p_T$  of PF candidates within a cone of  $\Delta R < 0.3$  around the photon candidate are  
 291 required to be below the bounds defined. Only the PF candidates that do not overlap with the  
 292 EM shower of the candidate photon are included in the isolation sums.

293 Two candidate definitions are employed. “Loose” photons are used to reject events with un-  
 294 wanted photons. These photons are required to pass the EGamma POG ‘loose’ identification  
 295 criteria [22], have a transverse momentum of at least 15 GeV and be within  $|\eta| < 2.5$ . The exact  
 296 identification criteria are summarized in Table 8. A “tight” photon definition is used for the  
 297 photon in the dedicated photon control region. These photons are required to be in the barrel  
 298 ( $|\eta| < 1.479$ ), and have  $p_T > 215$  GeV. The POG medium ID is employed, which is summarized  
 299 in Table 9.

Variable	Selection Barrel (Endcap)
Full 5x5 $\sigma_{i\eta i\eta}$	$< 0.0106 (< 0.0272)$
H/E	$< 0.04596 (< 0.0590)$
charged hadron isolation	$< 1.694 (< 2.089)$
neutral hadron isolation	$< 24.032(19.722) + 0.01512(0.0117) \times p_T + 2.259(2.3) \times 10^{-5} \times p_T^2$
photon isolation	$< 2.876(4.162) + 0.004017(0.0037) \times p_T$
Conversion safe electron veto	Yes (Yes)

Table 8: Loose photon identification criteria.

Variable	Selection Barrel
Full 5x5 $\sigma_{i\eta i\eta}$	$< 0.01015$
H/E	$< 0.02197$
charged hadron isolation	$< 1.141$
neutral hadron isolation	$< 1.189 + 0.01512 \times p_T + 2.259 \times 10^{-5} \times p_T^2$
photon isolation	$< 2.08 + 0.004017 \times p_T$
Conversion safe electron veto	Yes

Table 9: Tight photon identification criteria. The criteria are only given for the barrel region since endcap photons are not taken into account.

#### 4.4.1 Photon purity

Photons are reconstructed from ECAL clusters, and can be discriminated from other sources of ECAL deposits due to the properties of the cluster, as well as their lack of other associated signatures such as tracks or HCAL deposits. This discrimination is not perfect, however, and in some cases, non-photon objects will incorrectly be identified as photons (“fakes”). The leading source of fake photons is QCD production of multijet events, where a jet is misidentified. This process is relevant mainly because of its large cross-section, which yields non-negligible contributions to the photon selection even if the per-jet probability of misreconstruction is small.

To estimate the fake contribution to the photon control region, a purity measurement is performed. The photon purity is defined as the fraction of reconstructed photons that is actually due to an isolated photon from the hard scattering event, as opposed to a fake. The purity is obtained from a template fit to the distribution of the  $\sigma_{i\eta i\eta}$  variable in data, which  $\sigma_{i\eta i\eta}$  variable represents the width of the ECAL shower in the  $\eta$  direction. Due to the different shower behavior of photons and hadrons, the  $\sigma_{i\eta i\eta}$  distribution shows characteristically different behavior between these two classes of reconstructed photons: Real photons show a large peak with a cut-off around  $\sigma_{i\eta i\eta} \approx 0.01$ , while fake photons will have a smaller peak (stemming from actual photons inside jets), and an additional non-peaking bulk region at larger  $\sigma_{i\eta i\eta}$  (stemming from hadrons interacting in the ECAL). The inputs to the fit are defined as follows:

- Photons in data are selected by applying the same identification criteria as for the tight selection defined above, with the exception of the  $\sigma_{i\eta i\eta}$  requirement. By removing this requirement, the full  $\sigma_{i\eta i\eta}$  distribution can be observed.
- A real photon template is obtained from GJETS simulation. The same identification criteria are applied as in data.
- A fake photon template is obtained from data. In this case, the identification criteria are modified by requiring that at least one of the isolation criteria is not passed.

325 Therefore, the photons in this template do not overlap with the 'data' template,  
 326 and represent reconstructed photons inside jets.

327 The templates are derived in separate bins of the photon  $p_T$  and the measurement is performed  
 328 separately for 2017 and 2018. Events are selected using the following criteria:

- 329 • At least one jet with  $p_T > 100$  and  $|\eta| < 2.4$ , which is not overlapping with a photon  
 330 of interest.
- 331 • The event must pass the HLT\_Photon200 trigger, as well as the MET filters.
- 332 • The event must have  $p_T^{\text{miss}} < 60 \text{ GeV}$ .

333 The fits are shown in Figs. 7 and 8, and the resulting purity as a function of photon  $p_T$  is shown  
 334 in Fig. 9. While the fits are performed in the full range of  $\sigma_{\eta\eta\eta} < 0.15$ , the purity is evaluated  
 335 only taking into account the contributions with  $\sigma_{\eta\eta\eta} < 0.01015$ , which is the requirement posed  
 336 in the identification criteria used in the analysis. The resulting purity ranges between 2 and 4%,  
 337 with a decreasing trend with photon  $p_T$ .

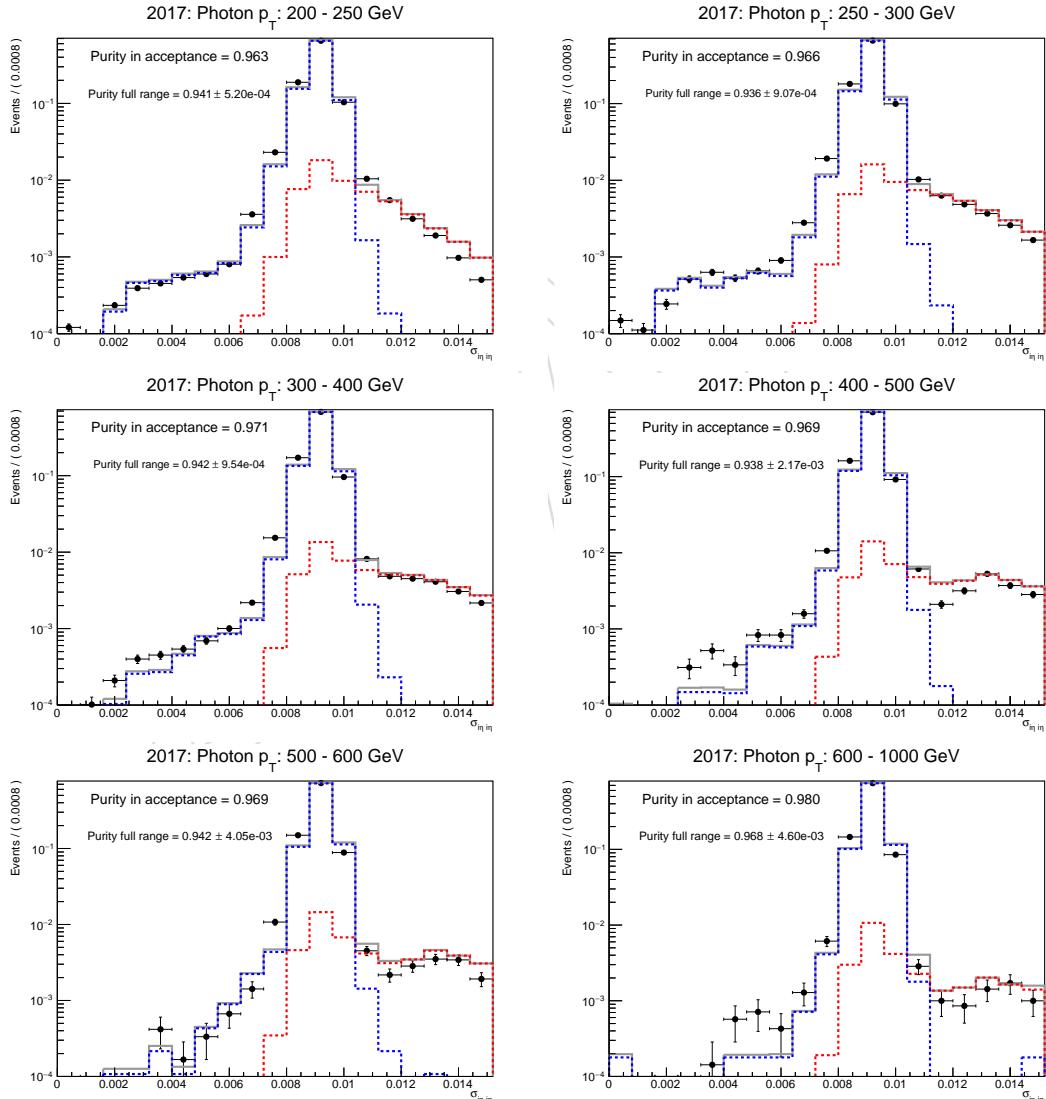


Figure 7: Template fits used to determine the photon purity in the 2017 dataset. The fits are shown in bins of photon  $p_T$ , increasing from left to right and top to bottom.

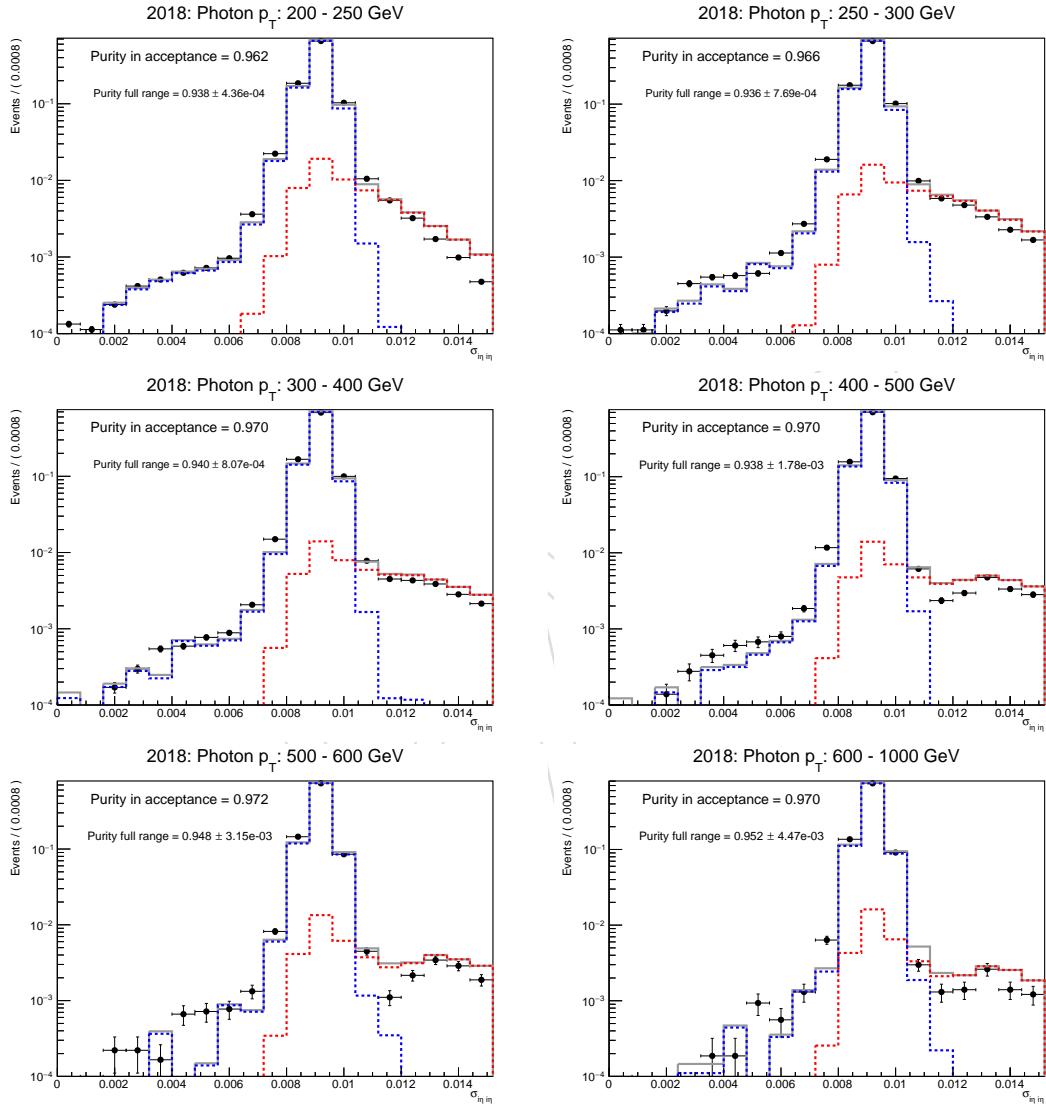


Figure 8: Template fits used to determine the photon purity in the 2018 dataset. The fits are shown in bins of photon  $p_T$ , increasing from left to right and top to bottom.

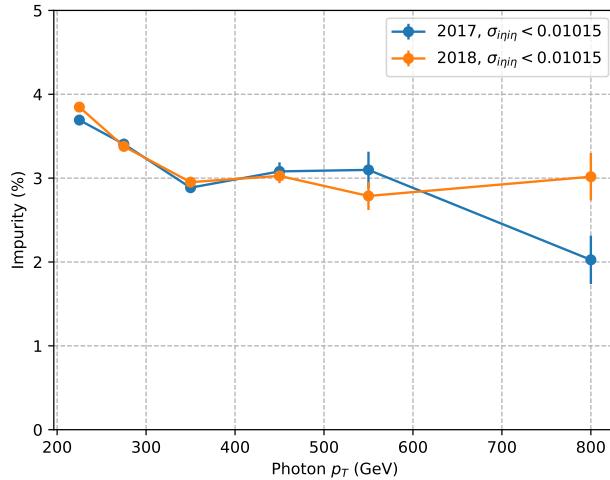


Figure 9: Photon purity as a function of photon  $p_T$ . This figure shows the aggregated results obtained from the fits shown in Figs. 7 and 8.

## 338 5 Reweighting of simulated events

339 Simulated signal and background samples are corrected for various effects through reweighting  
340 procedures outlined in this section.

### 341 5.1 Trigger efficiency reweighting

342 The efficiency is calculated as a function of the recoil and  $M_{jj}$ . The trigger is found to be more  
343 than 95% efficient for events with a recoil larger than 250 GeV, and more than 99% efficient for  
344 events with a recoil larger than 375 GeV. The MC-to-data scale factor is found to be within 1%  
345 of unity everywhere except for the lowest recoil bin at 250 GeV, where it is within 2% (c.f. sec.  
346 3.1).

### 347 5.2 Pileup reweighting

348 The pileup (PU) conditions in the simulated samples are not identical to the ones observed  
349 measured in data, and a reweighting is applied to remove the difference. The reweighting is  
350 performed by matching the true pileup distribution of each simulated sample with the pileup  
351 distribution in data, obtained through the pileupCalc tool assuming a minimum bias cross  
352 section of  $69.2 \pm 4.6\%$  mb, following the recommendations in Ref. [23]. The true pileup dis-  
353 tributions in data and simulation are shown in Fig. 10. The distribution of the number of  
354 reconstructed vertices for  $W \rightarrow \mu\nu$  events before and after PU reweighting is shown in Fig. 11.  
355 In this variable, the PU reweighting method leads to a worse overall agreement between data  
356 and simulation. To check this behavior, the distribution of the event energy density  $\rho$  is shown  
357 in Fig. 11, again before and after PU reweighting. Here, the agreement before PU reweighting  
358 is worse than in the primary vertex distribution and the PU reweighting clearly improves the  
359 agreement.

### 360 5.3 Lepton and photon identification/reconstruction efficiency reweighting

361 Data-to-simulation scale factors are applied to events in the control regions to account for dif-  
362 ferences in the reconstruction, identification and isolation of leptons between data and simu-  
363 lationn. These data-to-MC scale factors are derived from the efficiencies that are measured for

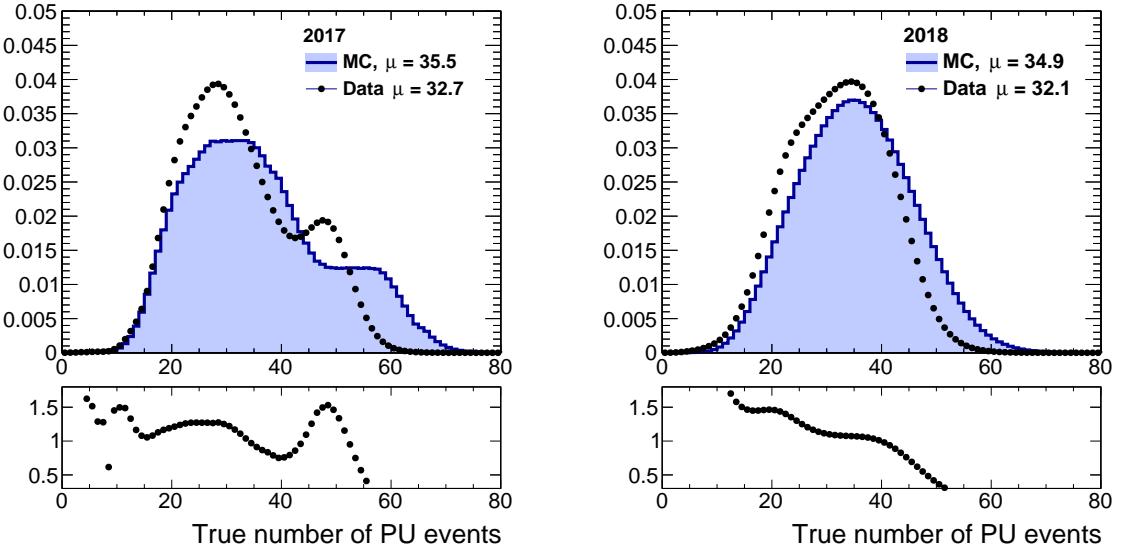


Figure 10: Distribution of the true number of PU events in data and simulation for 2017 (left) and 2018 (right). The distributions for data are extracted assuming a minimum bias cross section of 69.2 mb.

- 364 the electron and muon selections in bins of  $p_T$  and  $\eta$  in both data and simulation. These scale  
 365 factors are provided by the relevant POGs.  
 366 The reconstruction scale factors for electrons are shown in Fig. 13. The corresponding identifi-  
 367 cation scale factors for veto and tight electrons are shown in Fig. 14, and include the effect of  
 368 the isolation efficiency.  
 369 The identification scale factors for muons are shown in Fig. 15. Here, isolation scale factors  
 370 are applied separately and are shown in Fig. ???. The corresponding corrections for muons are  
 371 deemed negligible [24].  
 372 The scale factors for id and isolation for tight muons are shown in Fig. ??.

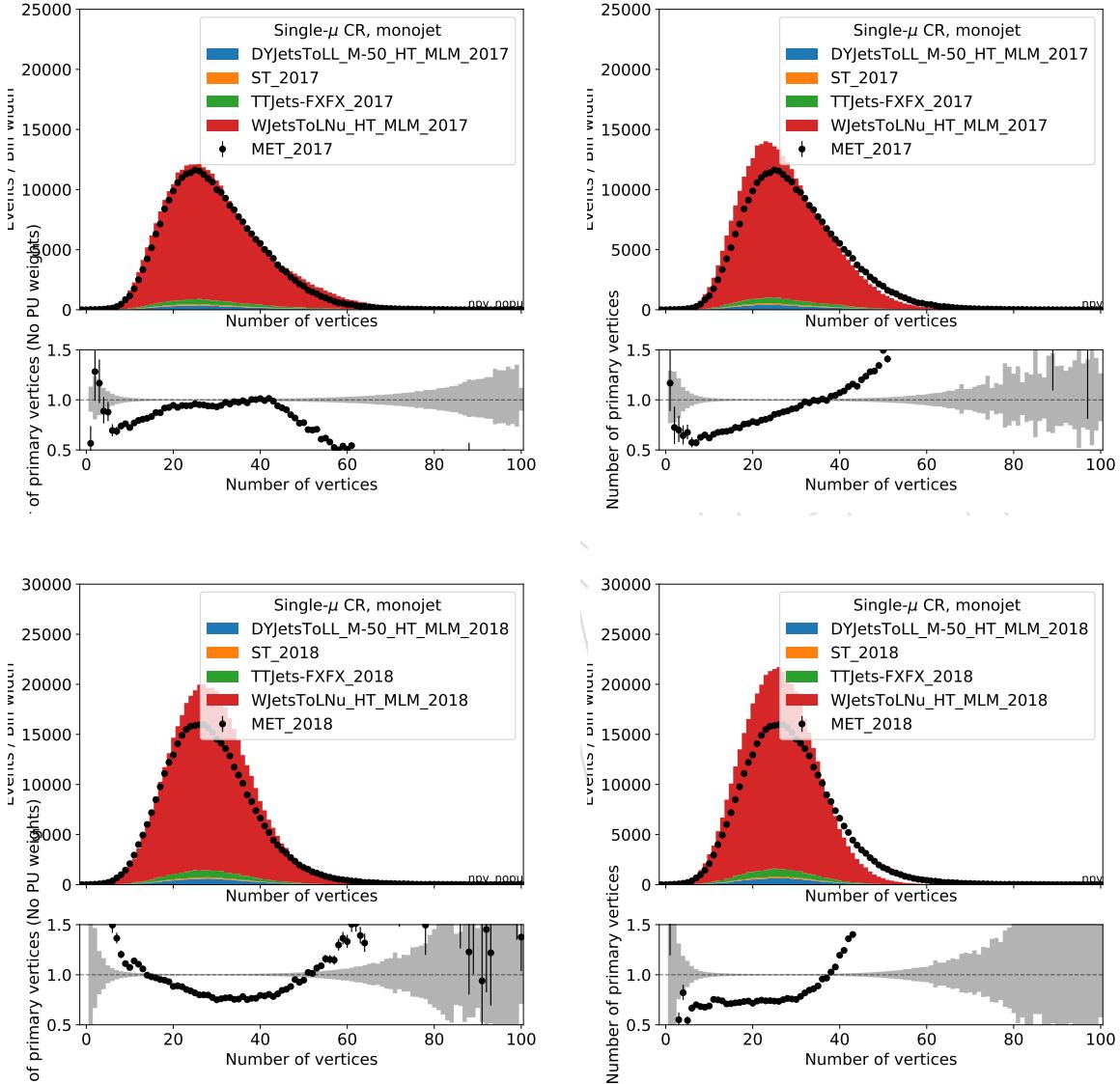


Figure 11: Distribution of the number of vertices in  $W \rightarrow \mu\nu$  events in data and simulation before pileup re-weighting (left) and after pileup reweighting (right). The Monte Carlo is normalized to the luminosity of  $41.53$  and  $59.7 \text{ fb}^{-1}$ , respectively for 2017 and 2018.

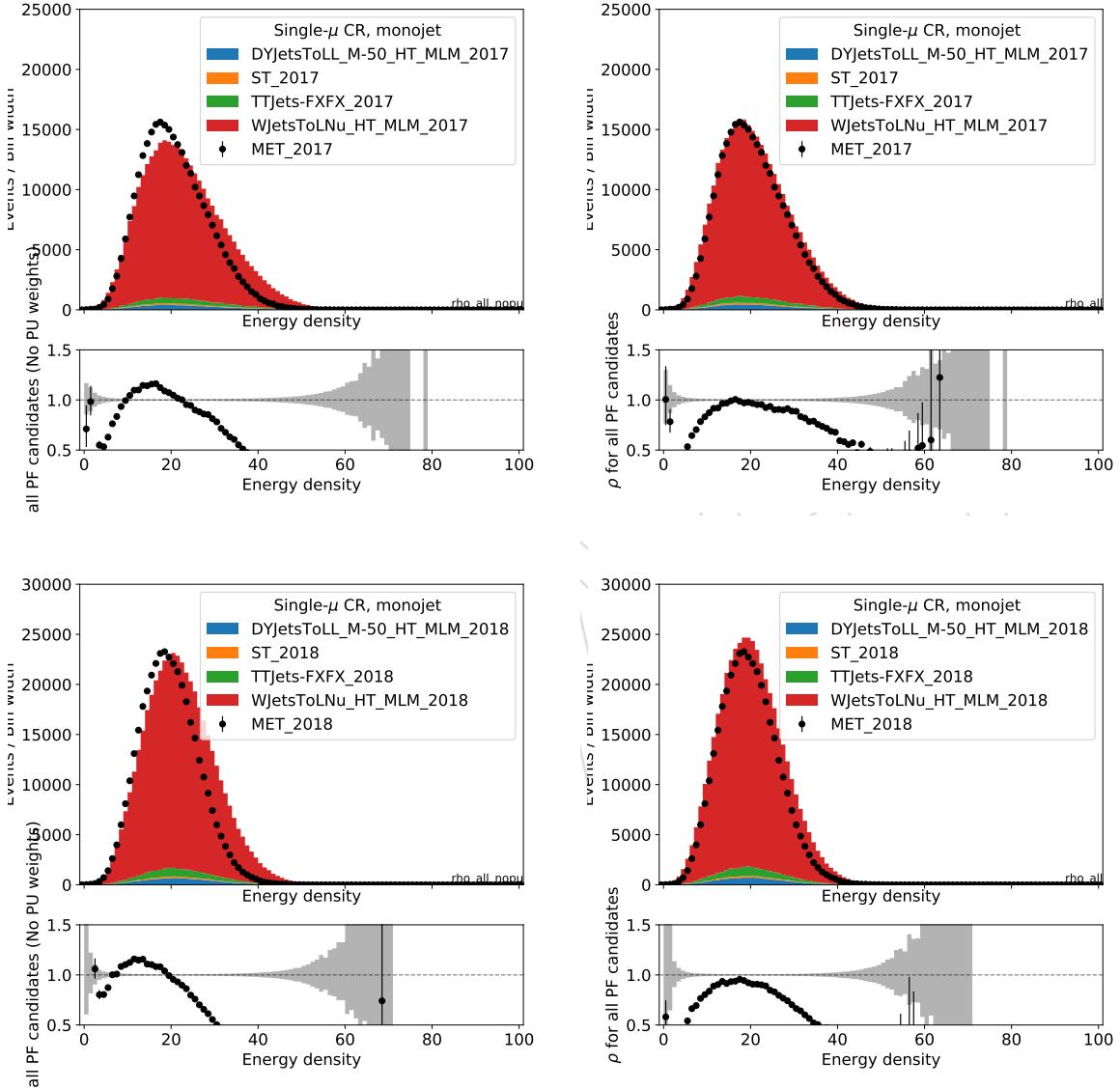


Figure 12: Distribution of the event energy density  $\rho$  in  $W \rightarrow \mu\nu$  events in data and simulation before pileup re-weighting (left) and after pileup reweighting (right). The Monte Carlo is normalized to the luminosity of  $41.53$  and  $59.7 \text{ fb}^{-1}$ , respectively for 2017 and 2018.

