

# CMS Draft Analysis Note

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

2020/02/26

Archive Hash: 7b24b77

Archive Date: 2020/02/19

## Measurement of differential cross-section in single top tW-channel at CMS with $\sqrt{s} = 13$ TeV

Enrique Palencia Cortezón, Víctor Rodríguez Bouza, and Sergio Sánchez Cruz  
Universidad de Oviedo – ICTEA

### Abstract

A measurement of the differential tW-channel single-top-quark cross section is presented in pp collisions at  $\sqrt{s} = 13$  TeV in dilepton events. The fiducial region is defined according to the detector acceptance cuts, and requiring the presence of exactly one jet, that must be  $b$ -tagged. The presence of lower energy jets is vetoed to reduce the contribution from the dominating backgrounds. Resulting distributions are unfolded to particle level. We also provide the results normalised to the fiducial cross section.

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

PDFAuthor: Victor Rodriguez Bouza, Enrique Palencia Cortezon, Sergio Sanchez Cruz  
pdftitle=Differential cross-section in single top tW-channel  
PDFTitle: "" PDFSubject:  
CMS  
PDFKeywords: CMS, top, singletop, differential

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



# 1 Introduction

*Editor: Sergio Sánchez*

**NOTE: results and analysis associated with commit with label 0f32431 from the 29th of April of 2019.**

**NOTE: results and analysis associated with commit with label**

**b8a0b682de5cb2454f98fd46f0d34d1d7c535ddd from the 3<sup>rd</sup> of February of 2020.**

**NOTE (latest): results and analysis associated with commit with label**

**eebfd4db20e844e4236d2c8bb66493a78782d6bd from the 19<sup>th</sup> of February of 2020.**

9

10 Electroweak production of single top quarks has been first observed by the CDF [1] and D0 [2]  
 11 experiments at the Tevatron. Single-top-quark production proceeds mainly via three processes:  
 12 the t-channel exchange of a virtual W boson, the s-channel production and decay of a virtual  
 13 W boson, and the associated production of a top quark and a W boson (tW). The latter chan-  
 14 nel, which has a negligible production cross section at the Tevatron, represents a significant  
 15 contribution to single-top-quark production at the Large Hadron Collider (LHC).

16 Associated tW production is a very interesting production mechanism because of its interfe-  
 17 rence with top quark pair production [3–5], its sensitivity to new physics [6–8] and its role as a  
 18 background to SUSY and Higgs searches.

19 The definition of tW production in perturbative QCD mixes with top quark pair production ( $t\bar{t}$ )  
 20 at next-to-leading order (NLO) [3, 5, 9]. Two schemes are proposed to describe the tW signal:  
 21 "diagram removal" (DR) [3], where all NLO diagrams which are doubly resonant, such as  
 22 those in Fig. 1, are excluded from the signal definition; and "diagram subtraction" (DS) [3, 10],  
 23 in which the differential cross section is modified with a gauge-invariant subtraction term, that  
 24 locally cancels the contribution of  $t\bar{t}$  diagrams. The DR scheme is used as the nominal model in  
 25 this analysis, however the differences between the two schemes in the response of the detector  
 26 are evaluated.

27 The production cross section for tW has been computed at approximate next-to-next-to-leading  
 28 order, the theoretical prediction of the cross section for tW in pp collisions at  $\sqrt{s} = 13$  TeV,  
 29 assuming a top-quark mass ( $m_t$ ) of 172.5 GeV, is  $71.7 \pm 1.8(\text{scale}) \pm 3.4(\text{PDF}) \text{ pb}$  [11], the first  
 30 uncertainty corresponds to scale variation and the second to parton distribution function (pdf)  
 31 sets. The leading order Feynman diagrams for tW production are shown in Fig. 2.

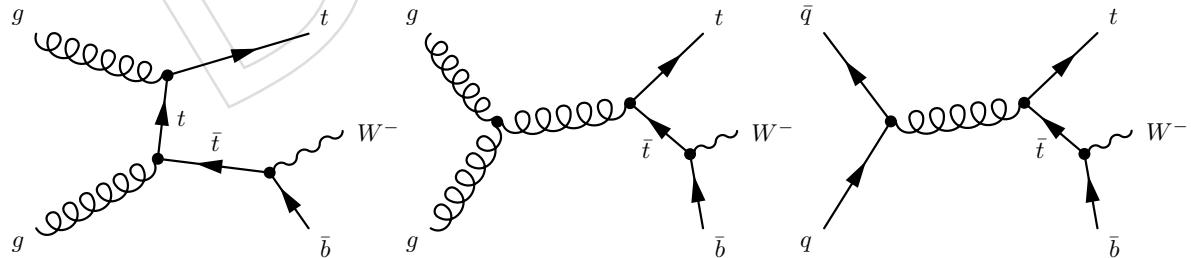


Figure 1: Feynman diagrams for tW single-top-quark production at next-to-leading order that are removed from the signal definition in the DR scheme, the charge-conjugate modes are implicitly included.

32 The tW channel was not accessible at the Tevatron due to its small cross-section in  $p\bar{p}$  collisions  
 33 at  $\sqrt{s} = 1.96$  TeV. At the LHC, however, evidence of this process with 7 TeV collision data  
 34 was presented by the CMS [12] and ATLAS [13] Collaborations. With 8 TeV collision data,

35 observations were made by the CMS [14] and the ATLAS [15] Collaborations with cross-section  
 36 measurements in good agreement with theoretical predictions.

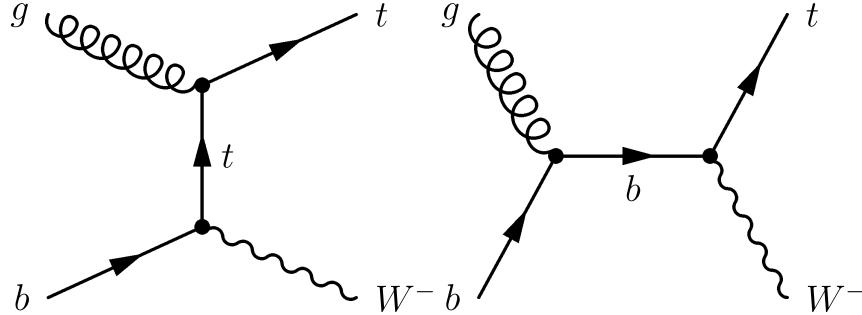


Figure 2: Leading order Feynman diagrams for single-top-quark production in the  $tW$  mode, the charge-conjugate modes are implicitly included.

37 The inclusive  $tW$  production cross-section has been measured by both CMS [16] and ATLAS [17]  
 38 Collaborations at  $\sqrt{s} = 13$  TeV with accuracies for the order of 10% and 30%, respectively. Both  
 39 measurements employ dileptonic ( $e^+e^-$ ,  $\mu^+\mu^-$  and  $\mu^\pm e^\mp$ , for ATLAS, and  $\mu^\pm e^\mp$  for CMS) chan-  
 40 nels and exploit the different ( $n_{jet}, n_{b-jet}$ ) distribution between signal and the dominating  $t\bar{t}$  back-  
 41 ground to define signal and control categories. Additionally, both analyses utilize multivariate  
 42 analysis techniques, boosted decision trees (BDT), to distinguish  $tW$  from  $t\bar{t}$ .

43 The measurement of the differential cross-section of this analysis is particularly challenging  
 44 due to the overwhelming presence of  $t\bar{t}$  in the most signal enriched region. The ATLAS Col-  
 45 laboration [18] has performed the first attempt to measure the differential cross-section of the  
 46  $tW$  production process. A cut-and-count analysis is performed in a signal-enriched region  
 47 defined by a cut on the BDT discriminant employed in the inclusive analysis. The analysis re-  
 48 ported on this note employs a selection on the number of low energy jets in the event to select  
 49 the signal-enriched region instead of a BDT cut. This represents the main difference between  
 50 the approach followed in the present analysis and the one followed by the ATLAS Collabo-  
 51 ration. This difference has several implications to the analysis, due to the fact that the current  
 52 analysis does not make use of a BDT. The cuts used to select a signal-enriched region are purely  
 53 topological and can be reproduced at the particle level. Therefore:

- 54 • any variable can be studied, while in the ATLAS approach only variables uncorre-  
 55 lated to the BDT can be studied.
- 56 • In the ATLAS approach an extrapolation from the observable phase space -given  
 57 by the cut on the BDT- to the fiducial space must be performed. This introduces  
 58 a model dependency in the analysis, and additional uncertainties associated to the  
 59 larger extrapolation. As the topological cuts imposed in the present analysis can be  
 60 reproduced at particle level, this limitation does not exist in the current analysis.

61 This note reports the first measurement of the differential cross-section measurement of  $tW$  pro-  
 62 duction in CMS. In this analysis, the measurement of the  $tW$  production cross-section as a func-  
 63 tion of the following observables:

- 64 •  $p_T$  of the leading lepton,
- 65 • the  $p_T$  of the jet, the difference in the  $\phi$  angle of the muon and the electron in the  
 66 event,
- 67 • the difference in the  $\phi$  angle between the two leptons of the events,
- 68 • the longitudinal component of the system formed by the muon, the electron and the

69 jet of the event ( $p_Z$ ),  
70 • the invariant mass of the system formed by the electron, the muon and the jet,  
71 • the transverse mass of the system of formed by the electron, the muon, the jet and  
72 the missing transverse energy of the event.

73 The first two variables provide central information regarding the kinematic properties of the  
74 events. Even more, they give another probe to the modellisation of the top  $p_T$ . The  $\Delta\varphi$  allows to  
75 explore correlations between both physical objects and explore spin-related properties. The  $p_Z$   
76 distribution can be used to probe the boost of the complete tW system. The final two invariant  
77 and transverse mass allow to apprehend mass/energy related properties of the whole system  
78 (in addition, these two masses are also measured in the analysis of the ATLAS collaboration,  
79 unlike the previous four, that they cannot consider).

80 The measurement is performed over a fiducial region enriched in the signal process. The results  
81 obtained are compared with the predictions of Powheg using both the DS and DR schemes and  
82 of aMC@NLO using the DR scheme.

83 The note is structured as follows. Section 2 gives a summary of the data and Monte Carlo  
84 samples used. The applied triggers, object and event selections are discussed in Sect. 4. The  
85 systematic uncertainties taken into account are described in Sect. 5. The description of the un-  
86 folding of the detector-level distributions to the particle-level distributions is shown in Sect. 6.  
87 The results are discussed in Sect. 7. Conclusions and a summary of the results are given in  
88 Sect. 8. In addition, an alternative approach to this analysis is explained in the appendix A.

## 89 2 Data and Monte Carlo samples

90 *Editor: Sergio Sánchez Cruz*

91 The analysis reported in this note is performed using data sets collected during the full LHC  
92 2016 run at 13 TeV. The data and simulated samples use to estimate the nominal signal accep-  
93 tance and background yields are fully documented in [19]. For the estimation of the systematic  
94 uncertainties related to  $t\bar{t}$  modeling several samples have been added in order to increase the  
95 amount of simulated events available. Those can be seen in Table 1.

## 96 3 Object selection and event reweighting

97 *Editor: Víctor Rodríguez*

98 In this section all the procedures for the processing of data and Monte Carlo samples are de-  
99 picted. Object and trigger selection, as well as event reweighting coincide with the ones used  
100 in the inclusive analysis [16, 19]. Small modifications regarding the text after comments from  
101 reviewers have been made.

### 102 3.1 Trigger selection

103 In the CMS design, the real time selection of events is achieved in two physical steps, namely  
104 the fast Level-1 Trigger and the High-Level Trigger (HLT) operating on longer timescales. The  
105 Level-1 trigger is built of mostly hardware level information of the detectors while the HLT se-  
106 lection is implemented as a sequence of reconstruction and filter steps of increasing complexity.  
107 High Level inclusive triggers designed to select events in the dilepton channel have been con-  
108 sidered, namely

Sample	Events
/TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8	29,084,029
/TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8 (_ext1-v1)	29,915,551
/TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8 (_ext1-v2)	90,763,410
/TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8 (_ext1-v1)	59,033,604
/TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8 (_ext2-v1)	97,436,211
/TT_TuneCUETP8M2T4_13TeV-powheg-fsrdown-pythia8	29,716,580
/TT_TuneCUETP8M2T4_13TeV-powheg-fsrdown-pythia8 (_ext1-v1)	29,590,326
/TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8	29,777,488
/TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8 (_ext1-v1)	29,453,411
/TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8 (_ext1-v2)	93,386,680
/TT_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8	29,117,820
/TT_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8 (_ext1-v1)	29,046,156
/TTTo2L2Nu_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8/	9,693,881
/TT_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8	29,689,380
/TT_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8 (_ext1-v1)	29,169,226
/TTTo2L2Nu_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8	9,649,574
/TT_TuneCUETP8M2T4up_13TeV-powheg-pythia8	29,310,620
/TT_TuneCUETP8M2T4down_13TeV-powheg-pythia8	28,354,188
/TT_TuneCUETP8M2T4_mtop1665_13TeV-powheg-pythia8	19,380,254
/TT_TuneCUETP8M2T4_mtop1695_13TeV-powheg-pythia8 (_backup)	29,173,030
/TT_TuneCUETP8M2T4_mtop1695_13TeV-powheg-pythia8 (_ext1-v1)	9,954,200
/TT_TuneCUETP8M2T4_mtop1695_13TeV-powheg-pythia8 (_ext2-v1)	19,415,360
/TT_TuneCUETP8M2T4_mtop1715_13TeV-powheg-pythia8	19,578,812
/TT_TuneCUETP8M2T4_mtop1735_13TeV-powheg-pythia8	19,419,050
/TT_TuneCUETP8M2T4_mtop1755_13TeV-powheg-pythia8	29,459,232
/TT_TuneCUETP8M2T4_mtop1755_13TeV-powheg-pythia8 (_ext1-v1)	9,930,052
/TT_TuneCUETP8M2T4_mtop1755_13TeV-powheg-pythia8 (_ext2-v1)	19,995,376
/TT_TuneCUETP8M2T4_mtop1785_13TeV-powheg-pythia8	16,377,176
/TT_TuneCUETP8M2T4_erdON_13TeV-powheg-pythia8	29,943,330
/TT_TuneCUETP8M2T4_erdON_13TeV-powheg-pythia8 (_ext1-v1)	29,938,880
/TTTo2L2Nu_TuneCUETP8M2T4_erdON_13TeV-powheg-pythia8	9,537,400
/TT_TuneCUETP8M2T4_QCDbasedCRTune_erdON_13TeV-powheg-pythia8	29,983,790
/TT_TuneCUETP8M2T4_QCDbasedCRTune_erdON_13TeV-powheg-pythia8 (_ext1-v1)	29,636,416
TTTo2L2Nu_TuneCUETP8M2T4_QCDbasedCRTune_erdON_13TeV-powheg-pythia8	9,816,448
/TT_TuneCUETP8M2T4_GluonMoveCRTune_13TeV-powheg-pythia8	59,037,234
/TTTo2L2Nu_TuneCUETP8M2T4_GluonMoveCRTune_13TeV-powheg-pythia8	9,862,990
/TT_TuneCUETP8M2T4_GluonMoveCRTune_erdON_13TeV-powheg-pythia8	56,168,970
/TTJets_TuneCUETP8M2T4_13TeV-amcatnloFXFX-pythia8	43,561,608
/TT_TuneEE5C_13TeV-powheg-herwigpp	29,412,687
/TT_TuneEE5C_13TeV-powheg-herwigpp (_ext2-v1)	19,762,915

Table 1: Simulated  $t\bar{t}$  background samples used in this analysis for the estimation of some systematic uncertainties, together with the number of events processed.

(\*/RunII Summer16 MiniAODv2-PUMoriond17\_80X\_mcRun2\_asymptotic\_2016\_TrancheIV\_v6-v1)

109 For MC and data (Runs B to G):

- 110 •  $e^\pm \mu^\mp$ : HLT\_Ele27\_WPTight\_Gsf\_v\* OR HLT\_IsoMu24\_v\* OR  
 111     HLT\_IsoTkMu24\_v\* OR HLT\_Mu8\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL\_v\*  
 112     OR HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL\_v\*
- 113 •  $e^+ e^-$ : HLT\_Ele23\_Ele12\_CaloIdL\_TrackIdL\_IsoVL\_DZ\_v\* OR HLT\_Ele27\_WPTight\_Gsf\_v\*
- 114 •  $\mu^+ \mu^-$ : HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_v\* OR  
 115     HLT\_Mu17\_TrkIsoVVL\_TkMu8\_TrkIsoVVL\_v\* OR HLT\_IsoMu24\_v\* OR HLT\_IsoTkMu24\_v\*

116 For data (Run H):

- $e^\pm \mu^\mp$ : HLT\_Ele27\_WPTight\_Gsf\_v\* OR HLT\_IsoMu24\_v\* OR HLT\_IsoTkMu24\_v\* OR HLT\_Mu8\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL\_DZ\_v\* OR HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL\_DZ\_v\*
- $e^+ e^-$ : same as above, HLT\_Ele23\_Ele12\_CaloIdL\_TrackIdL\_IsoVL\_DZ\_v\* OR HLT\_Ele27\_WPTight\_Gsf\_v\*
- $\mu^+ \mu^-$ : HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ\_v\* OR HLT\_Mu17\_TrkIsoVVL\_TkMu8\_TrkIsoVVL\_DZ\_v\* OR HLT\_IsoMu24\_v\* OR HLT\_IsoTkMu24\_v\*

## 3.2 Reconstruction of signal and background events

The tW dilepton final state is characterized by the presence of a high- $p_T$  isolated lepton pair (electron and muon) associated with missing transverse energy  $\cancel{E}_T$  and 1 b-quark jet. The reconstruction of the different objects is based on the Particle-Flow (PF) algorithm [20].

### 3.2.1 Identification of electrons

The selection criteria for electron candidates are the following.

- $p_T > 20$  GeV and  $|\eta| < 2.4$ ;
- veto of the transition region  $1.4442 < |\eta_{SubCluster}| < 1.5660$ ;
- cut based tight ID (SPRING16\_25ns\_v1), as detailed in Ref. [21], for a high purity sample.;
- relative electron isolation (REI): the isolation is computed following the equation below, considering isolated deposits in a cone of  $\Delta R < 0.3$ . Pile up is taken into account by using an effective area correction.

$$REI = \frac{(ChargedHadronIso + NeutralHadronIso + PhotonIso)}{p_T(e)}$$

### 3.2.2 Identification of muons

The selection criteria for muon candidates are the ones from tight muon ID [22] as recommended by the Muon POG. These are specifically the following.

- $p_T > 20$  GeV and  $|\eta| < 2.4$ ;
- is a GlobalMuon and PFMuon;
- number of matchedStations  $> 1$ ;
- number of valid pixel hits  $> 0$ ;
- number of valid hits in the inner tracker  $> 5$ ;
- number of muon hits  $> 0$ ;
- $\chi^2/ndof < 10$  for the global muon fit;
- transverse impact parameter w.r.t. the beam spot  $< 0.2$  cm and  $d_Z < 0.5$  cm, applied on the track from the inner tracker;
- relative muon isolation (RMI)  $< 0.15$ , based on particle flow candidates, defined as

$$RMI = \frac{(ChargedHadronIso + NeutralHadronIso + PhotonIso)}{p_T(\mu)}.$$

Each component is computed using PF isolated deposits in a cone of  $\Delta R < 0.4$  around the muon direction, where charged PF candidates from PU events are re-

148 moved (PF charged subtraction).  $p_T(\mu)$  represents the transverse momentum of the  
 149 muon. Delta- $\beta$  corrections are applied.

### 150 3.2.3 Identification of jets

151 The jets are reconstructed using PF candidates (without considering the charged PF particles  
 152 coming from PU) with the anti- $k_T$  algorithm within an opening angle of  $\Delta R < 0.4$ . The selection  
 153 criteria for jet candidates are the following.

- 154 • L1Fastjet corrections compatible with PFnoPU, Level 2 and Level 3 jet energy correc-  
 155 tions and L2L3Residual corrections for data. The latest set of JECs (Summer16\_23Sep2016V3\_MC  
 156 and Summer16\_23Sep2016V3\_DATA) are used;
- 157 •  $p_T > 30$  GeV and  $|\eta| < 2.4$ ;
- 158 • loose jet identification [23], which corresponds to a fraction of neutral hadronic en-  
 159 ergy  $< 0.99$ , a fraction of neutral electromagnetic energy  $< 0.99$ , and number of con-  
 160 stituents  $> 1$  for jets with  $|\eta| < 2.7$ . On top of that, if  $|\eta| < 2.4$ , we require a fraction  
 161 of charged hadronic energy  $> 0$ , charge multiplicity  $> 0$  and a fraction of charged  
 162 electromagnetic energy  $< 0.99$ ;
- 163 • instead of using PF top projection for lepton removal, we apply a jet lepton cleaning:  
 164 exclusion of jets overlapping with fully selected leptons (electron/muon) used in the  
 165 analysis if  $\Delta R(jet, lepton) < 0.4$ .

166 In addition of the selected jets defined before, “loose” jets are defined as jets that fulfil all the  
 167 requirements described before replacing the  $p_T$  cut by  $20 \text{ GeV} < p_T < 30 \text{ GeV}$ .

### 168 3.2.4 Identification of $b$ -quark jets

169 As  $tW$  signal events have one jet originating from a  $b$  quark, the Combined Secondary Ver-  
 170 tex tagger (CSVv2) is used to improve the background rejection. The medium working point  
 171 (CSVv2M = 0.8484) is used in order to optimize the efficiency and background rejection without  
 172 introducing larger uncertainties.

### 173 3.2.5 Missing transverse momentum

174 This analysis uses corrected (“Type-1”)  $p_T^{\text{miss}}$ . Recommended filters rejecting events where  
 175 known  $p_T^{\text{miss}}$  problems were present have been applied. In order to reduce the instrumental  
 176 noise in the detector, the following filters are applied: primary vertex, CSC beam halo, HBHE  
 177 noise, HBHEiso noise, eebadSC, and Ecal TP. The detailed description of these filters can be  
 178 found in the official page of the MET POG [24].

## 179 3.3 Event reweighting

### 180 3.3.1 Pileup

181 The target pileup distribution for data is generated using the instantaneous luminosity per  
 182 bunch crossing for each luminosity section, stored in the LumiDB database, and the total pp  
 183 inelastic cross section. A Poisson smearing is applied to model statistical fluctuations. The  
 184 source distribution is taken from the PileupInfo collection which stores the true number of  
 185 pileup events mixed with the particular hard interaction process in each MC event. A variation  
 186 of  $\pm 4.6\%$  on the minimum bias cross section is used to estimate the uncertainties due to pileup  
 187 modeling [25].

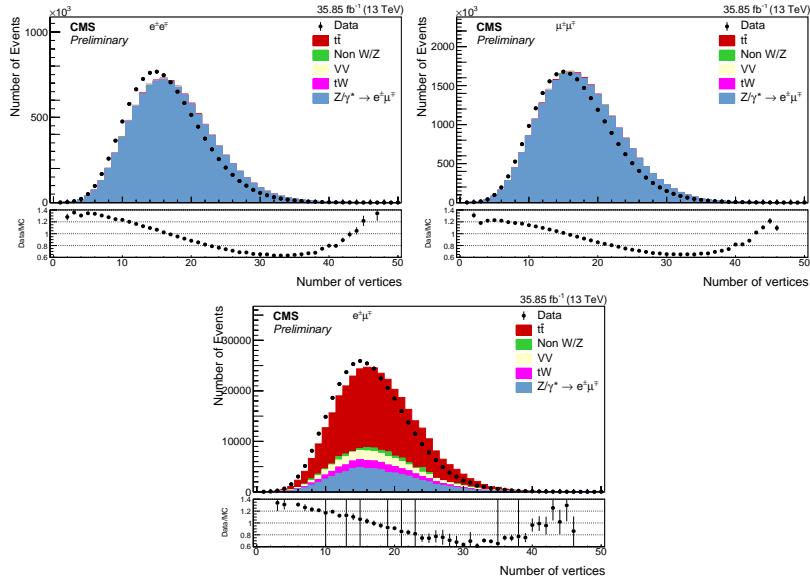


Figure 3: Distribution of the number of good vertices in the event.

188 **3.3.2 Trigger**

189 Scales factors derived in [26] are used, see Fig. 4.

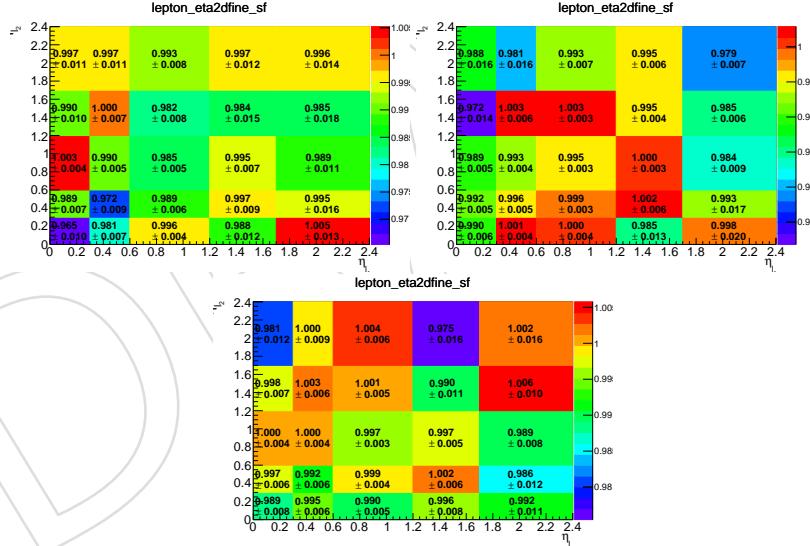


Figure 4: Trigger scale factors.

190 **3.3.3 Lepton ID/Iso**

191 Official muon [27] and electron [28] SFs from POG are used (see Figs. 5 and 6). In particular:

- 192 • Muon tracking efficiency: link.
- 193 • Muon tight ID (B - F): link (histogram MC\_NUM\_tightID\_DEN\_genTracks\_PAR\_pt\_eta / abseta\_pt\_ratio).
- 194 • Muon tight ID (G - H): link (histogram MC\_NUM\_tightID\_DEN\_genTracks\_PAR\_pt\_eta / abseta\_pt\_ratio).
- 195 • Muon tight isolation (B - F): link (histogram tightISO\_TightID\_pt\_eta / abseta\_pt\_ratio).
- 196 • Muon tight isolation (G - H): link (histogram tightISO\_TightID\_pt\_eta / abseta\_pt\_ratio).

- 197     • Electron reconstruction efficiency: link (histogram EGamma\_SF2D).  
 198     • Electron ID/Isolation: link (histogram EGamma\_SF2D).

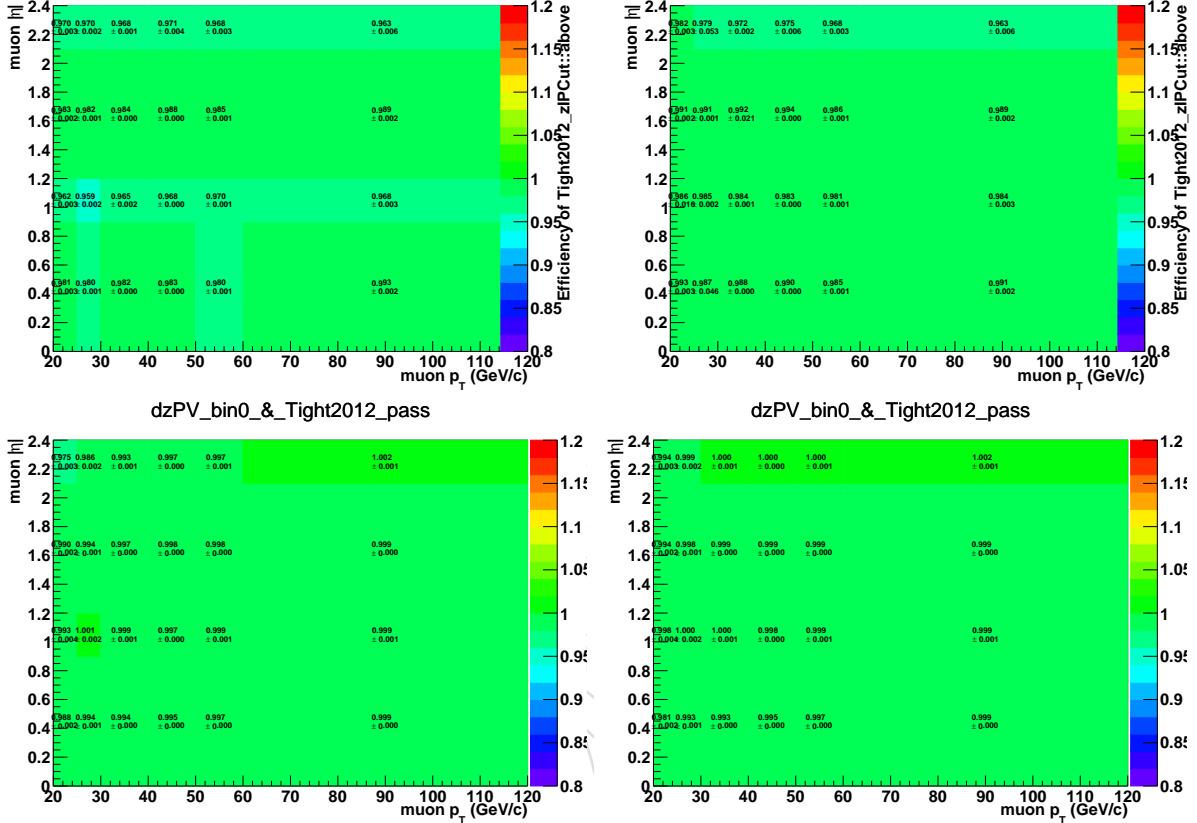


Figure 5: Muon identification (top) and isolation (bottom) SFs for run eras B-F (left) and G-H (right).

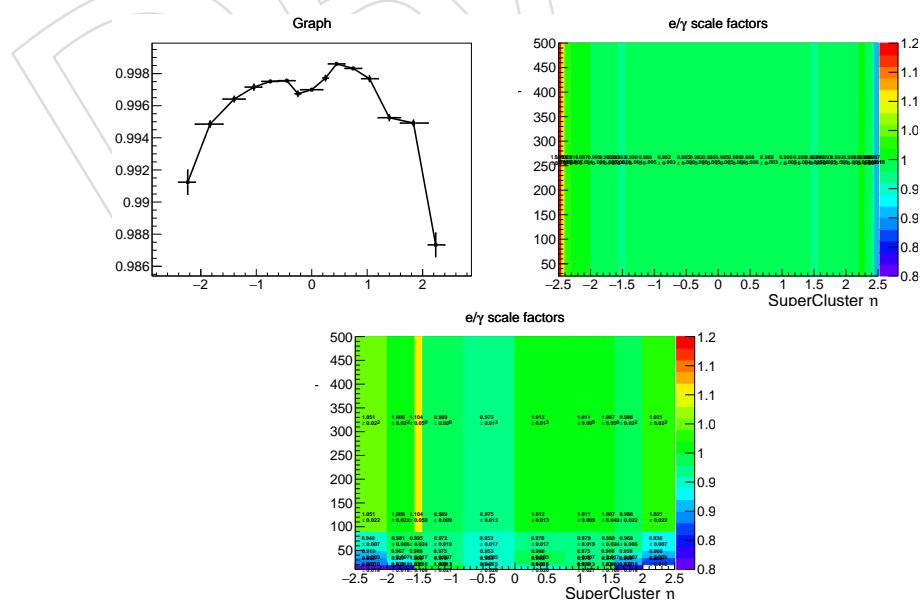


Figure 6: Muon tracker efficiency (left), electron reconstruction efficiency (center), and electron identification and isolation (right) SFs.

199 **3.3.4 B-tagging**

200 As the MC does not reproduce completely the  $b$ -tagging performance in data, the officially rec-  
 201 ommended data-to-MC scale factors [29] are applied:  
 202 [https://twiki.cern.ch/twiki/pub/CMS/BtagRecommendation80XReReco/CSVv2\\_Moriond17\\_B\\_H.csv](https://twiki.cern.ch/twiki/pub/CMS/BtagRecommendation80XReReco/CSVv2_Moriond17_B_H.csv).

204 **4 Event selection**

205 **4.1 Event selection criteria**

206 The selection is based in the requirements given in Ref. [30]. The events are firstly chosen if they  
 207 belong to the  $e^\pm\mu^\mp$  final state, i.e. if the two leading leptons of good quality are an electron and  
 208 a muon. This includes also electrons and muons from a  $\tau$  lepton with a non-hadronic decay.  
 209 We require the leading lepton to have  $p_T > 25$  GeV. In a second step, events are rejected  
 210 if these two leptons are not of opposite charge. To reduce the contamination from low mass  
 211 resonances, the invariant mass of the lepton pair is required to be greater than 20 GeV. On  
 212 top of that, exactly one reconstructed jet is required to be present in the event and it must be  
 213 identified as a  $b$ -jet. Finally, a veto in events containing additional “loose” jets is imposed. This  
 214 veto is the only difference with respect to the 1j1b region defined in [16, 19].

215 Measurements of the differential cross section for top quark pair production have shown that  
 216 the transverse momentum of the top quark is softer than predicted by the POWHEG simula-  
 217 tion [31, 32]. The effect of this mismodelling of the  $p_T$  spectrum is corrected by reweighting the  
 218  $t\bar{t}$  simulation to that measured in the data. This is done, as recommended by the TOP PAG,  
 219 using a function that comes from a fitting between the generated distributions of the top quark  
 220  $p_T$  with the latests theoretical approximations and data unfolded to parton level. This results in  
 221 another scale factor to apply to each MC event that comes from a  $t\bar{t}$  process. The overall extra  
 222 weight is  $\sqrt{SF(t)SF(\bar{t})}$ , where  $SF(t)$  is defined as in eq. 1.

$$SF(t) = e^{0.0615 - 0.0005 \cdot p_T(t)} \quad (1)$$

223 Apart from these requirements for our signal region, a control region is defined based in the  
 224 signal region, but inverting the “loose” jet veto. Figures 7 and 8 (Figs. 9 and 10 show the  
 225 same but with a more “descriptive” binning), data and Monte Carlo comparisons are shown  
 226 for events in the defined signal and control regions, respectively. Globally, both the signal and  
 227 control region show good agreement between data and the predictions given by the Monte  
 228 Carlo models.

229 **4.2 Yields**

230 The expected yields after applying the object and event selection of the signal region described  
 231 above are shown in Tab. 2, whereas the corresponding for the control region are in Tab. 3.

Source	$t\bar{t}$	Non W/Z	VV and $t\bar{t}$ V	Drell-Yan	Total background	tW (signal)	Total	Data
Amount	14215	79	157	363	14814	4160	18974	18025

Table 2: Total number of events observed in data and the number of signal and background events expected from simulation at 13 TeV in the signal region.

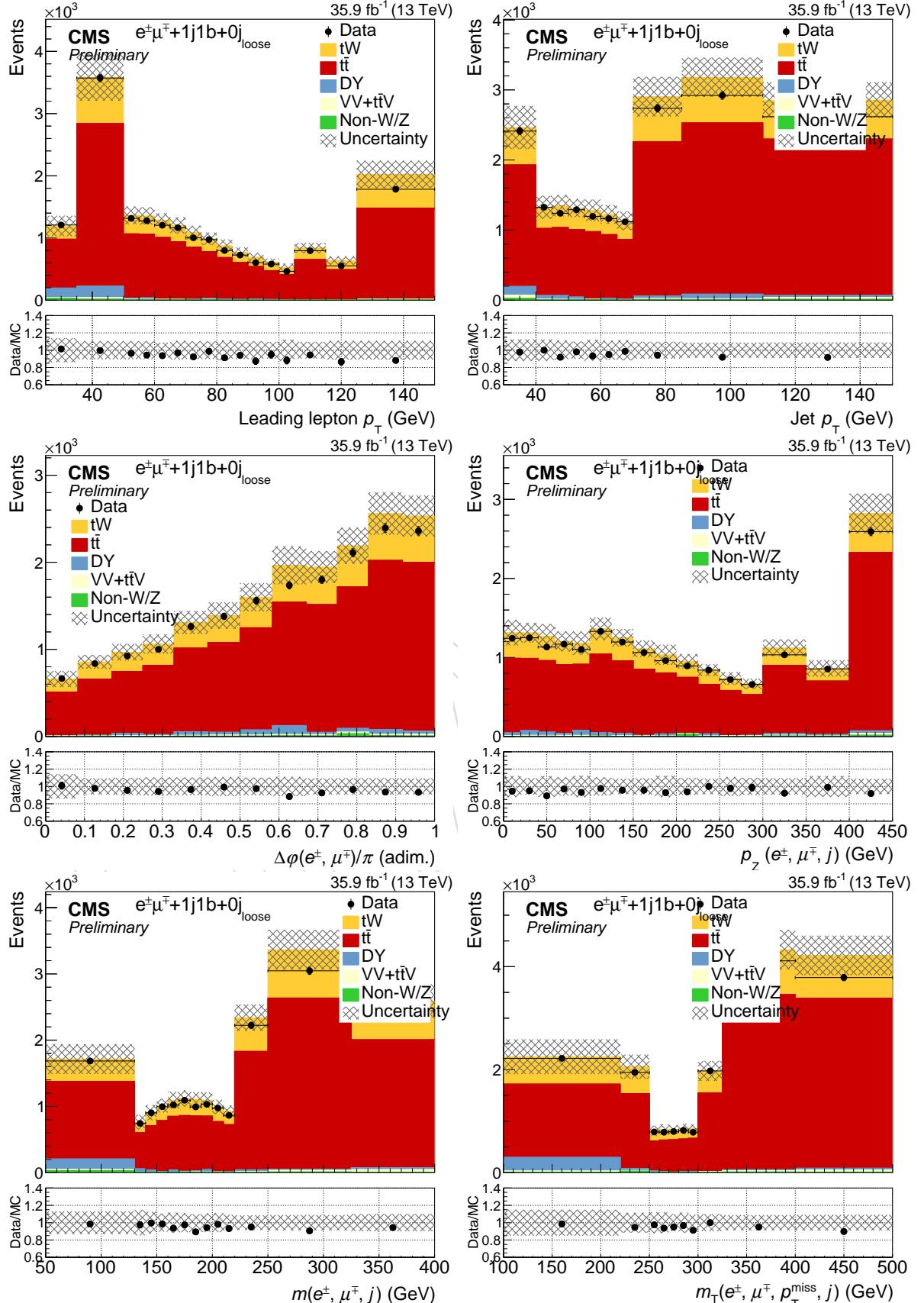


Figure 7: Distribution of the events for various observables inside the signal region. The last bin contains overflow events. The binning displayed is the chosen for the folded space / detector level results (see in the unfolding part how it is optimised).

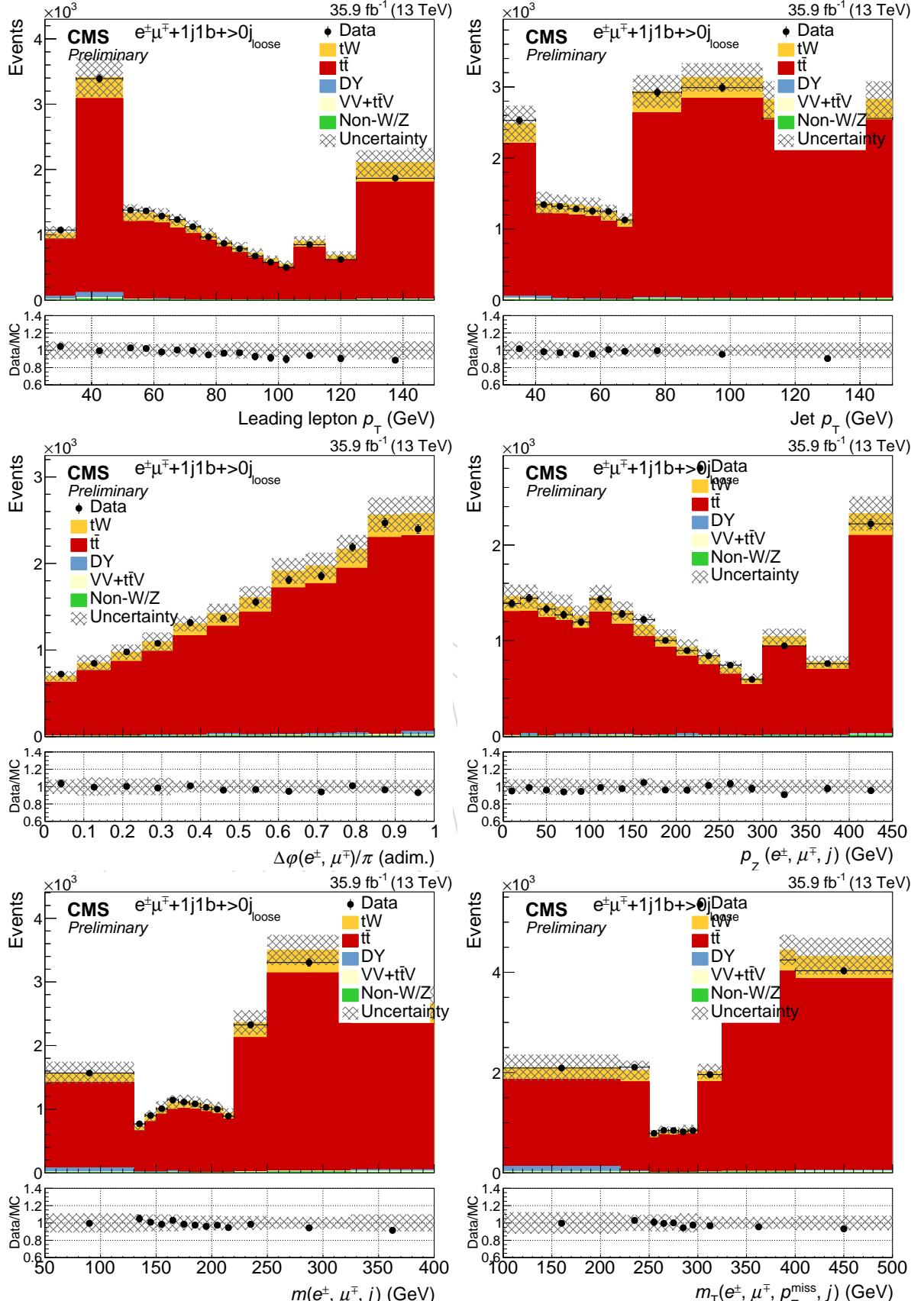


Figure 8: Distribution of the events for various observables inside the control region. The last bin contains overflow events. The binning displayed is the chosen for the folded space / detector level results (see in the unfolding part how it is optimised).

