

CMS Draft Analysis Note

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

2016/07/18

Head Id: 343919

Archive Id: 352930:352932

Archive Date: 2016/05/23

Archive Tag: trunk

Measurement of the differential cross section for $pp \rightarrow ZZ \rightarrow 4\ell$ produced in association with jets in pp collisions at $\sqrt{s} = 8$ TeV

N. Amapane¹, R. Bellan¹, R. Covarelli¹, L. Finco¹, R. Gomez Ambrosio¹, C. Mariotti², G. Pinna¹, and many other people³

¹ Università degli Studi di Torino and INFN.

² INFN Torino.

³ The Higgs to four leptons team.

Abstract

In this note we present the measurement of the differential cross section for two Z bosons produced in association with jets, extracted from pp collision data at $\sqrt{s} = 8$ TeV. The measurement is done considering the processes with leptonic decays of the Z bosons in either electrons or muons. The analysis is based on a data sample collected with the CMS Experiment at the LHC, corresponding to an integrated luminosity of 19.7 fb^{-1} . The measurements are performed in the leptonic decay modes $ZZ \rightarrow \ell\ell'\ell'\ell'$, where $\ell, \ell' = e, \mu$. The measured total cross section is $\sigma(pp \rightarrow ZZ) = 7.76 \pm 0.48(\text{stat.})^{+0.27}_{-0.29}(\text{syst.}) \text{ pb}$, for both Z bosons produced in the mass range $60 < m_Z < 120$ GeV and it is consistent with Standard Model predictions. The cross-section measured in the fiducial region corresponding to the visible phase space defined by the kinematic and geometrical acceptance of leptons is $\sigma_{pp \rightarrow ZZ \rightarrow 4\ell} = 20.50 \pm 1.26 \text{ (stat.)}^{+0.58}_{-0.64} \text{ (syst.)} \pm 0.53 \text{ (lumi.)}$, in agreement with the SM prediction. Differential cross sections are measured, in particular as a function of variables related to the jets produced in the event.

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

PDFAuthor: R. Bellan, L. Finco
PDFTitle: Measurement of the differential cross section for $ZZ \rightarrow 4\ell$ produced in association with jets in pp collisions at $\sqrt{s} = 8$ TeV
PDFSubject: CMS
PDFKeywords: CMS, physics, software, computing

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

DRAFT

1 Contents

1	Introduction	1
2	Dataset and Monte Carlo Samples	2
3	Event Selection	4
4	Background Estimation	5
5	Systematic Uncertainties	6
6	Results	8
7	6.1 Inclusive ZZ cross section measurement	8
8	6.2 The differential ZZ cross-section measurements	10
9	7 Enhanced EWK Region of ZZ + 2jets Events	52
10	8 Conclusions	53
11	A Unfolding	54
12	A.1 Unfolding procedure	54
13	A.2 Bin-to-bin migration	55
14	A.3 Unfolding validation test on Monte Carlo	60
15	A.4 Systematic Uncertainties	86

1 Introduction

The study of vector boson pair production in proton-proton collisions is crucial both in order to check the gauge structure of the Standard Model (SM) and to search for new physics. In addition to this, the production of vector boson pairs, and their decay particles, enter indeed as irreducible backgrounds in many Higgs boson analyses and new physics searches and it is thus extremely important to measure these processes with high precision. Both ATLAS and CMS Collaborations at the LHC measured the cross-section for the Z boson pair production decaying in charged-lepton final states. The CMS Collaboration performed the measurement in the $ZZ \rightarrow \ell\ell\ell'\ell'$ decay channels, where $\ell = e, \mu$ and $\ell' = e, \mu, \tau$ with the data corresponding to an integrated luminosity of 5.1 (19.6) fb^{-1} at $\sqrt{s} = 7$ (8) TeV [1, 2]. The measured total cross section is $\sigma(pp \rightarrow ZZ) = 6.24^{+0.86}_{-0.80} (\text{stat.})^{+0.41}_{-0.32} (\text{syst.}) \pm 0.14 (\text{lumi.}) \text{ pb}$ at $\sqrt{s} = 7$ TeV and $7.7 \pm 0.5 (\text{stat.})^{+0.5}_{-0.4} (\text{syst.}) \pm 0.4 (\text{theo.}) \pm 0.2 (\text{lumi.})$ at $\sqrt{s} = 8$ TeV, for both Z bosons in the mass range $60 < m_Z < 120$ GeV. The ATLAS Collaboration measured a total cross section of $6.7 \pm 0.7 (\text{stat.})^{+0.4}_{-0.3} (\text{syst.}) \pm 0.3 (\text{lumi.}) \text{ pb}$ using $ZZ \rightarrow \ell\ell\ell'\ell'$ and $ZZ \rightarrow \ell\ell\nu\nu$ final states with a data sample corresponding to an integrated luminosity of 4.6 fb^{-1} at $\sqrt{s} = 7$ TeV and $66 < m_Z < 116$ GeV. The ATLAS result of the total ZZ production cross section at $\sqrt{s} = 8$ TeV is $\sigma(pp \rightarrow ZZ) = 7.1^{+0.5}_{-0.4} (\text{stat.}) \pm 0.3 (\text{syst.}) \pm 0.2 (\text{lumi.}) \text{ pb}$, measured in the $ZZ \rightarrow \ell\ell\ell'\ell'$ final state, with $66 < m_Z < 116$ GeV and using a data sample of 20.3 fb^{-1} . Measurements of the ZZ cross sections performed at the Tevatron are summarized in [3, 4]. All measurements are in agreement with the corresponding SM predictions.

In this note we present the measurement of the differential cross section of two Z bosons produced in association with jets, extracted from pp collision data at $\sqrt{s} = 8$ TeV. The measurement is done considering the processes with leptonic decays of the Z bosons in either electrons or muons, using datasets containing an integrated luminosity of 19.7 fb^{-1} .

This analysis extends the measurement done in [2], using the latest reprocessed data, to event quantities that have jets as basic observable. The main goal is to measure the differential cross-section in the number of jets accompanying the two Z bosons, as well as the measurement of

the differential cross-section as a function of the invariant mass of the two energy-leading jets (m_{jj}), the difference in (pseudo-)rapidity ($\Delta\eta(j_1, j_2)$) between them, their transverse momentum and pseudorapidity ($p_T^{j1,j2}$ and $\eta^{j1,j2}$). Very useful information can be extracted from these differential cross-sections: the cross-section dependency to the jet multiplicity has never been measured so far, and can tell us how well we understand the QCD correction to the ZZ production. The measurement in the $\Delta\eta(j_1, j_2)$ variable will be instrumental for the vector boson scattering analysis in this channel, while the m_{jj} will be also the basis for future multi-boson final state searches (namely where an additional electroweak boson decays hadronically). The latter is also a key variable for the investigation of anomalous Quartic Gauge couplings (aQGC). Furthermore, the dependency to the transverse momentum and pseudorapidity of the leading and sub-leading jet is investigated in order to obtain further information.

The implant of the analysis is largely the same as [2] and, for what concerns the reconstruction of the two Z bosons, it is explicitly based on the algorithms and the techniques used for the $H \rightarrow ZZ \rightarrow 4\ell$ analysis [5] (as it was done for [2]), with the obvious change in the mass window ($60 < m_{Z_{1,2}} < 120$ GeV).

This note is organized as follow: section 2 summarizes the Monte Carlo and the dataset used for this analysis, section 3 describes the event selection, while section 4 deals with the background estimation and the related uncertainties. Section 5 lists the several sources of systematic uncertainty and, finally, in section 6 we present the result of this analysis, including the unfolding of the detector resolution for the differential cross-section that we measured.

2 Dataset and Monte Carlo Samples

Several Monte Carlo (MC) event generators are used to simulate the signal and background contributions. The MC samples are used to optimize the event selection, evaluate the signal efficiency and acceptance, extract the unfolding response matrices and estimate the irreducible background yields. The full set of MC samples used in this analysis is reported in table 1.

The $qq/qg \rightarrow ZZ$ and $gg \rightarrow ZZ$ (tree-level only) processes are generated at leading-order (LO) for 0, 1 and 2 jets with `MadGraph 5.1` [6], while the $gg \rightarrow ZZ$ (box diagrams) processes are generated at LO with `MCFM` [7]. For comparison we also use signal $qq/qg \rightarrow ZZ$ samples generated with `Powheg` (NLO for 0 jet) [8, 9] and `MadGraph5_aMCatNLO` (NLO for 1 jet) [10]. Vector boson scattering processes with two Z bosons produced in association with 2 forward and backward jets are simulated using `Phantom` [11]. Other diboson and triboson processes (WZ , Zg , ZZZ , WZZ , WWZ) and the Z + jets samples are generated at LO with `MadGraph`, like events from $t\bar{t}$ production. The `PYTHIA 6.4` [12] package is used for parton showering, hadronization, and the underlying event simulation. The default set of parton distribution functions (PDF) used for LO generators is `CTEQ6L` [13], whereas `CT10` [14] is used for NLO generators. The ZZ yields from simulation are scaled according to the theoretical cross sections calculated with `MCFM`.

The data samples used in this analysis correspond to an integrated luminosity of 19.7 fb^{-1} at $\sqrt{s} = 8$ TeV collected in 2012. The integrated luminosity is measured using data from the HF system and the pixel detector [15]. The uncertainty in the integrated luminosity measurement is 2.6%.

The data sets processed in this analysis are summarized in table 2, along with the trigger selection operated on them (based on the same selection presented in [2, 5]).

Table 1: List of signal and background samples used in the analysis.

Process	Cross Sections [pb]	Sample
Signal Samples		
$ZZ \rightarrow 4\ell$ (MadGraph)	0.071 (no τ)	ZZJetsTo4L_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM/PAT_CMG-V5_15_0/
$q\bar{q} \rightarrow ZZ \rightarrow 4\mu$ (Powheg)	0.07691	ZZTo4mu_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$q\bar{q} \rightarrow ZZ \rightarrow 4e$ (Powheg)	0.07691	ZZTo4e_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$q\bar{q} \rightarrow ZZ \rightarrow 2e2\mu$ (Powheg)	0.1767	ZZTo2e2mu_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$gg \rightarrow ZZ \rightarrow 4\mu$ (MCFM)	0.000592340	GlugluTo4mu-SMHContInInterf_M-125p6.8TeV-MCFM67-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$gg \rightarrow ZZ \rightarrow 4e$ (MCFM)	0.000592000	GlugluTo4e-SMHContInInterf_M-125p6.8TeV-MCFM67-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$gg \rightarrow ZZ \rightarrow 2e2\mu$ (MCFM)	0.00118409	GlugluTo2e2mu_SMHContInInterf_M-125p6.8TeV-MCFM67-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZ \rightarrow 4\mu + 2jets$ (Phantom)	2.49E-04	ZZTo4muJJ_SMHContInInterf_M-125p6.8TeV-phantom-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZ \rightarrow 4e + 2jets$ (Phantom)	2.48E-04	ZZTo4eJJ_SMHContInInterf_M-125p6.8TeV-phantom-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZ \rightarrow 2e2\mu + 2jets$ (Phantom)	5.43E-04	ZZTo2e2muJJ_SMHContInInterf_M-125p6.8TeV-phantom-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$WZZ + jets$ (MadGraph)	0.01968	WZZNogstarJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZZ + jets$ (MadGraph)	0.005527	ZZZNogstarJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$H + H$ (Pythia)	0.0001081681	TTbarH_HToZZTo4LM-126.8TeV-pythia6/Summer12_DR53X-PU_S10_START53_V7C-v1
Irreducible Background Samples		
$t\bar{t} + Z + jets$ (MadGraph)	0.2057	TTZJets_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1
$t\bar{t} + WW + jets$ (MadGraph)	0.00208989	TTWWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$WWZ + jets$ (MadGraph)	0.05795	WWZNogstarJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
Reducible Background Samples		
Z (MadGraph)	3503.71	DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
WZ (MadGraph)	1.057	WZJetsTo3LNu_TuneZ2Star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
WWW (MadGraph)	0.08058	WWWWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$t\bar{t} + \gamma + jets$ (MadGraph)	1.93236	TTGJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$t\bar{t} \rightarrow 2l2\tau 2B$ (Powheg)	23.64	TTTo2L2Nu2B_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$H + W + jets$ (MadGraph)	0.232	TTWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1

Table 2: List of data samples used in the analysis: 22Jan2013ReReco reprocessing of DoubleMu, DoubleEle, MuEG; total integrated luminosity of 19.7 fb^{-1} .

Data Samples
DoubleMu/Run2012A-22Jan2013-v1
DoubleMuParked/Run2012B-22Jan2013-v1
DoubleMuParked/Run2012C-22Jan2013-v1
DoubleMuParked/Run2012D-22Jan2013-v1
DoubleElectron/Run2012A-22Jan2013-v1
DoubleElectron/Run2012B-22Jan2013-v1
DoubleElectron/Run2012C-22Jan2013-v1
DoubleElectron/Run2012D-22Jan2013-v1
MuEG/Run2012A-22Jan2013-v1
MuEG/Run2012B-22Jan2013-v1
MuEG/Run2012C-22Jan2013-v1
MuEG/Run2012D-22Jan2013-v1

87 3 Event Selection

88 A complete reconstruction of the individual particles emerging from each collision event is
 89 obtained via a particle-flow (PF) technique [16], which uses the information from all CMS sub-
 90 detectors. The identification of the particles produced in the event is the very same as the one
 91 described in [2], including the final state radiation treatment. Also the logic to build up the ZZ
 92 candidate is the same as the Higgs boson analysis [5].

93 The CMS standard selection of runs and luminosity sections is applied, which requires high
 94 quality data with a good functioning of the different sub-detectors. Events are selected online
 95 from the presence of a pair of electrons or muons, or a triplet of electrons. Triggers requiring
 96 an electron and a muon are also used. The minimal transverse momenta (p_T) of the first and
 97 second lepton are 17 and 8 GeV, respectively, for the double lepton triggers, while they are
 98 15, 8 and 5 GeV for the triple electron trigger. More details can be found in [17]. The trigger
 99 efficiency for ZZ events within the acceptance of this analysis is greater than 98%.

100 Events with identified and isolated primary electrons or muons are first selected by the offline
 101 selection. We require a Z candidate formed with a pair of leptons of the same flavor and oppo-
 102 site charge. The FSR photons are kept if $|m_{\ell\ell\gamma} - m_Z| < |m_{\ell\ell} - m_Z|$ and $m_{\ell\ell\gamma} < 100 \text{ GeV}$ and in
 103 the following the presence of the photons in the $\ell\ell$ kinematics is implicit.

104 The lepton pair with the invariant mass closest to the nominal Z mass is denoted Z_1 . A second
 105 $\ell^+\ell^-$ pair is required and denoted Z_2 . If more than one Z_2 candidate satisfies all criteria, the
 106 pair of leptons with the highest scalar sum of p_T is chosen. Both Z candidates have to satisfy
 107 $60 < m_Z < 120 \text{ GeV}$. Among the four selected leptons forming the Z_1 and the Z_2 candidates
 108 at least one should have $p_T \geq 20 \text{ GeV}$ and another one have $p_T \geq 10 \text{ GeV}$ and any opposite-
 109 charge pair of leptons chosen among the four selected leptons satisfy $m_{\ell\ell} \geq 4 \text{ GeV}$.

110 Jets are reconstructed using the anti- k_T clustering algorithm [18], with a size parameter of
 111 $R = 0.5$, by summing the four-momenta of individual PF particles according to the FASTJET
 112 package of reference [19]. Once reconstructed, jets overlapping (within $\Delta R = 0.5$) with any of
 113 the two leptons coming form the decay of the Z boson are removed from the jet collection. Jets
 114 satisfy the selection criteria recommended by Jet-MET group [20]. Jet energy corrections are
 115 applied as a function of the jet p_T and η [21]. Moreover, they are required to have $p_T > 30 \text{ GeV}$,
 116 to reduce the pileup contamination as well as large uncertainty on the energy measurement,

and $|\eta| < 4.7$, to ensure a good quality of the tracking information. Since the jet energy resolution in data is worse than in simulation, the p_T values of simulated jets need to be smeared to describe data. Following the prescription of the CMS Jet-MET group [22], the corrected p_T of the reconstructed jets are randomly smeared using a Gaussian distribution with a width of $\sqrt{c^2 - 1} \cdot \sigma_{MC}$, where c is the scaling factor (see Table 3). This method only allows one to worsen the resolution ($c > 1$). To determine the jet resolution in simulation, σ_{MC} , recommended tools have been used [22].

Table 3: Scaling factors applied on MC jets to described the data resolution

$ \eta $ -region	0.0 – 0.5	0.5 – 1.1	1.1 – 1.7	1.7 – 2.3	2.3 – 2.8	2.8 – 3.2	3.2 – 5.0
c	1.079	1.099	1.121	1.208	1.254	1.395	1.056

123

Ancillary quantities, like data/MC efficiency scale factor and jet-to-lepton fake rate, are the same as the one measured in [2].

4 Background Estimation

The lepton identification and isolation requirements significantly suppress the background contribution, and the remnant portion of it is very small compared to the signal. We can identify two main background components: an irreducible background from genuine four high- p_T isolated leptons processes, such as $t\bar{t}Z$, WWZ and $t\bar{t}WW$, and a reducible background from processes with less than four leptons, but with jets that are misidentified as leptons.

The irreducible background is very small and it is estimated using MC samples (see Fig. 1).

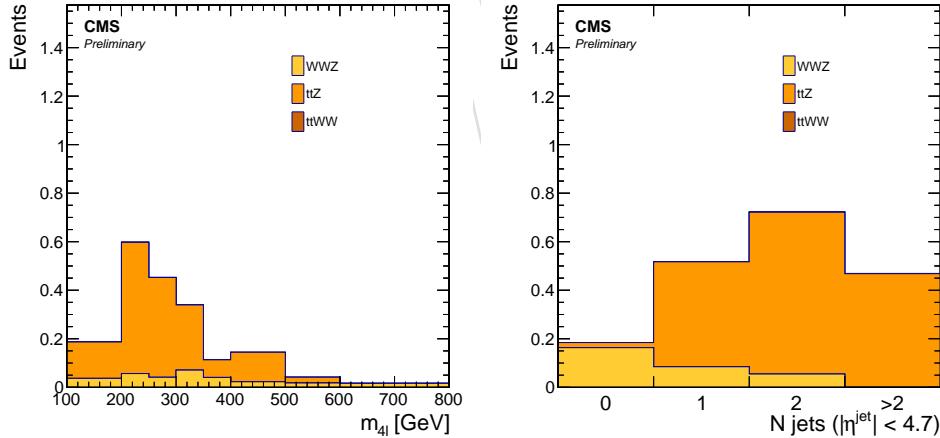


Figure 1: Number of events of the irreducible background component in the signal region as a function of the invariant mass of the 4 lepton system (left) and the reconstructed number of jets produced in the event (right).

132

The main background contribution that arises mostly from the Z in association with jets, as well as from $t\bar{t}$, WZ and $WWW + \text{jets}$, is estimated with data analyzed in dedicated control regions. For this type of background, a jet or a non-prompt lepton (i.e. lepton produced in heavy meson decays) is misidentified as an isolated electron or muon. These processes are referred to as reducible background and are estimated using a data driven method, known as “method A”

in the $H \rightarrow ZZ \rightarrow 4\ell$ analysis [5]. This method is based on two control samples, obtained as subsets of four lepton events, by demanding a Z_1 candidate reconstructed with the same requirement as the Z_1 built in the signal region, and two additional leptons, with opposite sign and same flavor ($e^\pm e^\mp$ or $\mu^\pm \mu^\mp$). The “3P1F” control sample contains events with only one of the two additional leptons failing identification and isolation criteria, the other being a good quality lepton, and it is used to estimate the contribution in the signal region of events with three genuine leptons, such as $WZ + \text{jets}$. In the “2P2F” control region, instead, the two additional leptons are both required to fail the final identification and isolation criteria. This sample is meant to be used to estimate both the contribution in the signal region and in the “3P1F” region of events with only two genuine leptons, such as $Z + \text{jets}$ and $t\bar{t}$. The number of reducible background events in the signal region is given by the following formula:

$$N_{\text{exp}}^{\text{red bkg}} = \sum_i^{N_{3\text{P}1\text{F}}} p_i - \sum_i^{N_{2\text{P}2\text{F}}} p_{i,1} p_{i,2}, \quad (1)$$

133 where:

- 134 • $N_{3\text{P}1\text{F}}$ and $N_{2\text{P}2\text{F}}$ are the number of events in the “3P1F” and “2P2F” control regions
135 respectively

- 136 • $p_i = \frac{f_i}{1-f_i}$

- 137 • $p_{i,1(2)} = \frac{f_{i,1(2)}}{1-f_{i,1(2)}}$

- 138 • f_i and $f_{i,1(2)}$ are the fake rates of the failing lepton in the “3P1F” sample and of the
139 first (second) failing lepton in the “2P2F” control sample.

140 The “fake rate” is the lepton misidentification probability $f(p_T^\ell, \eta^\ell)$ to extrapolate the back-
141 ground yields from the control region to the signal region. This probability is defined as the
142 fraction of non-signal leptons passing the lepton selection of the analysis. The fake rate is es-
143 timated in a sample composed by $Z_1 + 1\ell_{\text{loose}}$ events, where beside to the opposite sign/same
144 flavor pair which forms the Z_1 , exactly one additional lepton is reconstructed fulfilling all the
145 selection requirements, but for the identification and the isolation cuts.

146 The lepton misidentification probability ranges from 1% to 15% and has a mild dependence on
147 the pseudorapidities for the electrons.

148 In order to obtain the reducible background contribution as a function of the analyzed vari-
149 ables, the number of events in the “3P1F” and “2P2F” control regions is estimated in each bin
150 of each distribution separately, it is multiplied by the corresponding fake rate (depending on
151 leptons p_T and η) and summed, according to 1. In order to be consistent with the signal re-
152 gion, reconstructed jets that overlap with the loose leptons in the CRs are removed. In Figure 2
153 the contribution of the reducible background is reported as a function of the 4-lepton invari-
154 ant mass and the jet multiplicity, and it is estimated both from MC samples and by using the
155 “fake-rate” method. The lack of statistics makes the data-driven estimate necessary.

156 5 Systematic Uncertainties

157 Systematic uncertainties for trigger efficiency (1.5%) are evaluated from data [2]. Uncertainties
158 arising from lepton identification, isolation, tracking and impact parameter are 1-5% for muons
159 and electrons [2]. The uncertainty in the LHC integrated luminosity of the data sample is
160 2.6% [15]. Theoretical uncertainties in the $ZZ \rightarrow \ell\ell\ell'\ell'$ acceptance are evaluated using MCFM
161 and by varying the renormalization and factorization scales, up and down, by a factor of two
162 with respect to the default values $\mu_R = \mu_F = m_Z$. The variations in the acceptance are 0.1%

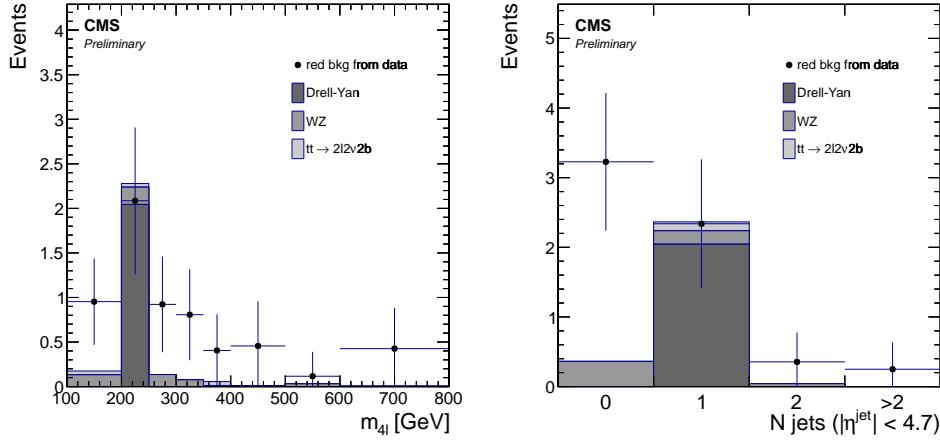


Figure 2: Reducible background component in the signal region as a function of the invariant mass of the 4 lepton system (left) and the reconstructed number of jets produced in the event (right). Points represent the data-driven estimate, the stacked histogram represents the Monte Carlo predictions, characterized by a very poor statistics.

(NLO $q\bar{q} \rightarrow ZZ$) and 0.4% ($gg \rightarrow ZZ$), and can be neglected. Uncertainties related to the choice of the PDF and the strong coupling constant are evaluated following the PDF4LHC [14, 23, 24] prescription and using CT10, MSTW08, and NNPDF [25] PDF sets. They are found to be 5% in the wider fiducial region and 1% in the tighter fiducial region (fiducial regions are defined in the next section).

The reducible background uncertainties in $Z + \text{jets}$, $WZ + \text{jets}$ and $t\bar{t}$ yields reflect the uncertainties in the measured values of the misidentification rates and the limited statistics of the control regions in the data, and vary between 30% and 70% on the background yield, which corresponds to a $\sim 1\%$ of uncertainty on the cross-section measurement. The irreducible background uncertainty is smaller and varies between 13% and 26% on the yield (less than 1% on the cross-section estimate). Table 4 summarizes the sources of systematic uncertainty and their values.

Table 4: Estimated systematic uncertainties on the signal yield used to compute the inclusive cross section measurement

Systematic source	Uncertainty value
Trigger	1.5%
Lepton ID, ISO and Tracking	1 - 5%
Luminosity	2.6%
Reducible background	$\sim 1\%$
Irreducible background	<1%
Acceptance for PDF	1 - 5%

175 6 Results

176 6.1 Inclusive ZZ cross section measurement

The inclusive cross-section is determined as

$$\sigma_{pp \rightarrow ZZ} = \frac{N_{data} - N_{bkg}}{\mathcal{L} \cdot A \cdot \epsilon \cdot \mathcal{B}(Z \rightarrow \ell^+ \ell^-) \cdot \mathcal{B}(Z \rightarrow \ell'^+ \ell'^-)},$$

177 where N_{data} and N_{bkg} are the total number of data and background events, A is the geometrical
 178 acceptance, ϵ is the signal efficiency, $\mathcal{L} = 19.7 \text{ fb}^{-1}$ is the integrated luminosity and $\mathcal{B}(Z \rightarrow$
 179 $\ell^+ \ell^-) = 3.363 \pm 0.004$ (3.366 ± 0.007)% is the branching fraction of a Z boson decaying into an
 180 electron (muon) pair [26].

181 The region in which the measurement is performed (called “fiducial region”) is defined by re-
 182 quiring the invariant mass of each Z boson between 60 and 120 GeV. The product of acceptance
 183 and efficiency ($A \cdot \epsilon$) is evaluated from simulation and defined as the ratio between the num-
 184 ber of events in the fiducial region that pass the analysis selection at reconstruction level and
 185 the number of events generated in that fiducial region. The requirements on p_T and η for the
 186 particles in the final state reduce the full possible phase space of the $ZZ \rightarrow 4\ell$ measurement by
 187 a factor within a range of 0.32-0.47 for the $4e$, 4μ and $2e2\mu$ (depending on the final state) with
 188 respect to all events generated in the mass window $60 < m_{Z_1}, m_{Z_2} < 120$ GeV. The $A \cdot \epsilon$ product
 189 estimated for each final state and for the different samples is listed in Tables 5 and 6.

Table 5: $A \cdot \epsilon$ for the three final states used in the $pp \rightarrow ZZ$ cross-section measurement. The values reported are a product of the detector geometrical acceptance and the object reconstruction and event identification efficiency.

Final State	$A \cdot \epsilon$
4μ	47%
$4e$	32%
$2e2\mu$	40%

190 The number of signal events used to evaluate the ZZ cross-section measurement is computed
 191 subtracting the irreducible and reducible background estimates from the selected data sam-
 192 ples. Values are reported in Table 7. In Fig. 3, 4 and 5 the distributions of the reconstructed
 193 four-lepton mass, the number of jets in the events, the invariant mass of the two most energetic
 194 jets (m_{jj}), the pseudorapidity between them ($\Delta\eta_{jj}$) and their p_T and η distributions are reported
 195 for both MC samples and data (for $m_{4\ell}$ the set of samples used is the one with Powheg, while
 196 for the other variables the one with MadGraph).

197

198 Table 8 lists the total cross-section obtained from each individual decay channel as well as the
 199 total cross-section based on the combination of all channels. The measured cross-section agrees
 200 with the theoretical value of $7.5 \pm 0.5 \text{ pb}$ calculated with MCFM 6.6. In this calculation, the
 201 contribution from $q\bar{q} \rightarrow ZZ$ is obtained at NLO, while the smaller contribution (approximately
 202 6%) from $gg \rightarrow ZZ$ is obtained at LO.

203 The fiducial region in which the cross-section is extracted, defined requiring two Z bosons
 204 with mass between 60 and 120 GeV, is much wider with respect to the region in which events
 205 are really measured, limited by the active volume of the detector, as shown by the low values
 206 of $A \cdot \epsilon$. In order to obtain a measurement closer to what is effectively reconstructed by the
 207 detector, a *tight fiducial region* is defined as follows:

Table 6: Signal yield and efficiency for the signal samples. The $A \cdot \epsilon$ values reported are a product of the detector geometrical acceptance and the object reconstruction and event identification efficiency.

Process	Signal yield			$A \cdot \epsilon$		
	$2e2\mu$	$4e$	4μ	$2e2\mu$	$4e$	4μ
$ZZ \rightarrow 4\ell$ (MadGraph)	125.3	51.7	51.7	0.39	0.32	0.46
$q\bar{q} \rightarrow ZZ \rightarrow 4\mu$ (Powheg)	-	-	69.6	-	-	0.46
$q\bar{q} \rightarrow ZZ \rightarrow 4e$ (Powheg)	-	48.0	-	-	0.32	-
$q\bar{q} \rightarrow ZZ \rightarrow 2e2\mu$ (Powheg)	117.6	-	-	0.39	-	-
$gg \rightarrow ZZ \rightarrow 4\mu$ (MCFM)	-	-	5.04	-	-	0.72
$gg \rightarrow ZZ \rightarrow 4e$ (MCFM)	-	3.47	-	-	0.49	-
$gg \rightarrow ZZ \rightarrow 2e2\mu$ (MCFM)	8.52	-	-	0.60	-	-
$ZZ \rightarrow 4\mu + 2jets$ (Phantom)	-	-	0.633	-	-	0.65
$ZZ \rightarrow 4e + 2jets$ (Phantom)	-	0.440	-	0.45	-	-
$ZZ \rightarrow 2e2\mu + 2jets$ (Phantom)	1.08	-	-	0.56	-	-
$ZZZ + jets$ (MadGraph)	0.284	0.121	0.192	0.42	0.35	0.54
$WZZ + jets$ (MadGraph)	0.328	0.138	0.239	0.44	0.33	0.49
$W + H$ (MadGraph)	0.015	0.0058	0.0092	0.40	0.35	0.53
$Z + H$ (MadGraph)	0.200	0.083	0.117	0.42	0.34	0.48
$t\bar{t} + H$ (MadGraph)	0.017	0.007	0.010	0.44	0.34	0.49
Total MadGraph Set	135.78	55.93	80.64	0.40	0.32	0.47
Total Powheg Set	128.25	52.33	75.89	0.40	0.32	0.47

Table 7: The signal yield is obtained by subtracting from the observed yield the expected background events. Statistical uncertainties are reported.

Final state	Observed data	Irreducible background	Reducible background	Final yield
$ZZ \rightarrow 4\mu$	80 ± 8.9	0.52 ± 0.10	0.92 ± 0.61	78.6 ± 8.9
$ZZ \rightarrow 4e$	56 ± 7.5	0.21 ± 0.06	1.93 ± 0.72	54.0 ± 7.5
$ZZ \rightarrow 2e2\mu$	152 ± 12	1.16 ± 0.15	3.3 ± 1.1	148 ± 12
$ZZ \rightarrow 4\ell$	288 ± 17	1.89 ± 0.19	6.2 ± 1.4	281 ± 17

- $60 < m_{Z_1}, m_{Z_2} < 120$ GeV;
- at least one lepton with $p_T > 20$ GeV and another one with $p_T > 10$ GeV;
- electrons must have $p_T^e > 7$ GeV and $|\eta^e| < 2.5$;
- muons must have $p_T^\mu > 5$ GeV and $|\eta^\mu| < 2.4$.

The last three requirements correspond to the η and p_T -acceptance of the detector and are the same selection criteria demanded when leptons are built, while the first requirement corresponds to the definition of the wider region applied for the previous inclusive measurement. The formula used in this case is

$$\sigma_{ZZ} = \frac{N_{data} - N_{bkg}}{\mathcal{L} \cdot A \cdot \epsilon},$$

where the branching-ratio factor is removed in order to obtain the cross-section of the $pp \rightarrow ZZ \rightarrow 4\ell$ process, and the new values of $A \cdot \epsilon$ are reported in Table 9. The cross-section measurements obtained in this tight region are listed in Table 10.

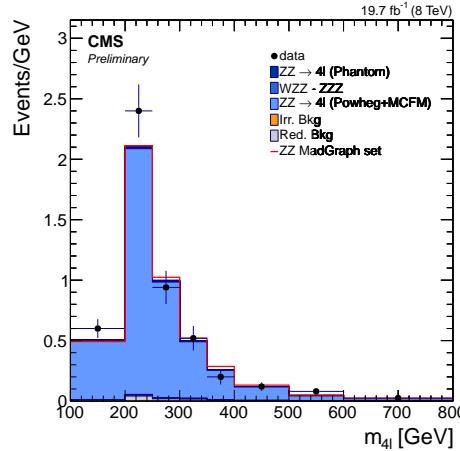


Figure 3: Different contributions to the distribution of the reconstructed four-lepton mass. Points represent the data, the stacked filled histogram represents the predictions for ZZ signal and background contributions using Powheg samples to describe $q\bar{q}(qg) \rightarrow ZZ \rightarrow 4\ell$ processes (while for the stacked histogram outlined in red the MadGraph simulation is used).

Table 8: The total ZZ production cross-section as measured in each decay channel and for the combination of all channels in the fiducial region $60 < m_{Z_1}, m_Z < 120$ GeV.

Process	Total cross-section [pb]
$pp \rightarrow ZZ(4\mu)$	7.46 ± 0.86 (stat.) $^{+0.45}_{-0.47}$ (syst.)
$pp \rightarrow ZZ(4e)$	7.32 ± 1.02 (stat.) $^{+0.59}_{-0.62}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	8.22 ± 0.70 (stat.) $^{+0.59}_{-0.61}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	7.76 ± 0.48 (stat.) $^{+0.47}_{-0.47}$ (syst.)

6.2 The differential ZZ cross-section measurements

The study of vector boson pair production is important both as check of the SM and in search for new physics. Non-resonant ZZ events are the dominant background in the $H \rightarrow ZZ \rightarrow 4\ell$ analysis and a good knowledge of these processes is very useful for the study of the associated production of a Z boson with the Higgs, especially at 13 TeV. However, a lot of information can still be extracted from data at 8 TeV, in particular as regards the associate production of a Z boson pair with jets. In addition to the invariant mass of the four lepton system, the differential ZZ cross-section is determined as a function of the number of jets produced in the event, the invariant mass of the two most energetic jets (m_{jj}), the $\Delta\eta$ between them ($\Delta\eta_{jj}$), their transverse momentum and pseudorapidity (leading jet $p_T(\eta)^{jet1}$ and sub-leading jet $p_T(\eta)^{jet2}$). Both jets

Table 9: $A \cdot \epsilon$ for the three final states used in the $pp \rightarrow ZZ \rightarrow 4\ell$ cross-section measurement in the tight fiducial region. The values reported are a product of the detector geometrical acceptance and the object reconstruction and event identification efficiency.

Final State	$A \cdot \epsilon$
4μ	84 %
$4e$	55%
$2e2\mu$	69%

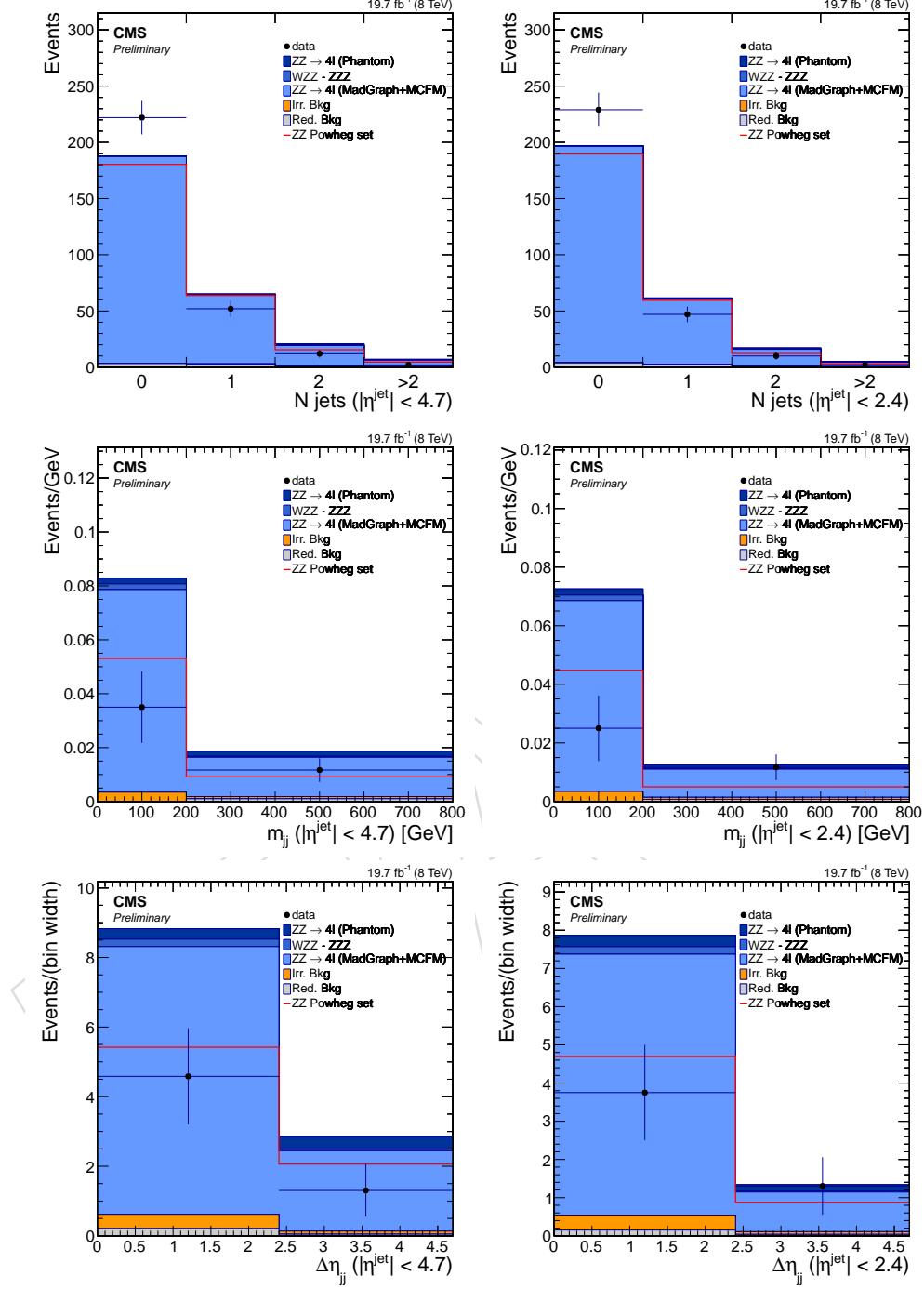


Figure 4: Different contributions to the distribution of the reconstructed number of jets produced in the event (top), the invariant mass of the two most energetic jets (center) and the pseudorapidity interval between them (bottom), with $\eta^{\text{jet}} < 4.7$ on the left and $\eta^{\text{jet}} < 2.4$ on the right. Points represent the data, the stacked filled histogram represents the predictions for ZZ signal and background contributions using `MadGraph` samples to describe $q\bar{q}(qg) \rightarrow ZZ \rightarrow 4\ell$ processes (while for the stacked histogram outlined in red the `Powheg` simulation is used).

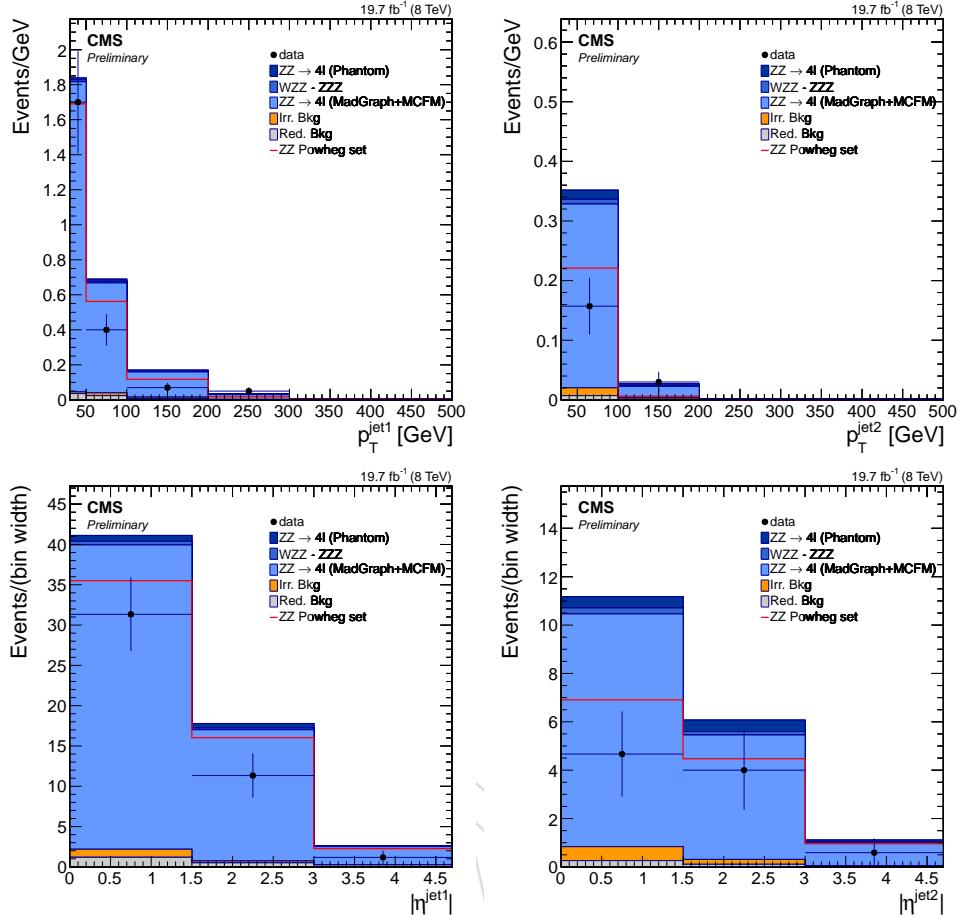


Figure 5: Different contributions to the distribution of the reconstructed transverse momentum (top) and pseudorapidity (bottom) of the leading (left) and sub-leading jet (right). Points represent the data, the stacked filled histogram represents the predictions for ZZ signal and background contributions using MadGraph samples to describe $q\bar{q}(qg) \rightarrow ZZ \rightarrow 4\ell$ processes (while for the stacked histogram outlined in red the Powheg simulation is used).