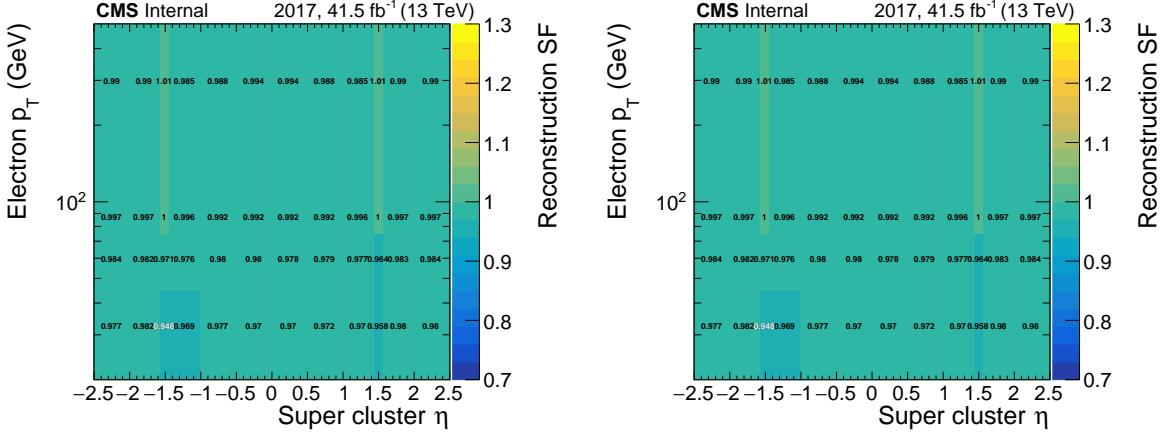


Figure 13: Scale factors for the reconstruction efficiency of electrons starting from a super cluster for 2017 (left) and 2018(right)

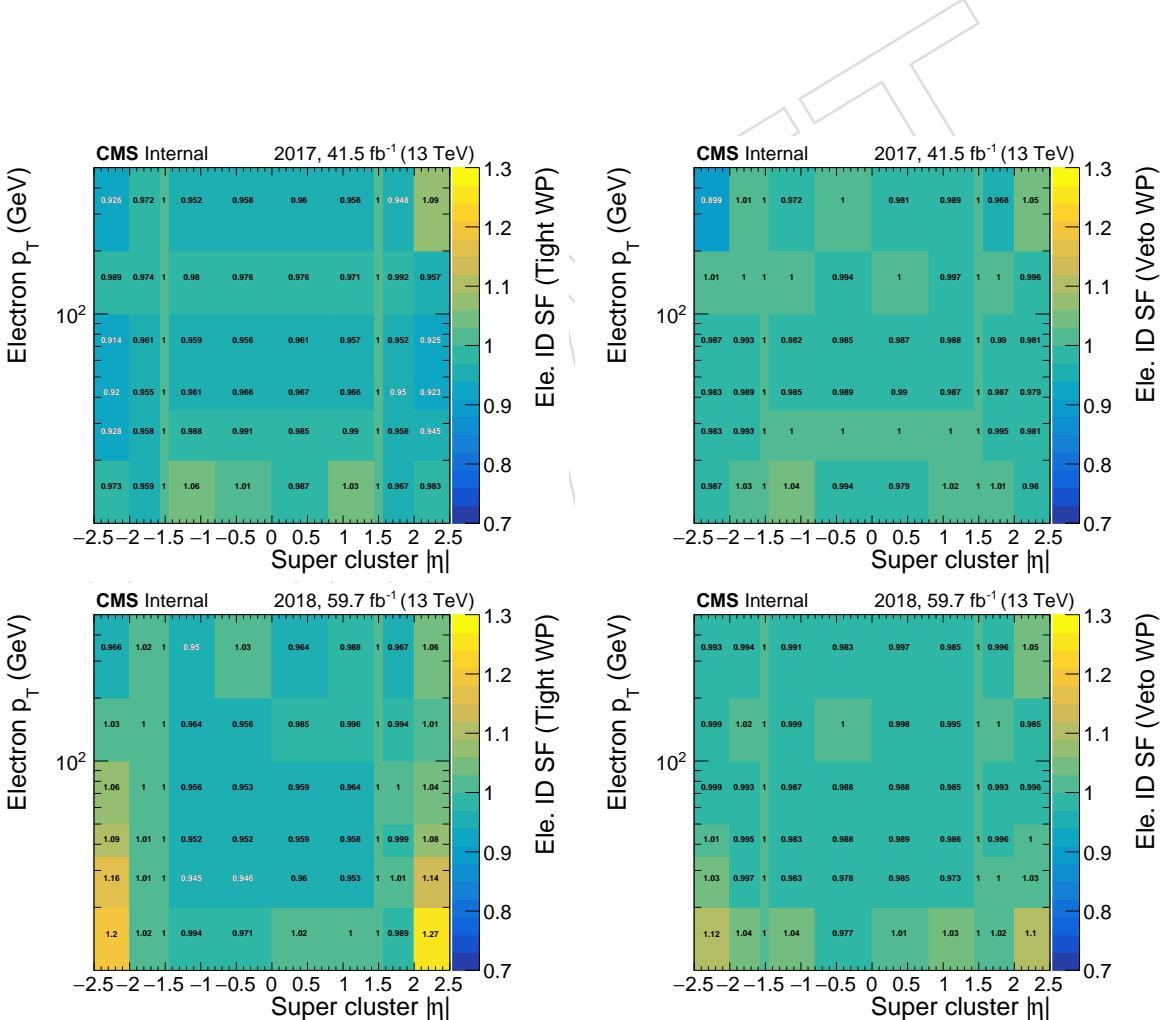


Figure 14: Scale factors for tight (left) and veto (right) electrons are shown for 2017 (top) and 2018 (bottom). The scale factors are provided in bins of electron  $p_T$  and  $\eta$ .

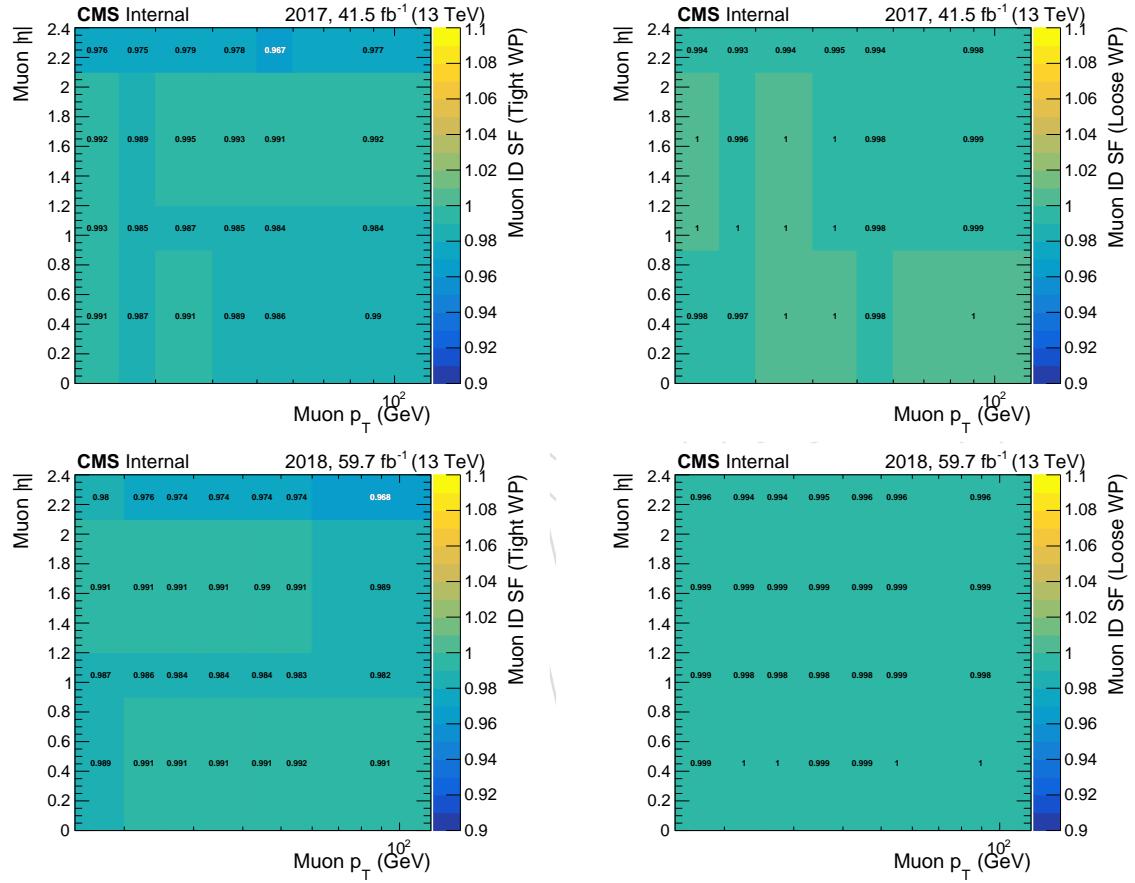


Figure 15: Scale factors for tight (left) and veto (right) muon identification are shown for 2017 (top) and 2018 (bottom). The scale factors are provided in bins of electron  $p_T$  and  $\eta$ .

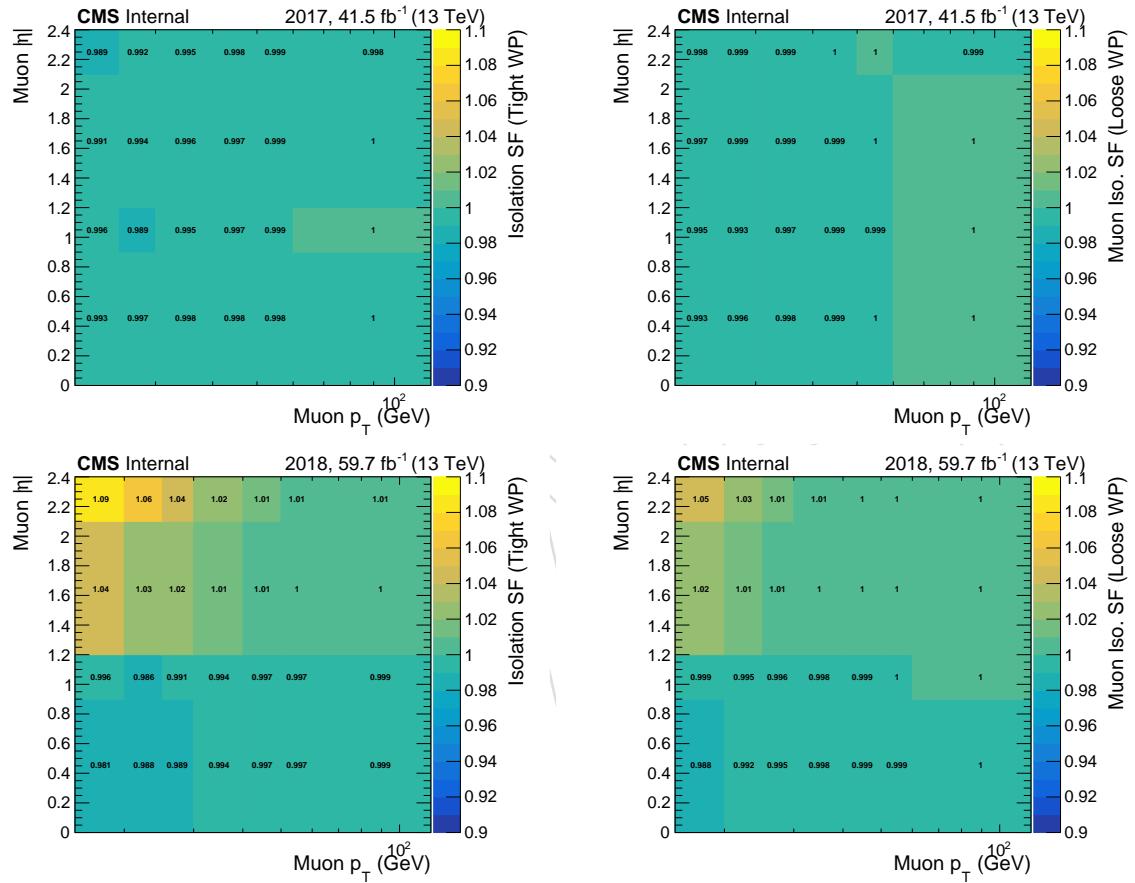


Figure 16: Scale factors for tight (left) and veto (right) muon isolation are shown for 2017 (top) and 2018 (bottom). The scale factors are provided in bins of electron  $p_T$  and  $\eta$ .

373 **5.4 Higher-order reweighting**

374 This analysis uses the ratios of the recoil distributions in signal and control regions to con-  
 375 strain the final background estimate in a partially data driven way. As signal and control re-  
 376 gions both have large statistical power, precise predictions of these ratios are necessary. To  
 377 achieve this goal, the LO simulation samples for the samples W, DY and photon backgrounds  
 378 are reweighted using higher-order corrections separately corresponding to NLO QCD, NLO  
 379 EW and NNLO QCD terms. The individual corrections are described in more detail in this  
 380 section. A concise overview of which corrections are applied to which processes is given in  
 381 Tab. 10.

Table 10: Summary of higher-order corrections applied to simulated samples. For each boson production process, separate samples and corrections are available for the EWK and QCD production modes. “MC order” reflects the perturbative order used in the generation of the simulation sample, while the further columns represent corrections applied on a per-event level in the analysis process.

Boson	production mode	MC order	NLO QCD	NNLO QCD	NLO EWK
Z	QCD	LO	✓	✓	✓
	EWK	LO	✓	–	–
W	QCD	LO	✓	✓	✓
	EWK	LO	✓	–	–
$\gamma$	QCD	LO	✓	✓	✓
	EWK	LO	–	–	–

382 **5.4.1 Generator-level boson construction**

383 All theory-based corrections of the W, DY and photon backgrounds are parametrized as a func-  
 384 tion of the generator-level  $p_T$  of the respective boson  $p_{T,V}$ . For each simulated event, this quan-  
 385 tity is calculated as follows. For DY and W samples, generator-level dilepton candidates are  
 386 built from:

- 387 1. “dressed” final-state electrons and muons. Lepton dressing means to collect all photons  
 388 radiated off the lepton within a cone of  $\Delta R < 0.1$  and adding their four-momenta back  
 389 to the lepton four-momentum. This procedure is meant to undo the effect of final state  
 390 photon radiation, which would otherwise distort the value of the reconstructed boson  
 391 four-momentum. This effect is especially relevant as electrons and muons follow differ-  
 392 ent radiation patterns. Lepton dressing is performed in central NanoAOD production  
 393 following the procedure used in the RIVET software.
- 394 2.  $\tau$  leptons with generator status 2. As  $\tau$  leptons are unstable, they are not present as final  
 395 state particles (status 1) in the generator record. The  $\tau$  lepton before its decay has status  
 396 2.
- 397 3. neutrinos with generator status 1.

- 398 The dilepton candidates are checked for flavour consistency with the desired boson candidate.  
 399 If multiple candidates are found in an event, the one with the highest invariant mass is used.
- 400 For photon events, the generator photon with highest  $p_T$  and status 1 is used.

401 **5.4.2 QCD NLO corrections to QCD V processes**

402 Scale factors corresponding to NLO QCD corrections for W and Z production are obtained  
 403 from centrally produced CMS samples. For the DY and W processes, samples from the “Fall17”  
 404 campaign are used, while “Summer16” samples are used for the  $\gamma$ +jets process. In both cases,  
 405 all samples are generated using `MadGraph5_amc@NLO`. The LO samples are binned in HT and  
 406 are equivalent to the ones used in the analysis, and are generated with up to four partons in  
 407 the matrix element. The NLO samples are generated with up to two additional partons in the  
 408 matrix element calculation. Further jet multiplicities are handled by the parton shower, which  
 409 in both cases is performed using `Pythia8` with tune `CP5 (CUETP8M1)` for Fall17 (Summer16).

410 The scale factors are derived by obtaining the distribution of interest at the generator-level in  
 411 both samples, normalizing the distributions to their respective cross sections, and then dividing  
 412 them as  $SF = NLO / LO$ . Identical selection criteria are applied to both samples based on the  
 413 generator-level boson and generator-level AK4 jets, which are clustered using all visible gen-  
 414 erator particles with status 1. Generator-level jets overlapping with a lepton or photon present  
 415 in the LHE record are rejected. The event selection requirements are:

- 416 1. At least two generator-level jets must be present, with the leading (trailing)  $p_T$  of at least  
   417 80 GeV (40 GeV) and  $|\eta| < 4.7$ .
- 418 2. The angular difference in the transverse plane between the two jets is required to be small:  
   419  $\Delta\phi(jj) < 1.5$ .
- 420 3. The two leading jets are required to be well separated in pseudorapidity,  $|\eta_1 - \eta_2| > 1$ ,  
   421 and be located in opposite hemispheres of the detector,  $\eta_1 \times \eta_2 < 0$ .
- 422 4. The difference in the azimuthal angle ( $\Delta\phi$ ) between the boson and the four leading jets in  
   423 the event is required to be larger than 0.5. Only jets with  $p_T > 30$  GeV and  $|\eta| < 2.4$  are  
   424 considered.

425 Compared to an inclusive derivation of the SF, the inclusion of the selection criteria leads to an  
 426 increase in the value of the SF of approximately five percentage points.

427 The scale factors are derived with a two-dimensional dependence on  $p_{T,V}$  and  $M_{jj}$ , and are  
 428 shown in Fig. 17 for W and DY, and in Fig. 18 for photon production.

429 **5.4.3 QCD NNLO corrections to QCD V processes**

430 NNLO corrections are obtained from the fixed-order calculations results in [25]. They are  
 431 parametrized as a function of the generator-level boson  $p_T$ . The correction is shown in Fig. 19.

432 **5.4.4 EW NLO corrections to QCD V processes**

433 Scale factors corresponding to NLO EW corrections are obtained from Ref. [25] and applied as  
 434 a function of the generator-level boson  $p_T$ . The scale factors are shown in Fig. 20.

435 **5.4.5 QCD NLO corrections to EWK V processes**

436 The QCD NLO corrections to EWK W and Z production have been calculated in Ref. [26] using  
 437 the VBF@NLO program. They are parametrized in  $p_{T,V}$  and  $M_{jj}$  and are shown in Fig. 21.

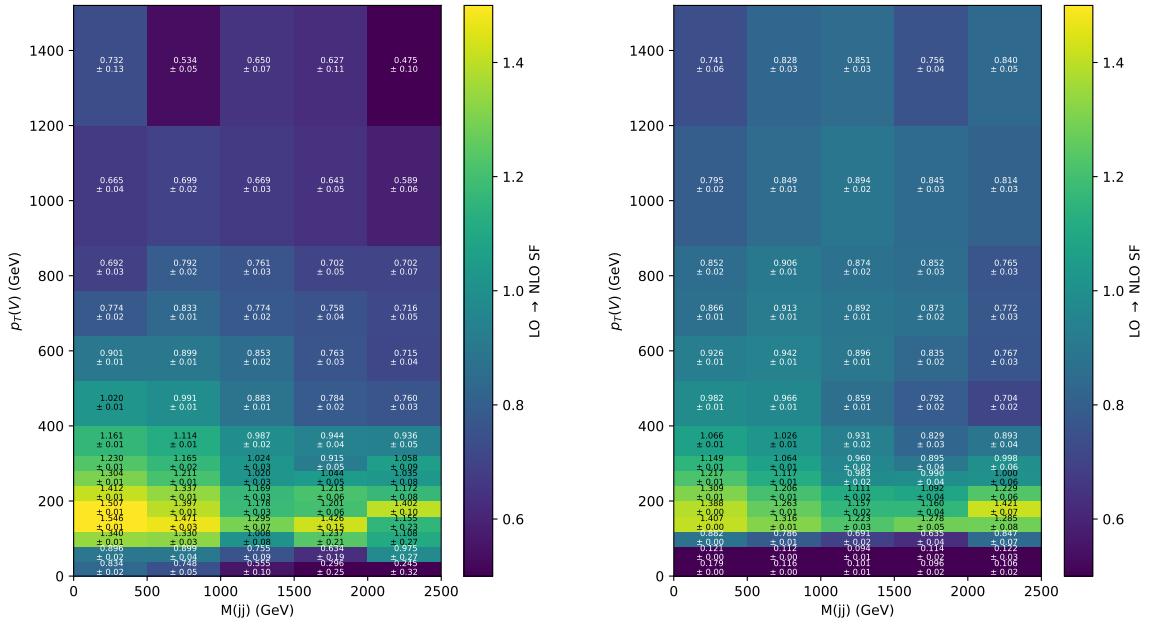


Figure 17: Same as Fig. ??, but now binned in two dimensions of the generator-level boson  $p_T$  and  $M_{jj}$ . The k factors are derived within the generator-level VBF selection described in the text. The uncertainties quoted in each bin are the statistical uncertainties due to the finite size of simulated samples.

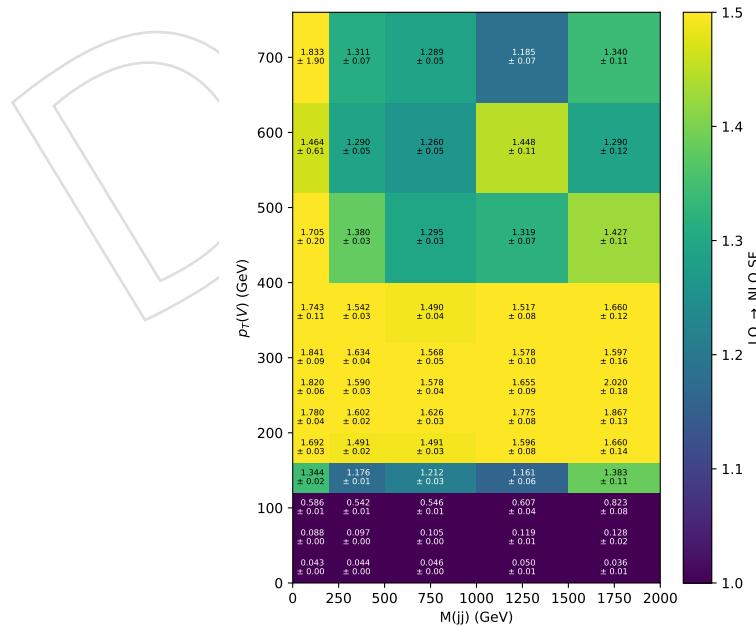


Figure 18: Same as Fig. 17, but for photon production.

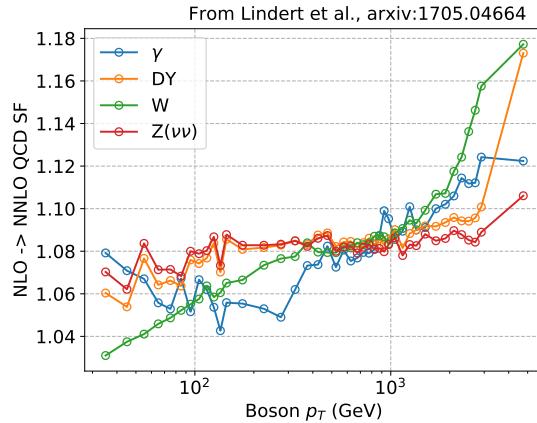


Figure 19: QCD NNLO scale factors for DY, W and photon production as a function of  $p_{T,V}$ .

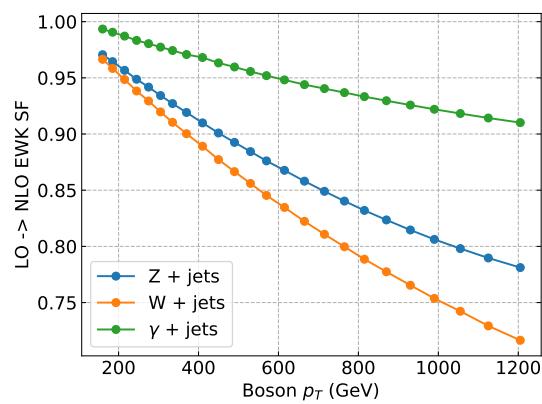


Figure 20: EW NLO scale factors for DY, W and photon production as a function of  $p_{T,V}$ .

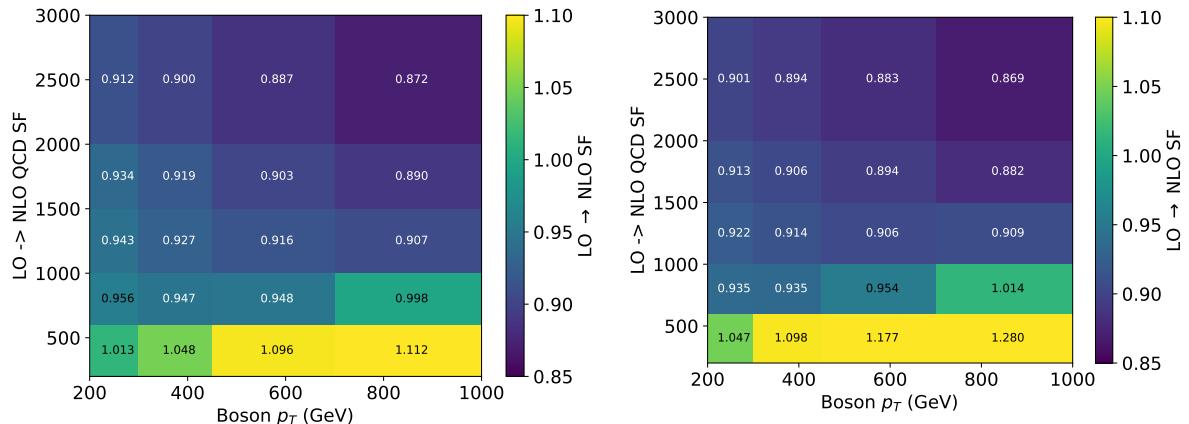


Figure 21: QCD NLO scale factors for EWK DY, W production of  $p_{T,V}$  and  $M_{jj}$ .

## 438 6 Event selection

### 439 6.1 Signal region selection

440 Signal region events are selected using triggers with thresholds of 120 GeV on both  $p_{\text{T},\text{trig}}^{\text{miss}}$  and  
 441  $H_{\text{T},\text{trig}}^{\text{miss}}$ . The  $p_{\text{T},\text{trig}}^{\text{miss}}$  corresponds to the magnitude of the vector  $\vec{p}_{\text{T}}$  sum of all the PF candidates  
 442 reconstructed at the trigger level, while the  $H_{\text{T},\text{trig}}^{\text{miss}}$  is computed as the magnitude of the vector  
 443  $\vec{p}_{\text{T}}$  sum of jets with  $p_{\text{T}} > 20$  GeV and  $|\eta| < 5.0$  reconstructed at the trigger level. The energy  
 444 fraction attributed to neutral hadrons in these jets is required to be smaller than 0.9. This re-  
 445 quirement suppresses anomalous events with jets originating from detector noise. To be able  
 446 to use the same triggers for selecting events in the muon control samples used for background  
 447 prediction, muon candidates are not included in the  $p_{\text{T},\text{trig}}^{\text{miss}}$  nor  $H_{\text{T},\text{trig}}^{\text{miss}}$  computation. The trigger  
 448 efficiency is measured to be 96% for events passing the analysis selection for  $p_{\text{T}}^{\text{miss}} > 250$  GeV  
 449 and becomes more than 99% efficient for events with  $p_{\text{T}}^{\text{miss}} > 350$  GeV.

450 Candidate events are required to have  $p_{\text{T}}^{\text{miss}} > 250$  GeV. The leading AK4 jet in the signal event  
 451 is required to have  $p_{\text{T}} > 80$  GeV and  $|\eta| < 5.0$ , and the subleading AK4 jet is required to have  
 452  $p_{\text{T}} > 40$  GeV and  $|\eta| < 5.0$ . In addition, the analysis employs various event filters to reduce  
 453 events with large misreconstructed  $p_{\text{T}}^{\text{miss}}$  [27] originating from noncollision backgrounds.

454 For the VBF signal events, two leading jets in opposite hemispheres are expected, with large  
 455 dijet mass. Furthermore, these jets are expected to have large rapidity separation and small  
 456 azimuthal separation. Therefore, this analysis employs several requirements on  $M_{jj}$ ,  $\Delta\eta_{jj}$  and  
 457  $\Delta\phi_{jj}$ , which can be found in Table 11.

458 The main background processes in this search are the  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  processes.  
 459 The  $Z(\nu\nu) + \text{jets}$  process is an irreducible background and constitutes the largest background  
 460 in the search. In contrast, the background from  $W(\ell\nu) + \text{jets}$  is suppressed by imposing a veto  
 461 on events containing one or more loose muons or electrons with  $p_{\text{T}} > 10$  GeV, or hadronically  
 462 decaying  $\tau$  leptons with  $p_{\text{T}} > 20$  GeV. Events that contain a loose, isolated photon with  $p_{\text{T}} >$   
 463 15 GeV and  $|\eta| < 2.5$  are also rejected. This helps to suppress electroweak (EW) backgrounds in  
 464 which a photon is radiated from the initial state. To reduce the contamination from top quark  
 465 backgrounds, events are rejected if they contain a b tagged jet with  $p_{\text{T}} > 20$  GeV and  $|\eta| <$   
 466 2.4. These jets are identified using the DeepCSV algorithm [11, 28], adopting the “medium”  
 467 working point, which corresponds to correctly identifying a jet originating from a bottom quark  
 468 with a probability of 80% and misidentifying a jet originating from a charm quark (light-flavor  
 469 jet) with a probability of 12 (2)%. Lastly, QCD multijet background with  $E_{\text{T}}^{\text{miss}}$  arising from  
 470 mismeasurements of the jet momenta is suppressed by requiring the minimum azimuthal angle  
 471 between the  $\vec{p}_{\text{T}}^{\text{miss}}$  direction and each of the first four leading jets with  $p_{\text{T}}$  greater than 30 GeV  
 472 and  $|\eta| < 4.7$  to be larger than 0.5 radians.

473 The selection requirements for this analysis are summarized in Table 11.

474 Figs. 22, 23 show the distribution of the  $E_{\text{T}}^{\text{miss}}$ , the number of jets,  $p_{\text{T}}$  and  $\eta$  distribution of the  
 475 leading AK4 jet for events in the VBF signal category respectively for 2017 and 2018 datasets.