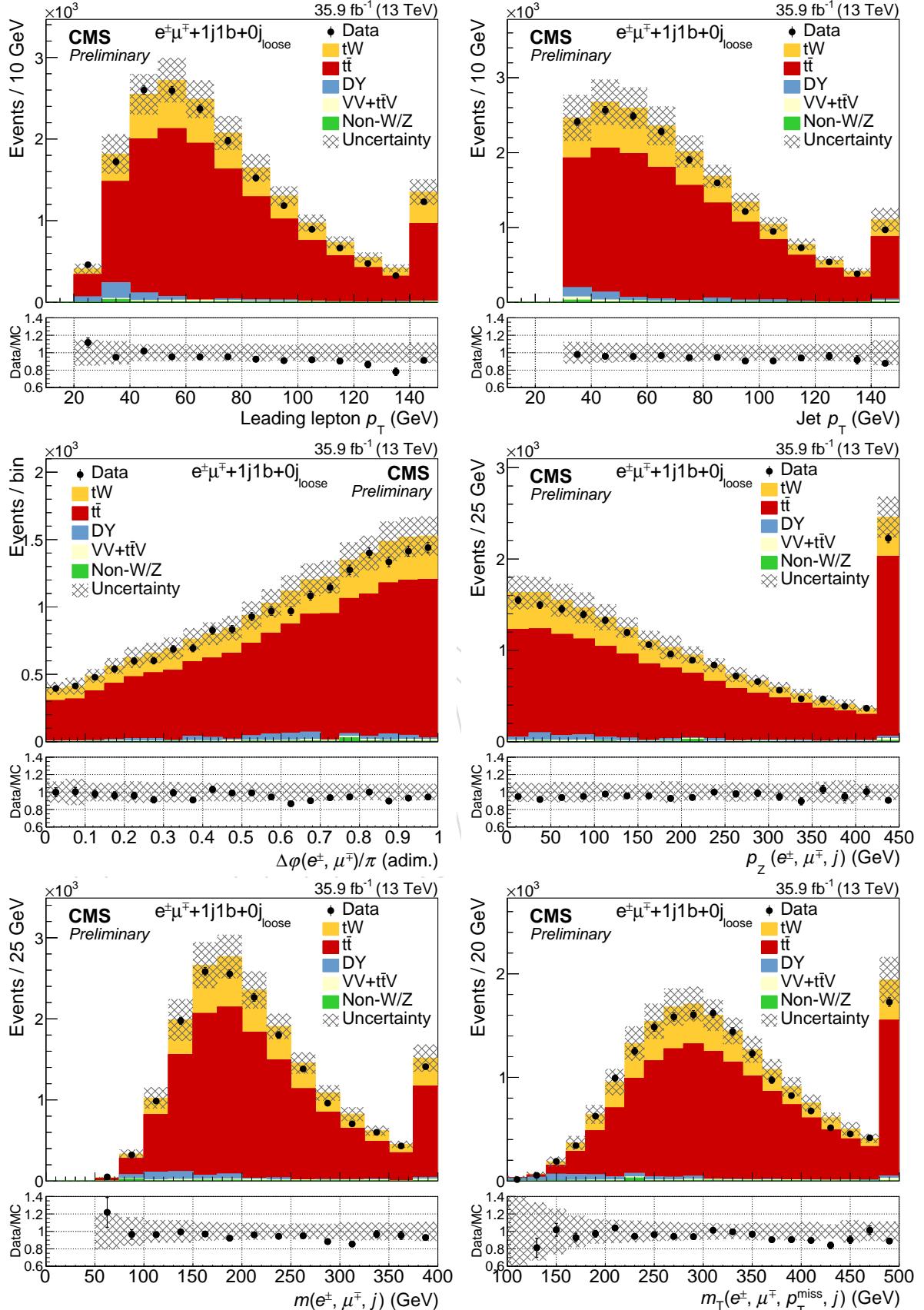


Figure 9: Distribution of the events for various observables inside the signal region. The last bin contains overflow events.

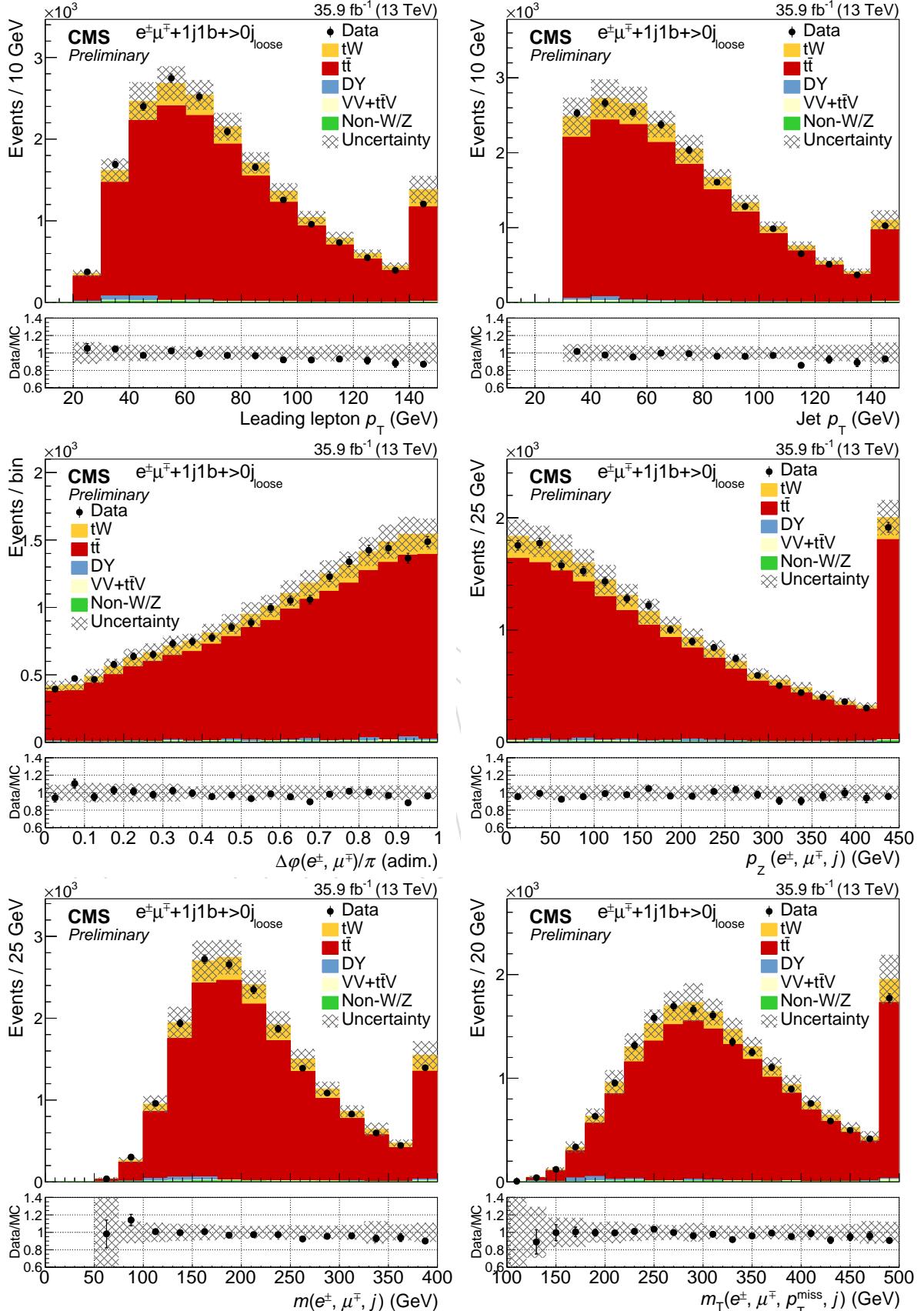


Figure 10: Distribution of the events for various observables inside the control region. The last bin contains overflow events.

Source	Non-W/Z	VV and $t\bar{t}$ V	Drell-Yan	$t\bar{t}$	Total background	tW (signal)	Total	Data
Amount	74	86	125	16875	17160	1989	19149	18581

Table 3: Total number of events observed in data and the number of signal and background events expected from simulation at 13 TeV in the control region.

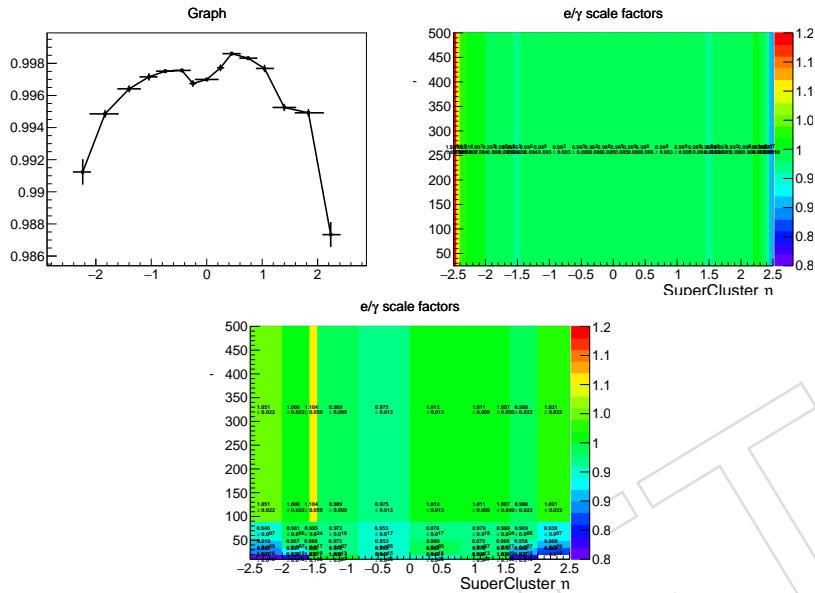


Figure 11: Muon tracker efficiency (left), electron reconstruction efficiency (center), and electron identification and isolation (right) SFs.

## 232 5 Systematic Uncertainties

233 *Editor: Sergio Sánchez*

234 The measurement of the differential tW cross-section is affected by systematic uncertainties  
235 that originate from both detector effects and theoretical assumptions. Due to the dominating  
236 presence of  $t\bar{t}$  events in the targeted signal region, the impact of these uncertainty sources  
237 into the measured cross-section is dominantly due to the effect of these uncertainties in the  
238  $t\bar{t}$  estimation, both in its normalization and shape.

239 There is a smaller contribution to the signal uncertainty of both experimental and theoretical  
240 sources. These affect to the evaluation of the response matrix used to extrapolate the measured  
241 distribution at reconstructed level to the distribution in particle level.

242 Each source of systematic uncertainty is assessed individually by suitable variations of the  
243 MC simulations or by variations of parameter values in the analysis within their estimated  
244 uncertainties. In order to handle properly correlations of the uncertainties between signal and  
245 background, the complete analysis procedure is done with the varied MC. The comparison  
246 between the nominal final result and varied distribution is then calculated for both variations  
247 (up and down) and the average of both differences is taken as the final uncertainty. This way  
248 of taking the asymmetric value was chosen once a study of the differences with the variations  
249 up and down and the average differences for all uncertainties was done. This showed that  
250 these three “distributions” had more or less the same shape and reasonably the same values in  
251 general.

252 In those cases where there are not two different variations (i.e. “up” and “down”), the only

253 variation is symmetrised. Then, the whole procedure is done and when systematic uncer-  
 254 tainties are calculated, the average as commented in the previous paragraph between the two  
 255 variations is taken).

## 256 5.1 Experimental uncertainties

- 257 • The uncertainties on the dilepton trigger (see [26]) and lepton identification efficien-  
 258 cies (see [22, 28]) in simulation are estimated by varying data-to-simulation scale  
 259 factors by their uncertainties. For muons, additional uncertainties of 1% for ID, 0.5%  
 260 for Isolation and 0.5% for the single muon triggers are added in quadrature [27].
- 261 • The uncertainty due to the limited knowledge of the jet energy scale (JES) is usually  
 262 determined by changes implemented in jet energy in bins of  $p_T$  and  $\eta$ , typically by a  
 263 few percent [33].
- 264 • The uncertainty due to the limited accuracy of the jet energy resolution (JER) is de-  
 265 termined by changing the JER correction from (1.109, 1.138, 1.114, 1.123, 1.084, 1.082,  
 266 1.140, 1.067, 1.177) for  $\eta$  bins (0.0 - 0.5, 0.5 - 0.8, 0.8 - 1.1, 1.1 - 1.3, 1.3 - 1.7, 1.7 - 1.9,  
 267 1.9 - 2.1, 2.1 - 2.3, 2.3 - 2.5) within their uncertainties (0.008, 0.013, 0.013, 0.024, 0.011,  
 268 0.035, 0.047, 0.053, 0.041), respectively [34, 35].
- 269 • The uncertainty due to the b-tagging efficiency is determined by varying the b-  
 270 tagging scale factors derived at 13 TeV according to its uncertainty [29].
- 271 • The effect of pileup events is evaluated by weighting the simulation to the minimum  
 272 bias cross section determined in data. The pileup model estimates the mean number  
 273 of additional pp interactions to be about 20 events. These estimates are based on the  
 274 total inelastic proton-proton cross section, which is determined to be 69.2 mb [25] at  
 275 13 TeV. The systematic uncertainty is estimated by varying the cross section within  
 276 its uncertainty of  $\pm 4.6\%$  and propagating the variation to the estimates of the mean  
 277 number of additional pp interactions.
- 278 • For ttV, VV, DY and nonW background contributions, a conservative normalization  
 279 uncertainty of  $\pm 50\%$  is assumed as done in [16, 19].
- 280 • For  $t\bar{t}$  a conservative 4% normalization uncertainty is assumed to take into account  
 281 uncertainties in the measurement of the inclusive  $t\bar{t}$  cross-section, as measured in [36].  
 282 There exists a more precise measurement of the  $t\bar{t}$  cross-section [37], but the selec-  
 283 tion that is performed overlaps with the current measurement. For that reason, the  
 284 precision measured in [36] is assumed.
- 285 • The uncertainty on the integrated luminosity is currently estimated to be 2.5% [38].

## 286 5.2 Modeling Uncertainties

287 The modeling of the  $t\bar{t}$  and tW events by the MC simulation is an important ingredient in  
 288 this measurement. The impact of theoretical assumptions in the modeling is determined by  
 289 repeating the analysis and replacing the standard POWHEG +PYTHIA  $t\bar{t}$  or tW simulation by  
 290 dedicated simulation samples with altered parameters.

291 The uncertainty on modeling of the hard-production process (ME scale) is assessed through  
 292 changes in the renormalization and factorization scales in the POWHEG samples by factors of  
 293 two and 0.5 relative to their common nominal value, which is set in POWHEG to  $Q^2 = m_t^2 + p_{T,t}^2$ ,  
 294 where  $p_{T,t}^2$  denotes the square transverse momentum of the top quark in the  $t\bar{t}$  zero-momentum  
 295 frame. This variation is performed separately for  $t\bar{t}$  and tW events.

296 In order to take into account parton-shower (PS) uncertainties, different effects are studied:

- Underlying event: non-perturbative QCD effects are taken into account by tuning PYTHIA to measurements of the underlying event [39, 40]. The parameters of the tune are varied up and down within their uncertainties in the generation of  $t\bar{t}$  events.
- ME/PS matching: the uncertainty in the combination of the matrix element calculation with the parton showers estimated from the variation of the POWHEG parameter  $h_{damp} = 1.58^{+0.66}_{-0.59} \cdot m_t$  in  $t\bar{t}$  events. This parameter regulates the damping of real emissions in the NLO calculation when matching to the PS [40].
- Initial (final) state radiation scale: the PS scale used for the simulation of the initial (final) state radiation is varied up and down by a factor of two ( $\sqrt{2}$ ) in  $t\bar{t}$  and tW events. These contribution is handled simultaneously for  $t\bar{t}$  and tW events, and separately for initial and final state radiation.
- Colour reconnection: the effect of multiple-parton interactions and the parameterization of color reconnection have been studied in [39] and are varied accordingly. In addition, we use a simulation with activated color reconnection of resonant decays. The uncertainties that arise from ambiguities in modeling color reconnection effects are estimated by comparing the default model in PYTHIA with two alternative models of color reconnection, a model with string formation beyond leading color [41] and a model in that the gluons can be moved to another string [42]. All models are tuned to measurements of the underlying event [39, 40]. In addition, the effects of color reconnection on the top decay products can be turned on in PYTHIA by enabling early resonance decays (turned off in the nominal sample).
- During run 1 and run 2 it was found that the  $p_T$  spectra of top quarks in data was significantly softer than those predicted by the various MC simulations based on either LO or NLO ME interfaced with parton showers. The effect was confirmed (albeit to a lesser degree) by ATLAS. Although predictions at NNLO+NNLL [43] and approx. NNNLO in pQCD [44] provide a much improved description, a residual discrepancy remains. The origin of this discrepancy is not clear (some ideas point to higher order QCD corrections, or EWK corrections, for example). As the properties of the top quark affect us indirectly through the b-jet produced after its decay, as well as through the physical objects product of the decay of the W boson, it is recommended inside the TOP PAG to take into account this discrepancy between expectations and actual data in the analysis. We add as a source of uncertainty the difference between the nominal results and the final results obtained after a reweighting is applied to the  $t\bar{t}$  sample. The weights applied to all events are derived using theory predictions at NNLO with Powheg+Pythia8 and comparing with data from Run 2 [45–48].
- The uncertainty in the mass of the top is added as a source of uncertainty. To take it into account, we use samples of tW and  $t\bar{t}$  that have been obtained with a different mass of the quark (namely  $m_{top} = 169.5$  GeV and  $m_{top} = 175.5$  GeV). As usually the difference consider is that of  $\pm 1$  GeV, we take as uncertainty the difference with the results obtained with the varied contribution of the  $t\bar{t}$  and tW processes but dividing the difference in those samples with respect to the nominal values by 3.
- The uncertainty associated to the missmodeling of the momentum of the top quark is taken as the difference with respect to the uncorrected shapes.

The uncertainty from the choice of PDFs is determined by reweighting the sample of simulated  $t\bar{t}$  events according to the 100 NNPDF3.0 error PDF sets [49].

Additionally, the difference between the two schemes defining the signal, diagram removal (DR) and diagram subtraction (DS) is taken into account as a systematic uncertainty.

## 344 6 Unfolding results to particle level

345 *Editor: Sergio Sánchez*

346 The measurement of the observables is distorted by the response and the acceptance of the  
 347 detector. Unfolding techniques must therefore be used to infer the actual distributions without  
 348 the detector effects in order to be compared with other theoretical predictions.

349 In this work, we unfold the measured distributions at reconstruction level to particle level.  
 350 Particle level is defined by the obtained particles after the generation of the strong process and  
 351 after the parton shower simulations. With this set-up, the definition of the fiducial region is  
 352 closer to the event selection that is applied in the analysis, yielding to smaller extrapolation,  
 353 reducing the uncertainties associated to it.

### 354 6.1 Definition of particle level

355 Objects are defined at particle level as particles produced by event generators after the radia-  
 356 tion and with a lifetime longer than 30 ps. Prompt leptons may have emitted photons, and this  
 357 effect is taken into account by clustering all close-by photons to leptons, i.e., they are “dressed”.  
 358 Similarly, the observables related to quarks and gluons are jets that are obtained when cluster-  
 359 ing the hadrons and non-prompt leptons produced in the parton showers. All the details on  
 360 the particle level definition are given in [50].

### 361 6.2 Unfolding

362 The procedure for obtaining the particle level distribution from the measured (smeared) dis-  
 363 tribution can be written as follows. Distributions are typically discretized as histograms, and  
 364 then the particle level and smeared distribution can be related as

$$y_i = A_{ij}x_j, \quad (2)$$

365 where  $y_i$  and  $x_j$  are the histograms of the smeared and particle level distributions, respectively.  
 366  $A_{ij}$  is the response matrix, that gives the probability for an event generated in the bin  $j$  of the  
 367 generator level distribution to be measured in the bin  $i$  of the smeared distribution.

368 In this analysis, the response matrix  $\{A_{ij}\}$  is estimated using simulated  $tW$  events in the signal  
 369 region defined in Sect. 4. The reconstructed events are corrected accordingly with the corre-  
 370 sponding scale factors and calibrations described in Sect. 4.

371 This problem could be uniquely solved by inverting the  $\{A_{ij}\}$  matrix. However, the response  
 372 matrix can have large dimension or be highly non-diagonal, so its diagonalization can be nu-  
 373 merically complex. Several approaches can be followed to tackle this problem. In this work,  
 374 the `TUnfold` [51] implementation of Tikhonov regularization is used.

375 The problem in equation (2) is solved by building an estimator, given by the maximization of  
 376 the likelihood

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \quad (3)$$

$$= (\mathbf{y} - A\mathbf{x})^T V_y (\mathbf{y} - A\mathbf{x}) \quad (4)$$

$$+ \tau^2 (\mathbf{x} - f_b \mathbf{x}_0)^T (L^T L) (\mathbf{x} - f_b \mathbf{x}_0) \quad (5)$$

$$+ \lambda (Y - e^T \mathbf{x}), \quad (6)$$

377 where  $Y = \sum_i y_i$  and  $e_j = \sum_i A_{ij}$ .  $V_y$  is the covariance matrix of  $\mathbf{y}$ , in this case, a diagonal matrix  
 378 with the square statistical uncertainty of each bin in the diagonal. Other uncertainty sources  
 379 are externalized from the unfolding procedure and are not included in  $V_y$ .

380 The first term in (3),  $\mathcal{L}_1$ , corresponds to the likelihood function that is minimized to obtain the  
 381 maximum likelihood estimator (MLE). The other two terms  $\mathcal{L}_2$  and  $\mathcal{L}_3$  introduce a bias with  
 382 respect to the unbiased estimator given by the least squares. This bias is introduced to reduce  
 383 the variance with respect to the least square minimization estimator. In particular, the term  
 384  $\mathcal{L}_2$  describes the regularization which penalizes large curvatures in  $\mathbf{y}$ , associated to numerical  
 385 instabilities, and is modeled by a regularization parameter  $\tau$ , that is chosen to regulate the bias  
 386 introduced by the method. The term  $\mathcal{L}_3$  is an area constrain, that ensures the normalization of  
 387 the smeared and generated distributions agree.

### 388 6.3 Choice of binning

389 The choice of the discretization that is used to define the equation (2) can be relevant for the nu-  
 390 mercial convergence of the method. A non-optimal binning can yield to numerical instabilities  
 391 that will result in large uncertainties. Even if the dominating uncertainties in this analysis arise  
 392 from the estimation of the background, a special care is taken to select the optimal binning.

393 First, we begin with the choice of the number of bins that we are going to use for each distri-  
 394 bution. To set that number, we have plotted with a reasonable bin width (taking care to not  
 395 let small bins exist where there appears to be few signal events) all distributions for 4, 5, 6,  
 396 7, 8, 9 and 10 bins using the expected (Asimov dataset) values. Afterwards, for each one of  
 397 the variables, the number of bins is chosen by checking the evolution of the uncertainties and  
 398 finding approximately the best compromise between a better description (more resolution, a  
 399 higher number of bins) and a relatively low uncertainties. The plots of the distributions for  
 400 those quoted bins are drawn in the Appendix B. The number of bins chosen for the different  
 401 distributions are the following.

402 **Leading lepton  $p_T$**  8 bins.

403 **Jet  $p_T$**  5 bins.

404  $\Delta\varphi(e^\pm, \mu^\mp)$  6 bins.

405  $p_Z(e^\pm, \mu^\mp, \text{jet})$  8 bins.

406  $m(e^\pm, \mu^\mp, \text{jet})$  6 bins.

407  $m_T(e^\pm, \mu^\mp, \text{jet}, p_T^{\text{miss}})$  5 bins.

408 Afterwards, we checked out that the response matrices are reasonably diagonal to reduce the  
 409 chances of having numerical problems in the unfolding step. To do so, we calculate the purity  
 410 of the reconstruction (or folded space or detector level) bins and the stability of the generation  
 411 (or unfolded space or particle level) bins. These quantities are defined as

$$s_i := \frac{\sum_{j=1}^{N_{\text{bins}}^{\text{fol.}}} n_{ij}}{n_i} \quad p_j := \frac{\sum_{i=1}^{N_{\text{bins}}^{\text{unf.}}} n_{ij}}{n_j^R}, \quad (7)$$

412 where  $n_{ij}$  are the number of reconstructed events whose value of the measured variable fall in  
 413 the folded-space bin  $j$  that had the simulated value of the variable in the unfolded-space bin  $i$ ,  
 414  $n_i$  are the number of events whose simulated value of the variable fell in the unfolded space bin  
 415  $i$  and  $n_j^R$  is the amount of simulated events in the folded space bin  $j$ . Both purity and stability,  
 416 when maximised simultaneously in all the spectre of the distribution, enhance the diagonality  
 417 of the response matrices, easing the numerical problems of the unfolding and reducing the  
 418 need of regularization.

419 Distributions of the stability and puritiy are shown in Fig. 12.

#### 420 6.4 Choice of regularization parameter

421 The choice of the regularization parameter,  $\tau$ , is performed by performing the L-curve method [52]  
 422 using for this check the Asimov dataset. In this method, a scan for values of the  $\tau$  parame-  
 423 ter is performed. For each value of  $\tau$ , the likelihood in (3) is maximized, and the values of  
 424  $L_x = \log \mathcal{L}_1$  and  $L_y = \log \mathcal{L}_2/\tau^2$  in the maximum are evaluated. The function obtained corre-  
 425 sponds to the L-curve, and gives the interplay between high bias/less variance (high  $L_x$ , low  
 426  $L_y$ ) and low bias/high variance (low  $L_x$ , high  $L_y$ ). The curves are shown in Fig. 13, and show a  
 427 kink, which is taken as the optimal point for the unfolding argument algorithm.

428 The value of the best  $\tau$  is also shown in Fig. 13 and found to be very small. In particular, the  
 429 ratio between the unfolded distribution obtained with the best choice of  $\tau$  and fixing  $\tau$  to zero  
 430 has been checked. Those distributions can be seen in Fig. 14 (again obtained with the Asimov  
 431 dataset), and the results are found to have negligible differences.

432 Therefore the analysis does not make use of regularization. This decision is justified by the  
 433 small effect of regularization in the analysis, and also by the fact that the response matrices  
 434 are very diagonal, as shown in 20. The latter point is also checked by obtaining the condition  
 435 number of the matrices, which is shown in Tab. 4. All of them are close to the unity, which  
 436 means that regularization without bias is possible.

#### 437 6.5 Effect of the area constraint

438 The area constraint included in the likelihood, the term  $\mathcal{L}_3$  in equation (3), is also a type of  
 439 regularization that ensures that the total event yield in the folded and unfolded distributions  
 440 are consistent. The effect of this introduced bias in the likelihood is assessed by comparing the  
 441 unfolded distribution obtained with and without this term in the maximized likelihood while  
 442 using the Asimov dataset results. Those distributions are shown in Fig. 15, showing no dif-  
 443 ference between applying and not applying this term. This, together the results discussed in  
 444 section 6.4 show that the problem in Eq. (2) can be solved in a numerically robust fashion with-  
 445 out introducing regularization terms. Because of this, the analysis therefore does not include  
 446 the bias terms in the unfolding distribution. Therefore the MLE is used.

#### 447 6.6 Unfolding bias estimation

448 When introducing a non-zero regularization parameter, a bias of the obtained unfolded dis-  
 449 tribution with respect to its true value may appear. This can be quantified by performing  
 450 the unfolding procedure by unfolding the true smeared distribution obtained from MC and  
 451 comparing with the generator level distribution given by the same generator. The difference  
 452 between the two is an estimator of the bias introduced.

453 In this analysis, no regularization is applied. Therefore this comparison is performed just as  
 454 a sanity check of the whole procedure for the determination of the response matrix and the

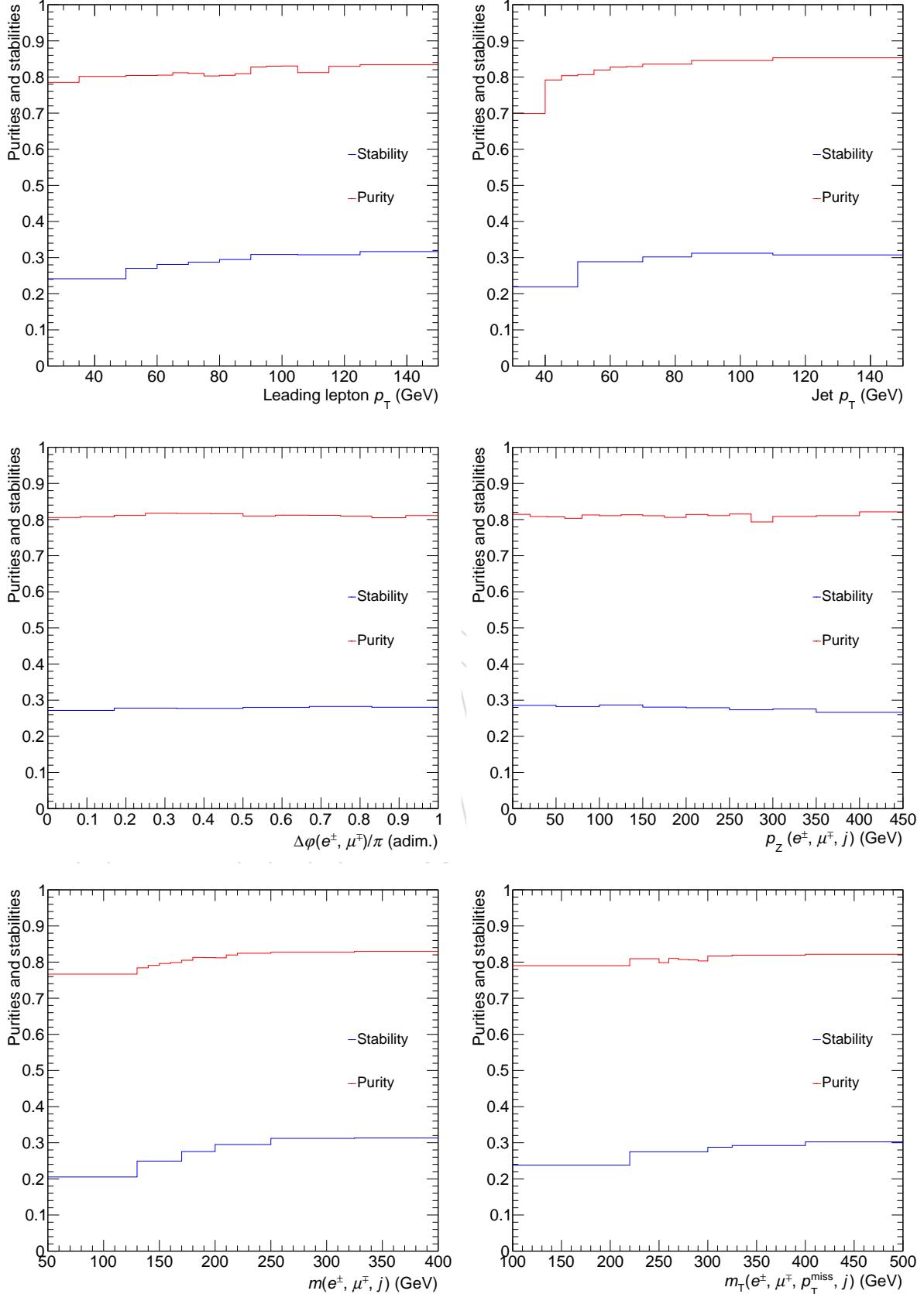


Figure 12: Graphs showing the purity and stability of each bin of the folded and unfolded (respectively) spaces for all the distributions chosen to be unfolded. It can be seen a relatively low stability, that is related with a low reconstruction efficiency for the events in our fiducial region.

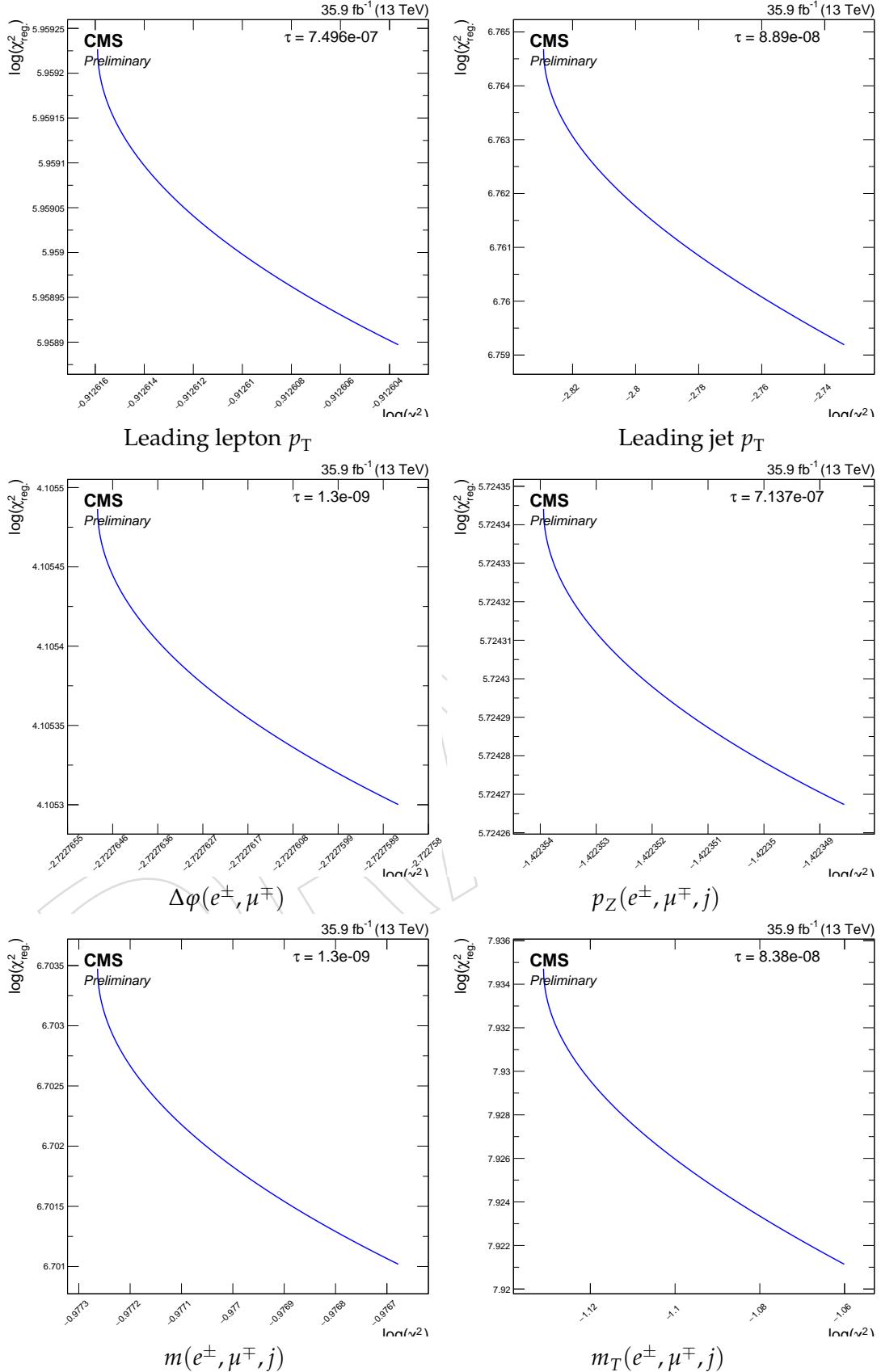


Figure 13: L-Curves of the variables chosen to be unfolded (of its nominal values). The optimal value of the tau parameter (the one corresponding to the point of maximum curvature) is also presented. The search for the  $\tau$  parameter was done with the Asimov dataset results.

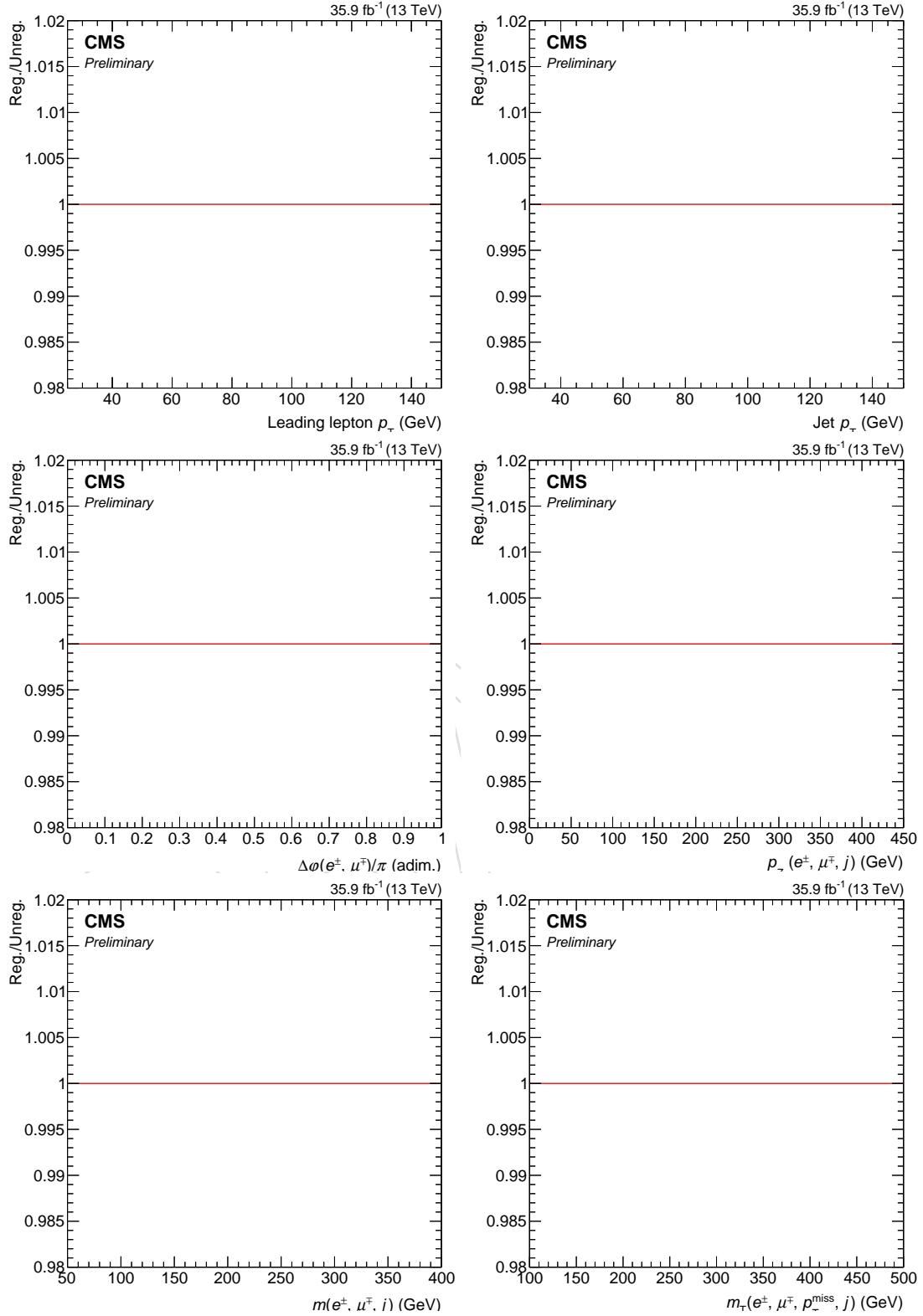


Figure 14: Ratio between the unfolded values of the variables under study when no regularisation is applied and when the  $\tau$  parameter chosen by the L-Curve method is used. Notice the scale of the y axis.