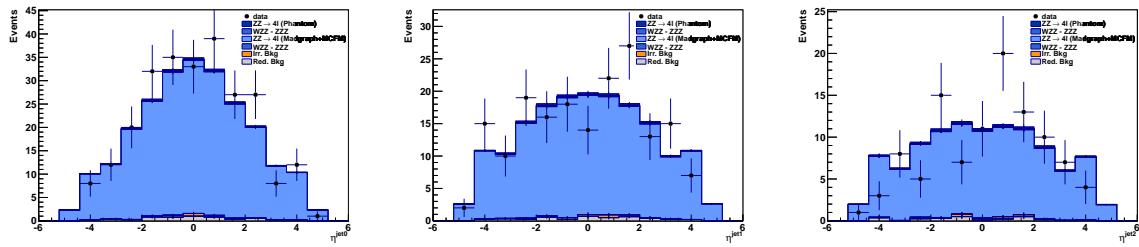


Figure 6: Different contributions to the η -distribution of the most energetic (left), second most energetic (center) and third most energetic (right) jet in the event, considering $p_T^{jet} > 10$ GeV

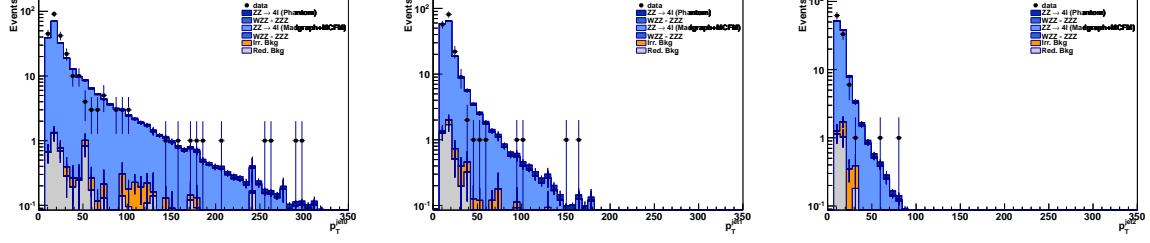


Figure 7: Different contributions to the p_T -distribution of the most energetic (left), second most energetic (center) and third most energetic (right) jet in the event, considering $p_T^{jet} > 10$ GeV.

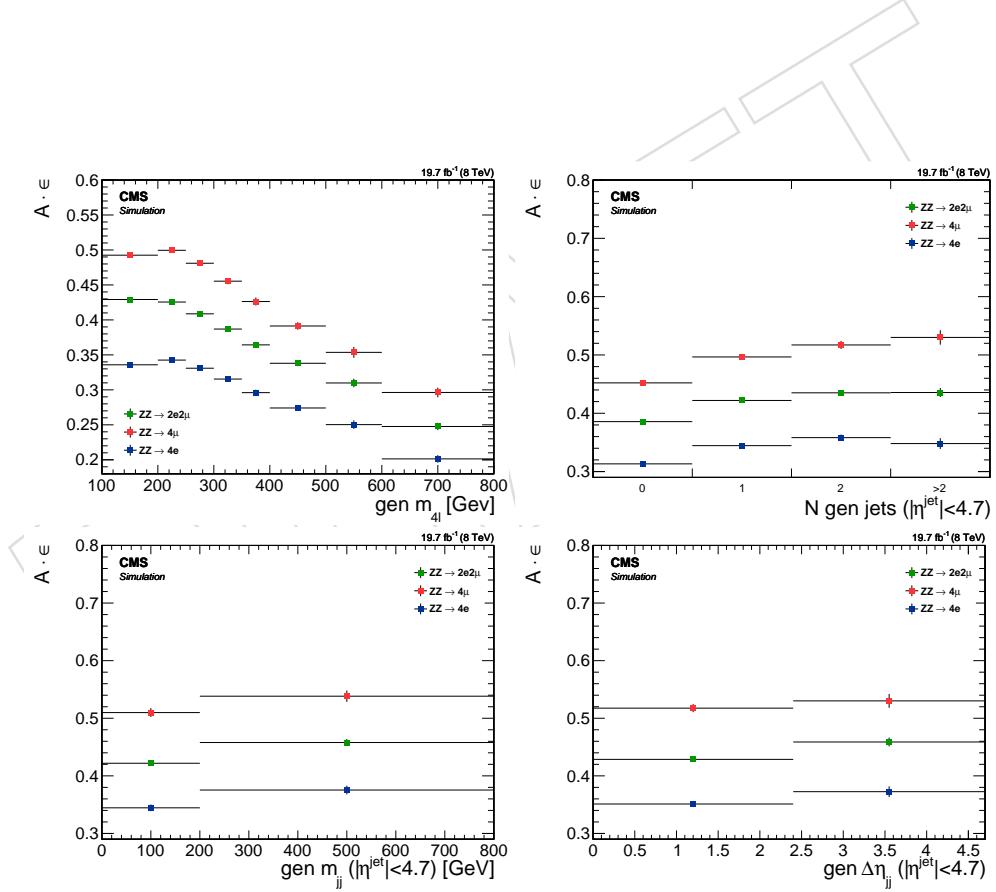


Figure 8: Acceptance ($A \cdot \epsilon$) computed in the wider fiducial region as a function of m_{ZZ} (top left), N jets (top right), m_{jj} (bottom left) and $\Delta\eta_{jj}$ distributions, according to the final state: 4μ (green), $4e$ (blue), $2e2\mu$ (red). Distributions are obtained using the Powheg set of samples for m_{ZZ} and the MadGraph set of samples for the other variables.

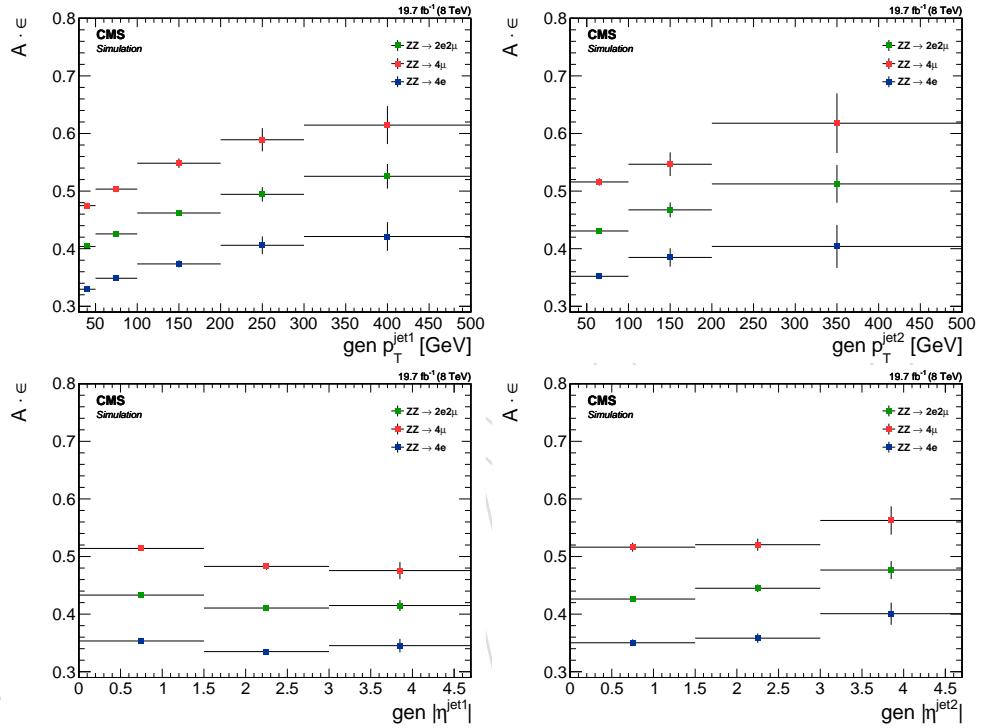


Figure 9: Acceptance ($A \cdot \epsilon$) computed in the wider fiducial region as a function of p_T^{jet1} (top left), p_T^{jet2} (top right), η^{jet1} (bottom left) and η^{jet2} distributions, according to the final state: 4μ (green), $4e$ (blue), $2e2\mu$ (red). Distributions are obtained using the `Powheg` set of samples for m_{ZZ} and the `MadGraph` set of samples for the other variables.

Table 10: The total ZZ production cross-section as measured in each decay channel and for the combination of all channels in the tight fiducial region.

Final state	Total cross-section [fb]
$ZZ \rightarrow 4\mu$	4.74 ± 0.54 (stat.) $^{+0.17}_{-0.18}$ (syst.)
$ZZ \rightarrow 4e$	4.99 ± 0.69 (stat.) $^{+0.32}_{-0.35}$ (syst.)
$ZZ \rightarrow 2e2\mu$	10.77 ± 0.90 (stat.) $^{+0.55}_{-0.59}$ (syst.)
$ZZ \rightarrow 4\ell$	20.50 ± 1.26 (stat.) $^{+0.82}_{-0.85}$ (syst.)

225 with $\eta^{jet} < 4.7$ and central jets with $\eta^{jet} < 2.4$ are considered.

226

227 The cross-section in each bin of each observable is determined from the event yields sub-
228 tracting the backgrounds. Each distribution is then corrected for event selection efficiencies
229 and for detector resolution effects in order to be compared with predictions from event gen-
230 erators. The correction procedure is based on unfolding techniques, as implemented in the
231 `RooUnfold` toolkit [27], which provides both singular value decomposition (SVD) [28] and the
232 d'Agostini [29] methods. Both algorithms use a response matrix that correlates the observable
233 with and without detector effects. Regularization parameters can be tuned to obtain results
234 that are robust against numerical instabilities and statistical fluctuations. The differential cross-
235 section is then derived by dividing the corrected number of events by the integrated luminosity,
236 the branching ratio and the bin width.

237

238 For each measured distribution, a response matrix is evaluated using two different sets of gen-
239 erators: the first one is composed of signal samples generated with `MadGraph` ($qq/qg/gg \rightarrow$
240 ZZ), `MCFM` ($gg \rightarrow ZZ$) and `Phantom` ($qq \rightarrow ZZ + 2jets$). The second one has the `Powheg` sam-
241 ple ($qq \rightarrow ZZ$) instead of the `MadGraph` one. The `MadGraph` set is the reference set for the
242 jet-related variables, while the `Powheg` one is used for check and comparison purposes. For
243 the m_{ZZ} distribution the role of the two sets of samples is switched.

244 As reported above, the signal definition of the generated events requires $60 < m_{Z_1}, m_{Z_2} < 120$
245 GeV. In order to minimize the model uncertainties due to unnecessary extrapolations of the
246 measurement outside experimentally well-described phase space regions, cross-section distri-
247 butions are also extracted in the tighter fiducial region, corresponding to the visible phase space
248 defined by the kinematic and geometrical acceptance of leptons.

249 The same selection as in the inclusive cross-section measurement is applied to the reconstructed
250 events.

251 Response matrices built using the reference set of samples and obtained in the tight fiducial
252 region are reported in Figures 10-18 for each variable.

253

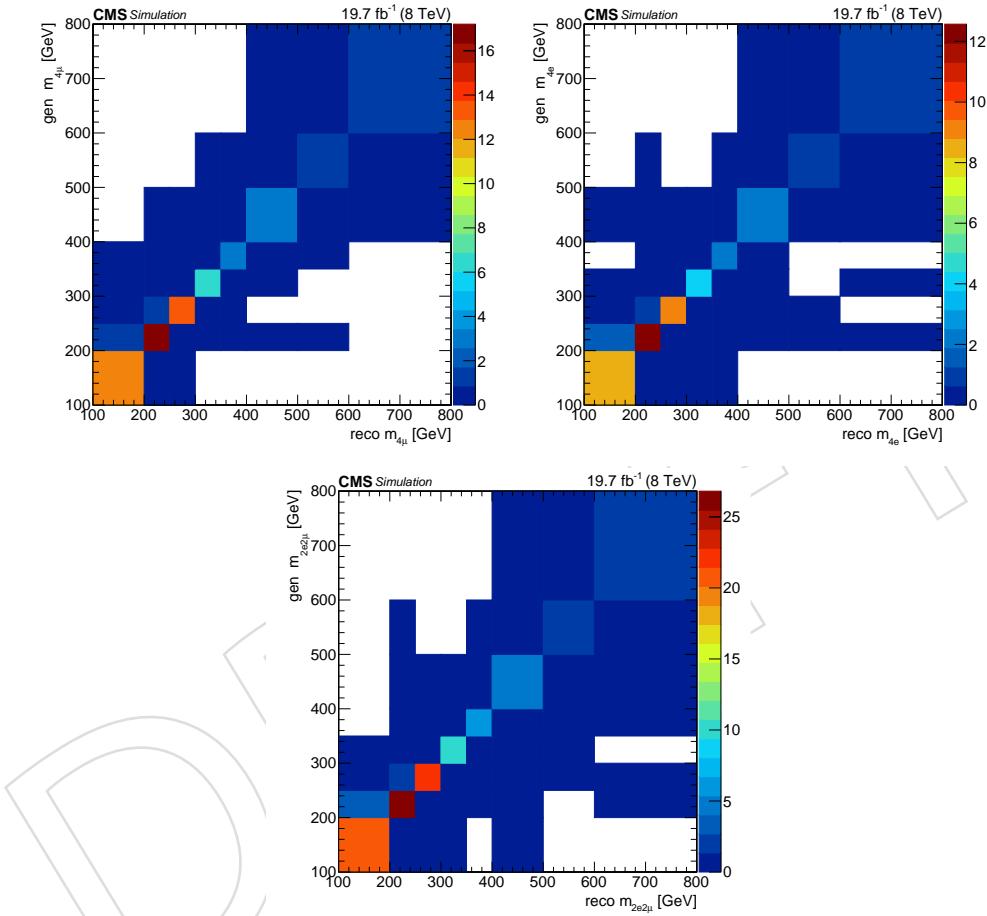


Figure 10: Response matrices for the m_{ZZ} distribution, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the Powheg set of samples. The tight fiducial region is considered.

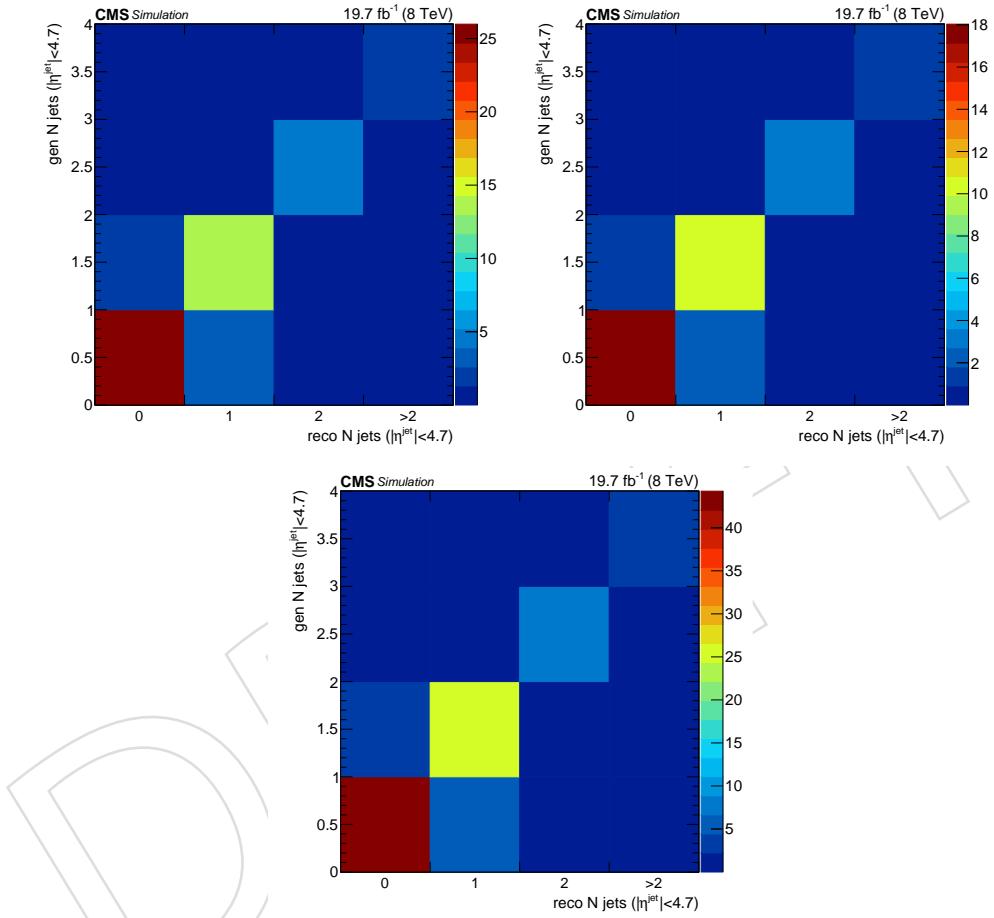


Figure 11: Response matrices for the N_{jets} distribution, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

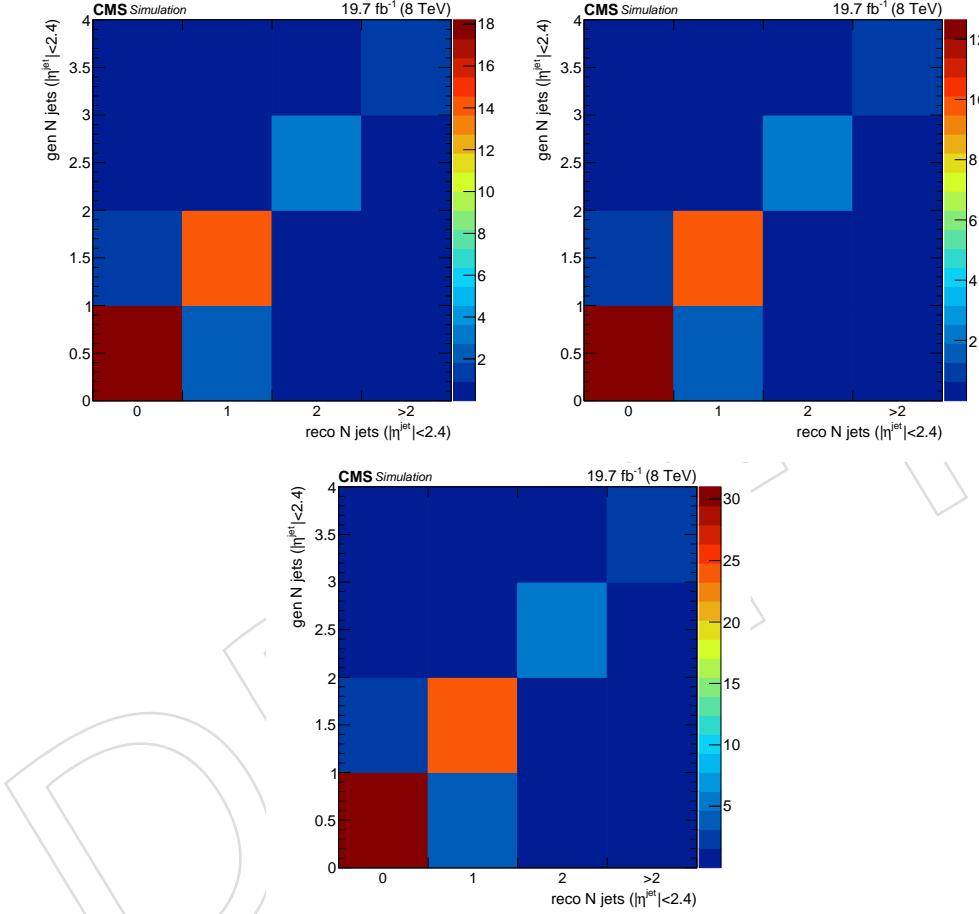


Figure 12: Response matrices for the N central jets distribution (with $\eta^{\text{jet}} < 2.4$) , according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

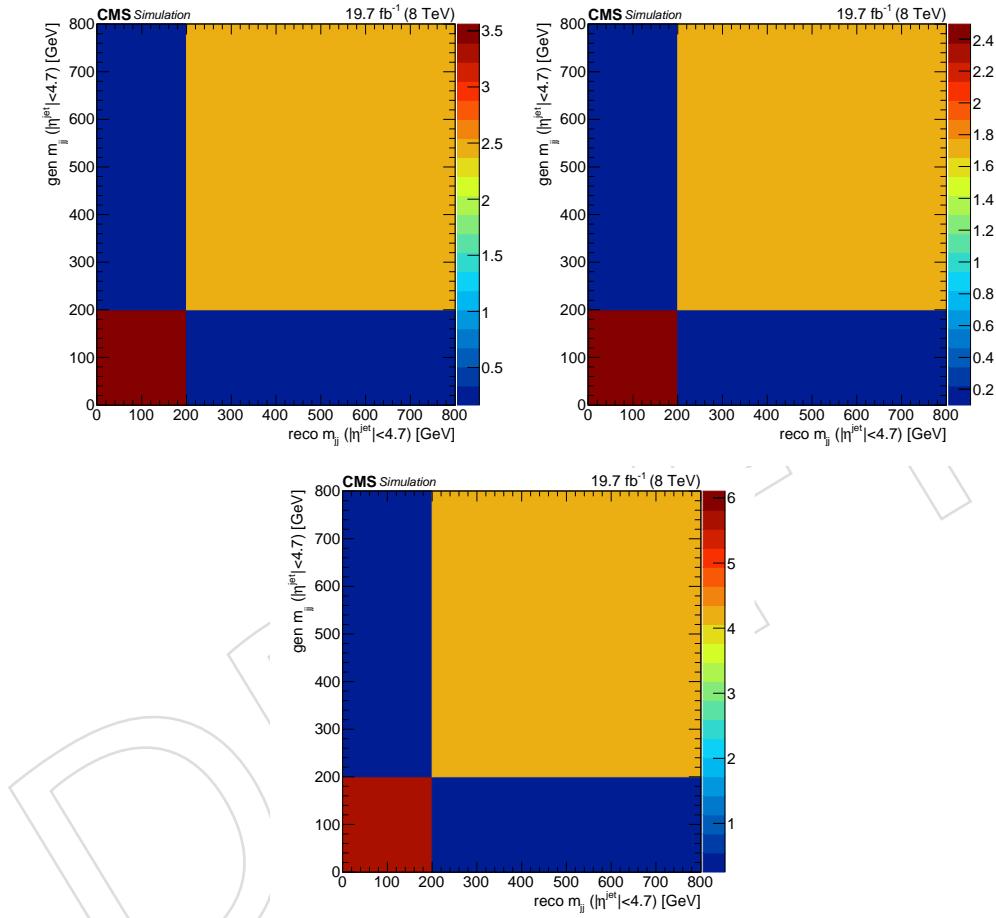


Figure 13: Response matrices for the m_{jj} distribution, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

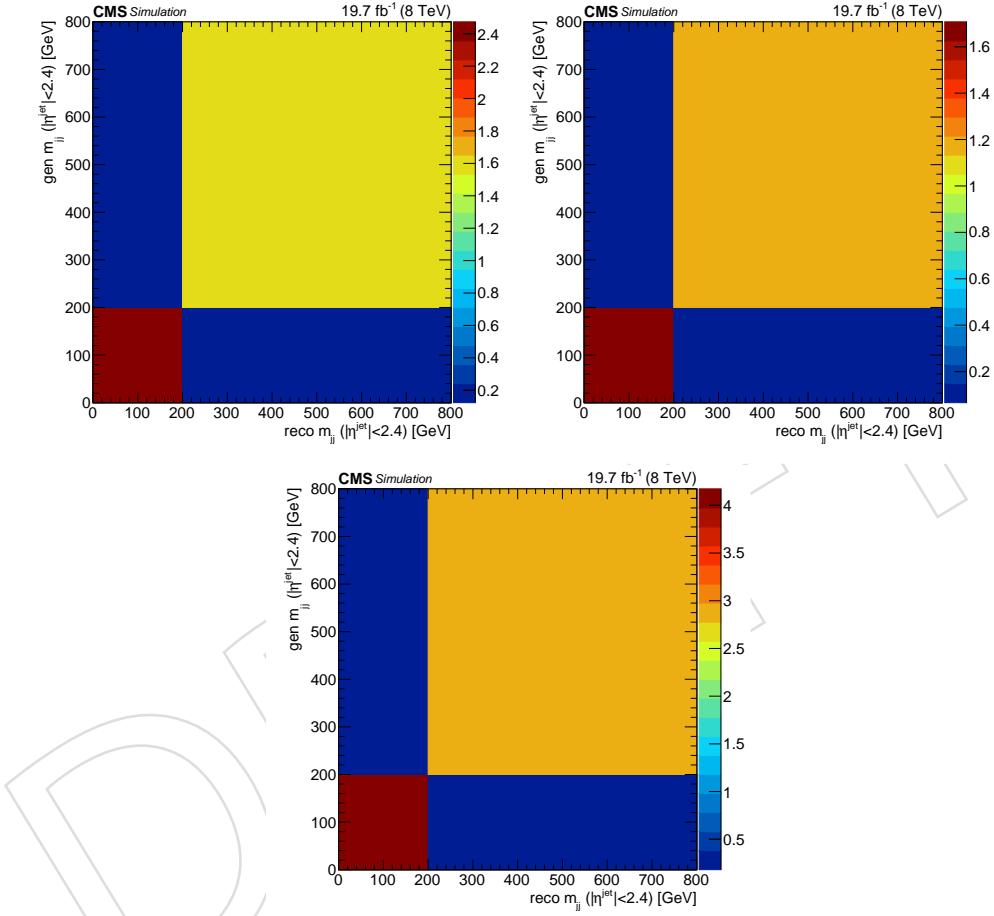


Figure 14: Response matrices for the m_{jj} distribution (with $\eta^{\text{jet}} < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

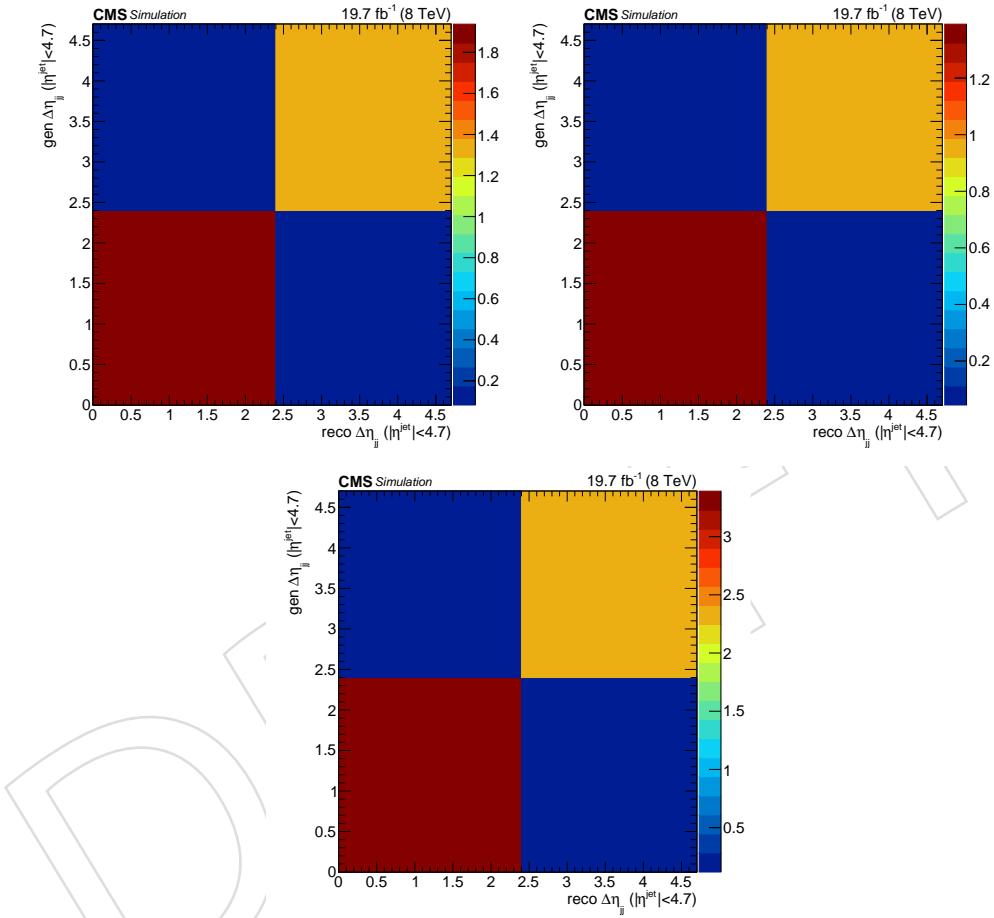


Figure 15: Response matrices for the $\Delta\eta_{jj}$ distribution, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

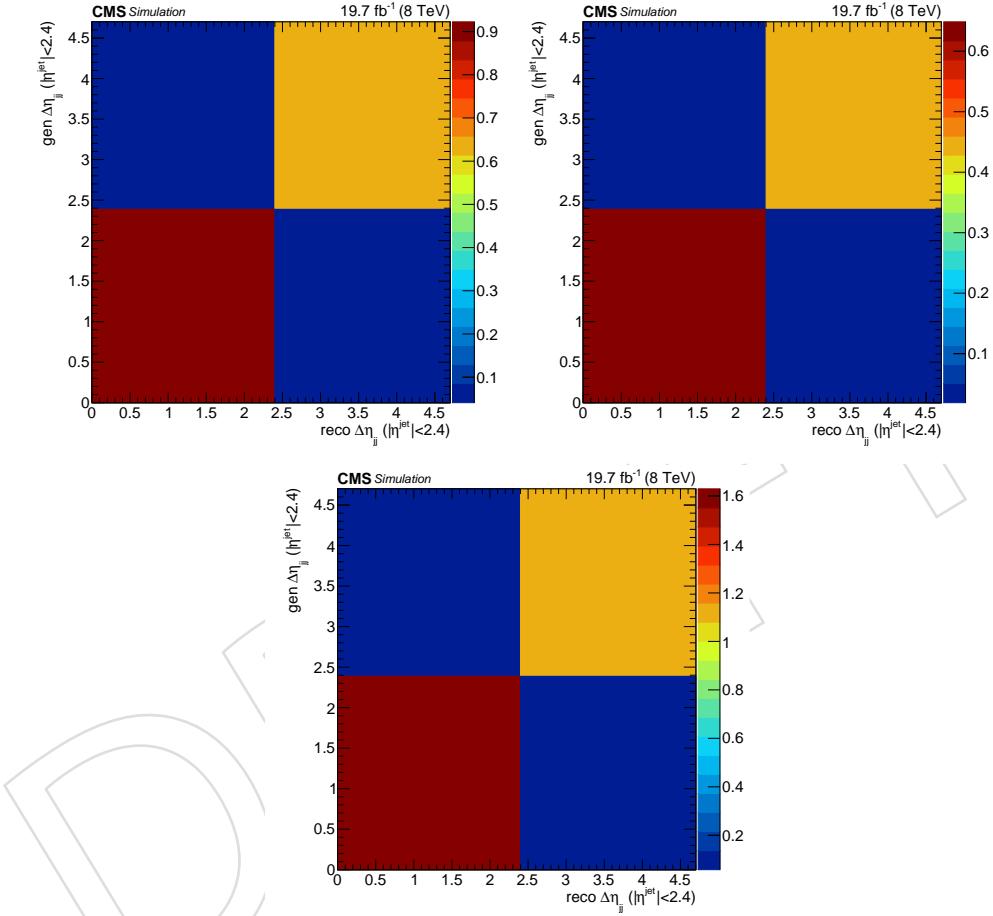


Figure 16: Response matrices for the $\Delta\eta_{jj}$ distribution (with $\eta^{jet} < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

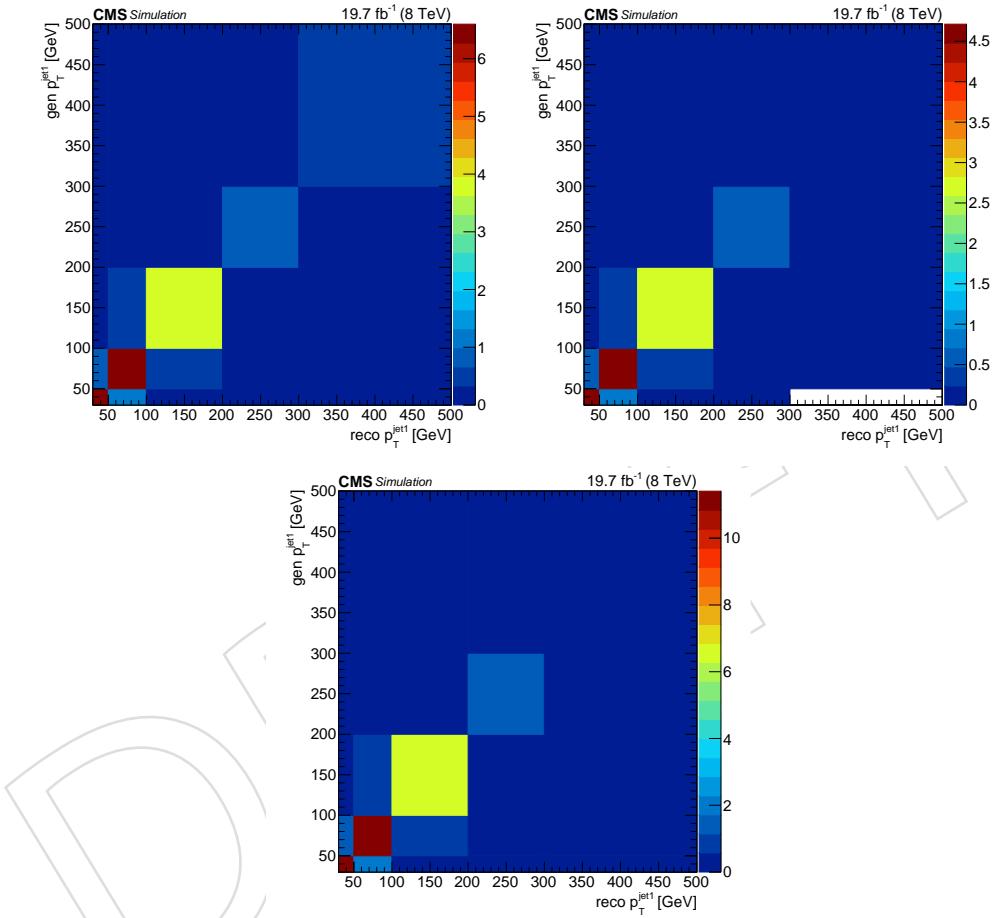


Figure 17: Response matrices for the p_T distribution of the leading jet, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

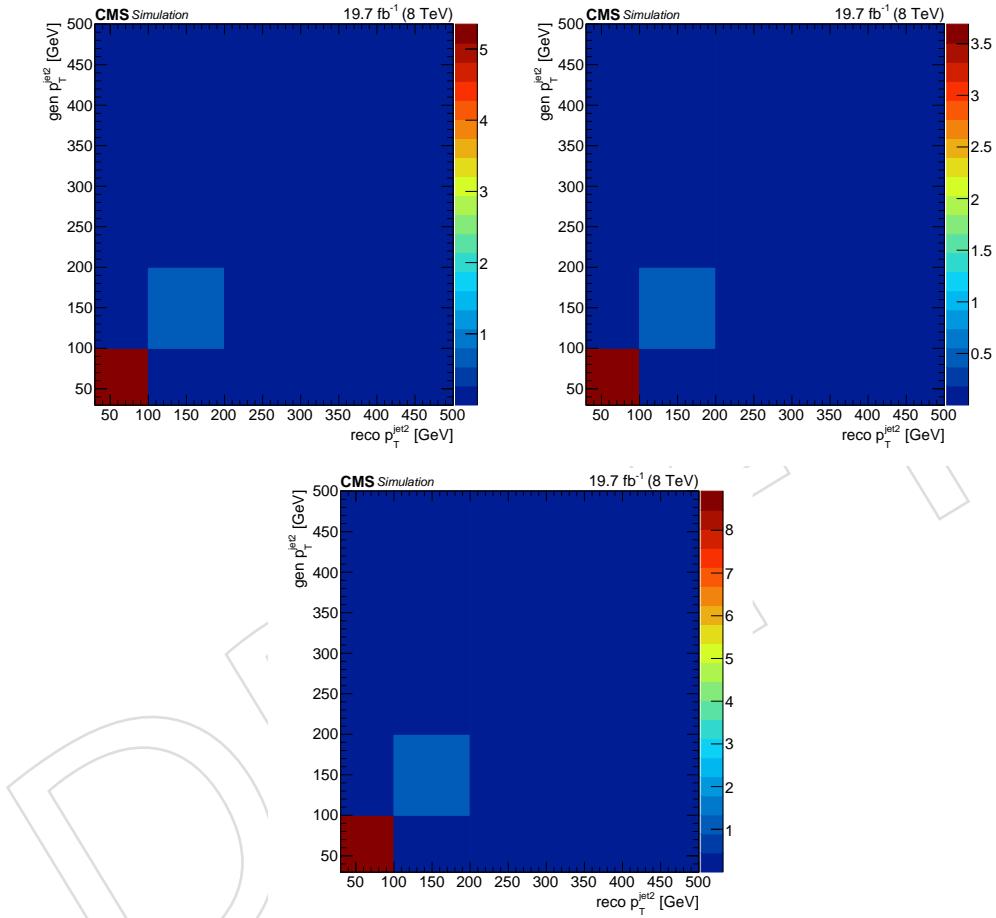


Figure 18: Response matrices for the p_T distribution of the sub-leading jet, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

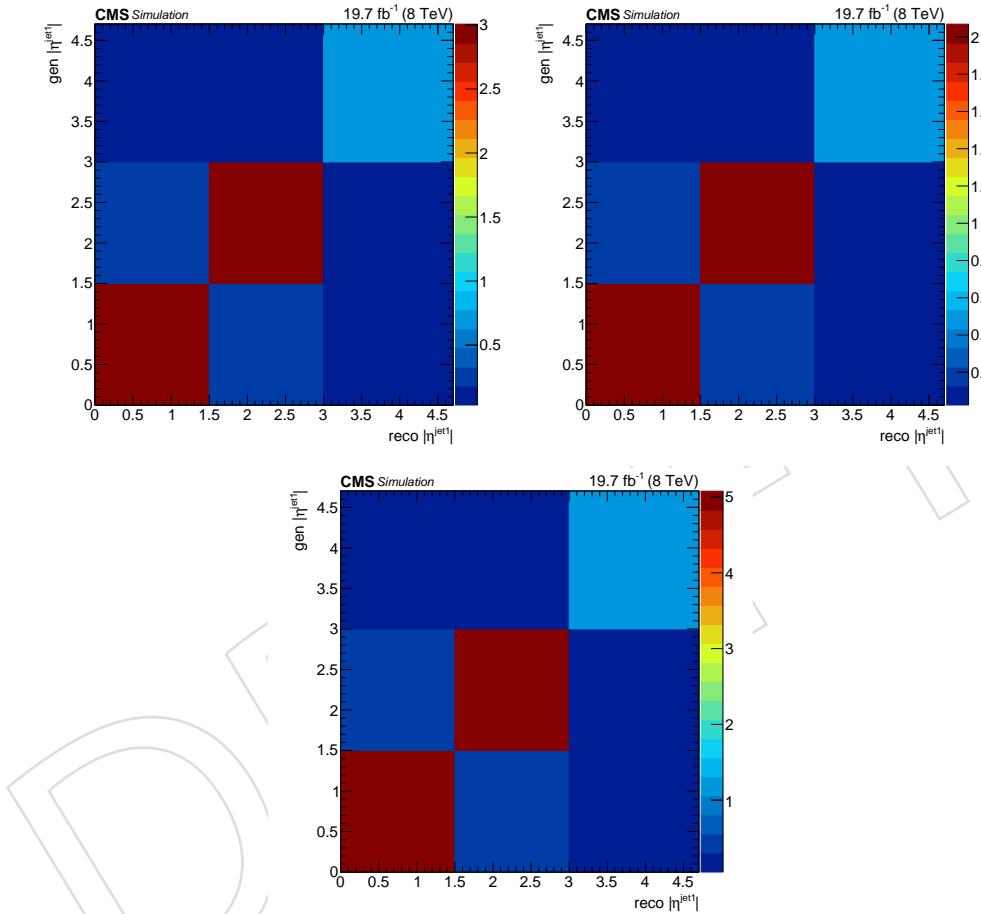


Figure 19: Response matrices for the η distribution of the leading jet, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

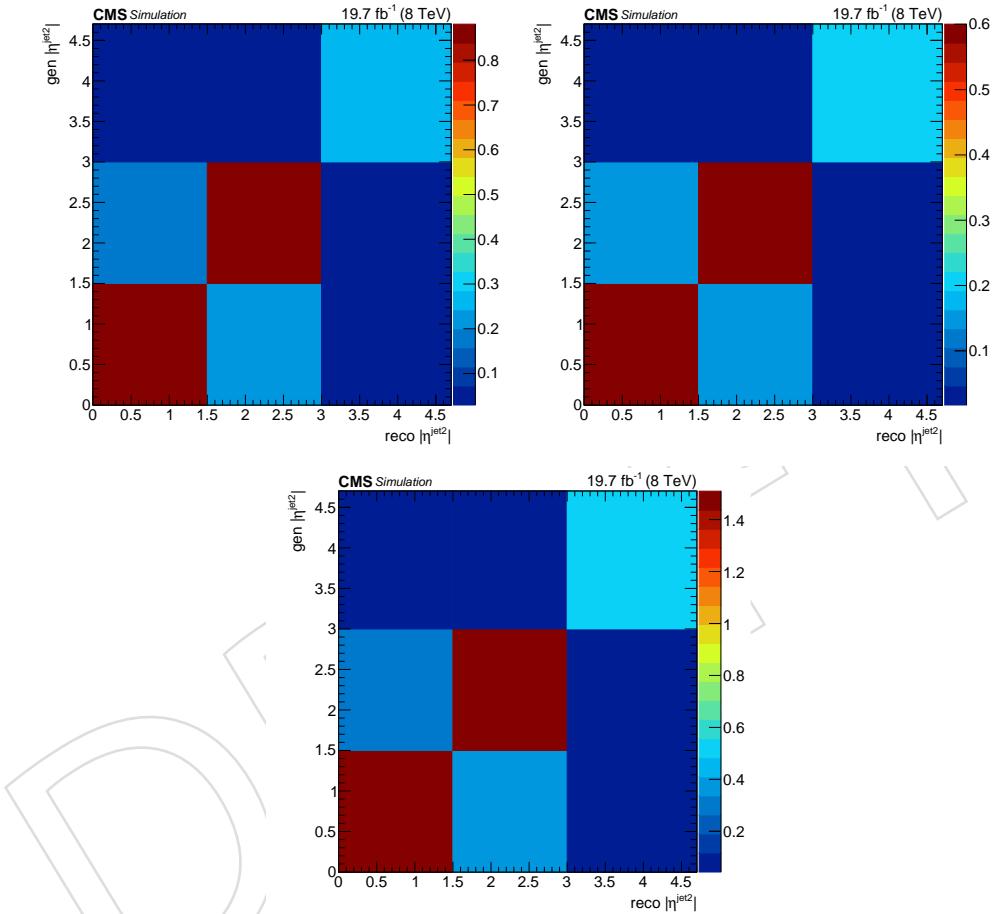


Figure 20: Response matrices for the η distribution of the sub-leading jet, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

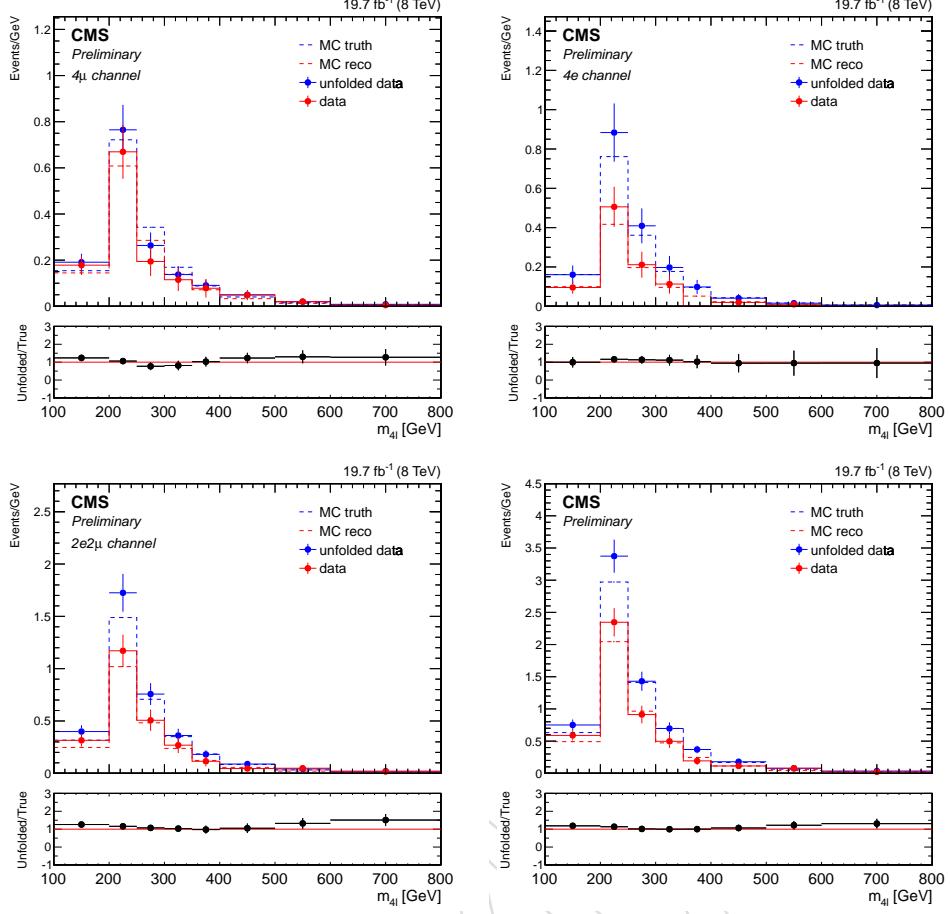


Figure 21: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) m_{ZZ} distributions, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). The unfolded distributions are obtained in the tight fiducial region using the SVD algorithm, with $k_{reg} = 4$, and compared to predictions from the Powheg set of samples.