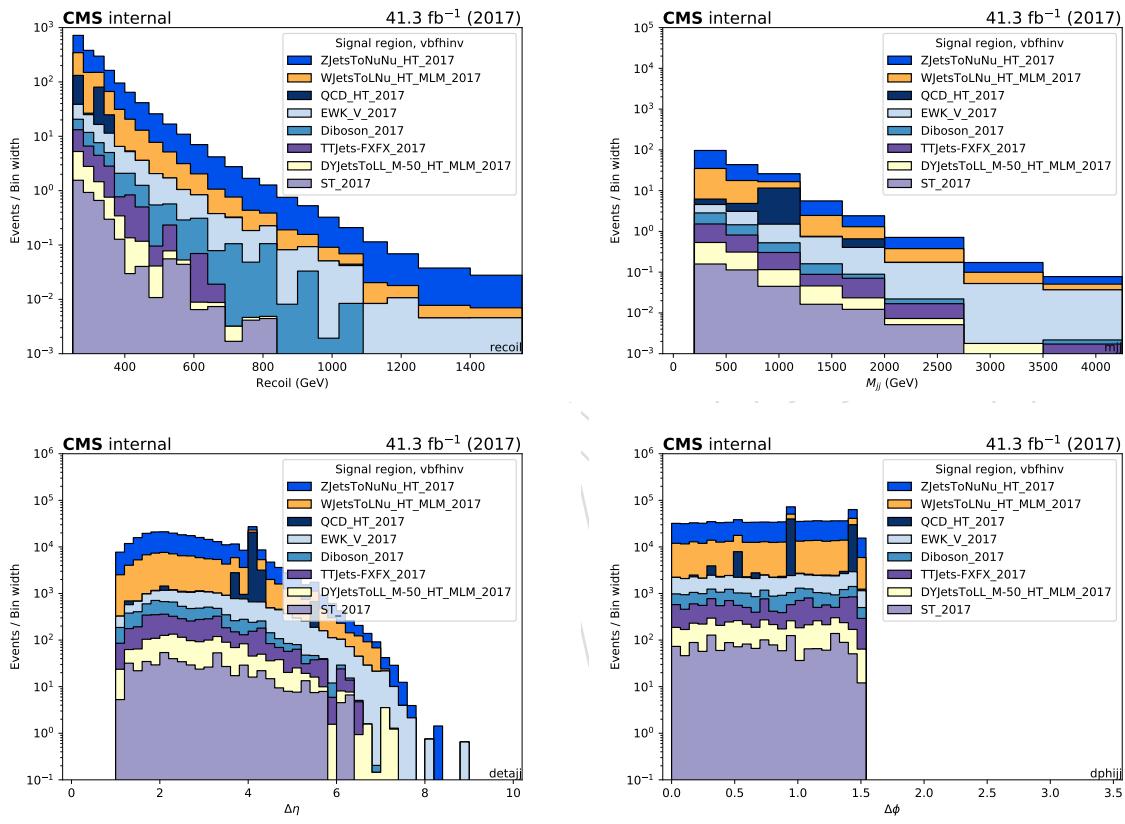


Figure 22: Recoil distribution,  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  and  $\Delta\phi_{jj}$  distribution in the VBF signal region, using 2017 dataset.

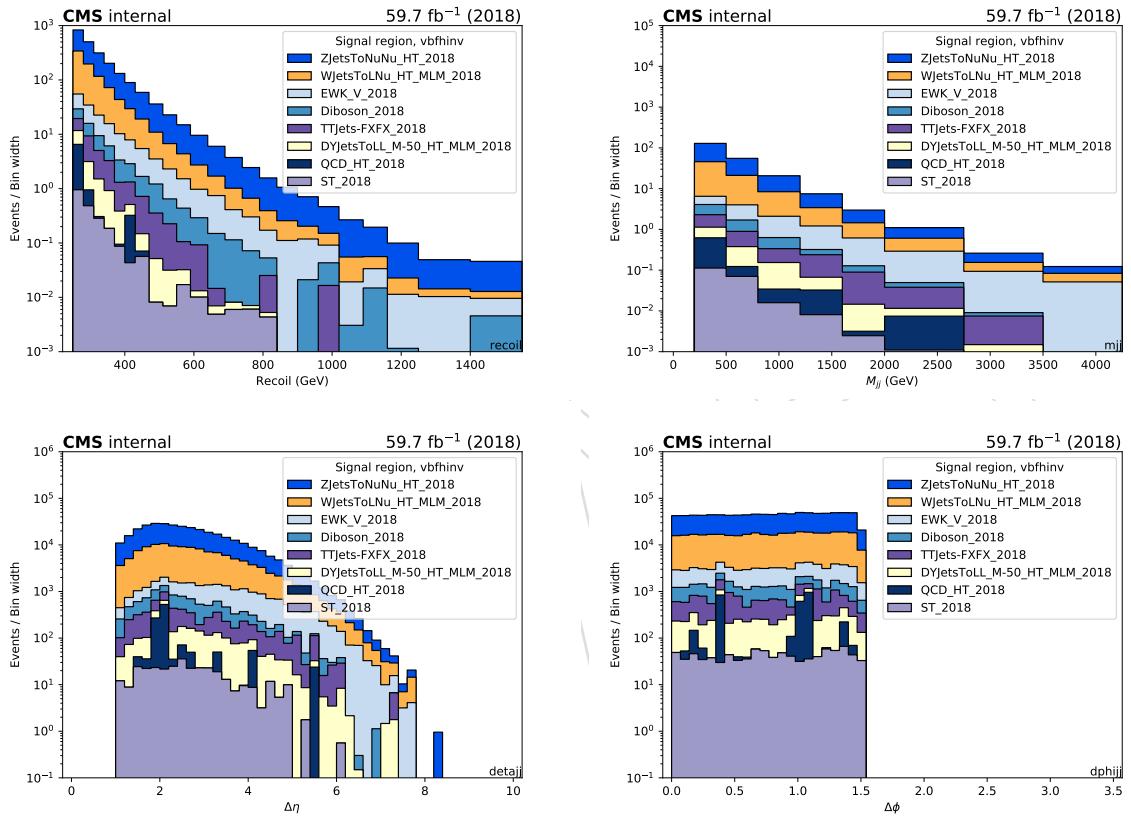


Figure 23: Recoil distribution,  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  and  $\Delta\phi_{jj}$  distribution in the VBF signal region, using 2018 dataset.

Table 11: Summary of the common selection requirements

Variable	Selection	Target background
Muon (electron) veto	$p_T > 10 \text{ GeV},  \eta  < 2.4(2.5)$	$Z(\ell\ell) + \text{jets}, W(\ell\nu) + \text{jets}$
$\tau$ lepton veto	$p_T > 20 \text{ GeV},  \eta  < 2.3$	$Z(\ell\ell) + \text{jets}, W(\ell\nu) + \text{jets}$
Photon veto	$p_T > 15 \text{ GeV},  \eta  < 2.5$	$\gamma + \text{jets}$
Bottom jet veto	DeepCSV medium $< 0.4941/0.4184$ (2017 / 2018) for all jets with $p_T > 20 \text{ GeV},  \eta  < 2.4$	Top quark
$p_T^{\text{miss}}$	$> 250 \text{ GeV}$	QCD, top quark, $Z(\ell\ell) + \text{jets}$
$\Delta\phi(\vec{p}_T^{\text{jet}}, \vec{p}_T^{\text{miss}})$	$> 0.5 \text{ radians}$	QCD
Leading AK4 jet $p_T$ and $\eta$	$> 80 \text{ GeV} \text{ and }  \eta  < 5.0$	All
Subleading AK4 jet $p_T$ and $\eta$	$> 40 \text{ GeV} \text{ and }  \eta  < 5.0$	All
$M_{jj}$	$> 200 \text{ GeV}$	
$\Delta\eta_{jj}$	$> 1.0$	
$\Delta\phi_{jj}$	$< 1.5$	

## 476 6.2 Single muon control region selection

477 Single-muon control sample events are selected using full signal region criteria of VBF selec-  
 478 tion with the exception of the muon veto. The  $p_T^{\text{miss}}$  requirement is replaced by an identical  
 479 requirement on the hadronic recoil, which is defined as the sum of  $\vec{p}_T^{\text{miss}}$  and the muon  $\vec{p}_T$ , and  
 480 thus corresponds to the distribution of the  $W p_T$ . In the single-muon control sample, exactly  
 481 one tightly identified, isolated muon with  $p_T > 20 \text{ GeV}$  is required. No additional loose muons  
 482 or electrons with  $p_T > 10 \text{ GeV}$  are allowed. In addition, the transverse mass of the muon- $\vec{p}_T^{\text{miss}}$   
 483 system is required to be smaller than than 160 GeV. The transverse mass ( $M_T$ ) is computed as  
 484  $M_T = \sqrt{2E_T^{\text{miss}} p_T^\mu (1 - \cos\Delta\phi)}$ , where  $p_T^\mu$  is the  $p_T$  of the muon, and  $\Delta\phi$  is the angle between  $\vec{p}_T^\mu$   
 485 and  $\vec{p}_T^{\text{miss}}$ .

486 Figs. 24 and 26 show the distributions of the recoil,  $M_{jj}$ ,  $\Delta\eta_{jj}$  and  $\Delta\phi_{jj}$  of the two leading AK4  
 487 jets for events in the single-muon control sample for the VBF category in 2017 and 2018 datasets,  
 488 respectively. Figs. 25 and 27 show the distributions of the leading muon  $p_T$  and  $\eta$ , as well as  
 489 the muon- $\vec{p}_T^{\text{miss}}$  transverse mass, again for 2017 and 2018, respectively.

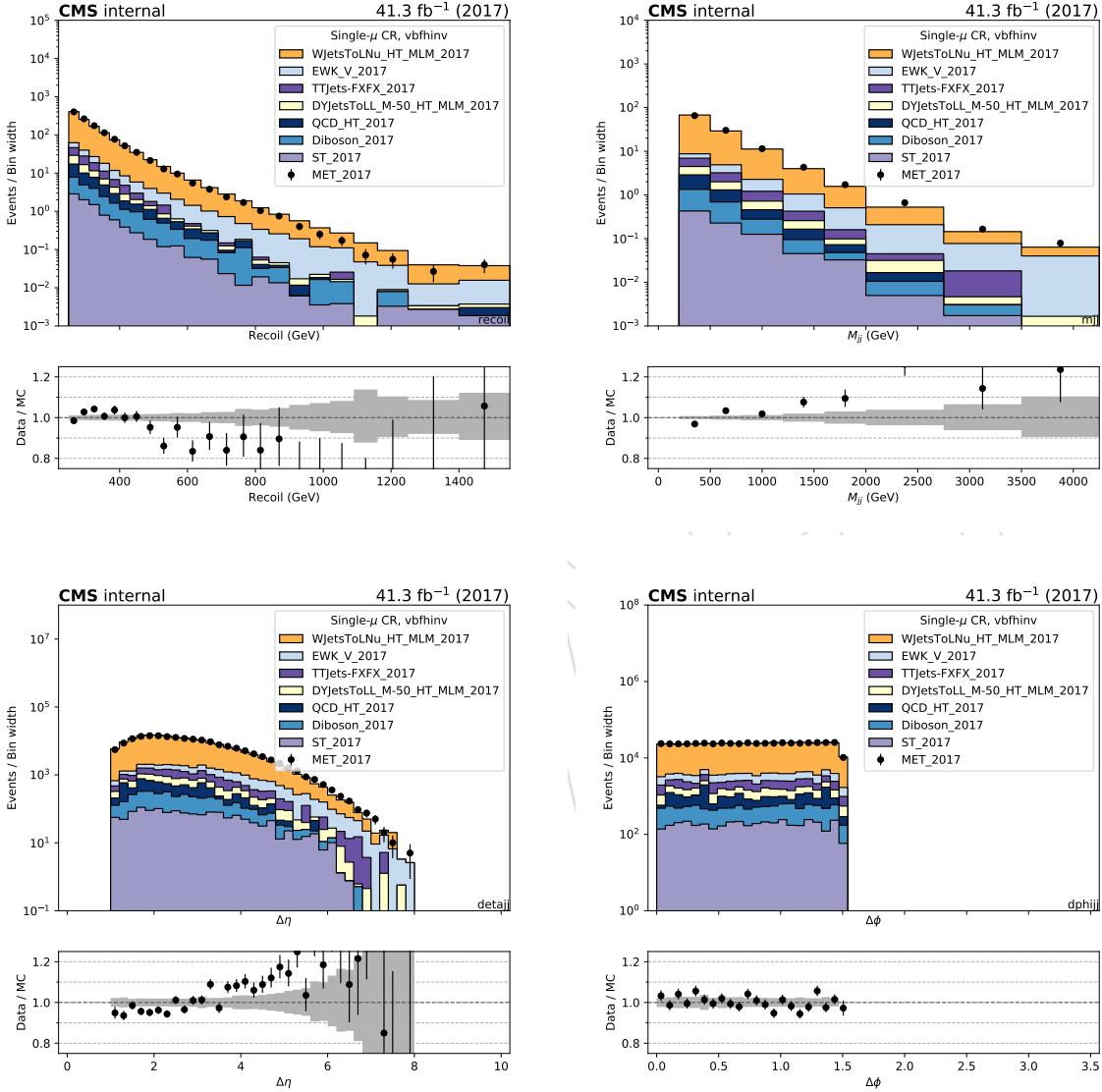


Figure 24: Comparison between 2017 data and Monte Carlo simulation in the single muon control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

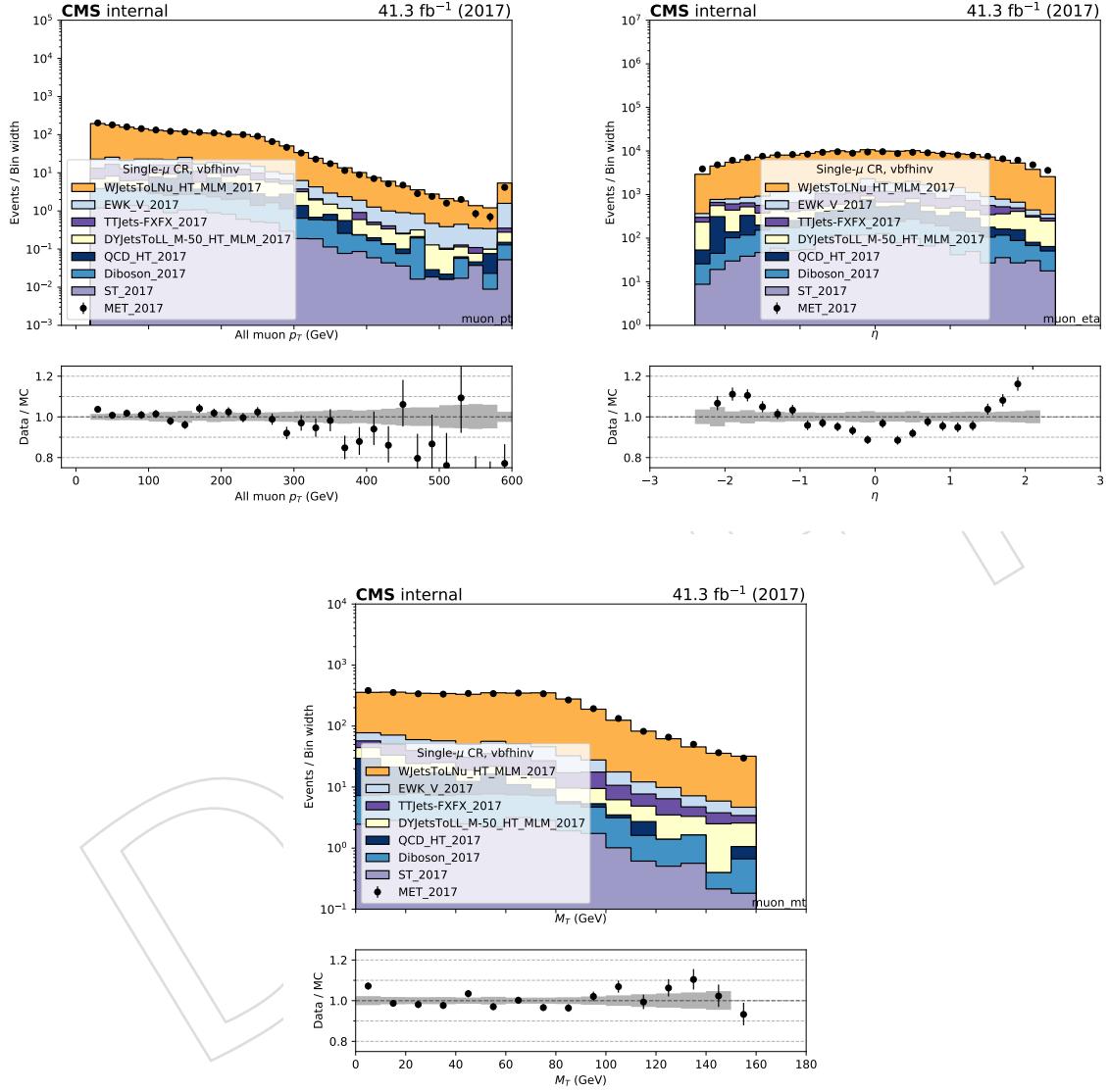


Figure 25: Comparison between 2017 data and Monte Carlo simulation in the single muon control sample for the  $p_T$  and  $\eta$  of the leading muon and the transverse mass distribution with the VBF selection.

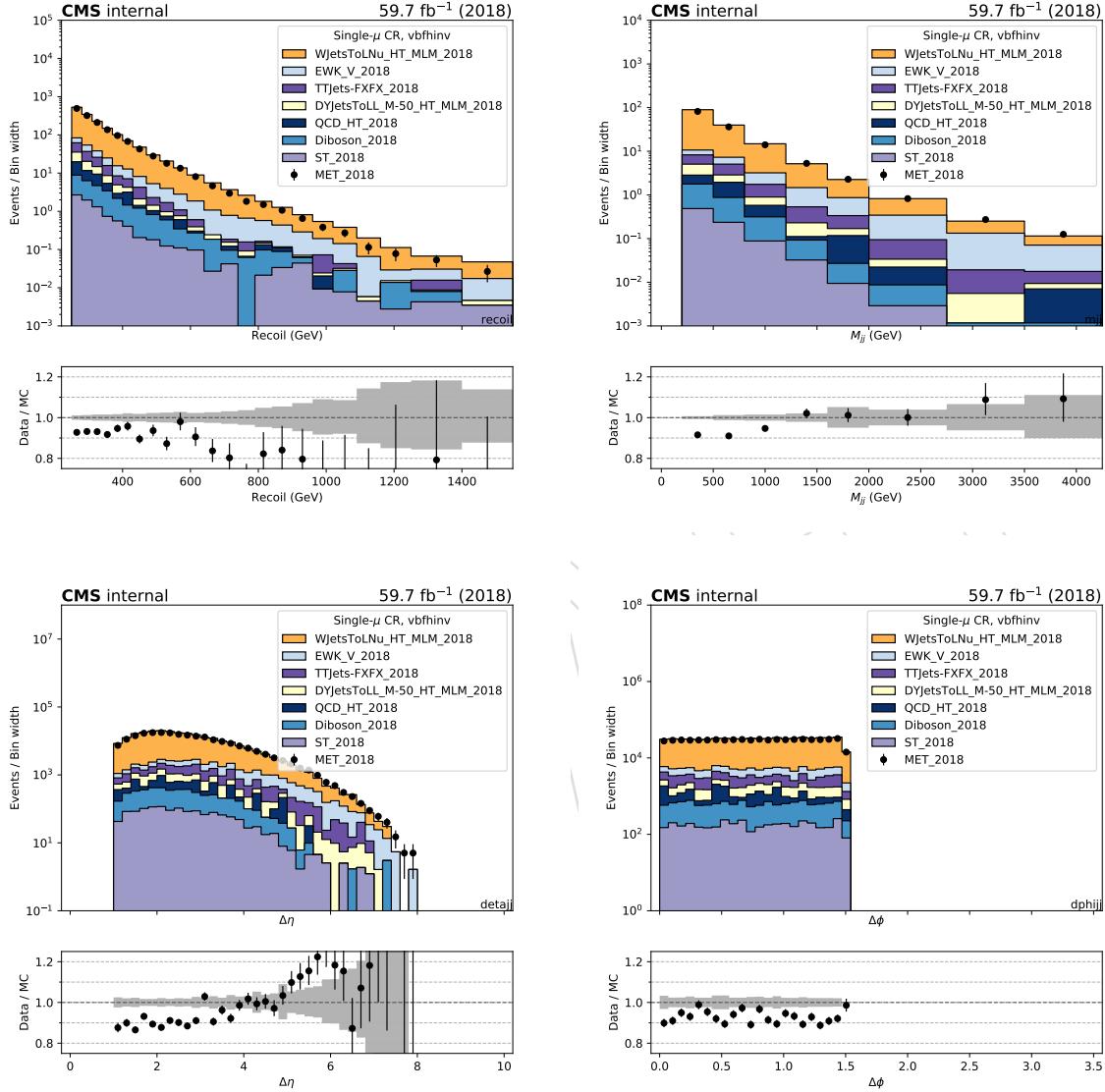


Figure 26: Comparison between 2018 data and Monte Carlo simulation in the single muon control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

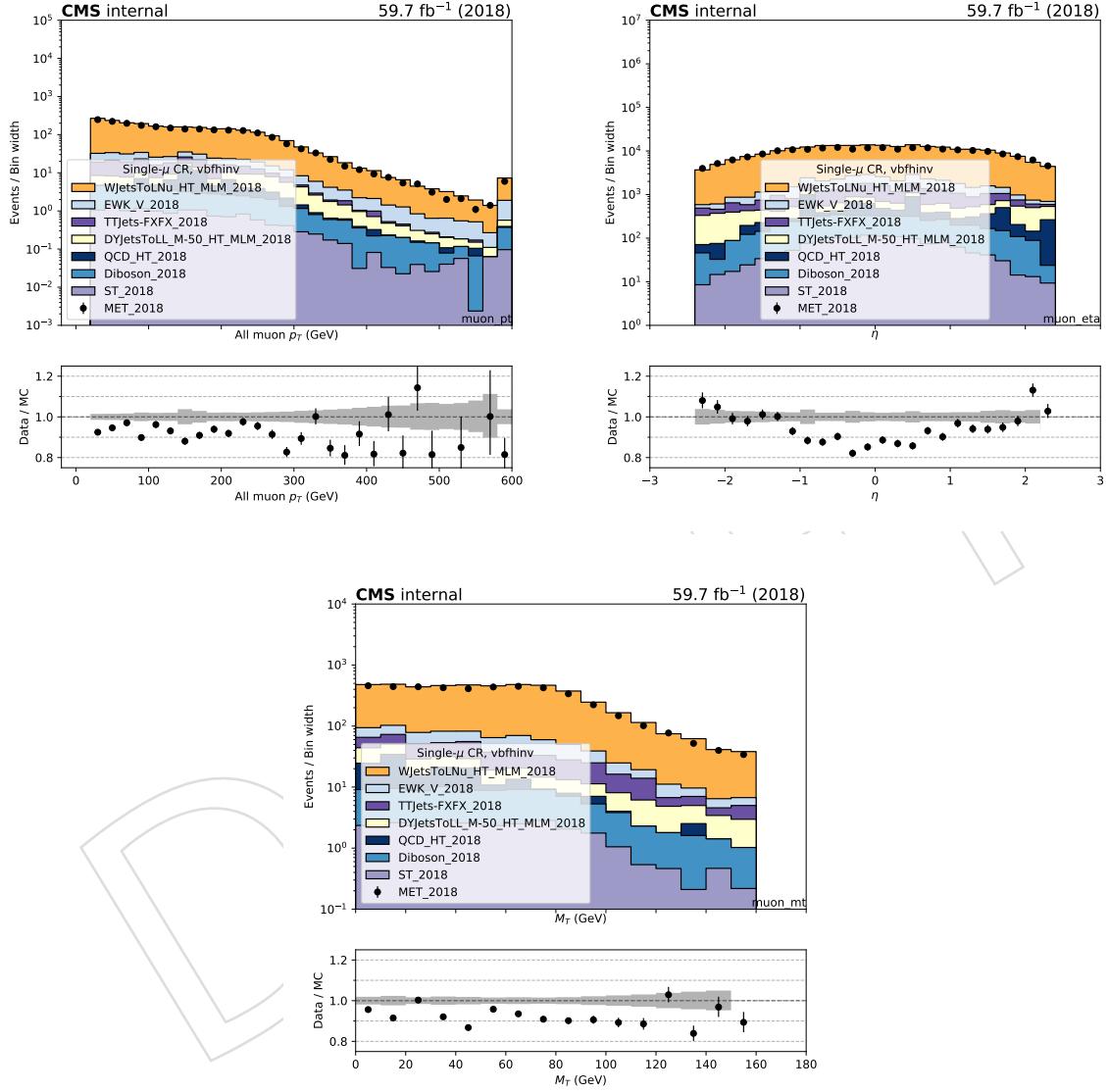


Figure 27: Comparison between 2018 data and Monte Carlo simulation in the single muon control sample for the  $p_T$  and  $\eta$  of the leading muon and the transverse mass distribution with the VBF selection.

**490 6.3 Single electron control region selection**

491 Events for the single-electron control sample are collected with the single-electron and photon  
492 triggers described in Sec. 2. The  $p_T^{\text{miss}}$  requirement is replaced with an identical requirement  
493 on the hadronic recoil, which is defined as the sum of  $\vec{p}_T^{\text{miss}}$  and the electron  $\vec{p}_T$ , and thus  
494 corresponds to the distribution of the  $W p_T$ . The events in the single-electron control sample  
495 are required to contain exactly one tightly identified and isolated electron with  $p_T > 40$  GeV.  
496 In addition, the contamination from QCD multijet events in this control sample is suppressed  
497 by requiring  $E_T^{\text{miss}} > 50$  GeV and  $M_T < 160$  GeV.

498 Figs. 28 and 30 show the distributions of the recoil,  $M_{jj}$ ,  $\Delta\eta_{jj}$  and  $\Delta\phi_{jj}$  of the two leading AK4  
499 jets for events in the single-electron control sample for the VBF category in 2017 and 2018  
500 datasets, respectively. Figs. 29 and 31 show the distributions of the leading electron  $p_T$  and  
501  $\eta$ , as well as the electron- $p_T^{\text{miss}}$  transverse mass, again for 2017 and 2018, respectively.