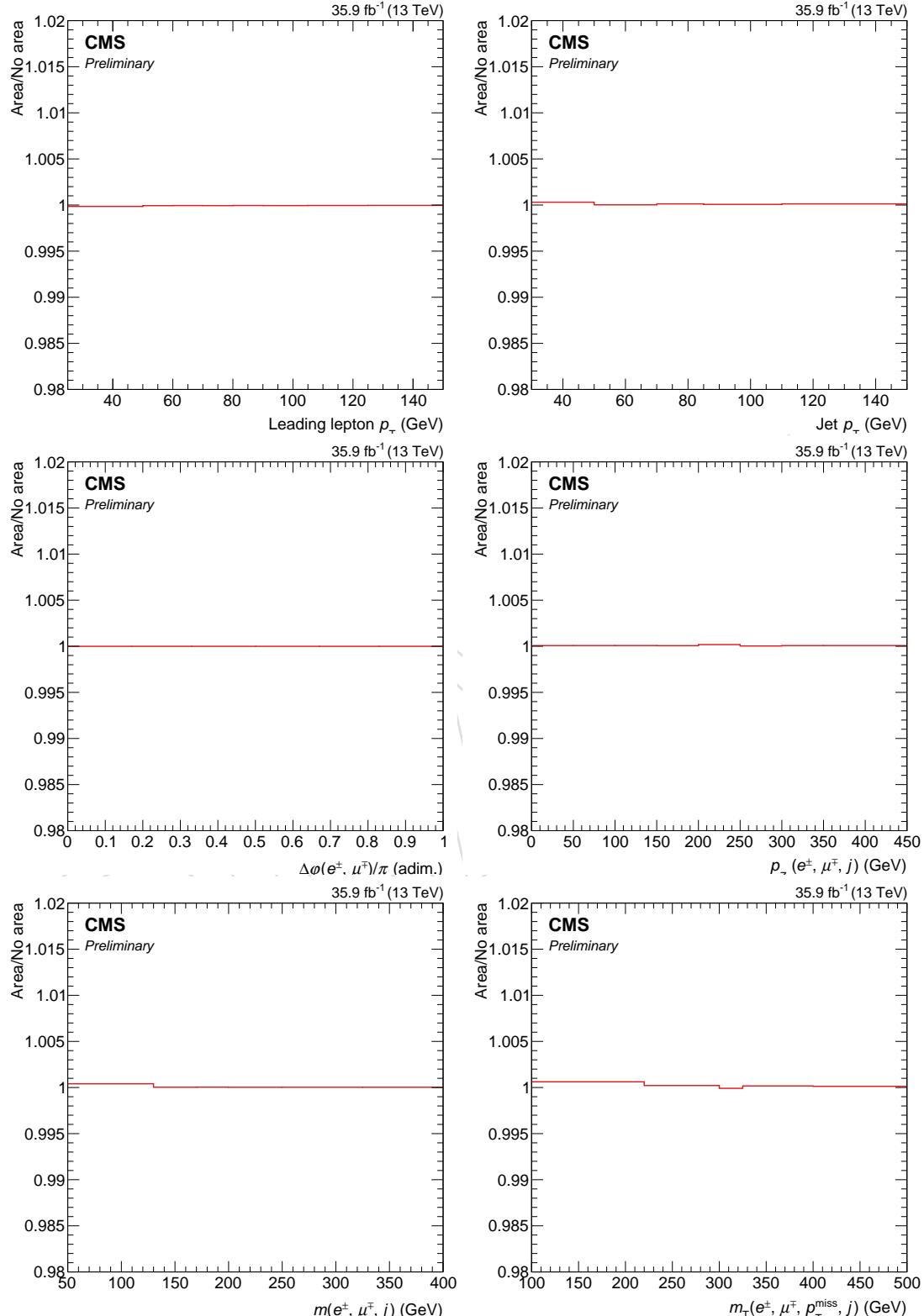


Figure 15: Ratio between the unfolded values of the variables under study when no area constraint is applied and when it is. Notice the scale of the y axis.

Syst./Variable	Leading lepton $p_T$ (GeV)	Jet $p_T$ (GeV)	$\Delta\varphi(e^\pm, \mu^\mp)/\pi$	$p_Z(e^\pm, \mu^\mp, j)$ (GeV)	$m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$ (GeV)	$m(e^\pm, \mu^\mp, j)$ (GeV)
JES up	2.05	4.59	1.04	2.05	6.05	3.33
JES down	2.07	4.27	1.03	2.02	5.31	3.35
JER	2.06	5.08	1.03	2.13	5.66	3.47
Electron eff. up	2.05	4.42	1.03	2.05	5.58	3.34
Electron eff. down	2.06	4.42	1.03	2.05	5.59	3.34
Muon eff. up	2.05	4.42	1.03	2.05	5.58	3.34
Muon eff. down	2.06	4.42	1.03	2.05	5.59	3.34
Trigger eff. up	2.05	4.42	1.03	2.05	5.59	3.34
Trigger eff. down	2.06	4.42	1.03	2.05	5.58	3.34
PU up	2.06	4.45	1.04	2.05	5.74	3.36
PU down	2.05	4.39	1.03	2.05	5.44	3.32
B-tagging up	2.05	4.42	1.03	2.05	5.58	3.34
B-tagging down	2.06	4.43	1.03	2.05	5.59	3.34
Mistagging up	2.05	4.42	1.03	2.05	5.58	3.34
Mistagging down	2.05	4.42	1.03	2.05	5.58	3.34
Top $p_T$ rew. up	2.05	4.42	1.03	2.05	5.58	3.34
Top $p_T$ rew. down	2.05	4.42	1.03	2.05	5.58	3.34
tW DS-DR	2.14	4.32	1.06	1.98	5.59	3.42
FSR up	2.06	4.41	1.05	1.97	5.35	3.34
FSR down	2.09	4.56	1.05	2.07	5.65	3.38
ISR up	2.07	4.16	1.04	1.99	5.56	3.27
ISR down	2.07	4.47	1.02	2.05	5.80	3.34
tW $\mu_R/\mu_F$ up	2.06	4.54	1.09	1.95	5.52	3.23
tW $\mu_R/\mu_F$ down	2.21	4.32	1.05	2.08	5.41	3.46
$m_{top}$ unc. up	2.07	4.39	1.04	2.00	5.48	3.30
$m_{top}$ unc. down	2.08	4.44	1.07	2.00	5.49	3.33
Nominal	2.05	4.42	1.03	2.05	5.58	3.34

Table 4: Condition number for the response matrices of the variables under study. These correspond to the matrices obtained with the nominal samples, as well as the varied samples depending on uncertainty source and, if applicable, variation.

455 unfolding procedure. The comparison between the truth generated and unfolded distribution  
 456 is shown in Fig. 16. The two distributions clearly agree.

## 457 7 Results

458 *Editor: Víctor Rodríguez Bouza*

459 The signal extraction is performed by subtracting to the data the estimated background con-  
 460 tribution in each bin of the measured distribution, obtaining then the results at detector level  
 461 or in the folded space. In Figs. 17, 18, 19 these values (already in terms of the differential cross  
 462 section, in pb), corresponding to the measured distributions in the smeared space, are shown  
 463 for the variables under study. Figs. 21, 22 and 23 the analogous is shown, but at particle level,  
 464 after the unfolding is performed. The response matrices used are depicted in Fig. 20, and they  
 465 are reasonably diagonal. In all cases the main uncertainties that affect the results are the ones  
 466 coming from the JES and JER, that are due to their effect in the background estimation.

467 The numerical results for all the observables are displayed in Table 6, with its corresponding  
 468 uncertainties. The full covariance matrices, including systematic uncertainties, before and after  
 469 the unfolding procedure are shown in Fig. 24 and 25 respectively, whereas in 27 the analogous  
 470 is shown when both normalisations to bin width and fiducial cross section are applied; in 26  
 471 the ones considering only binning normalisation at particle level are presented. In addition, we  
 472 display in the tables 5 a collection of the value of the statistic and the p-values corresponding to  
 473 Pearson's  $\chi^2$  tests of goodness of fit done between the final results and the Powheg DR, Powheg  
 474 DS and aMCatNLO distributions. The information of the table indicates a good agreement  
 475 between the distributions and the results.

476 In addition, in figures 28, 29 and 30 the results are shown, but normalised to the bin width.

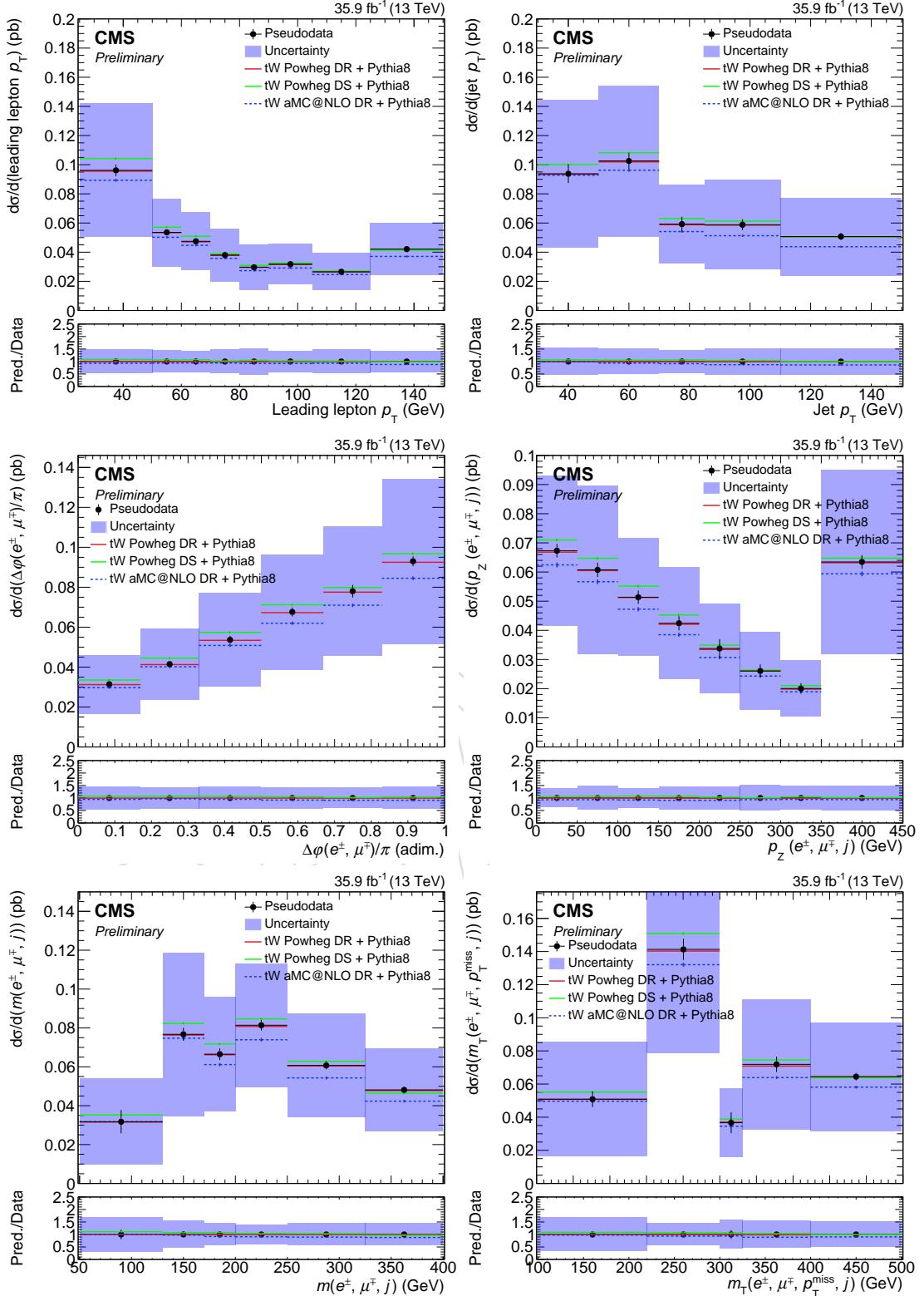


Figure 16: Unfolded space results obtained with the Asimov dataset (pseudodata) as a sanity check for the unfolding procedure. An agreement between the Asimov dataset and the true distribution is observed. The comparison with the generation values of other generator software (MG5\_AMC@NLO) are shown, as well as the comparison with the DS simulation.

Variable/MC sample	POWHEG DR	POWHEG DS	MG5_AMC@NLO
Leading lepton $p_T$	0.11 — 11.728	0.224 — 9.425	0.161 — 10.512
Jet $p_T$	0.732 — 2.022	0.797 — 1.667	0.863 — 1.292
$\Delta\phi(e^\pm, \mu^\mp)/\pi$	0.719 — 2.877	0.77 — 2.543	0.782 — 2.461
$p_Z(e^\pm, \mu^\mp, j)$	0.69 — 4.755	0.699 — 4.679	0.675 — 4.876
$m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$	0.383 — 4.17	0.476 — 3.512	0.441 — 3.749
$m(e^\pm, \mu^\mp, j)$	0.26 — 6.503	0.29 — 6.166	0.396 — 5.169

Table 5: Results (p-value – test statistic value) of the  $\chi^2$  goodness-of-fit tests performed to check the compatibility between data and the Powheg DR, Powheg DS and aMCatNLO models.

Leading lepton $p_T$ (GeV)	[25, 50]	[50, 60]	[60, 70]	[70, 80]	[80, 90]	[90, 105]	[105, 125]	[125, $\infty$ )
$d\sigma/d$ (Leading lepton $p_T$ ) (pb)	0.10	0.04	0.03	0.029	0.019	0.016	0.014	0.021
Uncertainty (pb)	0.05	0.02	0.02	0.018	0.016	0.014	0.013	0.018
Jet $p_T$ (GeV)	[30, 50]	[50, 70]	[70, 85]	[85, 110)	[110, $\infty$ )			
$d\sigma/d$ (Jet $p_T$ ) (pb)	0.09	0.07	0.05	0.03	0.02			
Uncertainty (pb)	0.06	0.05	0.03	0.03	0.03			
$\Delta\phi(e^\pm, \mu^\mp)/\pi$	[0.0, 0.17)	[0.17, 0.33)	[0.33, 0.5)	[0.5, 0.67)	[0.67, 0.83)	[0.83, 1]		
$d\sigma/d$ ( $\Delta\phi(e^\pm, \mu^\mp)/\pi$ ) (pb)	0.028	0.032	0.05	0.05	0.05	0.05	0.06	
Uncertainty (pb)	0.015	0.018	0.02	0.03	0.03	0.03	0.04	
$p_Z(e^\pm, \mu^\mp, j)$ (GeV)	[0, 50)	[50, 100)	[100, 150)	[150, 200)	[200, 250)	[250, 300)	[300, 350)	[350, $\infty$ )
$d\sigma/d$ ( $p_Z(e^\pm, \mu^\mp, j)$ ) (pb)	0.05	0.04	0.05	0.025	0.029	0.026	0.012	0.04
Uncertainty (pb)	0.03	0.03	0.02	0.019	0.016	0.015	0.011	0.03
$m(e^\pm, \mu^\mp, j)$ (GeV)	[50, 130)	[130, 170)	[170, 200)	[200, 250)	[250, 325)	[325, $\infty$ )		
$d\sigma/d$ ( $m(e^\pm, \mu^\mp, j)$ ) (pb)	0.03	0.08	0.04	0.07	0.03	0.04		
Uncertainty (pb)	0.02	0.04	0.03	0.03	0.03	0.03		
$m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$ (GeV)	[100, 220)	[220, 300)	[300, 325)	[325, 400)	[400, $\infty$ )			
$d\sigma/d$ ( $m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$ ) (pb)	0.06	0.10	0.05	0.05	0.02			
Uncertainty (pb)	0.04	0.06	0.03	0.04	0.03			

Table 6: Numerical results of the differential cross sections for all the observables considered, as well as the upper and lower uncertainties for each case.

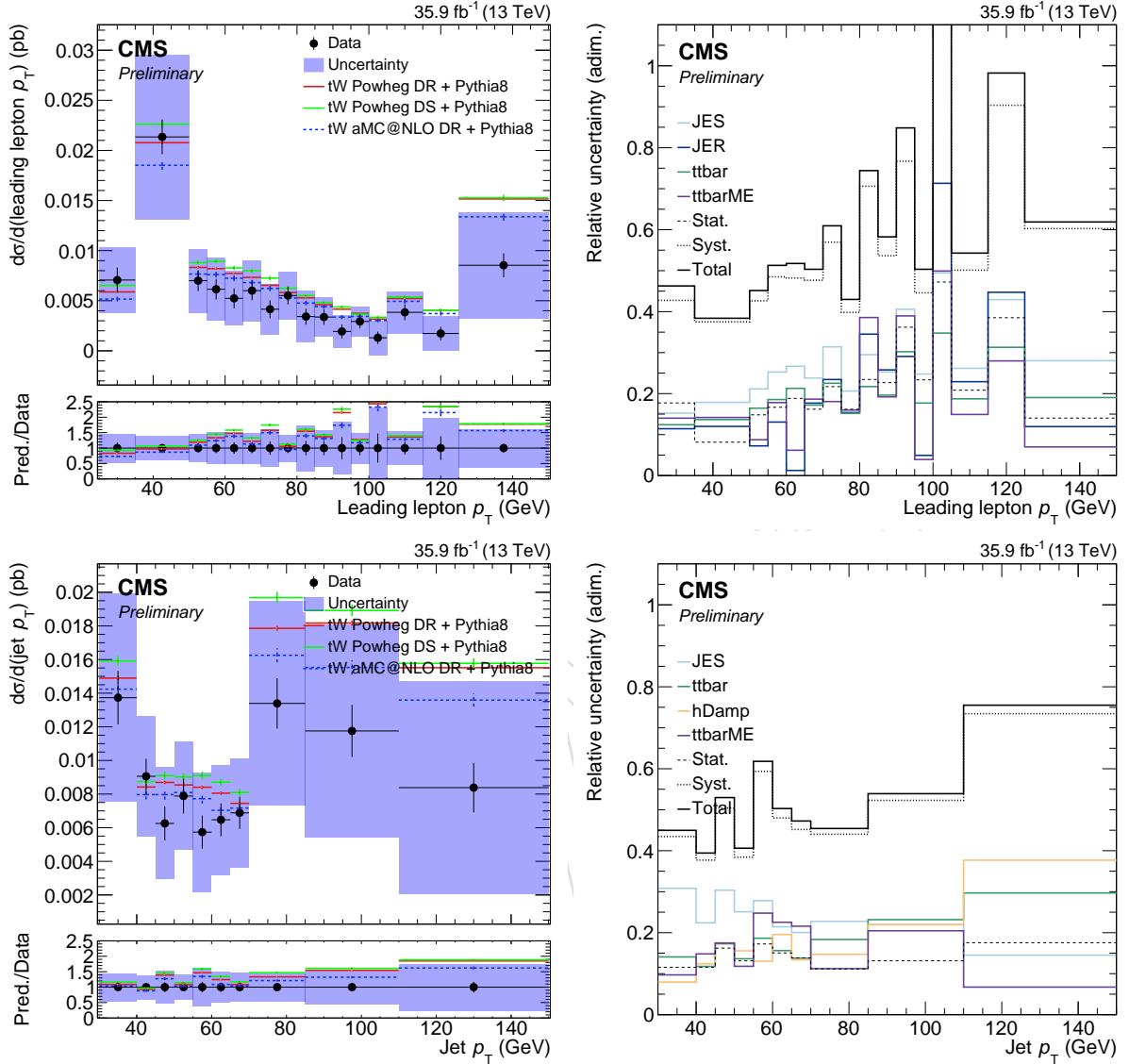


Figure 17: (Left) Differential tW production cross section as a function of the  $p_T$  of the leading lepton and of the jet of the process in the folded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

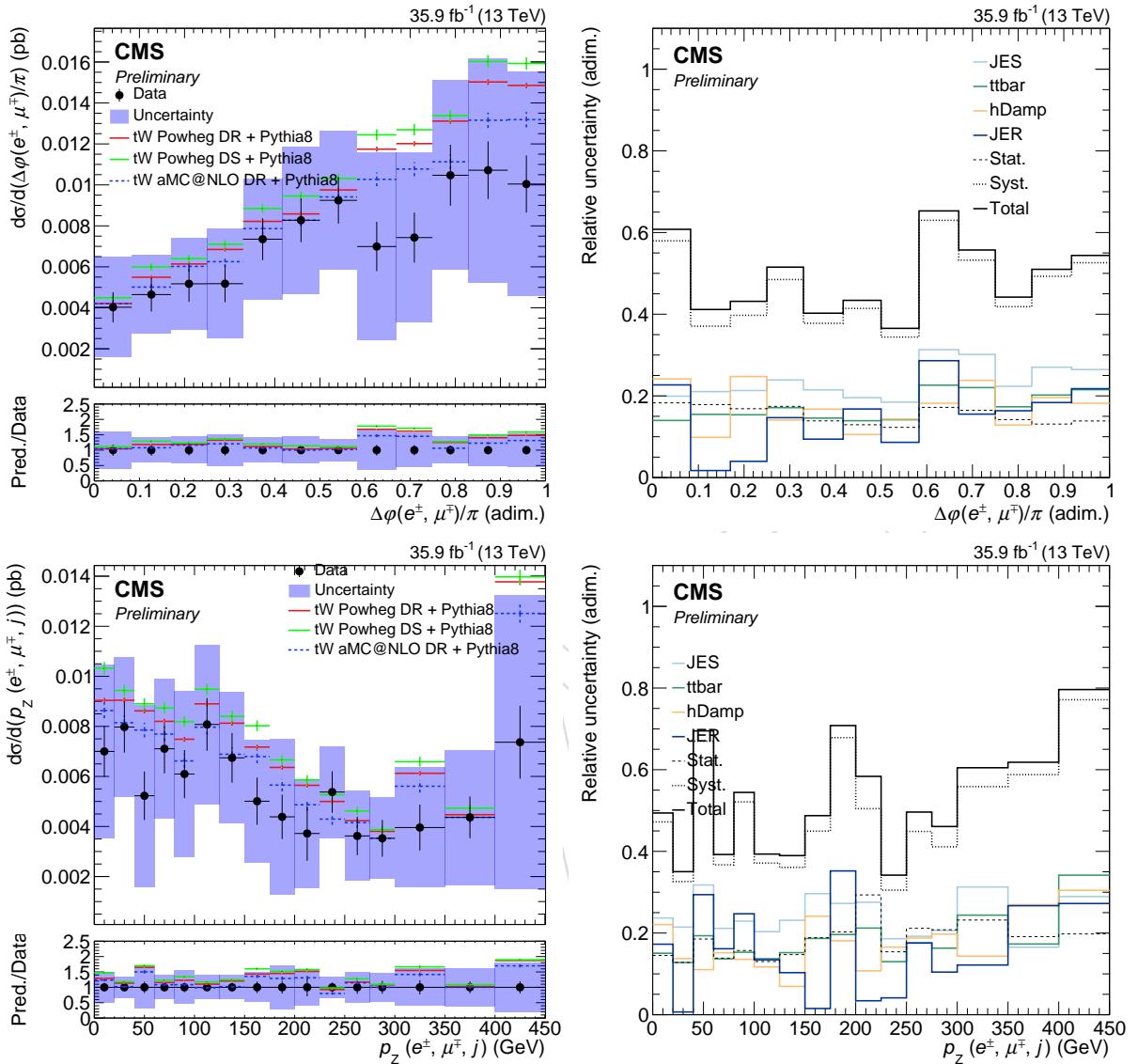


Figure 18: (Left) Differential  $tW$  production cross section as a function of  $\Delta\phi(e^\pm, \mu^\mp)$  and  $p_z(e^\pm, \mu^\mp, j)$  in the folded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

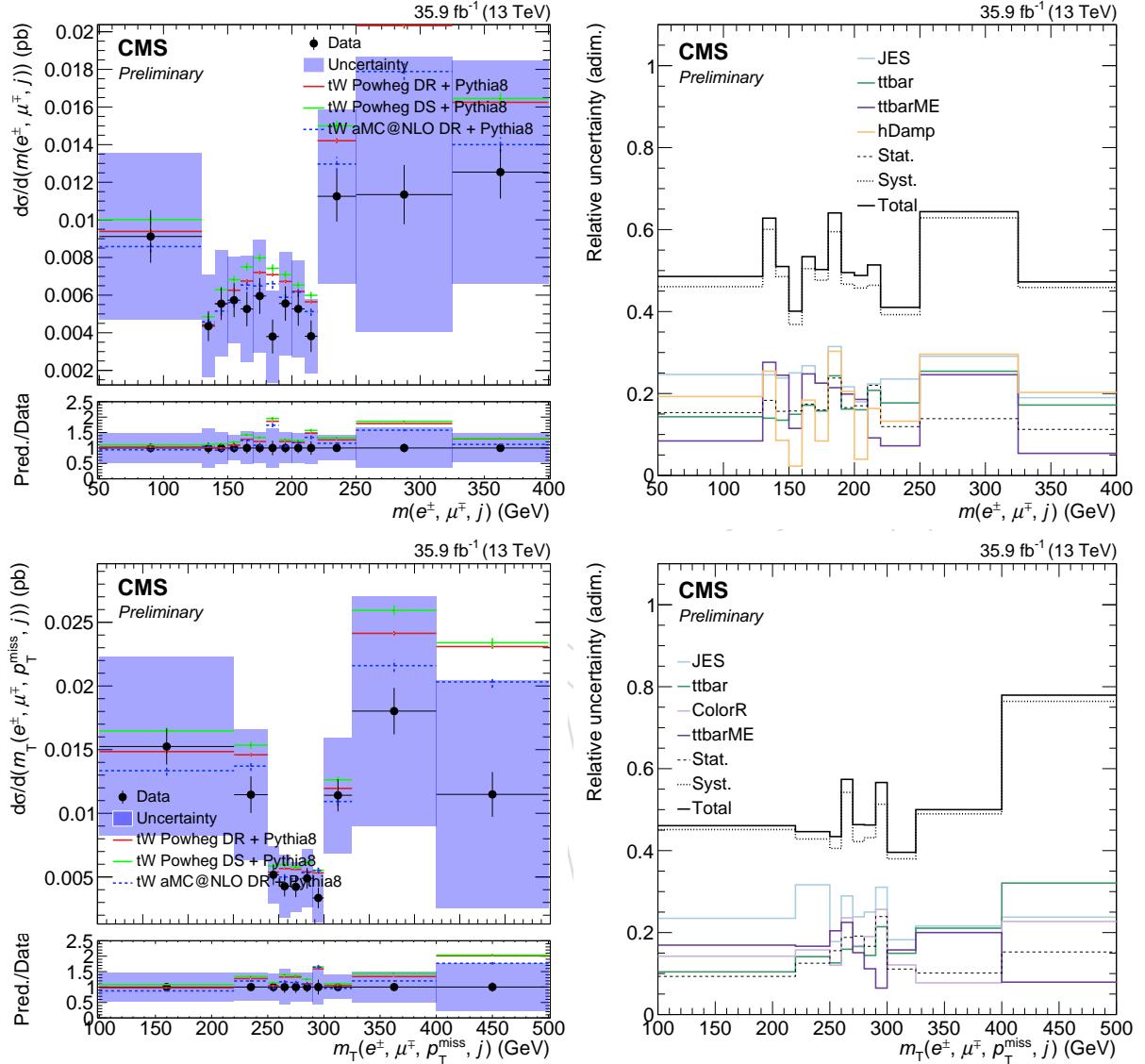


Figure 19: (Left) Differential tW production cross section as a function of  $m(e^\pm, \mu^\mp, j)$  and  $m_T(e^\pm, \mu^\mp, j, \cancel{E}_T)$  in the folded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

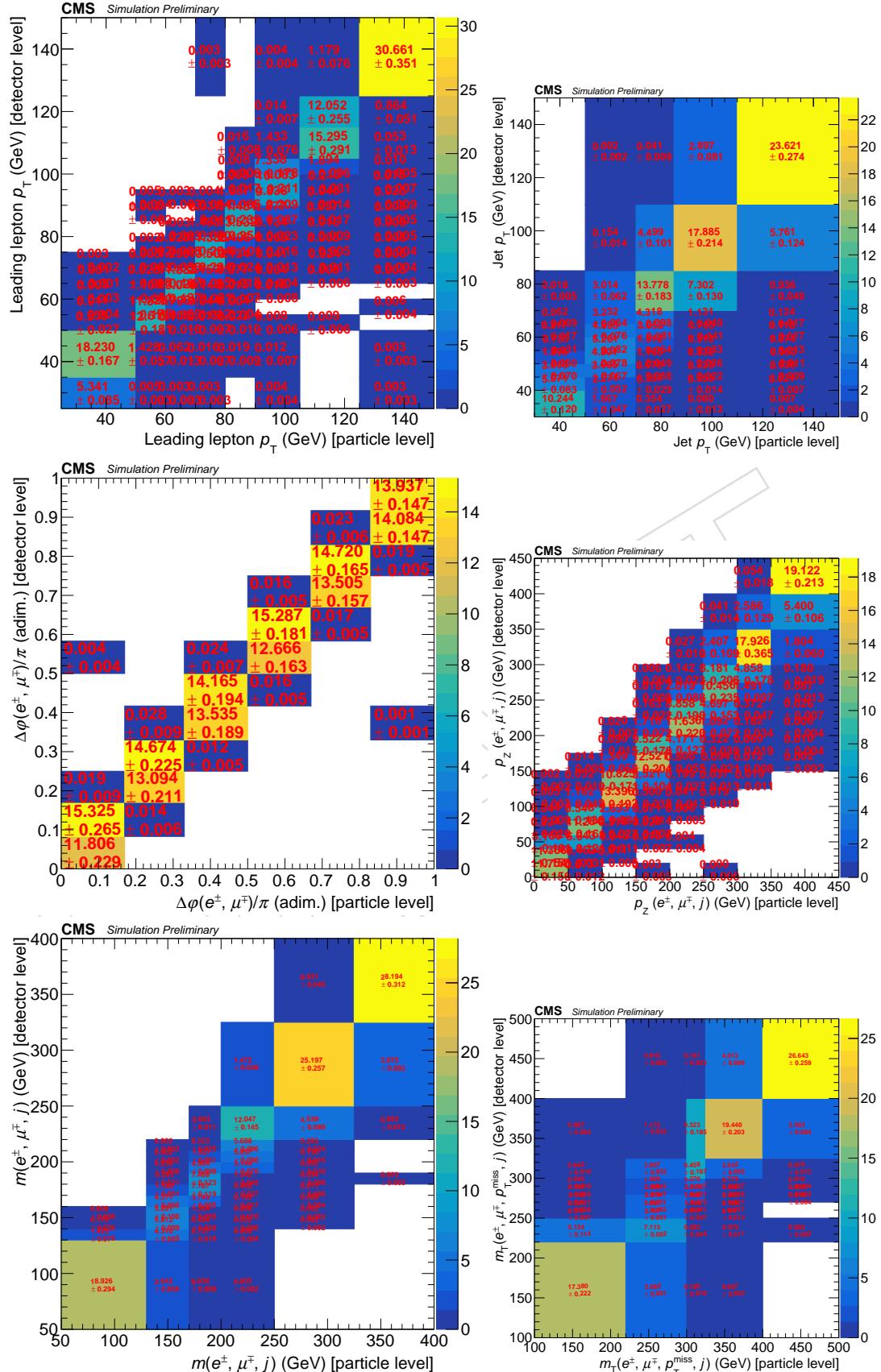


Figure 20: Response matrices of the variables under study: the reconstructed event axis is the folded space axis and the generated event axis, the unfolded. They are scaled to 100 (all contents multiplied by it).

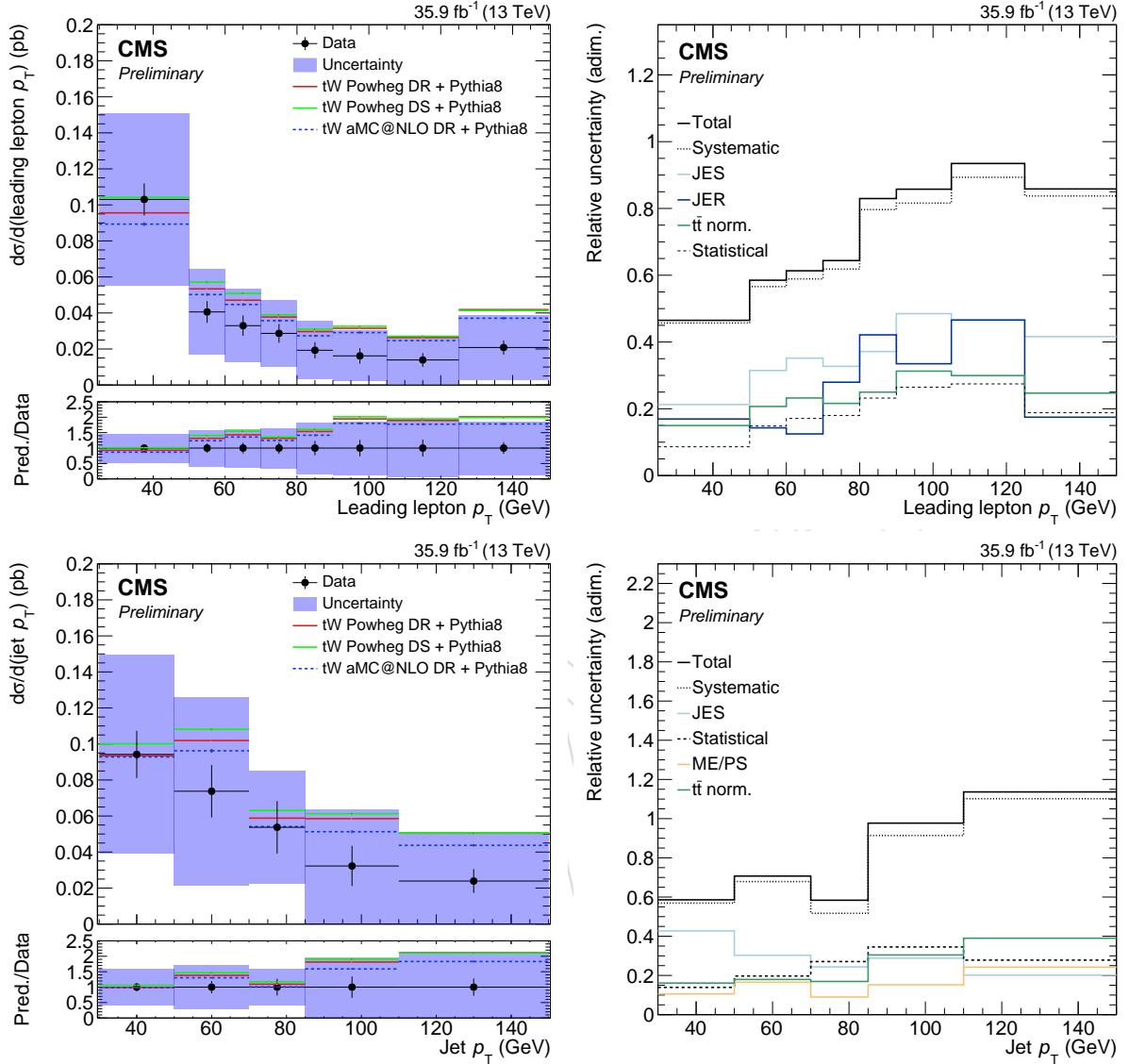


Figure 21: (Left) Differential tW production cross section as a function of the  $p_T$  of the leading lepton and of the jet of the process in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

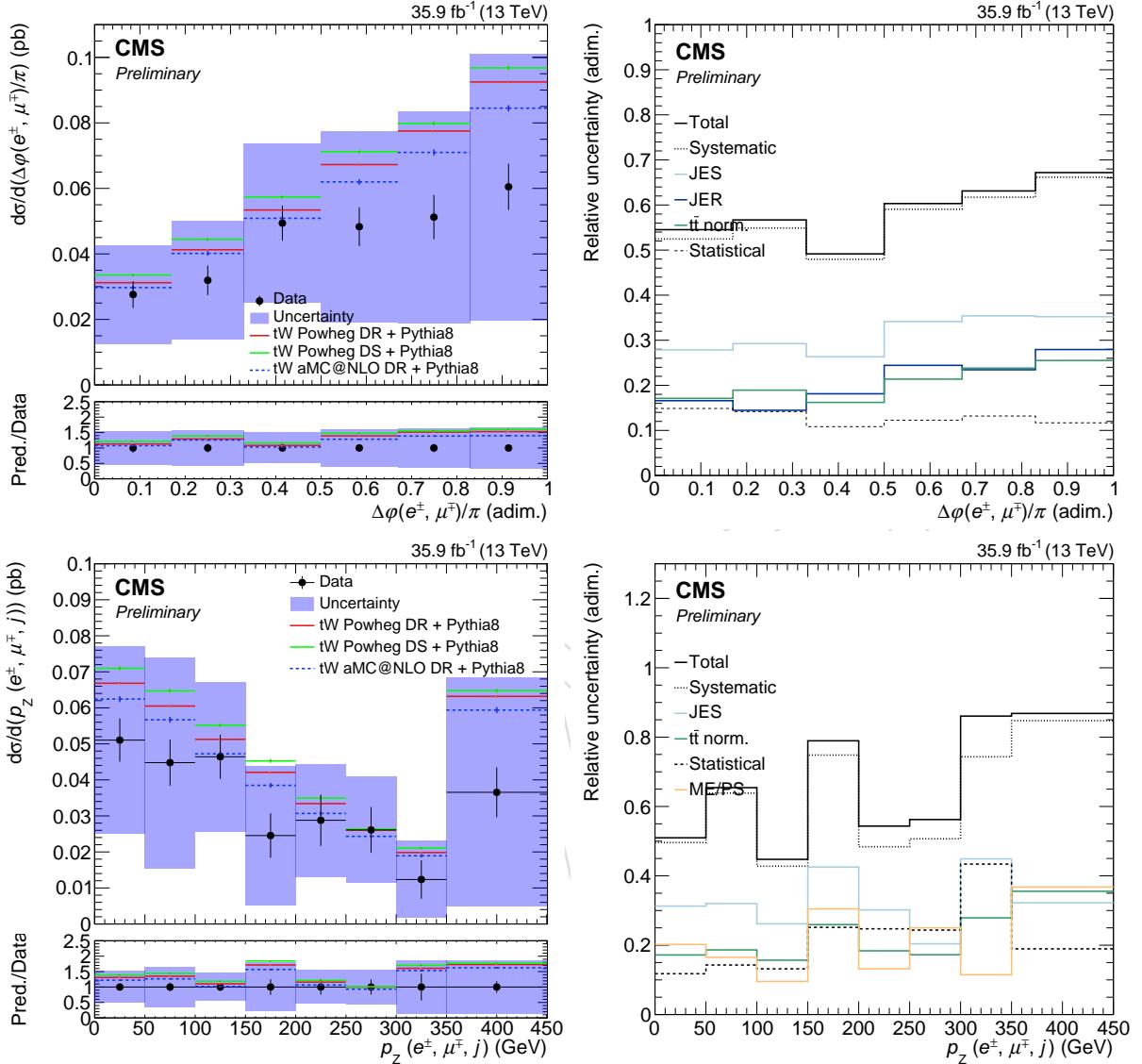


Figure 22: (Left) Differential tW production cross section as a function of  $\Delta\phi(e^\pm, \mu^\mp)$  and  $p_z(e^\pm, \mu^\mp, j)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