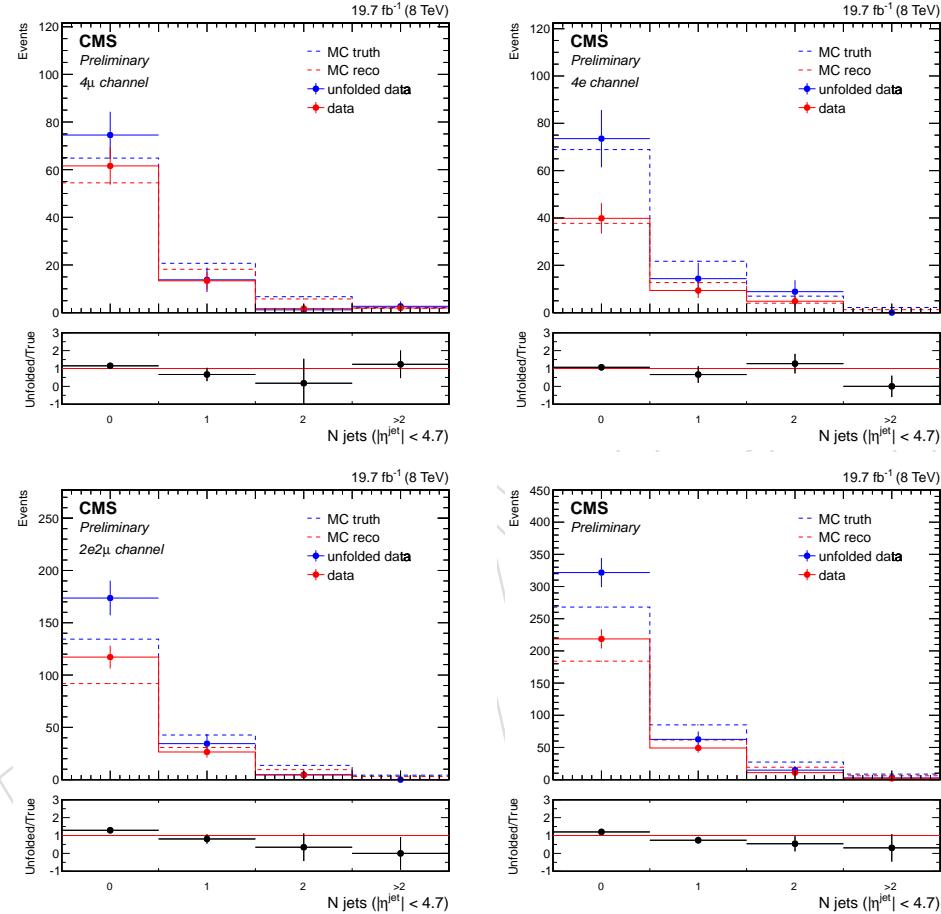


Figure 22: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) N_{jets} distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

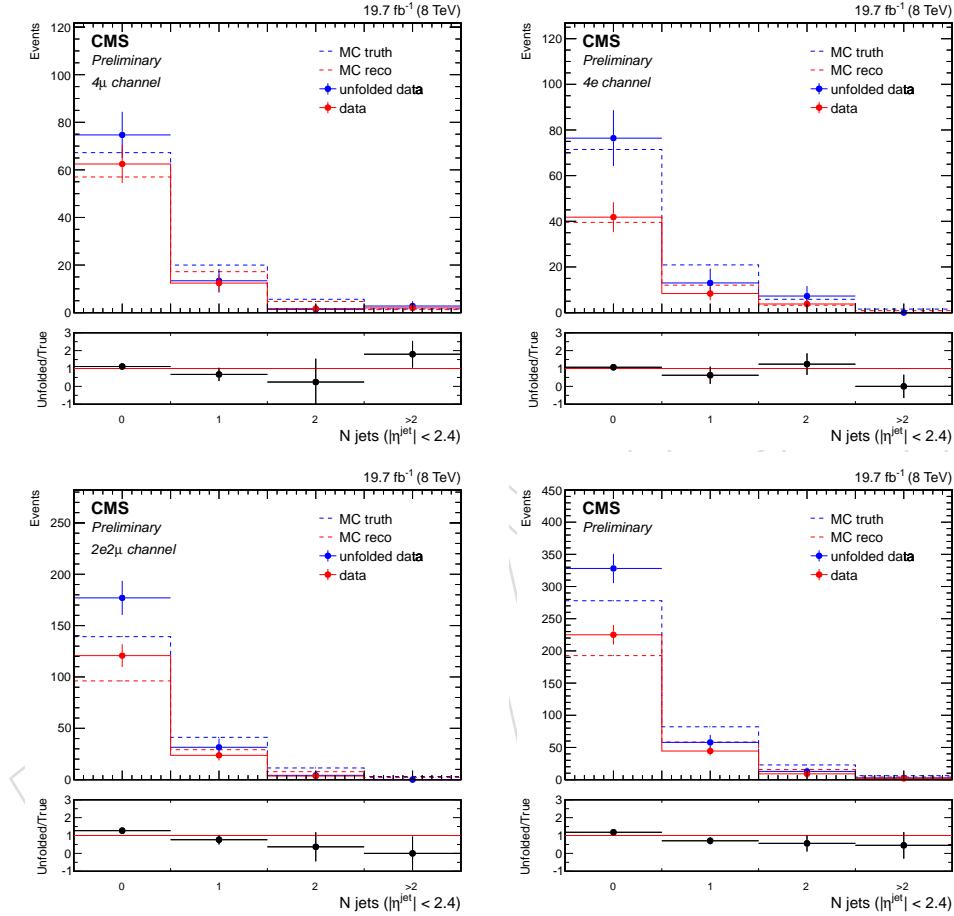


Figure 23: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) N central jets distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the `MadGraph` set of samples.

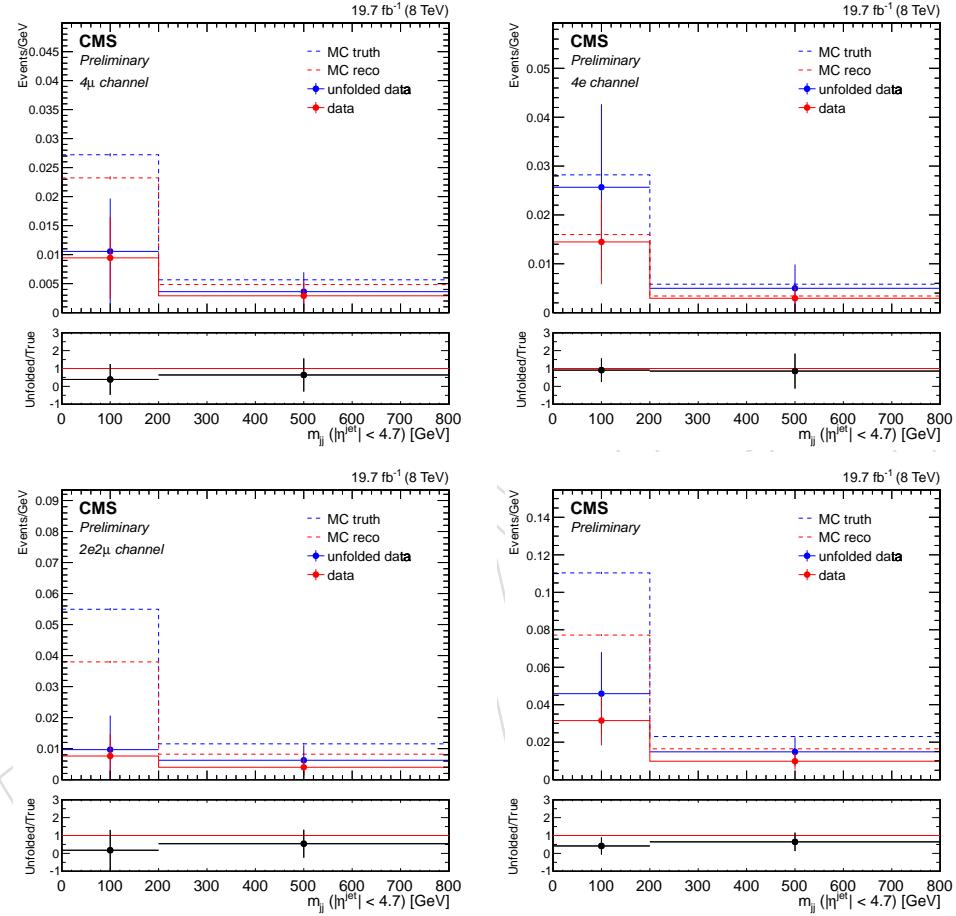


Figure 24: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) m_{jj} distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

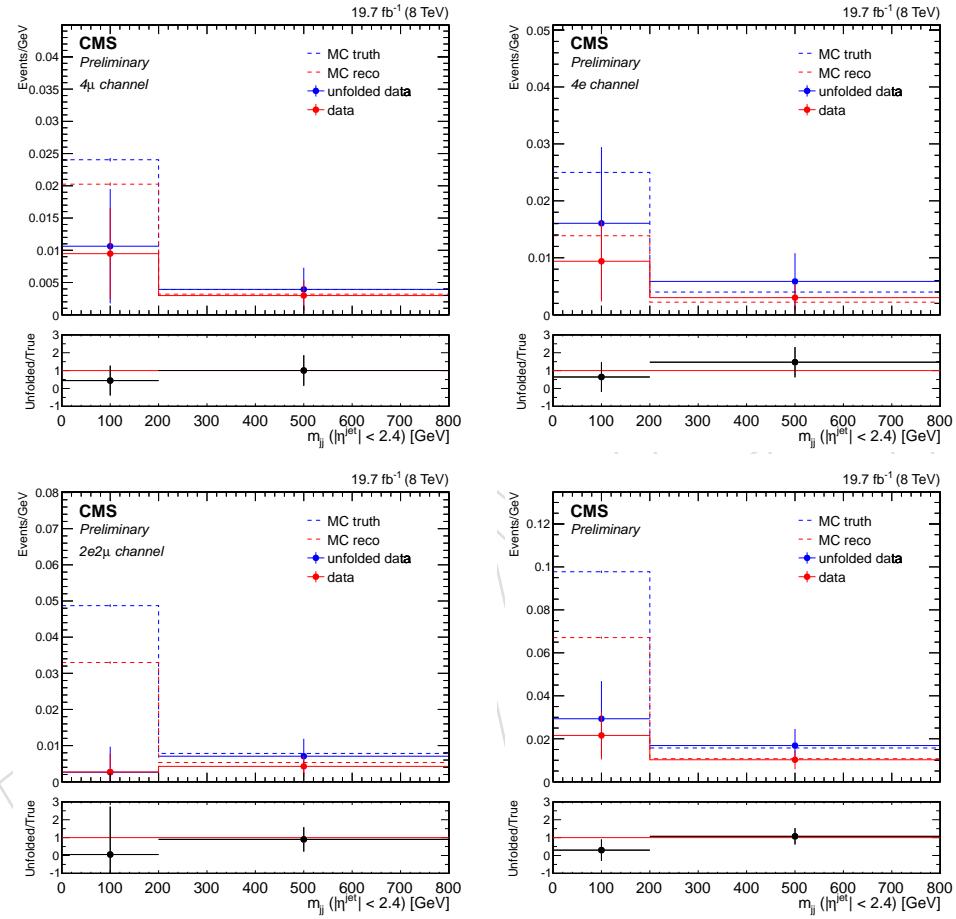


Figure 25: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) m_{jj} distributions (using central jets with $\eta^{jet} < 2.4$), according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

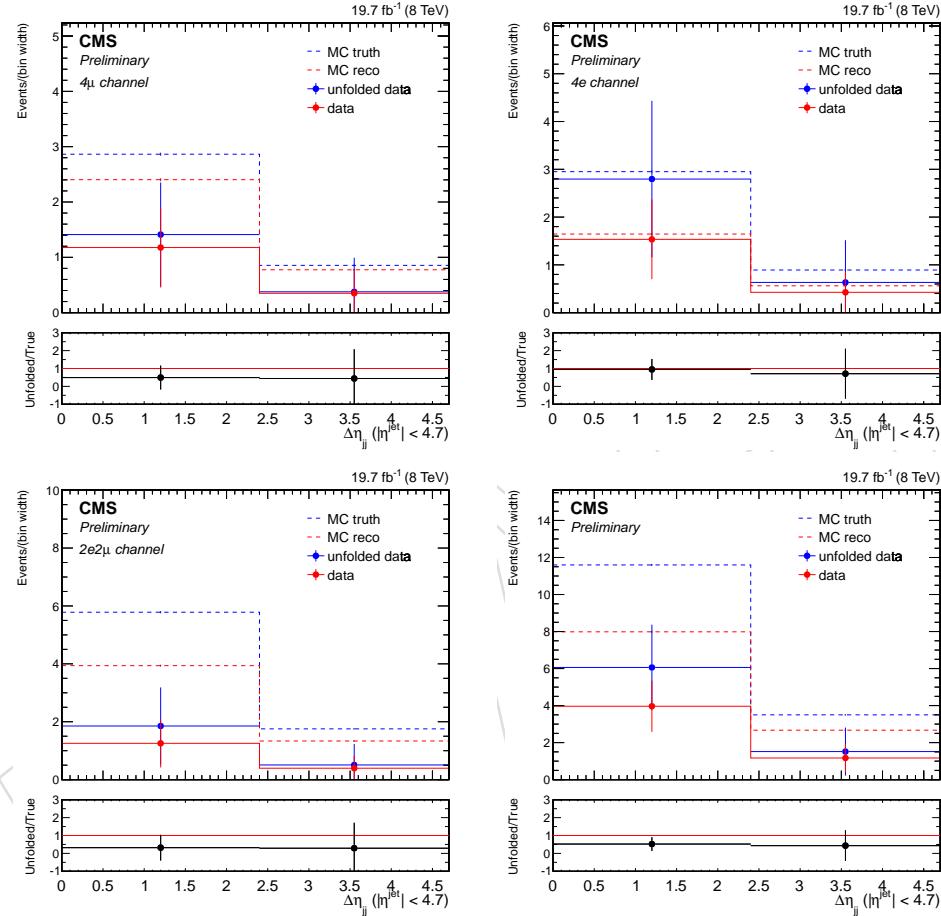


Figure 26: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) $\Delta\eta_{jj}$ distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

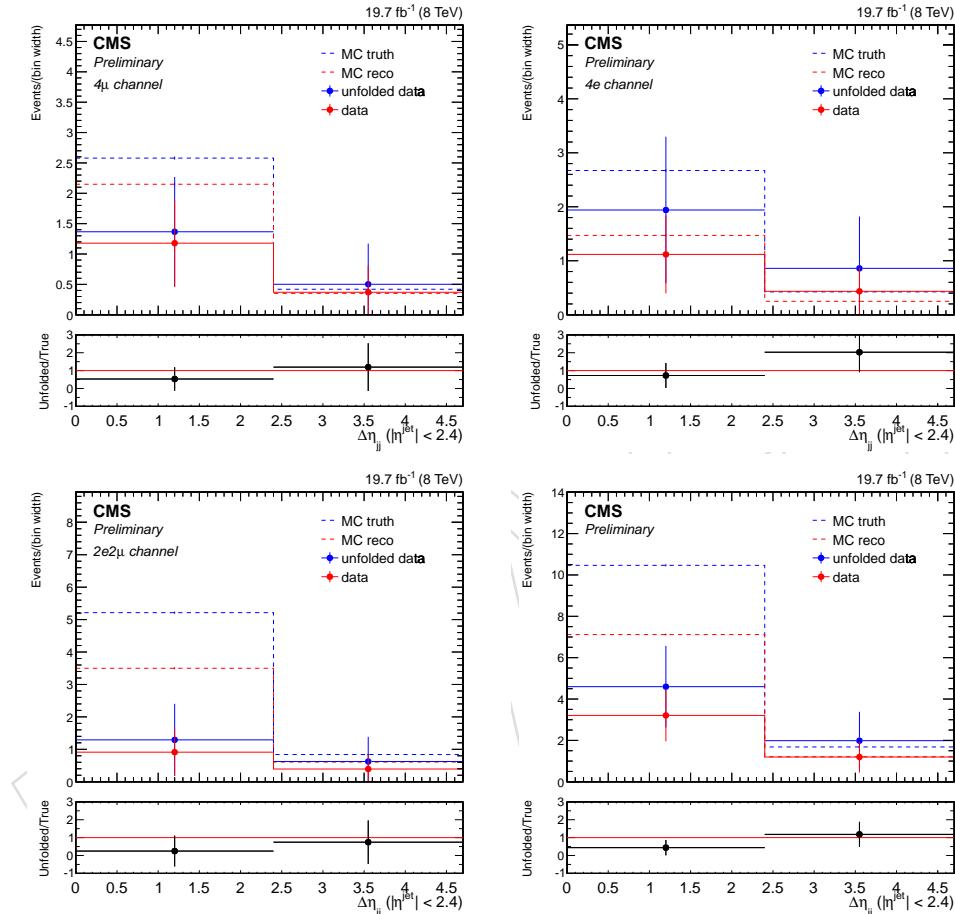


Figure 27: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) $\Delta\eta_{jj}$ distributions (using central jets with $\eta^{jet} < 2.4$), according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

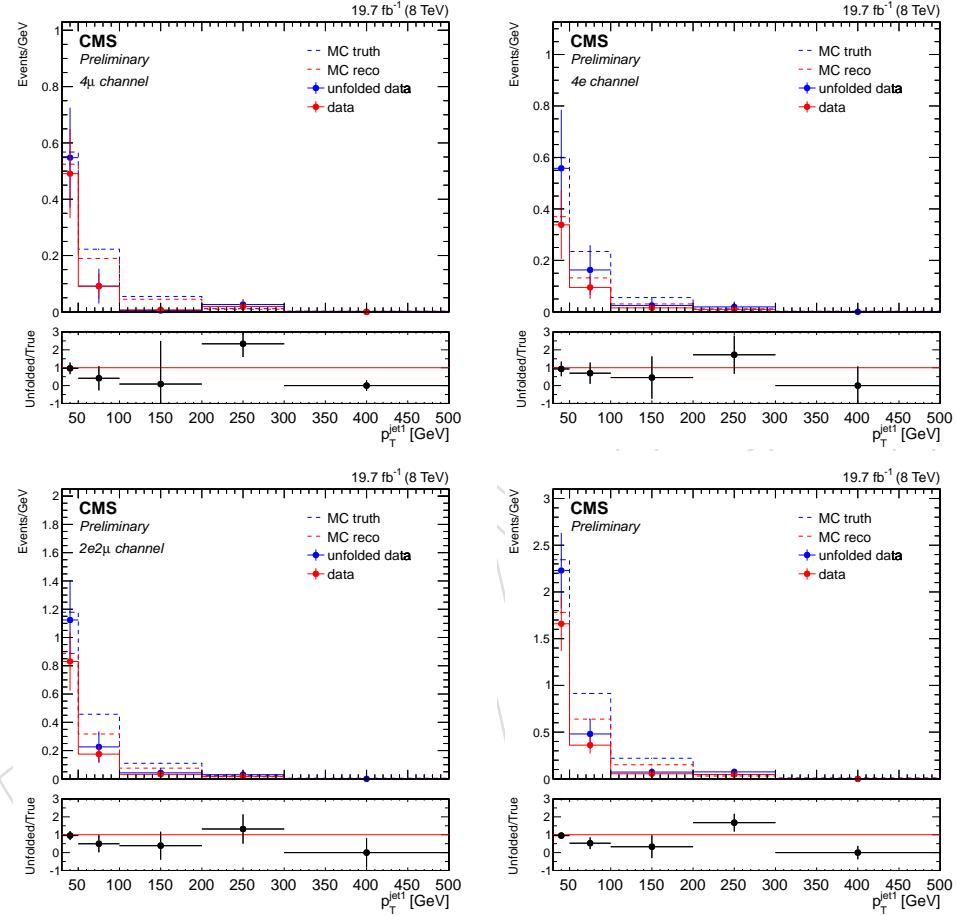


Figure 28: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) p_T^{jet1} distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D’Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

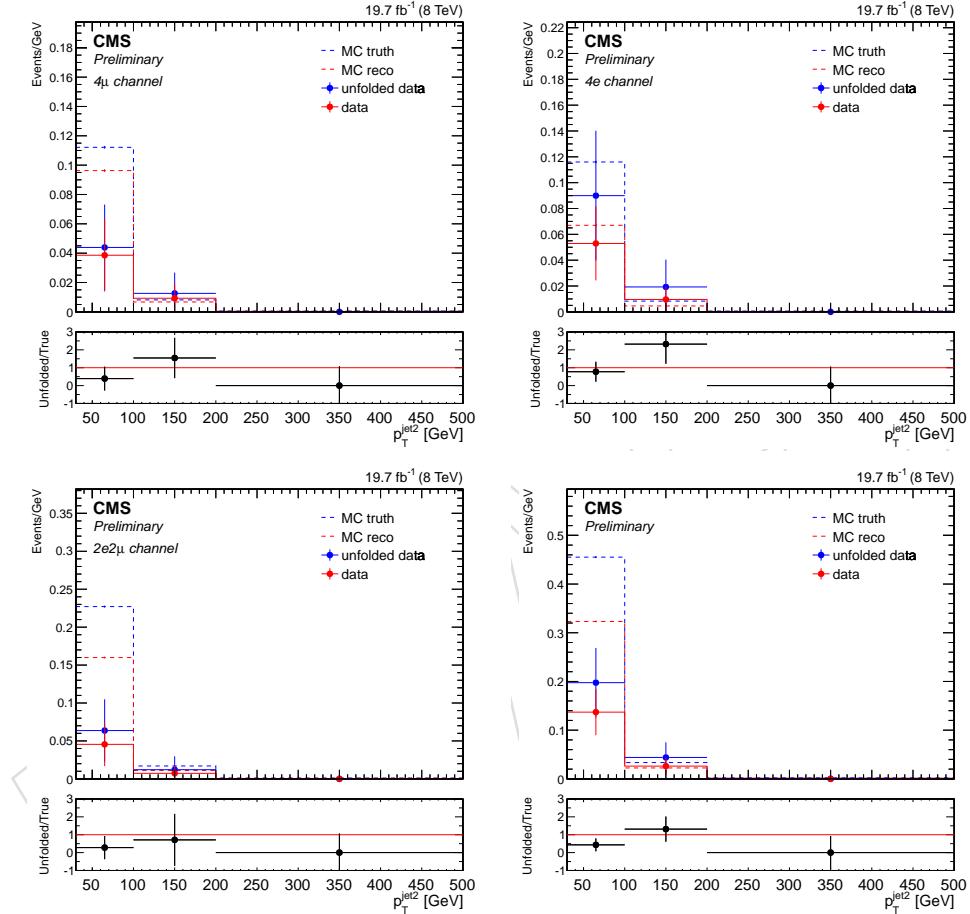


Figure 29: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) p_T^{jet2} distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

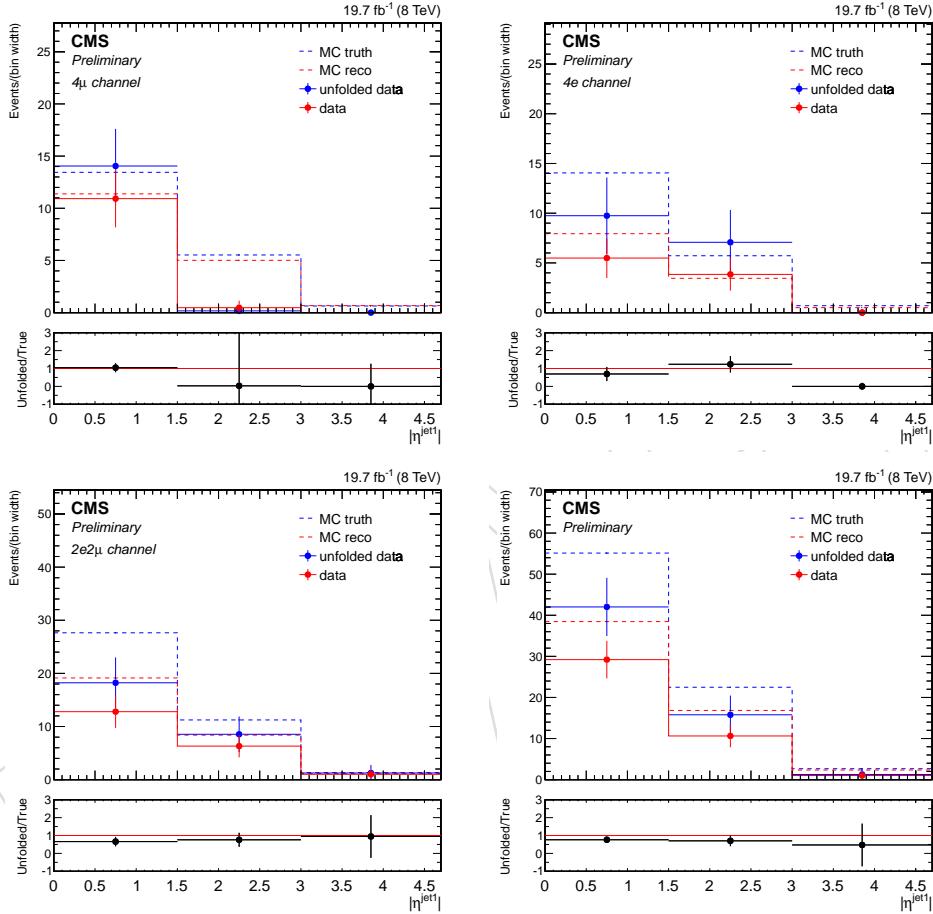


Figure 30: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) η_{jet^1} distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

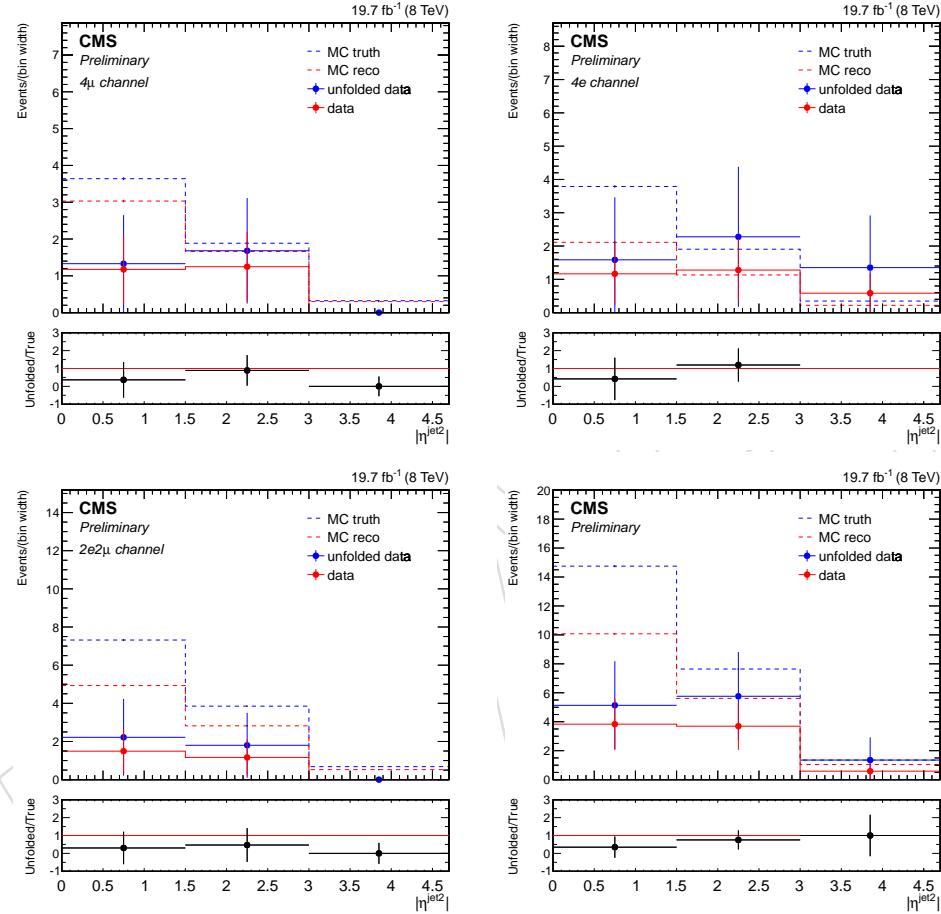


Figure 31: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) $\eta^{\text{jet}2}$ distributions, according to the final state: (top left) 4μ , (top right) $4e$, (bottom left) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 4 iterations, and compared to predictions from the MadGraph set of samples.

The systematic uncertainties in each bin are assessed from the variations of the nominal cross-section, by repeating the full analysis for every systematic variation. The difference with respect to the nominal value is taken as the systematic uncertainty for each bin and each measured observable. By using this method [30], the possible correlations of the systematic uncertainties between bins are taken into account. Due to the normalization, those systematic uncertainties that are correlated across all bins of the measurement, and therefore mainly affect the normalization, cancel out at least partly. The errors also include the statistical error propagation through the unfolding method using the covariance matrix and the difference in the response matrix from MadGraph and Powheg.

For the m_{ZZ} observable, the unfolded distributions are obtained with the SVD algorithm and the regularization parameter k_{reg} is chosen equal to 4. On the other hand, the other distributions are unfolded using the D'Agostini method, with 4 iterations. The unfolded distributions are presented from Figure 21 to Figure 31. The measurement is compared to the predictions from the MadGraph and Powheg sets of samples. The differential distributions normalized to the unity obtained in the tight fiducial region are presented from Figure 32 to Figure 42 for the 4μ , $4e$ and $2e2\mu$ decay channels and for their combination. The ratios between the measured distribution and the expected one from both the MadGraph and Powheg sets of samples are reported below each plot. Measurements are also compared with a set of samples in which MadGraph5_aMCatNLO is used to generate $q\bar{q} \rightarrow ZZ$ signal processes. The Powheg and MadGraph5_aMCatNLO theoretical uncertainties due to the scale choice are estimated varying independently μ_R and μ_F by a factor from 0.5 to 2.

The total ZZ production cross-section as measured in each decay channel and for the combination of all channels in the wide and tight fiducial regions classified with respect to the number of jets and central jets is reported in Tables 11, 12, 13 and 14.

Table 11: The total ZZ production cross-section as measured in each decay channel and for the combination of all channels in the fiducial region $60 < m_{Z_1}, m_{Z_2} < 120$ GeV classified with respect to the number of jets.

Process	Number of jets ($ \eta^{jet} < 4.7$)	Total cross-section [pb]
$pp \rightarrow ZZ(4\mu)$	0	6.27 ± 0.82 (stat.) ± 0.39 (syst.)
$pp \rightarrow ZZ(4e)$	0	5.84 ± 0.96 (stat.) ± 0.48 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	0	7.03 ± 0.67 (stat.) ± 0.56 (syst.)
$pp \rightarrow ZZ(4\ell)$	0	6.48 ± 0.46 (stat.) ± 0.40 (syst.)
$pp \rightarrow ZZ(4\mu)$	1	1.03 ± 0.38 (stat.) ± 0.14 (syst.)
$pp \rightarrow ZZ(4e)$	1	1.02 ± 0.47 (stat.) ± 0.16 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	1	1.25 ± 0.32 (stat.) ± 0.26 (syst.)
$pp \rightarrow ZZ(4\ell)$	1	1.11 ± 0.22 (stat.) ± 0.10 (syst.)
$pp \rightarrow ZZ(4\mu)$	2	0.08 ± 0.11 (stat.) ± 0.06 (syst.)
$pp \rightarrow ZZ(4e)$	2	0.60 ± 0.33 (stat.) ± 0.07 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	2	0.16 ± 0.13 (stat.) ± 0.08 (syst.)
$pp \rightarrow ZZ(4\ell)$	2	0.15 ± 0.08 (stat.) ± 0.03 (syst.)
$pp \rightarrow ZZ(4\mu)$	> 2	0.18 ± 0.14 (stat.) ± 0.06 (syst.)
$pp \rightarrow ZZ(4e)$	> 2	$(0.32 \pm 0.19$ (stat.) ± 0.43 (syst.)) $\cdot 10^{-3}$
$pp \rightarrow ZZ(2e2\mu)$	> 2	$(0.89 \pm 0.81$ (stat.) ± 1.23 (syst.)) $\cdot 10^{-4}$
$pp \rightarrow ZZ(4\ell)$	> 2	$(1.1 \pm 0.7$ (stat.) ± 1.2 (syst.)) $\cdot 10^{-4}$

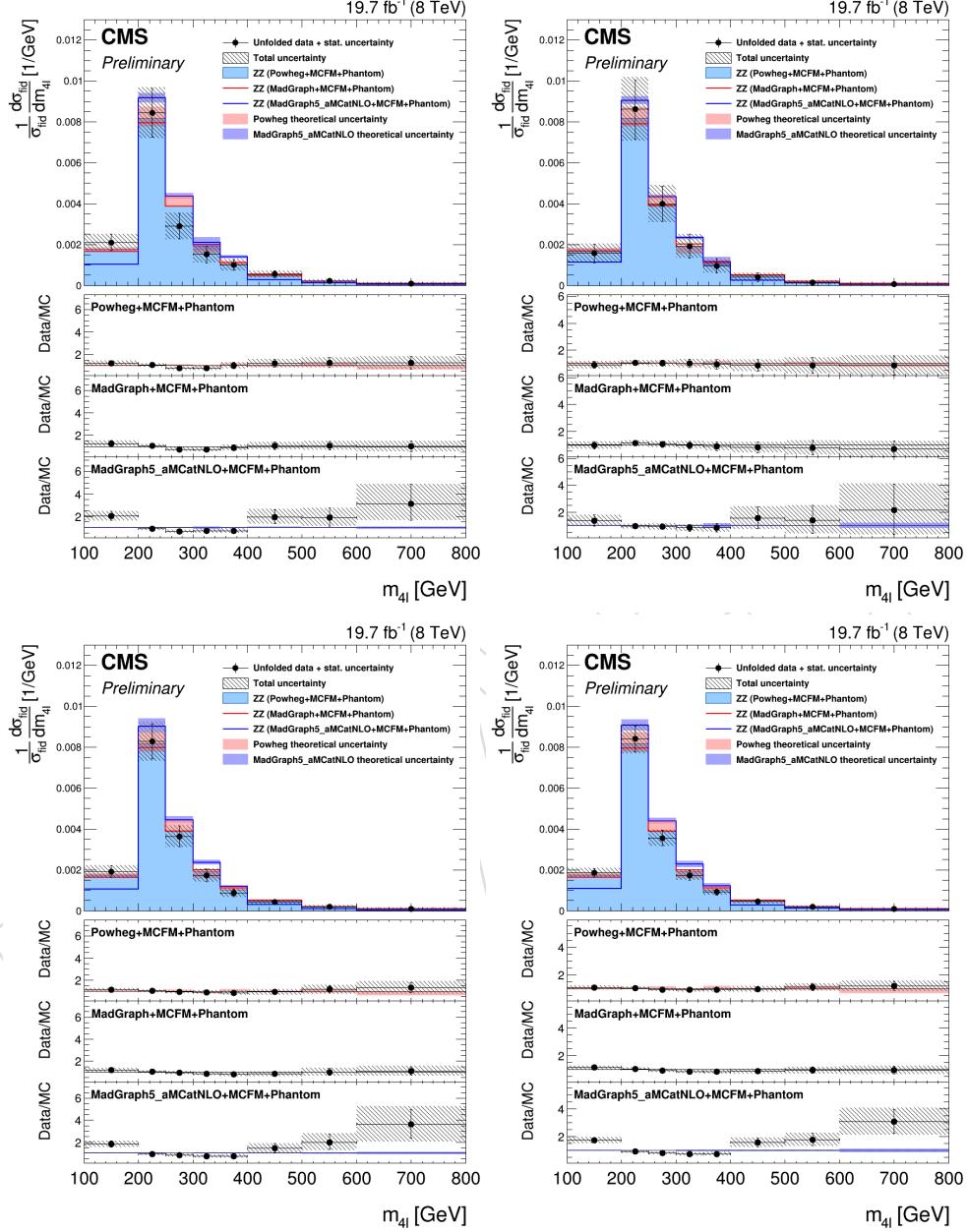


Figure 32: Normalized differential cross-sections as a function of the invariant mass of the 4 lepton system, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the Powheg, MadGraph and MadGraph5_aMCatNLO sets of samples.