DRAFT

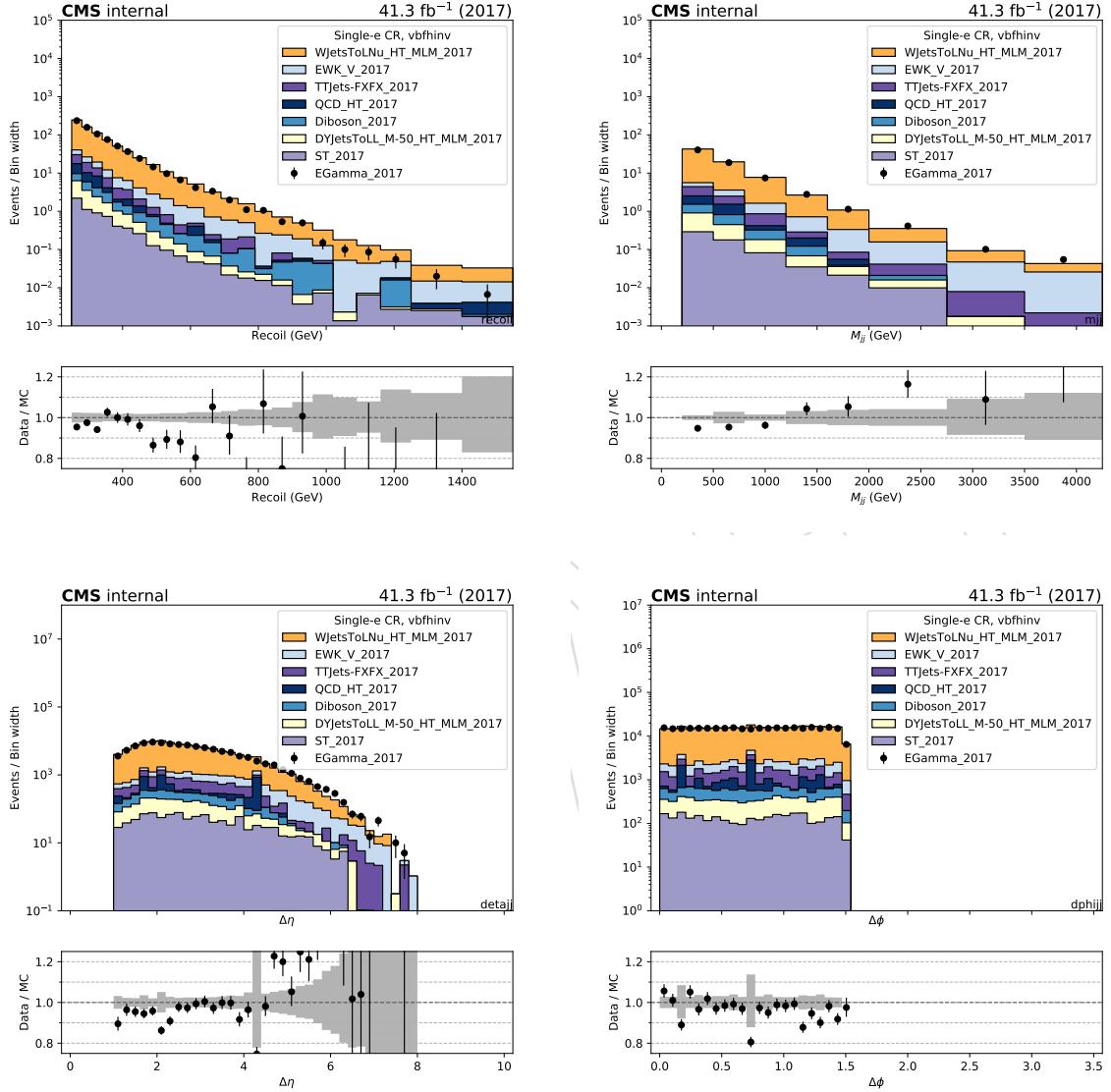


Figure 28: Comparison between 2017 data and Monte Carlo simulation in the single electron control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

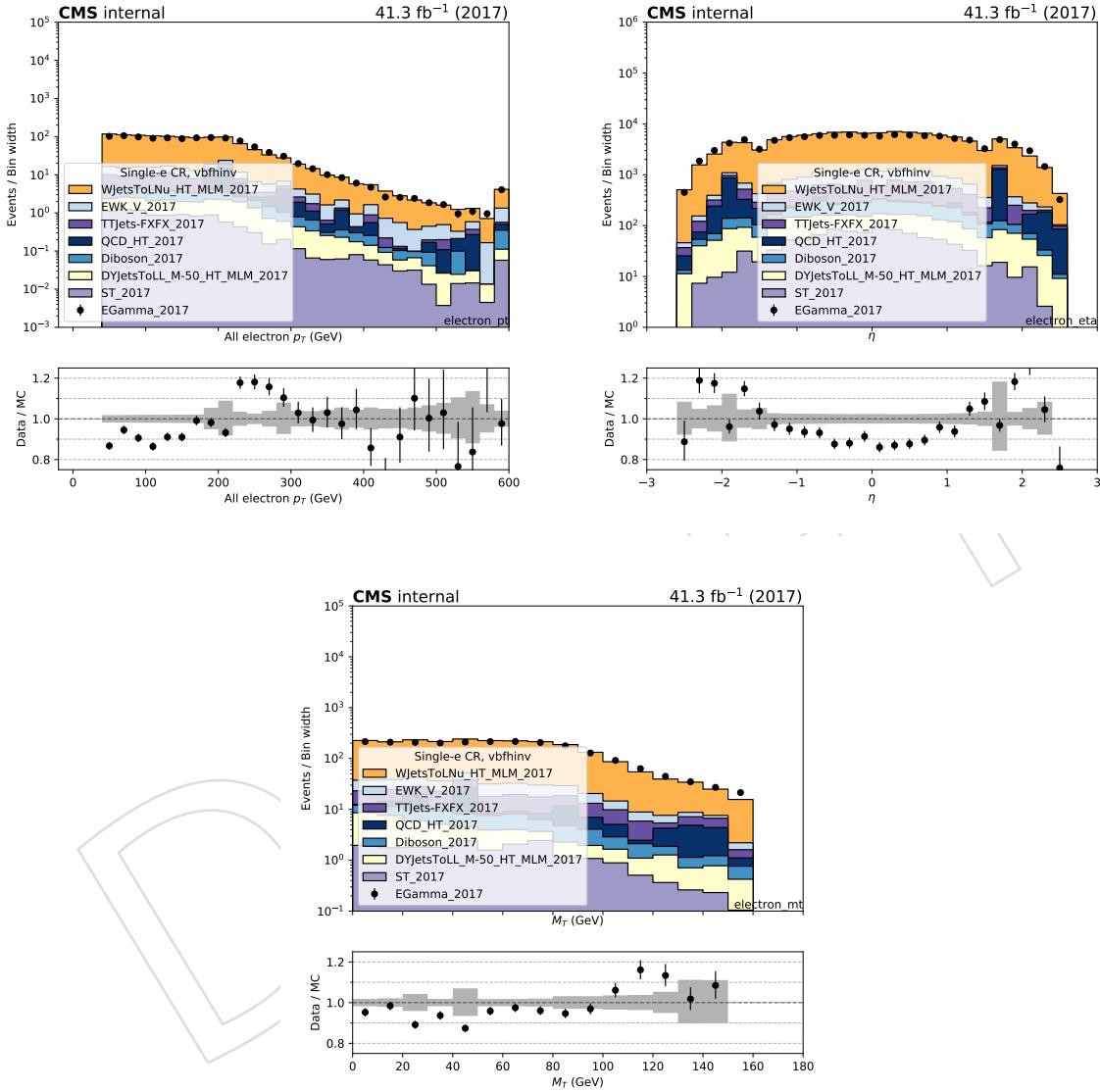


Figure 29: Comparison between 2017 data and Monte Carlo simulation in the single electron control sample for the  $p_T$  and  $\eta$  of the leading electron and the transverse mass distribution with the VBF selection.

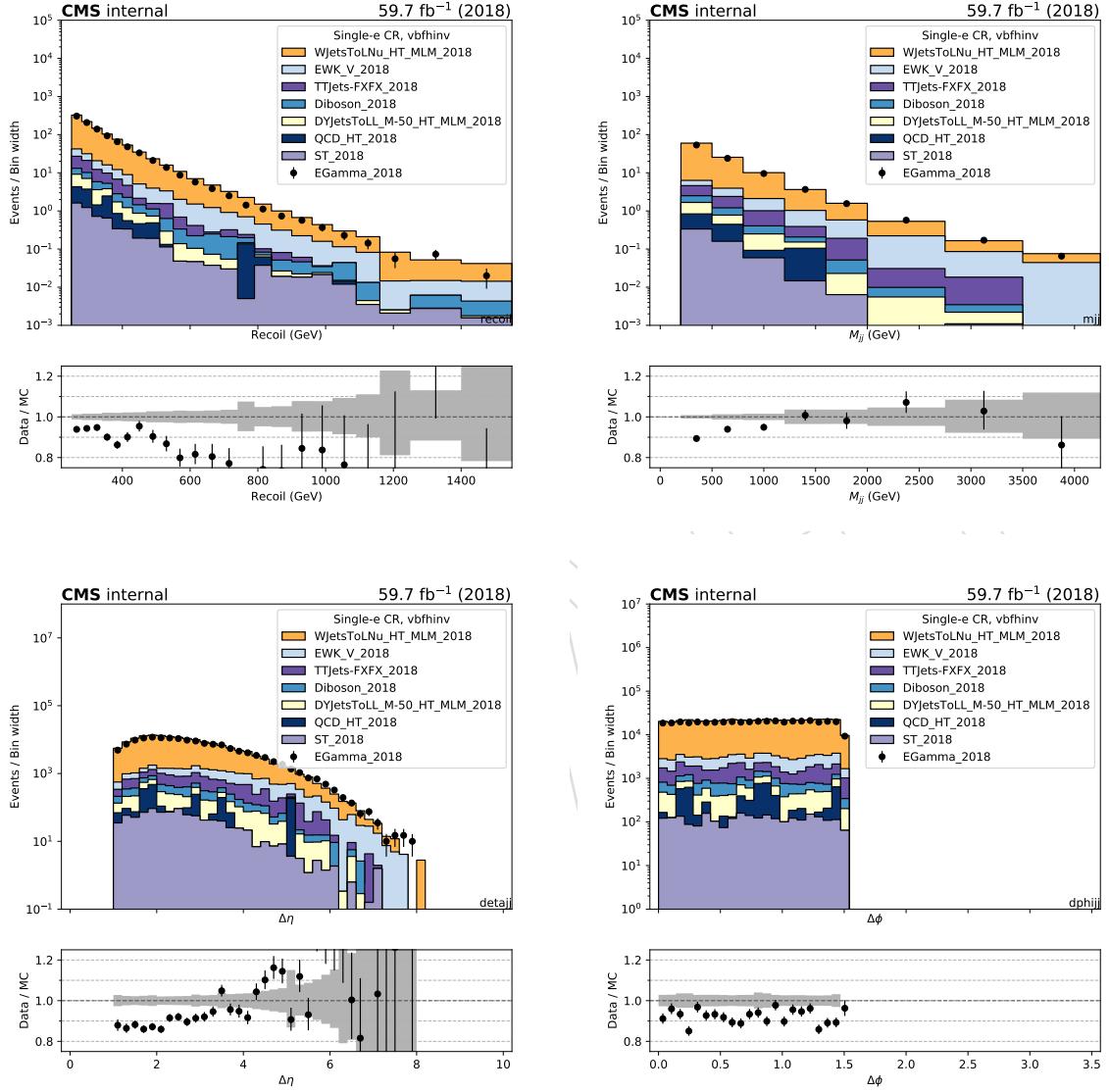


Figure 30: Comparison between 2018 data and Monte Carlo simulation in the single electron control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

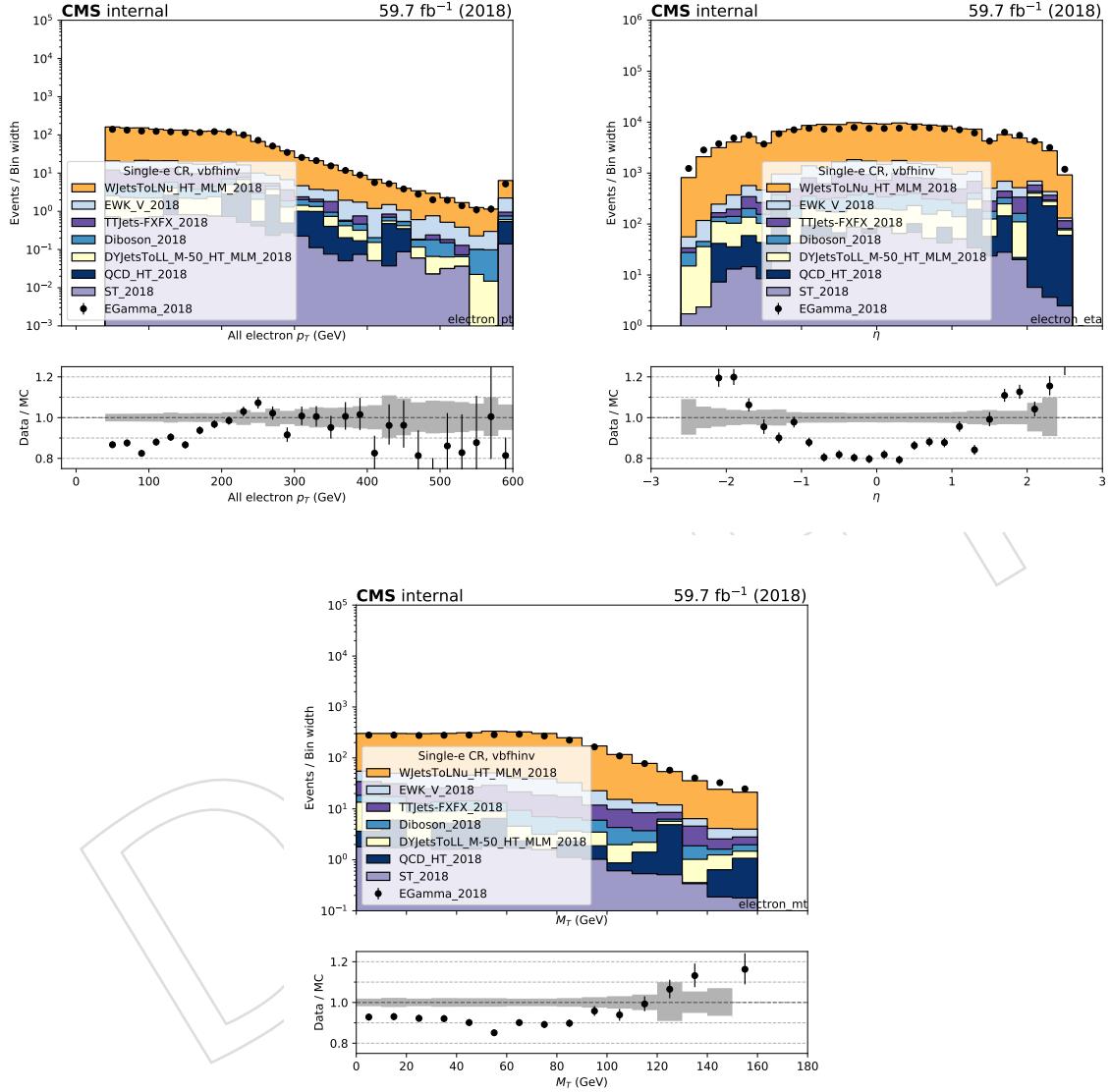


Figure 31: Comparison between 2018 data and Monte Carlo simulation in the single electron control sample for the  $p_T$  and  $\eta$  of the leading electron and the dilepton mass distribution with the VBF selection.

#### 502 6.4 Double muon control region selection

503 Double-muon control sample events are selected using full signal region criteria of VBF cate-  
504 gory with the exception of the muon veto. In the double-muon control sample, events are se-  
505 lected requiring leading (subleading) muon  $p_T$  greater than 20 (10) GeV and an invariant mass  
506 in the range 60 to 120 GeV, compatible with a Z boson decay. At least one of the two muons  
507 is required to pass the tight candidate definition. Events are rejected if there is an additional  
508 loose muon or electron with  $p_T > 10$  GeV. The SR  $p_T^{\text{miss}}$  requirement is replacement an identical  
509 requirement on the hadronic recoil, which is defined as the sum of  $\vec{p}_T^{\text{miss}}$  and the muon  $\vec{p}_T$ , and  
510 thus corresponds to the distribution of the Z  $p_T$  smeared with the  $p_T^{\text{miss}}$  resolution.

511 Figs. 32 and 34 shows the distributions of the recoil,  $M_{jj}$ ,  $\Delta\eta_{jj}$  and  $\Delta\phi_{jj}$  of the two leading  
512 AK4 jets for events in the double-muon control sample for the VBF category in 2017 and 2018  
513 datasets, respectively. Figs. 33 and 35 show the distributions of the leading muon  $p_T$  and  $\eta$ , as  
514 well as the dimuon mass and  $p_T$ , again for 2017 and 2018, respectively.

DRAFT

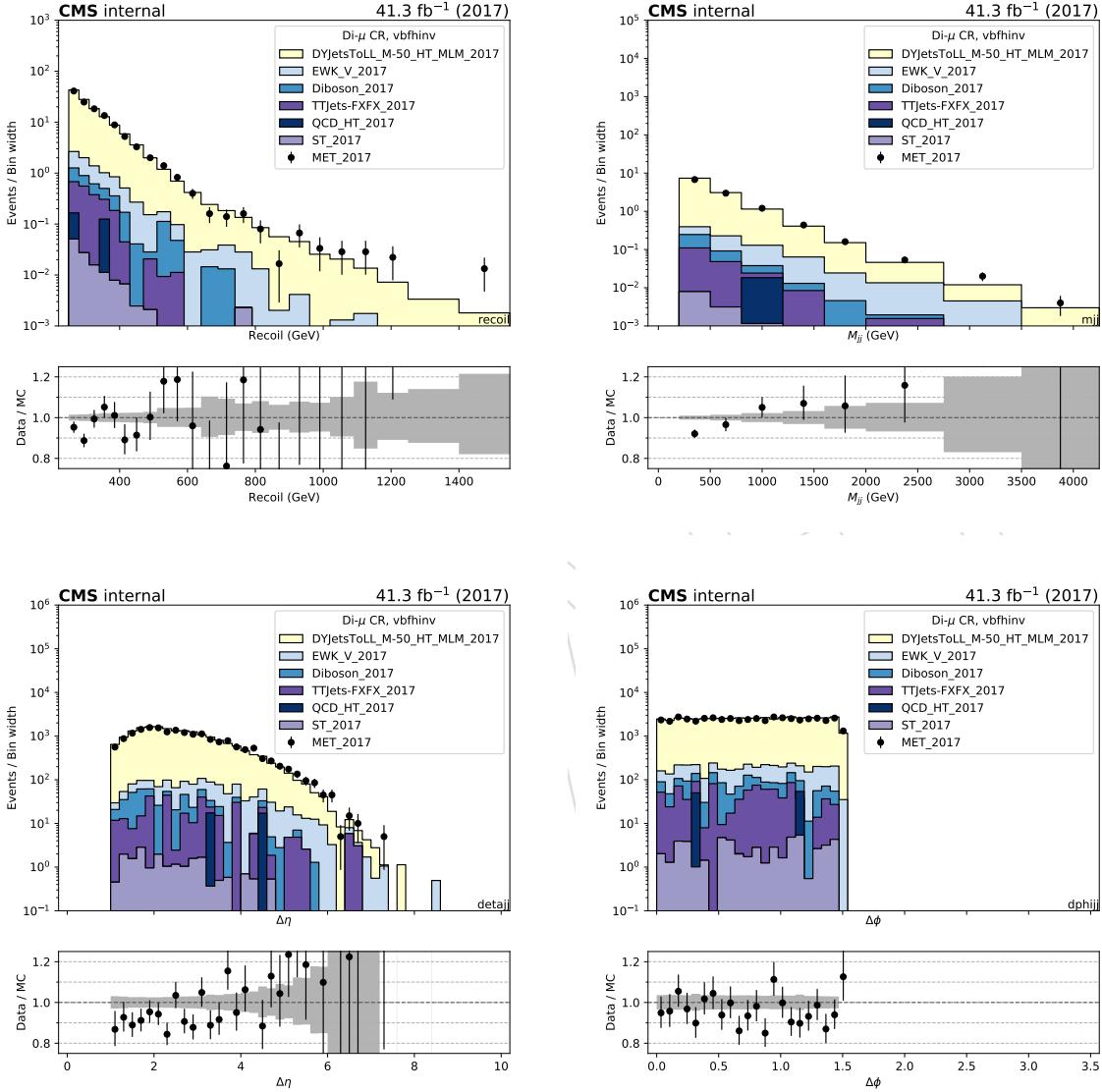


Figure 32: Comparison between 2017 data and Monte Carlo simulation in the double muon control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

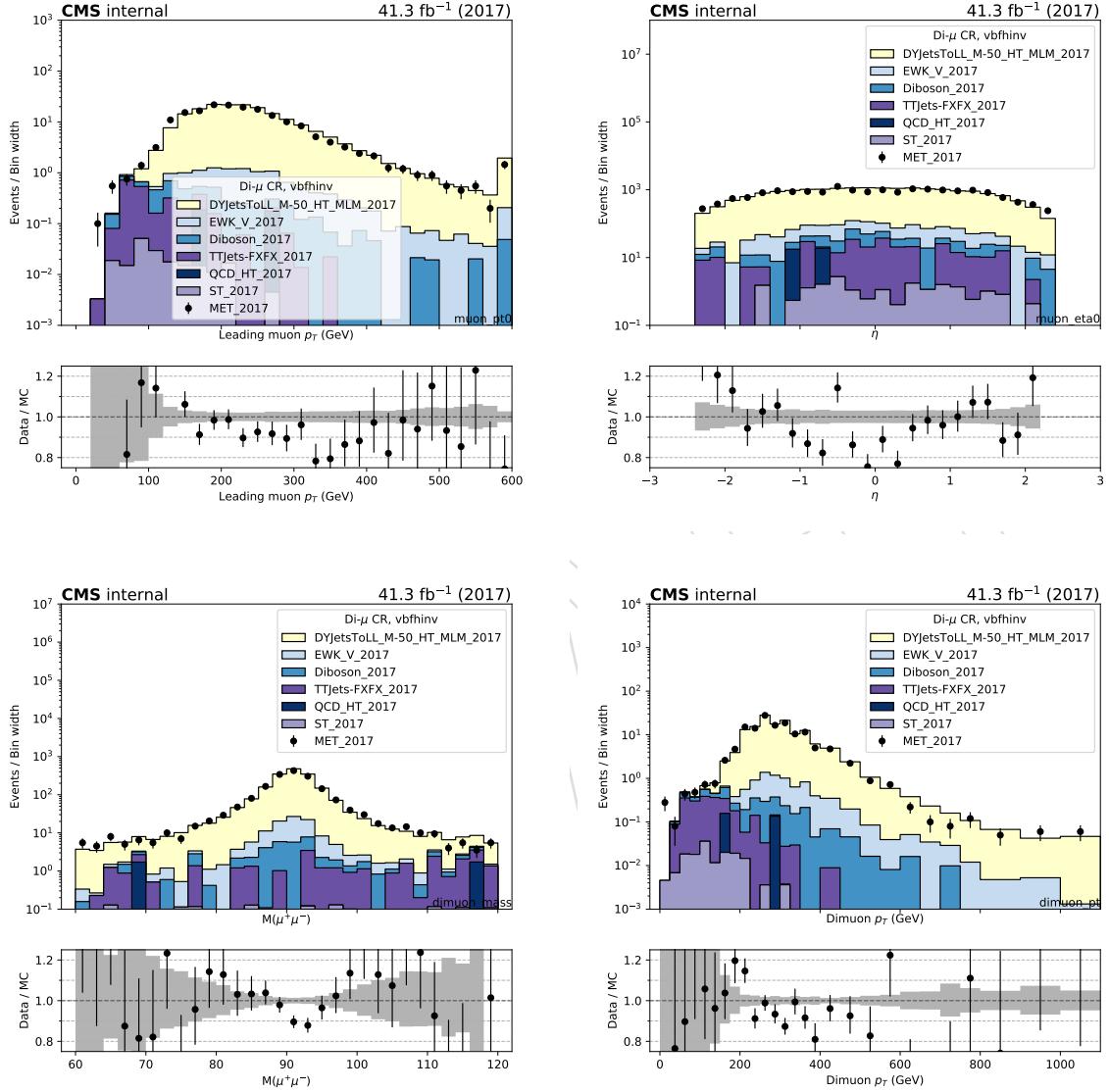


Figure 33: Comparison between 2017 data and Monte Carlo simulation in the double muon control sample for the  $p_T$  and  $\eta$  of the leading muon and the transverse mass and  $p_T$  of the dimuon candidate with the VBF selection.

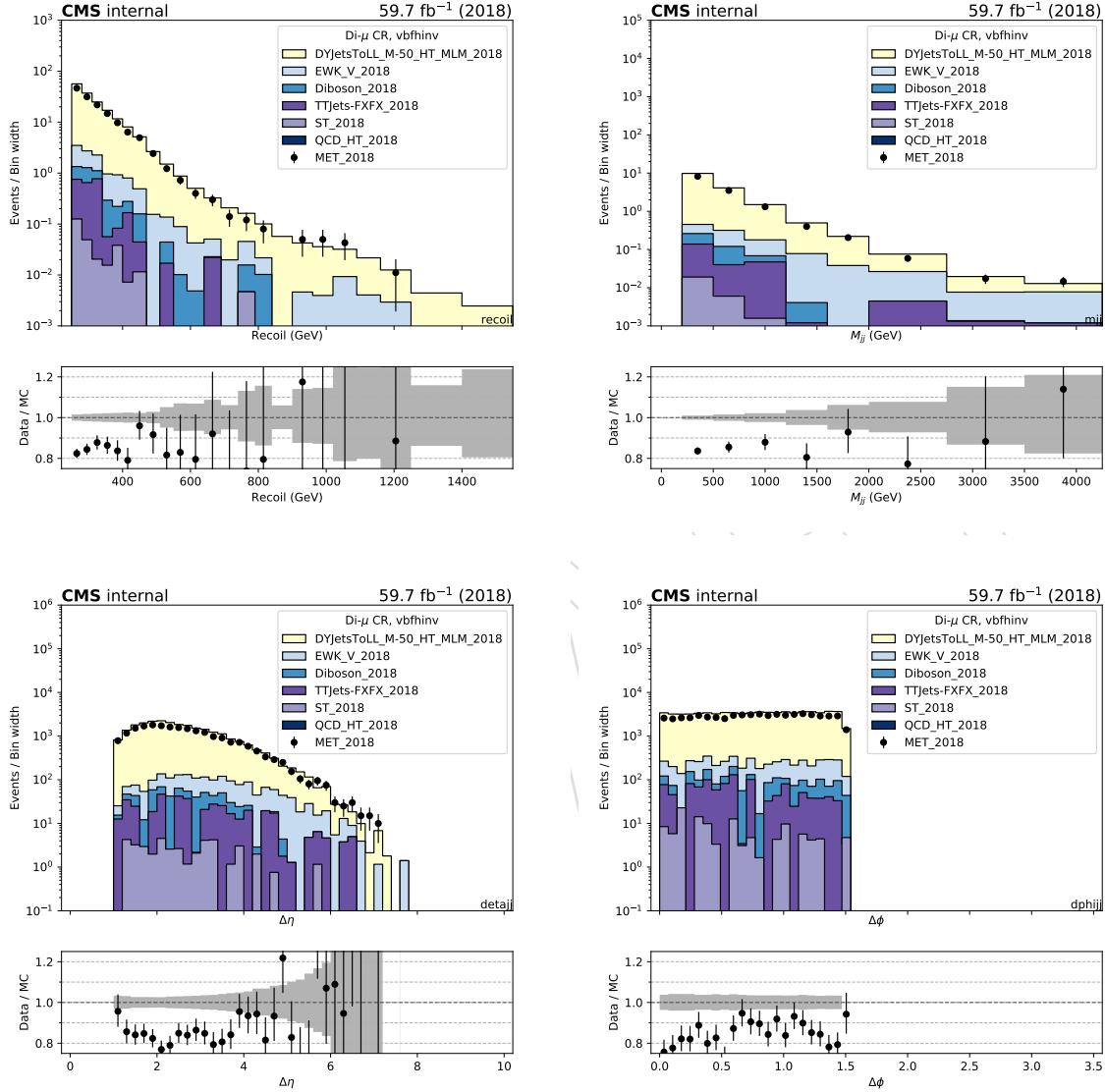


Figure 34: Comparison between 2018 data and Monte Carlo simulation in the double muon control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

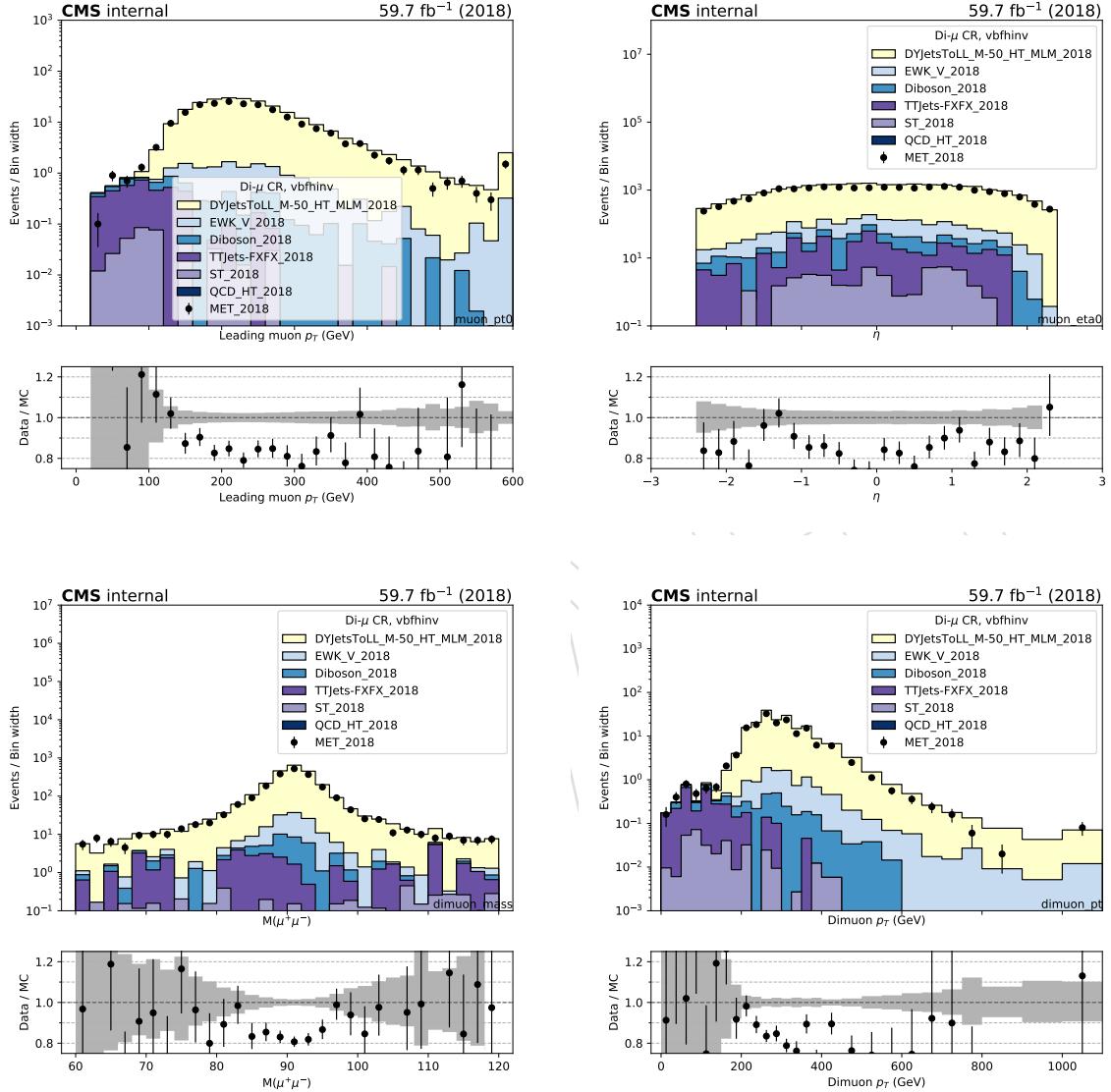


Figure 35: Comparison between 2018 data and monte carlo simulation in the double muon control sample for the  $p_T$  and  $\eta$  of the leading muon and the transverse mass and  $p_T$  of the dimuon candidate with the VBF selection.