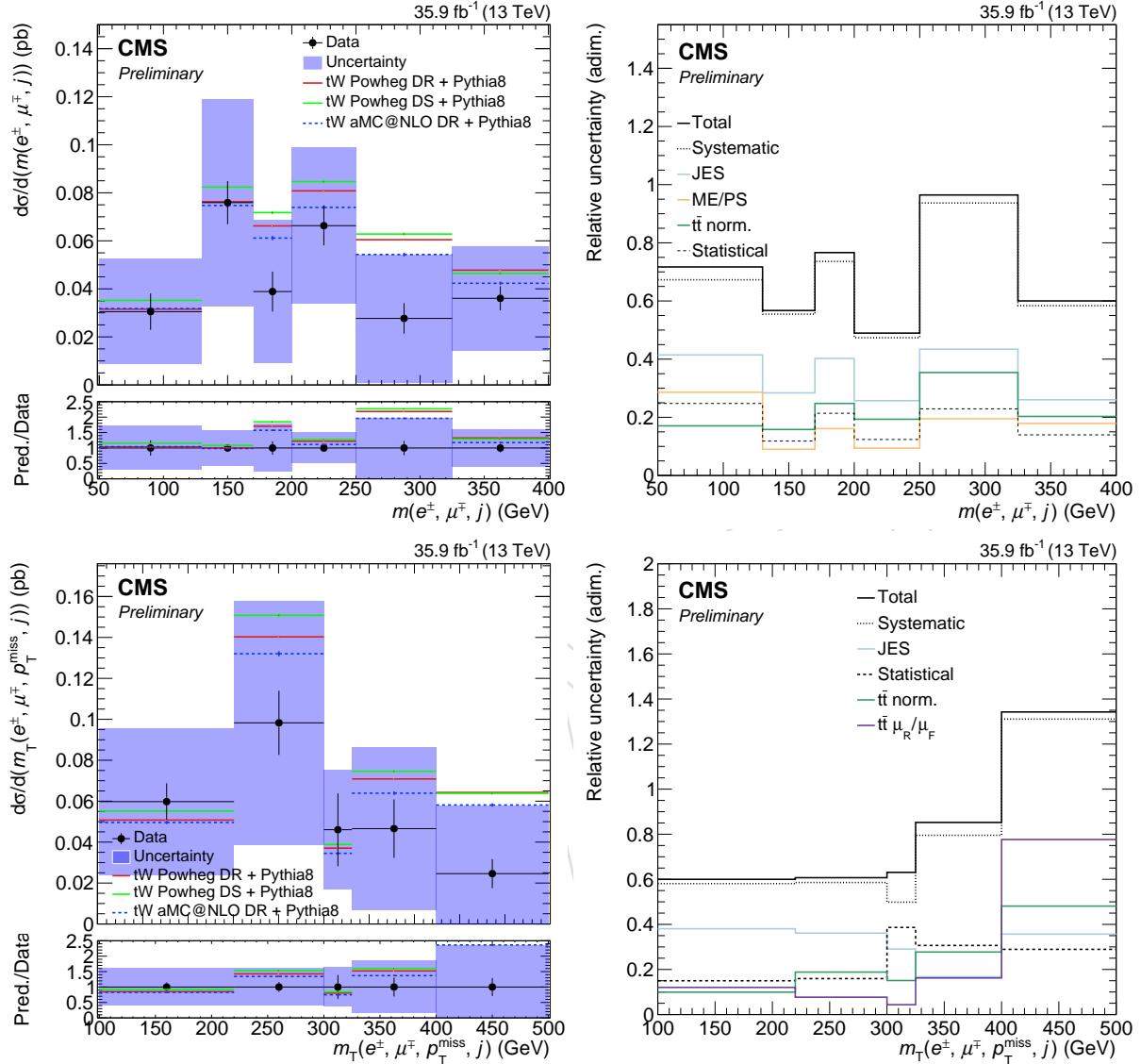


Figure 23: (Left) Differential  $tW$  production cross section as a function of  $m(e^\pm, \mu^\mp, j)$  and  $m_T(e^\pm, \mu^\mp, j, \cancel{E}_T)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

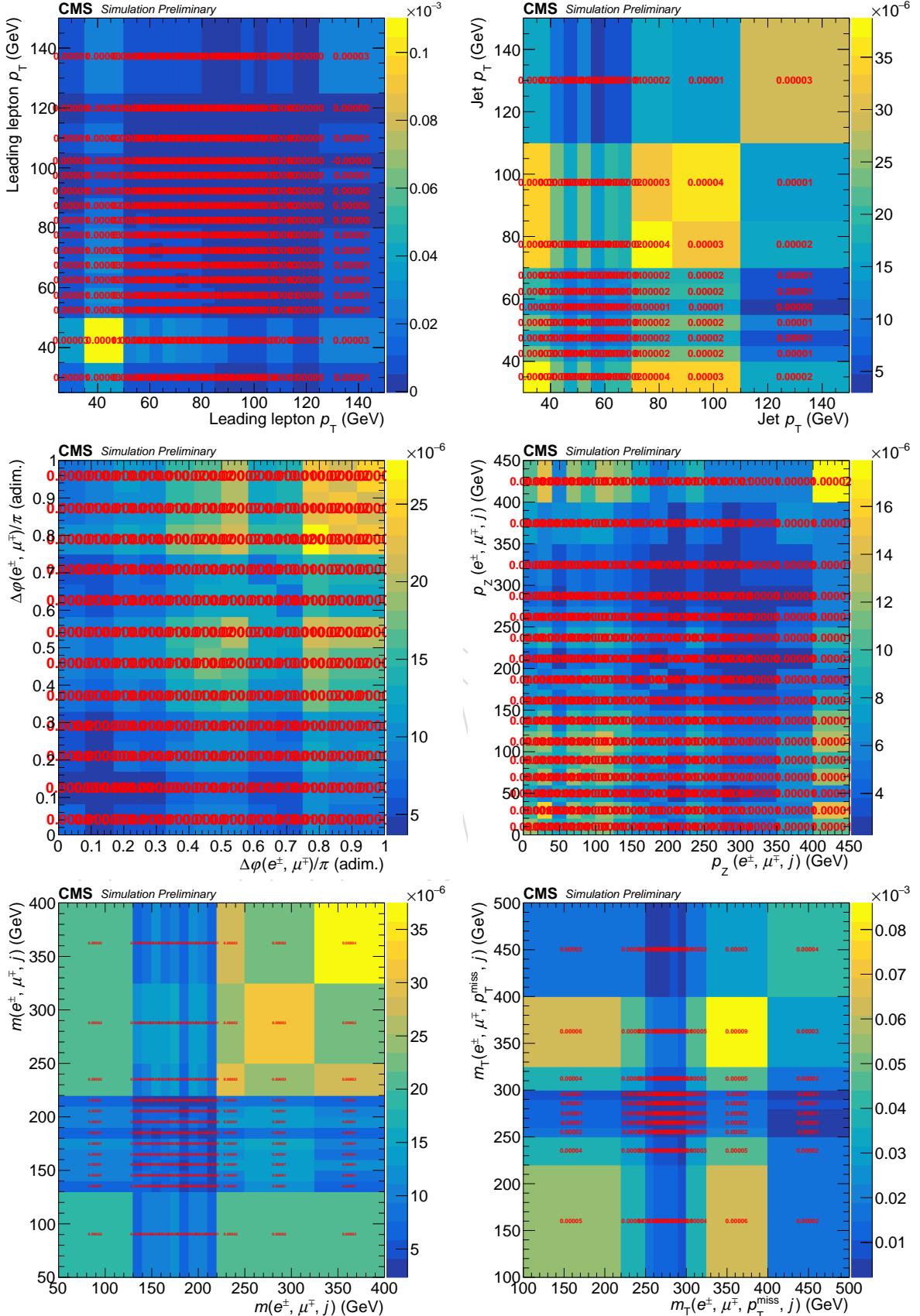


Figure 24: Covariance matrices between bins of the various observables in the folded space taking into consideration all the contributions from all uncertainty sources.

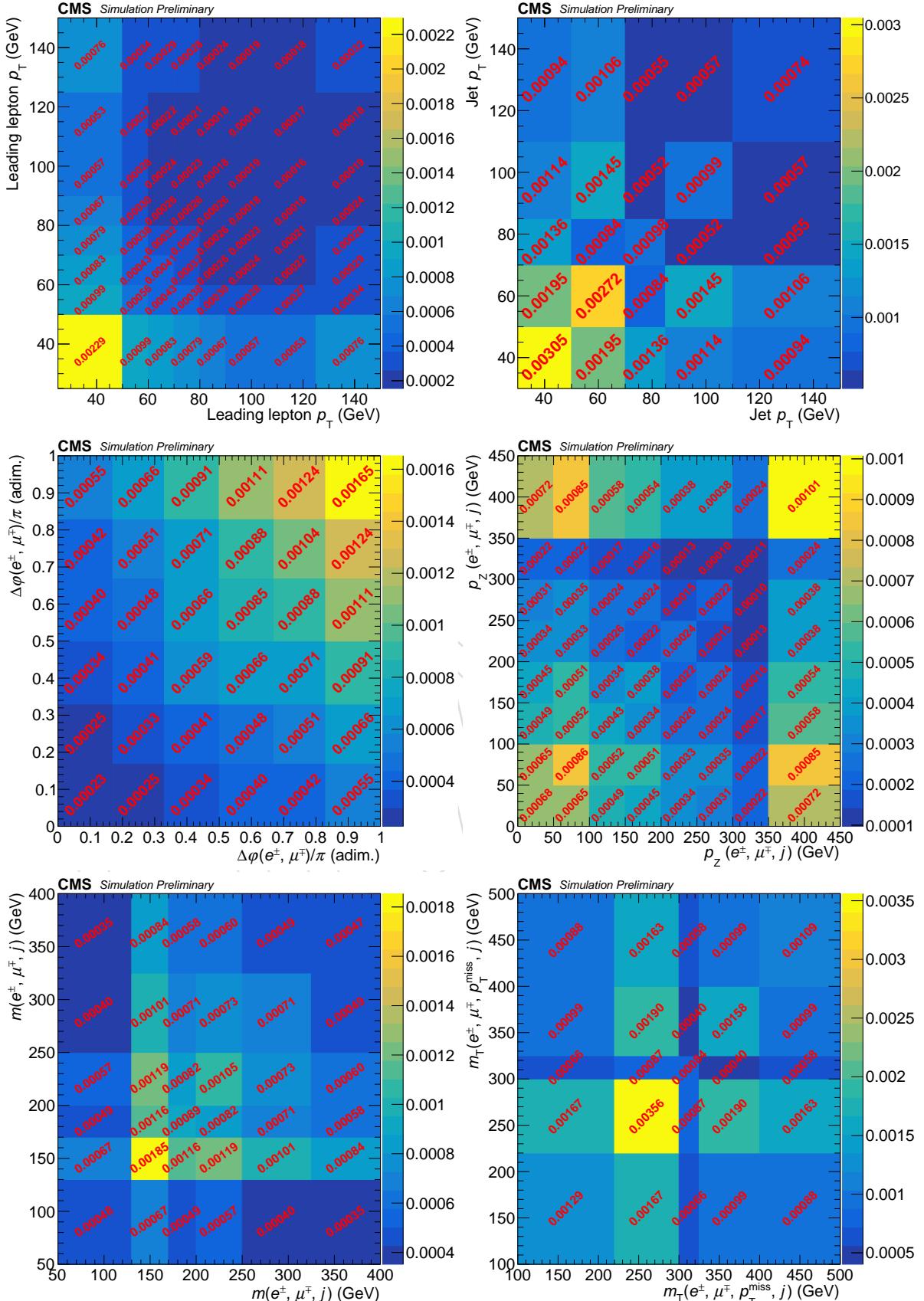


Figure 25: Covariance matrices between bins of the various observables in the unfolded space taking into consideration all the contributions from all uncertainty sources.

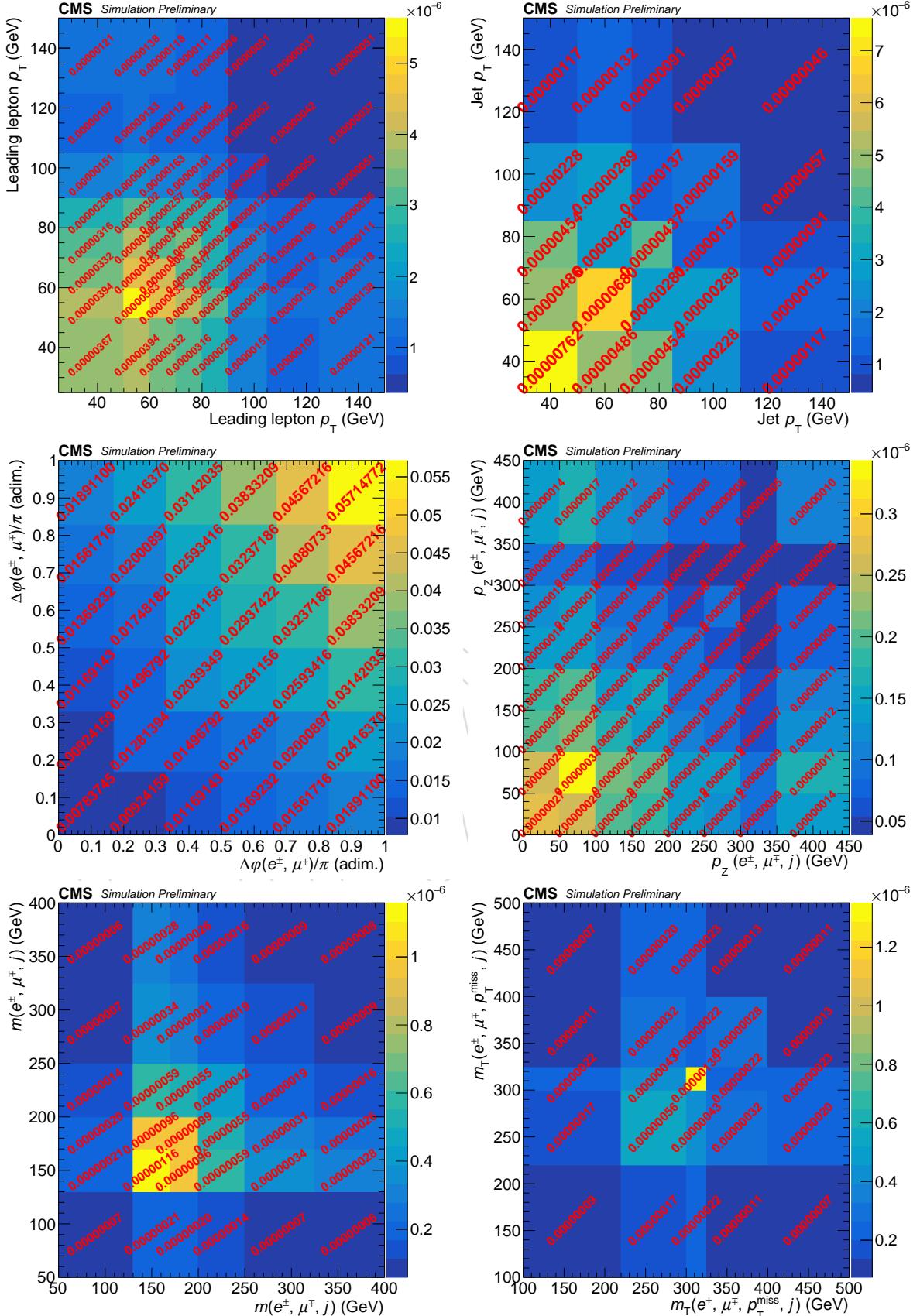


Figure 26: Covariance matrices between bins of the various observables in the unfolded space after normalising to the bin width taking into consideration all the contributions from all uncertainty sources.

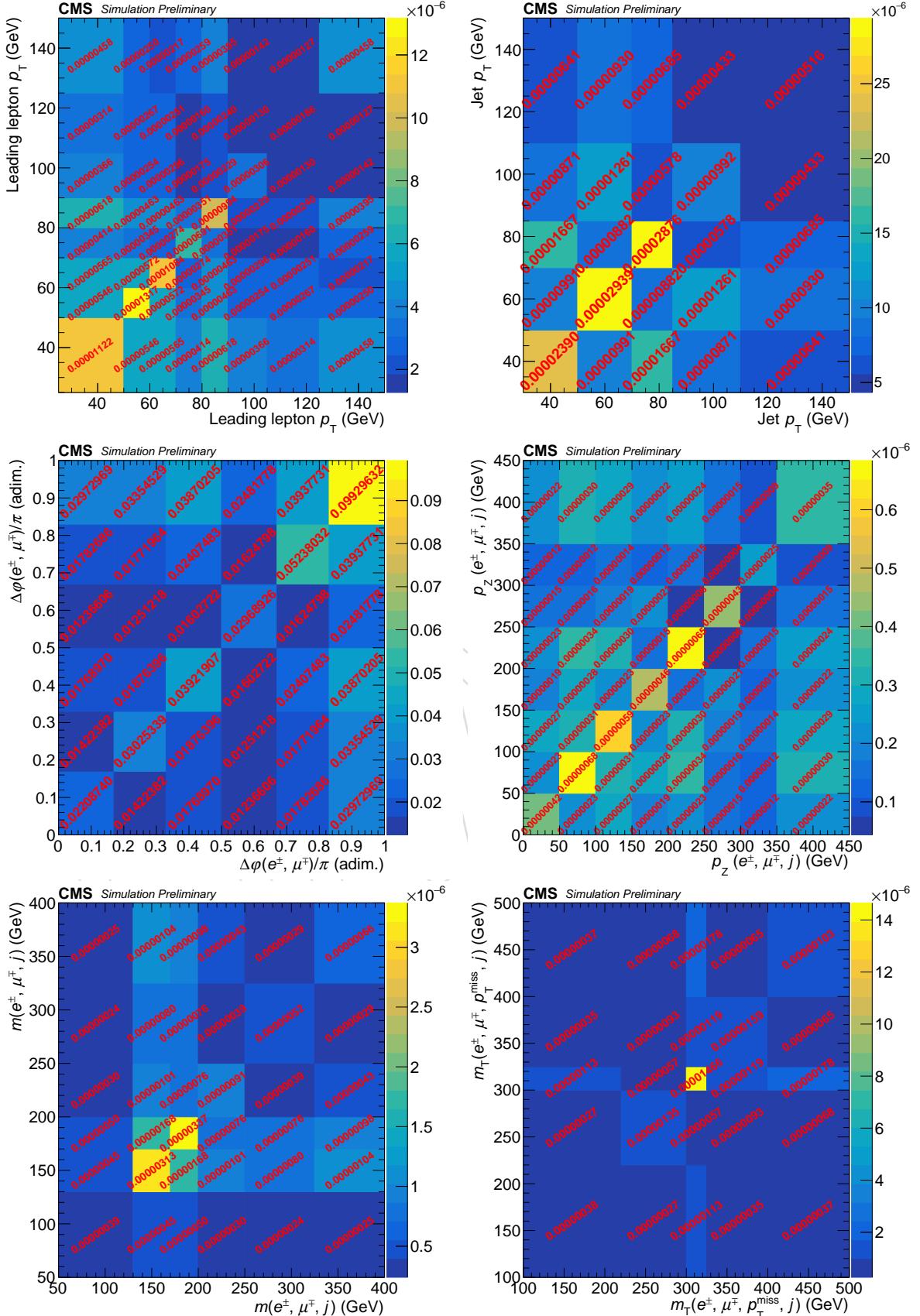


Figure 27: Covariance matrices between bins of the various observables in the unfolded space after normalising to the fiducial cross section and bin width taking into consideration all the contributions from all uncertainty sources.

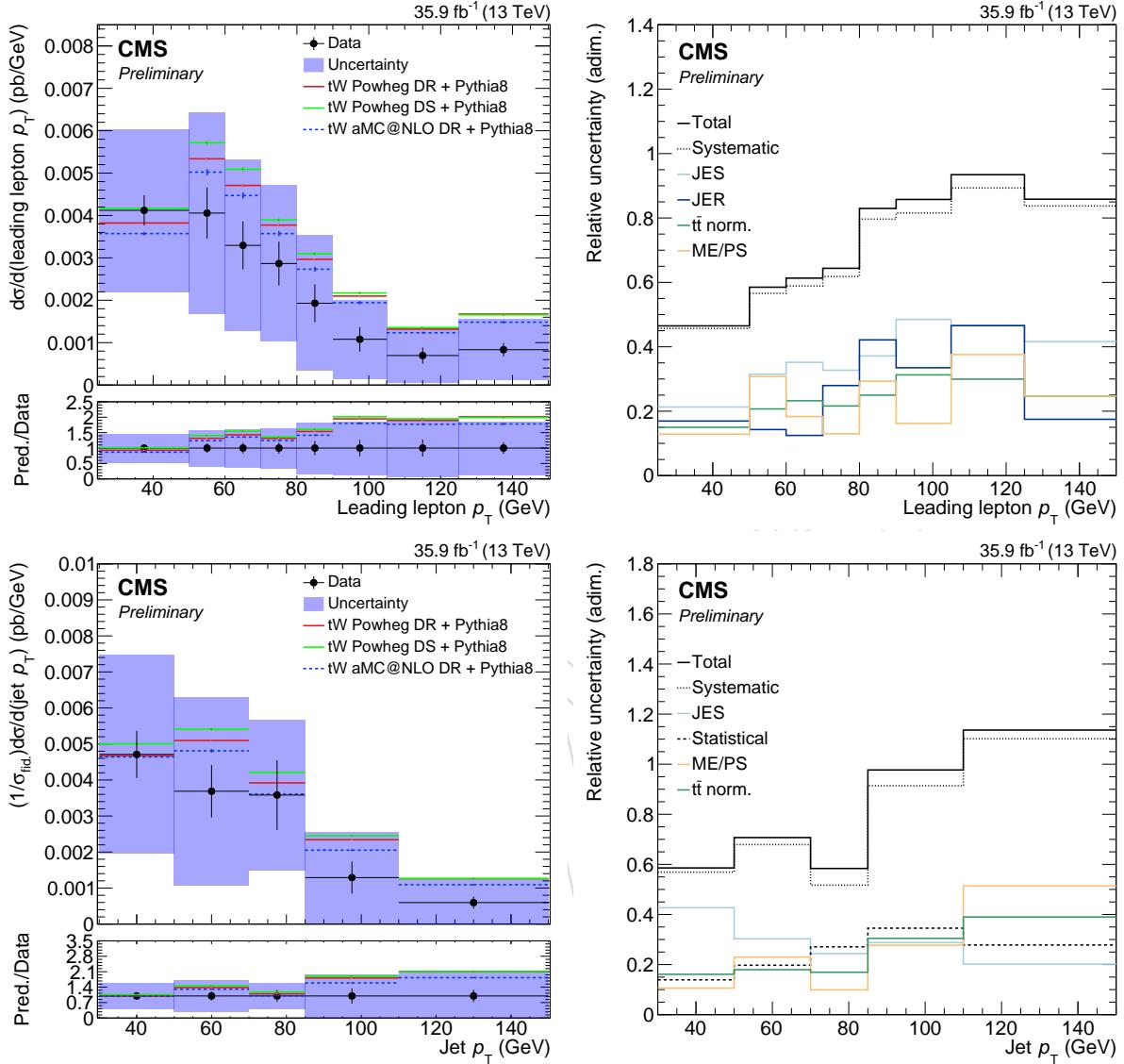


Figure 28: (Left) Differential tW production cross section normalised to the bin width as a function of the  $p_T$  of the leading lepton and of the jet of the process in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

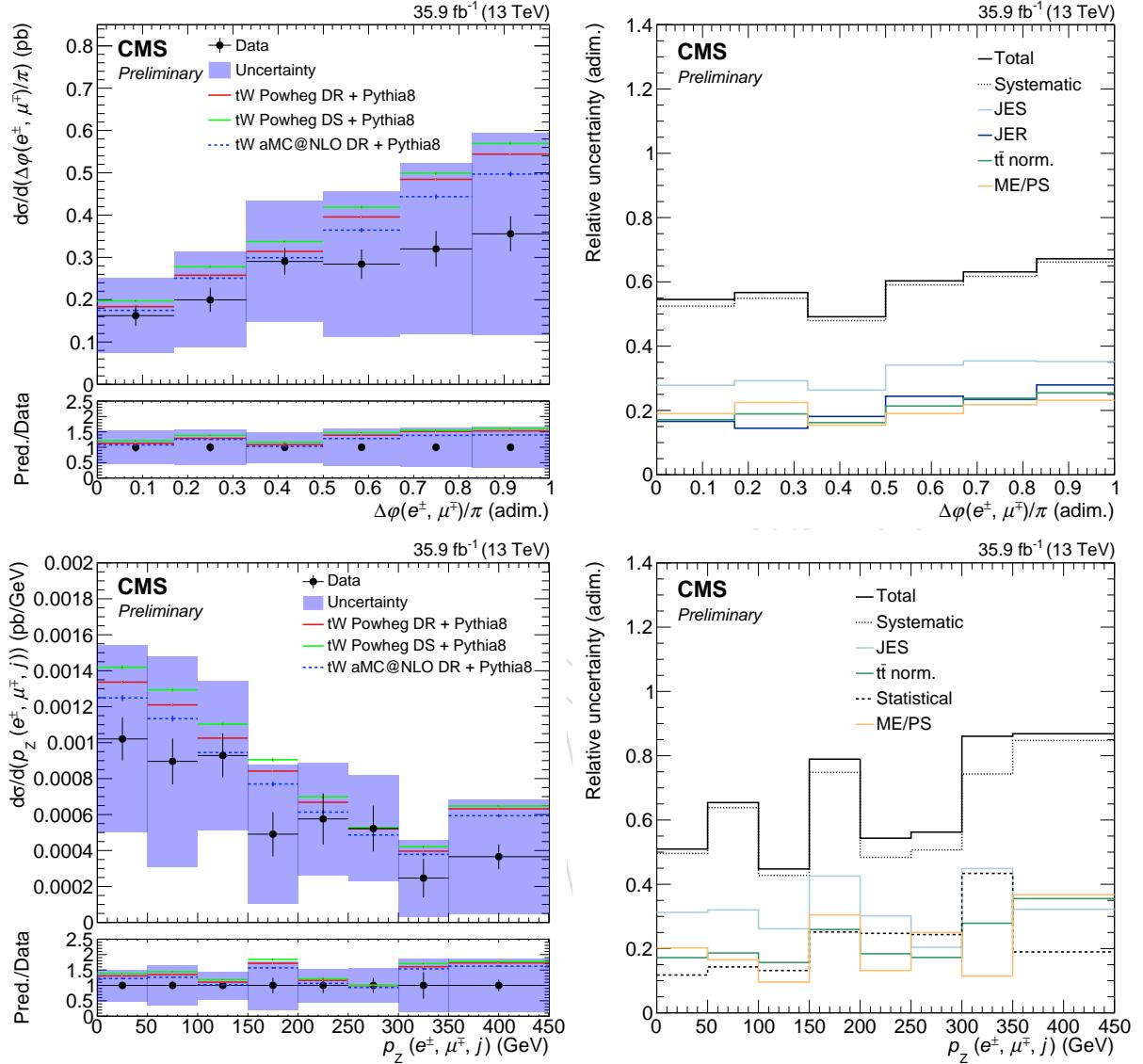


Figure 29: (Left) Differential tW production cross section normalised to the bin width as a function of  $\Delta\varphi(e^\pm, \mu^\mp)$  and  $p_Z(e^\pm, \mu^\mp, j)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

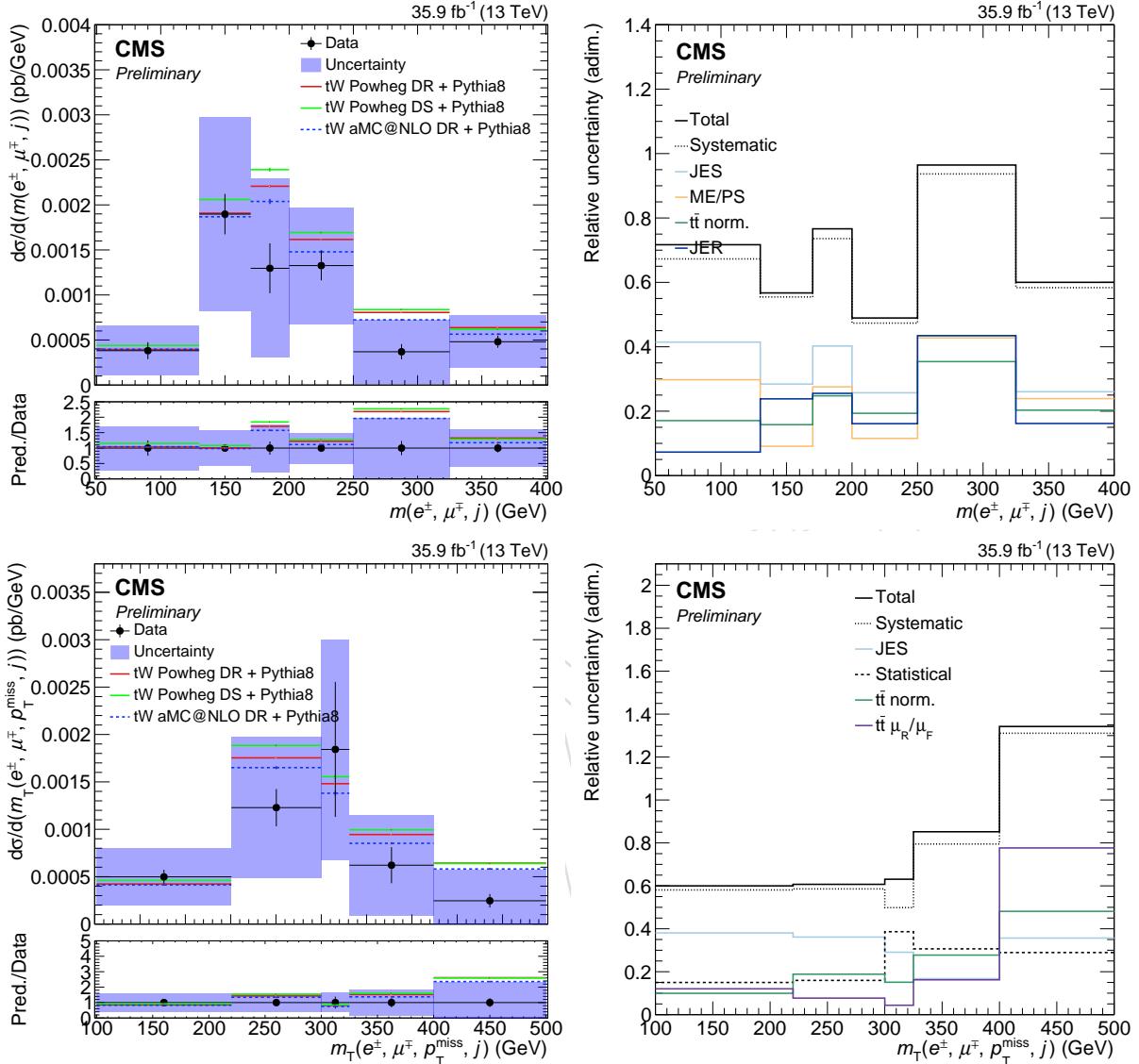


Figure 30: (Left) Differential tW production cross section normalised to the bin width as a function of  $m(e^\pm, \mu^\mp, j)$  and  $m_T(e^\pm, \mu^\mp, j, E_T)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

### 477 7.1 Normalised (to the fiducial region and also bin) results

478 In addition to the “absolute” results, we have calculated the differential cross section depending  
 479 on the same previous observables but normalising the final results to the values of the  
 480 fiducial cross section. This fiducial cross section ends having a value of

$$\sigma_{fid.} = 0.28^{+0.15}_{-0.14} \text{ pb} \quad (8)$$

481 The uncertainties considered are the same as the used for the various differential cross sections  
 482 showed in this section previously. They are propagated by measuring the fiducial cross sections  
 483 with all the various yields and taking the variation with the nominal ones as the uncertainty.

484 This yields the result plots showed in figures 31, 32 and 33. The numerical results are shown in  
 485 Tab. 7. In addition to this normalisation, we have also normalised the results to the bin width.  
 486 These other results are shown in figures 34, 35 and 36. The numerical results are shown in  
 487 Tab. 7 and the GOF tests of these final results are shown in Tab. 8.

Leading lepton $p_T$ (GeV)	[25, 50]	[50, 60]	[60, 70]	[70, 80]	[80, 90]	[90, 105]	[105, 125]	[125, $\infty$ )
$(1/\sigma_{fid})d\sigma/d$ (Leading lepton $p_T$ )	0.38	0.15	0.12	0.11	0.07	0.06	0.05	0.08
Uncertainty	0.08	0.04	0.03	0.03	0.03	0.03	0.03	0.05
Jet $p_T$ (GeV)	[30, 50]	[50, 70]	[70, 85]	[85, 110]	[110, $\infty$ )			
$(1/\sigma_{fid})d\sigma/d$ (Jet $p_T$ )	0.35	0.27	0.20	0.12	0.09			
Uncertainty	0.10	0.11	0.08	0.08	0.09			
$\Delta\varphi(e^\pm, \mu^\mp)/\pi$	[0.0, 0.17)	[0.17, 0.33)	[0.33, 0.5)	[0.5, 0.67)	[0.67, 0.83)	[0.83, 1]		
$(1/\sigma_{fid})d\sigma/d$ ( $\Delta\varphi(e^\pm, \mu^\mp)/\pi$ )	0.10	0.12	0.18	0.18	0.19	0.22		
Uncertainty	0.03	0.03	0.03	0.03	0.04	0.05		
$p_Z(e^\pm, \mu^\mp, j)$ (GeV)	[0, 50)	[50, 100)	[100, 150)	[150, 200)	[200, 250)	[250, 300)	[300, 350)	[350, $\infty$ )
$(1/\sigma_{fid})d\sigma/d$ ( $p_Z(e^\pm, \mu^\mp, j)$ )	0.19	0.17	0.17	0.09	0.11	0.10	0.05	0.14
Uncertainty	0.03	0.04	0.04	0.03	0.04	0.03	0.03	0.06
$m(e^\pm, \mu^\mp, j)$ (GeV)	[50, 130)	[130, 170)	[170, 200)	[200, 250)	[250, 325)	[325, $\infty$ )		
$(1/\sigma_{fid})d\sigma/d$ ( $m(e^\pm, \mu^\mp, j)$ )	0.11	0.28	0.14	0.25	0.10	0.13		
Uncertainty	0.05	0.07	0.06	0.05	0.05	0.06		
$m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$ (GeV)	[100, 220)	[220, 300)	[300, 325)	[325, 400)	[400, $\infty$ )			
$(1/\sigma_{fid})d\sigma/d$ ( $m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$ )	0.22	0.36	0.17	0.17	0.09			
Uncertainty	0.07	0.09	0.10	0.09	0.10			

Table 7: Numerical results of the normalised (fiducial cross section) differential cross sections for all the observables considered, as well as the upper and lower uncertainties for each case.

## 488 8 Summary

489 *Editor: Víctor Rodríguez*

490 In this document we have presented the measurement of the differential cross section of the

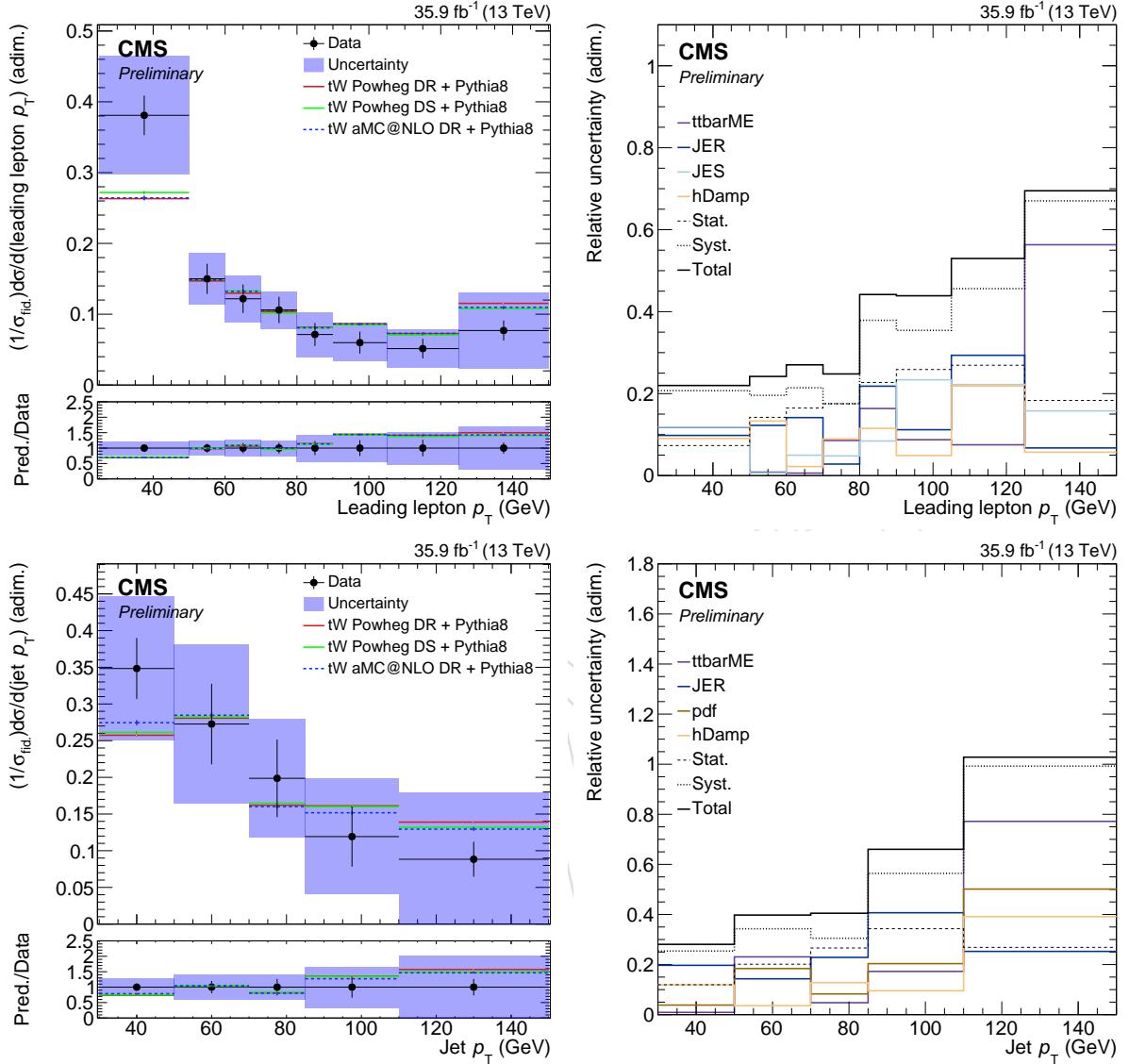


Figure 31: (Left) Normalised (fiducial cross section) differential  $tW$  production cross section as a function of the  $p_T$  of the leading lepton and of the jet of the process in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

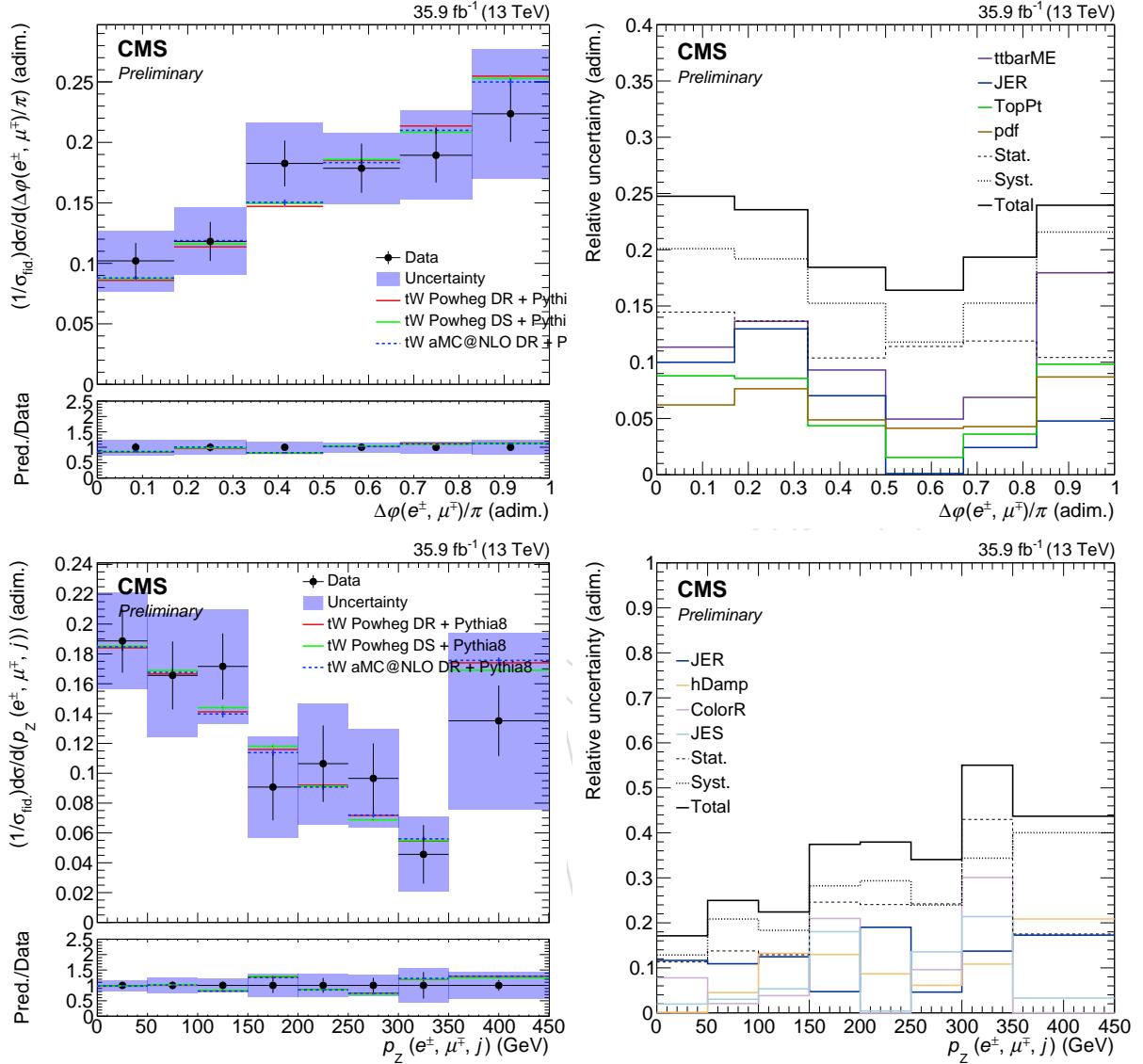


Figure 32: (Left) Normalised (fiducial cross section) differential tW production cross section as a function of  $\Delta\phi(e^\pm, \mu^\mp)$  and  $p_Z(e^\pm, \mu^\mp, j)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