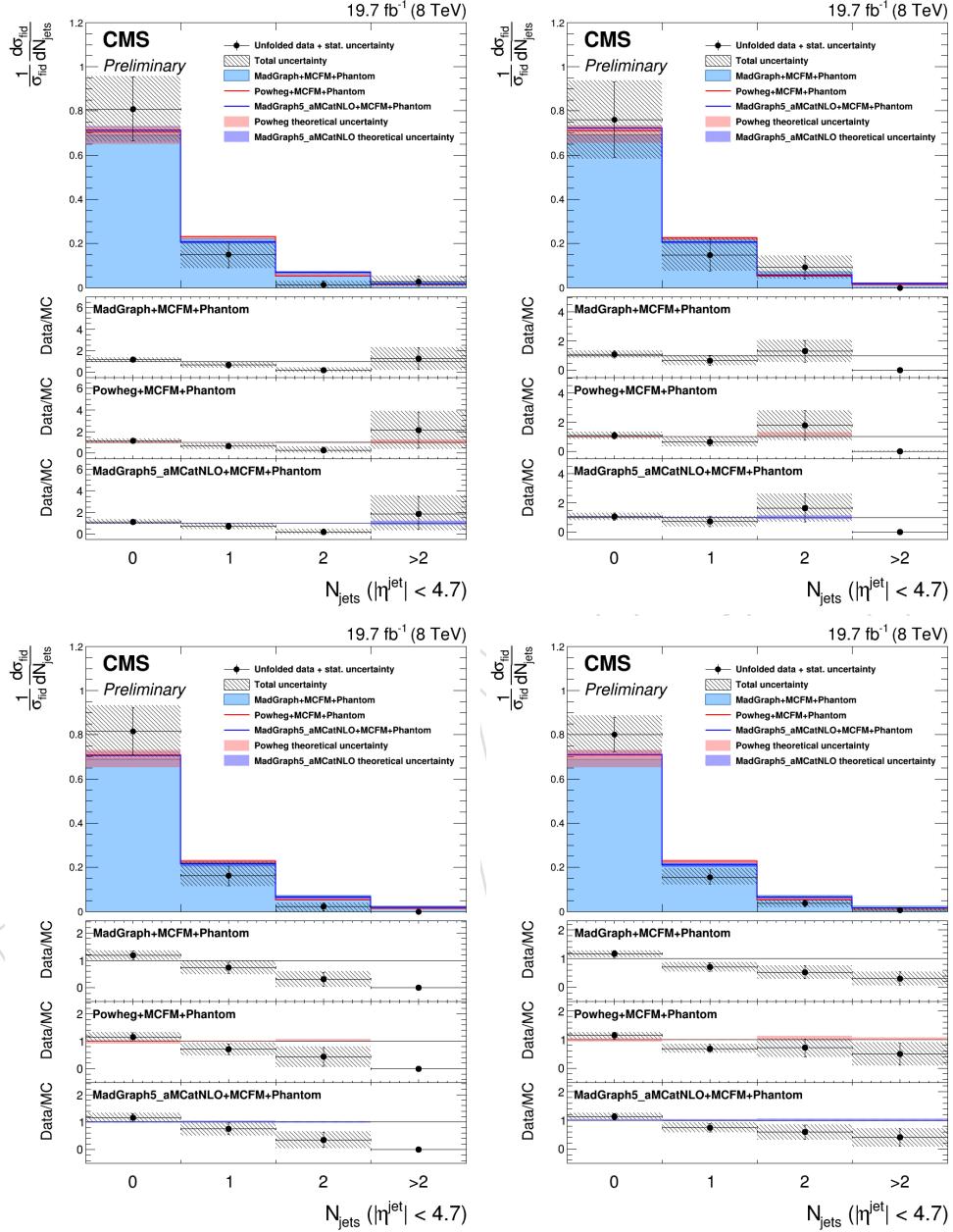


Figure 33: Normalized differential cross-sections as a function of the number of jets in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

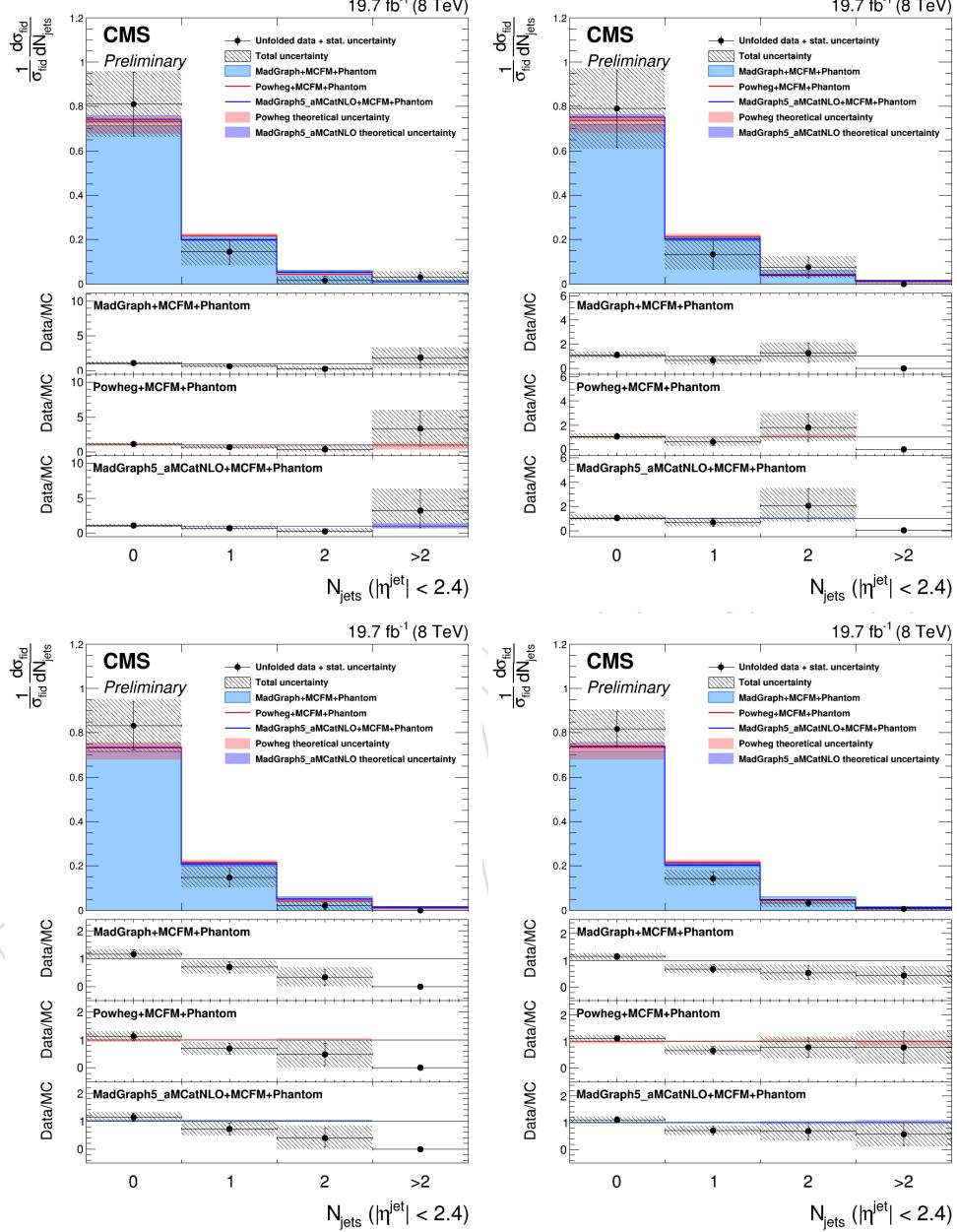


Figure 34: Normalized differential cross-sections as a function of the number of central jets in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

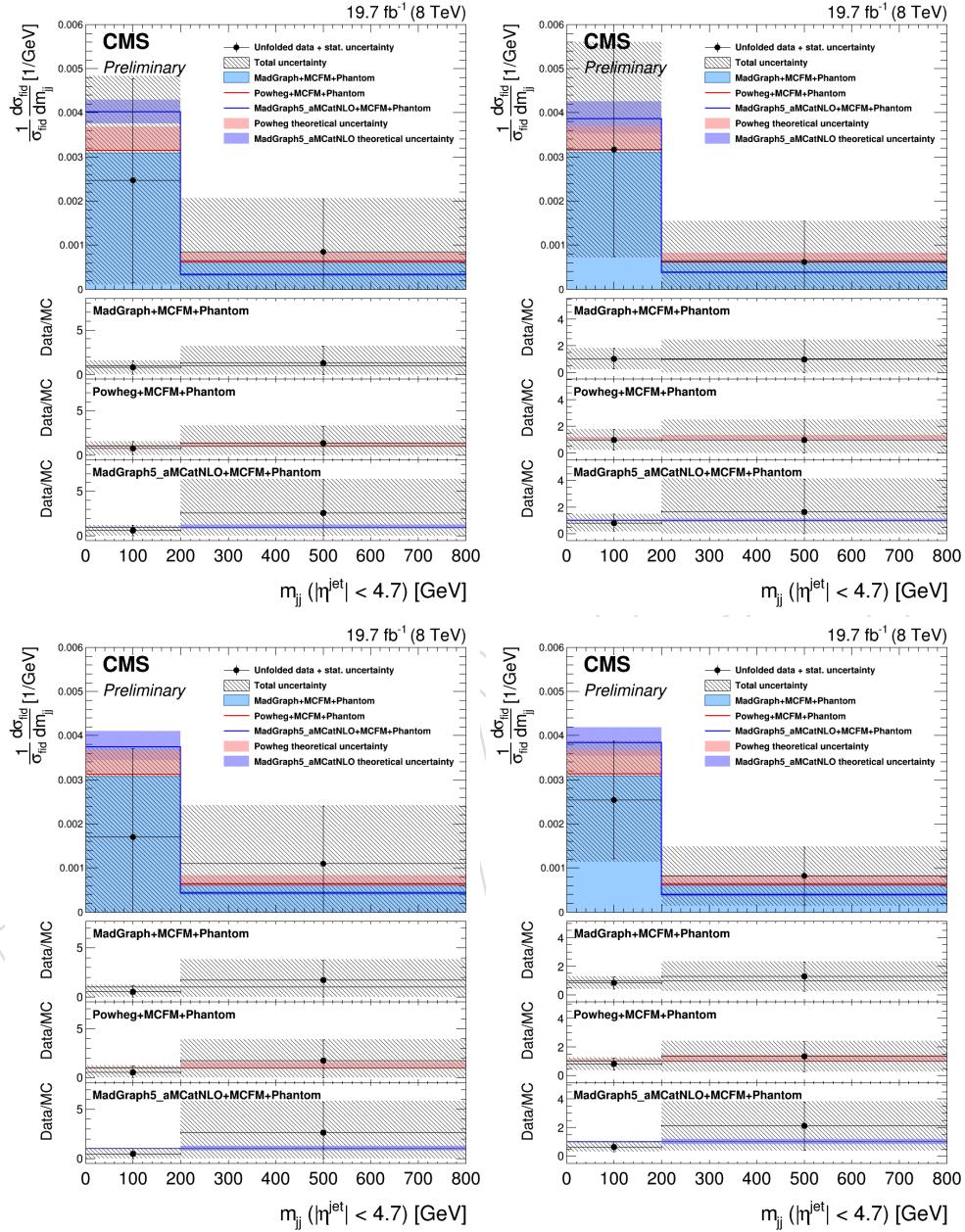


Figure 35: Normalized differential cross-sections as a function of the invariant mass of the most energetic jets in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

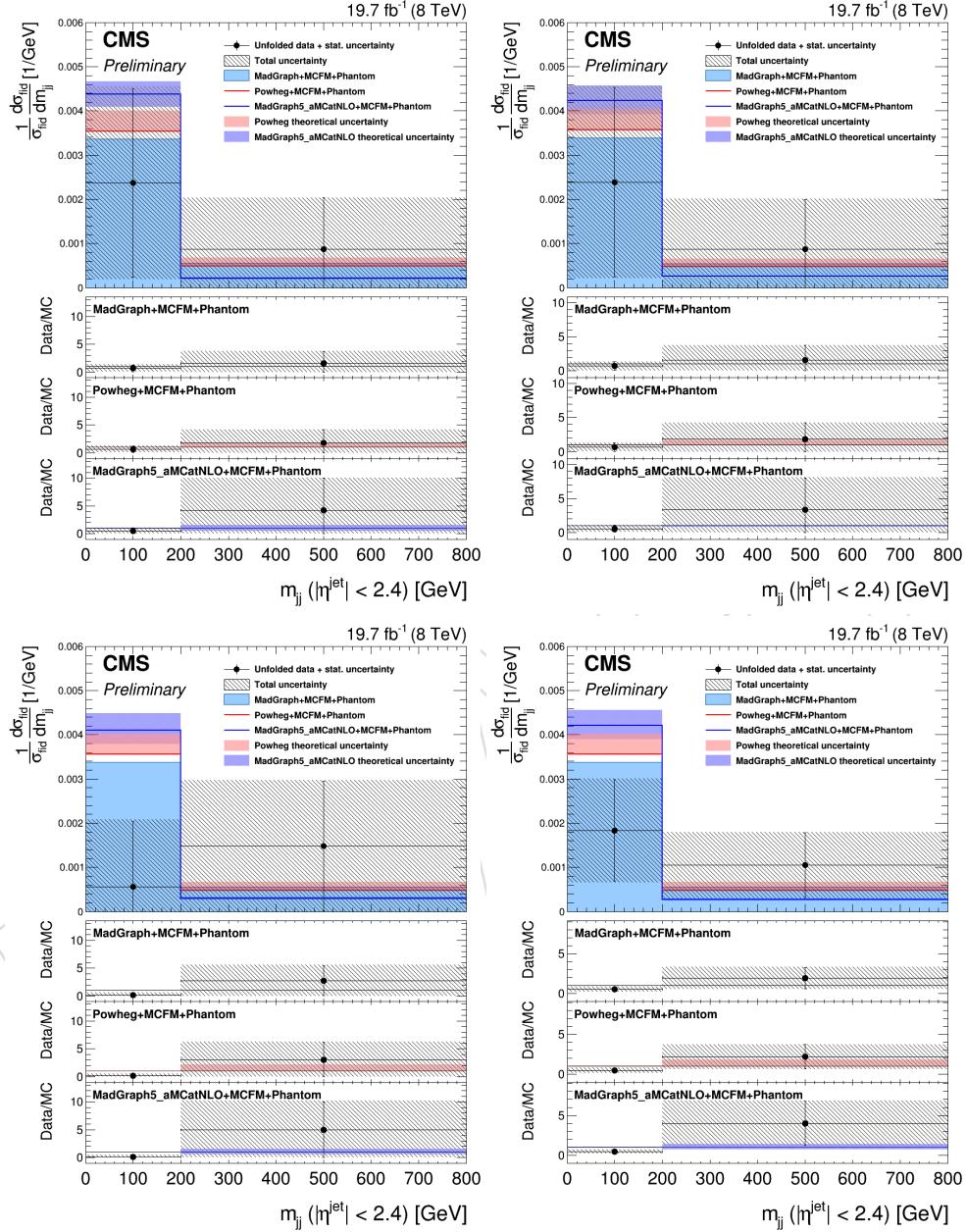


Figure 36: Normalized differential cross-sections as a function of the invariant mass of the most energetic central jets in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

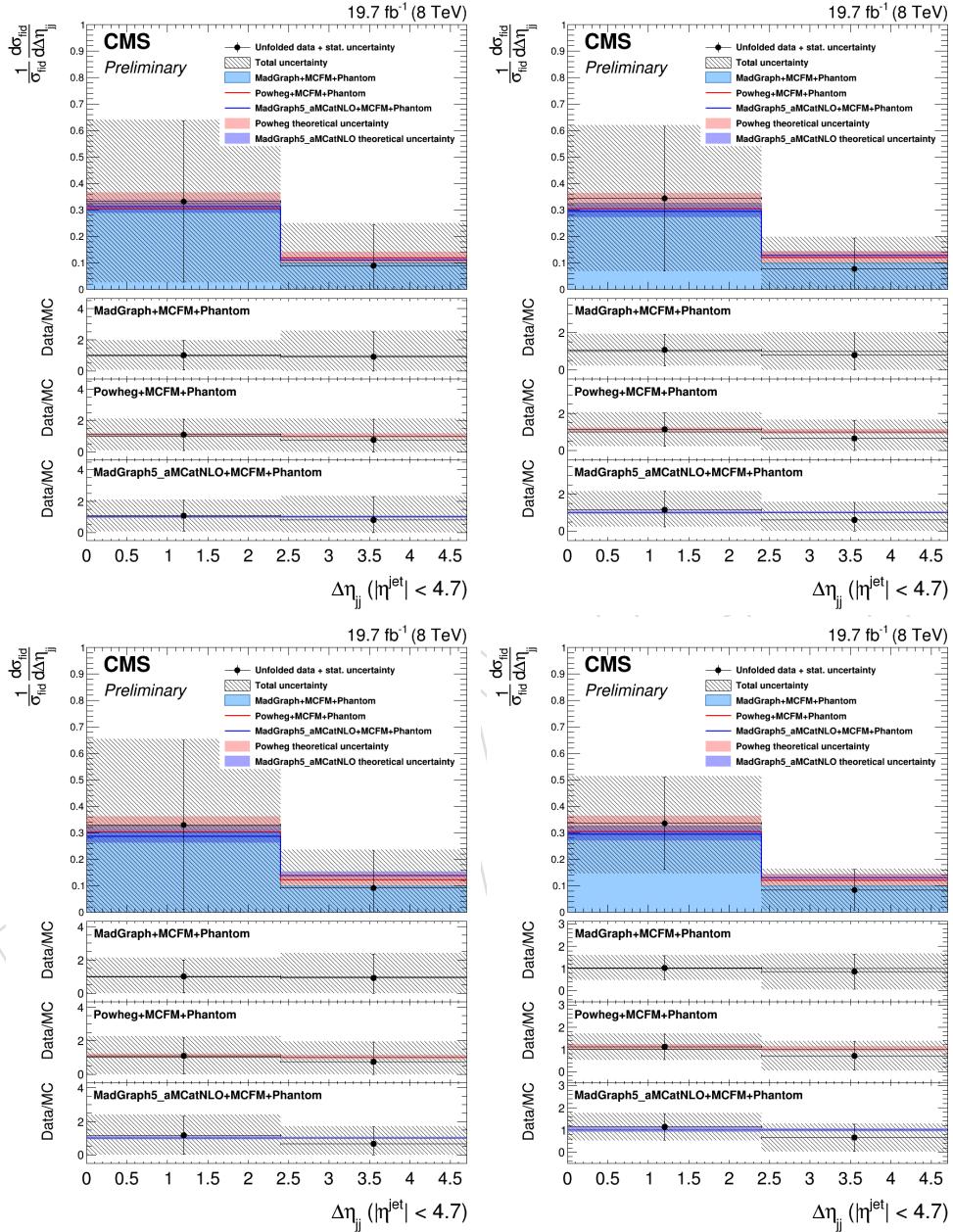


Figure 37: Normalized differential cross-sections as a function of the pseudorapidity interval between the most energetic jets in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

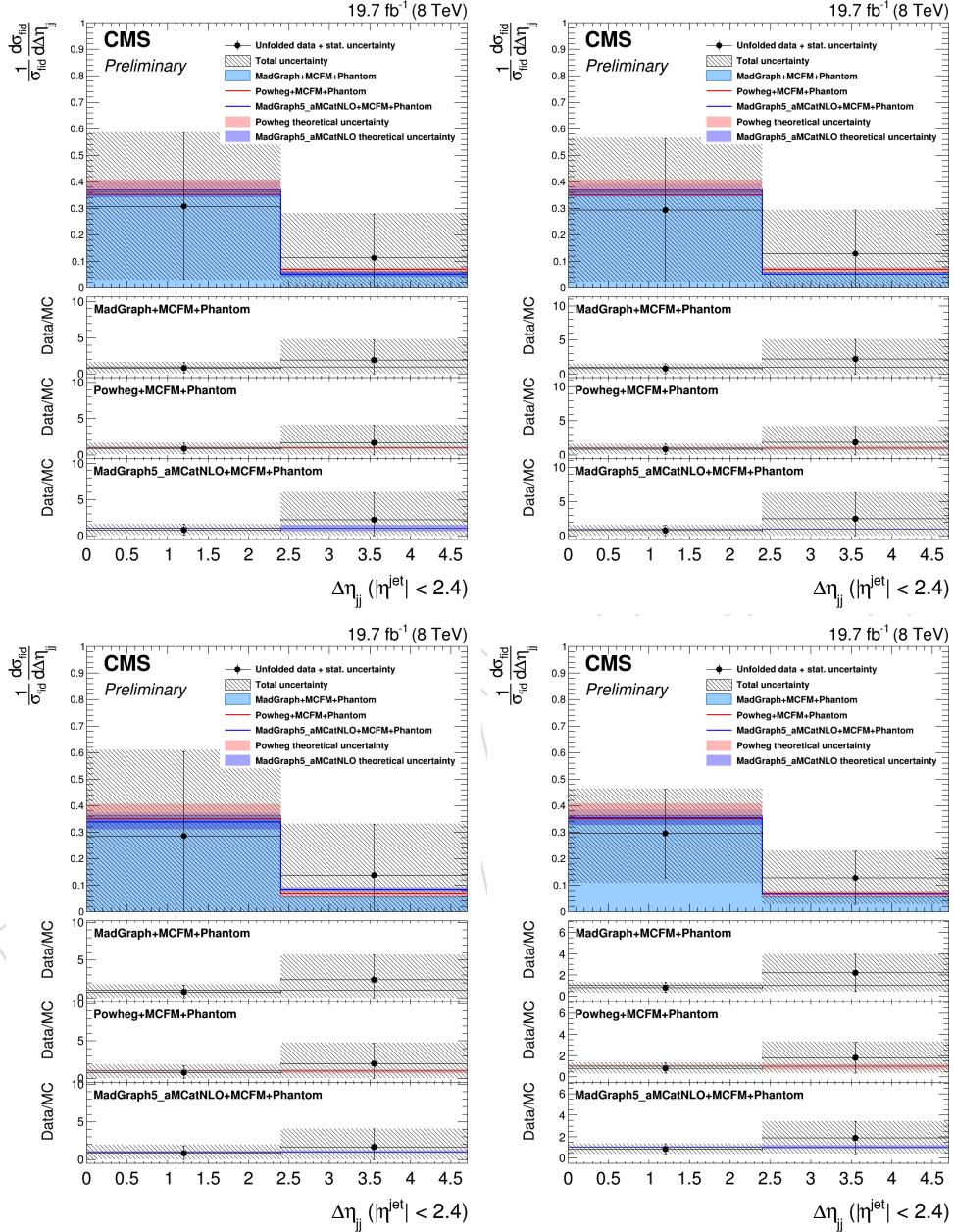


Figure 38: Normalized differential cross-sections as a function of the pseudorapidity interval between the most energetic central jets in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the `MadGraph`, `Powheg` and `MadGraph5_aMCatNLO` sets of samples.

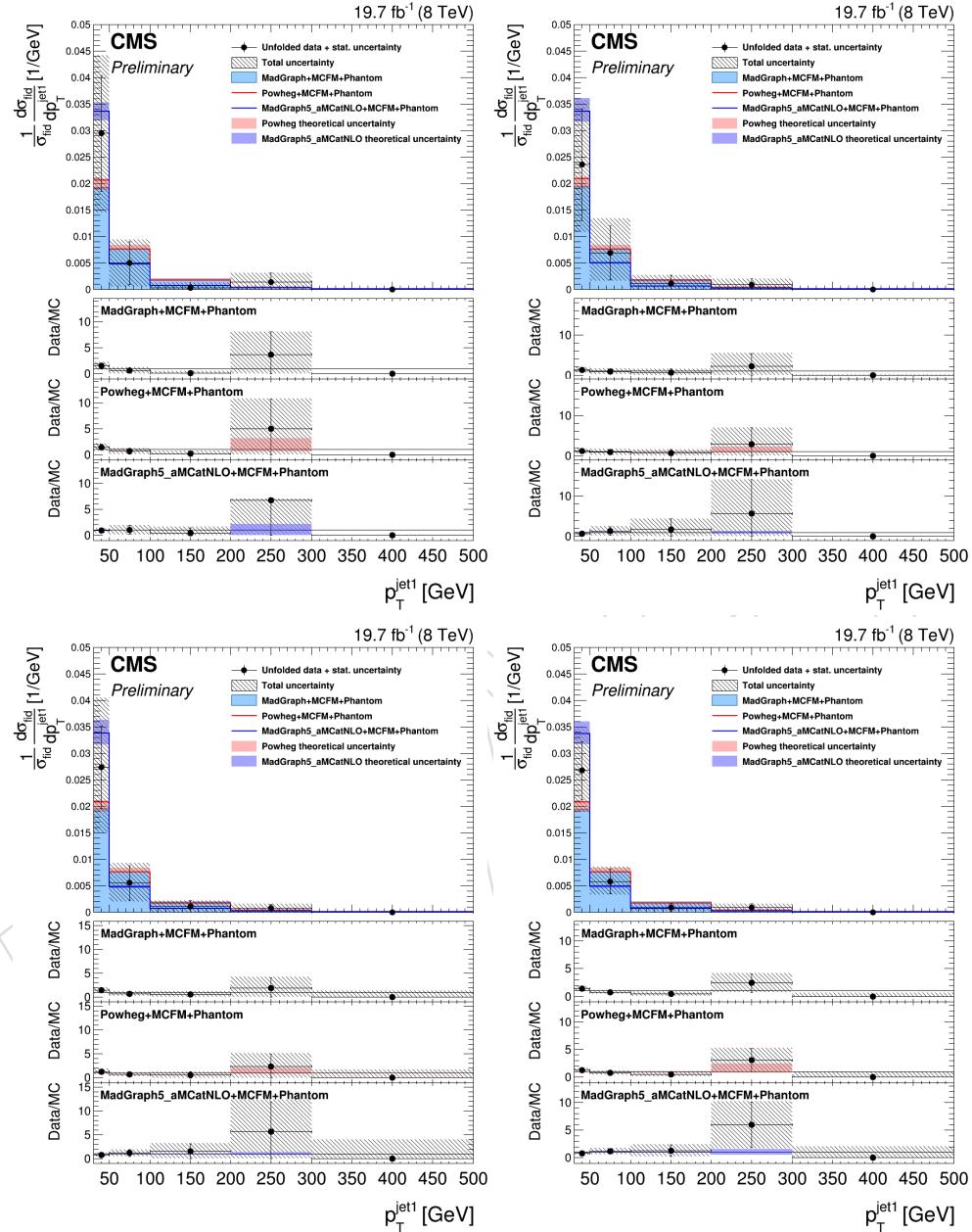


Figure 39: Normalized differential cross-sections as a function of the leading jet transverse momentum in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

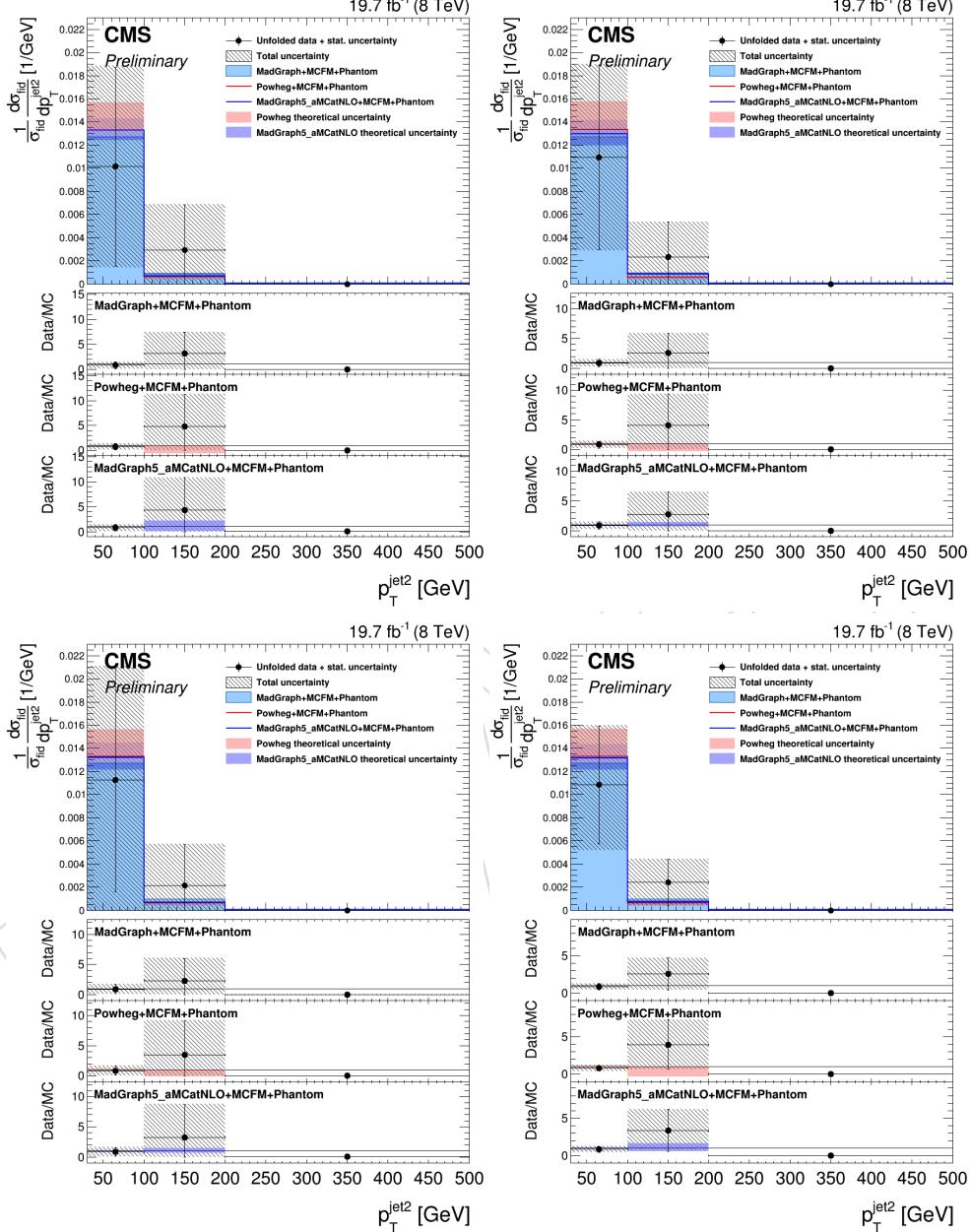


Figure 40: Normalized differential cross-sections as a function of the sub-leading jet transverse momentum in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

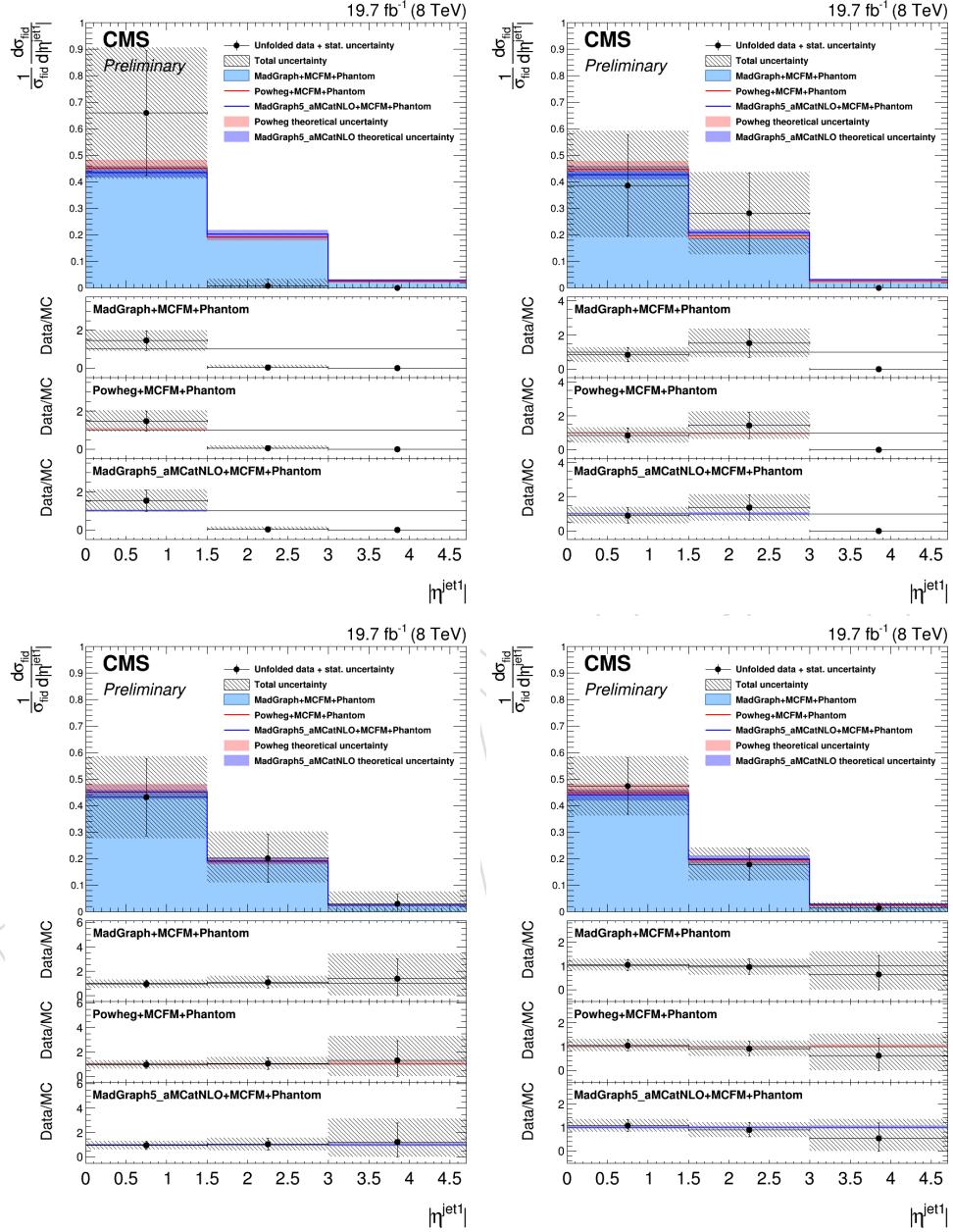


Figure 41: Normalized differential cross-sections as a function of the leading jet pseudorapidity in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

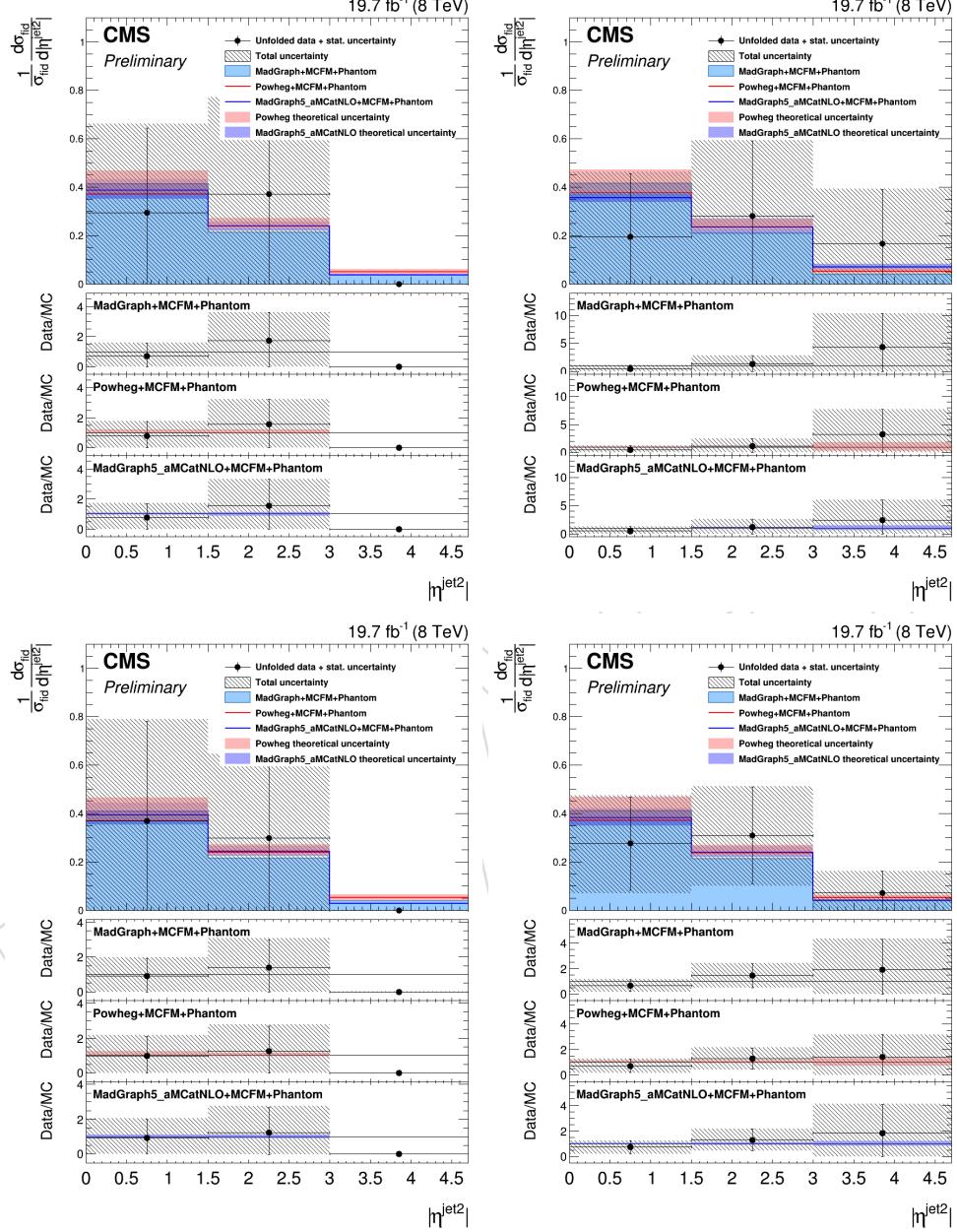


Figure 42: Normalized differential cross-sections as a function of the sub-leading jet pseudorapidity in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph, Powheg and MadGraph5_aMCatNLO sets of samples.

Table 12: The total ZZ production cross-section as measured in each decay channel and for the combination of all channels in the fiducial region $60 < m_{Z_1}, m_{Z_2} < 120$ GeV classified with respect to the number of central jets.

Process	Number of jets ($ \eta^{jet} < 2.4$)	Total cross-section [pb]
$pp \rightarrow ZZ(4\mu)$	0	6.27 ± 0.82 (stat.) ± 0.42 (syst.)
$pp \rightarrow ZZ(4e)$	0	6.05 ± 0.97 (stat.) ± 0.52 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	0	7.15 ± 0.67 (stat.) ± 0.58 (syst.)
$pp \rightarrow ZZ(4\ell)$	0	6.58 ± 0.46 (stat.) ± 0.42 (syst.)
$pp \rightarrow ZZ(4\mu)$	1	0.98 ± 0.36 (stat.) ± 0.15 (syst.)
$pp \rightarrow ZZ(4e)$	1	0.92 ± 0.44 (stat.) ± 0.17 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	1	1.13 ± 0.30 (stat.) ± 0.23 (syst.)
$pp \rightarrow ZZ(4\ell)$	1	1.02 ± 0.20 (stat.) ± 0.11 (syst.)
$pp \rightarrow ZZ(4\mu)$	2	0.10 ± 0.13 (stat.) ± 0.08 (syst.)
$pp \rightarrow ZZ(4e)$	2	0.49 ± 0.30 (stat.) ± 0.07 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	2	0.15 ± 0.12 (stat.) ± 0.10 (syst.)
$pp \rightarrow ZZ(4\ell)$	2	0.16 ± 0.08 (stat.) ± 0.04 (syst.)
$pp \rightarrow ZZ(4\mu)$	> 2	0.19 ± 0.14 (stat.) ± 0.06 (syst.)
$pp \rightarrow ZZ(4e)$	> 2	$(0.17 \pm 0.11$ (stat.) ± 0.65 (syst.)) \cdot 10^{-3}
$pp \rightarrow ZZ(2e2\mu)$	> 2	$(0.66 \pm 0.62$ (stat.) ± 0.96 (syst.)) \cdot 10^{-4}
$pp \rightarrow ZZ(4\ell)$	> 2	$(0.70 \pm 0.55$ (stat.) ± 0.93 (syst.)) \cdot 10^{-4}

Table 13: The total ZZ production cross-section as measured in each decay channel and for the combination of all channels in the tightt fiducial region classified with respect to the number of jets.

Process	Number of jets ($ \eta^{jet} < 4.7$)	Total cross-section [fb]
$pp \rightarrow ZZ(4\mu)$	0	3.78 ± 0.50 (stat.) ± 0.15 (syst.)
$pp \rightarrow ZZ(4e)$	0	3.73 ± 0.61 (stat.) ± 0.25 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	0	8.81 ± 0.84 (stat.) ± 0.55 (syst.)
$pp \rightarrow ZZ(4\ell)$	0	16.3 ± 1.2 (stat.) ± 0.8 (syst.)
$pp \rightarrow ZZ(4\mu)$	1	0.70 ± 0.26 (stat.) ± 0.08 (syst.)
$pp \rightarrow ZZ(4e)$	1	0.73 ± 0.34 (stat.) ± 0.10 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	1	1.75 ± 0.45 (stat.) ± 0.35 (syst.)
$pp \rightarrow ZZ(4\ell)$	1	3.17 ± 0.62 (stat.) ± 0.39 (syst.)
$pp \rightarrow ZZ(4\mu)$	2	0.06 ± 0.08 (stat.) ± 0.04 (syst.)
$pp \rightarrow ZZ(4e)$	2	0.45 ± 0.25 (stat.) ± 0.05 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	2	0.24 ± 0.18 (stat.) ± 0.12 (syst.)
$pp \rightarrow ZZ(4\ell)$	2	0.75 ± 0.32 (stat.) ± 0.14 (syst.)
$pp \rightarrow ZZ(4\mu)$	> 2	0.13 ± 0.10 (stat.) ± 0.05 (syst.)
$pp \rightarrow ZZ(4e)$	> 2	$(2.4 \pm 1.5$ (stat.) ± 3.3 (syst.)) \cdot 10^{-4}
$pp \rightarrow ZZ(2e2\mu)$	> 2	$(1.3 \pm 1.2$ (stat.) ± 1.9 (syst.)) \cdot 10^{-4}
$pp \rightarrow ZZ(4\ell)$	> 2	0.14 ± 0.10 (stat.) ± 0.5 (syst.)

Table 14: The total ZZ production cross-section as measured in each decay channel and for the combination of all channels in the `tightr` fiducial region classified with respect to the number of central jets.

Process	Number of jets ($ \eta^{jet} < 2.4$)	Total cross-section [fb]
$pp \rightarrow ZZ(4\mu)$	0	3.79 ± 0.49 (stat.) ± 0.17 (syst.)
$pp \rightarrow ZZ(4e)$	0	3.88 ± 0.62 (stat.) ± 0.27 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	0	8.98 ± 0.84 (stat.) ± 0.55 (syst.)
$pp \rightarrow ZZ(4\ell)$	0	16.6 ± 1.2 (stat.) ± 0.8 (syst.)
$pp \rightarrow ZZ(4\mu)$	1	0.68 ± 0.25 (stat.) ± 0.10 (syst.)
$pp \rightarrow ZZ(4e)$	1	0.66 ± 0.32 (stat.) ± 0.12 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	1	1.60 ± 0.42 (stat.) ± 0.32 (syst.)
$pp \rightarrow ZZ(4\ell)$	1	2.94 ± 0.59 (stat.) ± 0.36 (syst.)
$pp \rightarrow ZZ(4\mu)$	2	0.07 ± 0.09 (stat.) ± 0.05 (syst.)
$pp \rightarrow ZZ(4e)$	2	0.37 ± 0.22 (stat.) ± 0.05 (syst.)
$pp \rightarrow ZZ(2e2\mu)$	2	0.21 ± 0.17 (stat.) ± 0.15 (syst.)
$pp \rightarrow ZZ(4\ell)$	2	0.65 ± 0.30 (stat.) ± 0.17 (syst.)
$pp \rightarrow ZZ(4\mu)$	> 2	0.14 ± 0.10 (stat.) ± 0.04 (syst.)
$pp \rightarrow ZZ(4e)$	> 2	$(1.3 \pm 0.8$ (stat.) ± 4.9 (syst.) $) \cdot 10^{-4}$
$pp \rightarrow ZZ(2e2\mu)$	> 2	$(0.98 \pm 0.93$ (stat.) ± 1.48 (syst.) $) \cdot 10^{-4}$
$pp \rightarrow ZZ(4\ell)$	> 2	0.14 ± 0.10 (stat.) ± 0.4 (syst.)

278 7 Enhanced EWK Region of ZZ + 2jets Events

279 The scattering of two massive vector bosons (VBS) $VV \rightarrow VV$, with $V = W$ or Z , is the key
 280 process to probe the nature of the electroweak symmetry breaking (EWSB). In the absence of
 281 a Standard Model (SM) Higgs boson, the longitudinally polarized VBS amplitude increases
 282 as a function of the center-of-mass energy \sqrt{s} and violates unitarity at energies around 1 TeV.
 283 The recent discovery of a 125 GeV SM-like Higgs boson at LHC [31, 32] provides a plausible
 284 explanation for the mechanism that unitarizes this process. However, many physics scenarios
 285 predict enhancements in the VBS amplitude either from additional resonances, or if the ob-
 286 served SM-like Higgs boson is only partially responsible for the EWSB.

287 At hadron collider VBS can be represented by an interaction of gauge bosons radiated from
 288 initial state quarks yielding a final state with two bosons and two jets, in a purely electroweak
 289 process. Two classes of processes can generate a $VV + 2jet$ final state: the first class, that in-
 290 cludes also VBS processes, involves only electroweak interactions of order α_{EW}^4 (*electroweak pro-*
 291 *duction*), while the second class involves both strong and electroweak processes of order $\alpha_S^2 \alpha_{EW}^2$ (*strong pro-*
 292 *duction*). A fiducial region is defined in order to enhance the purity of electroweak
 293 $ZZ + 2jets$ and remove most of the strong events, which are considered as background in this
 294 analysis. This region is a subset of the one we used to measure the inclusive cross section of
 295 ZZ processes (see section 6), in which two Z bosons have a mass between 60 and 120 GeV. In
 296 addition to that inclusive region, it requires: at least two jets with an invariant mass (m_{jj}) larger
 297 than 300 GeV and separated in pseudorapidity by $\Delta\eta_{jj} > 2.4$, the leading(sub-leading) jet with
 298 $p_T > 100(70)$ GeV and the magnitude of the missing transverse energy $E_T < 60$ GeV. Only one
 299 event passing all selection requirements is observed in the data, compared to a SM expectation
 300 of 0.14 signal events and 0.47 background events. Figure 43 (left) shows the expected and ob-
 301 served $m_{4\ell}$ distribution after all fiducial region selection criteria are applied, except $\Delta\eta > 2.4$.
 302 In this region 4 events are observed and their kinematics is reported in Table 15. Figure 43
 303 (right) shows the m_{jj} distribution after the whole selection. All three dilepton channels are
 summed in both plots.

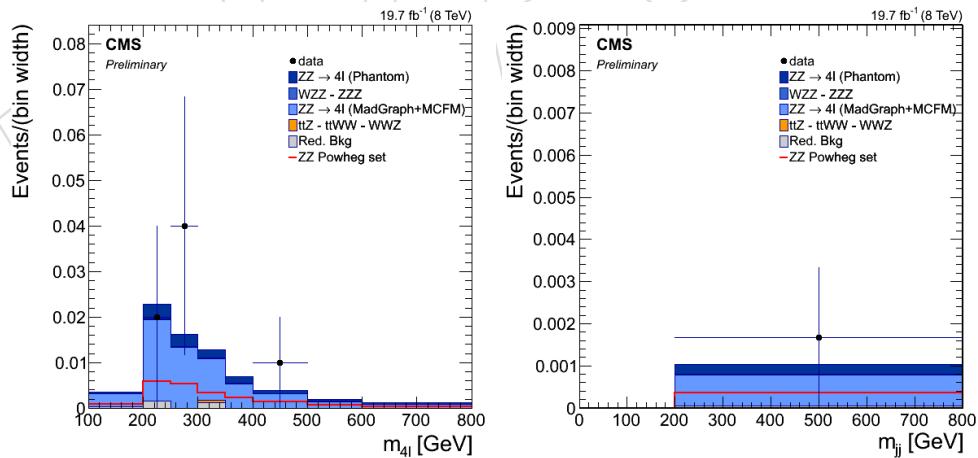


Figure 43: The $m_{4\ell}$ distribution (left) for events passing the region selections except for the $\Delta\eta_{jj} > 2.4$ selection. The m_{jj} distribution (right) for events passing all the region selections. Points represent the data, the stacked filled histogram represents the predictions for ZZ signal (from Phantom) and background contributions using MadGraph samples to describe $q\bar{q}(qg) \rightarrow ZZ \rightarrow 4\ell$ processes (while for the stacked histogram outlined in red the Powheg simulation is used).

Table 15: Kinematic values of events observed in the fiducial region without requiring $\Delta\eta_{jj} > 2.4$.

Event	$m_{4\ell}$	m_{jj}	$\Delta\eta_{jj}$	p_T^{jet1}	p_T^{jet2}
1	216.6	346.6	2.64	100.3	96.47
2	277.7	495.2	1.20	289.5	164.2
3	287.2	391.5	0.97	207.5	150.4
4	442.5	421.9	1.59	252.3	102.9

305 8 Conclusions

- 306 The measurement of the ZZ production total and differential cross-sections in pp collisions
 307 at $\sqrt{s} = 8$ TeV has been presented in the leptonic decay channels $pp \rightarrow ZZ \rightarrow \ell\ell\ell'\ell'$ with
 308 $\ell, \ell' = e, \mu$. The full 2012 dataset is used corresponding to an integrated luminosity of 19.7 fb^{-1} .
 309 Simple sequential sets of lepton reconstruction, identification and isolation cuts and a set of
 310 kinematic cuts are used following the selection introduced in the search for the Higgs boson
 311 decaying in 4 leptons (high mass selection). The main backgrounds are estimated using Monte
 312 Carlo samples and data-driven techniques and found to be very small for $60 < m_{Z_1} < 120$ and
 313 $60 < m_{Z_2} < 120$ GeV.
 314 The measured combined cross section is $\sigma_{pp \rightarrow ZZ \rightarrow 4\ell} = 20.50 \pm 1.26 \text{ (stat.)}^{+0.58}_{-0.64} \text{ (syst.)} \pm 0.53 \text{ (lumi.)}$,
 315 in agreement with the SM prediction of $20.21^{+3.3\%}_{-2.6\%} \text{ fb}$ from [33].
 316 Differential distributions for jet-related variables are presented and no significant discrepancy
 317 with respect to predictions is found.
 318 A tighter fiducial region is defined in order to select electroweak $ZZ + 2$ jets events. The obser-
 319 vation of one event in this region agrees with the SM expectation.