515 **6.5 Double electron control region selection**

516 Events for the double-electron control sample are collected with the single-electron and photon  
 517 triggers described in Sec. 2. In the offline analysis, events in the dielectron control sample are  
 518 required to contain exactly two oppositely charged electrons with leading (trailing) electron  $p_T$   
 519 greater than 40 (10) GeV, with at least one of the two passing the tight candidate definition. The  
 520 SR  $p_T^{\text{miss}}$  requirement is replacement an identical requirement on the hadronic recoil, which is  
 521 defined as the sum of  $\vec{p}_T^{\text{miss}}$  and the muon  $\vec{p}_T$ , and thus corresponds to the distribution of the Z  
 522  $p_T$  smeared with the  $p_T^{\text{miss}}$  resolution. Similar to the dimuon control sample case, the invariant  
 523 mass of the dielectron system is required to be between 60 and 120 GeV to be consistent with a  
 524 Z boson decay.

525 Figs. 36 and 38 shows the distributions of the recoil,  $M_{jj}$ ,  $\Delta\eta_{jj}$  and  $\Delta\phi_{jj}$  for the two leading  
 526 AK4 jets for events in the double-electron control sample for the VBF category in 2017 and 2018  
 527 datasets, respectively. Figs. 37 and 39 show the distributions of the leading electron  $p_T$  and  $\eta$ ,  
 528 as well as the dielectron mass and  $p_T$ , again for 2017 and 2018, respectively.

DRAFT

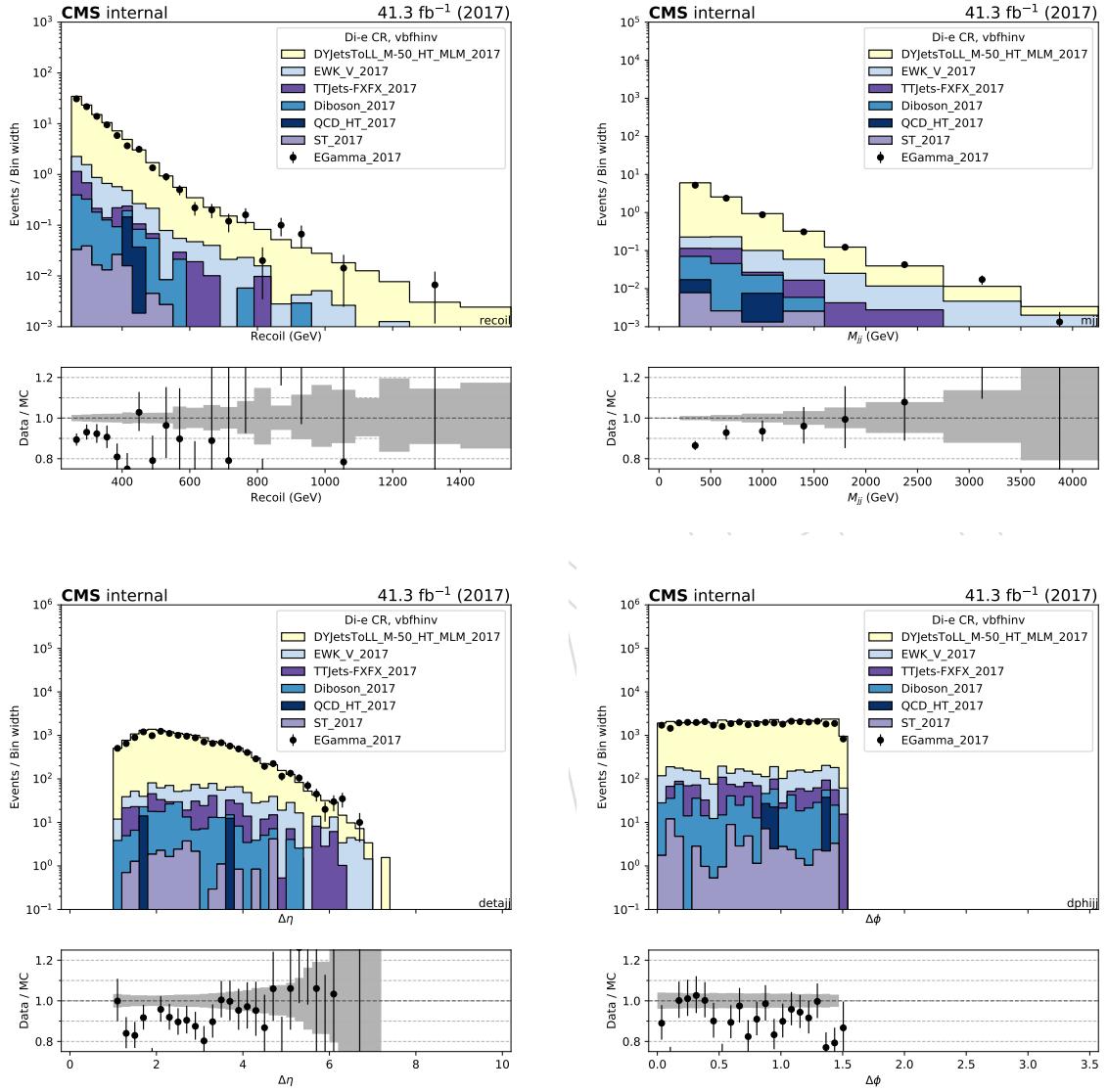


Figure 36: Comparison between 2017 data and Monte Carlo simulation in the double electron control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

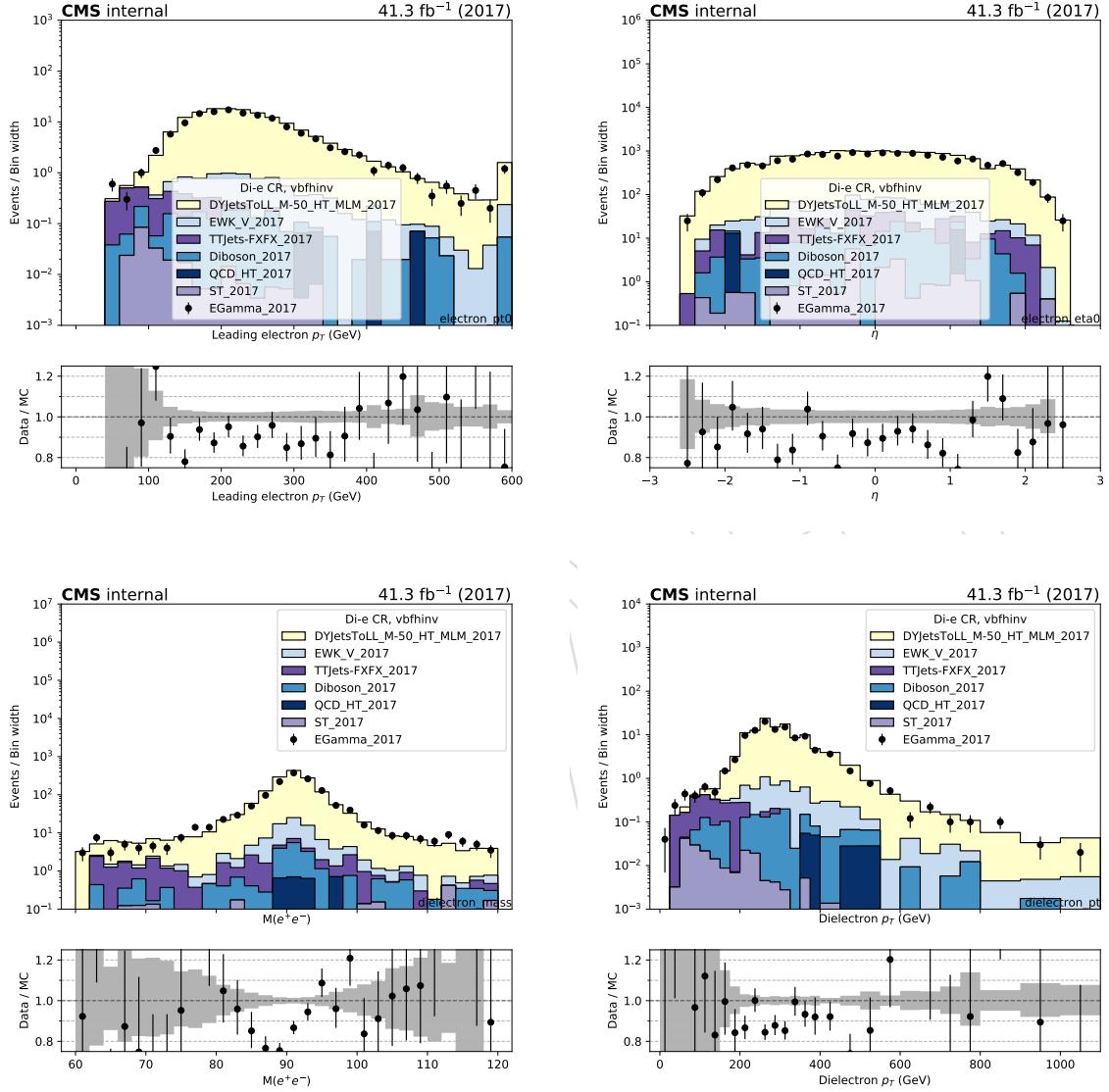


Figure 37: Comparison between 2017 data and Monte Carlo simulation in the double electron control sample for the  $p_T$  and  $\eta$  of the leading electron and the transverse mass and  $p_T$  of the dielectron candidate with the VBF selection.

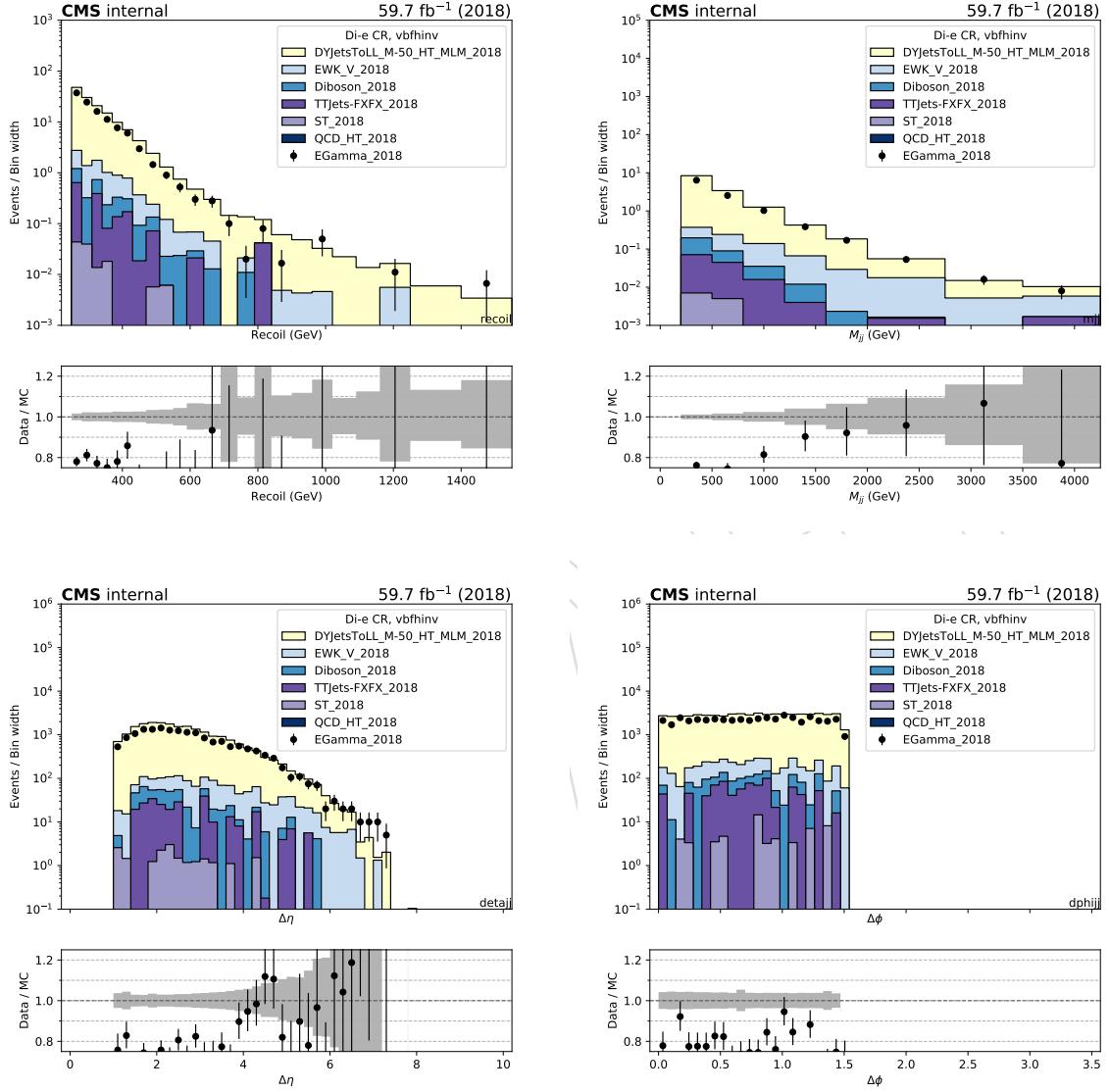


Figure 38: Comparison between 2018 data and Monte Carlo simulation in the double electron control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

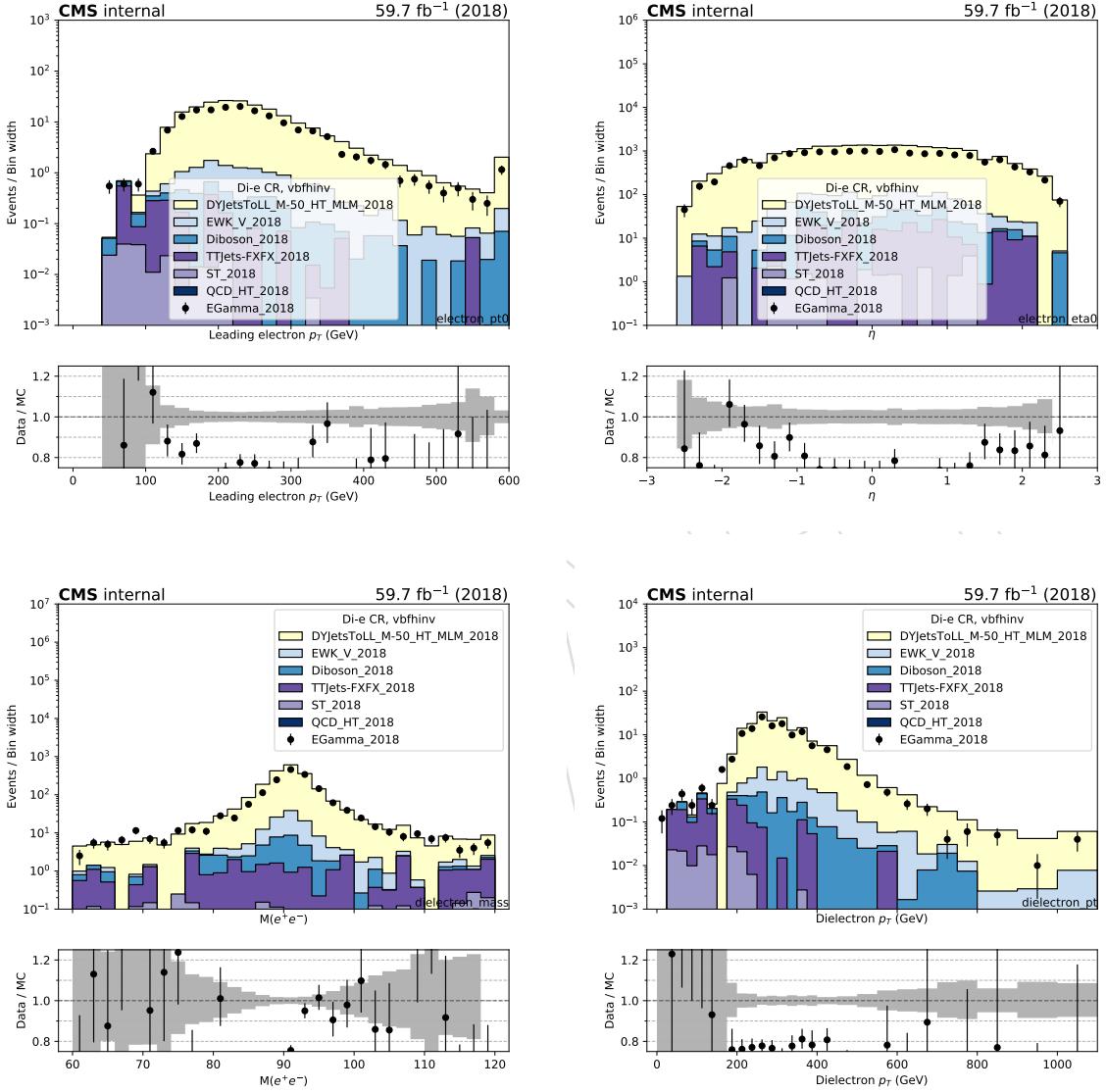


Figure 39: Comparison between 2018 data and Monte Carlo simulation in the double electron control sample for the  $p_T$  and  $\eta$  of the leading electron and the transverse mass and  $p_T$  of the dielectron candidate with the VBF selection.

## 529 6.6 Photon control region

530 The  $\gamma + \text{jets}$  control sample is selected using events with one high- $p_T$  photon collected using  
 531 single-photon triggers with  $p_T$  thresholds of 165 or 175 GeV, depending on the data taking  
 532 conditions. The photon is required to have  $p_T > 175$  GeV and to pass tight identification and  
 533 isolation criteria, to ensure a high trigger efficiency of 98%.

534 Figs. 40 and 40 show the distributions of the recoil,  $M_{jj}$ ,  $\Delta\eta_{jj}$  and  $\Delta\phi_{jj}$  distribution of the two  
 535 leading AK4 jets for events in the photon control sample for the VBF category in the 2017 and  
 536 2018 datasets, respectively. Similarly, Figs. 41 and 43 show the distributions of the photon  $p_T$   
 537 and  $\eta$ .

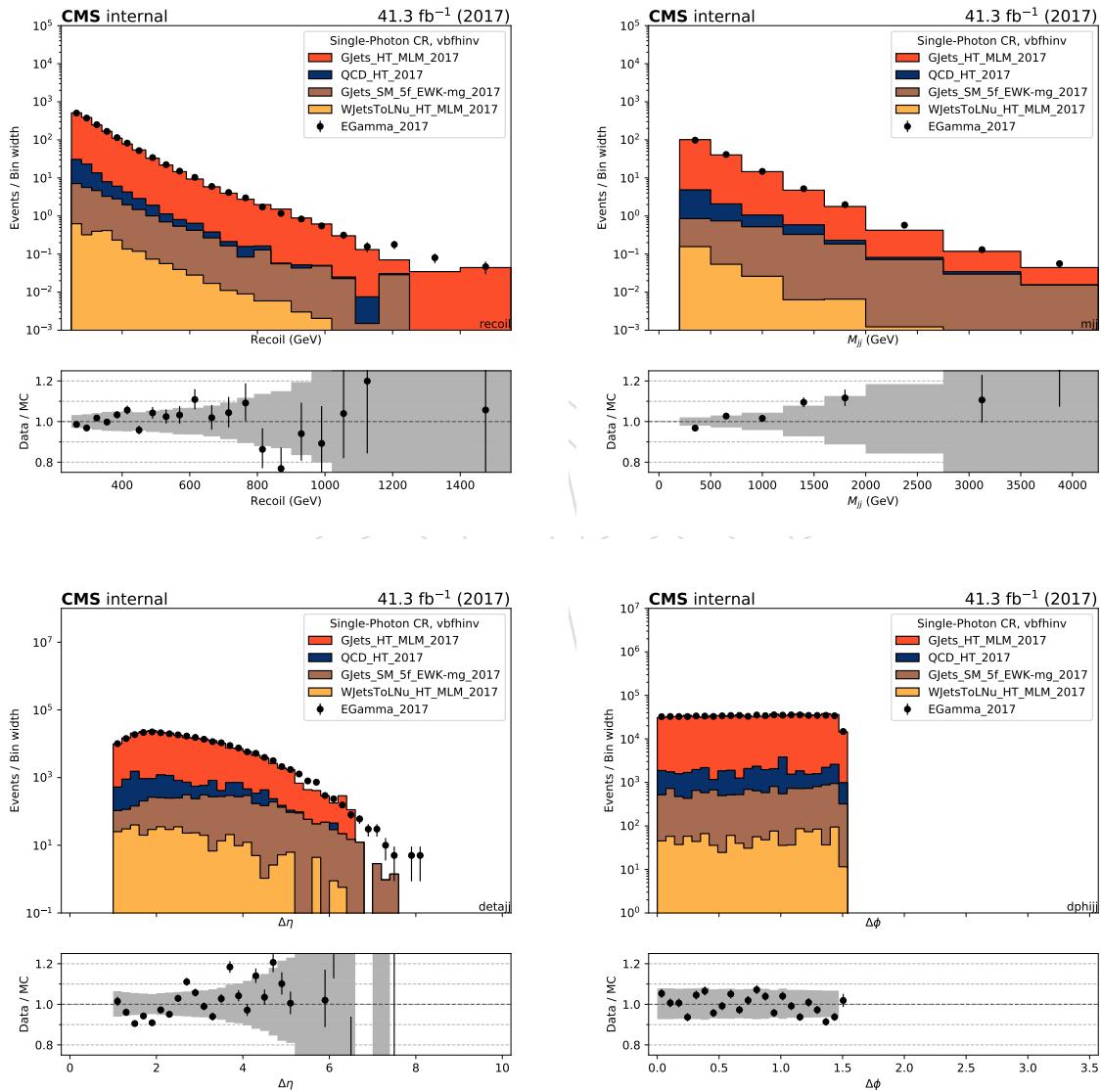


Figure 40: Comparison between 2017 data and Monte Carlo simulation in the photon control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

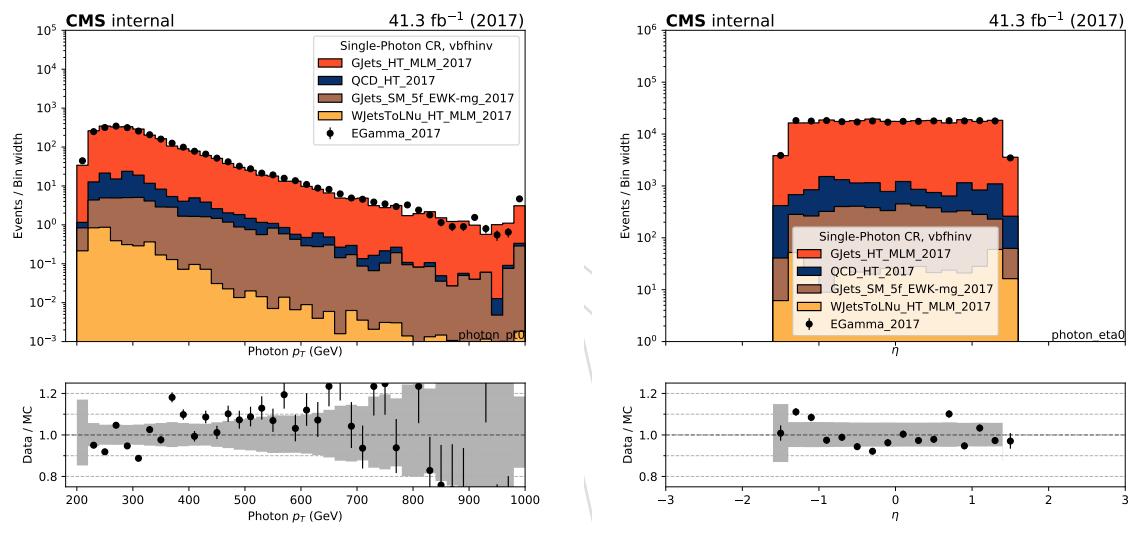


Figure 41: Comparison between 2017 data and Monte Carlo simulation in the photon control sample for the  $p_T$  and  $\eta$  of the leading photon with the VBF selection.

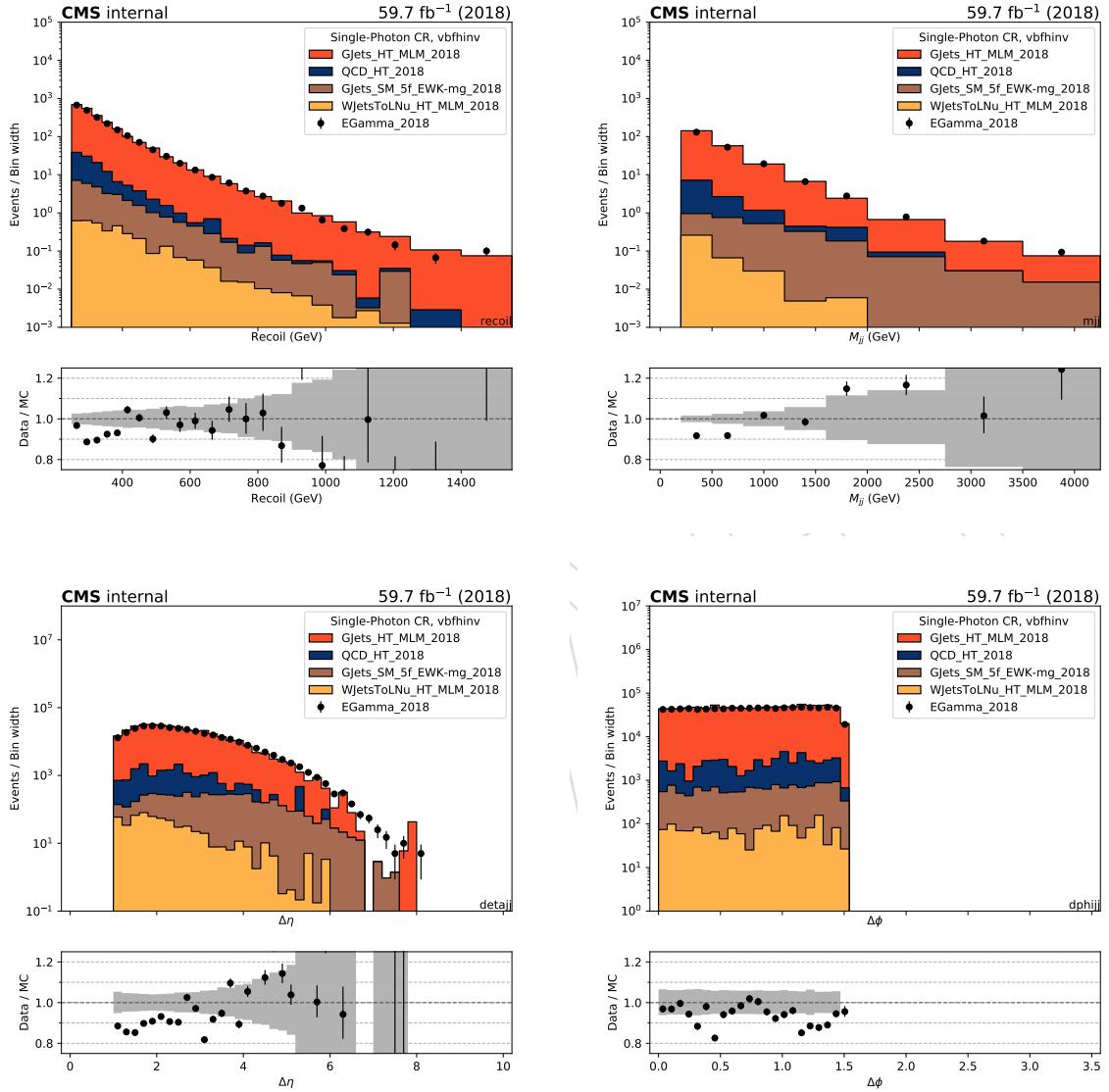


Figure 42: Comparison between 2018 data and Monte Carlo simulation in the photon control sample for the recoil distribution, the  $M_{jj}$  distribution,  $\Delta\eta_{jj}$  distribution and  $\Delta\phi_{jj}$  distribution for the two leading AK4 jets with the VBF selection.