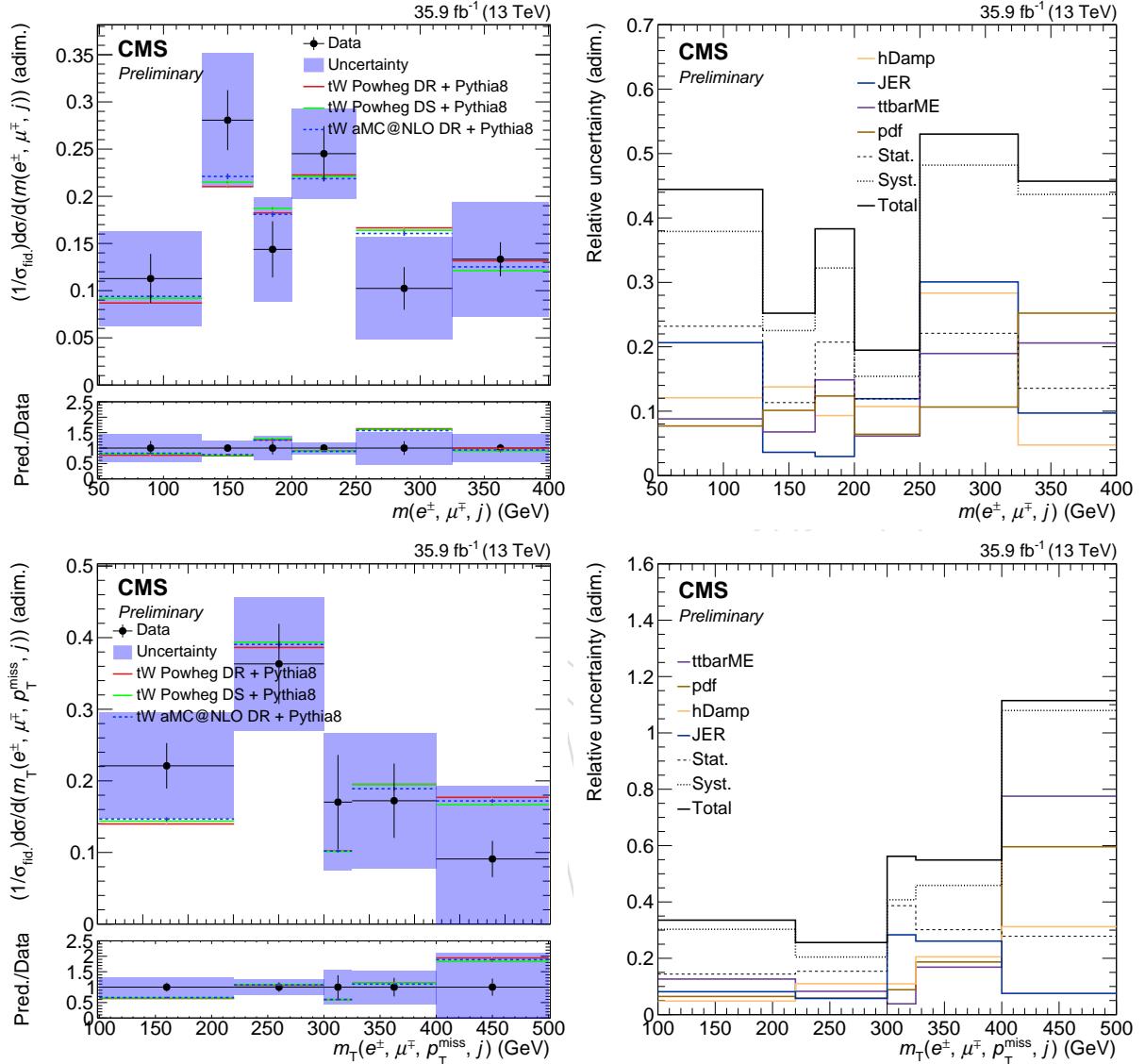


Figure 33: (Left) Normalised (fiducial cross section) differential tW production cross section as a function of  $m(e^\pm, \mu^\mp, j)$  and  $m_T(e^\pm, \mu^\mp, j, \cancel{E}_T)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

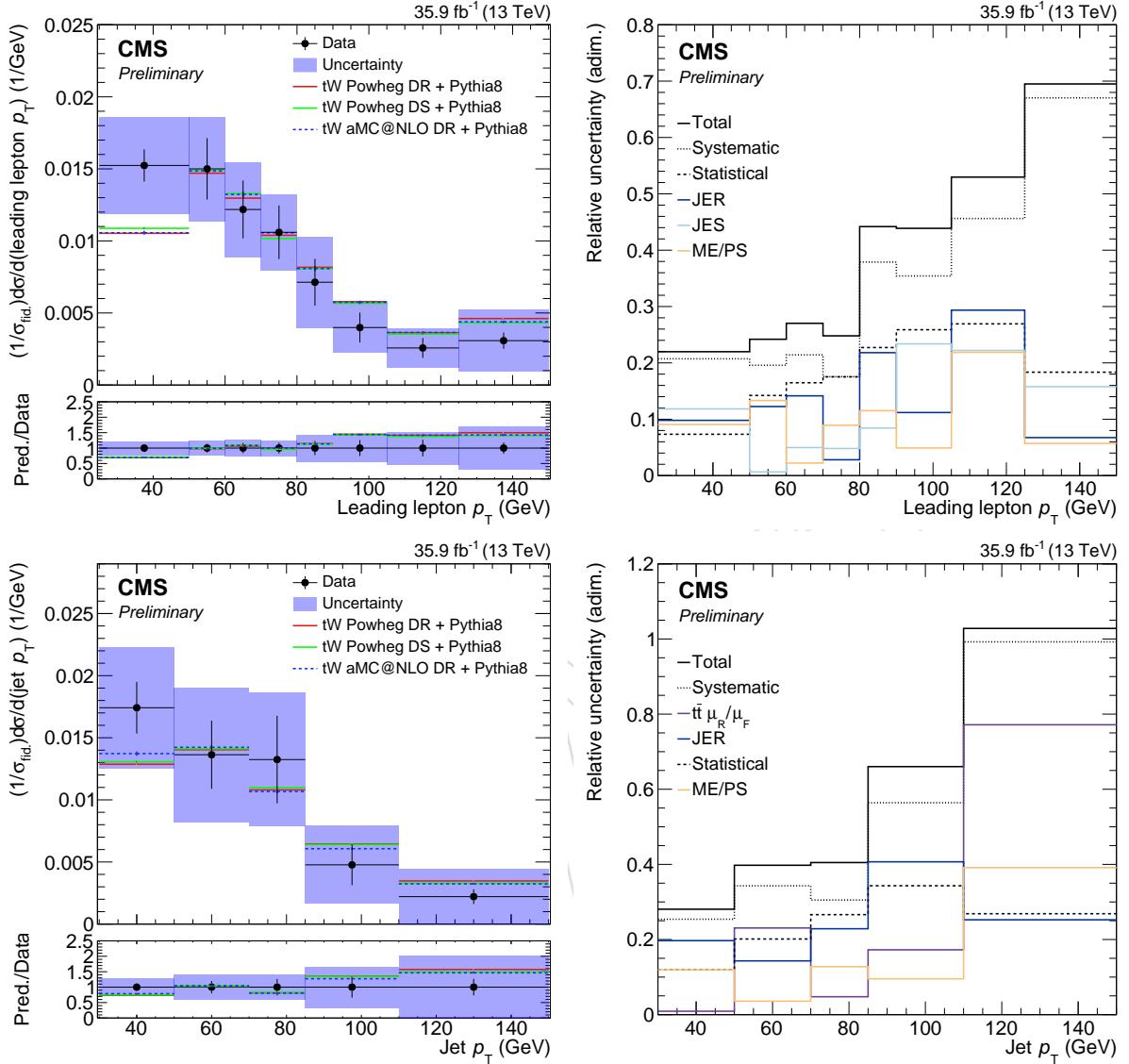


Figure 34: (Left) Normalised (fiducial cross section and bin) differential tW production cross section as a function of the  $p_T$  of the leading lepton and of the jet of the process in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

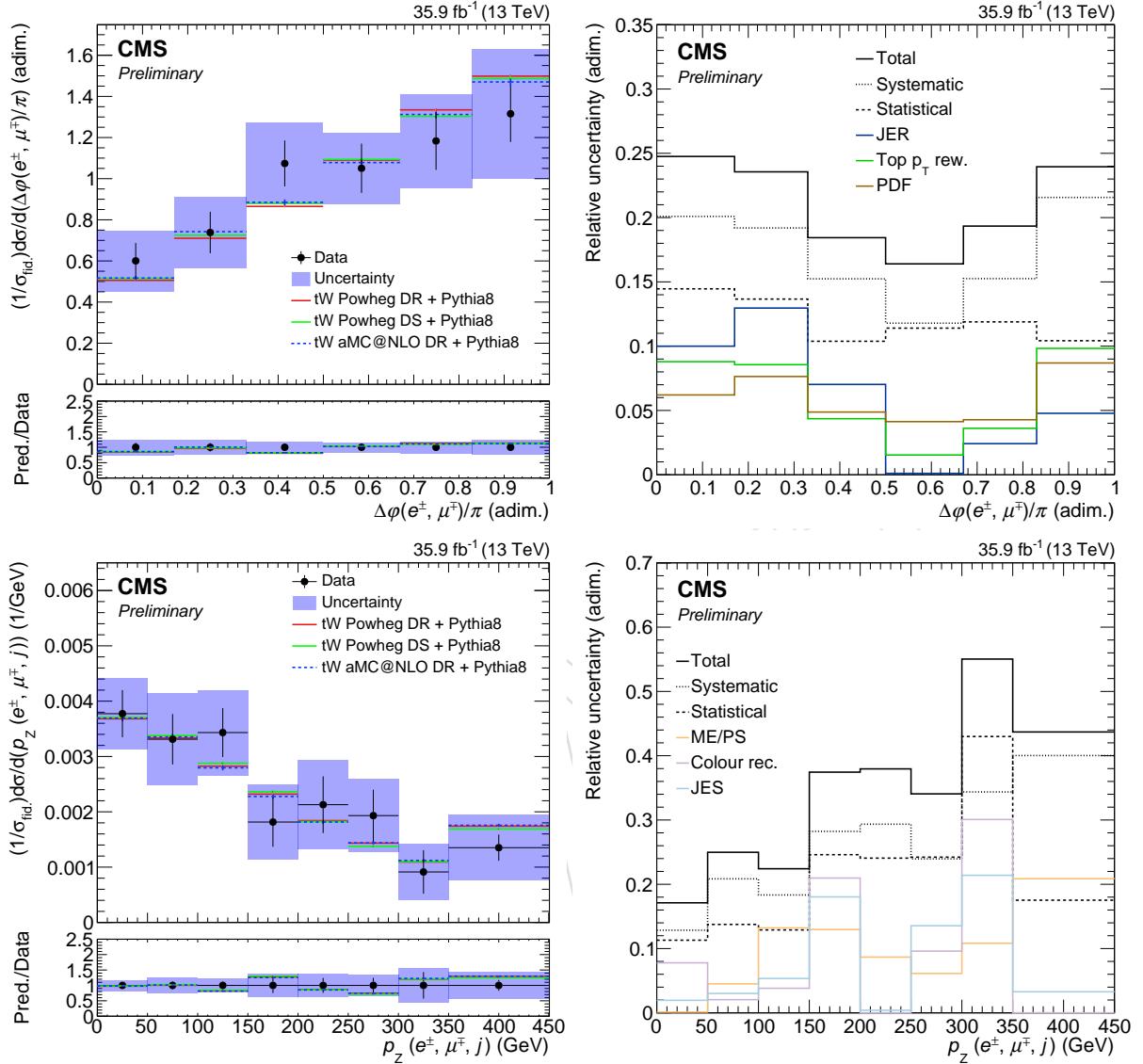


Figure 35: (Left) Normalised (fiducial cross section and bin) differential tW production cross section as a function of  $\Delta\phi(e^\pm, \mu^\mp)$  and  $p_Z(e^\pm, \mu^\mp, j)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

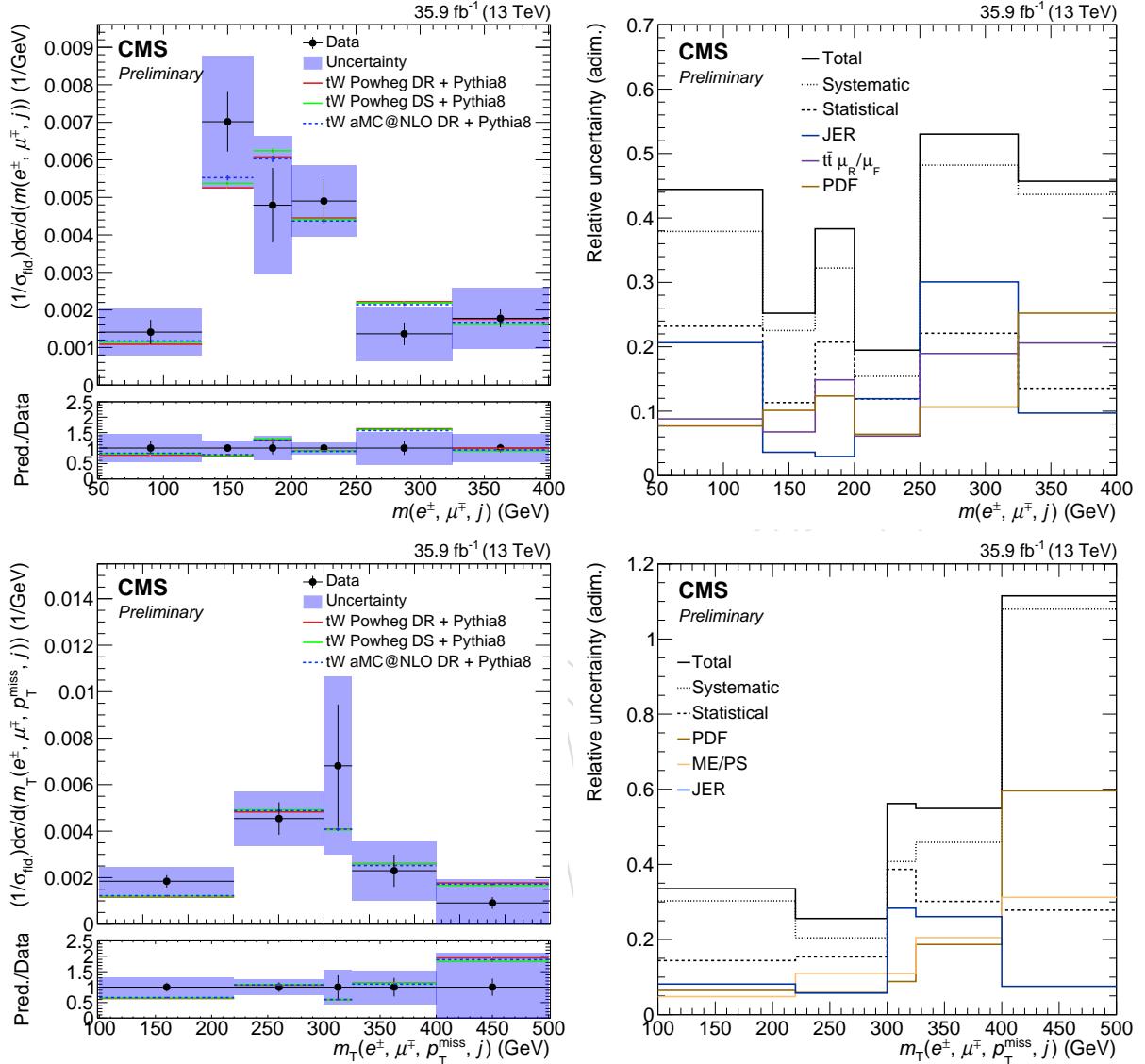


Figure 36: (Left) Normalised (fiducial cross section and bin) differential tW production cross section as a function of  $m(e^\pm, \mu^\mp, j)$  and  $m_T(e^\pm, \mu^\mp, j, E_T)$  in the unfolded space. For each variable the panel in the top shows the measured cross section, together with the statistical uncertainty depicted as bars. The blue band represents the total uncertainty. A comparison with the MG5\_AMC@NLO values and also the ones from the POWHEG sample obtained using the DS scheme are displayed. In the bottom panel, the ratio between data and the POWHEG models are shown. The statistical uncertainty is shown in a light blue, while the total uncertainty is shown in darker blue. (Right) Six leading (by average of the relative unc. in all bins) sources of uncertainty of the differential cross section measurement as a function of each observable, in addition of the total uncertainty (in black).

Variable/MC sample	POWHEG DR	POWHEG DS	MG5_AMC@NLO
Leading lepton $p_T$	0.017 — 17.096	0.046 — 14.276	0.025 — 16.054
Jet $p_T$	0.297 — 4.903	0.392 — 4.106	0.587 — 2.828
$\Delta\varphi(e^\pm, \mu^\mp)/\pi$	0.345 — 5.618	0.469 — 4.578	0.513 — 4.259
$p_Z(e^\pm, \mu^\mp, j)$	0.653 — 5.054	0.67 — 4.916	0.635 — 5.202
$m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$	0.278 — 5.095	0.361 — 4.348	0.357 — 4.38
$m(e^\pm, \mu^\mp, j)$	0.103 — 9.165	0.168 — 7.786	0.238 — 6.768

Table 8: Results (p-value – test statistic value) of the  $\chi^2$  goodness-of-fit tests performed to check the compatibility between data and the Powheg DR, Powheg DS and aMCatNLO models for the normalised (fiducial cross section and bin) results.

Table 9: Numerical results of the normalised (fiducial cross section and bin) differential cross sections for all the observables considered, as well as the upper and lower uncertainties for each case.

Leading lepton $p_T$ (GeV)	[25, 50]	[50, 60]	[60, 70]	[70, 80]	[80, 90]	[90, 105]	[105, 125]	[125, $\infty$ )
$(1/\sigma_{\text{fid}})d\sigma/d(\text{Leading lepton } p_T)$ (1/GeV)	0.015	0.015	0.012	0.011	0.007	0.0040	0.0026	0.003
Uncertainty (1/GeV)	0.003	0.004	0.003	0.003	0.003	0.0018	0.0014	0.002
Jet $p_T$ (GeV)	[30, 50]	[50, 70]	[70, 85]	[85, 110]	[110, $\infty$ )			
$(1/\sigma_{\text{fid}})d\sigma/d(\text{Jet } p_T)$ (1/GeV)	0.017	0.014	0.013	0.005	0.002			
Uncertainty (1/GeV)	0.005	0.005	0.005	0.003	0.002			
$\Delta\varphi(e^\pm, \mu^\mp)/\pi$	[0.0, 0.17]	[0.17, 0.33]	[0.33, 0.5]	[0.5, 0.67]	[0.67, 0.83]	[0.83, 1]		
$(1/\sigma_{\text{fid}})d\sigma/d(\Delta\varphi(e^\pm, \mu^\mp)/\pi)$ (1/GeV)	0.60	0.74	1.07	1.05	1.2	1.3		
Uncertainty (1/GeV)	0.15	0.17	0.20	0.17	0.2	0.3		
$p_Z(e^\pm, \mu^\mp, j)$ (GeV)	[0, 50]	[50, 100]	[100, 150]	[150, 200]	[200, 250]	[250, 300]	[300, 350]	[350, $\infty$ )
$(1/\sigma_{\text{fid}})d\sigma/d(p_Z(e^\pm, \mu^\mp, j))$ (1/GeV)	0.0038	0.0033	0.0034	0.0018	0.0021	0.0019	0.0009	0.0014
Uncertainty (1/GeV)	0.0006	0.0008	0.0008	0.0007	0.0008	0.0007	0.0005	0.0006
$m(e^\pm, \mu^\mp, j)$ (GeV)	[50, 130]	[130, 170]	[170, 200]	[200, 250]	[250, 325]	[325, $\infty$ )		
$(1/\sigma_{\text{fid}})d\sigma/d(m(e^\pm, \mu^\mp, j))$ (1/GeV)	0.0014	0.0070	0.0048	0.0049	0.0014	0.0018		
Uncertainty (1/GeV)	0.0006	0.0018	0.0018	0.0010	0.0007	0.0008		
$m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}})$ (GeV)	[100, 220)	[220, 300)	[300, 325)	[325, 400)	[400, $\infty$ )			
$(1/\sigma_{\text{fid}})d\sigma/d(m_T(e^\pm, \mu^\mp, j, p_T^{\text{miss}}))$ (1/GeV)	0.0018	0.0045	0.007	0.0023	0.0009			
Uncertainty (1/GeV)	0.0006	0.0012	0.004	0.0013	0.0010			

491 production of a top quark along a W boson using  $35.9 \text{ fb}^{-1}$  of data from 2016. This observable  
492 has been measured depending on various properties of the events: the  $p_T$  of the leading lepton,  
493 the  $p_T$  of the jet, the difference in the  $\varphi$  angle of the muon and the electron present in the events,  
494 the component in the Z axis of the muon, the electron and the jet of the event, the invariant mass  
495 of the muon, electron and the jet of the events and finally the transverse mass of the electron,  
496 the muon, the jet and the missing transverse energy of the event. The analysis is absolutely  
497 dominated by the overwhelming presence of the  $t\bar{t}$  process that acts as the main background.

498 The results obtained are consistent with the two models studied, POWHEG and MG5\_AMC@NLO,  
499 although some differences appear especially in the tails of the distributions. The comparison  
500 with the POWHEG simulated samples that use the DS approach to model the tW process is also  
501 considered, and are in all cases within the uncertainty of the results, except again in the tails of  
502 some distributions.

503 The uncertainties of the results are influenced the most by the effects of the jet detections (i.e.  
504 JER, JES), what makes sense considering the definition of our signal region and the large pres-  
505 ence of  $t\bar{t}$  background. Also modeling uncertainties (such as the ME scale or the matching  
506 between the ME and the PS) that concern the  $t\bar{t}$  background. Regarding the total uncertainties,  
507 they are about the  $\sim 40 - 75\%$  in the bulks of the not normalised distributions, reaching higher  
508 values in the tails.

509 When considering the normalised results, the uncertainties are reduced due to the fact that  
510 the variations in the fiducial cross section and the differentials can be relatively compensated  
511 when computing the ratio. This yields results that in some cases make the comparison with  
512 the predictions (POWHEG and MG5\_AMC@NLO) out the total uncertainty bar. Roughly, they  
513 sit in a  $25 - 35\%$ , being the most predominant, as in the absolute values, the ones related with  
514 the jets identifications (e.g. JER, JES) and others related with the modeling of the simulations  
515 (e.g. the matching, hDamp; the FSR/ISR...). The uncertainties are, exactly as with the absolute  
516 values, predominantly systematic (roughly always over the  $90 - 95\%$ , except in some bins).

## 517 References

- 518 [1] CDF Collaboration, “First observation of electroweak single top quark production”,  
519 *Phys. Rev. Lett.* **103** (2009) 092002, doi:[10.1103/PhysRevLett.103.092002](https://doi.org/10.1103/PhysRevLett.103.092002),  
520 arXiv:[0903.0885](https://arxiv.org/abs/0903.0885).
- 521 [2] D0 Collaboration, “Observation of single top quark production”, *Phys. Rev. Lett.* **103**  
522 (2009) 092001, doi:[10.1103/PhysRevLett.103.092001](https://doi.org/10.1103/PhysRevLett.103.092001), arXiv:[0903.0850](https://arxiv.org/abs/0903.0850).
- 523 [3] S. Frixione et al., “Single-top hadroproduction in association with a W boson”, *JHEP* **07**  
524 (2008) 029, doi:[10.1088/1126-6708/2008/07/029](https://doi.org/10.1088/1126-6708/2008/07/029), arXiv:[0805.3067](https://arxiv.org/abs/0805.3067).
- 525 [4] A. S. Belyaev, E. E. Boos, and L. V. Dudko, “Single top quark at future hadron colliders:  
526 Complete signal and background study”, *Phys. Rev. D* **59** (1999) 075001,  
527 doi:[10.1103/PhysRevD.59.075001](https://doi.org/10.1103/PhysRevD.59.075001), arXiv:[hep-ph/9806332](https://arxiv.org/abs/hep-ph/9806332).
- 528 [5] C. D. White, S. Frixione, E. Laenen, and F. Maltoni, “Isolating Wt production at the LHC”,  
529 *JHEP* **11** (2009) 074, doi:[10.1088/1126-6708/2009/11/074](https://doi.org/10.1088/1126-6708/2009/11/074), arXiv:[0908.0631](https://arxiv.org/abs/0908.0631).
- 530 [6] T. M. P. Tait and C. P. Yuan, “Single top quark production as a window to physics beyond  
531 the standard model”, *Phys. Rev. D* **63** (2000) 014018,  
532 doi:[10.1103/PhysRevD.63.014018](https://doi.org/10.1103/PhysRevD.63.014018), arXiv:[hep-ph/0007298](https://arxiv.org/abs/hep-ph/0007298).

- [7] Q.-H. Cao, J. Wudka, and C. P. Yuan, "Search for new physics via single top production at the LHC", *Phys. Lett. B* **658** (2007) 50, doi:10.1016/j.physletb.2007.10.057, arXiv:0704.2809.
- [8] V. Barger, M. McCaskey, and G. Shaughnessy, "Single top and higgs associated production at the lhc", *Phys. Rev. D* **81** (Feb, 2010) 034020, doi:10.1103/PhysRevD.81.034020.
- [9] A. S. Belyaev, E. E. Boos, and L. V. Dudko, "Single top quark at future hadron colliders: Complete signal and background study", *Phys. Rev. D* **59** (Feb, 1999) 075001, doi:10.1103/PhysRevD.59.075001.
- [10] T. M. P. Tait, " $tW^-$  mode of single top quark production", *Phys. Rev. D* **61** (1999) 034001, doi:10.1103/PhysRevD.61.034001.
- [11] N. Kidonakis, "Theoretical results for electroweak-boson and single-top production", in *Proceedings, 23rd International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2015): Dallas, Texas, USA, April 27-May 01, 2015*, volume DIS2015, p. 170. 2015. arXiv:1506.04072.
- [12] CMS Collaboration, "Evidence for associated production of a single top quark and W boson in pp collisions at  $\sqrt{s} = 7$  TeV", *Phys. Rev. Lett.* **110** (2013) 022003, doi:10.1103/PhysRevLett.110.022003, arXiv:1209.3489.
- [13] ATLAS Collaboration, "Evidence for the associated production of a W boson and a top quark in ATLAS at  $\sqrt{s} = 7$  TeV", *Phys. Lett. B* **716** (2012) 142, doi:10.1016/j.physletb.2012.08.011, arXiv:1205.5764.
- [14] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at  $\sqrt{s} = 8$  TeV", *Phys. Rev. Lett.* **112** (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- [15] ATLAS Collaboration, "Measurement of the production cross-section of a single top quark in association with a W boson at 8 TeV with the ATLAS experiment", *JHEP* **01** (2016) 064, doi:10.1007/JHEP01(2016)064, arXiv:1510.03752.
- [16] CMS Collaboration, "Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at  $\sqrt{s} = 13$  TeV", *JHEP* **10** (2018) 117, doi:10.1007/JHEP10(2018)117, arXiv:1805.07399.
- [17] ATLAS Collaboration, "Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at  $\sqrt{s} = 13$  TeV with ATLAS", *JHEP* **01** (2018) 063, doi:10.1007/JHEP01(2018)063, arXiv:1612.07231.
- [18] ATLAS Collaboration, "Measurement of differential cross-sections of a single top quark produced in association with a W boson at  $\sqrt{s} = 13$  TeV with ATLAS", *Eur. Phys. J. C* **78** (2018) 186, doi:10.1140/epjc/s10052-018-5649-8, arXiv:1712.01602.
- [19] E. Palencia, S. Sánchez et al., "Single top associated  $tW$  production at 13 TeV", *CMS AN-2017/113* (2017).
- [20] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

- 574 [21] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2>.
- 575
- 576 [22] [https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Tight\\_Muon](https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Tight_Muon).
- 577
- 578 [23] [https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID#Recommendations\\_for\\_13\\_TeV\\_2016](https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID#Recommendations_for_13_TeV_2016).
- 579
- 580 [24] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2>.
- 581
- 582 [25] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData>.
- 583
- 584 [26] T. Arndt et al., “Measurement of the 2016 Trigger Efficiencies for a dilepton Selection for a  $t\bar{t}$  analysis”, *CMS AN-2016/392* (2016).
- 585
- 586 [27] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2>.
- 587
- 588 [28] [https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaIDRecipesRun2#Cut\\_based\\_electron\\_identificatio](https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaIDRecipesRun2#Cut_based_electron_identificatio).
- 589
- 590 [29] [https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation#Recommendation\\_for\\_13\\_TeV\\_Data](https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation#Recommendation_for_13_TeV_Data).
- 591
- 592 [30] CMS collaboration Collaboration, “Site: <https://twiki.cern.ch/twiki/bin/view/CMS/TWikiTopRefEventSel>”,.
- 593
- 594 [31] CMS Collaboration, “Measurement of normalized differential  $t\bar{t}$  cross sections in the dilepton channel from pp collisions at  $\sqrt{s} = 13$  TeV”, *JHEP* **04** (2018) 060, doi:10.1007/JHEP04(2018)060, arXiv:1708.07638.
- 595
- 596
- 597 [32] CMS Collaboration, “Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV”, *Phys. Rev. D* **95** (2017) 092001, doi:10.1103/PhysRevD.95.092001, arXiv:1610.04191.
- 598
- 599
- 600 [33] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* **12** (2017), no. 02, P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- 601
- 602
- 603 [34] CMS Collaboration, “Jet algorithms performance in 13 TeV data”, CMS Physics Analysis Summary CMS-PAS-JME-16-003, CERN, 2017.
- 604
- 605 [35] [https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution#JER\\_Scaling\\_factors\\_and\\_Uncertai](https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution#JER_Scaling_factors_and_Uncertai).
- 606
- 607 [36] J. R. Gonzalez, et al., “Measurement of the Top-Quark Pair Production Cross Section in the Dilepton Channel with  $35.9 \text{ fb}^{-1}$  of 13 TeV data using the Cut and Count Method”, *CMS AN-2017/039* (2017).
- 608
- 609
- 610 [37] M. Defranchis, T. Arndt, et al., “Measurement of the inclusive top-quark pair production cross section and top quark mass with the full 2016 dataset using a likelihood fit”, *CMS AN-2016/447* (2016).
- 611
- 612

- 613 [38] CMS Collaboration, "CMS luminosity measurements for the 2016 data taking period",  
614 CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, 2017.
- 615 [39] CMS Collaboration, "Investigations of the impact of the parton shower tuning in Pythia 8  
616 in the modelling of  $t\bar{t}$  at  $\sqrt{s} = 8$  and 13 TeV", CMS Physics Analysis Summary  
617 CMS-PAS-TOP-16-021, CERN, 2016.
- 618 [40] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur.  
619 Phys. J. C* **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y,  
620 arXiv:1404.5630.
- 621 [41] J. R. Christiansen and P. Z. Skands, "String formation beyond leading colour", *JHEP* **08**  
622 (2015) 003, doi:10.1007/JHEP08(2015)003, arXiv:1505.01681.
- 623 [42] S. Argyropoulos and T. Sjöstrand, "Effects of color reconnection on  $t\bar{t}$  final states at the  
624 LHC", *JHEP* **11** (2014) 043, doi:10.1007/JHEP11(2014)043, arXiv:1407.6653.
- 625 [43] M. Czakon, D. Heymes, and A. Mitov, "High-precision differential predictions for  
626 top-quark pairs at the lhc", *Phys. Rev. Lett.* **116** (Feb, 2016) 082003,  
627 doi:10.1103/PhysRevLett.116.082003.
- 628 [44] N. Kidonakis, "Nnll threshold resummation for top-pair and single-top production",  
629 arXiv:arXiv:1210.7813.
- 630 [45] C. Collaboration, "Measurement of the differential cross section for top quark pair  
631 production in pp collisions at  $\sqrt{s} = 8$  tev",  
632 doi:10.1140/epjc/s10052-015-3709-x, arXiv:arXiv:1505.04480.
- 633 [46] CMS Collaboration Collaboration, "Measurement of the differential cross section for  $t\bar{t}$   
634 production in the dilepton final state at  $\sqrt{s} = 13$  TeV", Technical Report  
635 CMS-PAS-TOP-16-011, CERN, Geneva, 2016.
- 636 [47] CMS Collaboration Collaboration, "Measurement of the inclusive and differential  $t\bar{t}$   
637 production cross sections in lepton + jets final states at 13 TeV", Technical Report  
638 CMS-PAS-TOP-16-008, CERN, Geneva, 2016.
- 639 [48] C. Collaboration, "Measurement of the  $t$ - $\bar{t}$  production cross section in the all-jets final  
640 state in pp collisions at  $\sqrt{s} = 8$  tev", doi:10.1140/epjc/s10052-016-3956-5,  
641 arXiv:arXiv:1509.06076.
- 642 [49] J. Gao et al., "CT10 next-to-next-to-leading order global analysis of QCD", *Phys. Rev. D*  
643 **89** (2014) 033009, doi:10.1103/PhysRevD.89.033009, arXiv:1302.6246.
- 644 [50] CMS Collaboration Collaboration, "Object definitions for top quark analyses at the  
645 particle level", Technical Report CMS-NOTE-2017-004. CERN-CMS-NOTE-2017-004,  
646 CERN, Geneva, Jun, 2017.
- 647 [51] S. Schmitt, "TUnfold: an algorithm for correcting migration effects in high energy  
648 physics", *JINST* **7** (2012) T10003, doi:10.1088/1748-0221/7/10/T10003,  
649 arXiv:1205.6201.
- 650 [52] C. Hansen, "Analysis of discrete ill-posed problems by means of the l-curve", *SIAM  
651 Review Vol.34, No.4, pp. 561-580* (1992).

- 652 [53] J. H. Friedman, "Stochastic gradient boosting", *Comput. Stat. Data Anal.* **38** (2002)  
653 367–378, doi:10.1016/S0167-9473(01)00065-2.
- 654 [54] R. E. Schapire, "The strength of weak learnability", *Machine Learning* **5** (1990), no. 2,  
655 197–227, doi:10.1007/BF00116037.
- 656 [55] H. Voss, A. Höcker, J. Stelzer, and F. Tegenfeldt, "TMVA, the toolkit for multivariate data  
657 analysis with ROOT", in *XI International Workshop on Advanced Computing and Analysis  
658 Techniques in Physics Research*, p. 040. SISSA, 2007. arXiv:physics/0703039.  
659 PoS(ACAT2007)040.
- 660 [56] <https://cms-hcomb.gitbooks.io/combine/content/part4/>.
- 661 [57] S. Schmitt, "Tunfold: an algorithm for correcting migration effects in high energy  
662 physics", doi:10.1088/1748-0221/7/10/T10003, arXiv:1205.6201.
- 663 [58] CMS Collaboration Collaboration, "Measurement of differential cross sections and  
664 charge ratios for t-channel single top quark production at 13 TeV", Technical Report  
665 CMS-PAS-TOP-17-023, CERN, Geneva, 2019.
- 666 [59] CMS Collaboration Collaboration, "Search for heavy Higgs bosons decaying to a top  
667 quark pair in proton-proton collisions at  $\sqrt{s} = 13$  TeV", Technical Report  
668 CMS-PAS-HIG-17-027, CERN, Geneva, 2019.

DRAFT

## 669 A Multivariate analysis method

670 Beginning from the same event selection (except from the veto of loose jets: here we do not  
 671 apply it) as in the main method described in the previous sections, an alternative way of per-  
 672 forming the step of signal extraction (as well as the treatment of the uncertainties) was also  
 673 developed based on multivariate analysis (MVA) techniques. This different approach is de-  
 674 scribed in this appendix.

### 675 A.1 Event selection

676 The only difference in the previous steps to the signal extraction is the definition of the signal  
 677 region, that here looses the requirement of not having any loose ( $20 < p_T < 30$ ) jet. This yields  
 678 to the exact same signal region (of only one jet that must be b-tagged, i.e. 1j1t) as used in the  
 679 inclusive cross section measurement [19]. In the unfolded space results from this alternative  
 680 method, we have chosen the following variables to measure the differential cross section of this  
 681 process depending on them.

- 682 •  $p_T(\ell_1)$ : the transverse momentum of the lepton in the event with the highest one.
- 683 •  $p_Z(\ell_1, \ell_2, j)$ : the momentum in the Z axis (the one of the pipe of the LHC) of the  
 684 system formed by the two leptons and the jet.
- 685 •  $\Delta\varphi(\ell_1, \ell_2)$ : the difference in the  $\varphi$  angle between the two leptons.
- 686 •  $m(\ell_1, j)$ : the invariant mass of the system of the jet ( $j$ ) and the lepton with the highest  
 687  $p_T(\ell_1)$ .

688 In Fig. 37 the distributions of the selected variables in the quoted 1j1t region are shown. The  
 689 agreement between data and simulations is fairly good (36606 observed events and 37736 sim-  
 690 ultated events), except from the case of  $p_T(\ell_1)$ , where a slightly decreasing trend is observed.

#### 691 A.1.1 Uncertainties

692 The same uncertainties that are taken into account in the main method of the analysis, de-  
 693 scribed in Sect. 4, are considered also in this alternative approach.

## 694 A.2 Signal extraction

### 695 A.2.1 Boosted decision tree (BDT)

696 In this method, a boosting decision tree (BDT) [53] is trained to separate signal tW events  
 697 from those that come from the  $t\bar{t}$  background. This technique creates a collection of decision  
 698 trees that weakly separates events into signal and background based on a number of binary  
 699 decisions considering a single observable at a time. A boosting algorithm is then used to assign  
 700 weights to each tree such that the ensemble of weak classifiers performs as a strong classifier  
 701 [54]. In this analysis, the BDT implementation is provided by the TMVA package [55], using the  
 702 GradientBoost algorithm.

703 The training as well as the details of the samples used are inherited also from the inclusive  
 704 cross section measurement of this very same process [19] for the one jet one b-tagged jet (1j1t)  
 705 region. In a nutshell, we use dedicated samples (different from the ones used for the signal  
 706 extraction) to train them. The input variables that are used are reproduced in the following list.

- 707 • Number of loose jets (i.e.  $20 < p_T < 30$  GeV,  $|\eta| < 2.4$ ).
- 708 • Number of b-tagged loose jets ( $20 < p_T < 30$  GeV,  $|\eta| < 2.4$  and CSVM b-tagged).
- 709 •  $p_T^{sys}$ : vector sum of  $p_T$  of leptons, jet, and  $E_T$ .

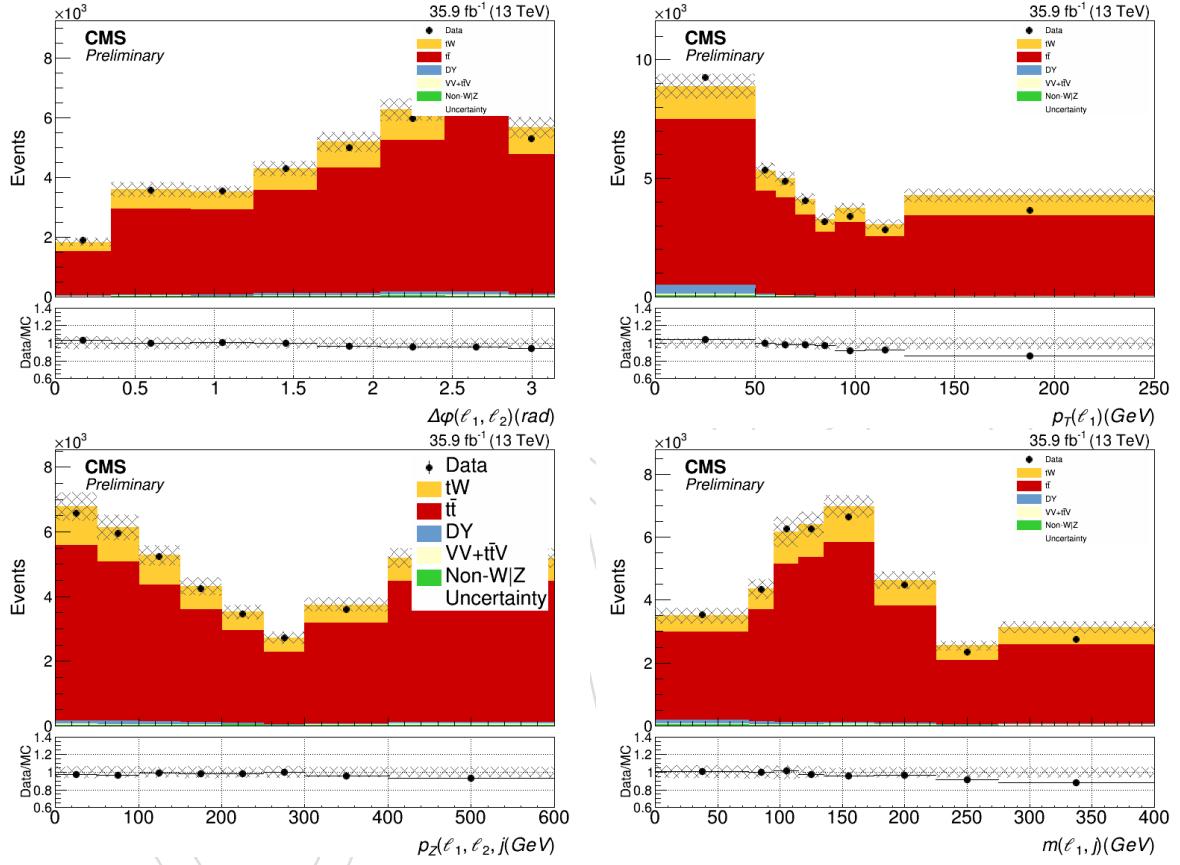


Figure 37: Distribution of the chosen variables to be unfolded. The agreement between data and Monte Carlo is fairly good, taking into account known deviations such as the one from  $p_T(\ell_1)$ . The binning shown is the selected for the unfolding (i.e. the folded space binning).