320 A Unfolding

321 In high energy physics the measurement of physical observables is usually distorted by several
 322 effects, such as finite resolution and limited acceptance of the detector. The observed spectrum
 323 of physics quantities is thus considered as a “noisy distortion” of the true one, i.e. the distri-
 324 bution one would observed under idealized conditions (ideal detector, no backgrounds...). An
 325 important task of the experimental method is therefore to extract the true distribution from the
 326 observed one, correcting for distortion and noise. This can be done by an unfolding procedure,
 327 that allows a direct comparison of the data obtained using different detectors with each other
 328 and with theoretical predictions.

329
 330 The measurement discussed here is based on the full data sample collected in proton-proton
 331 collisions during 2012 with the Compact Muon Solenoid Experiment (CMS) at the Large Hadron
 332 Collider (LHC), corresponding to the integrated luminosity of 19.7 fb^{-1} at a center of mass en-
 333 ergy of $\sqrt{s} = 8 \text{ TeV}$. The ZZ production cross section is measured differentially as a function
 334 of the invariant mass of the four-lepton system, the number of jets produced in the event, the
 335 invariant mass of the two most energetic jets (m_{jj}), the difference of pseudorapidity between
 336 them ($\Delta\eta_{jj}$) and their transverse momentum and pseudorapidity. The measurements are per-
 337 formed in the leptonic decay modes $ZZ \rightarrow \ell\ell\ell'\ell'$ channel ($\ell, \ell' = e, \mu$).

338
 339 The unfolding procedure is based on the so-called “response matrix”, derived using simulated
 340 Monte Carlo, which is a mapping between the true value (generated) of the observable and the
 341 reconstructed (measured) value, distorted by detector effects. Because of inevitable statistical
 342 uncertainties in the measured distribution, the exact solution that one would obtain inverting
 343 the response matrix (if it exists) can lead to unacceptable results, wildly oscillating and use-
 344 less. In order to avoid this unstable behavior, a “regularization condition” can be imposed,
 345 requiring a smooth true distribution. The regularized unfolding methods investigated here,
 346 as implemented in the RooUnfold package [27], are the Singular Value Decomposition (SVD)
 347 method [28] and the iterative Bayesian unfolding [29].

348 A.1 Unfolding procedure

The distribution of a measured observable be stored in a vector N_{rec} of dimension n , where the i th coordinate of the vector contains the number of entries in the corresponding bin of the histogram [28, 30]. The measurement is affected by the finite experimental resolution and/or the limited acceptance of the detector, so that each event from the true distribution may find itself in a range of (not necessarily) adjacent bins, or nowhere at all.

We generate the distribution N_{gen} of dimension m according to some idea of the physical process under study, and perform the detector simulation. At this stage, every event in a measured bin i can be directly matched to the generated one j . A well defined system of linear equations is thus determined, describing the relations between the simulated true and measured distributions:

$$\sum_j A_{ij} N_{gen}^j = N_{rec}^i.$$

The A_{ij} matrix is the response matrix that maps the binned generated distribution onto the measured one: each element (i, j) is indeed related to the probability that the observable generated in the j th bin would be measured in the i th bin. The unfolding procedure applies the response matrix to the measured data distribution (in which background is subtracted), taking into account the measurement uncertainties due to statistical fluctuations in the finite measured

sample through the covariance matrix:

$$\sum_j A_{ij} N_{gen}^j = N_{sig}^i := N_{data}^i - N_{bkg}^i.$$

349 The linear system of equations can be solved using the exact inverse of the response matrix
 350 on measured histogram and obtaining a data distribution “at particle level”. However a direct
 351 inversion of the matrix usually leads to completely unacceptable wildly oscillating results. In
 352 order to overcome this problem, different regularized algorithms can be used, such as the Sin-
 353 gular Value Decomposition (SVD) and the iterative “Baysian” method, well described in [28]
 354 and [29].

355

356 The response matrices used in the analysis are obtained from Monte Carlo samples that con-
 357 tain signal-only events with pile-up reweighting, lepton and trigger efficiency applied. Two
 358 sets of samples are employed, the Madgraph + MCFM + Phantom (as reference) and Powheg +
 359 MCFM + Phantom (as check) sets (for the m_{ZZ} distribution the role of the two sets of samples
 360 is switched). Leptons are generated requiring $m_{\ell^+\ell^-} > 4$ GeV in all samples but MadGraph,
 361 in which $m_{\ell^+\ell^-} > 12$ GeV, and reconstructed following the same steps of [2, 5]. Jets are gen-
 362 erated with $p_T > 10$ GeV and reconstructed following the criteria recommended by Jet-MET
 363 group [20]. Both generated events not measured because of detection or selection inefficiency
 364 and reconstructed events not generated as signal are also taken into account in the unfolding
 365 procedure.

366

367 In the following, first the choice of the binning of the distributions is discussed, in order to
 368 control migration from one bin to others. Then the performance of the unfolding is validated
 369 through studies on Monte Carlo samples and applied on data distribution. Finally the different
 370 sources of systematic errors are investigated and uncertainties are evaluated.

371 A.2 Bin-to-bin migration

Because of the finite experimental resolution, events that are actually produced (generated) in one bin might be measured (reconstructed) in another bin, leading to migrations of events across bin boundaries. This bin-to-bin migration is studied in term of purity and stability in each bin. As in [30], the purity p_i is defined as the number of events that are generated and correctly reconstructed in a given bin i divided by the number of events that are reconstructed in bin i , but generated anywhere. On the other hand, the stability is defined as the number of events that are generated and correctly reconstructed in a given bin i divided by the number of events that are generated in bin i , but reconstructed anywhere:

$$p_i = \frac{N_{gen\&reco}^i}{N_{reco}^i}; \quad s_i = \frac{N_{gen\&reco}^i}{N_{gen}^i},$$

372 where *reco* refers to reconstructed events fulfilling the full selection requirements described
 373 above and *gen* refers to generated events satisfying the phase space requirements. Without mi-
 374 gration effects, purity and stability would be equal to one. The purity (stability) is sensitive to
 375 migrations into (out of) the bin. In order to keep bin-to-bin migrations acceptably small, the bin
 376 widths for each observable are optimized such that for each bin purity and stability are greater
 377 than about 70%. Figures 44, 45, 46, 47, 48 and 49 show the purity and stability for the final
 378 binning definition. The binning selection for each observable is define as:

379

- 380 • $m_{4\ell}$: $\{100, 200, 250, 300, 350, 400, 500, 600, \geq 800\}$
- 381 • N jets: $\{0, 1, 2, > 2\}$
- 382 • m_{jj} : $\{0, 200, \geq 800\}$
- 383 • $\Delta\eta_{jj}$: $\{0, 2.4, \geq 4.7\}$
- 384 • p_T^{jet1} : $\{30, 50, 100, 200, 300, \geq 500\}$
- 385 • p_T^{jet2} : $\{30, 100, 200, \geq 500\}$
- 386 • η^{jet1}, η^{jet2} : $\{0, 1.5, 3, 4.7\}$

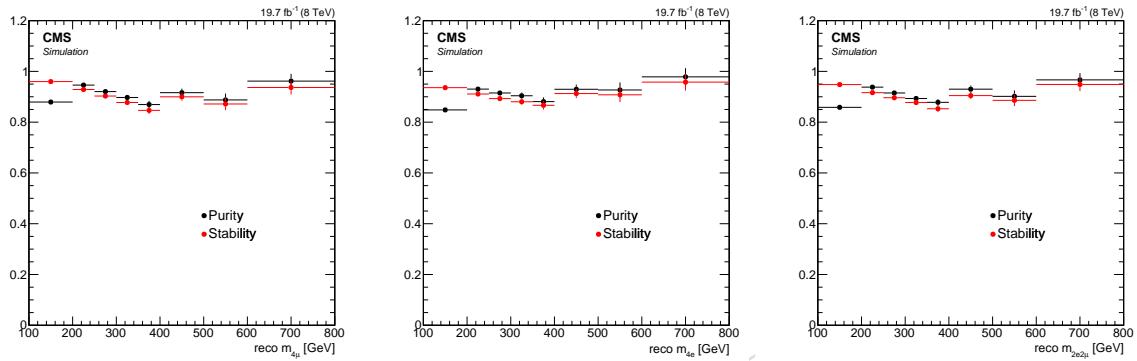


Figure 44: Purity and stability as a function of the 4-lepton invariant mass, for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states.

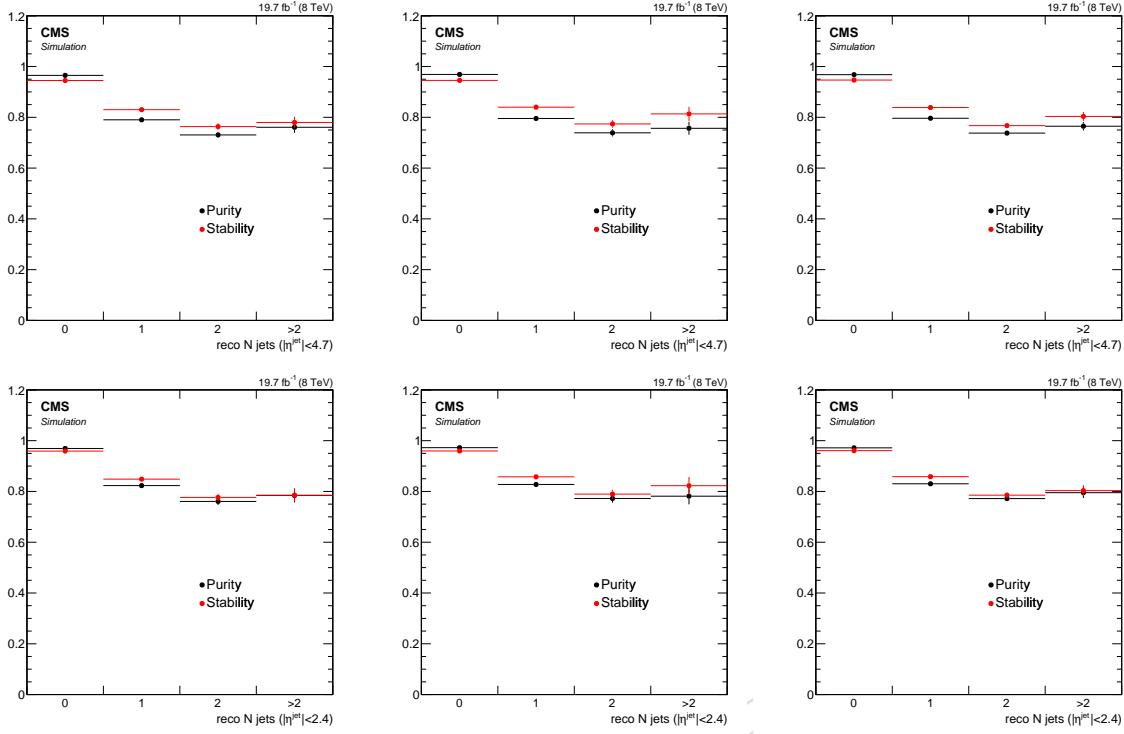


Figure 45: Purity and stability as a function of the number of jets (top) and central jets (bottom) in the event, for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states.

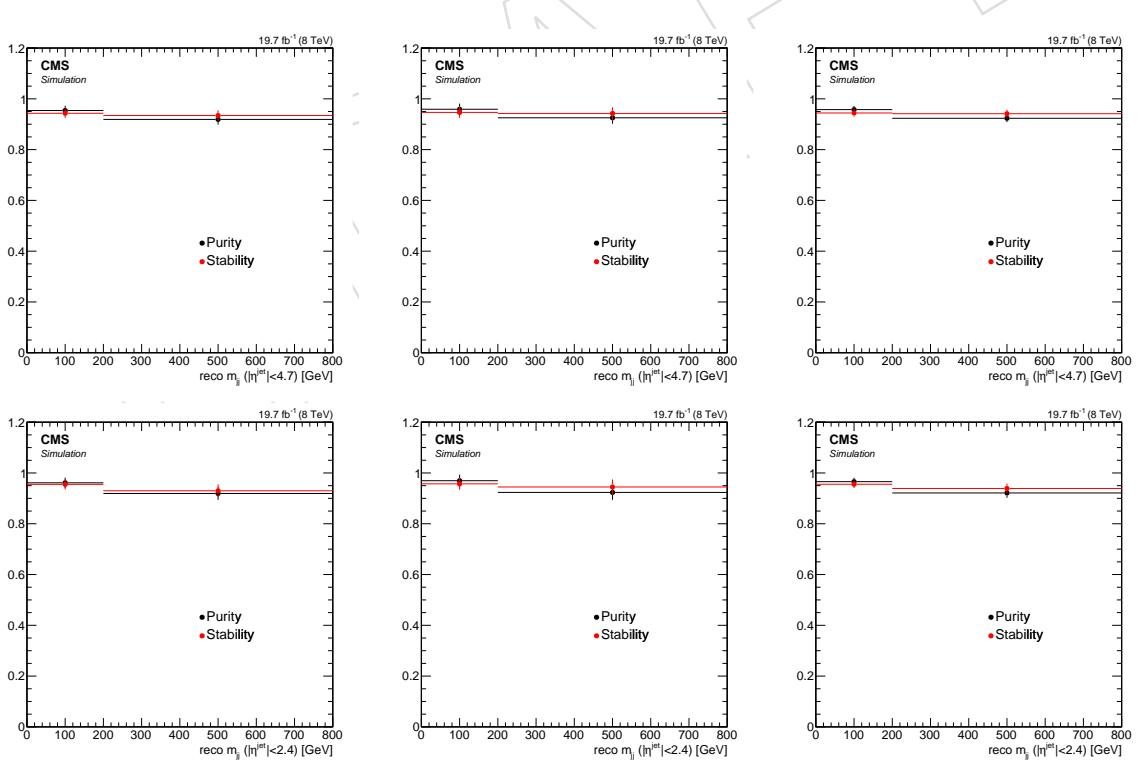


Figure 46: Purity and stability as a function of the invariant mass of the two most energetic jets (top) and central jets (bottom) in the event, for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states.

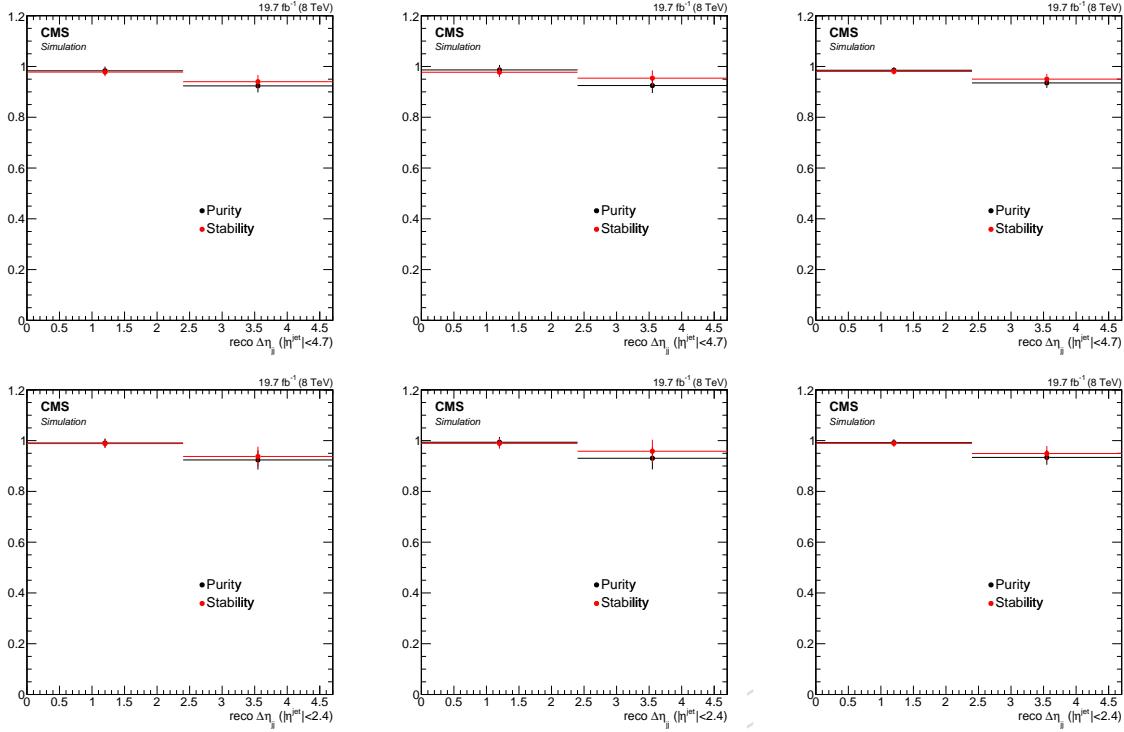


Figure 47: Purity and stability as a function of the $\Delta\eta$ between the two most energetic jets (top) and central jets (bottom) in the event, for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states.

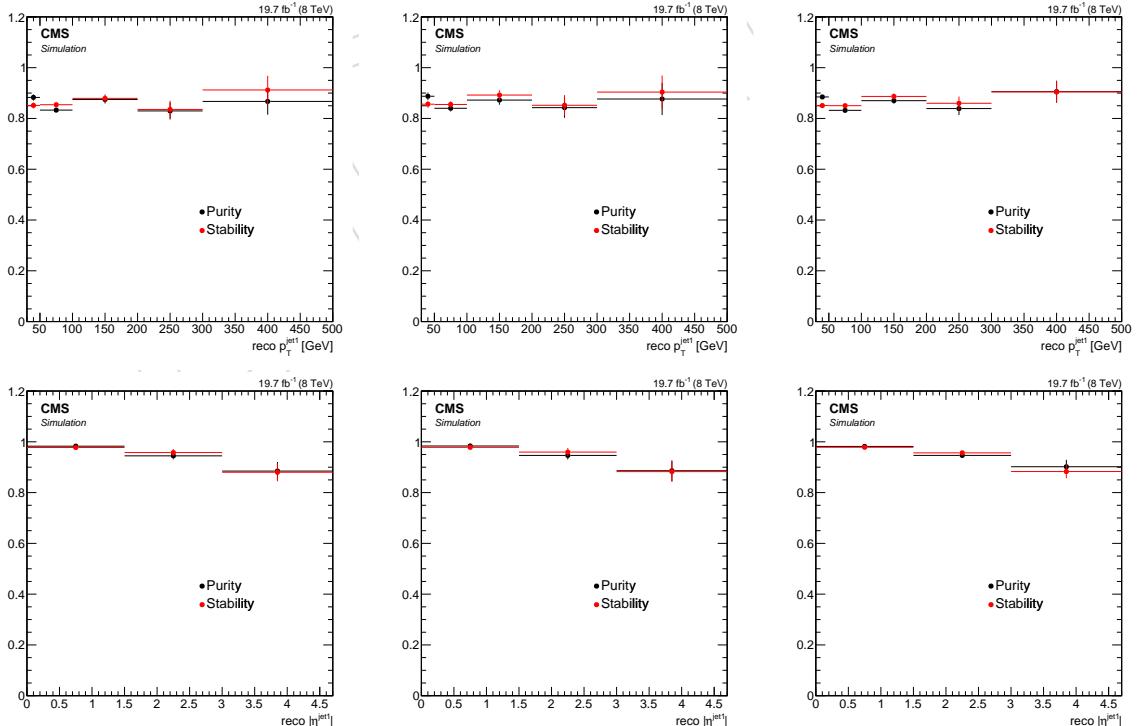


Figure 48: Purity and stability as a function of the p_T (top) and η (bottom) of the most energetic jet in the event, for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states.

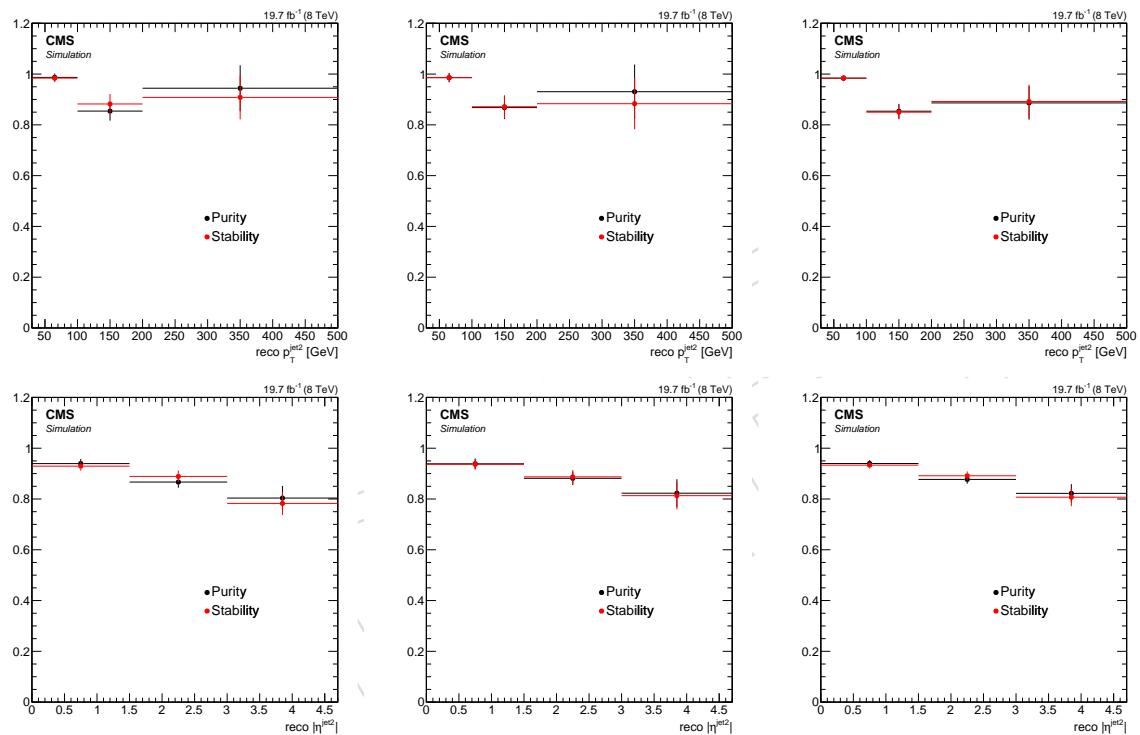
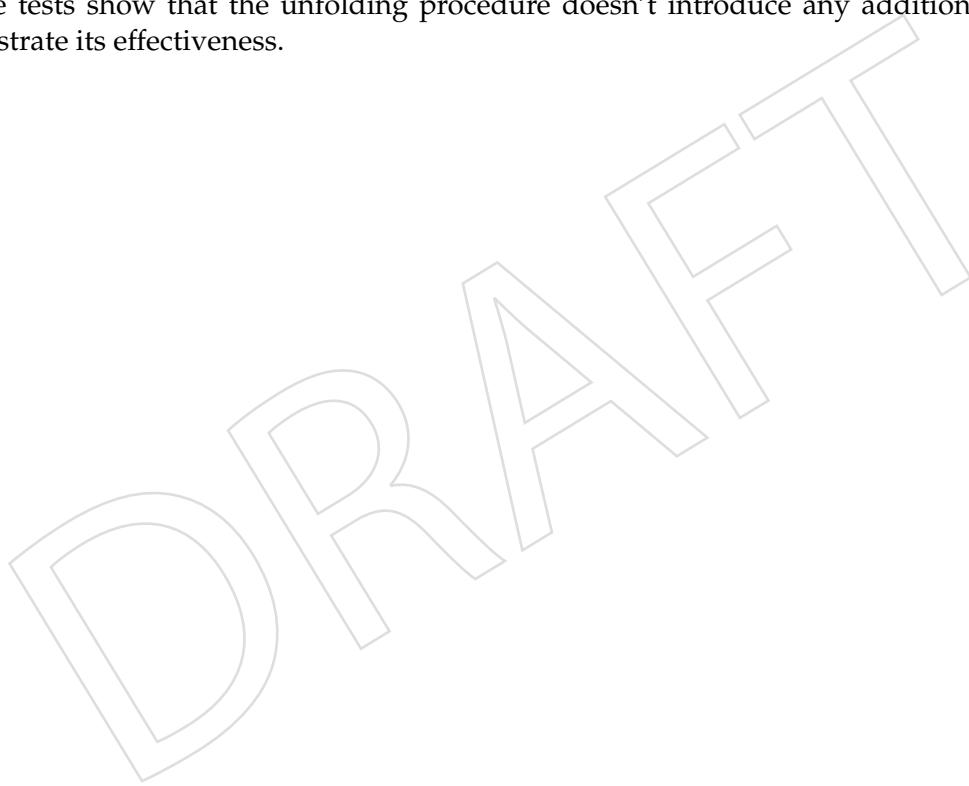


Figure 49: Purity and stability as a function of the p_T (top) and η (bottom) of the second most energetic jet in the event, for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states.

A.3 Unfolding validation test on Monte Carlo

Before looking at the data, it is recommended to test the unfolding procedure on MC events alone. As first step the consistency of the whole process is checked using the full MadGraph (or Powheg) set of samples, both for the distribution to be unfolded and the response matrix. If everything is correctly implemented the unfolded distribution and the generated one should be exactly the same. Moreover, in order to get meaningful results, the distribution that has to be unfolded must be statistically independent from the 2-dimensional response histogram. The MadGraph(Powheg) set is thus split into two samples: one of them is used to fill the response matrix, while the other one is used to build the distribution to unfold. Finally, in order to compare the effect of employing different signal samples and to be sure the procedure is independent of the choice of a particular MC, the response matrix from the MadGraph set is applied on the distribution obtained using the Powheg set and vice versa. Tests are performed both in the standard and tight fiducial regions and results for the latter case are reported from Figure 50 to Figure 71 for the considered distributions, in the three different final states. In each plot the ratio of unfolded over generated distribution is shown and, as expected, it is unity. Closure tests show that the unfolding procedure doesn't introduce any additional bias and demonstrate its effectiveness.



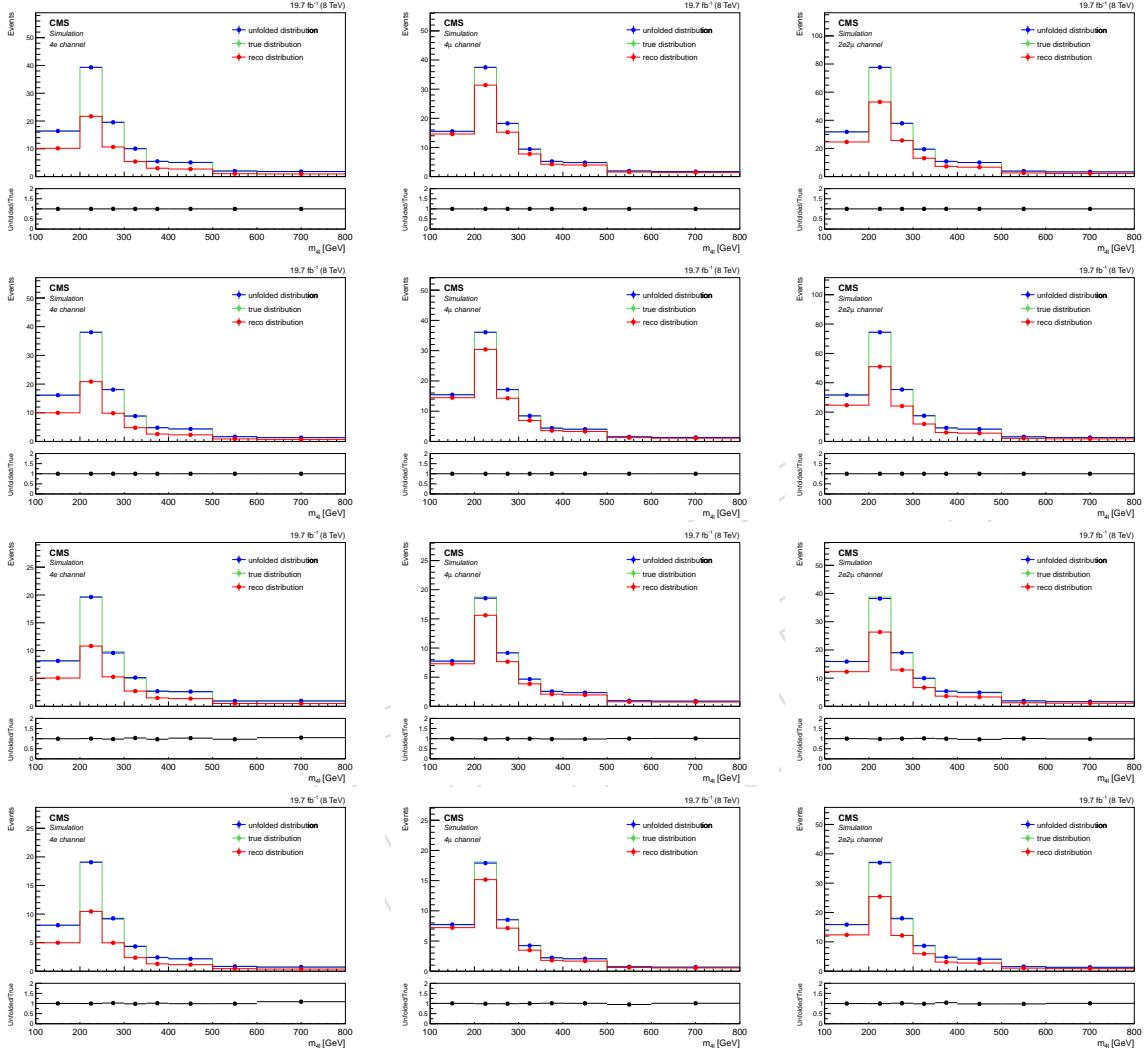


Figure 50: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of the 4-lepton mass system for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

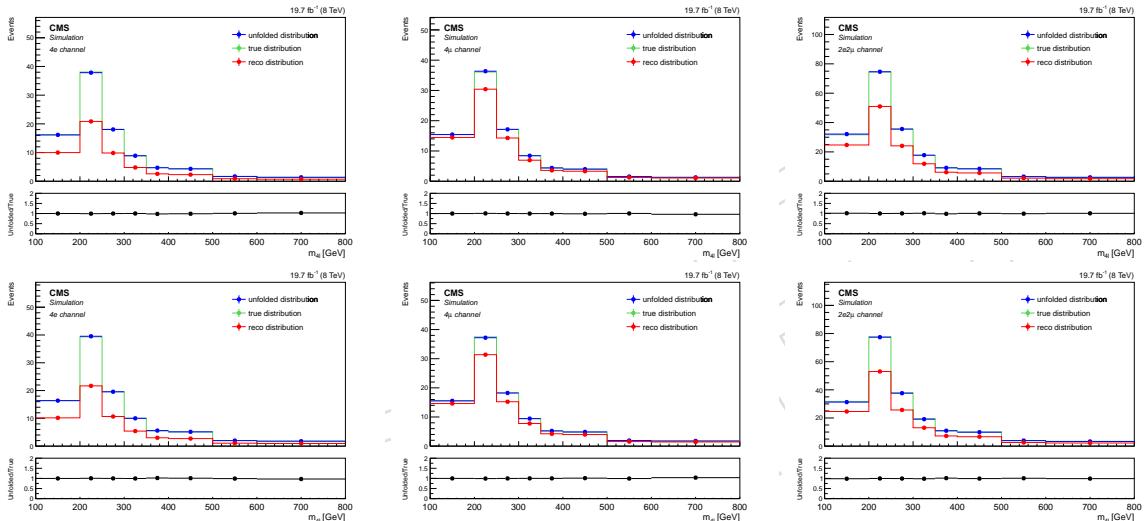


Figure 51: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of the 4-lepton mass system for the 4e (left), 4 μ (center) and 2e2 μ (right) final states.

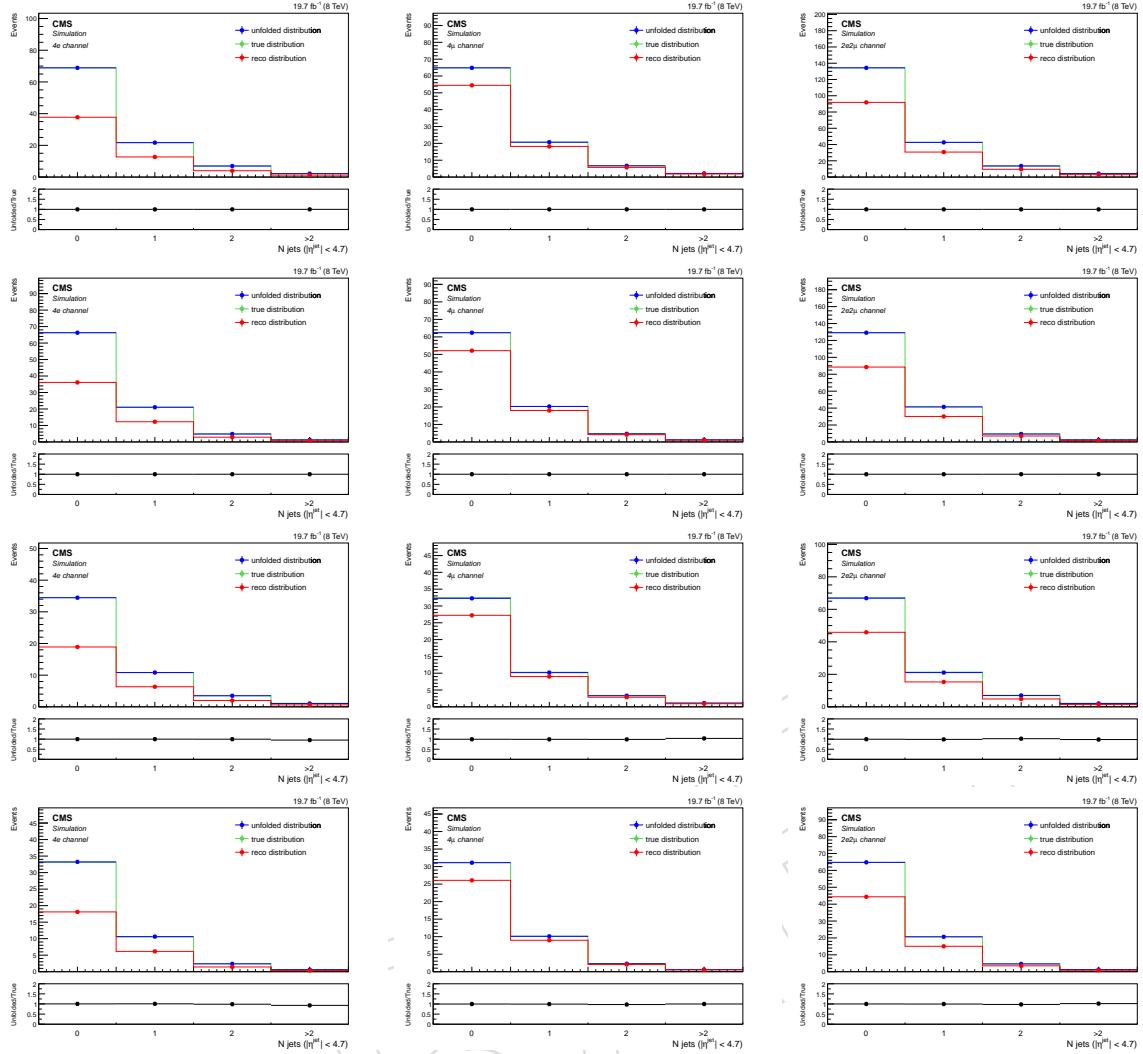


Figure 52: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of the number of jets for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

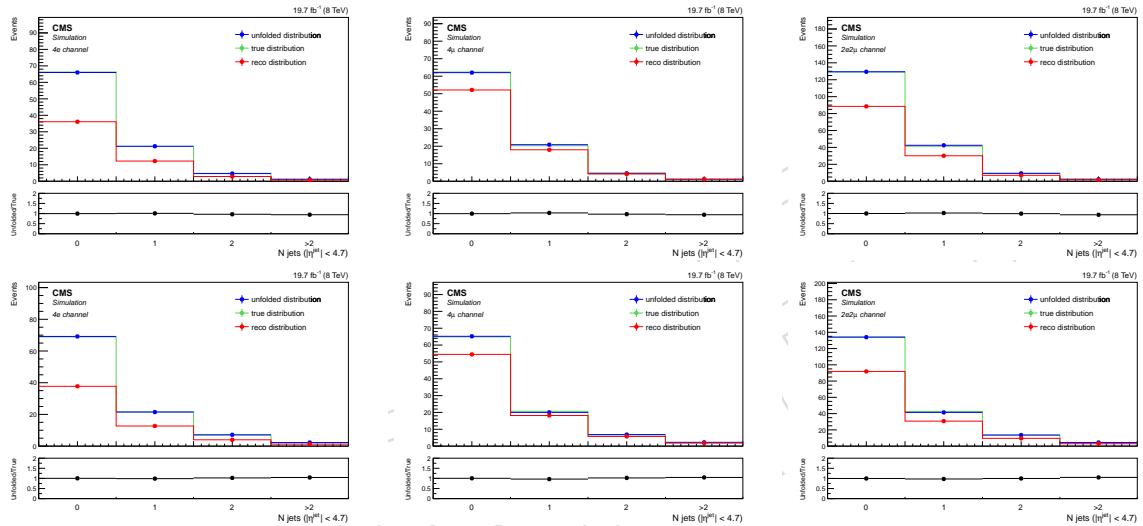


Figure 53: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of the number of jets for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

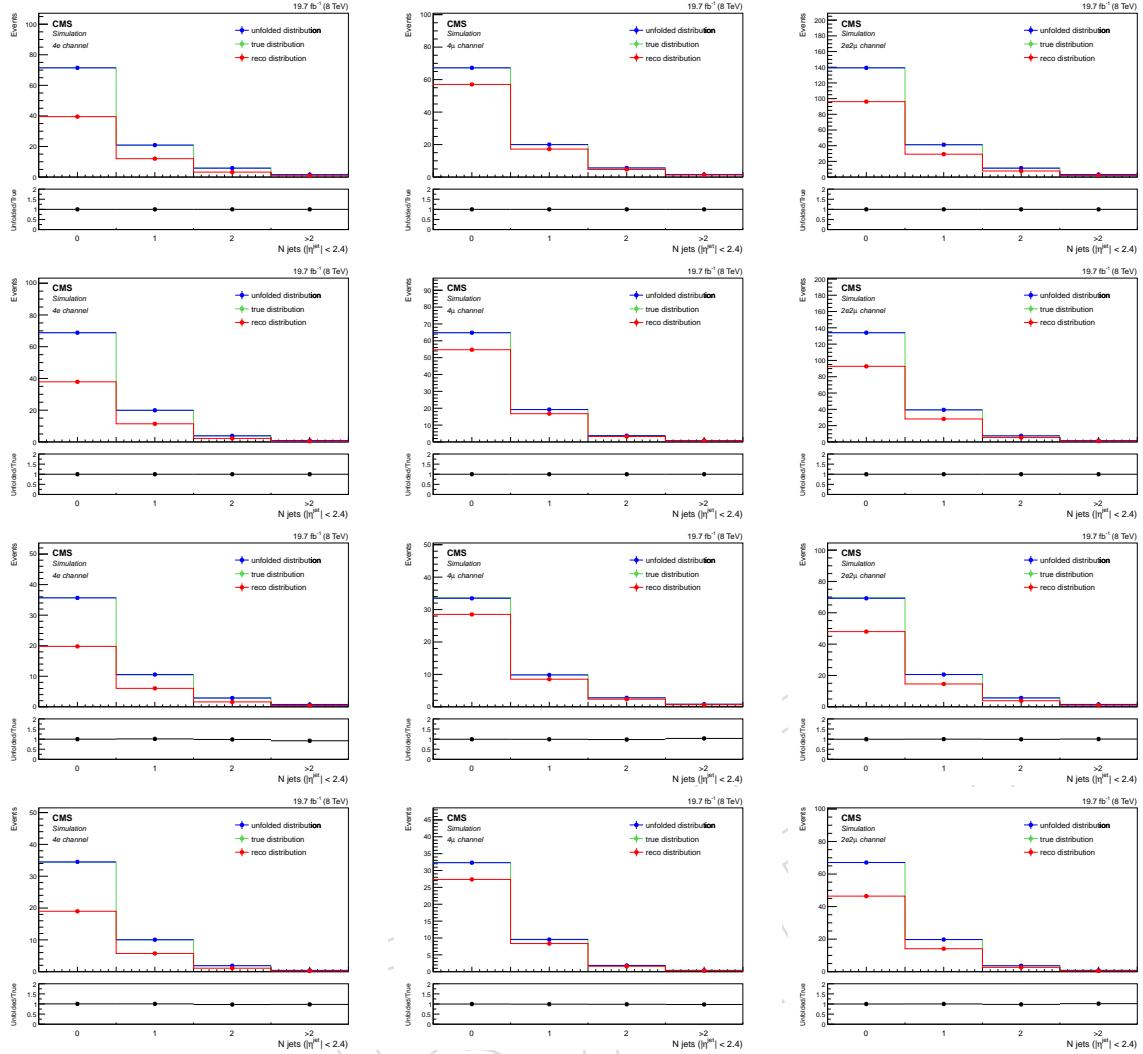


Figure 54: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of the number of central jets for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

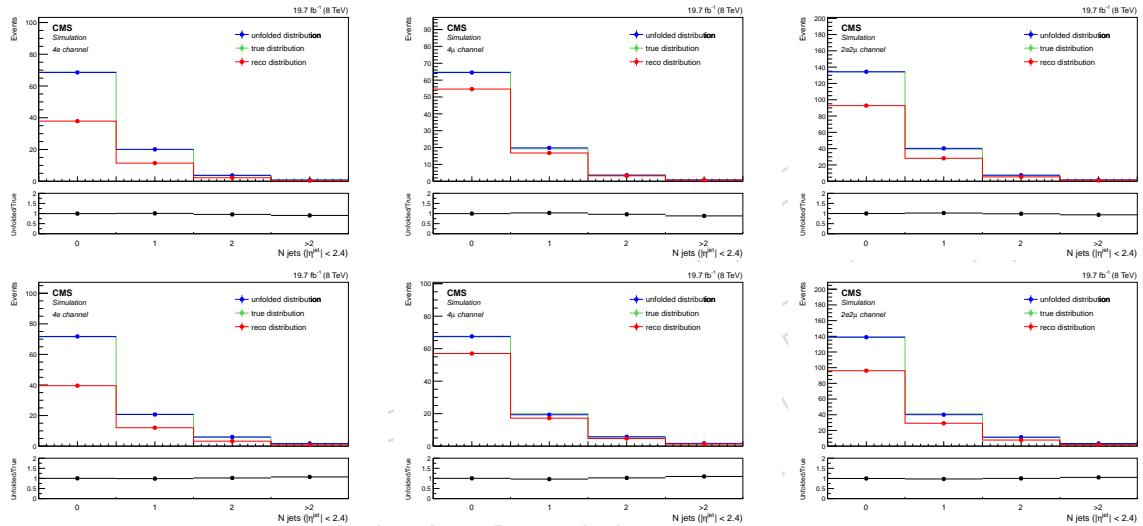


Figure 55: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of the number of central jets for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