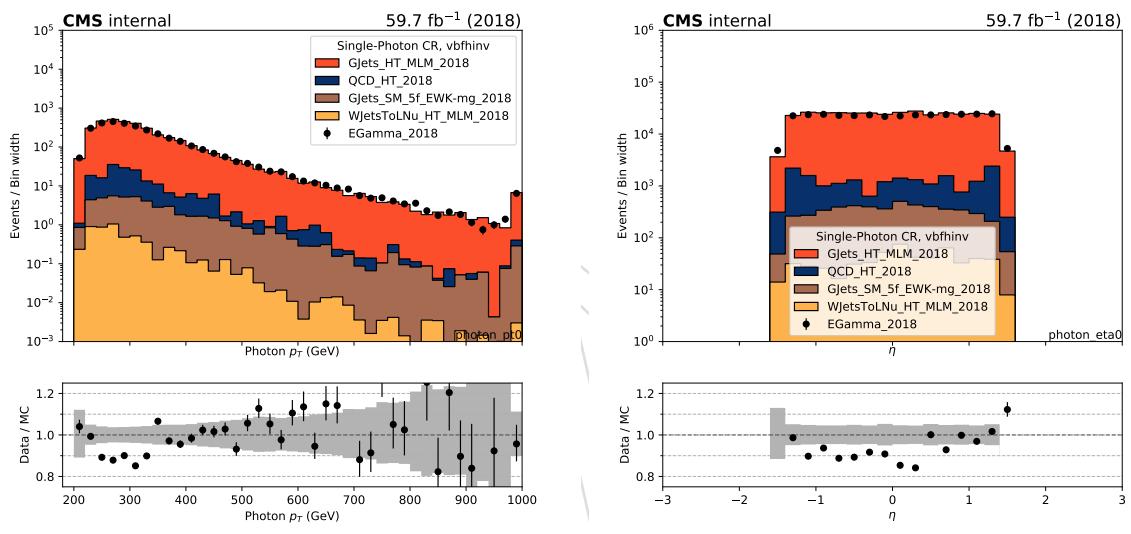


Figure 43: Comparison between 2018 data and Monte Carlo simulation in the photon control sample for the  $p_T$  and  $\eta$  of the leading photon with the VBF selection.

## 538 7 Background estimation

539 The largest background contributions from  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  processes are esti-  
 540 mated using data from four mutually exclusive control samples selected from dimuon, dielec-  
 541 tron, single-muon, and single-electron states, as explained below.

542 The remaining backgrounds that contribute to the total event yield in the signal region are  
 543 much smaller than those from  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  processes. These backgrounds in-  
 544 clude QCD multijet events which are measured from data using a  $\Delta\phi$  extrapolation method [29,  
 545 30], and top-quark and diboson processes, which are obtained directly from simulation and are  
 546 explained in the previous sections.

### 547 7.1 Signal extraction strategy

548 A binned likelihood fit to the data (constructed as a product of poisson probabilities) as pre-  
 549 sented here:

$$\mathcal{L}(\mu^{Z \rightarrow \nu\nu}, \mu, \theta) = \prod_i \text{Pois} \left( d_i | B_i(\theta) + (1 + f_i(\theta)_{\text{QCD}}) \mu_i^{Z \rightarrow \nu\nu} + R_i^{Z \rightarrow \circ \circ} (1 + f_i(\theta)_{\text{EW}}) \mu_i^{Z \rightarrow \nu\nu} + \mu S_i(\theta) \right) \times \\ \prod_i \text{Pois} \left( d_i^Z | B_i^Z(\theta) + \frac{\mu_i^{Z \rightarrow \nu\nu}}{R_i^Z(\theta)_{\text{QCD}}} + \frac{\mu_i^{Z \rightarrow \nu\nu}}{R_i^Z(\theta)_{\text{EW}}} \right) \\ \prod_i \text{Pois} \left( d_i^W | B_i^W(\theta) + \frac{f_i(\theta)_{\text{QCD}} \mu_i^{Z \rightarrow \nu\nu}}{R_i^W(\theta)_{\text{QCD}}} + \frac{f_i(\theta)_{\text{EW}} \mu_i^{Z \rightarrow \nu\nu}}{R_i^W(\theta)_{\text{EW}}} \right) \times \quad (3)$$

550 The fit is performed simultaneously in the four different control samples and in the signal re-  
 551 gion, for events selected in the vbf category, to estimate the  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  rate in  
 552 each  $m_{jj}$  bin. In this likelihood, the expected number of  $Z(\nu\nu) + \text{jets}$  events in each bin of  $m_{jj}$  are  
 553 the free parameters of the fit. The  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  rates are estimated separately  
 554 for the QCD and EW components separately in each  $m_{jj}$  bin. However the fit is constrained  
 555 using the  $R_i^{Z \rightarrow \circ \circ}$  which demonstrates the ratio between the QCD and EW components of the  
 556  $Z(\nu\nu) + \text{jets}$  background. This ratio does not have any additional uncertainty. The systematic  
 557 uncertainties ( $\theta$ ) enter the likelihood as additive perturbations to the transfer factors  $R_i^{Z/W}$ ,  
 558 and are modeled as Gaussians. The parameter  $\mu_i^{Z \rightarrow \nu\nu}$  represents the yield of the  $Z(\nu\nu) + \text{jets}$   
 559 background in the  $i$  dijet mass bin into the signal region, and is left freely floating in the fit. The  
 560 function  $f_i(\theta)$  is the transfer factor between the  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  backgrounds in  
 561 the signal region and represents a constraint between these backgrounds. The likelihood also  
 562 includes the signal region with  $B_i$  representing all the backgrounds,  $S$  representing the nominal  
 563 signal prediction, and  $\mu$  being the signal strength parameter also left floating in the fit.

### 564 7.2 Transfer factors

Transfer factors, derived from simulation, are used to link the yields of the  $Z(\ell\ell) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  processes in the control regions with the  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  background estimates in the signal region. These transfer factors are defined as the ratio of expected yields of the target process in the signal region and the process being measured in the control sample. As an example:

$$R_i^Z = \frac{N_{i,MC}^{Z \rightarrow \mu^+ \mu^-}}{N_{i,MC}^{Z \rightarrow \nu\nu}} \quad (4)$$

565 where  $N_i$  is the number of events in bin  $i$  of the dijet mass distribution,  $R_i$  is the transfer fac-  
 566 tor between the dimuon control region and  $Z \rightarrow \nu\nu$  background. Other transfer factors are  
 567 constructed in similar manner.

568 To estimate the  $W(\ell\nu) + \text{jets}$  background in the signal region, the transfer factors between  
 569 the  $W(\mu\nu) + \text{jets}$  and  $W(e\nu) + \text{jets}$  event yields in the single-lepton control samples and the  
 570  $W(\ell\nu) + \text{jets}$  background estimates in the signal region are constructed. These transfer factors  
 571 take into account the impact of lepton acceptances and efficiencies, lepton veto efficiencies, and  
 572 the difference in the trigger efficiencies in the case of the single-electron control sample. These  
 573 transfer factors are shown in Figure 44. The dotted red (blue) line shows the ratio of the pro-  
 574 cesses where the V-boson is produced via EW (QCD) production where as the solid lines show  
 575 the ratio of the processes where the V-boson is produced both via EW and QCD production.

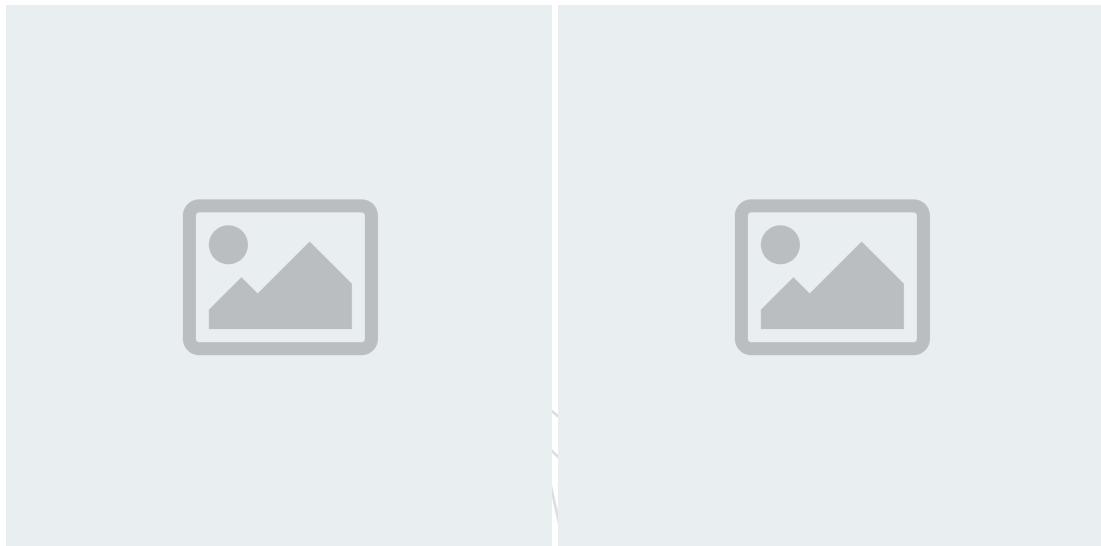


Figure 44: Transfer factors for the  $W \rightarrow \ell\nu$  background as a function of the dijet mass using the single muon and single electron control regions for the vbf final state. The grey band shows the statistical and systematic uncertainties on the ratios.

576 The  $Z \rightarrow \nu\nu$  background prediction in the signal region is connected to the yields of  $Z \rightarrow$   
 577  $\mu^+ \mu^-$  and  $Z \rightarrow e^+ e^-$  events in the dilepton control samples. The associated transfer factors ac-  
 578 count for the differences in the branching ratio of Z bosons to charged leptons relative to neutrinos and the impact of lepton acceptance and selection efficiencies. In the case of dielectron  
 579 events, the transfer factor also takes into account the difference in the trigger efficiencies. The  
 580 resulting constraint on the  $Z \rightarrow \nu\nu$  background from the dilepton control samples is limited by  
 581 the statistical uncertainty in the dilepton control samples due to the large branching fraction  
 582 difference of the Z boson decays to muons and electrons compared to that to neutrinos. These  
 583 transfer factors are shown in Figure 45. The dotted red (blue) line shows the ratio of the pro-  
 584 cesses where the V-boson is produced via EW (QCD) production where as the solid lines show  
 585 the ratio of the processes where the V-boson is produced both via EW and QCD production.

586 Finally, a transfer factor is also defined to connect the  $Z(\nu\nu) + \text{jets}$  and  $W(\ell\nu) + \text{jets}$  background  
 587 yields in the signal region to further benefit from larger statistical power that  $W(\ell\nu) + \text{jets}$  back-  
 588 ground has making it possible to experimentally constrain  $Z(\nu\nu) + \text{jets}$  production at high dijet  
 589 masses. These transfer factors are shown in Figure 46. The dotted red (blue) line shows the  
 590 ratio of the processes where the V-boson is produced via EW (QCD) production where as the  
 591 solid lines show the ratio of the processes where the V-boson is produced both via EW and  
 592 QCD production.

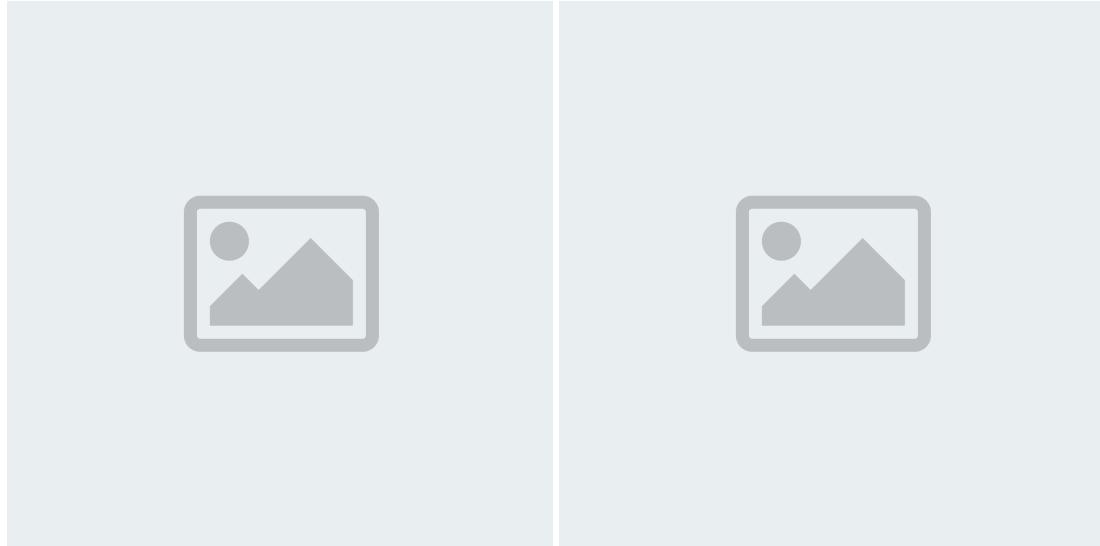


Figure 45: Transfer factors for the  $Z(\nu\nu) + \text{jets}$  background as a function of the dijet mass using the dimuon, and dielectron control regions in vbf final state. The grey band shows the statistical and systematic uncertainties on the ratios.

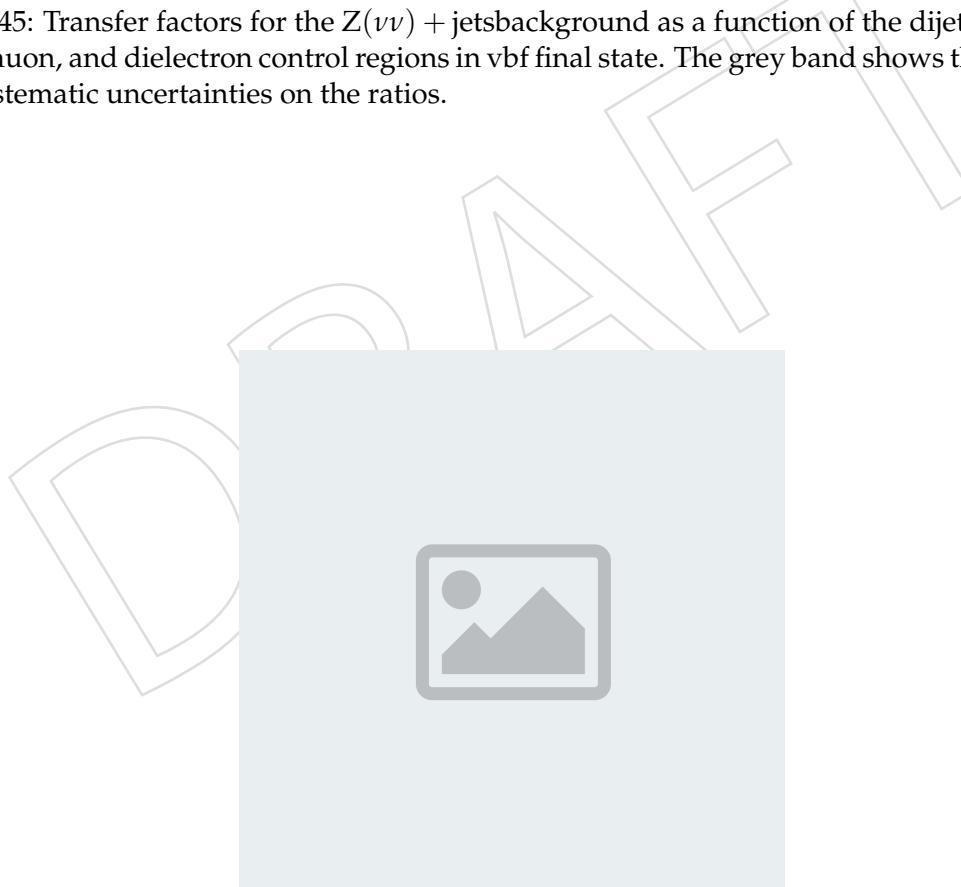


Figure 46: Transfer factors for the to estimate the  $Z(\nu\nu)+\text{jets}$  background from  $W+\text{jets}$  in the signal region as a function of the dijet mass for the vbf final state. The grey band shows the statistical and systematic uncertainties on the ratios.

594 **7.3 Systematic uncertainties**

595 Systematic uncertainties in the transfer factors are modeled as constrained nuisance parameters  
596 and include both experimental and theoretical uncertainties in the  $W + \text{jets}$  to  $Z + \text{jets}$  differen-  
597 tial cross section ratios.

598 **7.3.1 Theoretical uncertainties**

599 Theoretical uncertainties in  $V + \text{jets}$  processes include effects from QCD and EW higher-order  
600 corrections along with PDF modeling uncertainty. One of the uncertainties considered comes  
601 from the variations around the central renormalization and factorization scale choice. It is eval-  
602 uated by taking the differences in the NLO cross section as a function of boson  $p_T$  after chang-  
603 ing the renormalization and factorization scales by a factor of two and a factor of one-half with  
604 respect to the default value. These constant scale variations mainly affect the overall normal-  
605 ization of the boson  $p_T$  distributions. This uncertainty is treated to be uncorrelated across the  
606  $Z + \text{jets}$ ,  $W + \text{jets}$  processes, but correlated across the bins of the dijet mass distribution.

607 The PDF uncertainty has been estimated using the standard deviations of weights provided in  
608 the NNPDF3.0 parton distribution set and using the RMS of each bin of the distribution after  
609 varying the full spectra by these weights. This uncertainty is treated to be correlated across the  
610  $Z + \text{jets}$ ,  $W + \text{jets}$  processes, and the bins of the dijet mass distribution.

611 The uncertainty due the EW corrections is assumed to be the full correction itself as a conser-  
612 vative approach. The uncertainty is treated correlated across the processes and across the bins  
613 of the dijet mass spectra.

614 These uncertainties are applied to the QCD and EW  $V + \text{jets}$  processes, but are assumed to be  
615 uncorrelated. That is, the QCD  $W + \text{jets}$  component has a factorization scale uncertainty that is  
616 the same size as the factorization scale uncertainty on the EW  $W + \text{jets}$  component, but the two  
617 are treated as separate nuisances in the signal extraction fit.

618 The summary of the aforementioned theoretical uncertainties including their magnitude per  
619 process and on the ratio are shown in Figure 47

620 **7.3.2 Experimental uncertainties**

621 **Check numbers for 2017/18** Experimental uncertainties including the reconstruction efficiency  
622 (1% per muon or electron), and selection efficiencies of leptons (1% per muon and 2% per  
623 electron), and hadronically decaying  $\tau$  leptons (5%) are incorporated to the analysis. These  
624 reconstruction and selection efficiencies further translate into an uncertainty in the lepton veto  
625 efficiency of 3%. The lepton veto uncertainties in the transfer factors and are estimated through  
626 propagating the overall uncertainty on the tagging scale factor (loose-muon ID, veto-electron  
627 ID and loose MVA-tau ID) into the vetoed selection based on both the flavour composition  
628 of the  $W + \text{jets}$  process the acceptance of the lepton. The overall magnitude of the lepton-  
629 veto uncertainty is found to be around 3 (4)% and is found to be dominated by the  $\tau$ -veto  
630 uncertainty.

631 The uncertainty in the modeling of  $E_T^{\text{miss}}$  in simulation [31] is estimated to be 1-2% on the ratios  
632 and is dominated by the uncertainty on the jet energy scale. The jet energy scale uncertainties  
633 found to be not-cancelling in the ratio due to the differences in the jet kinematics between  
634 control and signal regions. The resulting jet energy scale uncertainties in the ratios can be seen  
635 in Figure 48

636 Lastly, uncertainties in the efficiency of the electron (2%), and  $p_T^{\text{miss}}$  (2%) triggers, are included

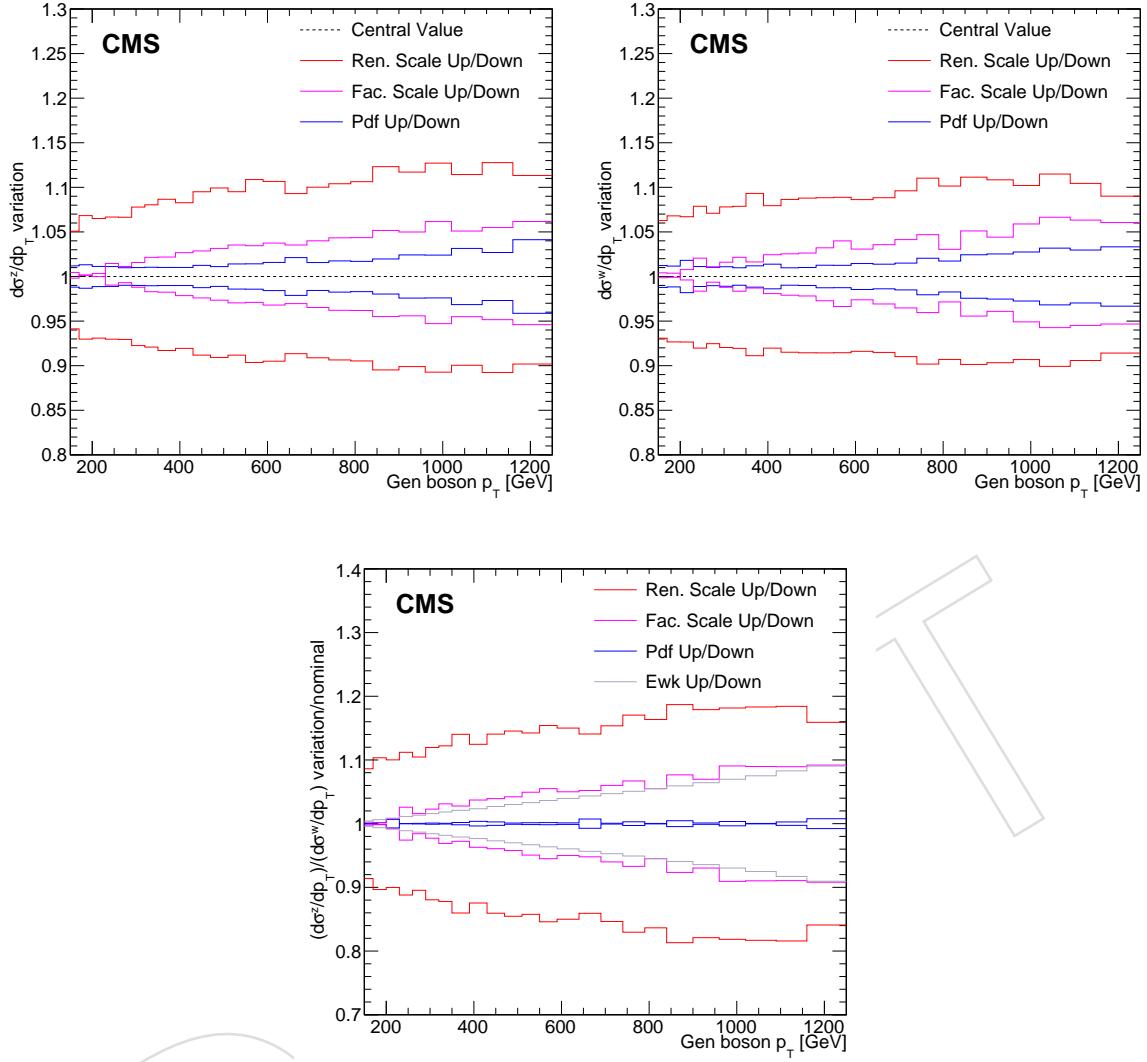


Figure 47: Theoretical uncertainties due to QCD and EW higher-order corrections and the PDF variation is shown for individual processes and for the ratio.

and are fully correlated across all the bins of dijet mass distribution.

The remaining uncertainties are for the monte carlo based backgrounds. A systematic uncertainty of 10% for the top quark background due to the modeling of the top quark  $p_T$  distribution in simulation. The uncertainty in the efficiency of the b jet veto is estimated to be 3% (1%) for the top quark (diboson) background. In addition, systematic uncertainties of 10 and 20% are included in the normalizations of the top quark [32] and diboson backgrounds [33, 34], respectively, to account for the uncertainties in their cross sections in the relevant kinematic phase space. Lastly, the uncertainty in the QCD multijet background estimate is found to be between 50–150% due to the variations of the jet response and the statistical uncertainty of the extrapolation factors.

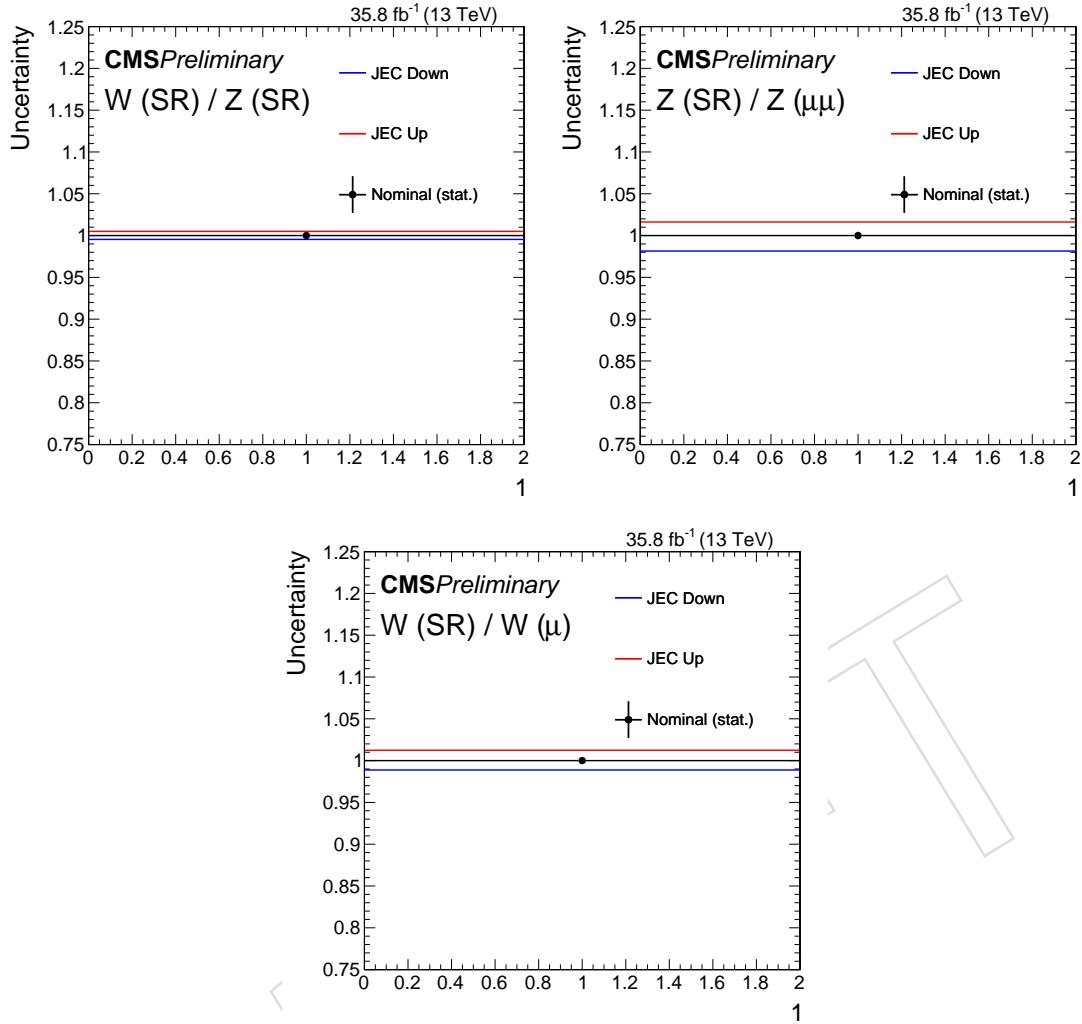


Figure 48: Redo for 2017/18 The resulting jet energy scale uncertainties in the ratios of  $W_{SR}/Z_{vv}$ ,  $Z_{vv}/Z_{\mu\mu}$  and  $W_{SR}/W_{\mu\mu}$ , respectively.

#### 647 7.4 Control sample validation

648 An important cross-check of the application of  $p_T$ -dependent NLO QCD and EW corrections is  
 649 represented by the agreement between data and simulation in the ratio of  $Z + \text{jets}$  events and  
 650  $W + \text{jets}$  events in the control samples as a function of  $m_{jj}$ .

651 Figure 49 shows the ratio between the dijet mass spectra in the single and double lepton control  
 652 regions, which is a proxy for the agreement of the  $Z/W$  ratio in data and simulation. Good  
 653 agreement is observed between data and simulation after the application of all corrections.