- $H_T$ : scalar sum of  $p_T$  of leptons, jet, and  $\cancel{E}_T$ .
  - Ratio of  $p_T^{sys}$  to  $H_T$  for the event.
  - $p_T$  of the leading, tight, b-tagged jet.
  - $p_T$  of leading loose jet, defined as 0 for events with no loose jets present.
  - $M_{sys}$ : Invariant mass of the combination of the leptons, jet, and  $\cancel{E}_T$ .
  - Centrality (ratio between the transverse momentum and the total momentum) of the system of the jet and the two leptons.
  - Ratio of scalar sum of  $p_T$  of the leptons to the  $H_T$  of full system.
  - Vector sum of  $p_T$  of jet and leptons.
- The result of the BDT discriminant for all the events of the signal region is reproduced in Fig. 38 (also from [19]).

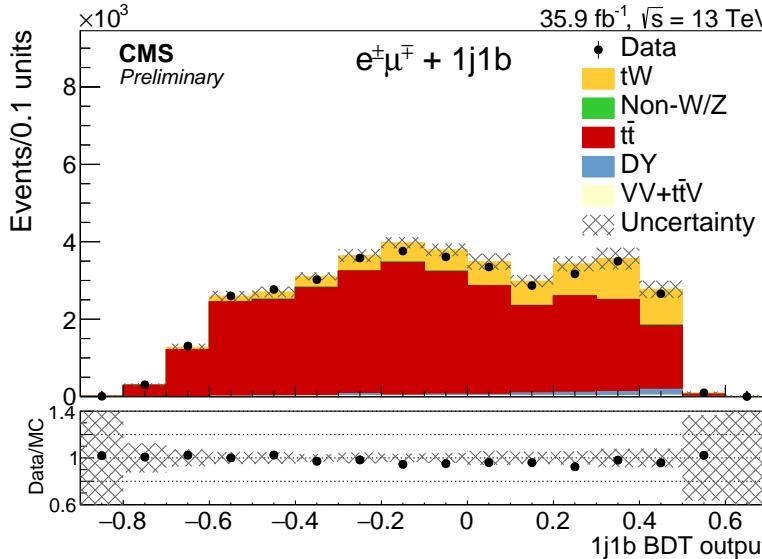


Figure 38: Results of the BDT discriminant for our signal region where the comparison between the signal and the background depending on it can be clearly seen.

720

### 721 A.2.2 Maximum likelihood (ML) fit

722 Due to the overwhelming presence of the  $t\bar{t}$  background, the signal extraction is enhanced by  
 723 exploiting our statistical knowledge of the binned distributions through the use of a maximum  
 724 likelihood fit (MLF). This technique is implemented as follows using the interface to RooStats  
 725 developed by the Higgs Combine Group (see the software's documentation, [56], for more  
 726 information).

727 For each observable (depending on which the diff. cross section will be put) listed in the pre-  
 728 vious section we obtain the distribution of the BDT discriminant for each of the bins of the  
 729 variable (bins that belong to the folded space) with a fixed number of bins for all of them. The  
 730 limits of each bin are chosen so that the amount of the  $t\bar{t}$  background group is equal in all of  
 731 them: this way we expect to slightly enhance the differences in signal between all the bins (that  
 732 must exist by construction of the BDT: from less signal in the negative values, to more in the  
 733 positive). This is done for each of the bins of the variable folded space: once all the distribu-  
 734 tions are done, the fit is performed to all of them at the same time. In the fit, the measured  
 735 differentail cross-section is parametrized as a set of parameters-of-interest (POIs),  $\vec{\mu}$ , that are

736 the signal strength modifiers of the signal in each bin of the measured distribution. These POIs  
 737 are estimators of the signal yield in each bin of the variable's folded space, that allow us to  
 738 extract the signal.

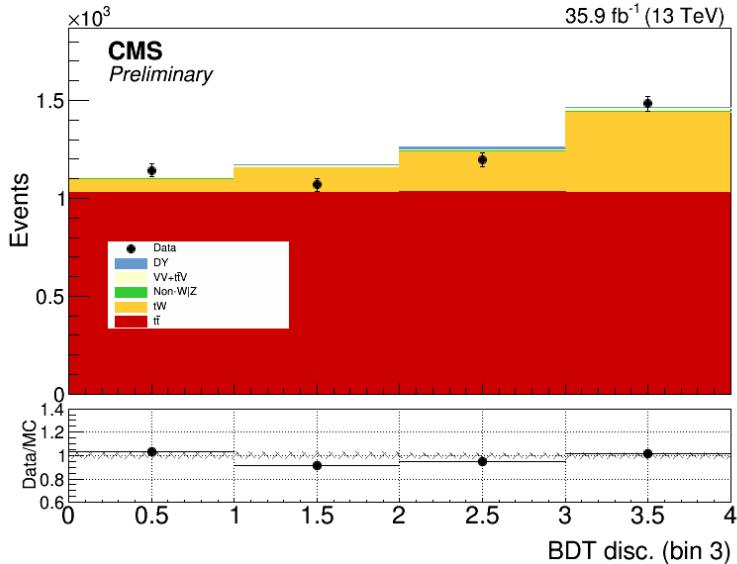


Figure 39: This is an example of the BDT distribution extracted from one particular bin of one variable. In this case, this is the third bin of the distribution of the transverse momentum of the lepton with the highest one: from this bin, the BDT distribution has been extracted as commented before, choosing four bins and an equal content of the  $t\bar{t}$  process (the X axis shows arbitrary values: they are not the values of the discriminant). The uncertainties shown here are only of statistical nature.

739 Apart from the  $\mu$  parameters (one for each bin), all uncertainties are introduced in the fit as nu-  
 740 nance parameters, except from the modelling ones. The modelling uncertainties are propagated  
 741 by taking its varied samples and performing the fit to them, as if it were the nominal ones. It is  
 742 said that the modelling uncertainties are *externalised* (a dedicated fit is performed with the var-  
 743 ied distribution to asses the uncertainty), whereas the others are parametrized with nuisance  
 744 parameters of the fit by using a mixed frequentist-Bayesian likelihood that houses Gaussian  
 745 priors for those *internalised* uncertainties.

746 During the development of this analysis we faced an obstacle, as we observed that when the  
 747 number of bins of the BDT discriminant distribution was changed, large and random variations  
 748 in the predominance of the various modelling uncertainties in the "post-fit" results appeared.  
 749 A careful check in those BDT discriminant distributions convinced us that the problem was  
 750 one of the amount of simulated events in the Monte Carlo samples, as the low quantity of them  
 751 made that the post-fit values had unnatural large or small signal amount because of statistical  
 752 fluctuations in the bins. This forced us to reduce largely the number of them, until we noticed  
 753 that there were not any kind of strange phenomena due to statistical fluctuations. The highest  
 754 number of bins that fulfilled this conditions was four, and it was the chosen one. In Fig. 39 one  
 755 of those distributions is shown.

### 756 A.3 Unfolding

757 After the signal extraction procedure, we perform the unfolding on our variable. To do so, we  
 758 use the TUnfold library [57], that implements the whole procedure in a comfortable way. The  
 759 response matrices are calculated previously from dedicated samples. The exact definition of

760 the fiducial region in the unfolded or particle level space is the following.

- 761 • There must be two leptons that must be an electron and a muon of opposite charge.  
 762 All leptons must have  $p_T > 20$  GeV and  $|\eta| < 2.4$ .
- 763 • One jet with  $p_T > 30$  GeV must be present and it must come from a b quark. The  
 764 identified jets must have  $|\eta| < 2.4$  and must be cleaned of all identified leptons within  
 765 a cone in the  $(\eta, \varphi)$  space of  $\Delta R < 0.4$ .
- 766 • The invariant mass of both leptons must be lower than 20 GeV.
- 767 • The leading lepton (the one with higher  $p_T$ ) must fulfil  $p_T > 25$  GeV.

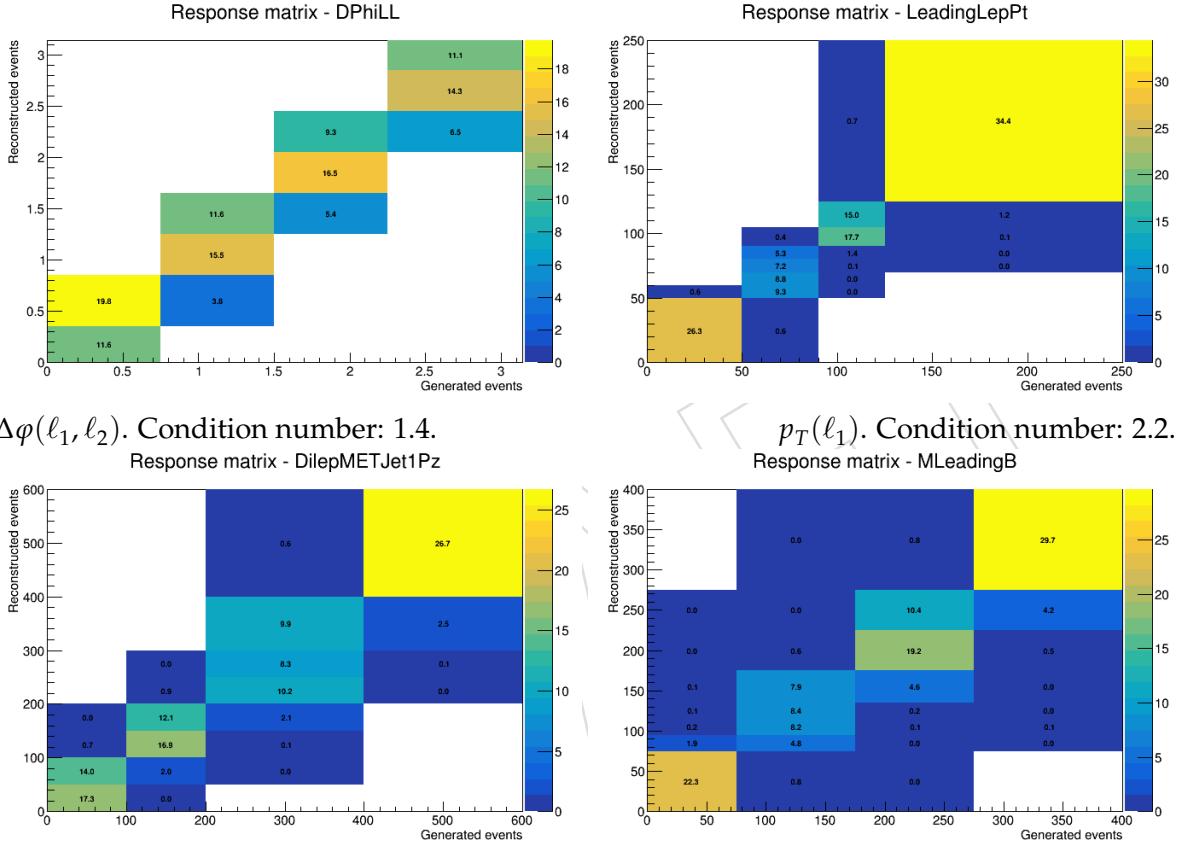


Figure 40: Response matrices of the variables chosen to be unfolded (of its nominal values): the reconstructed events axis is the folded space axis and the generated events axis, the unfolded. The condition numbers of each one are shown, and are all of them of order  $\sim 1$ .

768 The mentioned response matrices can be seen in Fig. 40. The binning in both folded and  
 769 unfolded spaces was optimised by checking the purities and stabilities of each bin. In figure 41  
 770 the plots of both for the chosen distribution appear. These quantities are defined as follows,

$$s_i := \frac{\sum_{j=1}^{N_{\text{bins}}^{\text{fol.}}} n_{ij}}{n_i} \quad p_j := \frac{\sum_{i=1}^{N_{\text{bins}}^{\text{unf.}}} n_{ij}}{n_j^R}, \quad (9)$$

771 where  $n_{ij}$  are the number of reconstructed events whose value of the measured variable fall in  
 772 the folded-space bin  $j$  that had the simulated value of the variable in the unfolded-space bin  $i$ ,

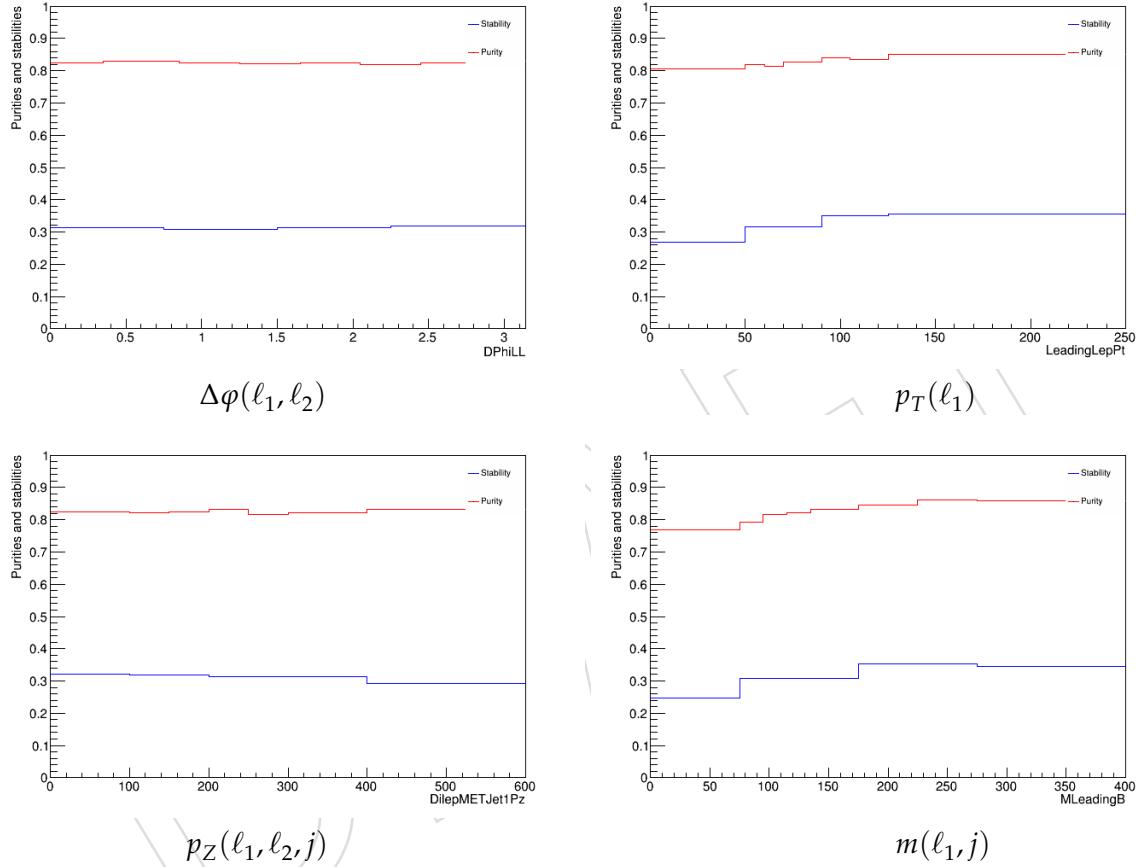


Figure 41: Graphs showing the purities and stabilities of each bin of the folded and unfolded (respectively) spaces for all the distributions chosen to be unfolded. It can be seen a relatively low stability, that is related with a low reconstruction efficiency for the events in our fiducial region.

773  $n_i$  are the number of events whose simulated value of the variable fell in the unfolded space  
 774 bin  $i$  and  $n_j^R$  is the amount of simulated events in the folded space bin  $j$ . Both purities and  
 775 stabilities, when maximised, essentially help enhance the diagonality of the response matrices,  
 776 easing the numerical problems of the unfolding and reducing the need of regularisation.

777 For the variables chosen to be unfolded, no regularisation was needed to perform the unfold-  
 778 ing. This was estimated calculating the condition number of each response matrix, that gave  
 779 us low values (of order  $\sim 1$ ). The numbers for the nominal response matrices are shown also in  
 780 Fig. 40. We also performed scans in the L-curve using Tykhonov regularisation to confirm this,  
 781 under the assumption that if no regularisation was needed, then the values of the  $\tau$  parameter  
 782 should be low, as they were for all the considered variables. The L-curves of all the variables  
 783 are plotted in Fig. 42, and the optimum (by the described criteria) tau parameter value is writ-  
 784 ten in each of them. We do consider an area constraint in this approach, implemented through  
 785 the use of TUnfold.

786 The internalised uncertainties from the signal extraction are propagated to this new fit by im-  
 787 porting the covariance matrix from that fit, whereas the externalised uncertainties (the mod-  
 788elling ones) are propagated by performing the unfolding for each of the variations of the distri-  
 789 butions. There are also other uncertainties related to the unfolding procedure itself. In order to  
 790 estimate the uncertainties in the response matrices themselves, we also calculate the response  
 791 matrices when varying each distribution by the source of uncertainty (e.g. we also calculate  
 792 a response matrix with the varied distributions of the jet energy scale). The uncertainties due  
 793 to the statistics of the sample that is used to obtain the response matrix are taken into consid-  
 794 eration in the procedure and added to the final uncertainty. To asses that our model does not  
 795 bias us deeply, we compare the final results with the generation values from simulations of  
 796 different software: in our case, we use tW samples of the MG5\_AMC@NLO generator as well  
 797 as our nominal POWHEG (using PYTHIA in addition of these two to model the parton shower).  
 798 As no regularisation is imposed, no uncertainties concerning it are needed.

799 At the end, the uncertainties that came from the signal extraction, in addition of those of the  
 800 unfolding procedure itself, are grouped and joint in the covariance matrix of the fit, which  
 801 will be shown in the results plots as one group called “Fit”. The other uncertainties, that were  
 802 externalised (the modelling ones) are presented and propagated asymmetrically for each bin  
 803 of the unfolded space of the variable. The total uncertainty in each of them is obtained by the  
 804 quadratic sum of all of them.

805 Closure checks done comparing with simulations and with the results (folded and unfolded)  
 806 have been done using the Asimov dataset. Other check done to the unfolding procedure was  
 807 done by checking the response to smearing the Asimov dataset and applying then the unfold-  
 808 ing, obtaining values not very far from the particle level information.

#### 809 A.4 Results

810 In Figs. 43 and 44 we show the smeared results or the folded space results, already in terms  
 811 of differential cross section expressed in picobarns, as well as a graph where a comparison  
 812 between some of all the uncertainty sources is presented (in addition to the total one), as de-  
 813 scribed in the figures’ caption. These uncertainties are presented asymmetrically in each bin. In  
 814 addition, and as a check, the comparison with the information of two Monte Carlo generators  
 815 is shown. It can be seen that the results fit with both generators within the uncertainties ( $\pm 1\sigma$ ),  
 816 or near them, with the exception of the  $p_T$  of the lepton with highest  $p_T$ . In some distributions,  
 817 such as  $\Delta\phi(\ell_1, \ell_2)$ , a phenomenon of small statistical fluctuations can be seen in the relative un-  
 818 certainties. Overall, although there is agreement with the Monte Carlo generators, the global

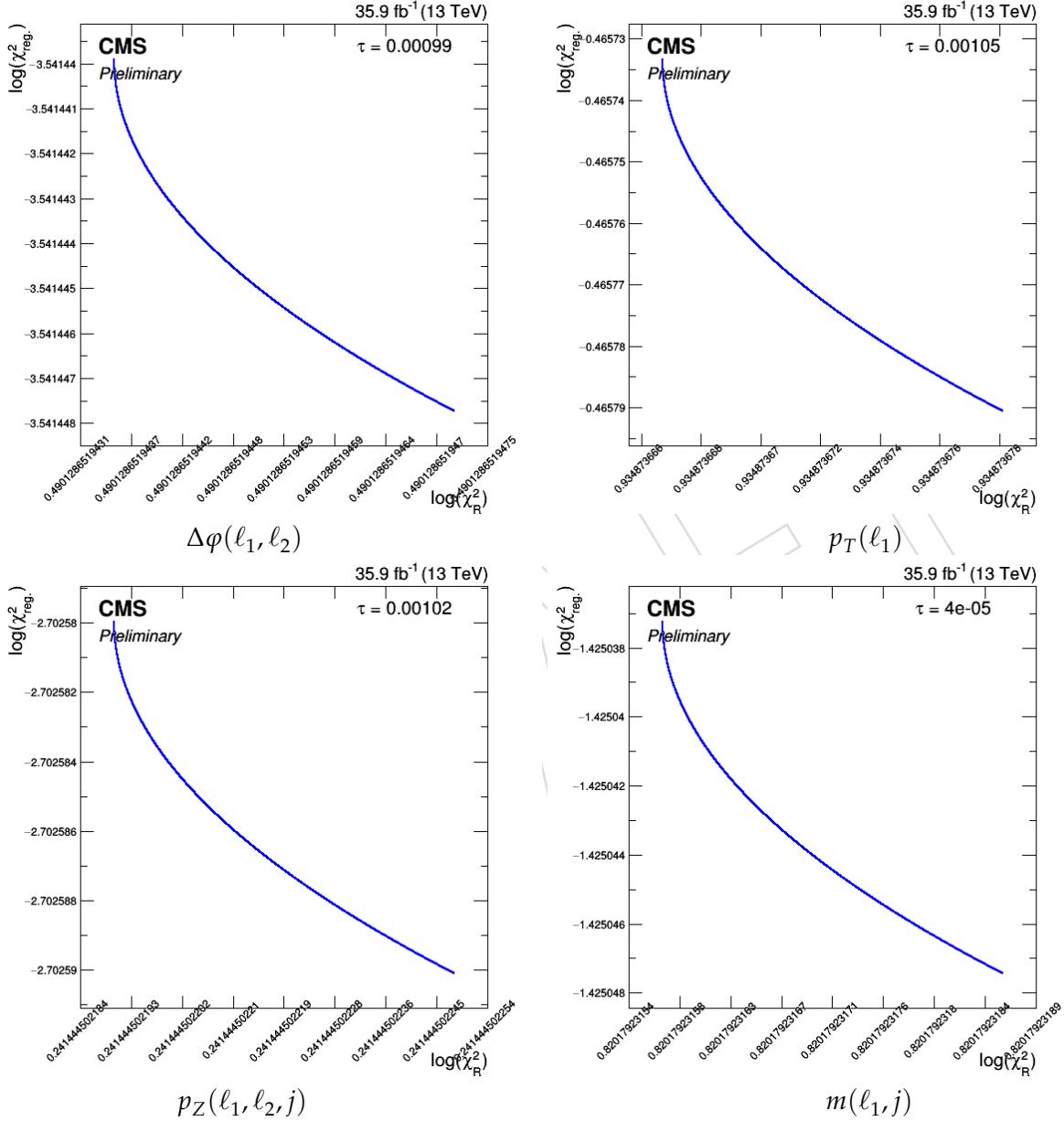


Figure 42: L-Curves of the variables chosen to be unfolded (of its nominal values). The optimum value of the tau parameter (the one corresponding to the point of maximum curvature) is also presented.

819 uncertainties are moderately large. The main sources are in general common for all, and are  
820 some that we can expect: the final and/or initial state radiation or the grouped item “fit”, which  
821 encompasses uncertainties such as the jet energy scale or jet energy resolution. This can be un-  
822 derstood as our selection of events chooses strictly one jet that also must be b-tagged: thus,  
823 uncertainties related with the jets (usually higher than those related with leptons) are expected  
824 to be predominant.

825 Regarding the last step, the unfolding, we found that no regularisation at all was needed as  
826 explained in the previous section. After the unfolding procedure, we obtain the results of  
827 figures 45 and 46. Again, the check after comparing with the particle level information from  
828 the Monte Carlo simulations is relatively good, with the exception of the transverse momentum  
829 of  $\ell_1$ . The uncertainties, asymmetrically presented again, are overall large, even reaching more  
830 than the 100% of the nominal value in some cases. the main sources of the are more or less  
831 the same as for the post-fit results in the folded space, although a bit larger as expected after  
832 performing the unfolding.

### 833 Discussion

834 In this appendix we have showed the unfolded results of a measurement of the differential  
835 cross section depending on the transverse momentum of the lepton in the event with the high-  
836 est one, the momentum in the Z axis (the one of the pipe of the LHC) of the system formed by  
837 the two leptons and the jet, the difference in the  $\varphi$  angle between the two leptons, the invariant  
838 mass of the system of the jet ( $j$ ) and the lepton with the highest  $p_T$  ( $\ell_1$ ) and the invariant mass  
839 of the system of the jet ( $j$ ) and the lepton with the lowest  $p_T$  ( $\ell_2$ ).

840 The closure checks done by comparing with the particle level information from simulations,  
841 and other tests performed to it were successful. The final results in the selected variables are  
842 found to be in agreement with predictions, specifically with the two generators (POWHEG and  
843 aMC@NLO) that we have considered. The main sources of uncertainties are those concerning  
844 the characteristics of the jets, as well as those related with the predominant background: the  
845 pair production of top quarks. The most relevant drawback of this approach is that the global  
846 uncertainties are large, as can be seen in Figs. 45 and 46. This also makes us unable to give  
847 preference both physic models used to compare with this alternative approach’s results.

848 We consider that this method to measure the differential cross section might be a priori offer  
849 a more beneficial signal extraction, due to the huge  $t\bar{t}$  background and the use of ML fits and  
850 MVA techniques. However, there are issues regarding the statistics of the Monte Carlo samples,  
851 already mentioned, that forced us to reduce the sensitivity by reducing the number of bins of  
852 the BDT discriminant distribution. As a result, a good part of the potential of this approach is  
853 lost, as the comparison between these results and the ones from the main analysis easily show.

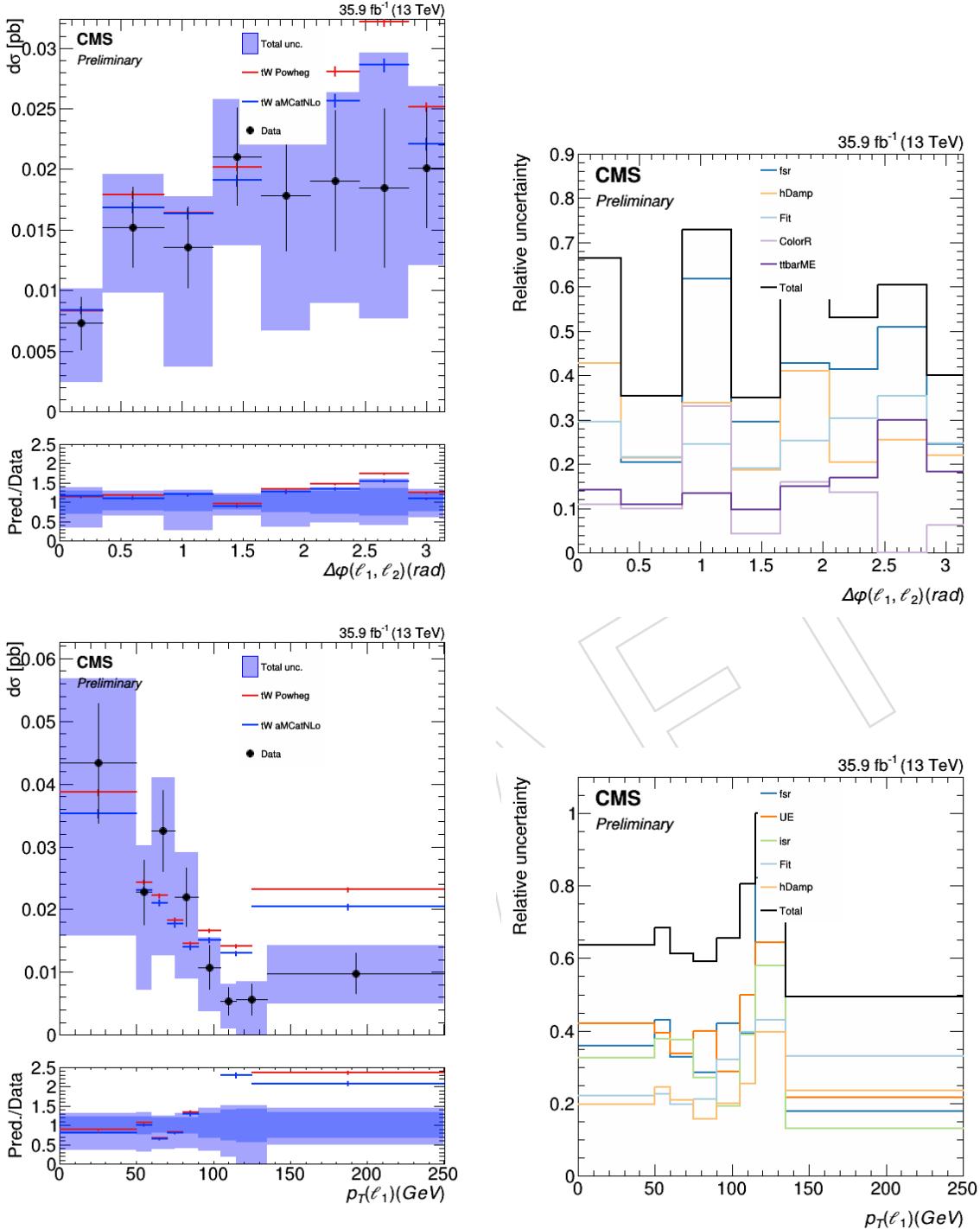


Figure 43: Distributions of  $\Delta\phi(\ell_1, \ell_2)$  and  $p_T(\ell_1)$  after the maximum likelihood fit (the signal extraction), already expressed as a differential cross section, but in the folded space. For each variable the first plot shows the results themselves, with the total uncertainty. In the ratio plot, the group of uncertainties “Fit” (that represent those uncertainties carried during the ML fit) are shown as a lighter blue, as well as the total in the same colour as in the main plot. The other graph represents, in relative terms (respect the nominal values), the five sources of uncertainty ordered by the maximum uncertainty (taking both variations) in all the bins of the variable, in addition of the total uncertainty. The total uncertainties shown are asymmetrical.

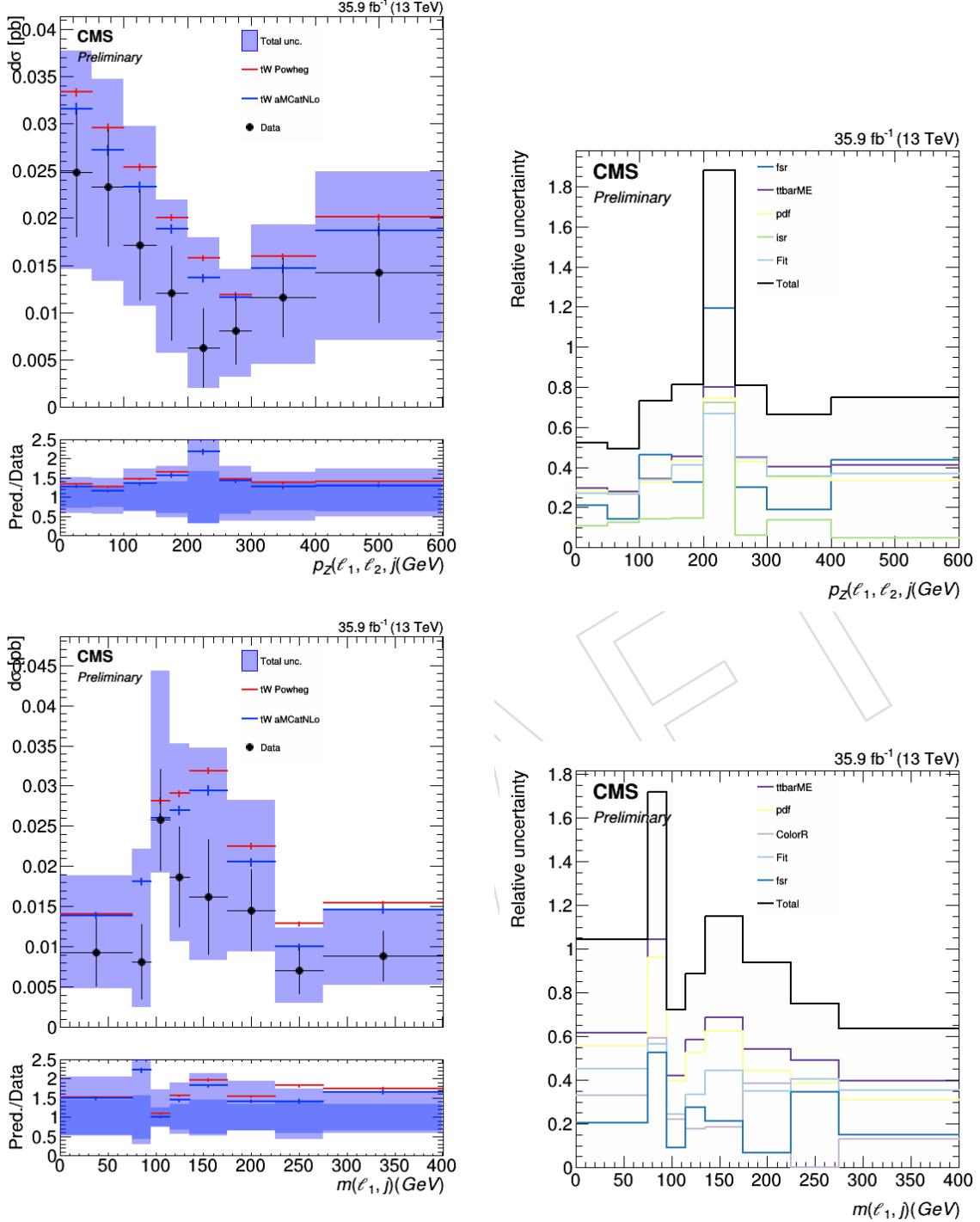


Figure 44: Distributions of  $p_Z(\ell_1, \ell_2, j)$  and  $m(\ell_1, j)$  after the maximum likelihood fit (the signal extraction), already expressed as a differential cross section, but in the folded space. For each variable the first plot shows the results themselves, with the total uncertainty. In the ratio plot, the group of uncertainties “Fit” (that represent those uncertainties carried during the ML fit) are shown as a lighter blue, as well as the total in the same colour as in the main plot. The other graph represents, in relative terms (respect the nominal values), the five sources of uncertainty ordered by the maximum uncertainty (taking both variations) in all the bins of the variable, in addition of the total uncertainty. The total uncertainties shown are asymmetrical.

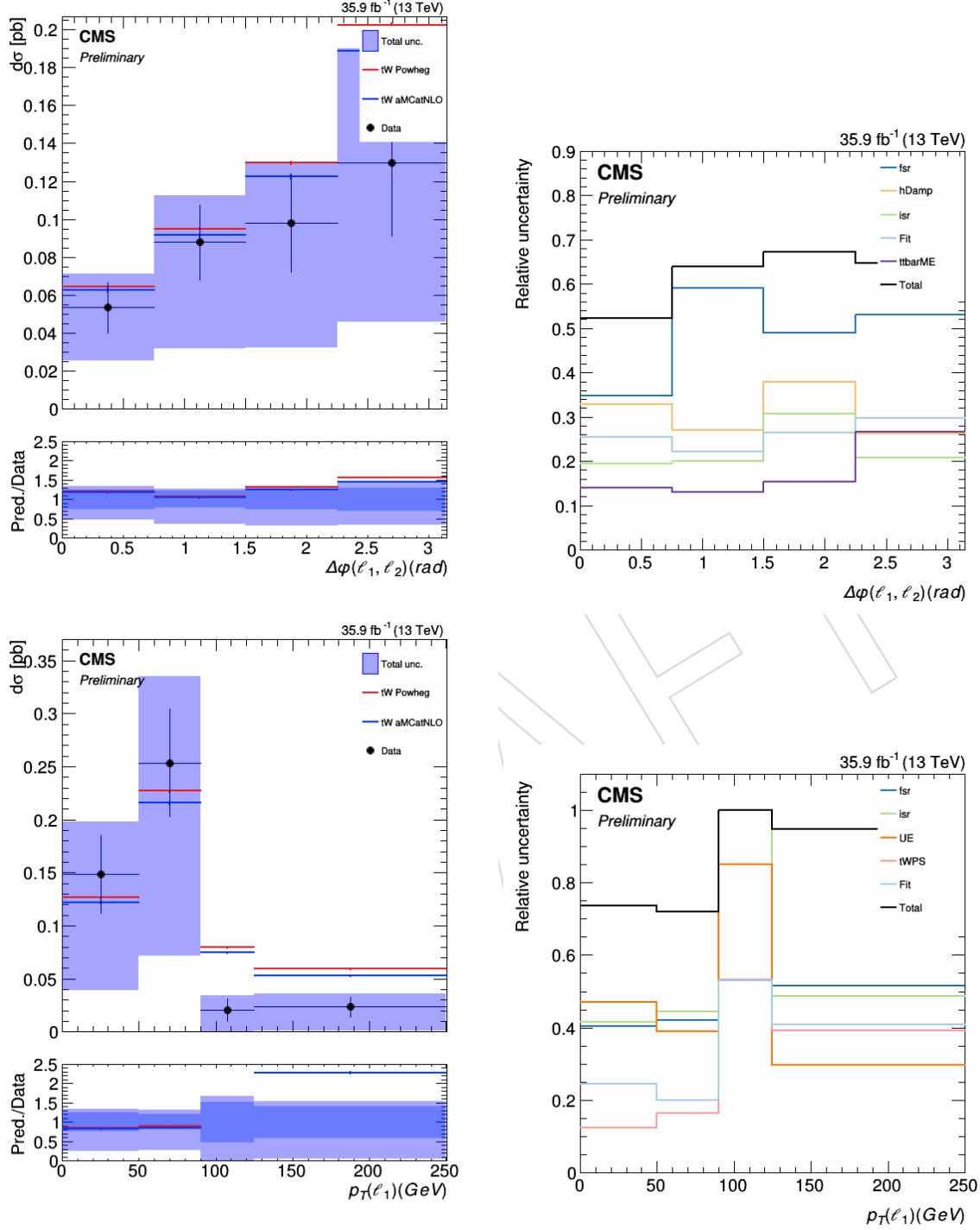


Figure 45: Final measurements of the differential cross sections depending on  $\Delta\phi(\ell_1, \ell_2)$  and  $p_T(\ell_1)$ . As with the post-fit results, for each distribution the first plot shows the results themselves, with the total uncertainty. In the ratio plot, the group of uncertainties “fit” (that represent those uncertainties carried during the fit and also the unfolding) are shown, as well as the total. The other graph represents the same as in the post-fit results. The uncertainties shown are asymmetrical.