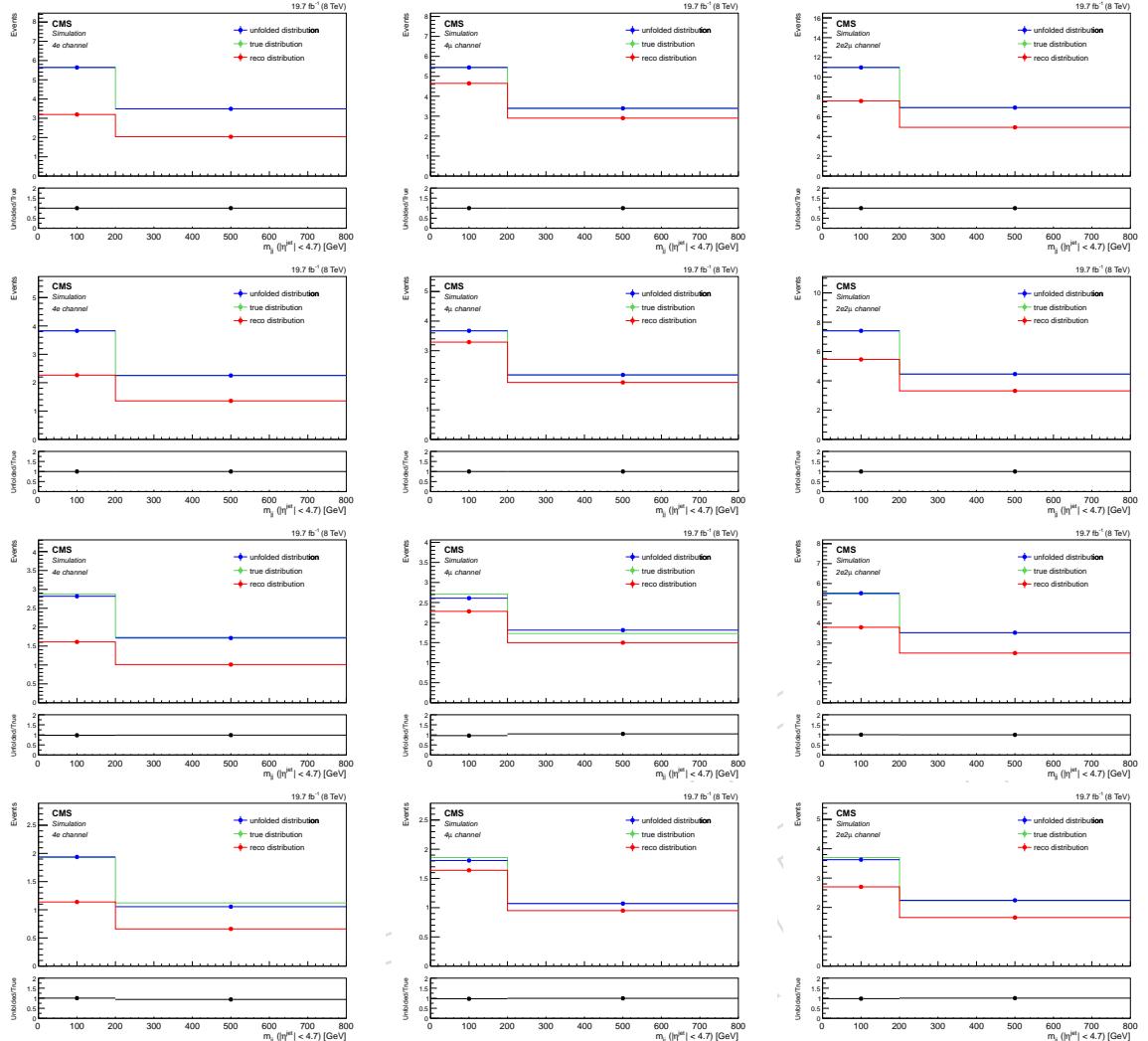


Figure 56: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of m_{jj} for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

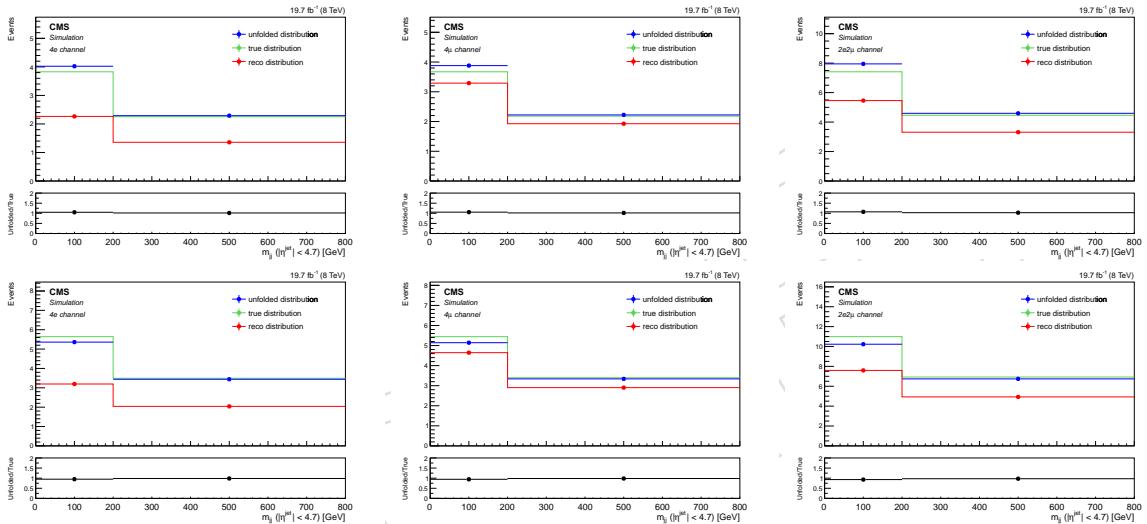


Figure 57: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of m_{jj} for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

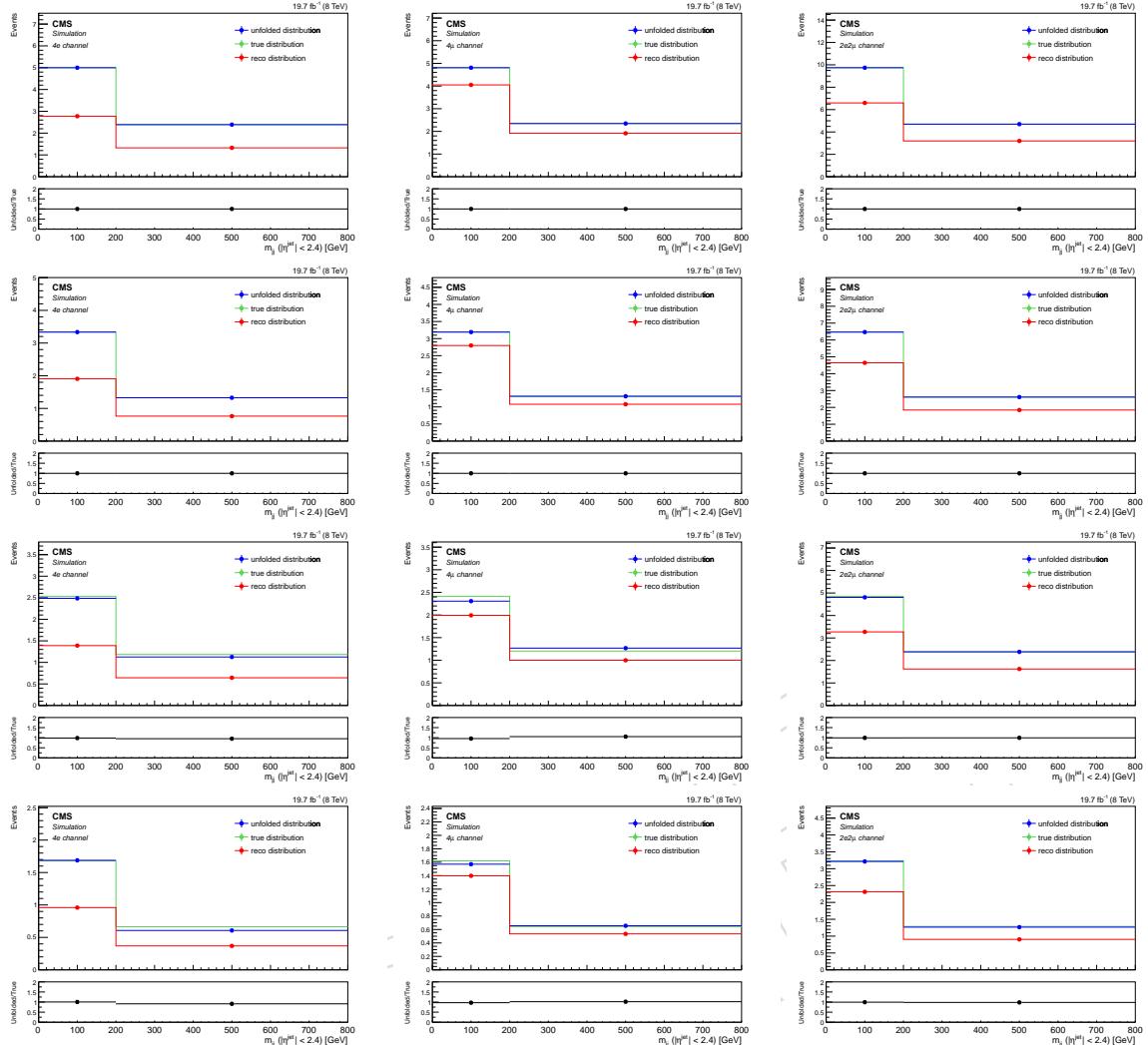


Figure 58: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of m_{jj} (with $|\eta^{jet}| < 2.4$) for the 4e (left), 4 μ (center) and 2e2 μ (right) final states.

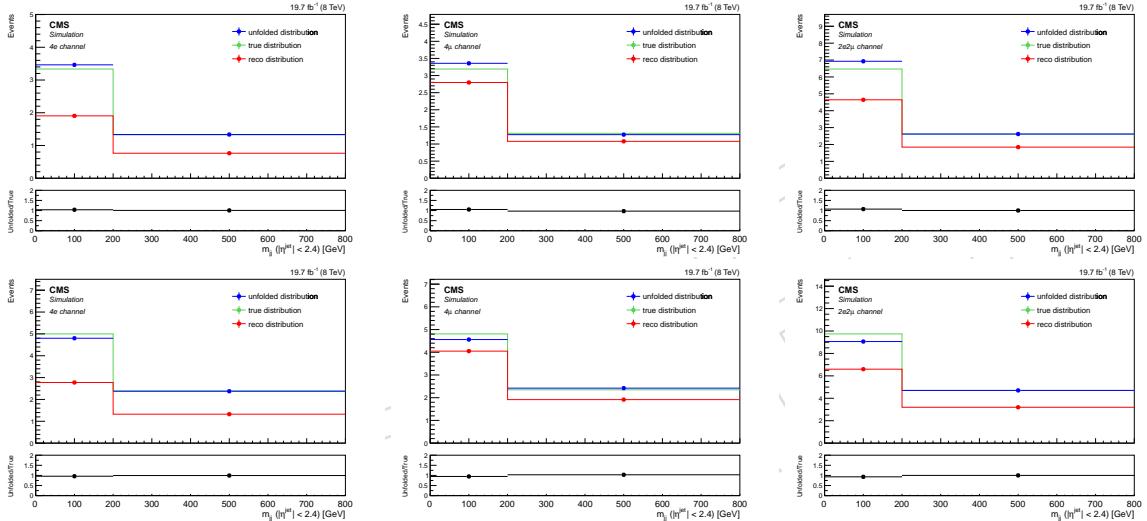


Figure 59: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of m_{jj} (with $|\eta^{jet}| < 2.4$) for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

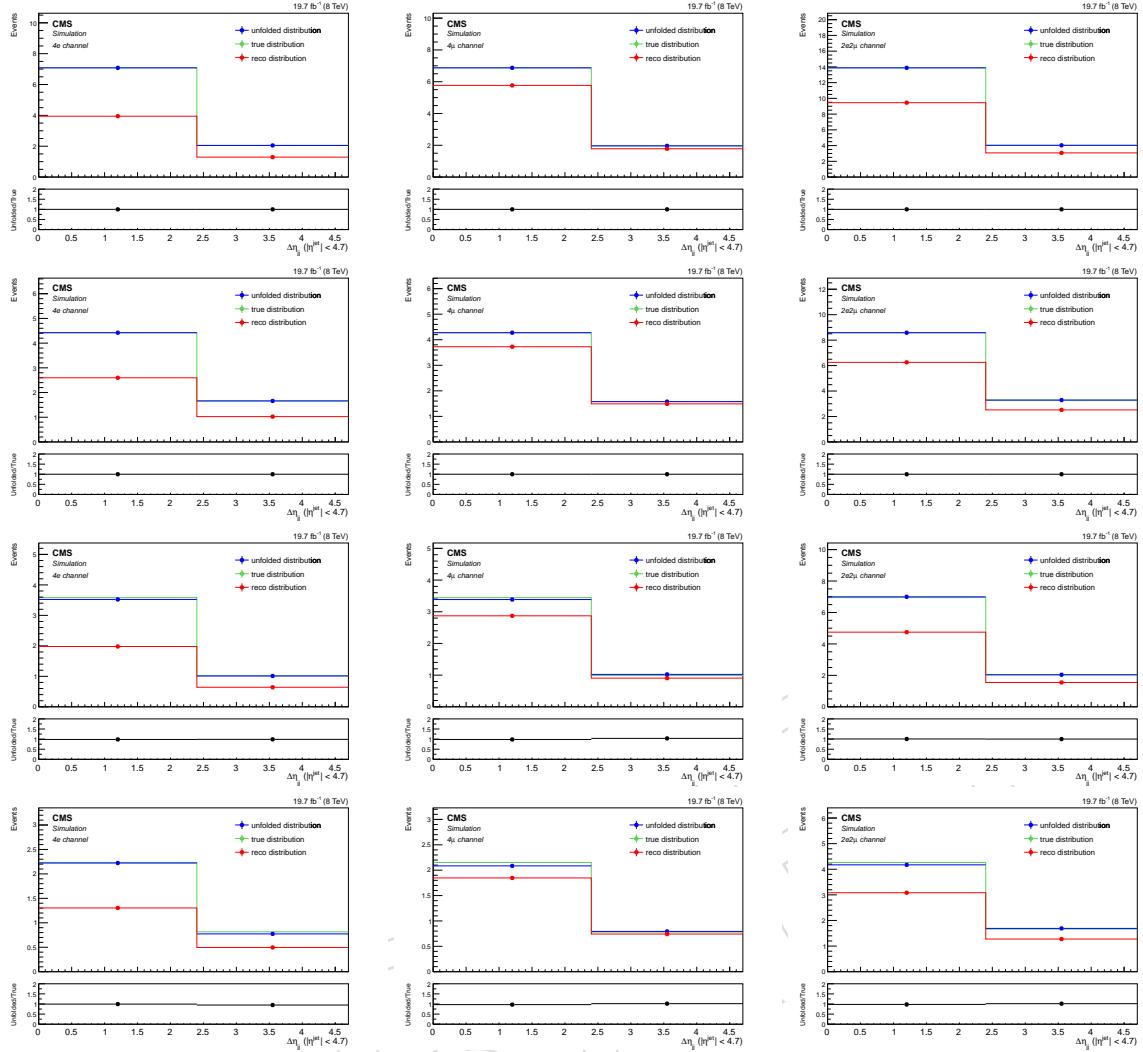


Figure 60: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of $\Delta\eta_{jj}$ for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

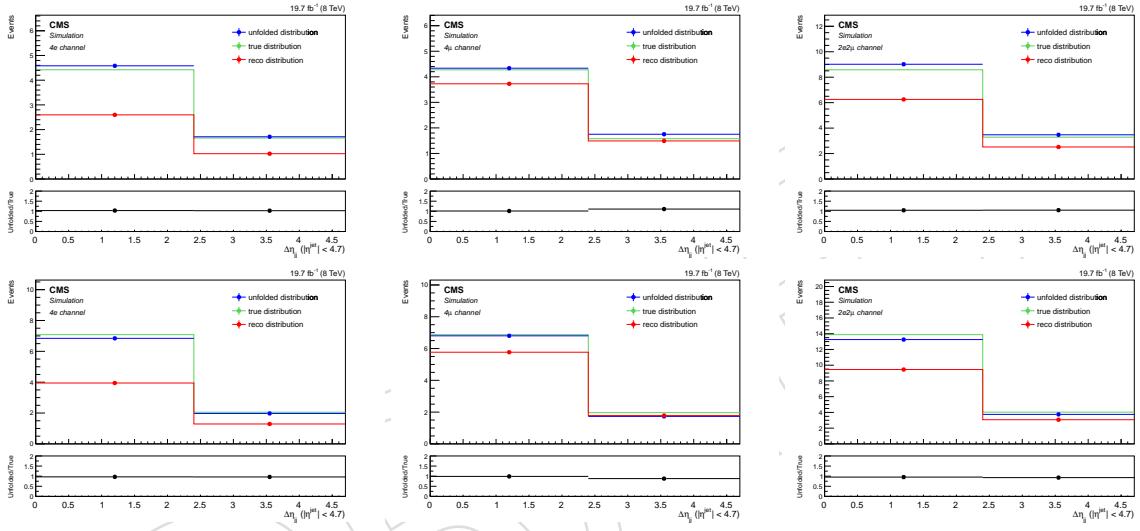


Figure 61: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of $\Delta\eta_{jj}$ for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

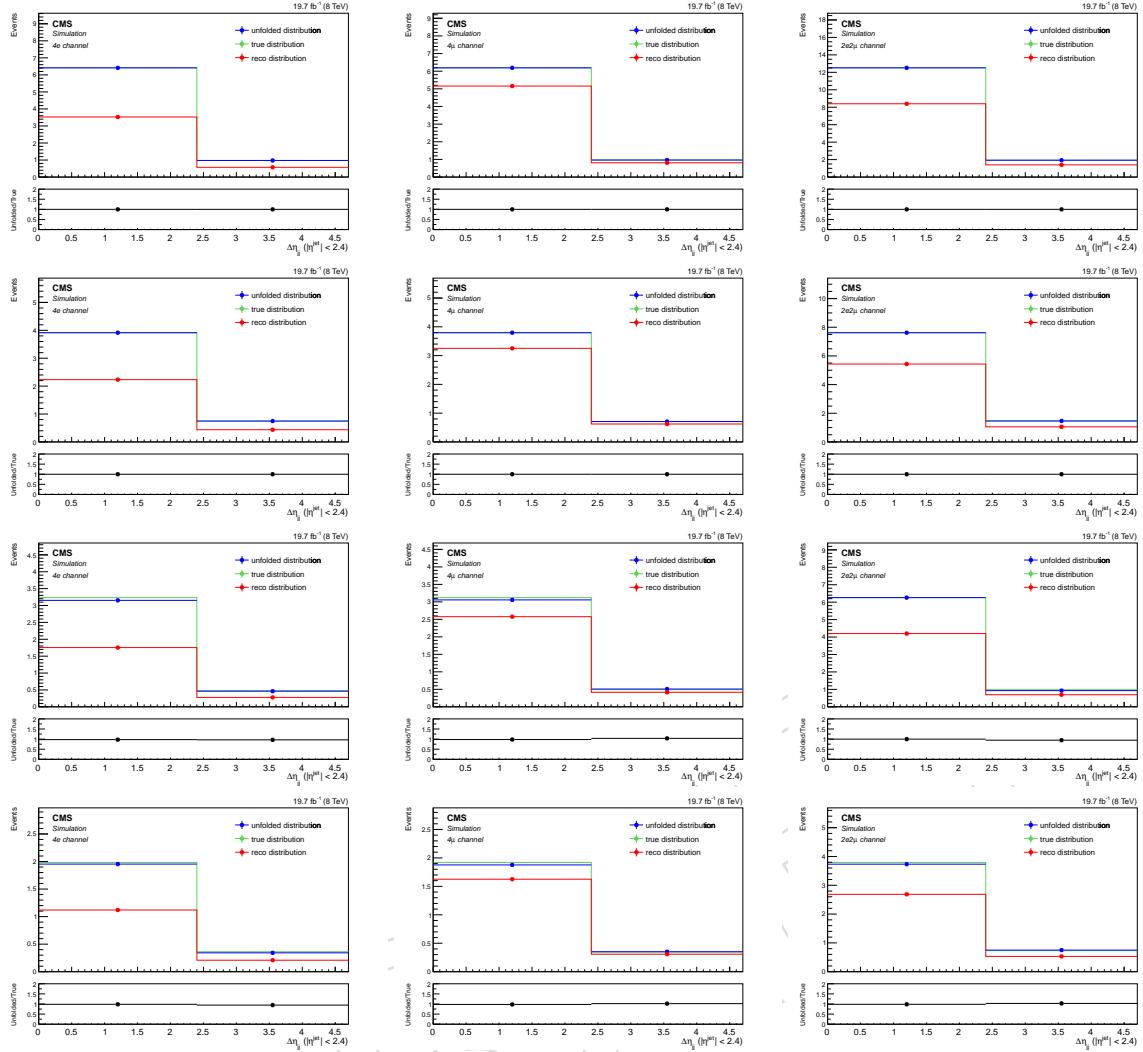


Figure 62: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of $\Delta\eta_{jj}$ (with $|\eta^{jet}| < 2.4$) for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

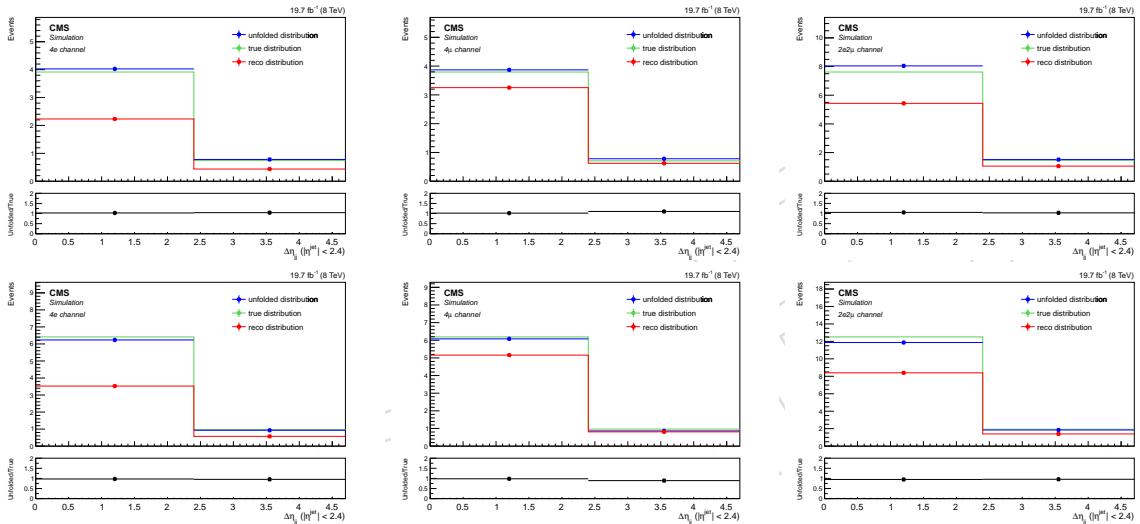


Figure 63: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of $\Delta\eta_{jj}$ (with $|\eta^{jet}| < 2.4$) for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

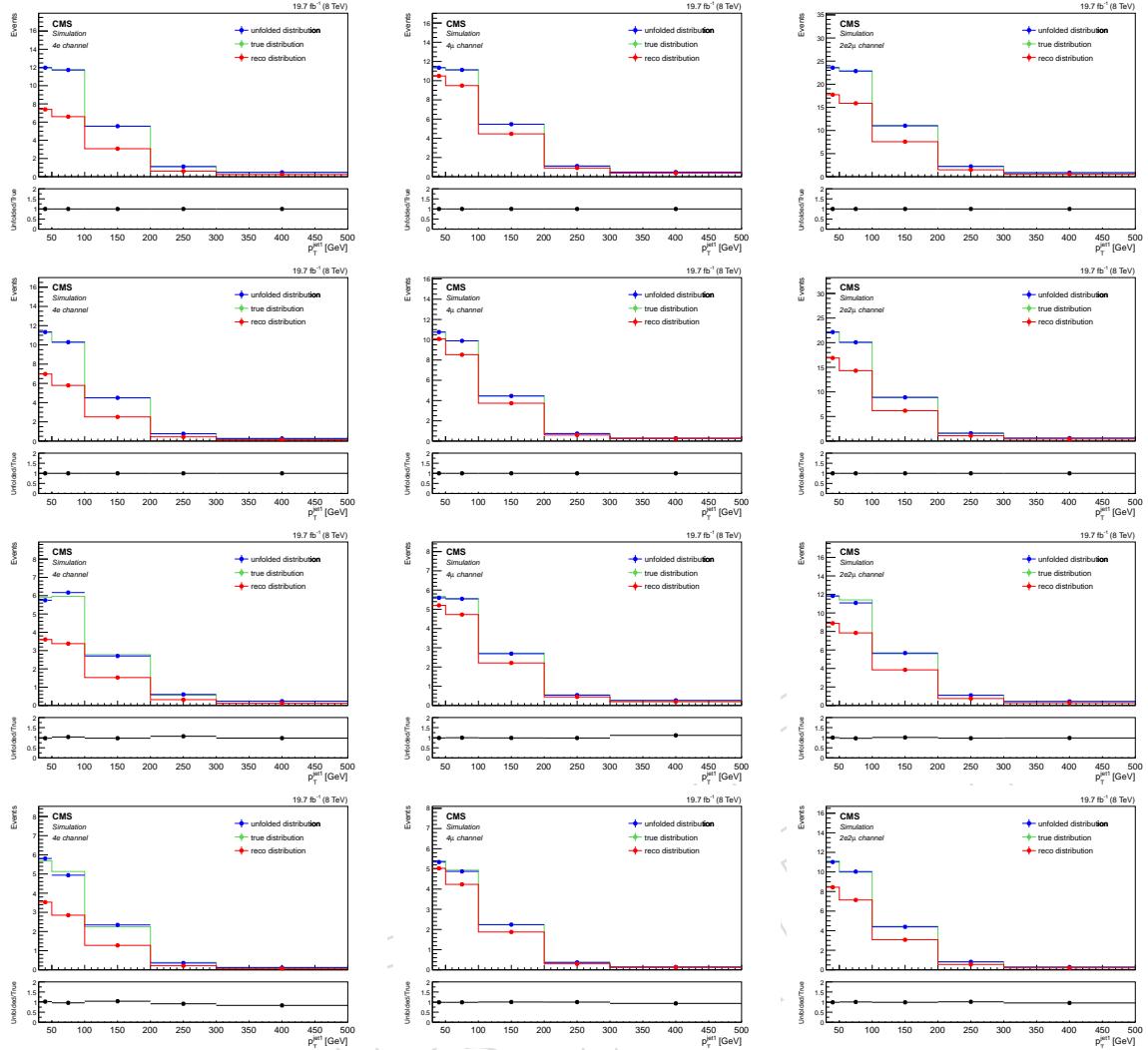


Figure 64: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of the p_T of the leading jet, for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

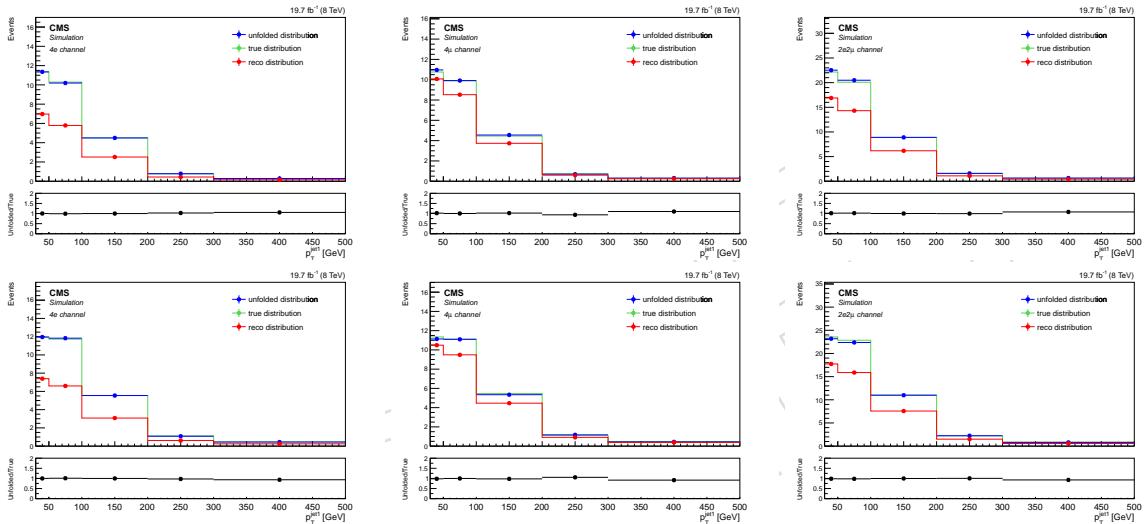


Figure 65: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of the p_T of the leading jet, for the 4e (left), 4 μ (center) and 2e2 μ (right) final states.

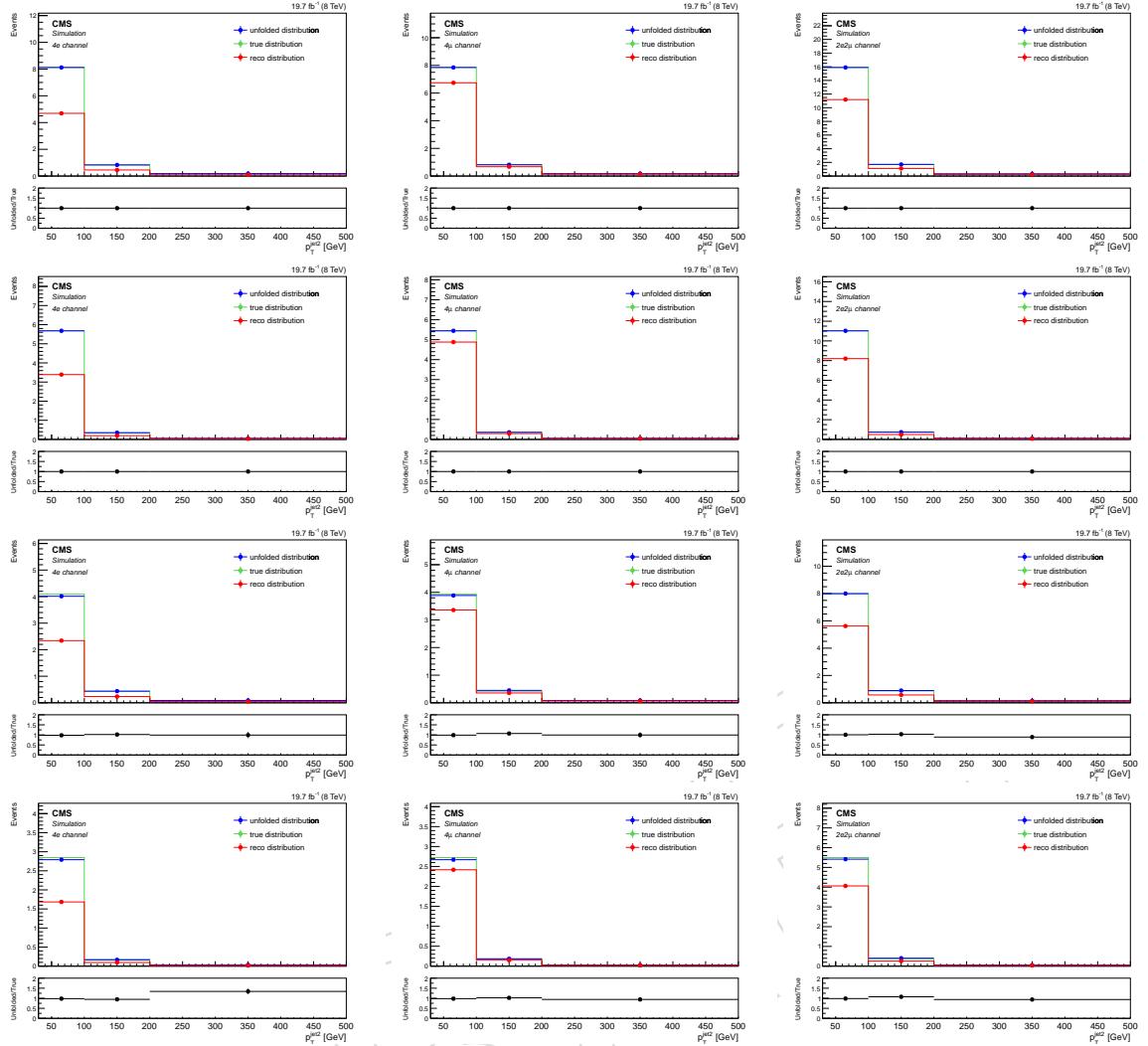


Figure 66: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of the p_T of the sub-leading jet, for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

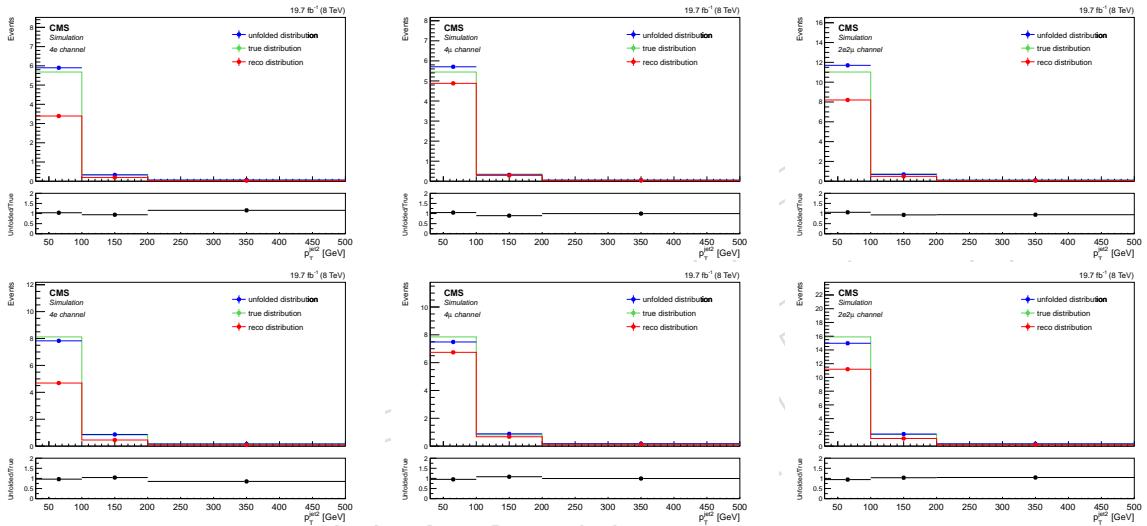


Figure 67: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of the p_T of the sub-leading jet, for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

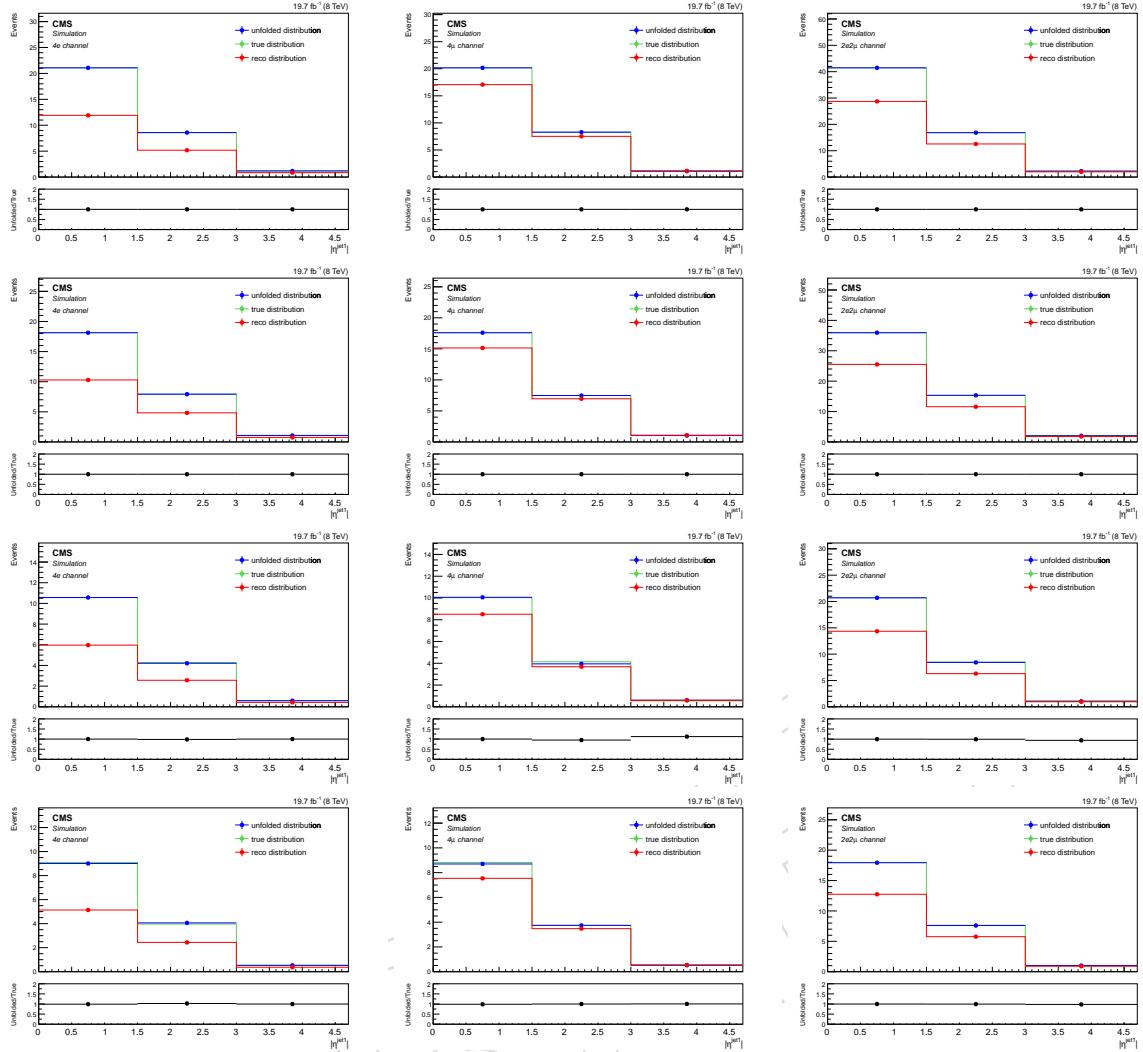


Figure 68: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of the η of the leading jet, for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

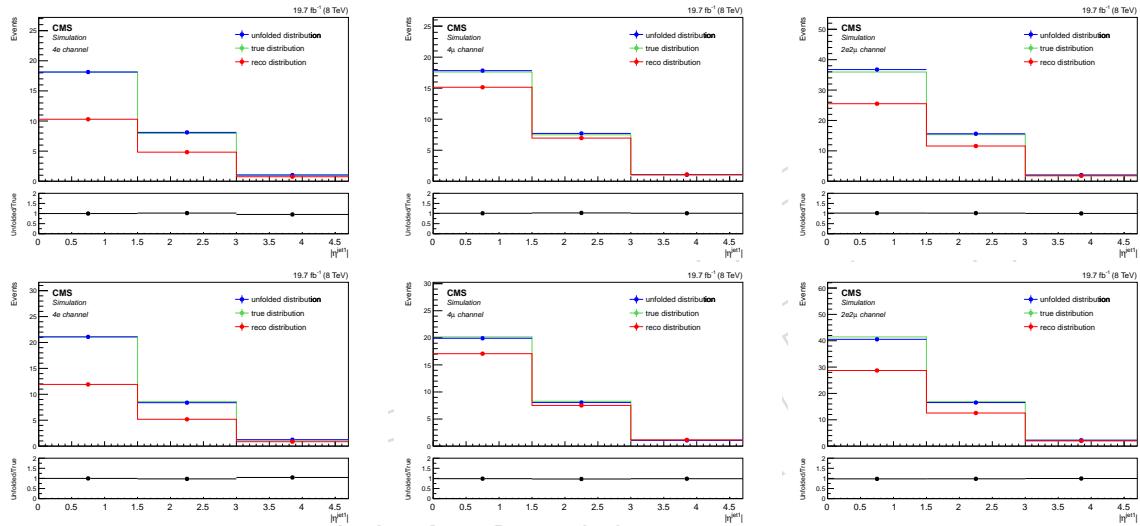


Figure 69: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of the η of the leading jet, for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

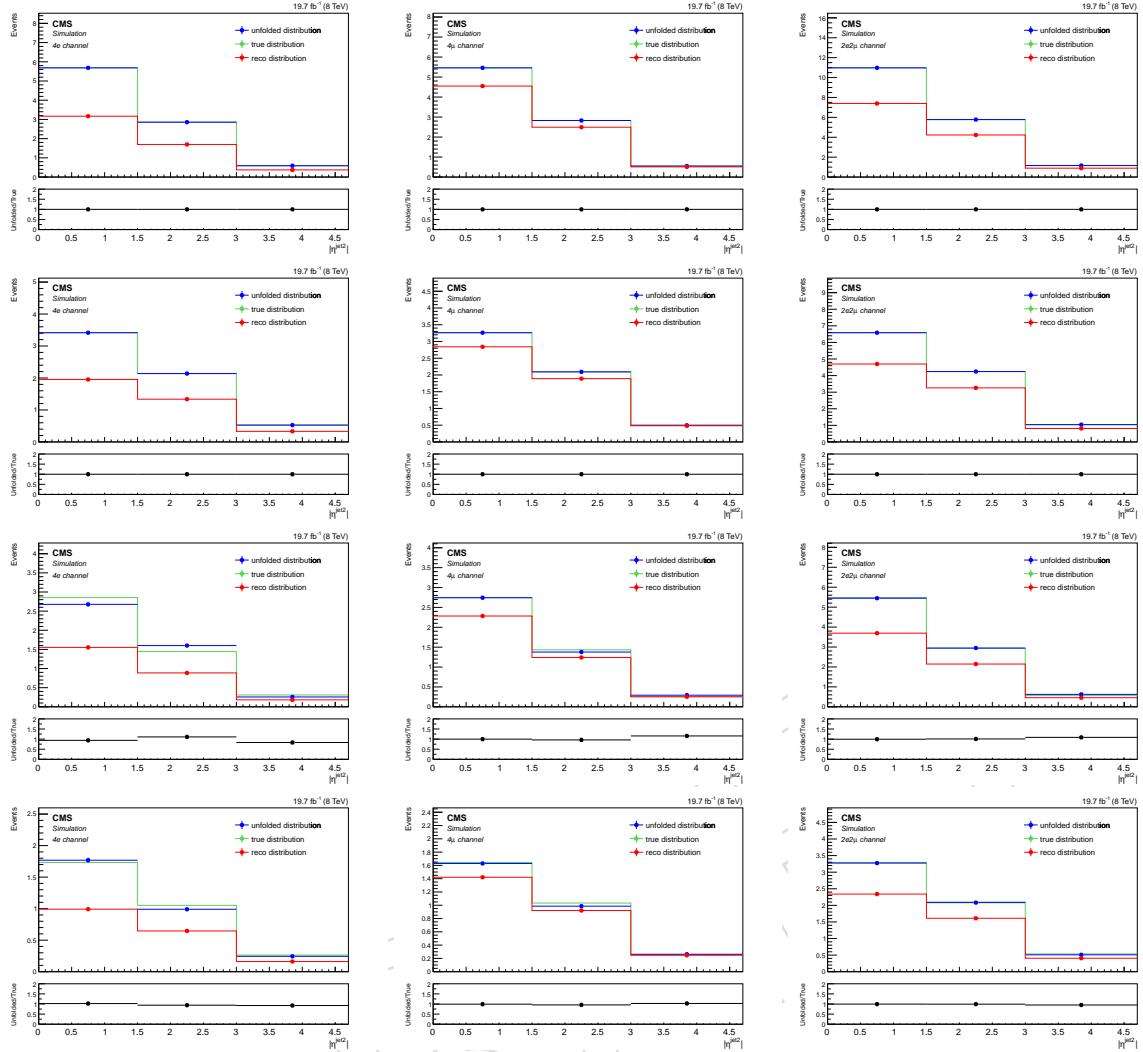


Figure 70: Unfolding tests. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on Powheg distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on Powheg distribution using the two different halves of the total sample set. Results are reported as a function of the η of the sub-leading jet, for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

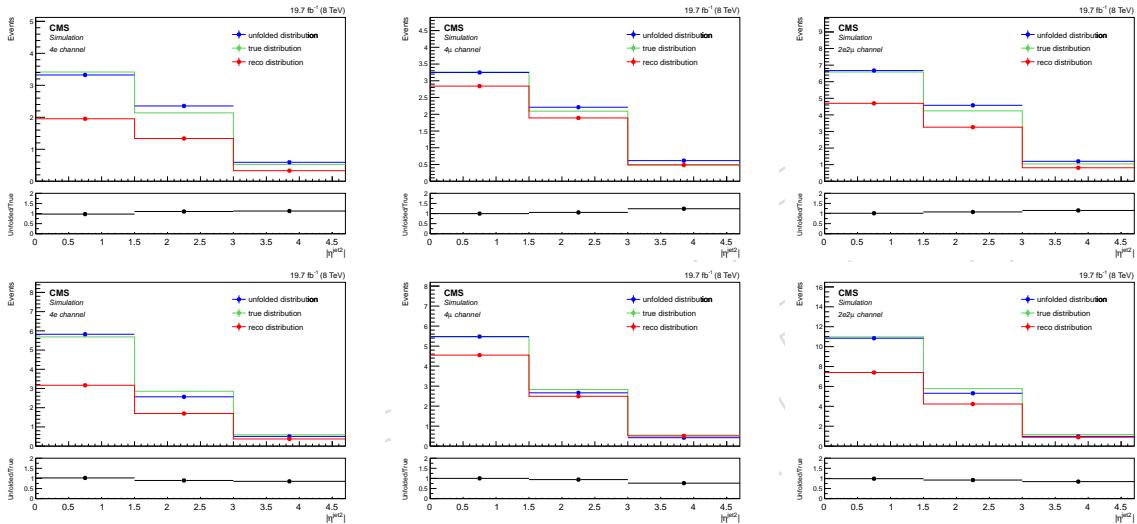


Figure 71: Unfolding tests. MadGraph matrix applied on Powheg distribution using the full set (top), Powheg matrix applied on MadGraph distribution using the full set (bottom). Results are reported as a function of the η of the sub-leading jet, for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final states.