#### 654 7.5 QCD multijet estimation

655 Update for 2017/18

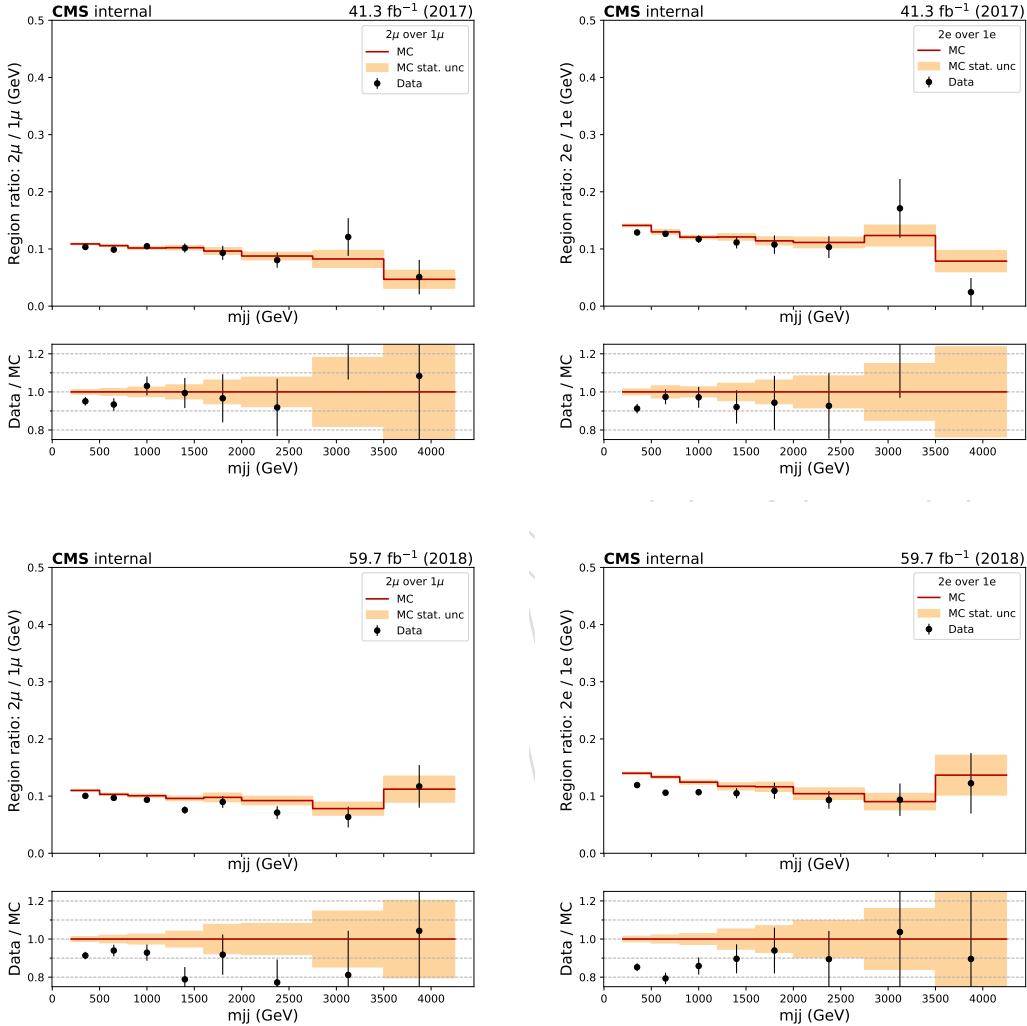


Figure 49: Comparison between data and MC simulation for the ratios of the dijet mass spectra between the single and double lepton regions. The ratios are shown for the muon regions (left) and electron regions (right), and are split by data-taking year with 2017 in the top, and 2018 in the bottom row. The uncertainty bands includes the statistical uncertainty in the simulation.

Source	Process	Uncertainty
Electron trigger	$W_{SR}/W_{ev}, Z_{vv}/Z_{ee}$	1%
$E_T^{\text{miss}}$ trigger	$W_{SR}/W_{CR}, Z_{vv}/Z_{CR}, Z/W,$	2%
Muon-reco efficiency	$W_{SR}/W_{\mu\nu}, Z_{vv}/Z_{\mu\mu}$	1% (per leg)
Muon-ID efficiency	$W_{SR}/W_{\mu\nu}, Z_{vv}/Z_{\mu\mu}$	1% (per leg)
Muon-Iso efficiency	$W_{SR}/W_{\mu\nu}, Z_{vv}/Z_{\mu\mu}$	0.5% (per leg)
Electron-reco efficiency	$W_{SR}/W_{ev}, Z_{vv}/Z_{ee}$	1% (per leg)
Electron-IDiso efficiency	$W_{SR}/W_{ev}, Z_{vv}/Z_{ee}$	1.5% (per leg)
Lepton veto	$W_{LSR}$	3% (QCD), 5.0% (EW)
Jet energy scale	$W_{CR}/W_{SR}$	2%
Jet energy scale	$Z_{CR}/Z_{SR}$	1%

Table 12: **Check numbers for 2017/18** Summary of experimental uncertainties affecting used in the analysis.

656    **8   Results**

DRAFT

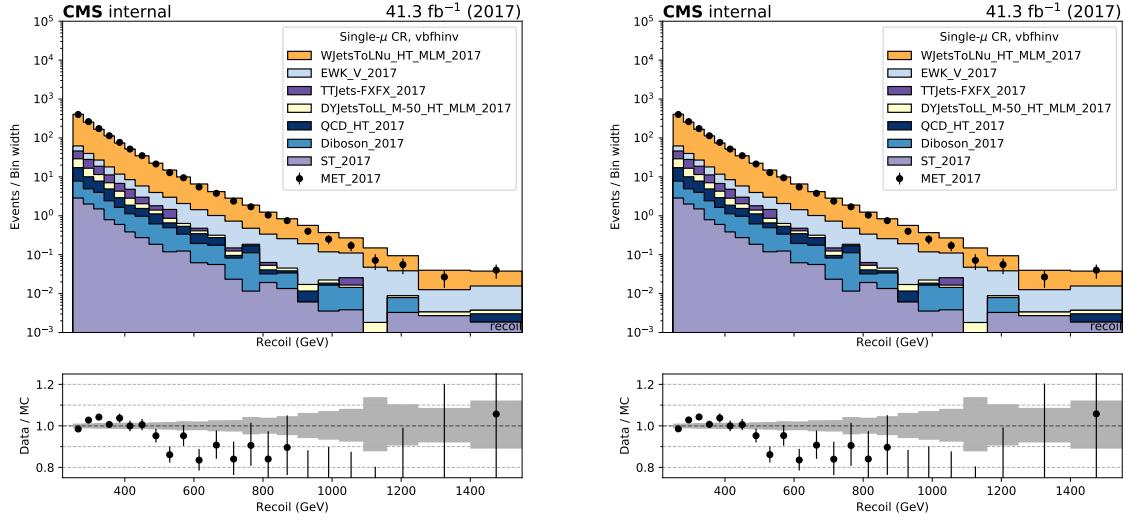


Figure 50: The recoil distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2017 samples.

## A Comparison of 1D and 2D QCD k-factors

In this appendix, several distributions such as recoil,  $M_{jj}$  and  $\Delta\eta_{jj}$  are compared for two cases in the single muon control region:

- Plots with 1D k-factors applied: Function of gen-boson  $p_T$
- Plots with 2D k-factors applied: Function of gen-boson  $p_T$  and  $M_{jj}$

Both k-factors are derived from samples in which the VBF selections are applied at the gen level. 1D k-factor is a function of generator level vector boson  $p_T$ , whereas 2D k-factor is a function of both generator level boson  $p_T$  and  $M_{jj}$ . The distributions are shown in the following pages.

Fig. 50 shows the comparison of the recoil distribution with 2017 samples, while Fig. 51 shows the same comparison with 2018 samples. Fig. 52 shows the comparison of the  $M_{jj}$  distribution with 2017 samples, while Fig. 53 shows the same comparison with 2018 samples. Fig. 54 shows the comparison of the  $\Delta\eta_{jj}$  distribution with 2017 samples, while Fig. 55 shows the same comparison with 2018 samples. Fig. 56 shows the comparison of the muon  $p_T$  distribution with 2017 samples, while Fig. 57 shows the same comparison with 2018 samples. Fig. 58 shows the comparison of the muon  $\eta$  distribution with 2017 samples, while Fig. 59 shows the same comparison with 2018 samples.

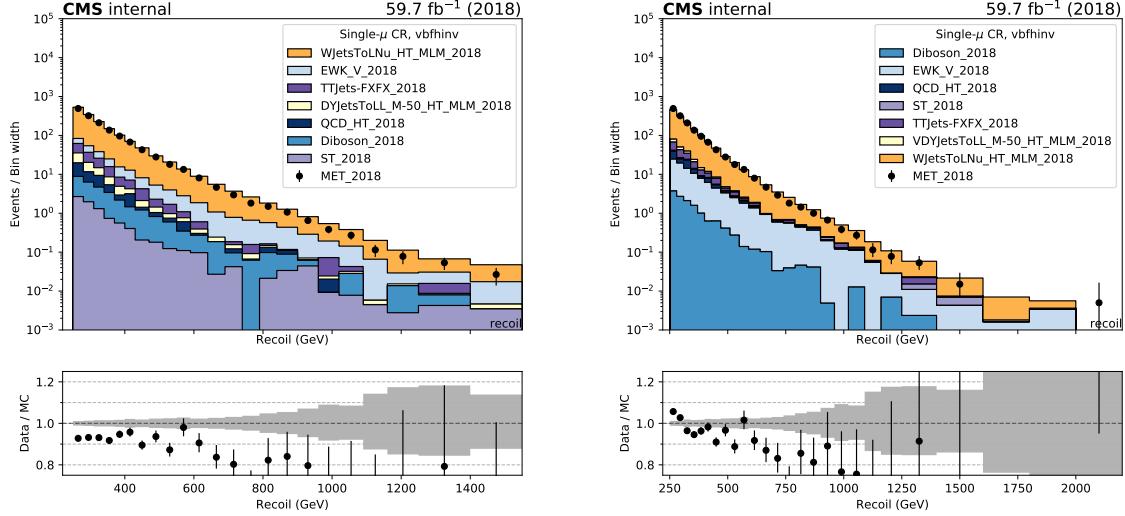


Figure 51: The recoil distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2018 samples.

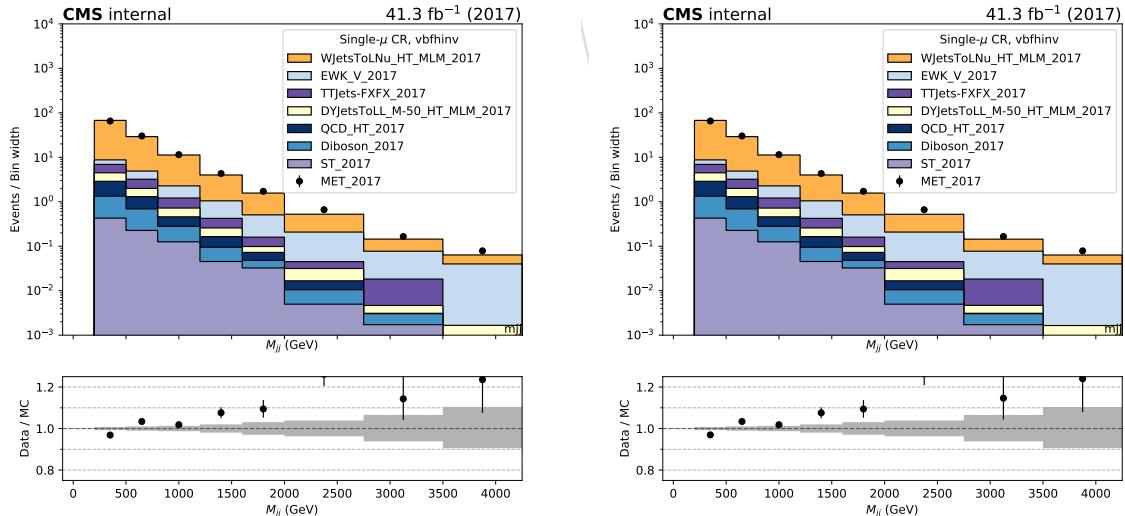


Figure 52: The  $M_{jj}$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2017 samples.

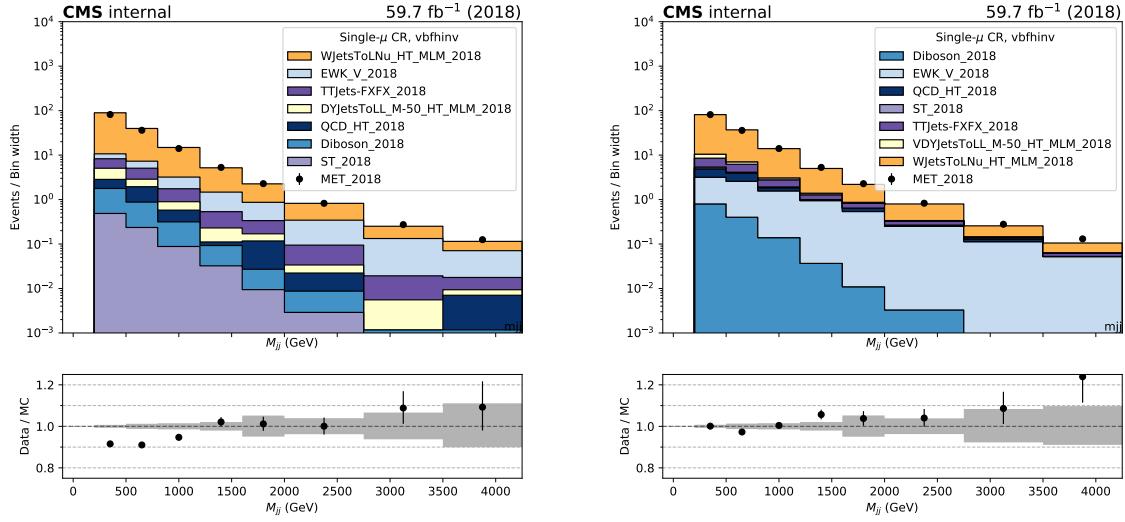


Figure 53: The  $M_{jj}$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2018 samples.

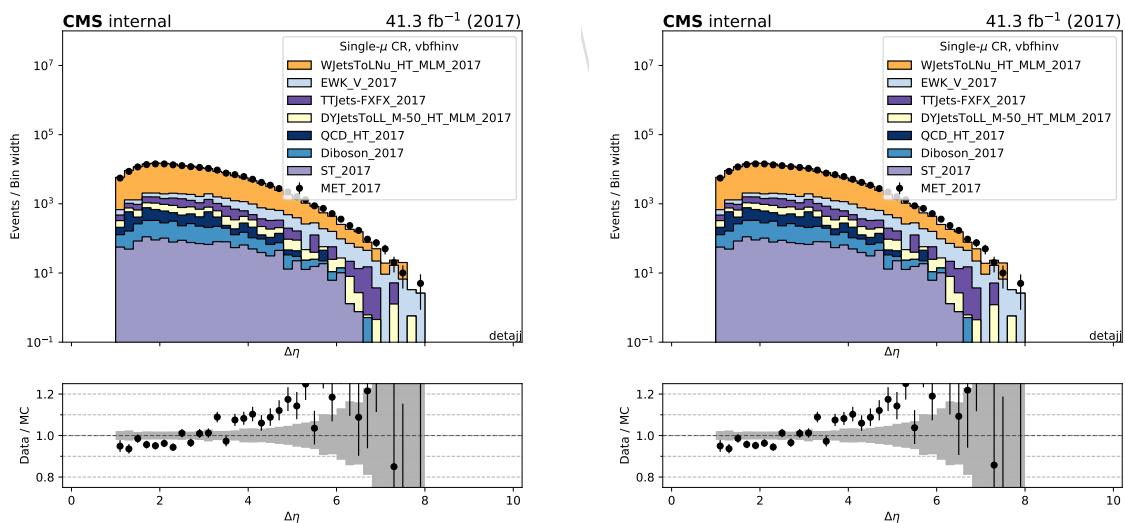


Figure 54: The  $\Delta\eta_{jj}$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2017 samples.

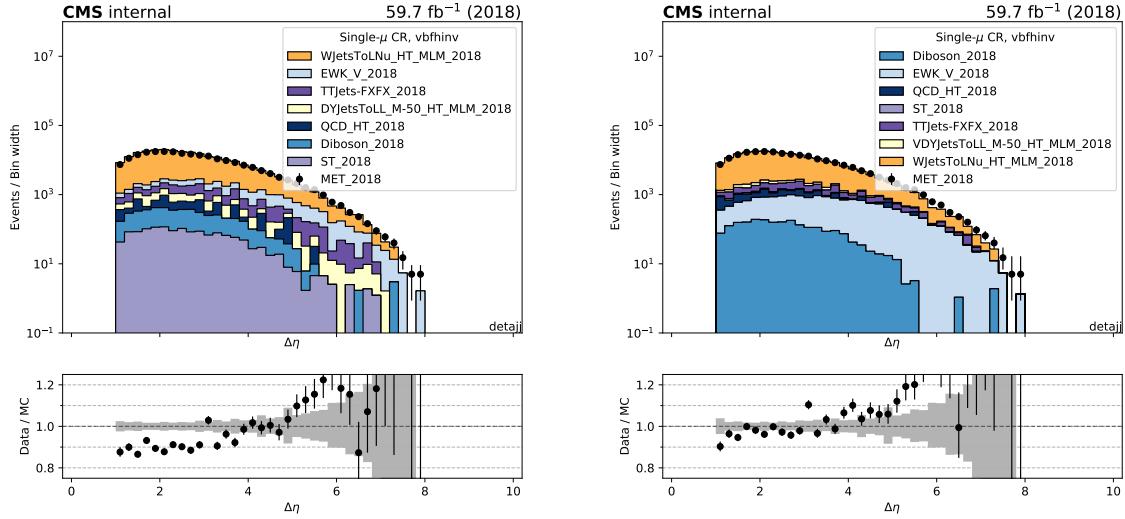


Figure 55: The  $\Delta\eta_{jj}$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2018 samples.

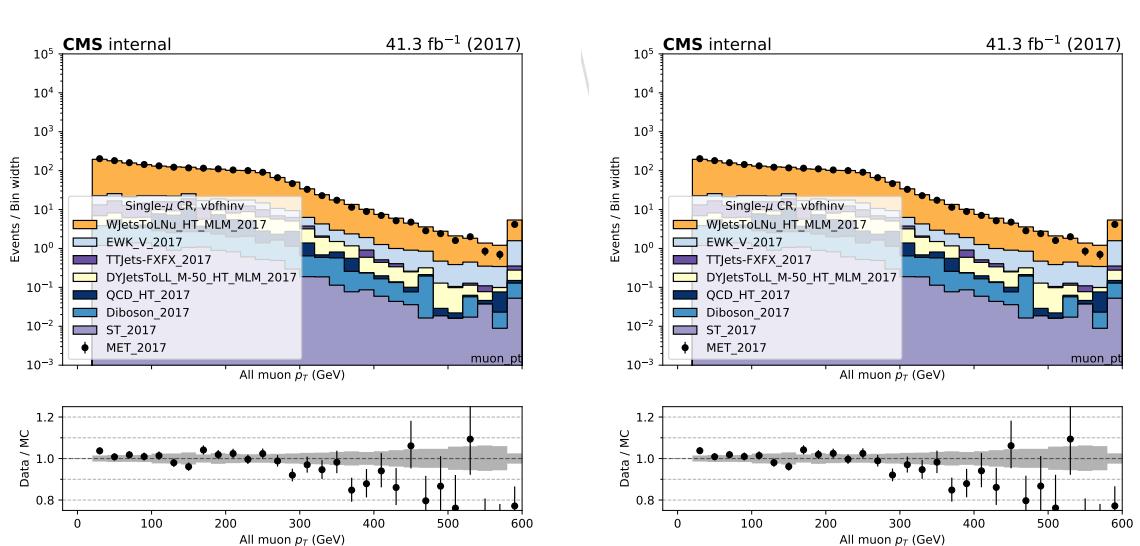


Figure 56: The muon  $p_T$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2017 samples.

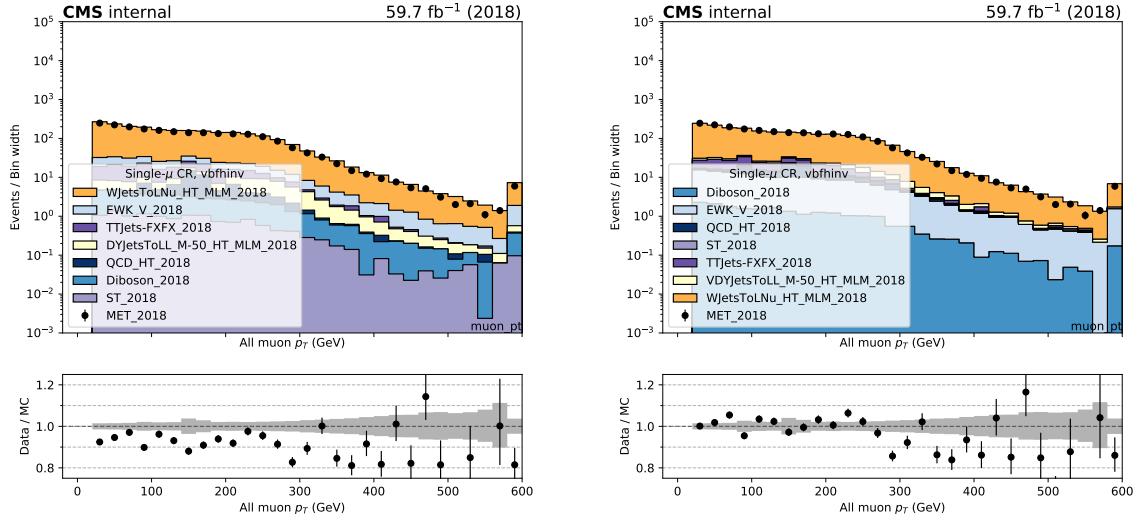


Figure 57: The muon  $p_T$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2018 samples.

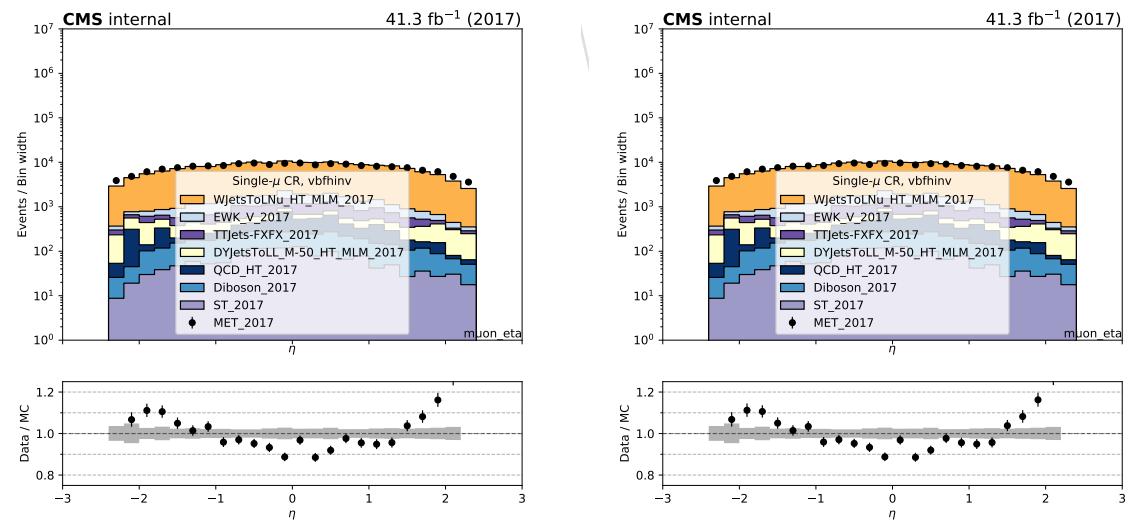


Figure 58: The muon  $\eta$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2017 samples.

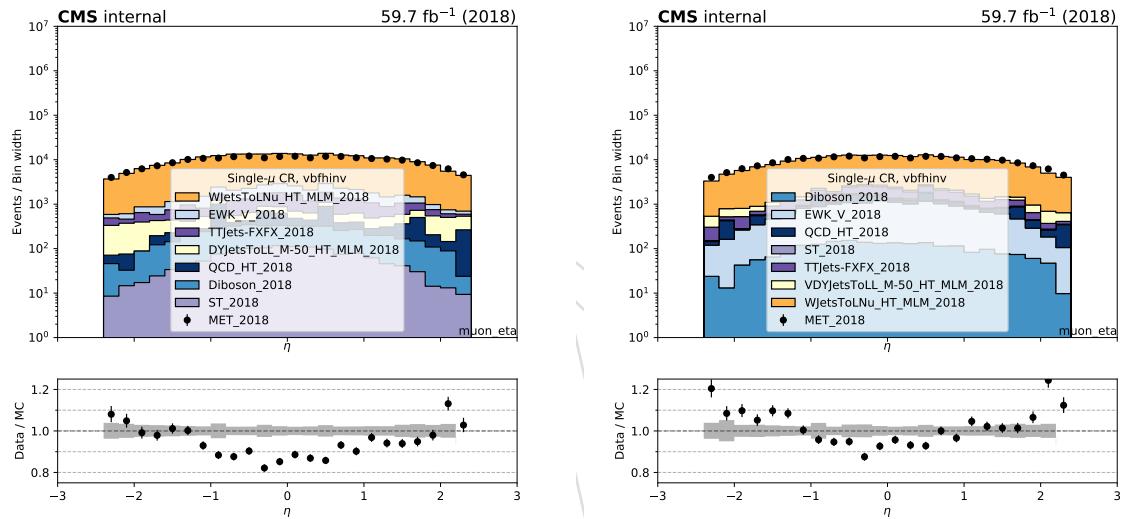


Figure 59: The muon  $\eta$  distribution in the single muon CR for the case in which 1D k-factors are used (left), compared to the case in which 2D k-factors are used (right), using 2018 samples.

## 673 B Trigger efficiencies in double muon CR

674 In this section, the efficiencies of  $p_T^{\text{miss}} + H_T^{\text{miss}}$  triggers measured in double muon control region  
 675 are presented. In addition, efficiencies of these triggers as a function of  $m_{jj}$  both in single muon  
 676 and double muon control region are also presented.

677 Fig. 60 shows the efficiencies as a function of  $M_{jj}$  in data and MC for the three categories in  
 678 2017, whereas Fig. 61 shows the efficiencies as a function of  $M_{jj}$  in data and MC for the three  
 679 categories in 2018.

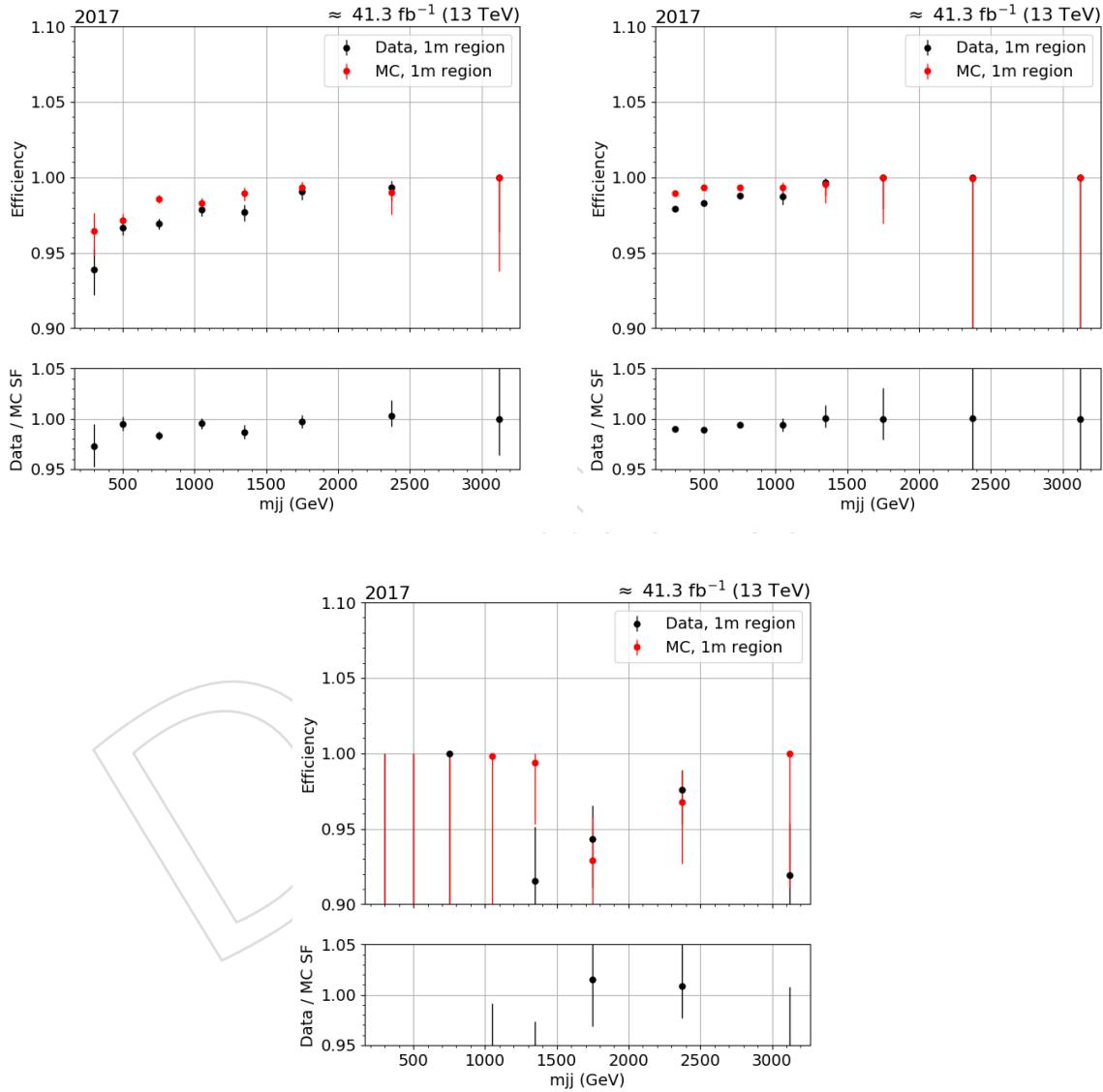


Figure 60: MET trigger efficiency as a function of  $m_{jj}$  in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2017 data and MC samples with the selection of single muon events.