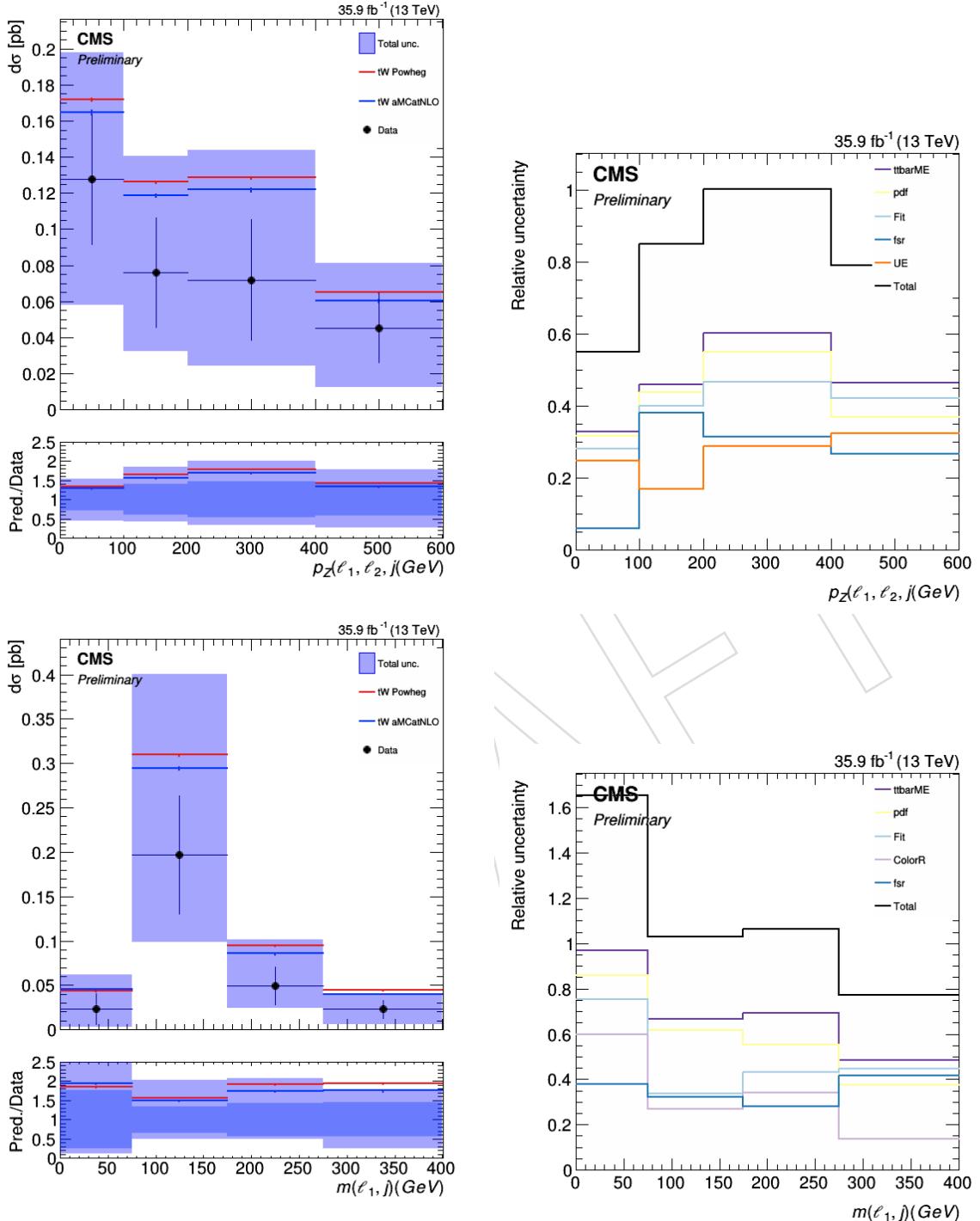


Figure 46: Final measurements of the differential cross sections depending on  $p_Z(\ell_1, \ell_2, j)$  and  $m(\ell_1, j)$ . As with the post-fit results, for each distribution the first plot shows the results themselves, with the total uncertainty. In the ratio plot, the group of uncertainties “fit” (that represent those uncertainties carried during the fit and also the unfolding) are shown, as well as the total. The other graph represents the same as in the post-fit results. The uncertainties shown are asymmetrical.

**A.5 Shortcomings of this approach**

Despite of being a powerful method, that has been applied sucessfully to other analyses, such as [58], the fit method has a few shortcomings due to the low purity of the signal regions, that stems from the difficulty to distinguish the signal from the  $t\bar{t}$  background. Because of this, and since this is the first result delivered by CMS on this channel, the authors have preferred to use the more robust methodology that is described in the body of this document. This section describes the investigations that were made on the stability of the fit.

These investigations were triggered by the large uncertainties observed in the experimental measurements. They are shown in the previously mentioned figures 45 and 46, and are above 50% in all the cases, up to almost 200% in some extreme bins, larger than the cut-and-count approach documented in the body of the analysis note, dominated by the systematic sources that are externalized from the fit. In the following lines, as an example, we study the impact of the PowhegertON variation in the jet  $p_T$  as an example of an uncertainty source that is dominating without a clear physical origin. This study has also been performed with some of the other uncertainties that are externalised, obtaining a similar result.

Figures 47, 48 and 49 show the prefit and postfit distributions of the BDT discriminant for each bin of the detector level (folded space) jet  $p_T$  distribution for the varied distribution to the expected nominal dataset. The three figures indicate that the variation is compatible with the nominal sample within the MC statistical uncertainties. Figure 50 shows the postfit distribution of the nuisances and parameters-of-interest of the fit, showing that the statistical fluctuations of the varied samples are not fully absorbed by the associated nuisance parameters, and that non-negligible pulls are then induced in other nuisance parameters. All these variations are propagated to the final signal strengths (last eight bins of the plots in Fig. 50), which are associated to large uncertainties in the differential cross-section (through the propagation of the uncertainty). As the Monte Carlo samples are consistent within statistical uncertainties, there is no physical origin to these uncertainties. These variations, enhanced by the low purity of the signal region, can also be seen as a clear increase in the  $tW$  contribution when checking the plots of figures 47, 48 and 49 (e.g. check the graphs for the seventh and eighth bins of the detector level distribution of the jet  $p_T$  and compare them between pre-fit and post-fit).

A procedure used in other analysis ([59]), smoothing the variations, is not applicable in our case because the smoothed variations would not give a precise estimate of the total uncertainty, given that the uncertainties that show an unphysical behaviour are the leading ones. We also decided not to include control regions in the fitting procedure (by adding more bins from orthogonal phase space regions) because it is not guaranteed that it would solve the problem for the existing signal region. In addition, this is the first differential cross section measurement in the  $tW$  channel. Thus, a more conservative cut-and-count, with an acceptable performance, was preferred.

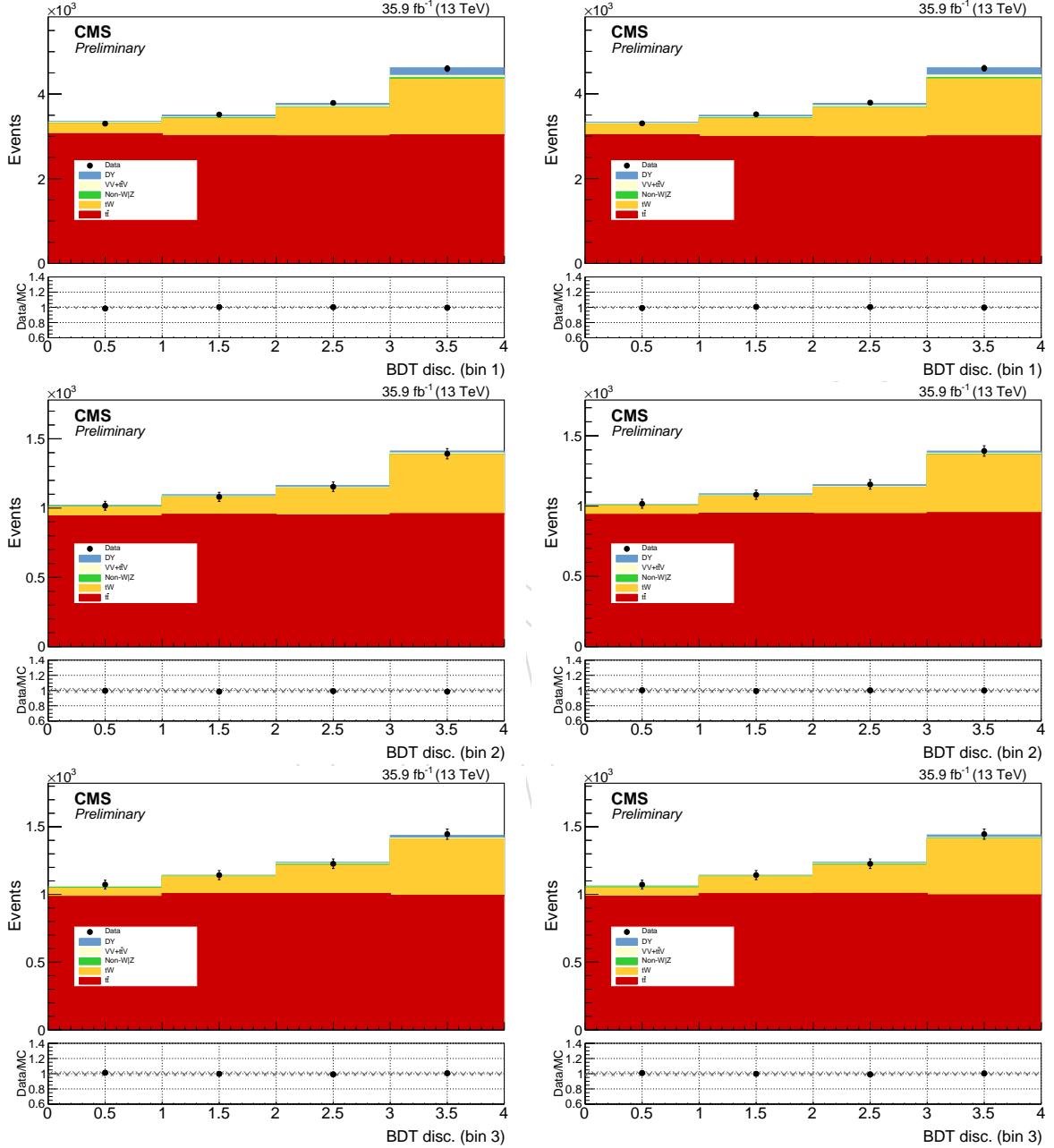


Figure 47: Comparison between the pre-fit (left column of plots) and post-fit (right column) distributions of the BDT discriminant of the contents of the first three bins of the detector level distribution of the jet  $p_T$ .

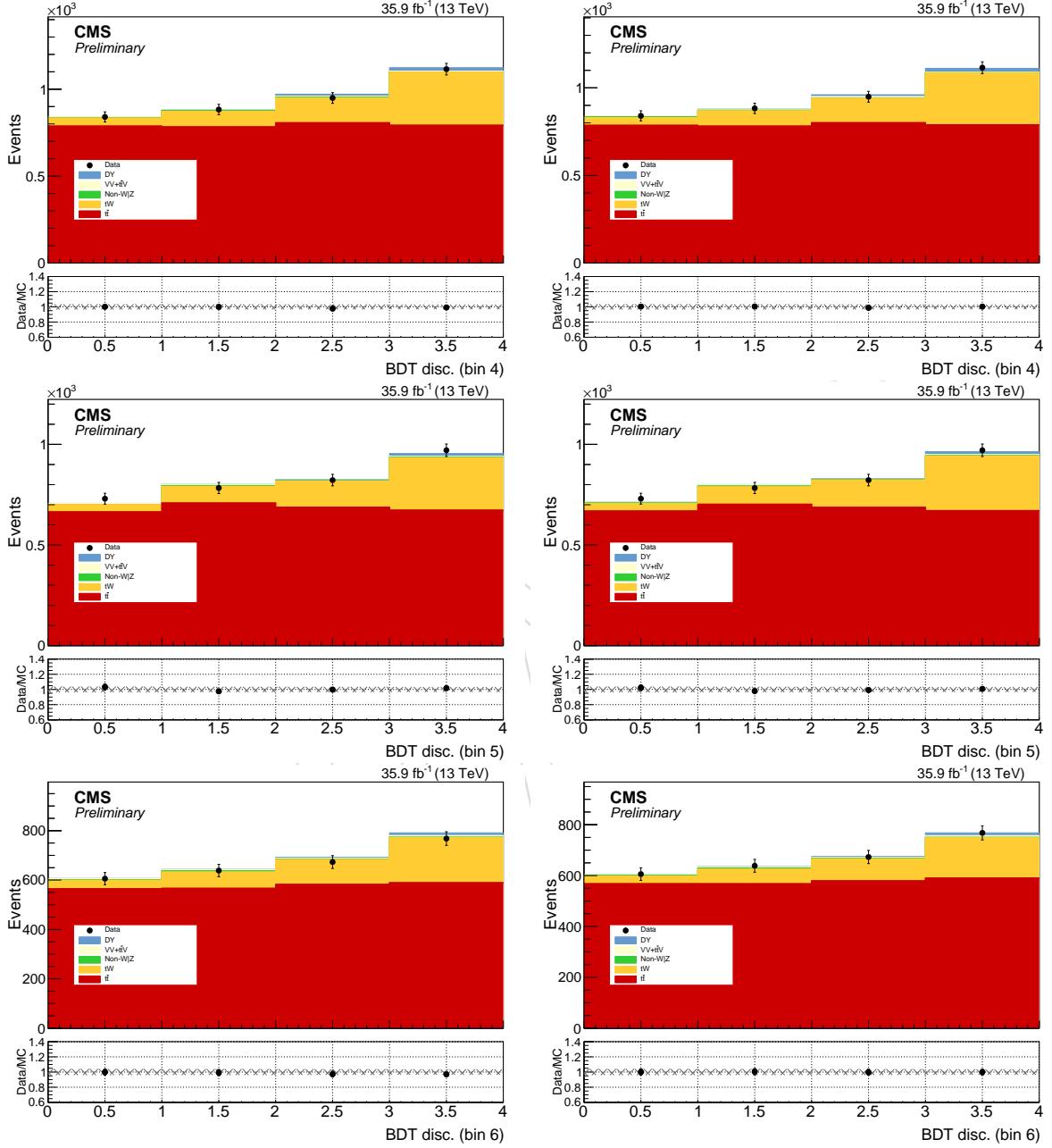


Figure 48: Comparison between the pre-fit (left column of plots) and post-fit (right column) distributions of the BDT discriminant of the contents of the fourth, fifth and sixth bins of the detector level distribution of the jet  $p_T$ .

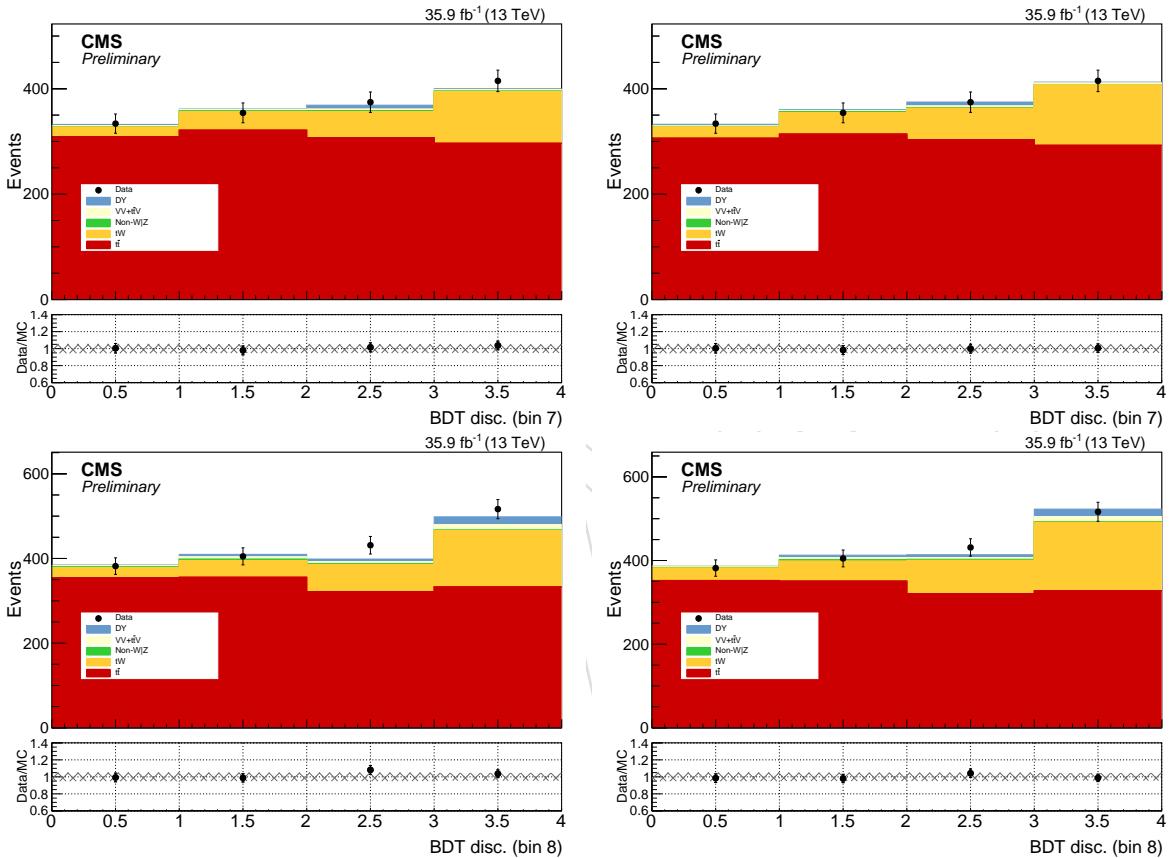


Figure 49: Comparison between the pre-fit (left column of plots) and post-fit (right column) distributions of the BDT discriminant of the contents of the last two bins of the detector level distribution of the jet  $p_T$ .

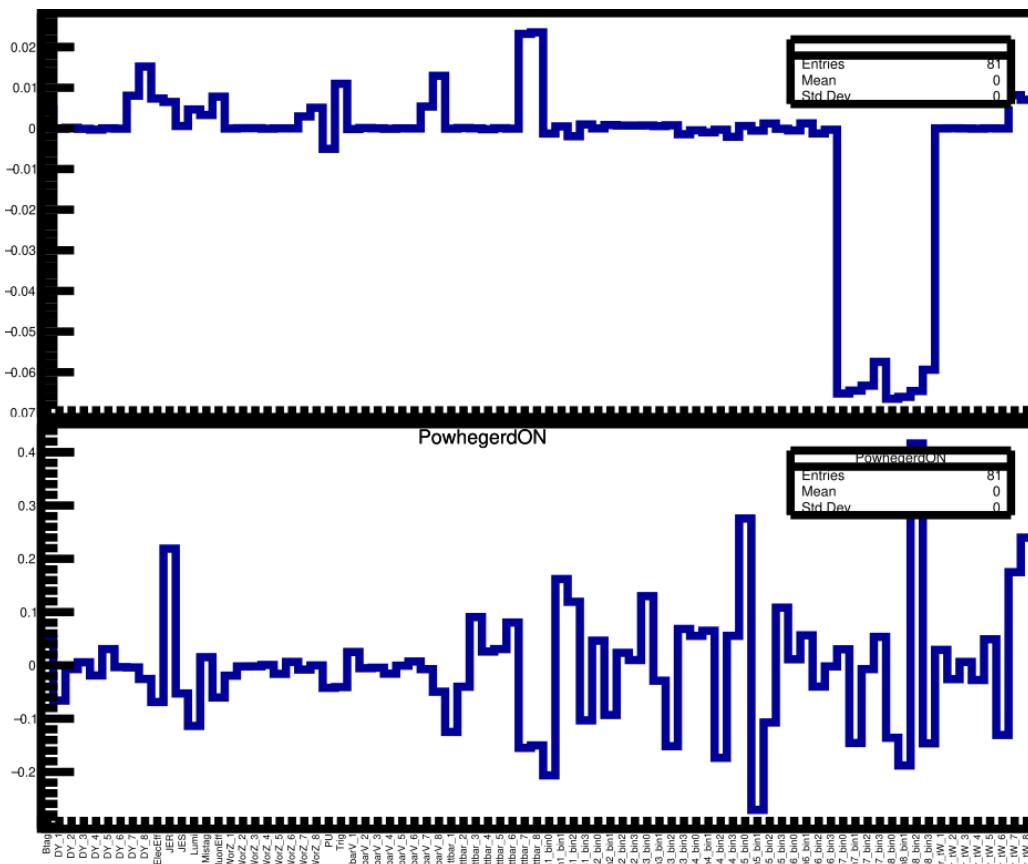


Figure 50: Pulls of the different nuisance parameters of the fit for the case of the nominal fit and the variation of the externalised uncertainty source PowhegON.

---

## 891 **B Choice of binning**

892 In order to study what binning to choose, we decided to plot the absolute results for each  
 893 variable with the Asimov dataset for different number of bins, going from 4 bins to 10 bins.  
 894 The final choice for each variable is the following:

895 **Leading lepton  $p_T$**  8 bins.

896 **Jet  $p_T$**  5 bins.

897  $\Delta\varphi(e^\pm, \mu^\mp)$  6 bins.

898  $p_Z(e^\pm, \mu^\mp, \text{jet})$  8 bins.

899  $m(e^\pm, \mu^\mp, \text{jet})$  6 bins.

900  $m_T(e^\pm, \mu^\mp, \text{jet}, p_T^{\text{miss}})$  5 bins.

901 The procedure for choosing each number of bins was to, using simulations, try to estimate the  
 902 best compromise between a lower (which would lead to a lower number of bins) and “stable”  
 903 (that meaning, without reducing so much the width of bins due to the statistical size of each  
 904 one) uncertainties and the higher resolution that means having more bins.

905 In Figs. 51, 52 (leading lepton  $p_T$ ), 53, 54 (jet  $p_T$ ), 55, 56 ( $\Delta\varphi(e^\pm, \mu^\mp)$ ), 57, 58 ( $p_Z(e^\pm, \mu^\mp, \text{jet})$ ), 59, 60  
 906 ( $m(e^\pm, \mu^\mp, \text{jet})$ ), 61 and 62 ( $m_T(e^\pm, \mu^\mp, \text{jet}, p_T^{\text{miss}})$ ) the absolute results (and relative uncertainties)  
 907 for each variable are shown for binnings 4 to 10.

DRAFT

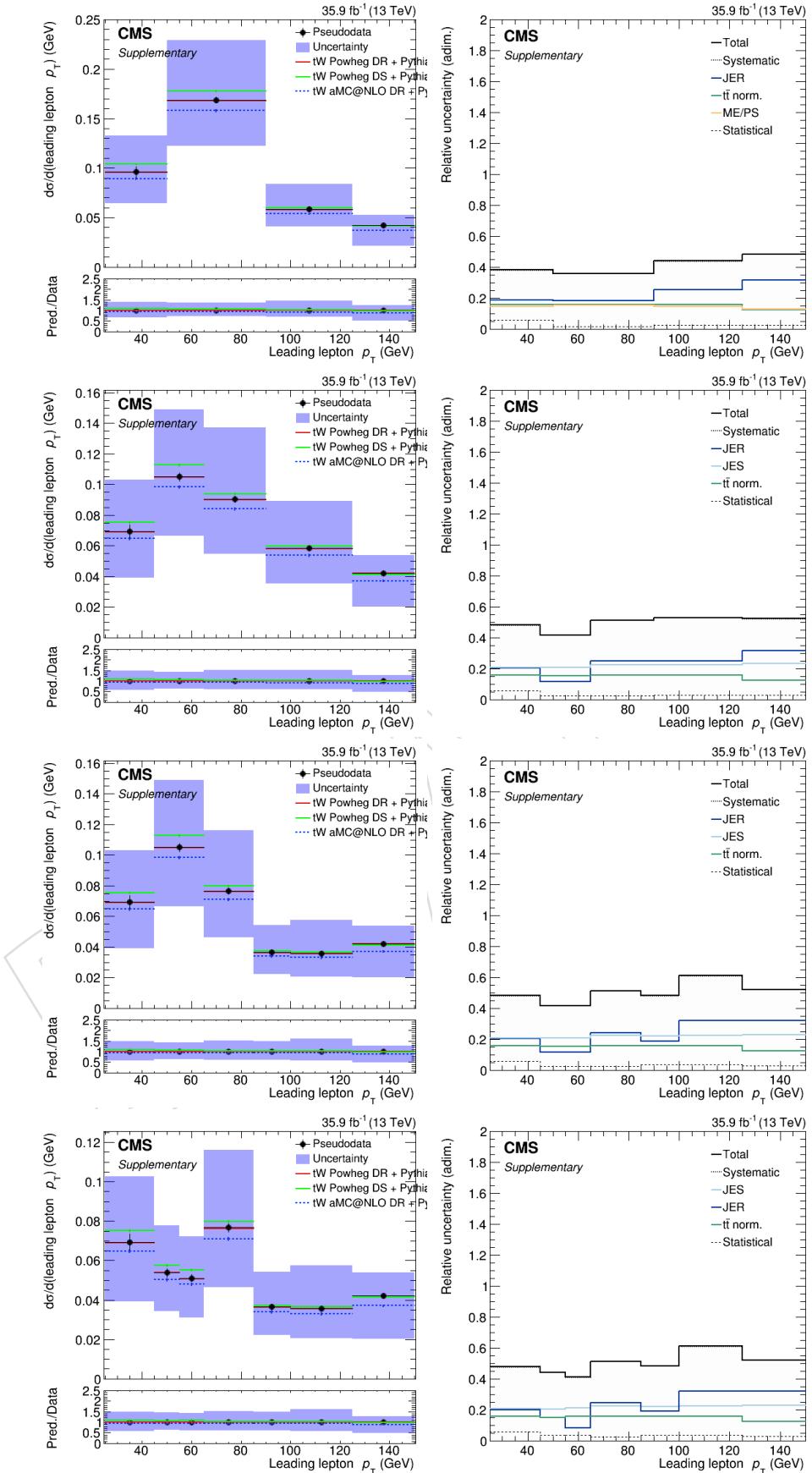


Figure 51: Absolute distributions and relative uncertainties with 4, 5, 6 and 7 bins for the leading lepton  $p_T$ .

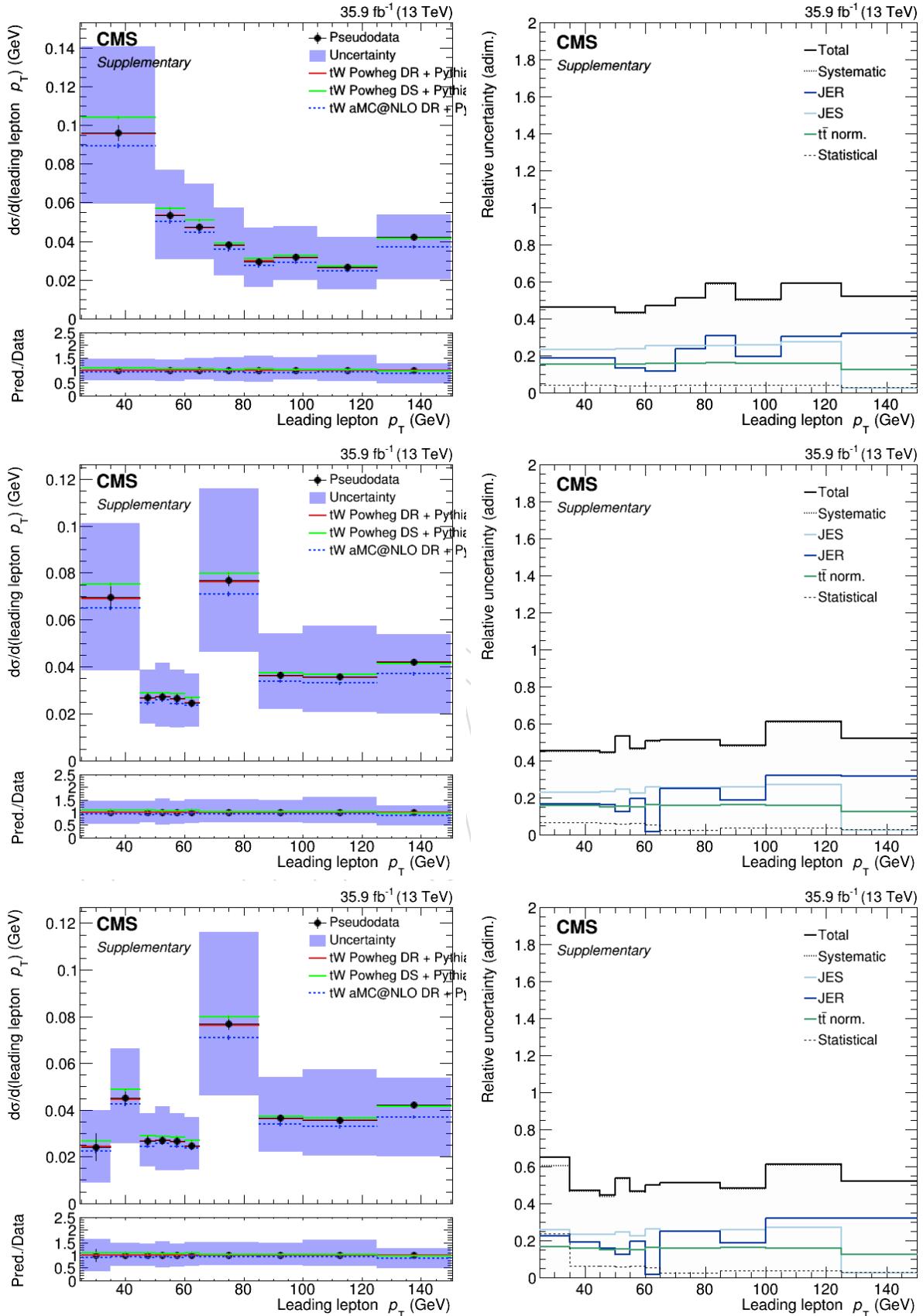
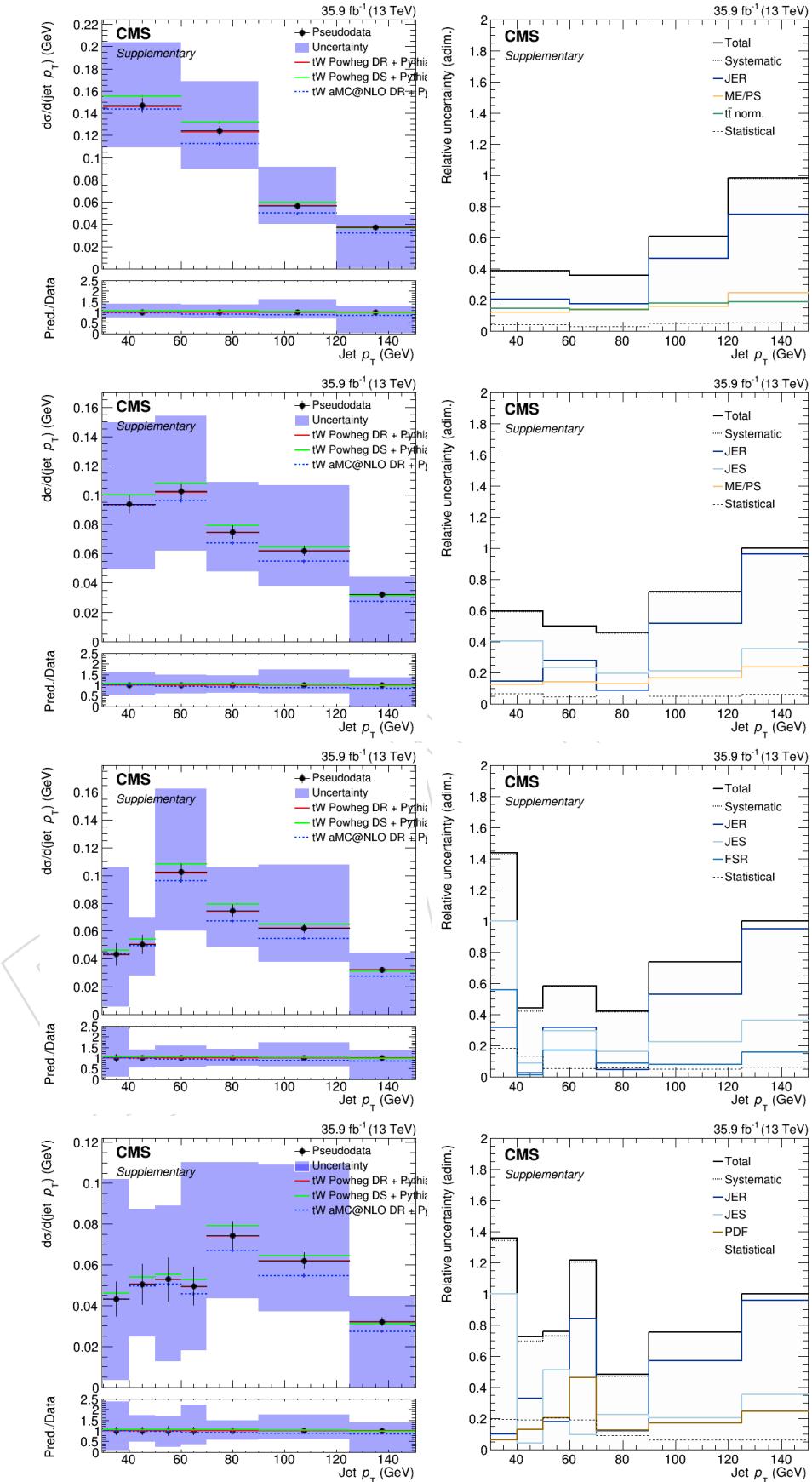


Figure 52: Absolute distributions and relative uncertainties with 8, 9 and 10 bins for the leading lepton  $p_T$ .

Figure 53: Absolute distributions and relative uncertainties with 4, 5, 6 and 7 bins for the jet  $p_T$ .

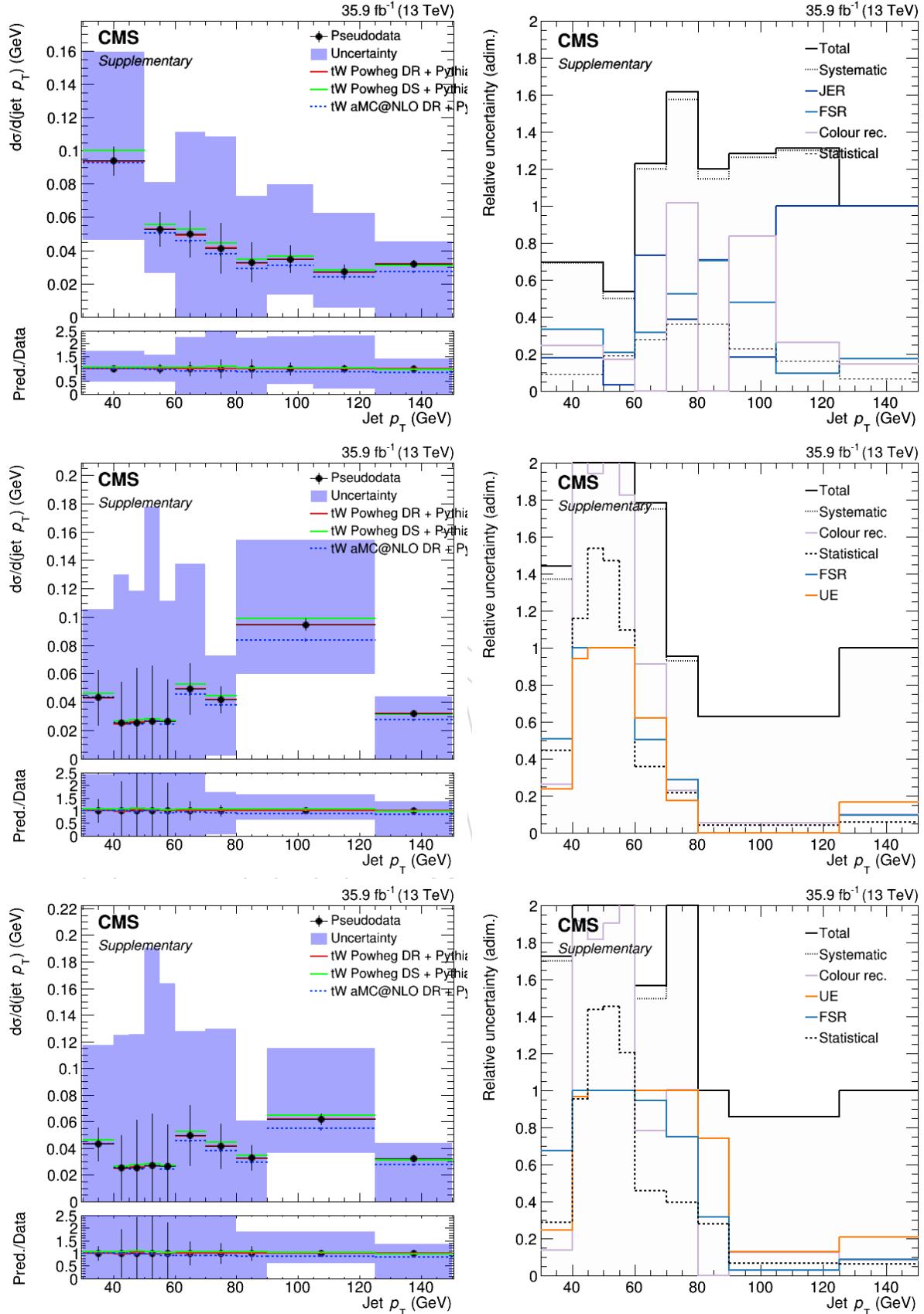


Figure 54: Absolute distributions and relative uncertainties with 8, 9 and 10 bins for the jet  $p_T$ .

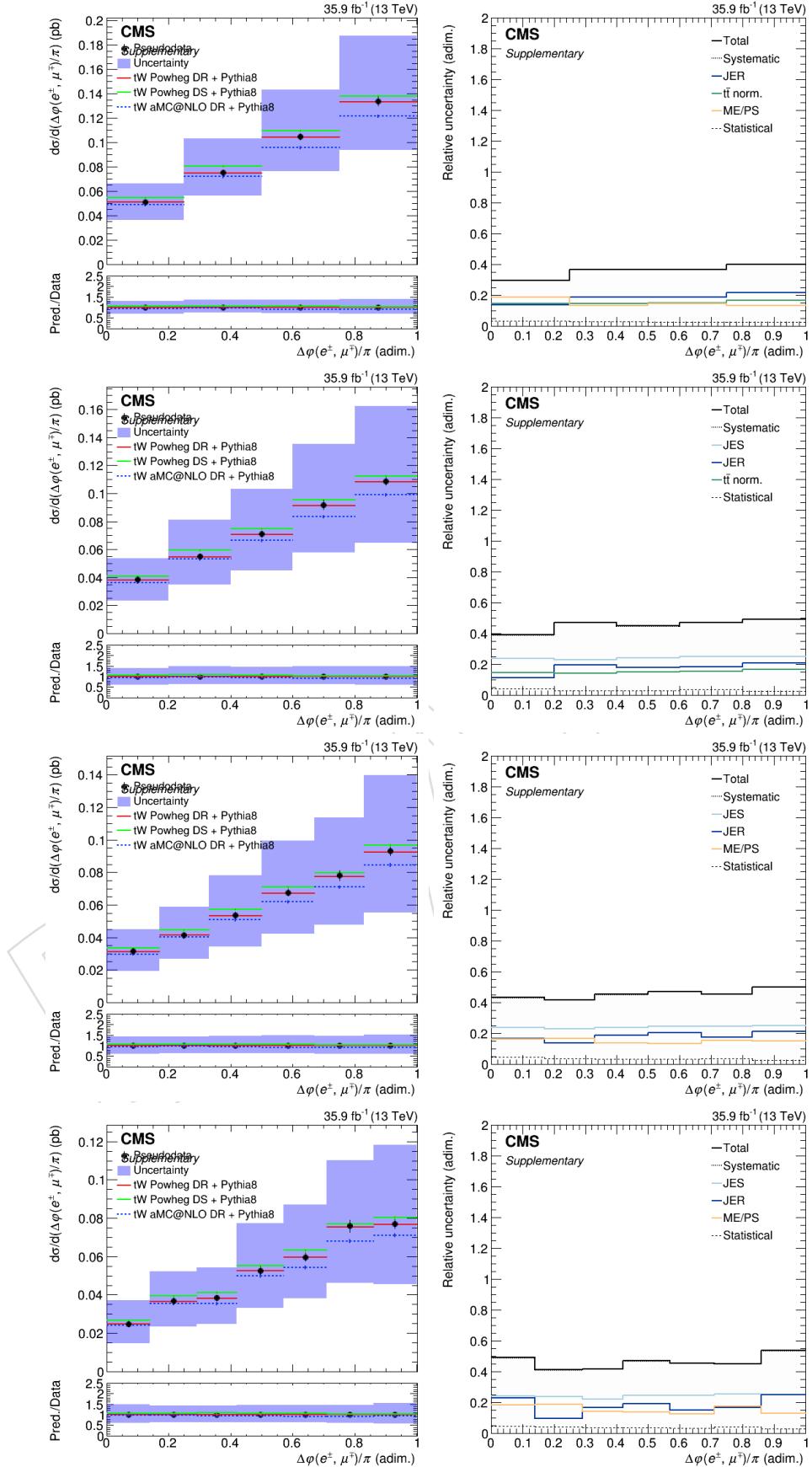


Figure 55: Absolute distributions and relative uncertainties with 4, 5, 6 and 7 bins for  $\Delta\phi(e^\pm, \mu^\mp)$ .