404 **A.3.1 Choice of the unfolding algorithm**

405 In order to choose the proper algorithm, results obtained using the SVD method and the so-
 406 called iterative “Bayesian” method are compared.

407

408 The SVD method depends on the choice of the regularization parameter (k_{reg}), which is needed
 409 since matrix inversion is sensitive to the statistical fluctuations and is in the range $[1, N_{bin}]$. The
 410 optimal value of k_{reg} is the one for which the errors associated to the unfolding procedure are
 411 small when compared to the statistical ones. This parameter can be seen as a cut-off for quickly
 412 oscillating terms corresponding to data statistical fluctuations, as described in detail in [28].
 413 Choosing a too small regularization parameter gets ride of these spurious fluctuations but the
 414 result tends to be biased by the MC truth. On the other hand, a too large regularization param-
 415 eter will decrease the MC dependence but give too large importance to data fluctuations which
 416 will be interpreted as real shape.

417

418 The Bayes method using an iterative approach requires a choice for the maximum number of
 419 iterations to be done that, as discussed in [29], should not lead to different results. In this anal-
 420 ysis, the iteration procedure is stopped as soon as the new unfolded distribution is compatible
 421 with the one obtained from the previous step. This is quantified by computing the $\chi^2/n.d.f.$
 422 of the change in the unfolded distribution for every iteration step. The number of iterations in
 423 the Bayes unfolding procedure is set as the one when the next iteration is compatible with the
 424 current one by computing:

- 425 • $\sum_{i=1}^{N_{bin}} (b_n^i - b_{n+1}^i)^2$, the sum of the deviations
- 426 • $\chi_n^2 = \sum_{i=1}^{N_{bin}} \frac{(b_n^i - b_{n+1}^i)^2}{(db_n^i)^2}$
- 427 • $\chi_{n+1}^2 = \sum_{i=1}^{N_{bin}} \frac{(b_n^i - b_{n+1}^i)^2}{(db_{n+1}^i)^2}$

428 where b_n^i is the content of the i th-bin of the histogram obtained using n iterations and db_n^i is the
 429 systematic uncertainty associated to b_n^i .

430

431 The unfolded m_{ZZ} -distributions are obtained using the SVD algorithm. In this case the reg-
 432 ularization parameter is chosen to be $k_{reg} = 4$, the default value ($N_{bin}/2$), since as shown in
 433 Figure 72 the distributions obtained using $k_{reg} = 2$ are clearly biased by the MC truth. On the
 434 other hand, all the other jet-related variables distributions are unfolded using the Bayesian al-
 435 gorithm, with four iterations. As shown in Figure 73, this choice is not biased by the MC truth
 436 distribution.

437

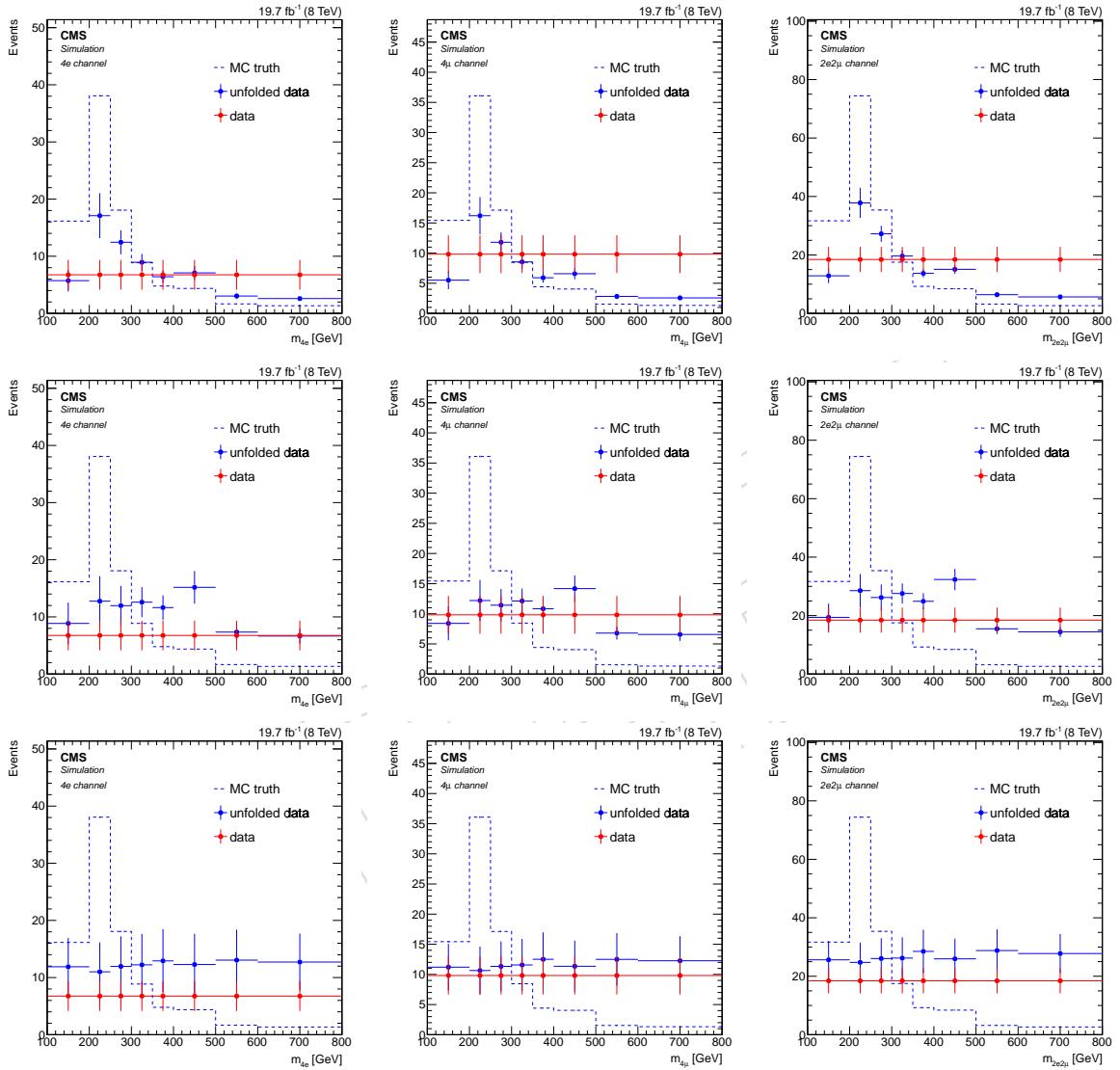


Figure 72: A flat distribution (in red) is unfolded using the SVD algorithm with $k_{reg} = 2$ (top) and $k_{reg} = 4$ (center) and using the D'Agostini method with 4 iterations (bottom). Results are reported as a function of the invariant mass of the 4-lepton system and for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final state.

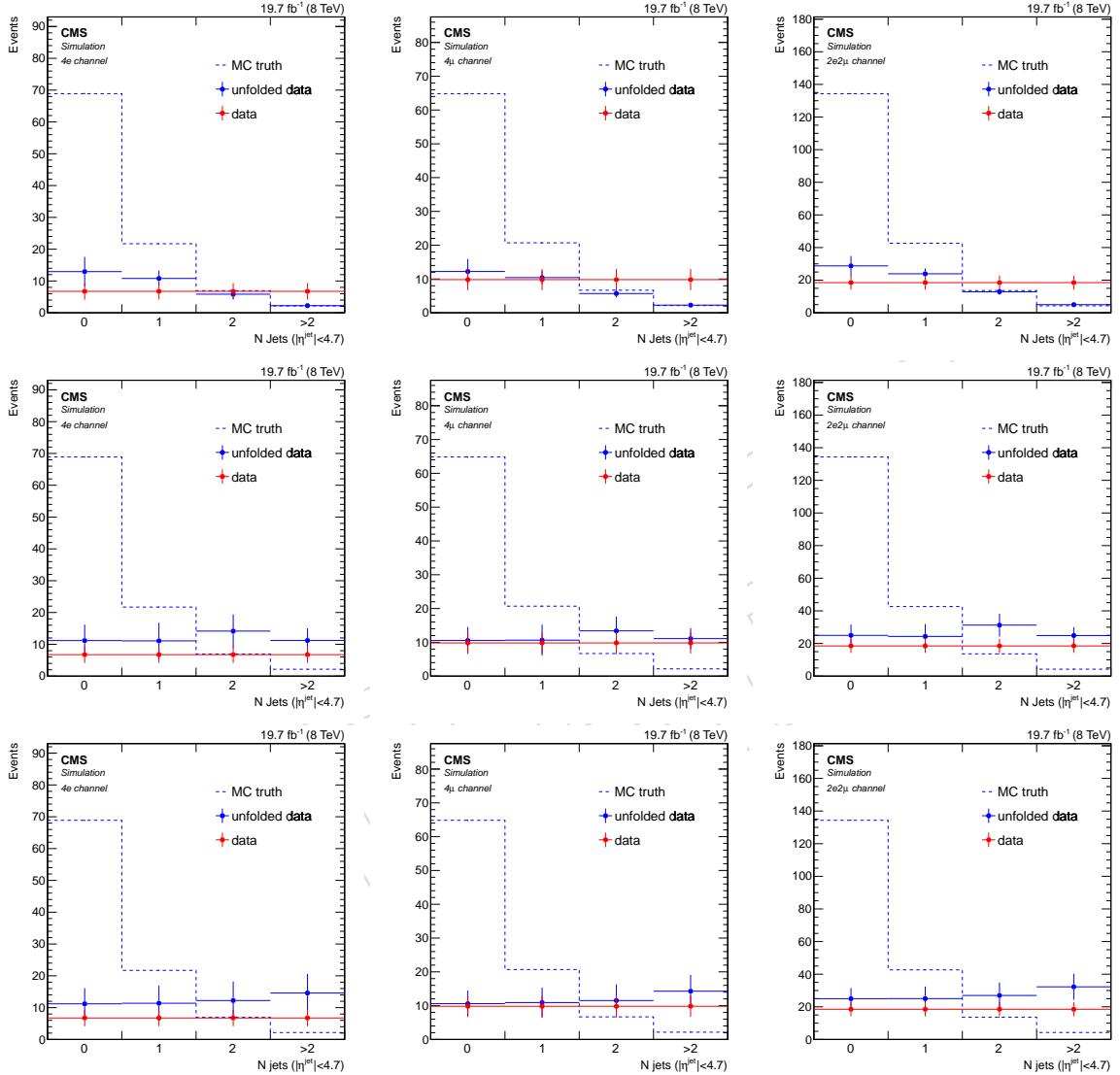


Figure 73: A flat distribution (in red) is unfolded using the SVD algorithm with $k_{reg} = 2$ (top) and $k_{reg} = 4$ (center) and using the D'Agostini method with 4 iterations (bottom). Results are reported as a function of the number of jets and for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final state.

438 **A.4 Systematic Uncertainties**

439 To evaluate the effects of the different sources of systematic uncertainty on the unfolded dis-
 440 tributions, the whole unfolding procedure is repeated for every contribution. The following
 441 systematic uncertainties are considered for the result and their contributions are individually
 442 reported in Figures from 74 to 106, for all variables and final states.

443 **A.4.1 Unfolding (measurement and response matrix uncertainties)**

444 Uncertainties due to the unfolded procedure are computed by the `RooUnfold` toolkit and con-
 445 sist of two types, the “measurement uncertainties” and the “response matrix uncertainties”.
 446 The former come from the propagation of the statistical errors of the distribution that has to be
 447 unfolded through the unfolding matrix, in particular bin-to-bin correlations, and must be con-
 448 sidered as statistical uncertainties. The latter are uncertainties on the response matrix elements
 449 due to the limited MC statistics but can be neglected.

450 **A.4.2 σ_{qq} and σ_{gg} ratio**

451 The response matrix is built adding the MCFM sample, which describes $gg \rightarrow ZZ \rightarrow 4\ell$ pro-
 452 cesses, to the MadGraph (or Powheg) one, that contains $qq/qg \rightarrow ZZ \rightarrow 4\ell$ events. The
 453 cross-sections used to simulate these processes are estimated with MCFM and are affected by
 454 an uncertainties ($d\sigma_{qq \rightarrow ZZ} = \pm 4.44\%$ and $d\sigma_{gg \rightarrow ZZ} = \pm 25.36\%$). This leads to a systematic
 455 effect that is evaluated building up new response matrices by varying the cross-sections by
 456 $\pm d\sigma_{qq/gg}$. The combinations returning the largest discrepancy are those in which the correc-
 457 tions of σ_{qq} and σ_{gg} have opposite sign, since in this case the difference of shapes is enhanced.
 458 The unfolded procedure is repeated using these new matrices and the systematic uncertainty is
 459 evaluated taking the difference between the largest and the smallest results for each bin. This
 460 contribution is found to be less than 1%.

461 **A.4.3 Choice of MC generator**

462 This systematic uncertainty is computed comparing unfolded data distributions obtained ap-
 463 plying the `Madgraph+MCFM+Phantom` or the `Powheg+MCFM+Phantom` response matrix and
 464 taking the difference between the two results. It varies between 1% and 10%.

465 **A.4.4 Irreducible and reducible backgrounds**

466 Uncertainties due to background estimate are evaluated by creating a new “data-background”
 467 distribution, in which the background is shifted up and down by its uncertainties, that is un-
 468 folded using standard matrices. This procedure is done for both reducible (1-16%) and irre-
 469ducible (0.1 - 5.5%) contributions.

470 **A.4.5 Unfolded Data over MC Truth ratio**

471 This systematic uncertainty is evaluated building up a new response matrix whose elements
 472 are weighted for the ratio between the unfolded data and the generator level information. The
 473 difference between the unfolded results obtained by the application of this weighted response
 474 matrix and the standard one is considered as a systematic error [34]. This uncertainty ranges
 475 from 1% to 29%, in the last bins where the statistics is very low, and it is thus one of the largest
 476 contributions.

A.4.6 Lepton Efficiency

This uncertainty comes from the measurement of data/MC efficiency scale factors used to weight MC distributions for lepton reconstruction. It is evaluated shifting up and down the efficiency by its error and building up two new response matrices to be used in the unfolding procedure. The difference between the unfolded results obtained applying these two “up-” and “down-”matrices is taken as the systematic uncertainty. Its contributions is found to ranges between 1% and 6%.

A.4.7 Jet energy resolutions (JER)

This correction has been estimated for data and MC in [21]. MC slightly overestimates the resolution compared to data. The effect is propagated accordingly by smearing the samples used to build the response matrices. This uncertainty is less than 3.5%.

A.4.8 Jet energy scale corrections (JES)

This uncertainty is calculated by rescaling the jet p_T spectrum up and down by one standard deviation of the measured jet energy correction. This is done separately in data [35] and in MC reconstructed distributions used for the response matrix. The distributions obtained using the standard corrections and its variations are then compared and their difference is taken as systematic uncertainty. When the jet transverse momenta of reconstructed MC distributions are shifted, this uncertainty ranges from 1.4% to 6.9%, while in the case of modified data distributions, it grows up to 66% in the last bins and it is thus the largest contribution affecting the measurement.

DRAFT

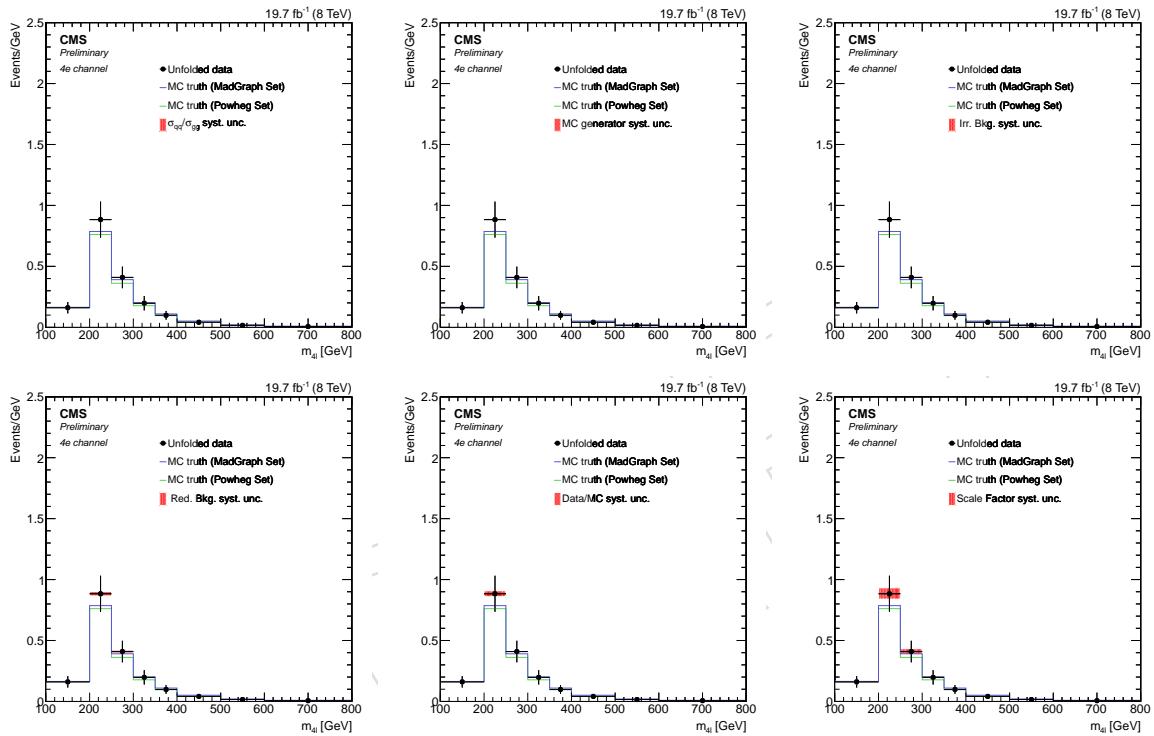


Figure 74: Effects of the different sources of systematic uncertainty on the unfolded distributions of the 4ℓ mass, for the $4e$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

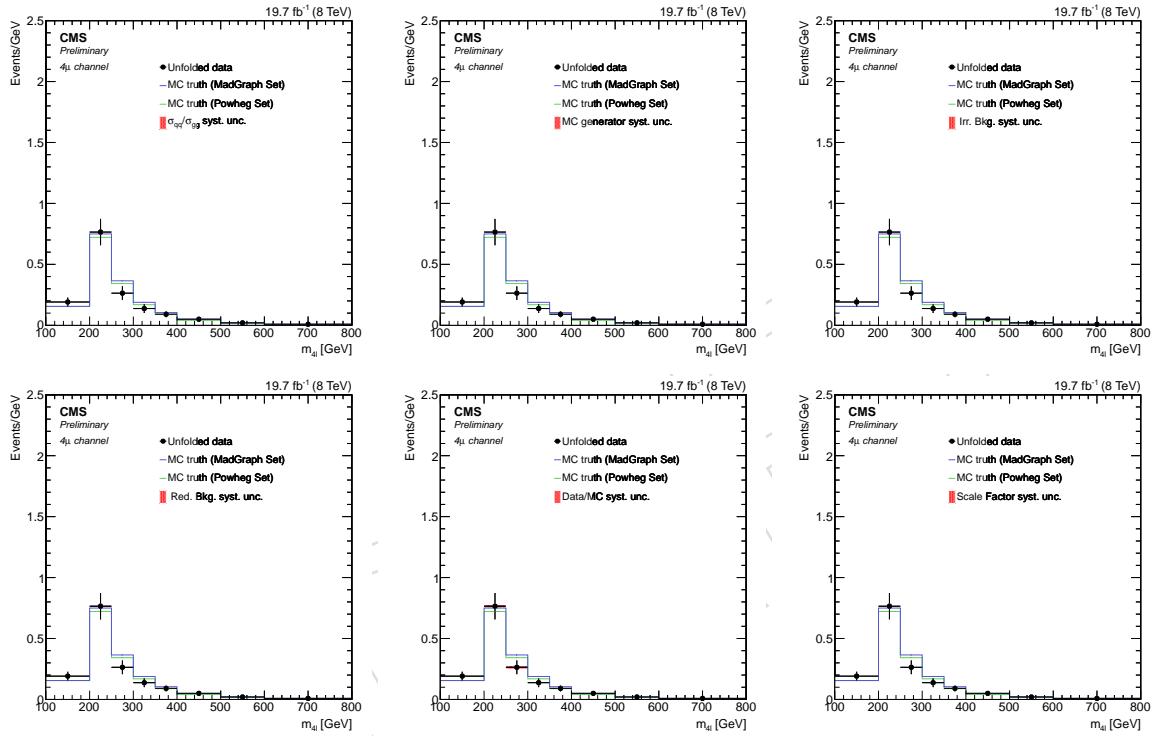


Figure 75: Effects of the different sources of systematic uncertainty on the unfolded distributions of the 4ℓ mass, for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

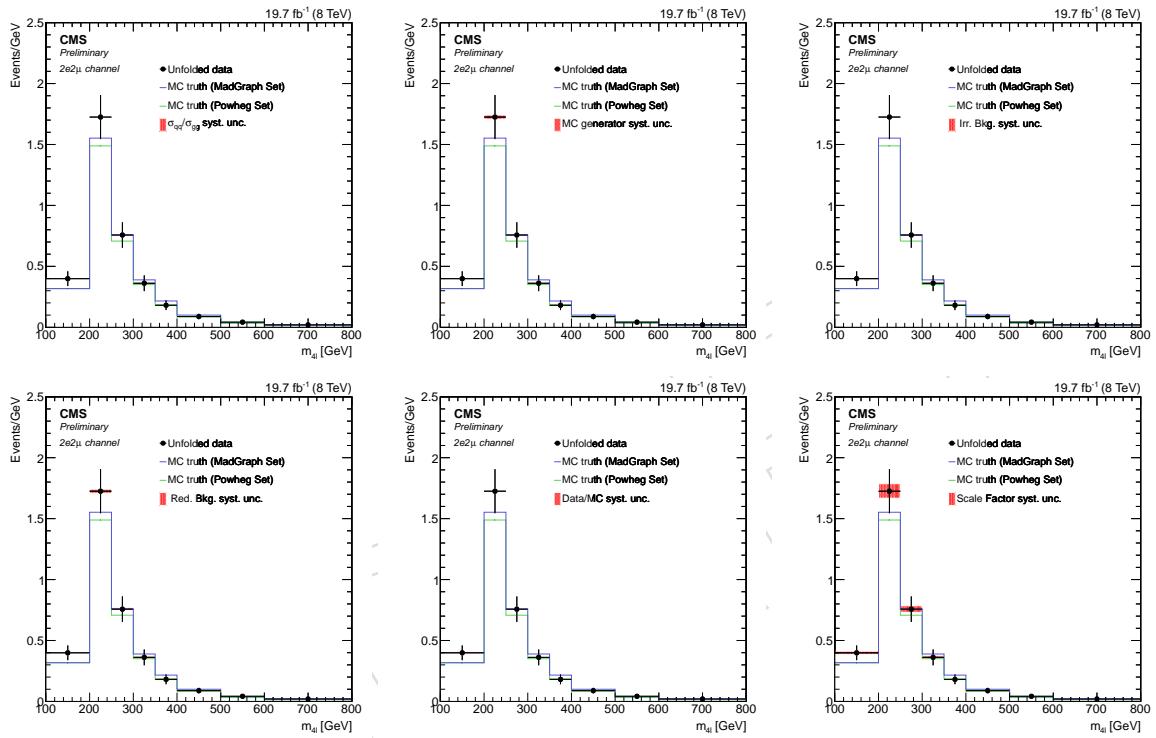


Figure 76: Effects of the different sources of systematic uncertainty on the unfolded distributions of the 4ℓ mass, for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

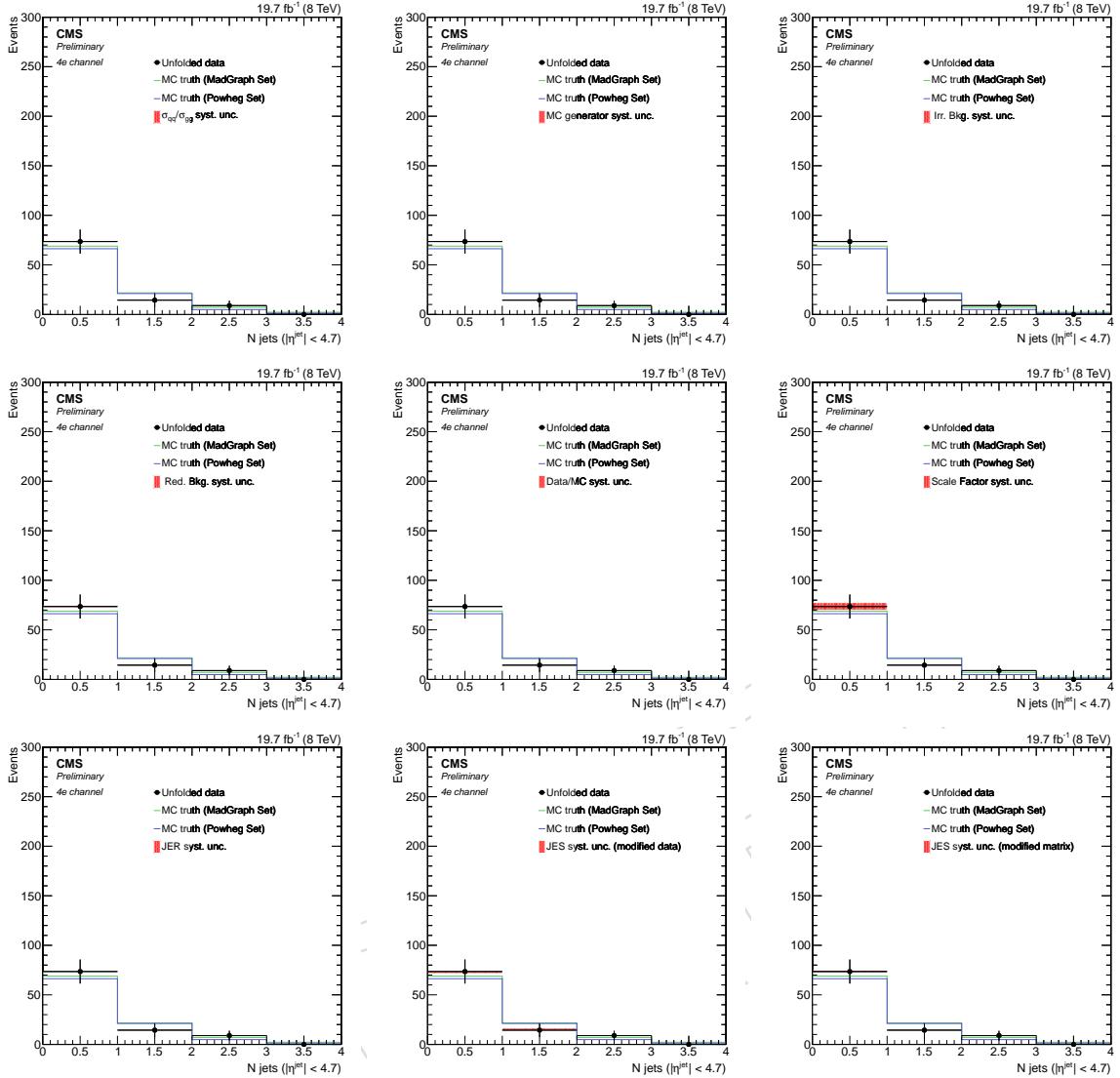


Figure 77: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of jets (with $|\eta^{\text{jet}}| < 4.7$), for the $4e$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

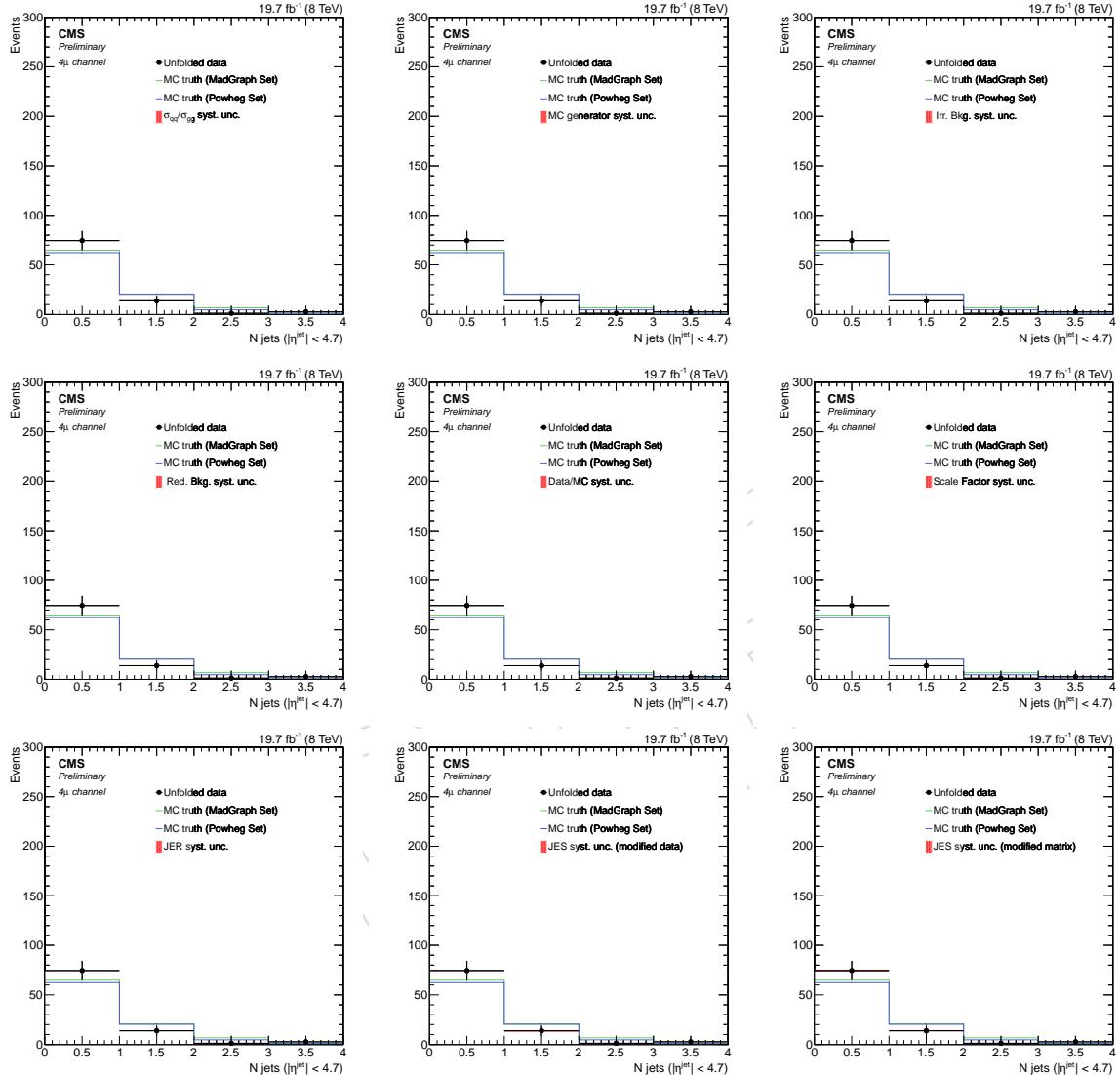


Figure 78: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of jets (with $|\eta^{\text{jet}}| < 4.7$), for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

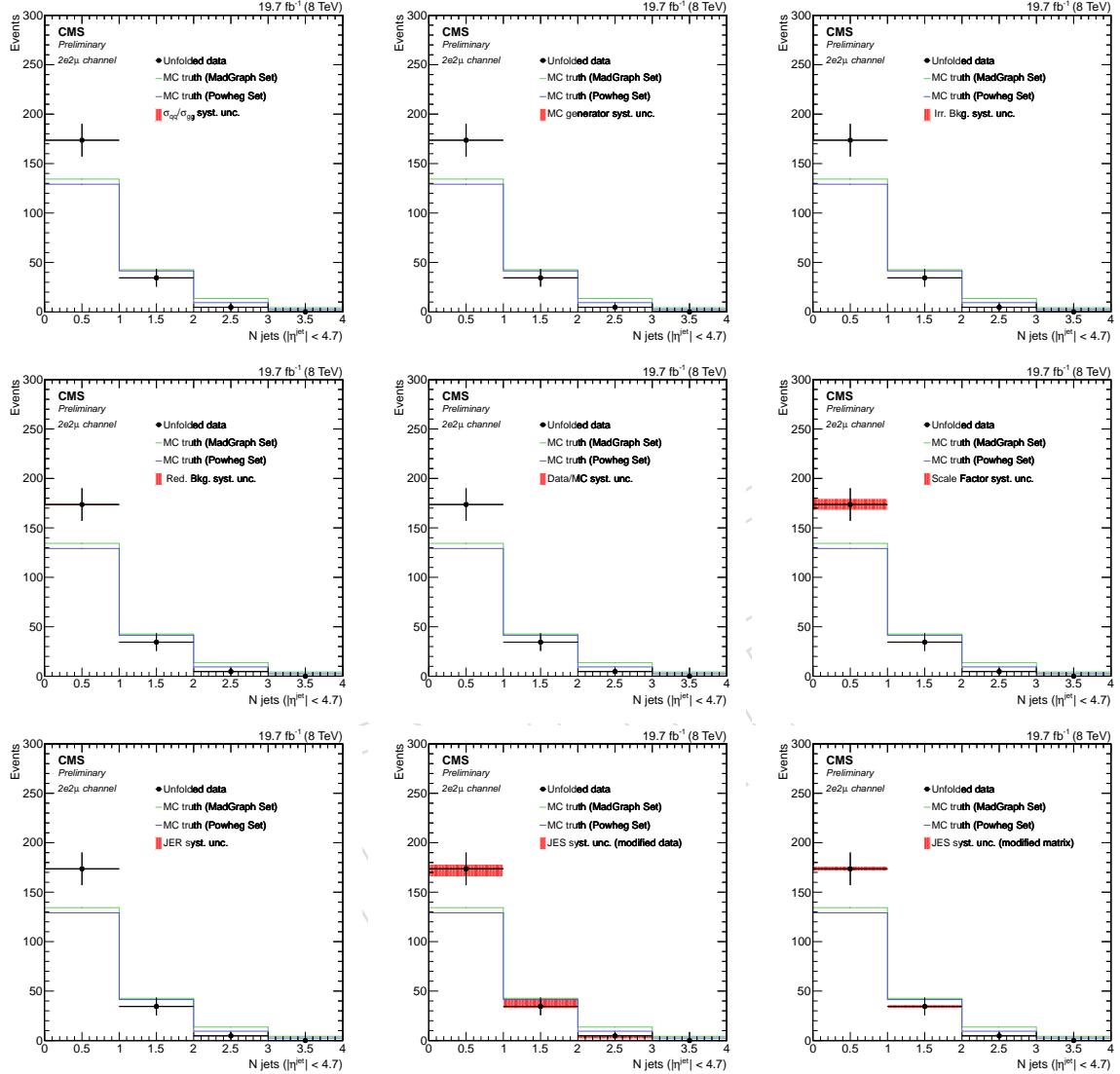


Figure 79: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of jets (with $|\eta^{\text{jet}}| < 4.7$), for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

A Unfolding

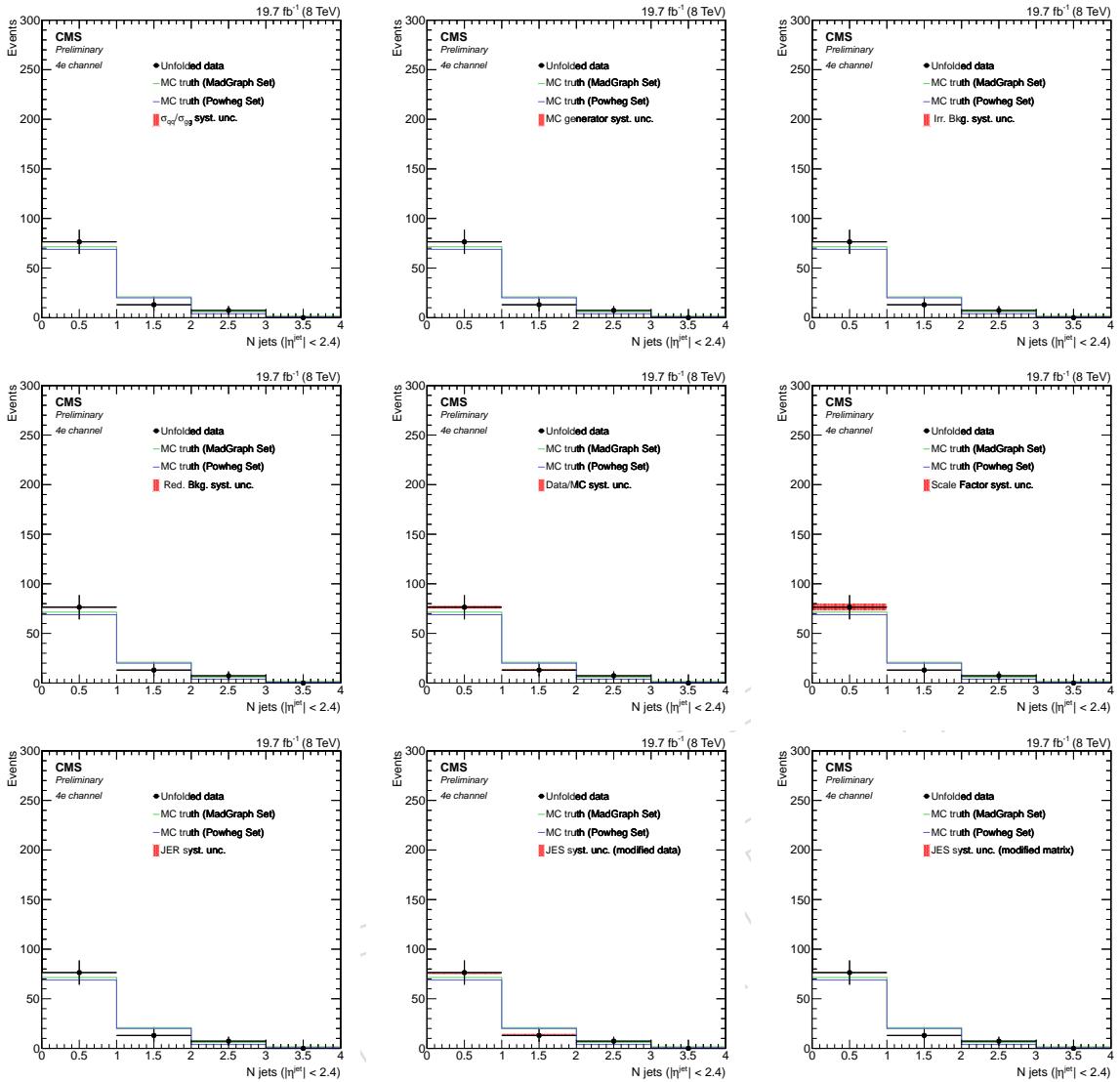


Figure 80: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of central jets (with $|\eta^{\text{jet}}| < 2.4$), for the $4e$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

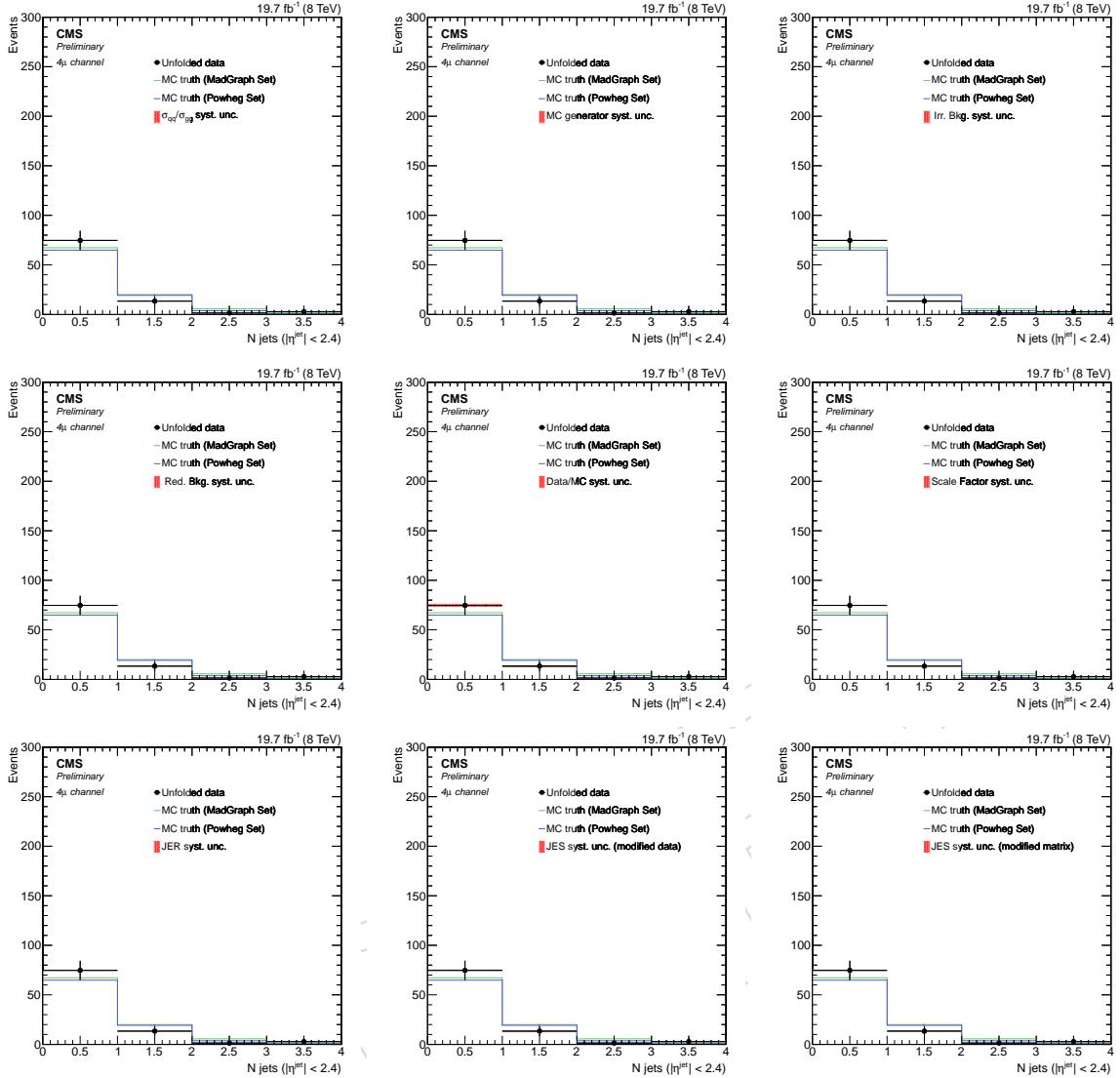


Figure 81: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of central jets (with $|\eta^{\text{jet}}| < 2.4$), for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