680 Figs. 62 and ?? show the efficiencies obtained in double muon CR as a function of recoil in data  
 681 and MC for the three jet categories in 2017 and 2018, respectively. Figs. 64 and ?? show the  
 682 efficiencies obtained in double muon CR as a function of  $M_{jj}$  in data and MC for the three jet

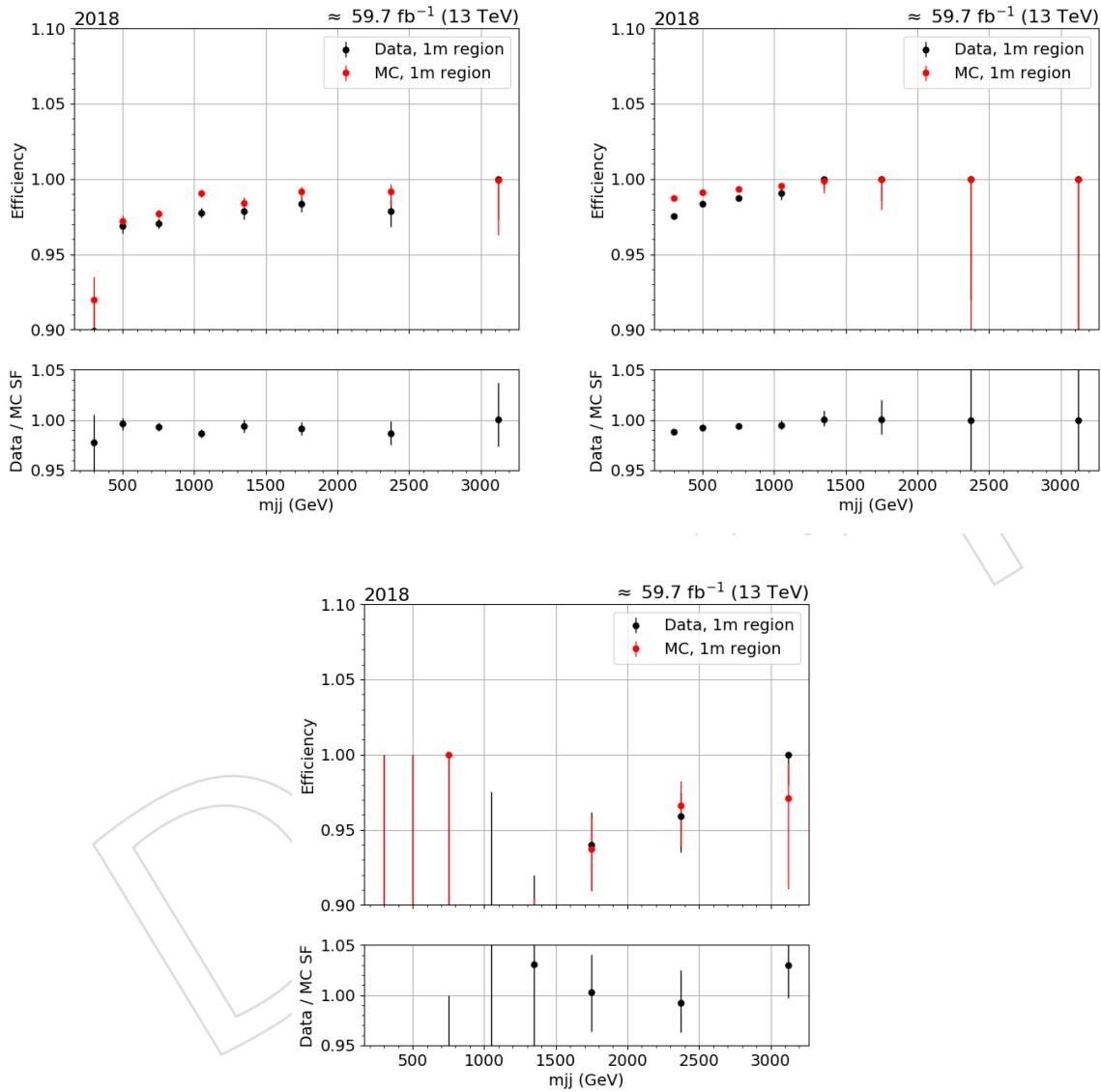


Figure 61: MET trigger efficiency as a function of  $m_{jj}$  in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2018 data and MC samples with the selection of single muon events.

683 categories in 2017 and 2018, respectively.

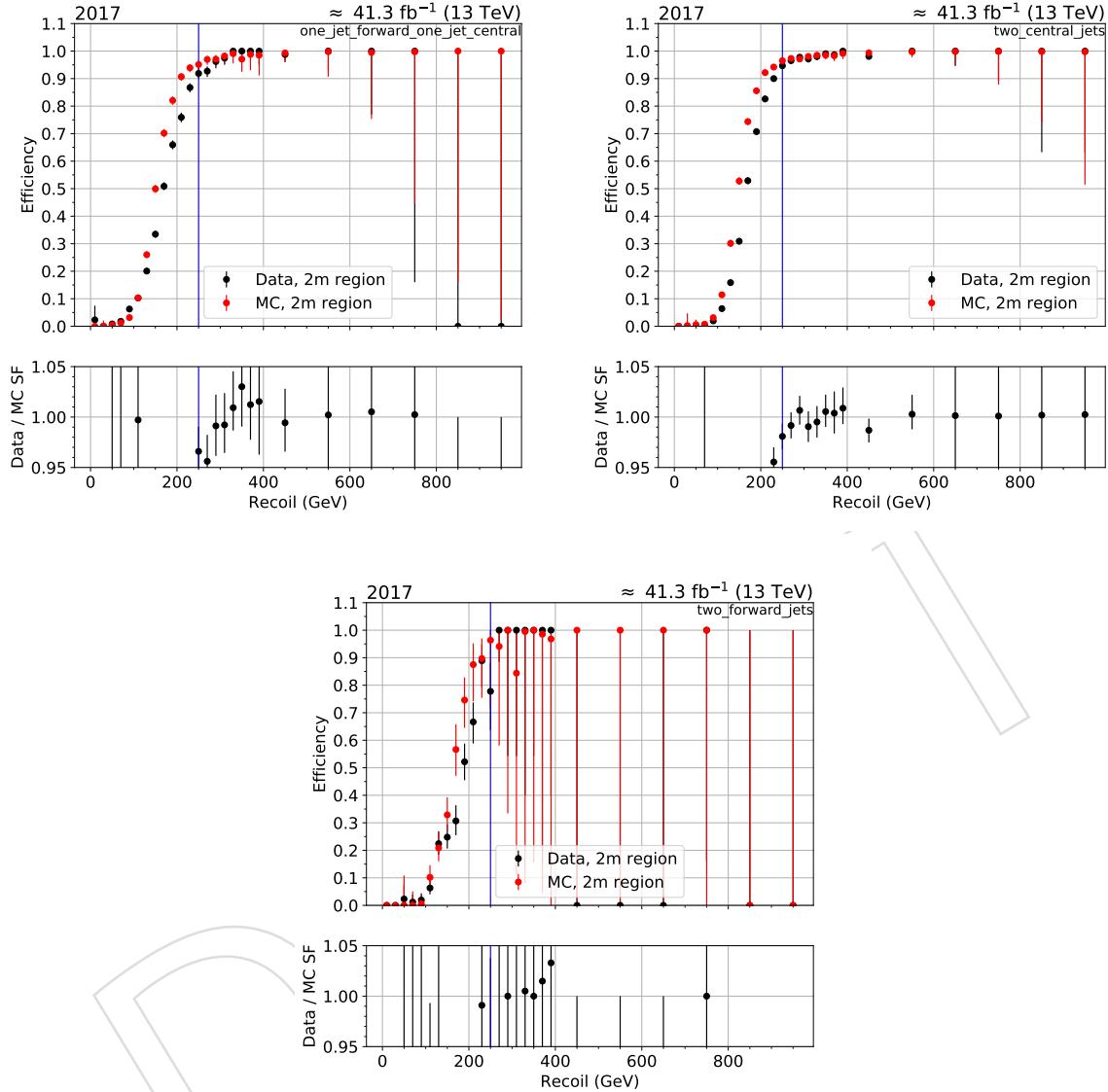


Figure 62: MET trigger efficiency as a function of recoil in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2017 data and MC samples with the selection of double muon events.

684 Fig. 66 shows the scale factors as a function of recoil in double muon control region, using 2017  
 685 (left) and 2018 (right) samples.

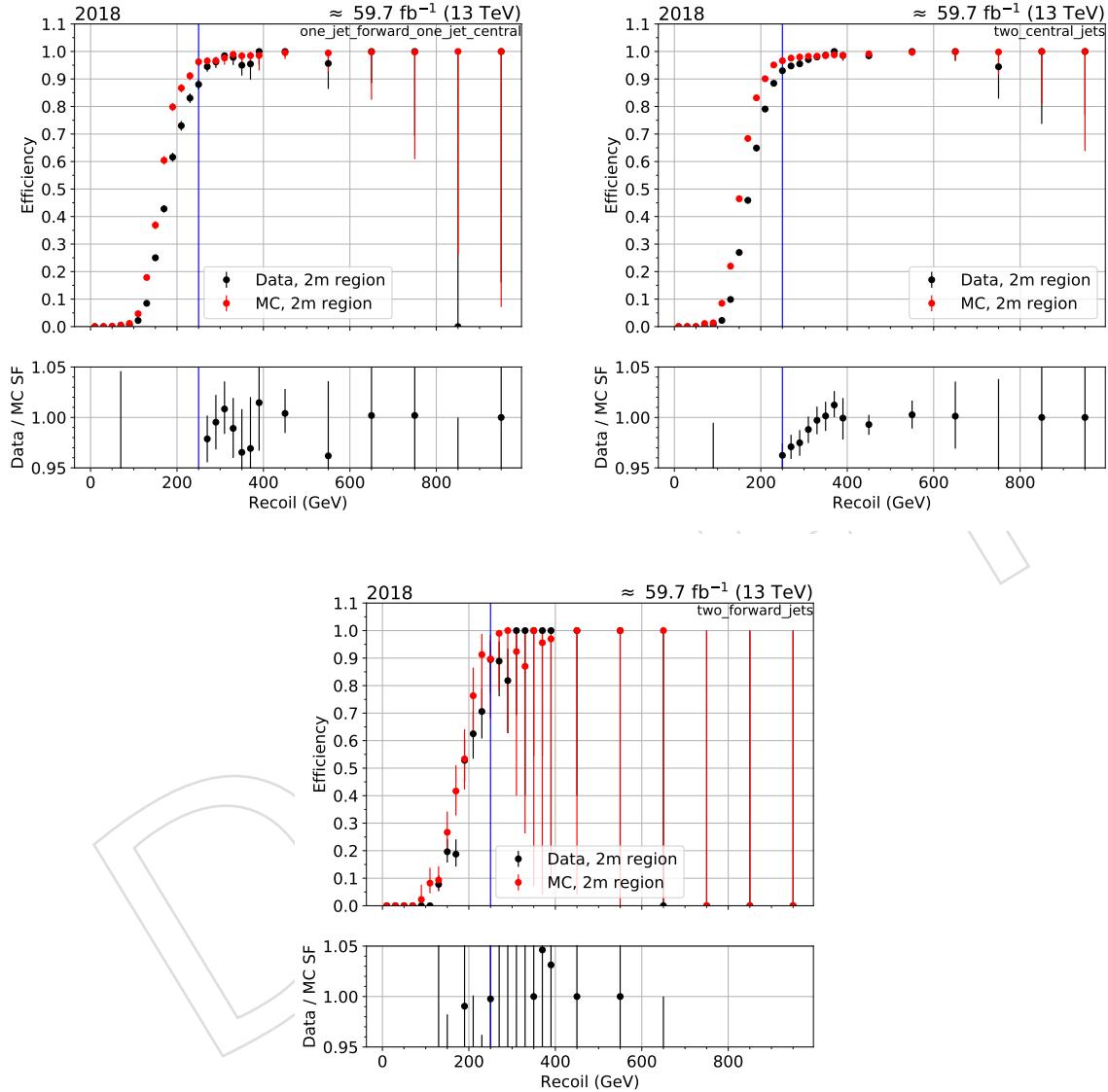


Figure 63: MET trigger efficiency as a function of recoil in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2018 data and MC samples with the selection of double muon events.

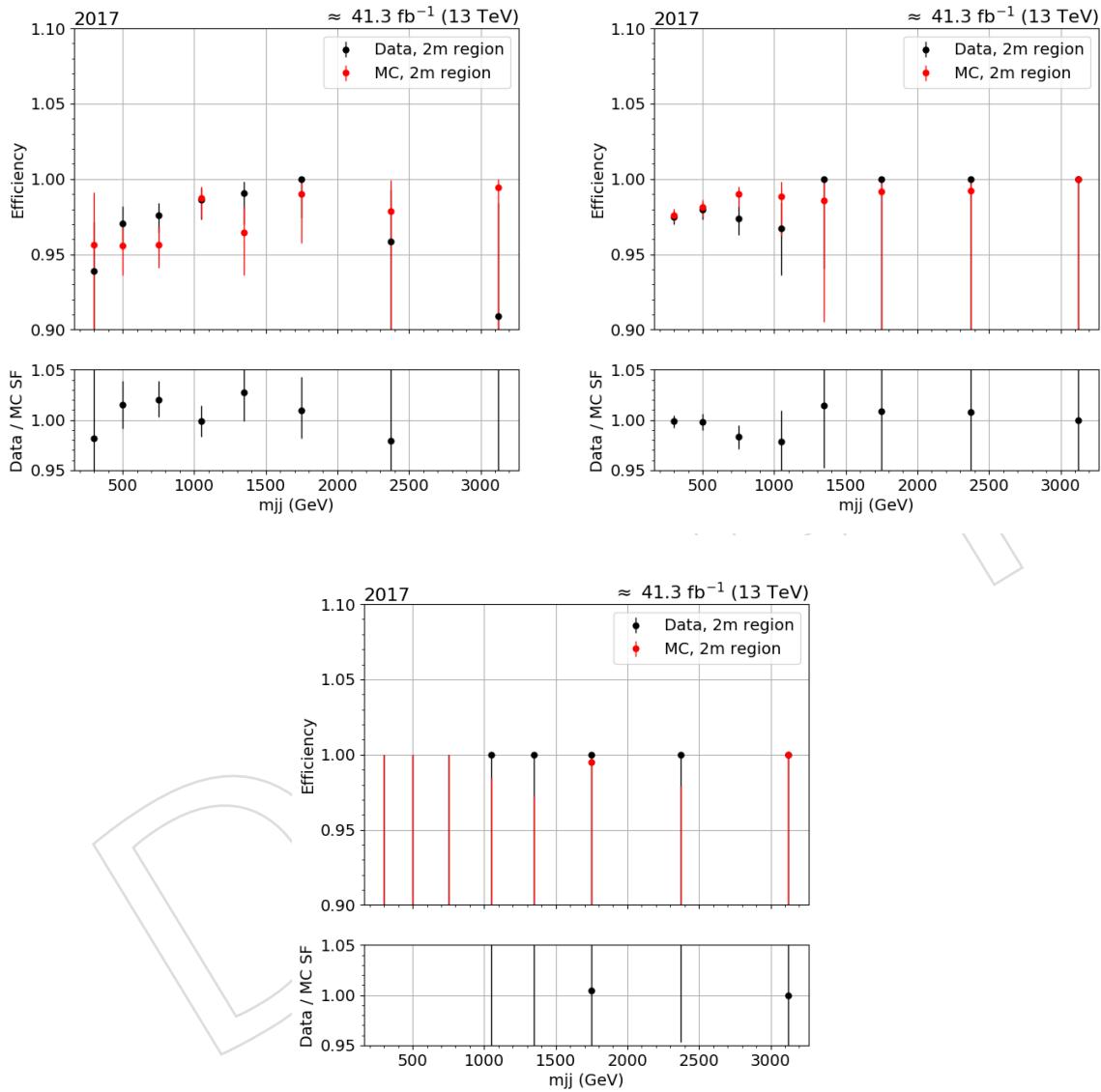


Figure 64: MET trigger efficiency as a function of  $m_{jj}$  in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2017 data and MC samples with the selection of double muon events.

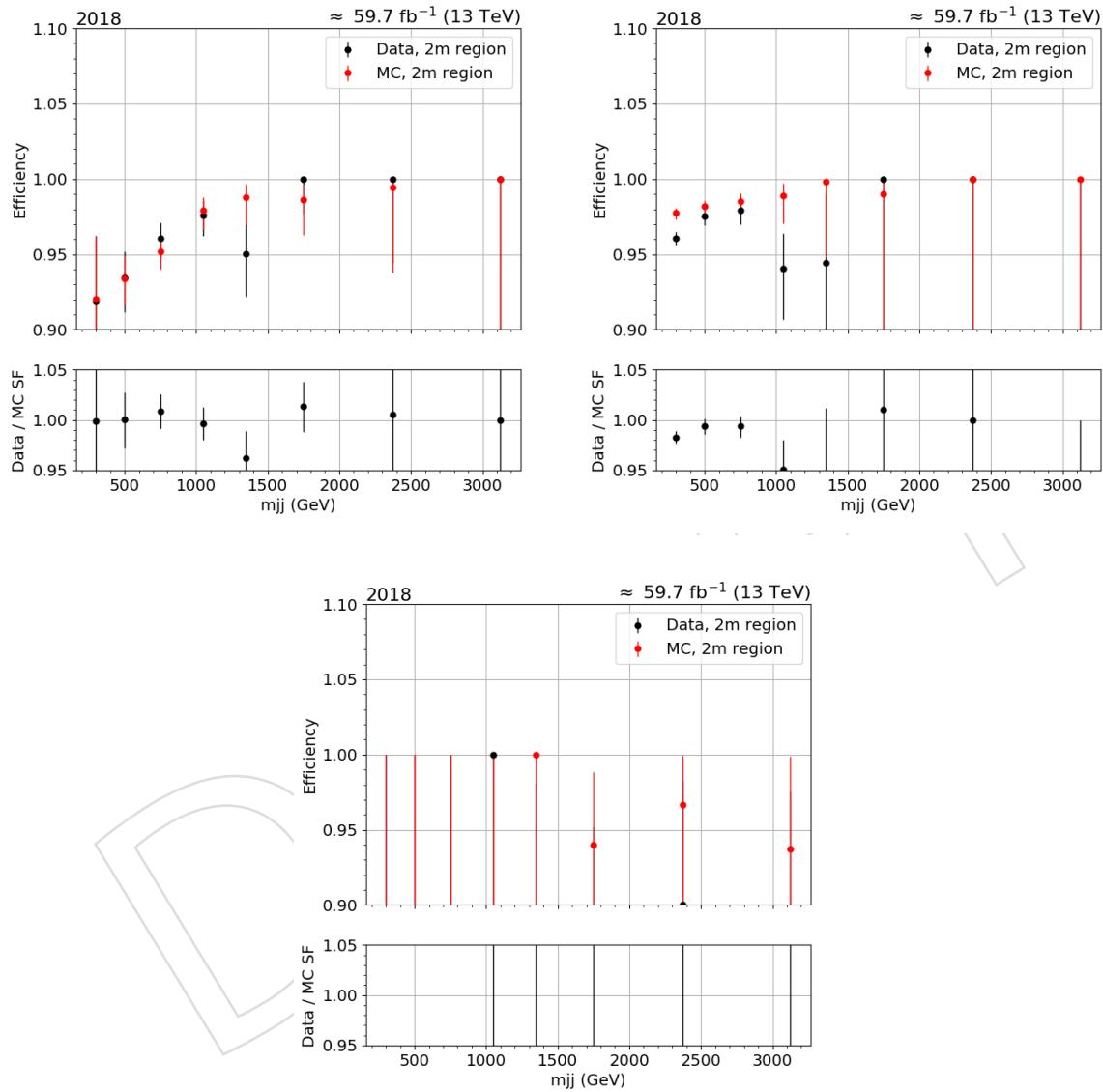


Figure 65: MET trigger efficiency as a function of  $m_{jj}$  in three categories: One forward jet and one central jet, two central jets and two forward jets. These results are obtained from 2018 data and MC samples with the selection of double muon events.

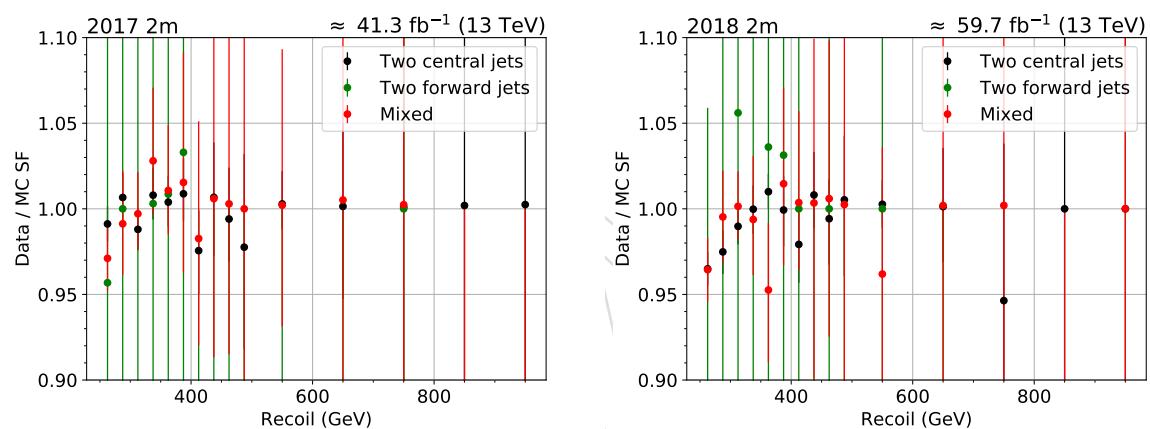


Figure 66: Data-MC scale factors for the three jet kinematic cases in the double muon control region, using 2017 (left) and 2018 (right) samples.

## 686 Acknowledgments

## 687 References

- 688 [1] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA8 tunes from  
689 underlying-event measurements”, arXiv:1903.12179.
- 690 [2] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, “Matching matrix elements and  
691 shower evolution for top-quark production in hadronic collisions”, *JHEP* **01** (2007) 013,  
692 doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.
- 693 [3] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012)  
694 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- 695 [4] NNPDF Collaboration, “Parton distributions from high-precision collider data”, *Eur.  
696 Phys. J.* **C77** (2017), no. 10, 663, doi:10.1140/epjc/s10052-017-5199-5,  
697 arXiv:1706.00428.
- 698 [5] GEANT4 Collaboration, “GEANT4: A simulation toolkit”, *Nucl. Instrum. Meth. A* **506**  
699 (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- 700 [6] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs cross  
701 sections: 4. deciphering the nature of the Higgs sector”,  
702 doi:10.2172/1345634, 10.23731/CYRM-2017-002, arXiv:1610.07922.
- 703 [7] CMS EGamma POG, “EgHLTRunIISummary”.  
704 <https://twiki.cern.ch/twiki/bin/view/CMS/EgHLTRunIISummary>.
- 705 [8] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm”, *JHEP* **04**  
706 (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- 707 [9] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, *Eur. Phys. J. C* **72** (2012)  
708 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- 709 [10] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp  
710 collisions at 8 TeV”, *JINST* **12** (2017) P02014,  
711 doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- 712 [11] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp  
713 collisions at 13 TeV”, *JINST* **13** (2018), no. 05, P05011,  
714 doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- 715 [12] CMS JME POG. <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2>.
- 717 [13] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS  
718 detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015) P06005,  
719 doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- 720 [14] CMS Collaboration, “Description and performance of track and primary-vertex  
721 reconstruction with the CMS tracker”, *JINST* **9** (2014) P10009,  
722 doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- 723 [15] CMS EGamma POG. [https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2#Spring15\\_selection\\_25ns](https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2#Spring15_selection_25ns).

- [725] [16] CMS Collaboration, "Performance of CMS muon reconstruction in pp  
 726 collision events at  $\sqrt{s} = 7 \text{ TeV}$ ", *JINST* **7** (2012) P10002,  
 727 doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [728] [17] CMS Muon POG. [https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Loose\\_Muon](https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Loose_Muon).
- [730] [18] CMS Muon POG. [https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Tight\\_Muon](https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Tight_Muon).
- [732] [19] CMS Collaboration, "Reconstruction and identification of  $\tau$  lepton decays to hadrons and  
 733  $\nu_\tau$  at CMS", *JINST* **11** (2016) P01019, doi:10.1088/1748-0221/11/01/P01019,  
 734 arXiv:1510.07488.
- [735] [20] tau-POG. [https://twiki.cern.ch/twiki/bin/view/CMS/TauIDRecommendation13TeV#Measurement\\_in\\_Z\\_tautau\\_events](https://twiki.cern.ch/twiki/bin/view/CMS/TauIDRecommendation13TeV#Measurement_in_Z_tautau_events).
- [737] [21] CMS Collaboration, "Performance of photon reconstruction and identification with the  
 738 CMS detector in proton-proton collisions at  $\sqrt{s} = 8 \text{ TeV}$ ", *JINST* **10** (2015) P08010,  
 739 doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.
- [740] [22] CMS EGamma POG. [https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedPhotonIdentificationRun2#SPRING15\\_selections\\_bunch\\_crossi](https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedPhotonIdentificationRun2#SPRING15_selections_bunch_crossi).
- [742] [23] CMS Lumi POG. [https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONfileforData#Pileup\\_JSON\\_Files\\_For\\_Run\\_II](https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONfileforData#Pileup_JSON_Files_For_Run_II).
- [744] [24] CMS Muon POG. <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceSelectionAndCalibrationsRun2>.
- [746] [25] J. M. Lindert et al., "Precise predictions for V+jets dark matter backgrounds", (2017).  
 747 arXiv:1705.04664.
- [748] [26] "Search for vbf higgs boson decays to invisible particles with 2016 dataset", 2017. CMS  
 749 Analysis Note : AN-17-267. [http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2017\\_267\\_v4.pdf](http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2017_267_v4.pdf).
- [751] [27] CMS Collaboration, "Performance of missing transverse momentum reconstruction in  
 752 proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$  using the CMS detector", *Submitted to: JINST*  
 753 (2019) arXiv:1903.06078.
- [754] [28] P. Lenzi, C. A. Palmer, J. Thomas-Wilsker et al., "Performance of flavour tagging  
 755 algorithms at 13 tev with 2018 data", CMS Note 2018/323, 2019.
- [756] [29] CMS Collaboration, "Search for New Physics with Jets and Missing Transverse  
 757 Momentum in  $pp$  collisions at  $\sqrt{s} = 7 \text{ TeV}$ ", *JHEP* **08** (2011) 155,  
 758 doi:10.1007/JHEP08(2011)155, arXiv:1106.4503.
- [759] [30] CMS Collaboration, "Search for dark matter produced with an energetic jet or a  
 760 hadronically decaying W or Z boson at  $\sqrt{s} = 13 \text{ TeV}$ ", *JHEP* **07** (2017) 014,  
 761 doi:10.1007/JHEP07(2017)014, arXiv:1703.01651.
- [762] [31] CMS Collaboration, "Performance of the CMS missing transverse momentum  
 763 reconstruction in pp data at  $\sqrt{s} = 8 \text{ TeV}$ ", *JINST* **10** (2015) P02006,  
 764 doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

- 765 [32] CMS Collaboration, “Measurement of the top quark pair production cross section in  
766 proton-proton collisions at  $\sqrt{s} = 13$  TeV”, *Phys. Rev. Lett.* **116** (2016) 052002,  
767 doi:10.1103/PhysRevLett.116.052002, arXiv:1510.05302.
- 768 [33] CMS Collaboration, “Measurement of the ZZ production cross section and Z  
769  $\rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$  branching fraction in pp collisions at  $\sqrt{s}=13$  TeV”, *Phys. Lett. B* **763** (2016)  
770 280–303, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.
- 771 [34] CMS Collaboration, “Measurement of the WZ production cross section in pp collisions at  
772  $\sqrt{s}=13$  TeV”, *Phys. Lett. B* **766** (2017) 268–290,  
773 doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.

DRAFT