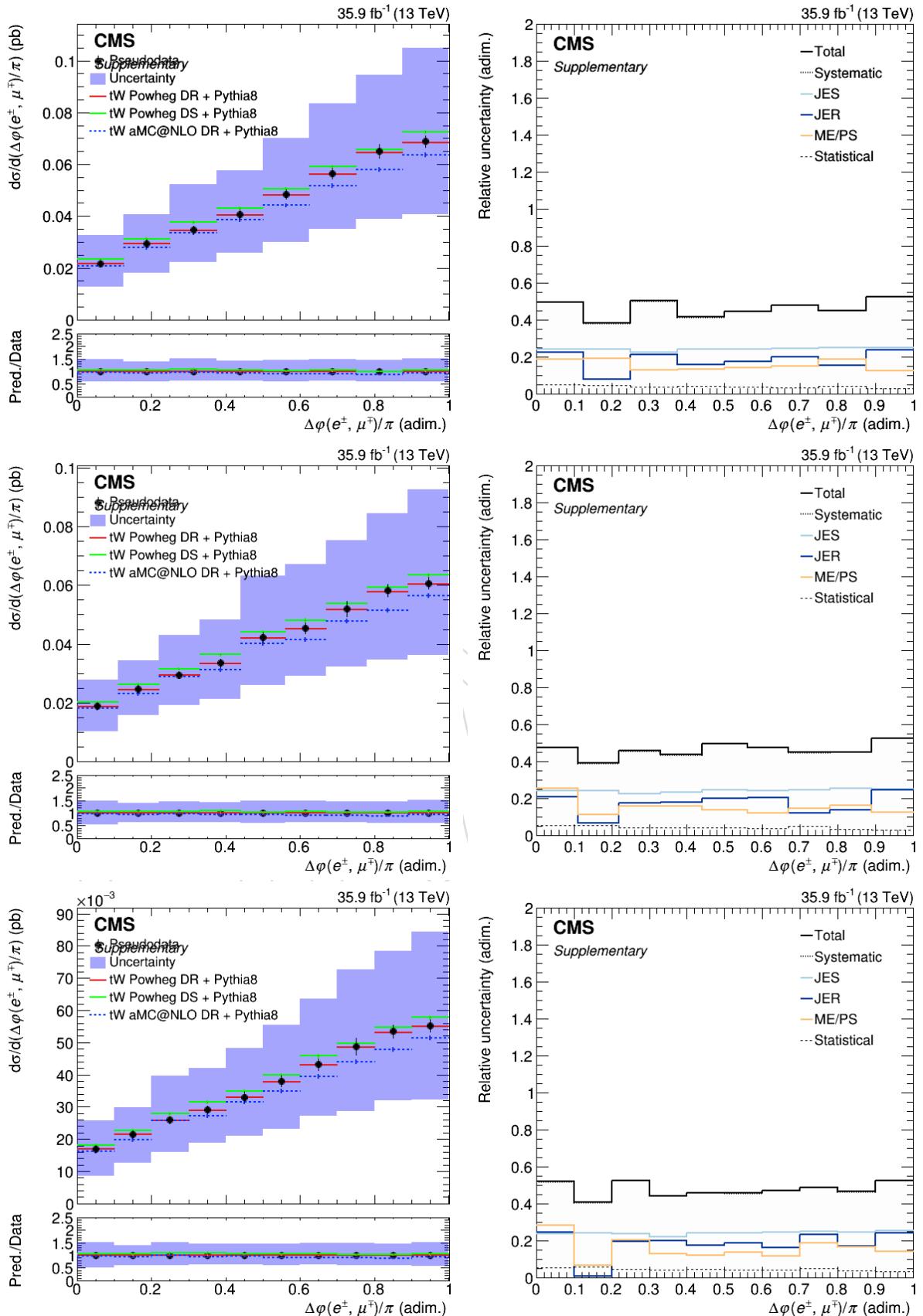


Figure 56: Absolute distributions and relative uncertainties with 8, 9 and 10 bins for  $\Delta\varphi(e^\pm, \mu^\mp)$ .

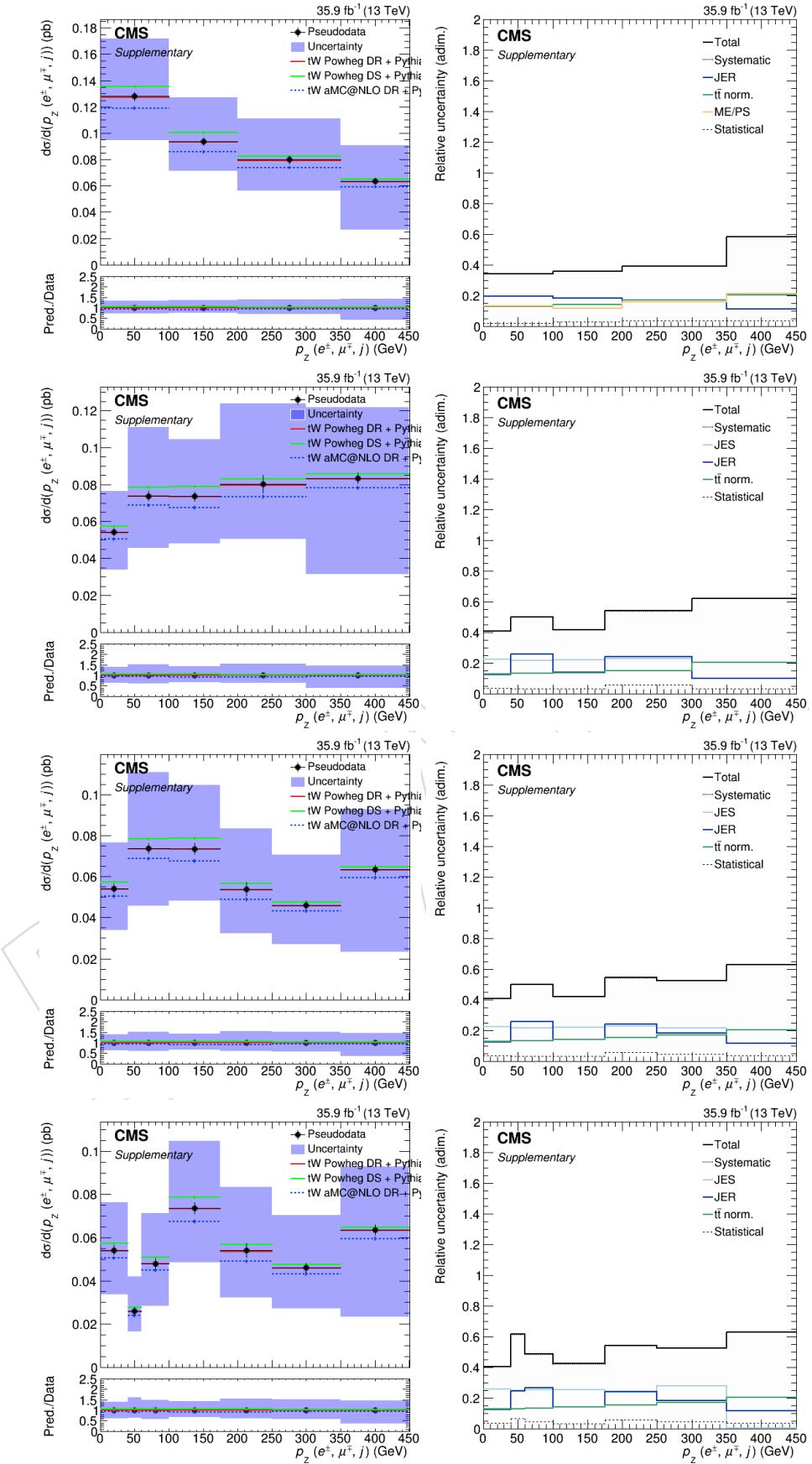


Figure 57: Absolute distributions and relative uncertainties with 4, 5, 6 and 7 bins for  $p_z(e^\pm, \mu^\mp, \text{jet})$ .

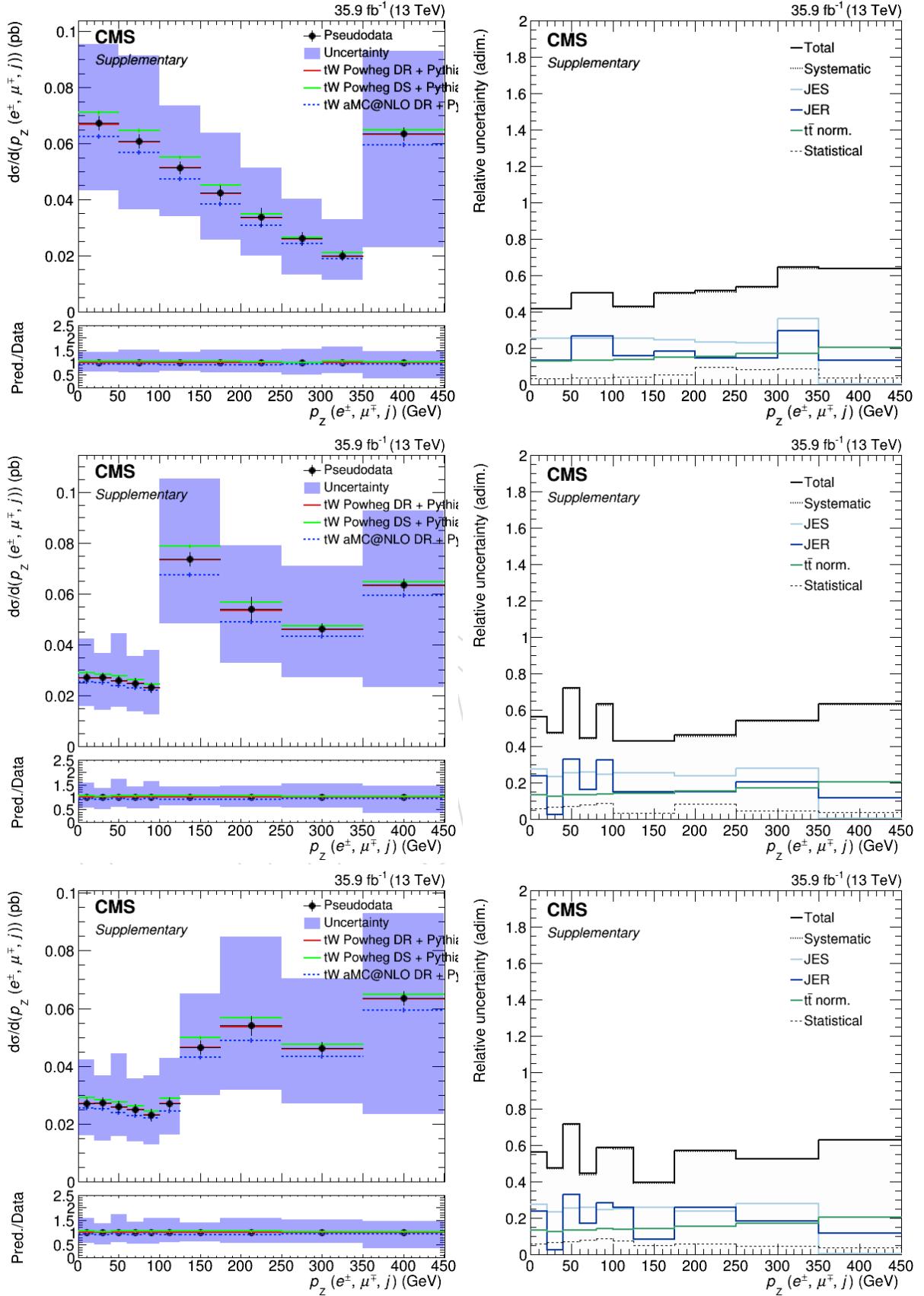


Figure 58: Absolute distributions and relative uncertainties with 8, 9 and 10 bins for  $p_z(e^\pm, \mu^\mp, \text{jet})$ .

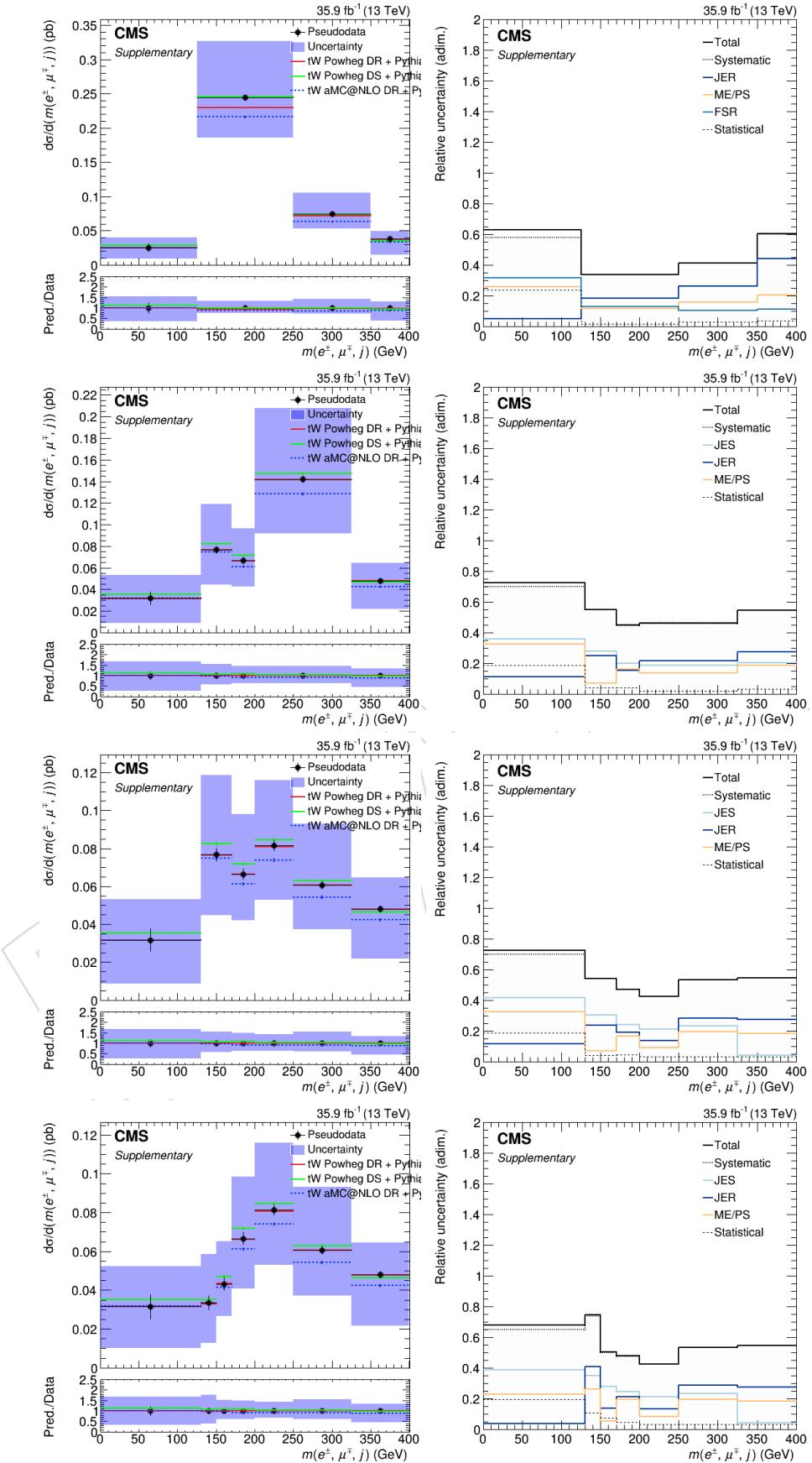


Figure 59: Absolute distributions and relative uncertainties with 4, 5, 6 and 7 bins for  $m(e^+, \mu^-, \text{jet})$ .

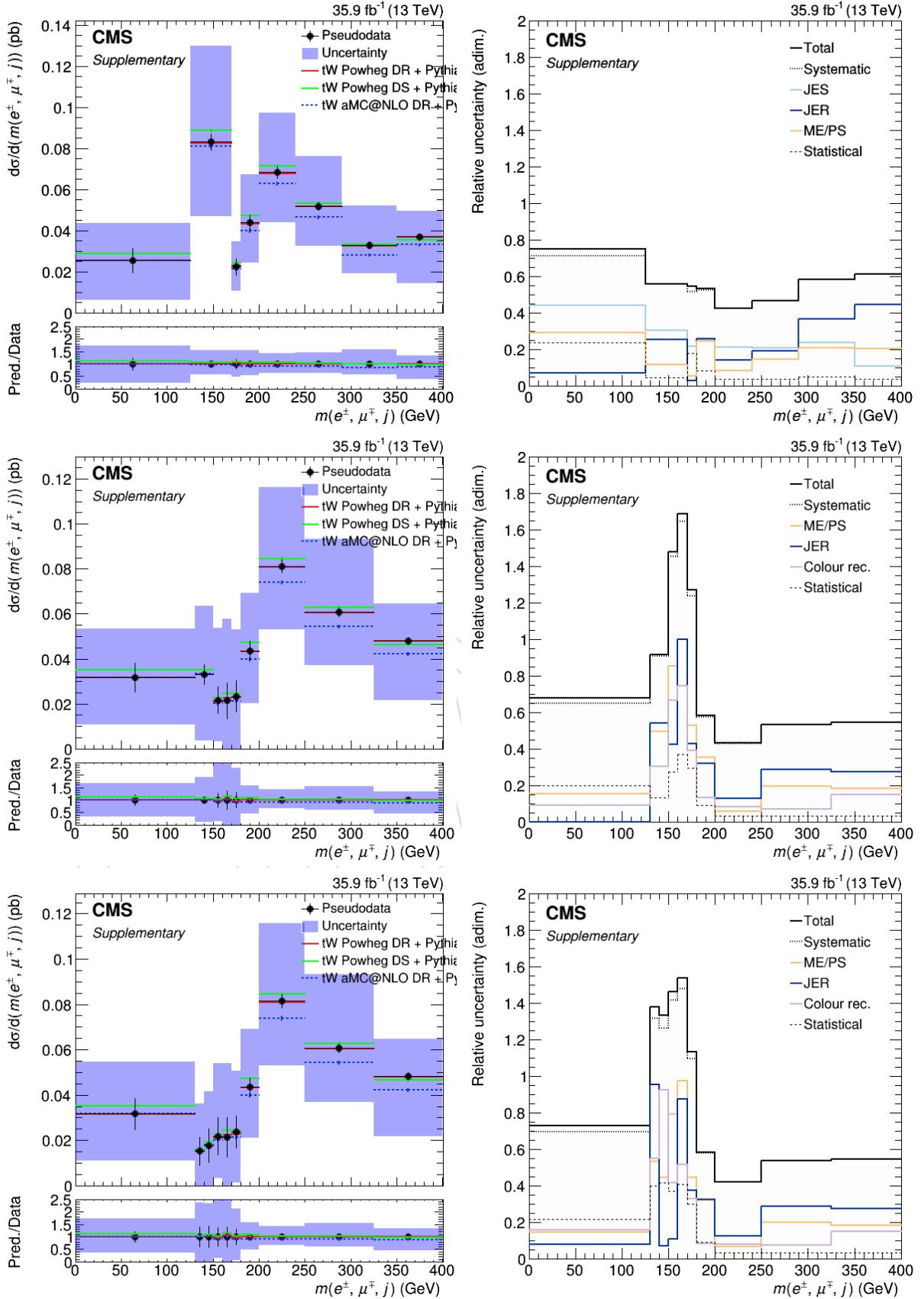


Figure 60: Absolute distributions and relative uncertainties with 8, 9 and 10 bins for  $m(e^+, \mu^-, \text{jet})$ .

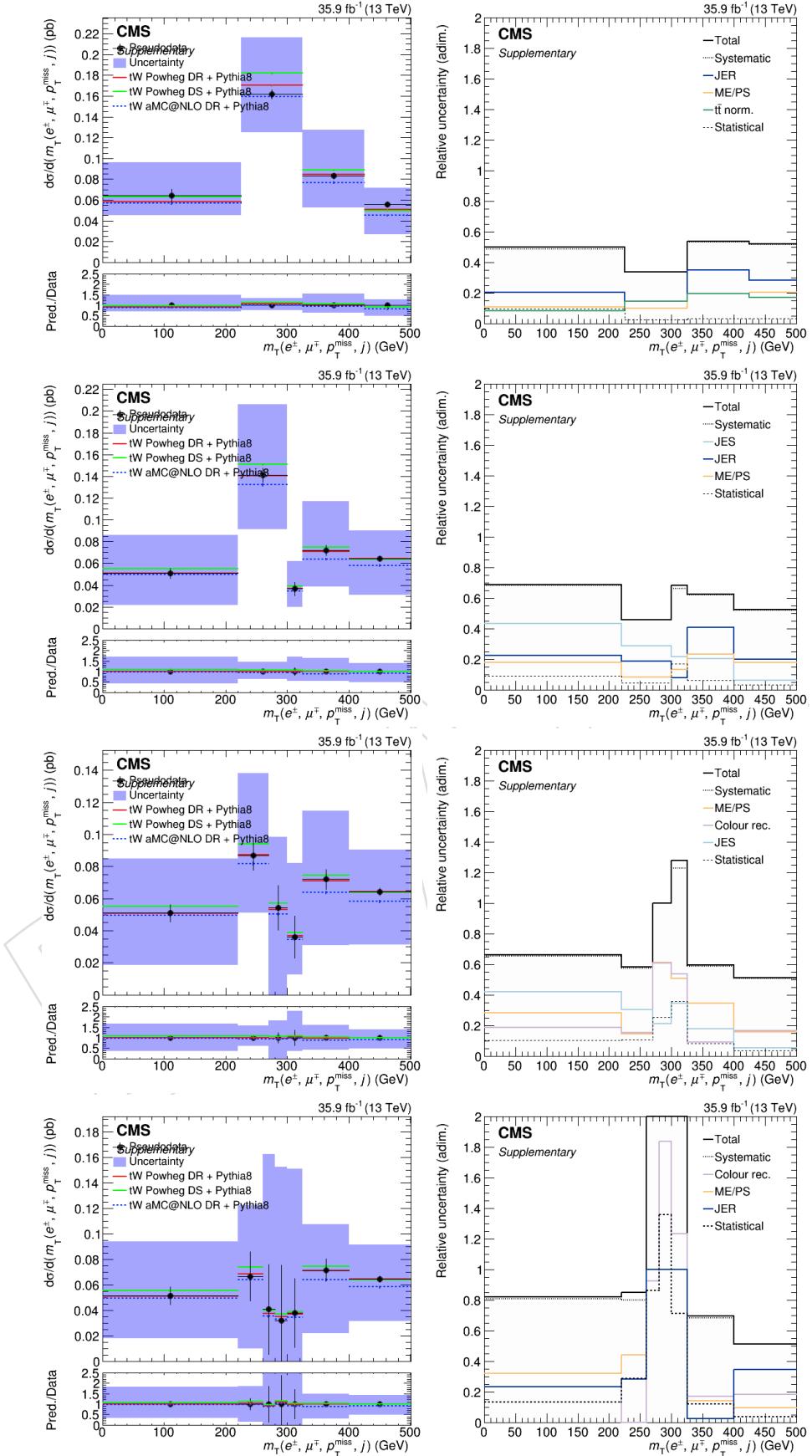


Figure 61: Absolute distributions and relative uncertainties with 4, 5, 6 and 7 bins for  $m_T(e^\pm, \mu^\mp, \text{jet}, p_T^{\text{miss}})$ .

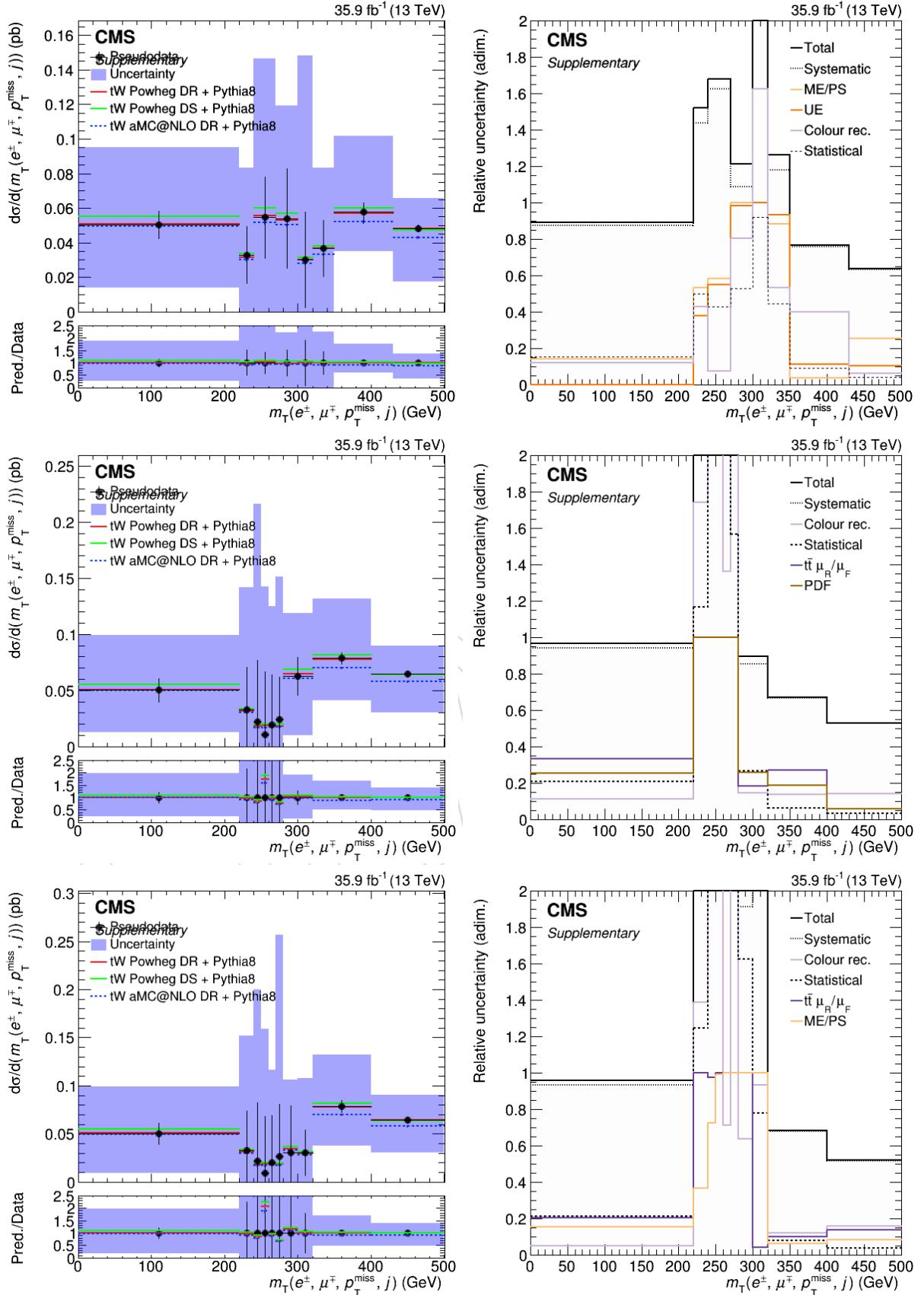


Figure 62: Absolute distributions and relative uncertainties with 8, 9 and 10 bins for  $m_T(e^\pm, \mu^\mp, \text{jet}, p_T^{\text{miss}})$ .

## 908 C Abundance of background in signal region and uncertainties

909 When checking the plots from the Results section (Sec. 7) one can notice an effect when mov-  
910 ing from the bulk to the tails of the distributions: that some uncertainties grow. This can be, a  
911 priori, logical for some cases, as in the tails there might be some “extreme” cases that might not  
912 be good modeled, or we might have problems of statistical nature.

913

914 However, some of the uncertainties that grow apparently do not have such reasons to do so,  
915 such as, for example, JES and JER, which in various plots can be seen growing in the tails.  
916 After carefully checking the folded results (because this effect was also seen in them and that  
917 way we can distangle it from the procedure of unfolding), we discovered that the reason for  
918 this increment appears to rely in the proportion between signal and background in all variables,  
919 that systematically decreases when going from the bulk to the tails of the distributions. This can  
920 be seen (with the old binning with 4 bins in unfolded space, that is 8 bins in folded space) in  
921 Figs. 63, 64, where we are taking as an example the mentioned JER and JES. The increase in the  
922 presence of background (which is essentially  $t\bar{t}$ ) can explain the growing of these uncertainties  
923 when moving to the tails of the distributions because the JES and JER uncertainties, at the stage  
924 of folded results, come only from the  $t\bar{t}$  background. The only case to consider where this does  
925 not happen is the last bin of the leading lepton  $p_T$ , where it increases (and this is translated also  
926 into the uncertainties of the folded results).

DRAFT

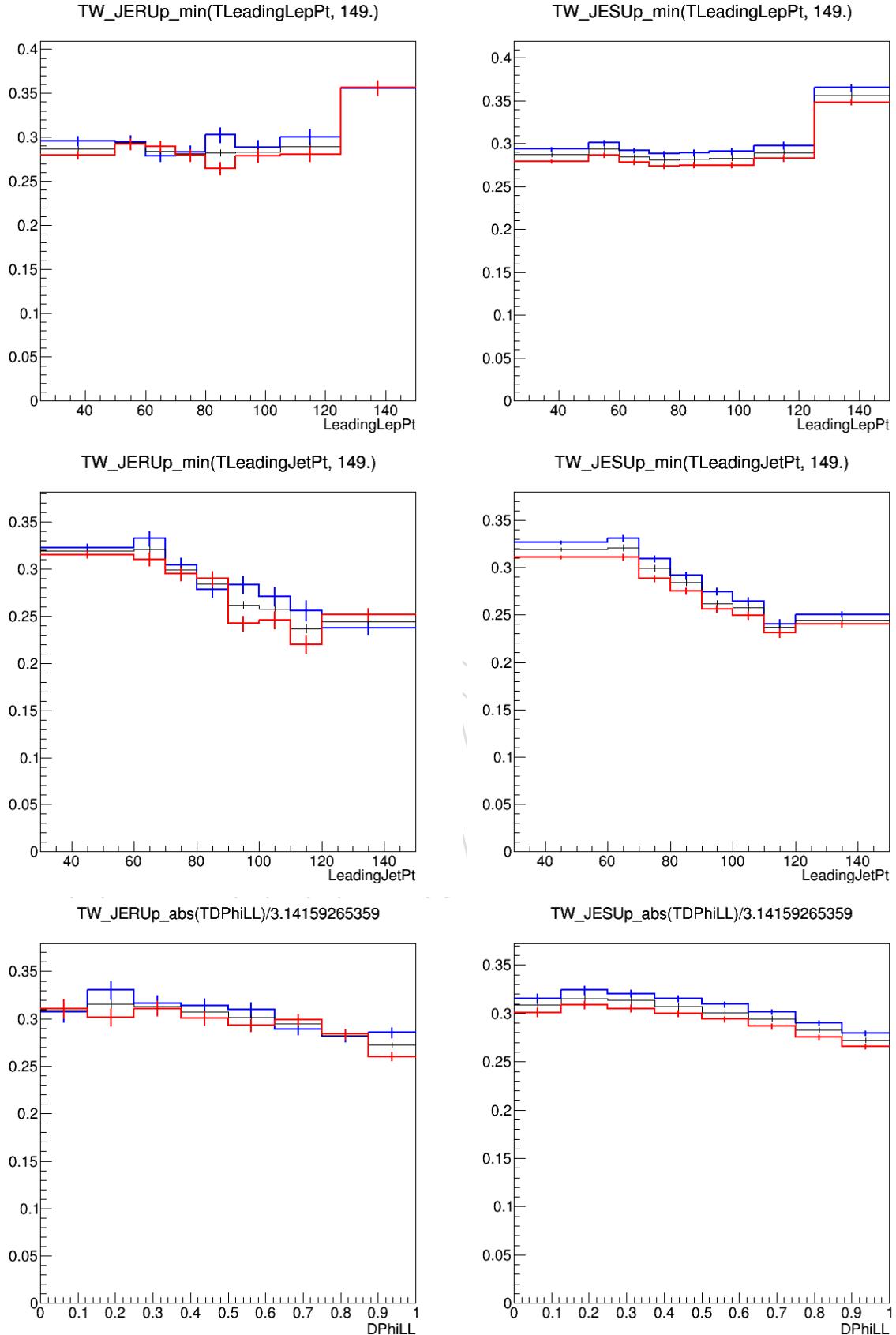


Figure 63: Distributions of the ratio between signal and background for the leading lepton  $p_T$ , the jet  $p_T$  and the  $\Delta\varphi(e^\pm, \mu^\mp)$ . In blue, the ratio of the upper variations of the uncertainties are shown, in red the down variations and in black the ratio of the nominal values.

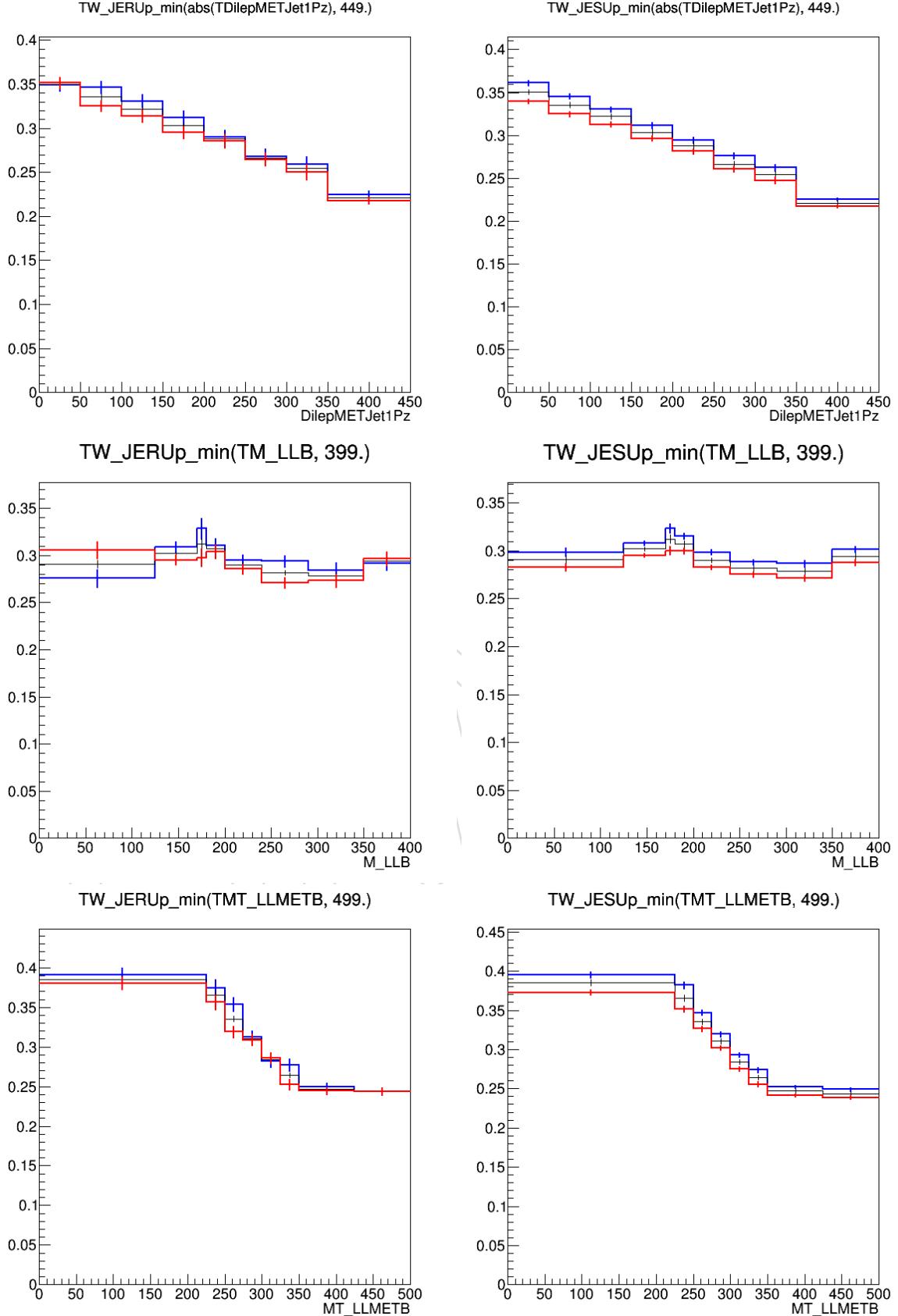


Figure 64: Distributions of the ratio between signal and background for  $p_Z(e^\pm, \mu^\mp, j)$ ,  $m(e^\pm, \mu^\mp, j)$  and  $m_T(e^\pm, \mu^\mp, j, p_T^{miss})$ . In blue, the ratio of the upper variations of the uncertainties are shown, in red the down variations and in black the ratio of the nominal values.

---

## 927 D Purities and stabilities w/o reconstruction efficiency

928 In this appendix we add for comparison purposes the plots of purities and stabilities for the six  
 929 observables but removing the reconstruction efficiency from the stabilities. To be precise, the  
 930 stability definition for a particle level bin  $i$  is the following in the plots of this appendix:

$$s_i := \frac{\sum_{j=1}^{N_{\text{bins}}^{\text{fol.}}} n_{ij}}{n'_i}. \quad (10)$$

931 Here,  $n_{ij}$  is the number of reconstructed events whose value of the measured variable fall in the  
 932 folded-space bin  $j$  that had the simulated value of the variable in the unfolded-space bin  $i$  and  
 933  $n'_i$  is the number of events whose simulated value of the variable fell in the unfolded space bin  
 934  $i$  that have been also reconstructed.

935 In Fig. 65 you can check the purities and stabilities calculated this way for the six observables.



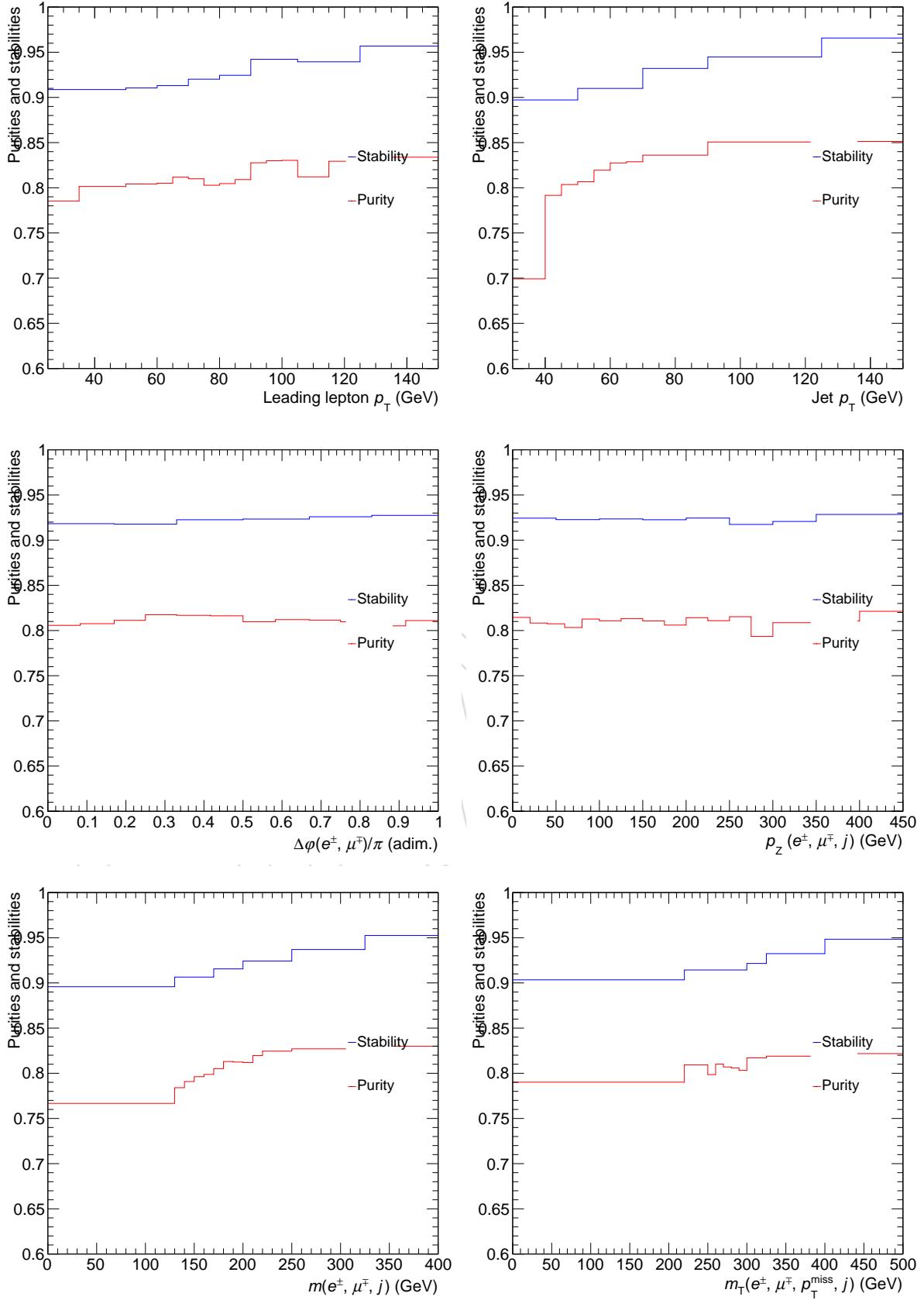


Figure 65: Graphs showing the purity and stability of each bin of the folded and unfolded (respectively) spaces for all the distributions chosen to be unfolded. The effect of reconstruction efficiency has been disentangled for the purity values.