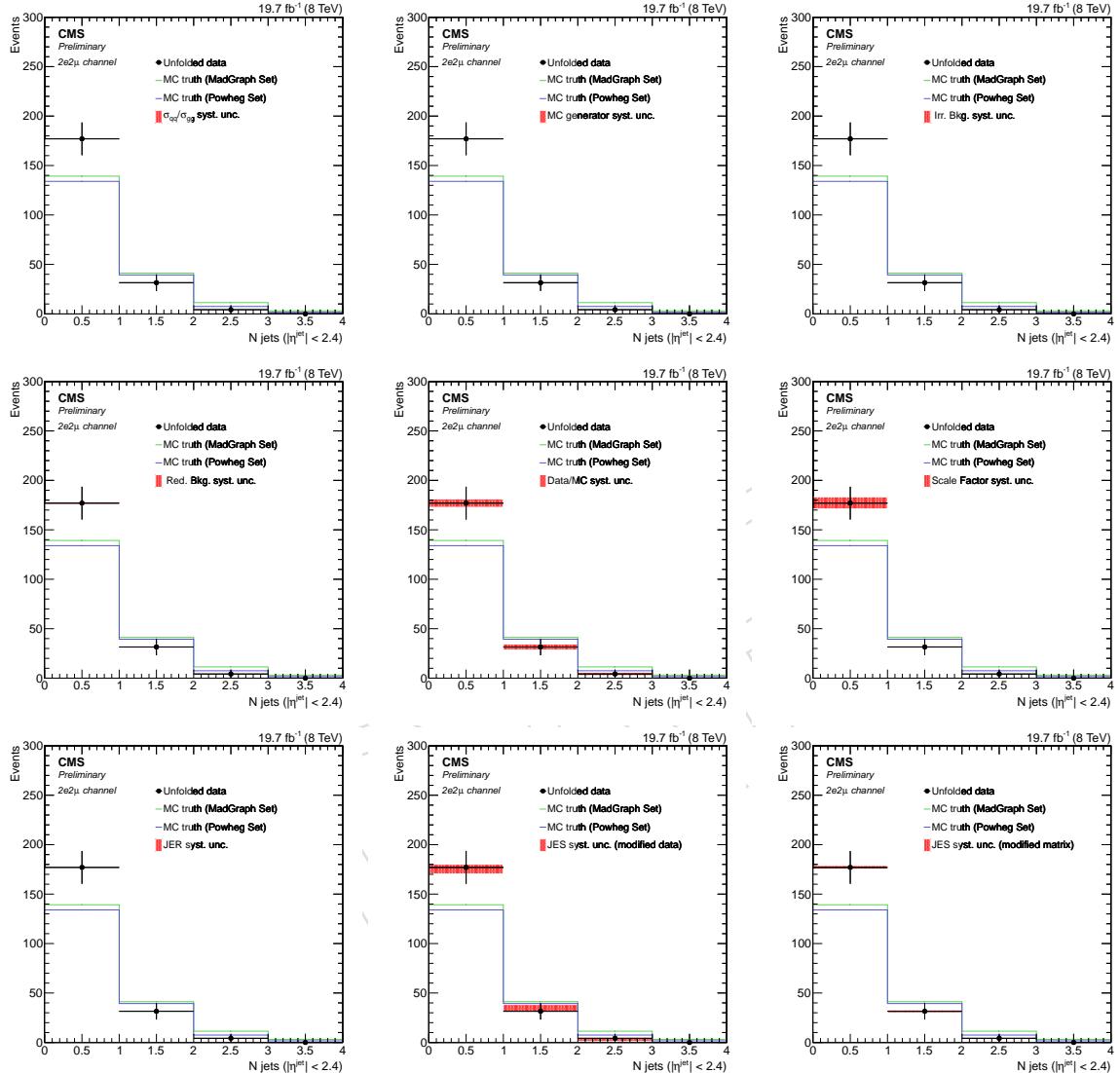


Figure 82: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of central jets (with $|\eta^{\text{jet}}| < 2.4$), for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

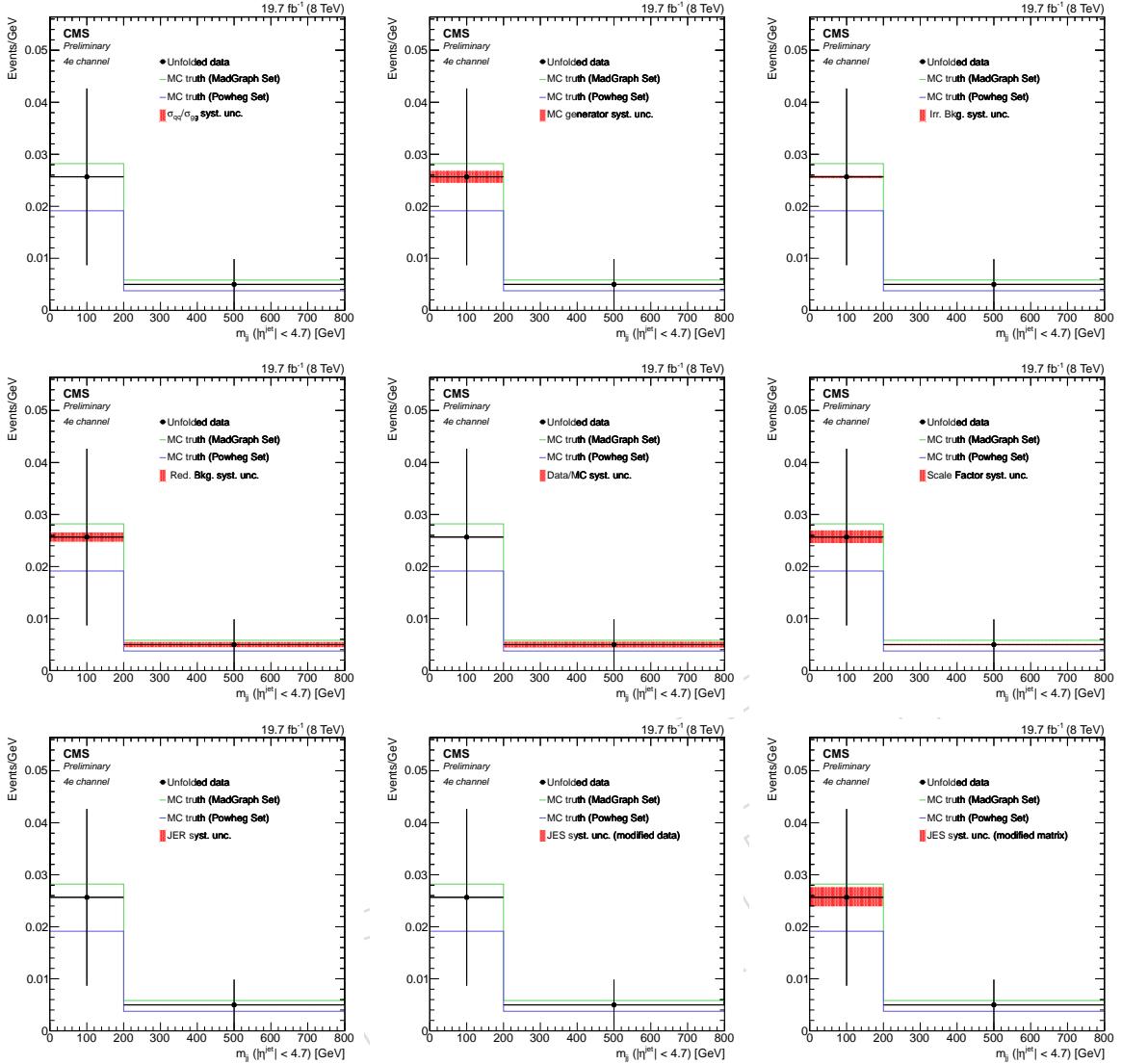


Figure 83: Effects of the different sources of systematic uncertainty on the unfolded distributions of m_{jj} (with $|\eta^{jet}| < 4.7$), for the $4e$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

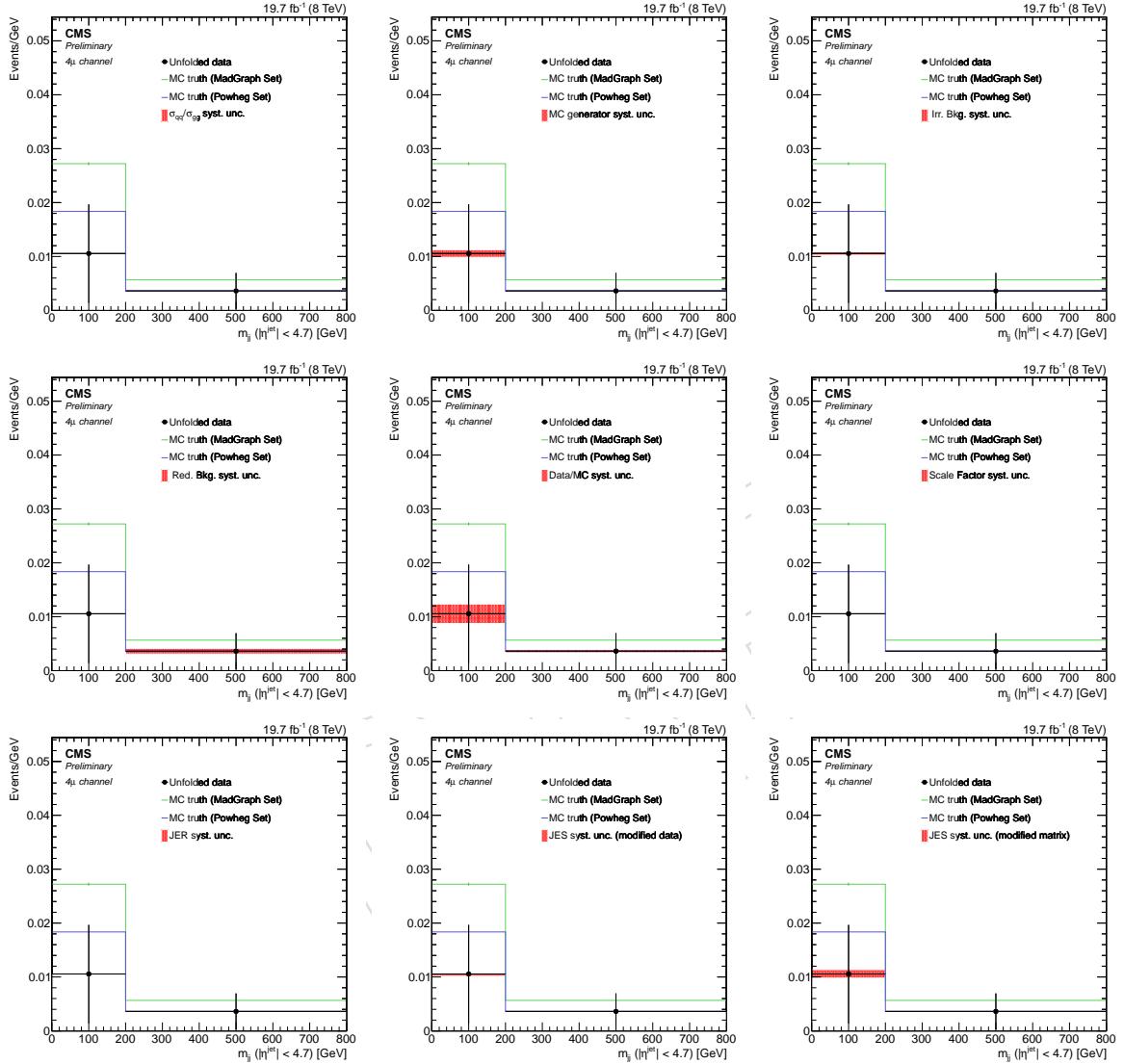


Figure 84: Effects of the different sources of systematic uncertainty on the unfolded distributions of m_{jj} (with $|\eta^{\text{jet}}| < 4.7$), for the 4μ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The shifted distributions due to the systematic effects are reported, together with the nominal one and the MC truth.

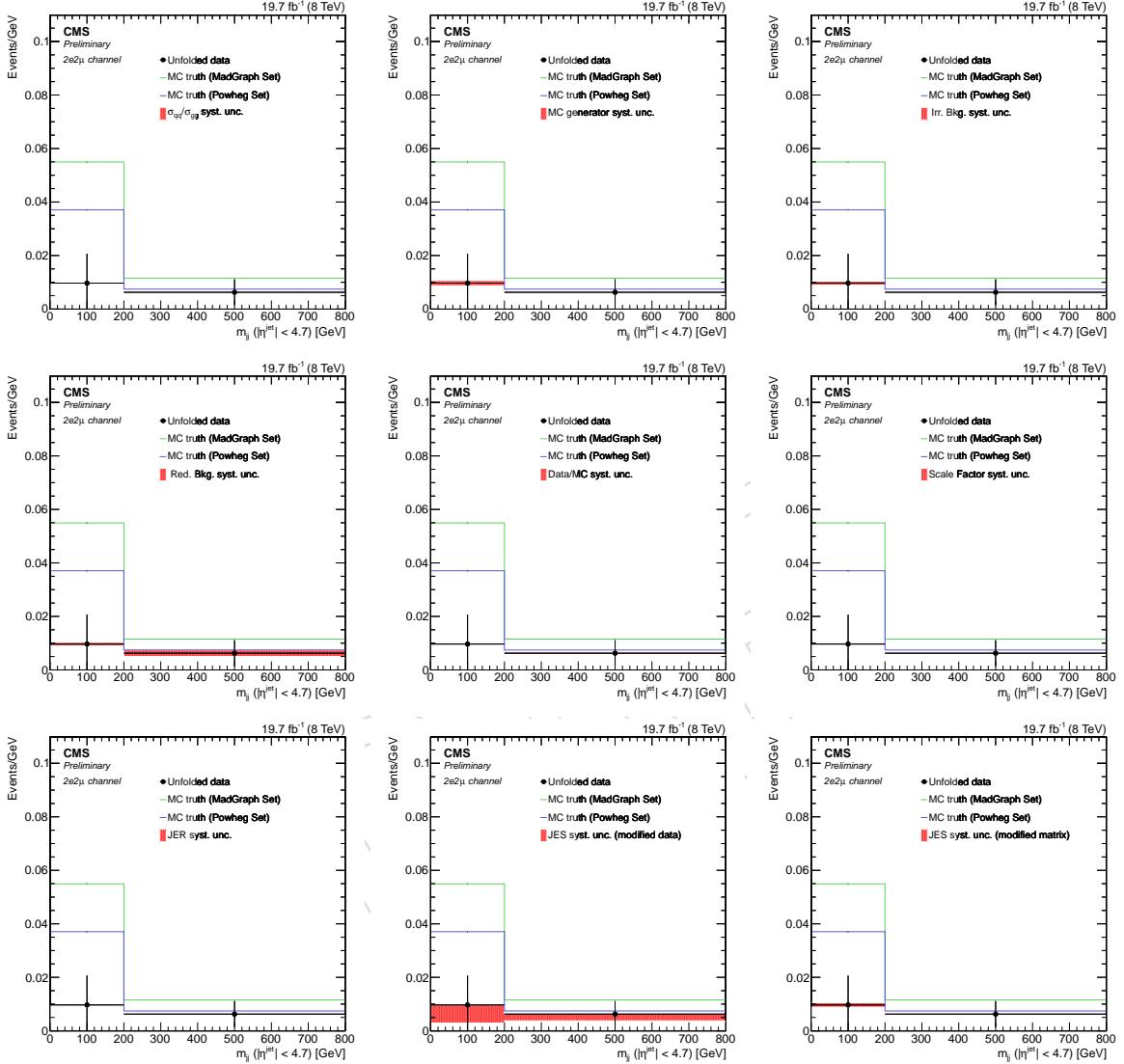


Figure 85: Effects of the different sources of systematic uncertainty on the unfolded distributions of m_{jj} (with $|\eta^{jet}| < 4.7$), for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The shifted distributions due to the systematic effects are reported, together with the nominal one and the MC truth.

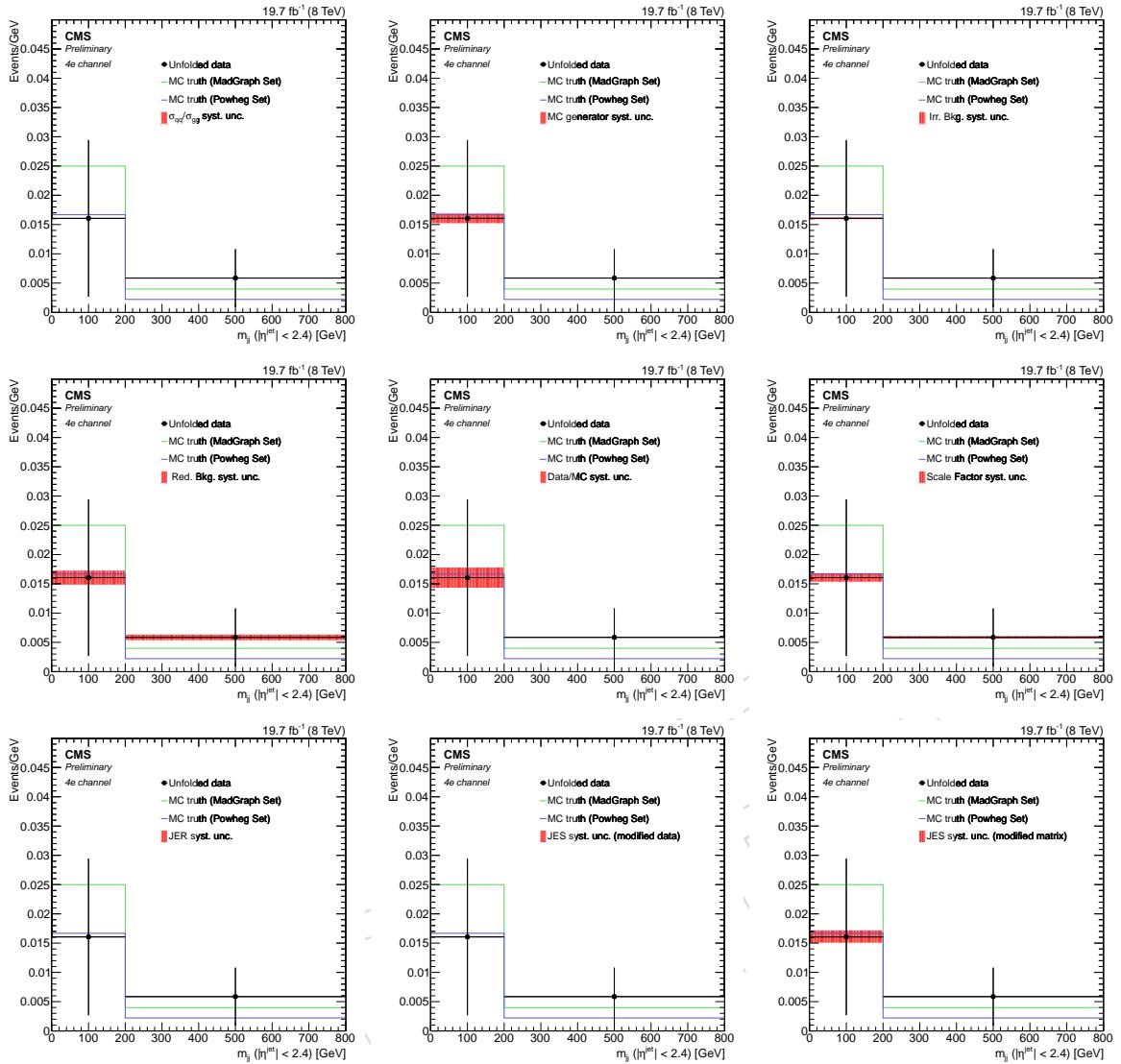


Figure 86: Effects of the different sources of systematic uncertainty on the unfolded distributions of m_{jj} (with $|\eta^{jet}| < 2.4$), for the $4e$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

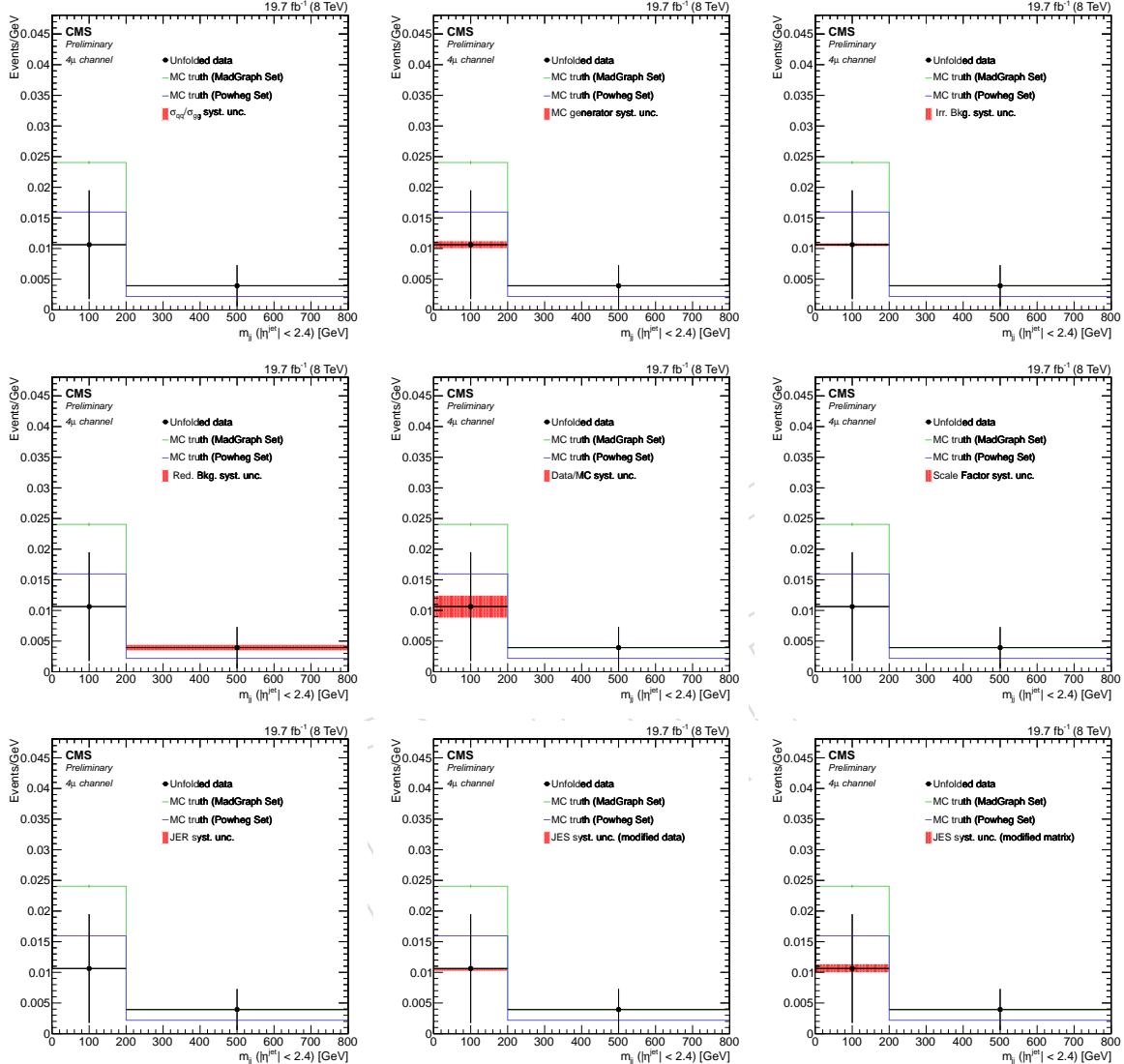


Figure 87: Effects of the different sources of systematic uncertainty on the unfolded distributions of m_{jj} (with $|\eta^{jet}| < 2.4$), for the 4μ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JES modifying data distribution, JES modifying the response matrix. The shifted distributions due to the systematic effects are reported, together with the nominal one and the MC truth.

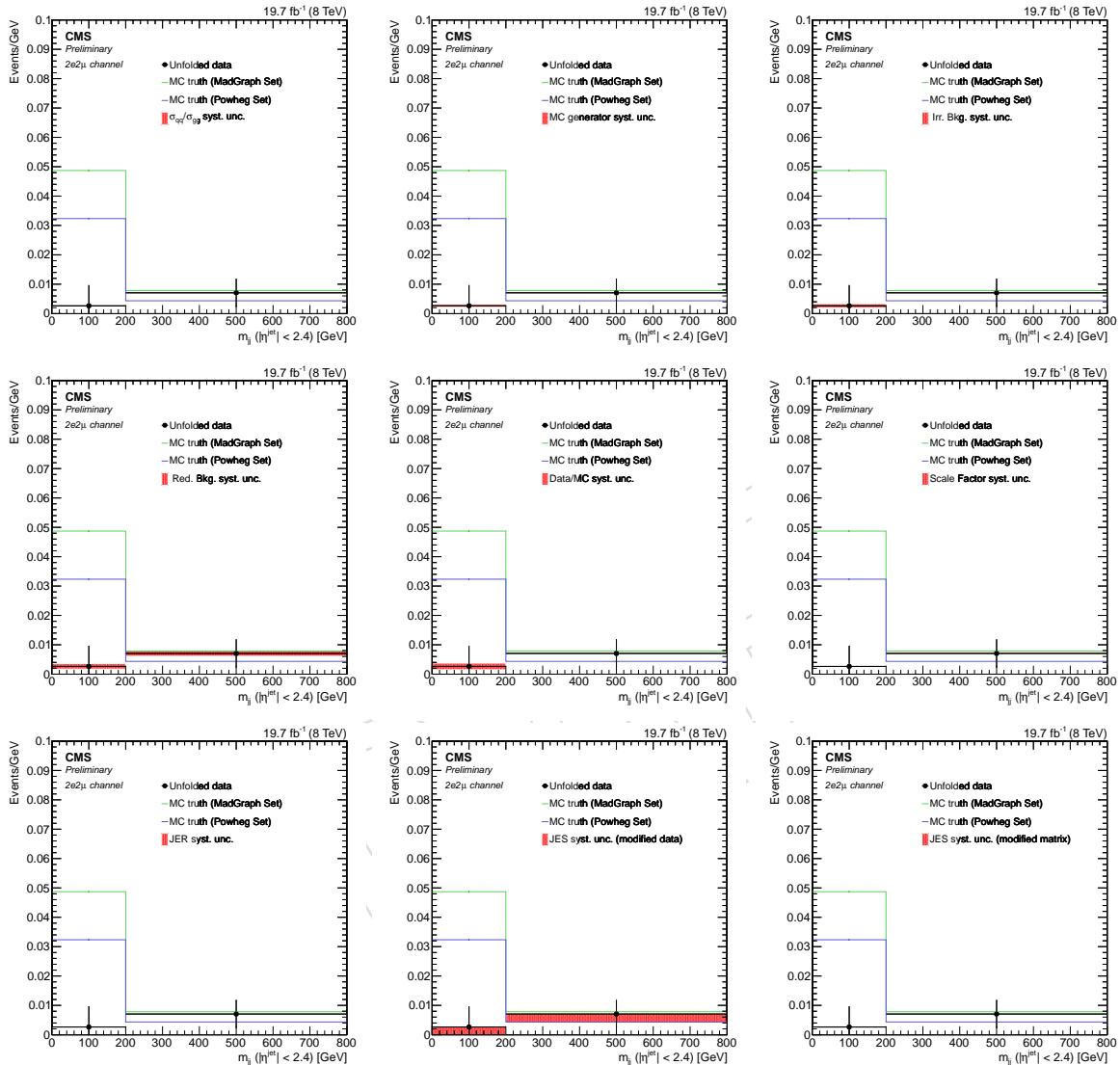


Figure 88: Effects of the different sources of systematic uncertainty on the unfolded distributions of m_{jj} (with $|\eta^{jet}| < 2.4$), for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The shifted distributions due to the systematic effects are reported, together with the nominal one and the MC truth.

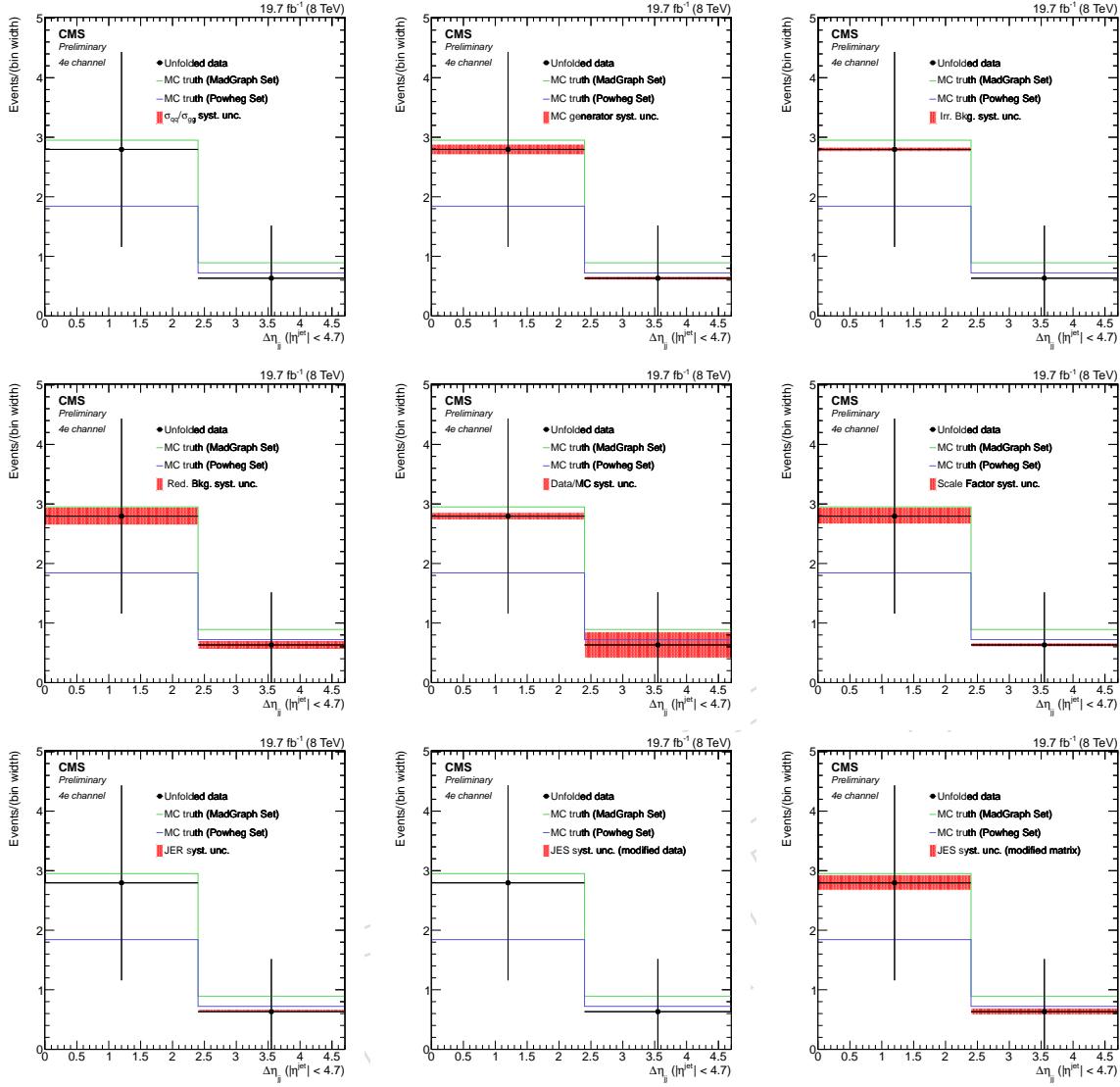


Figure 89: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\Delta\eta_{jj}$ (with $|\eta^{\text{jet}}| < 4.7$), for the $4e$ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

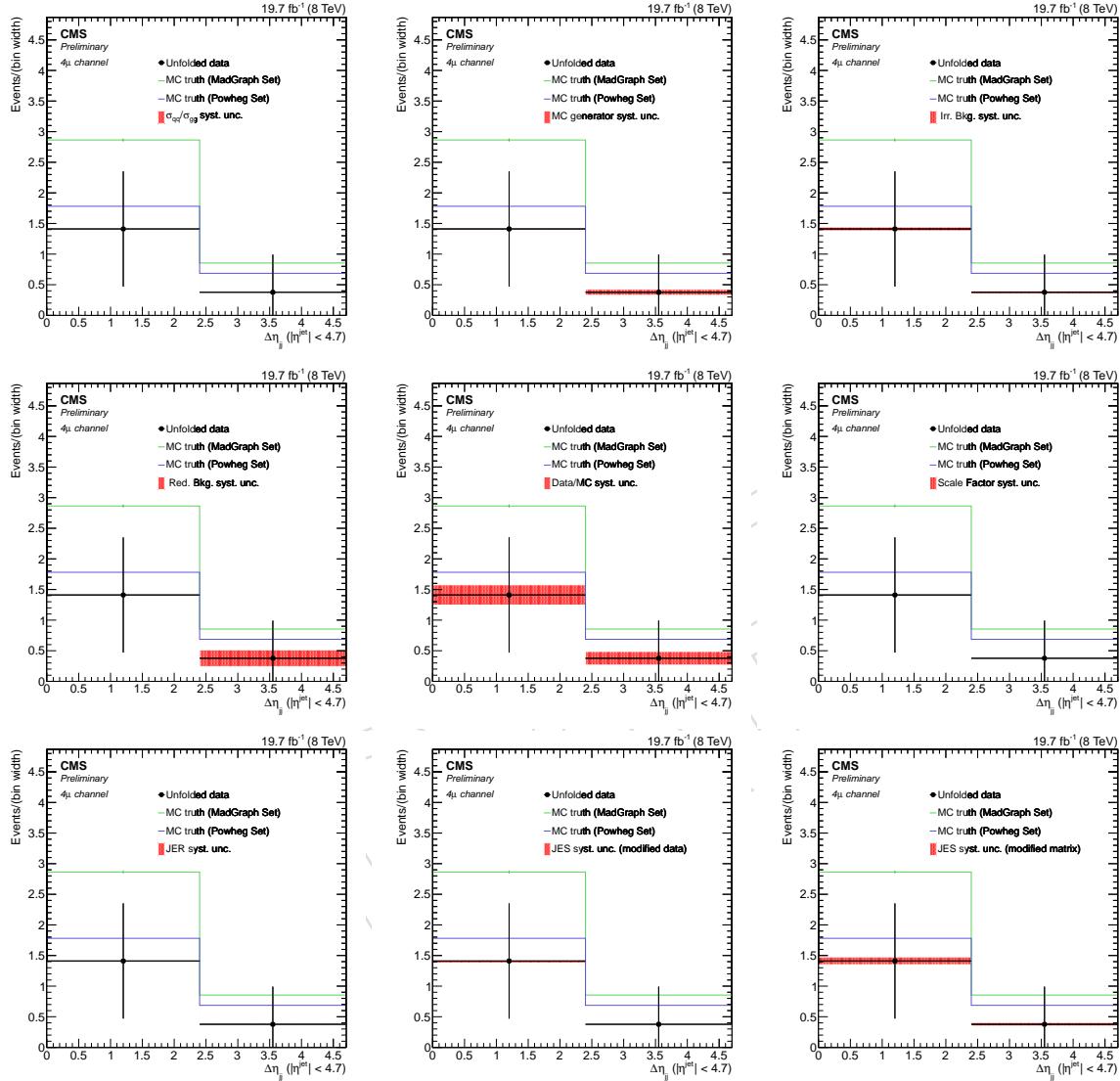


Figure 90: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\Delta\eta_{jj}$ (with $|\eta^{\text{jet}}| < 4.7$), for the 4μ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/true ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

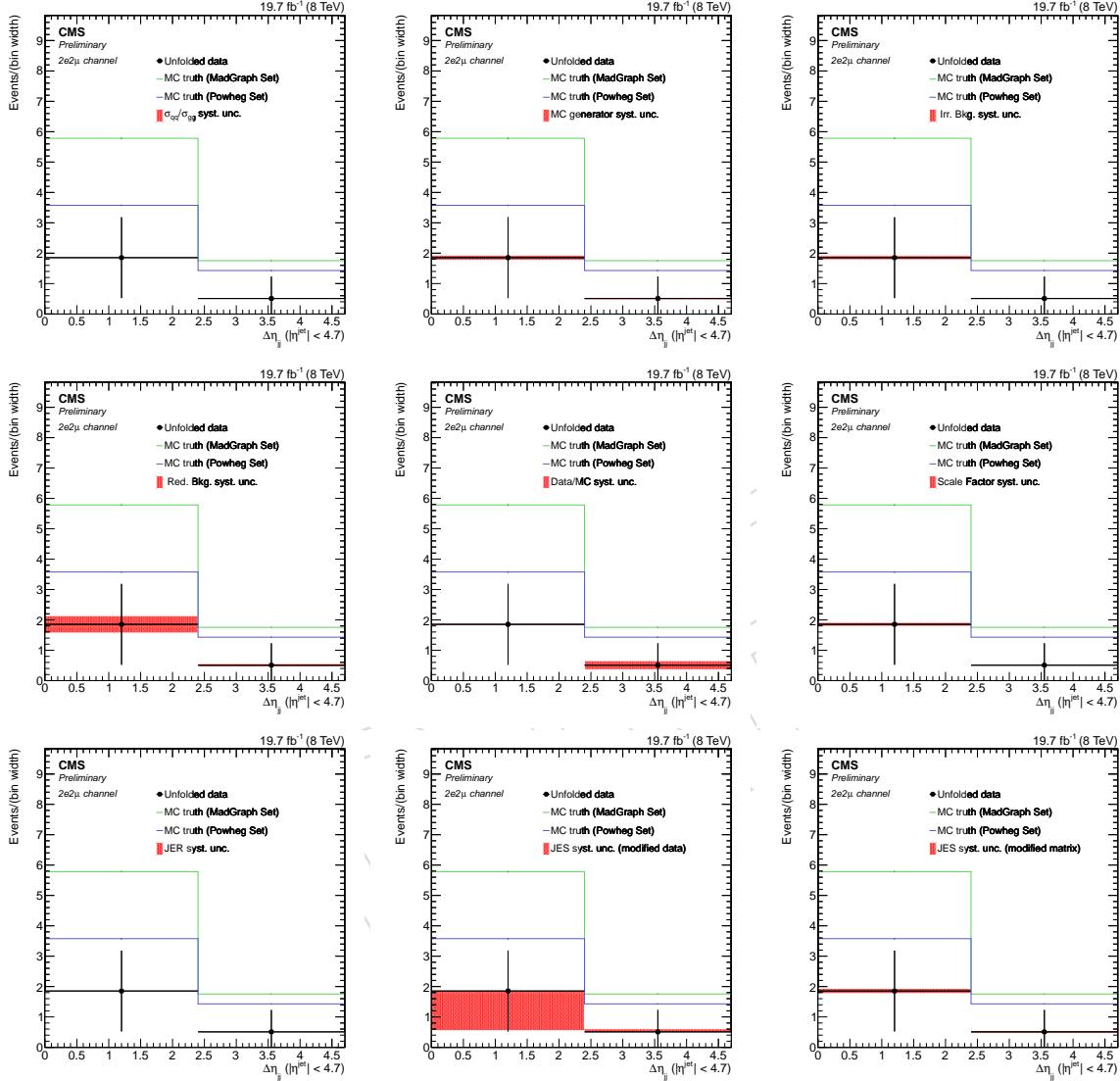


Figure 91: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\Delta\eta_{jj}$ (with $|\eta^{jet}| < 4.7$), for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

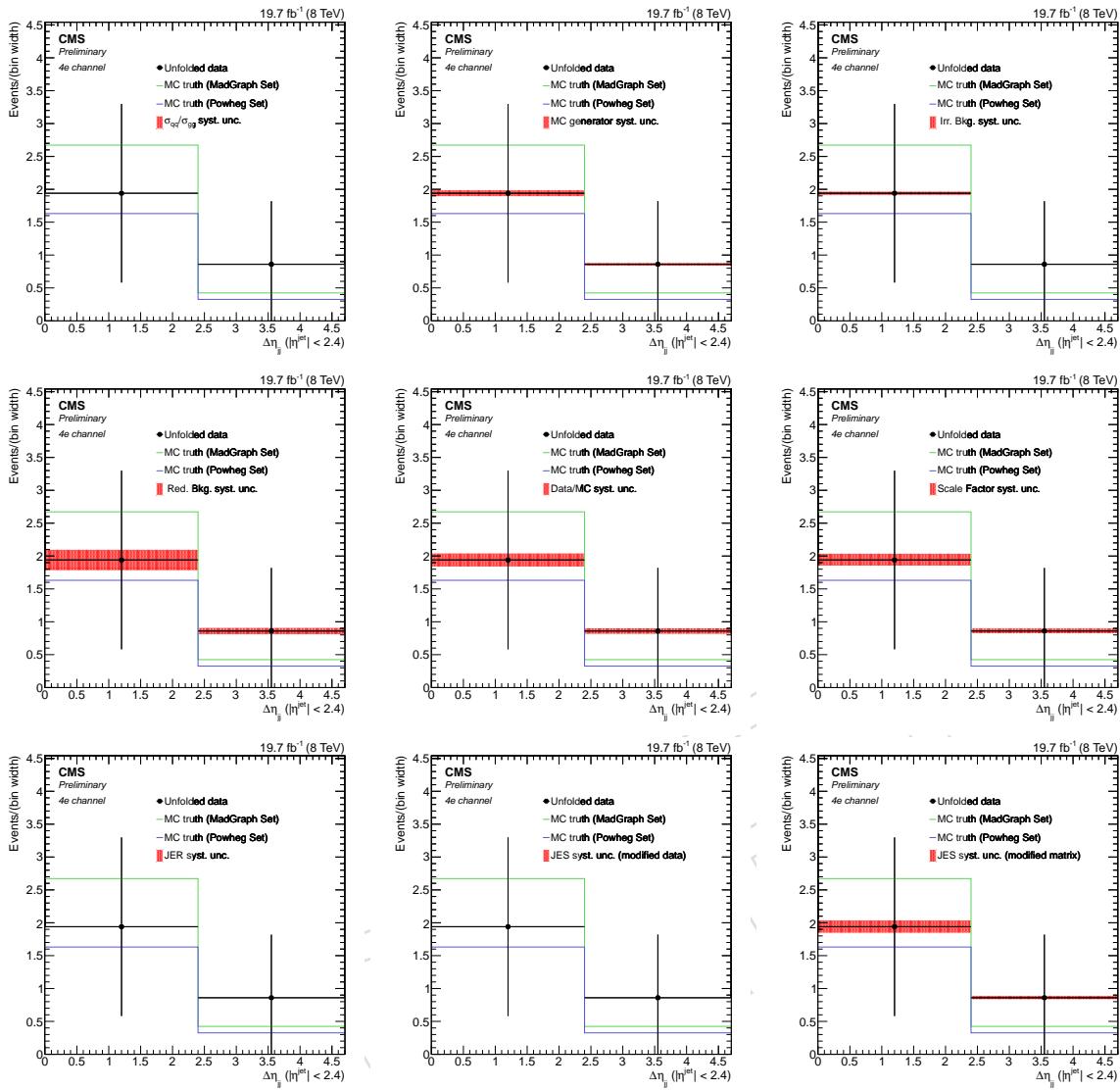
A Unfolding

Figure 92: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\Delta\eta_{jj}$ (with $|\eta^{\text{jet}}| < 2.4$), for the $4e$ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/true ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

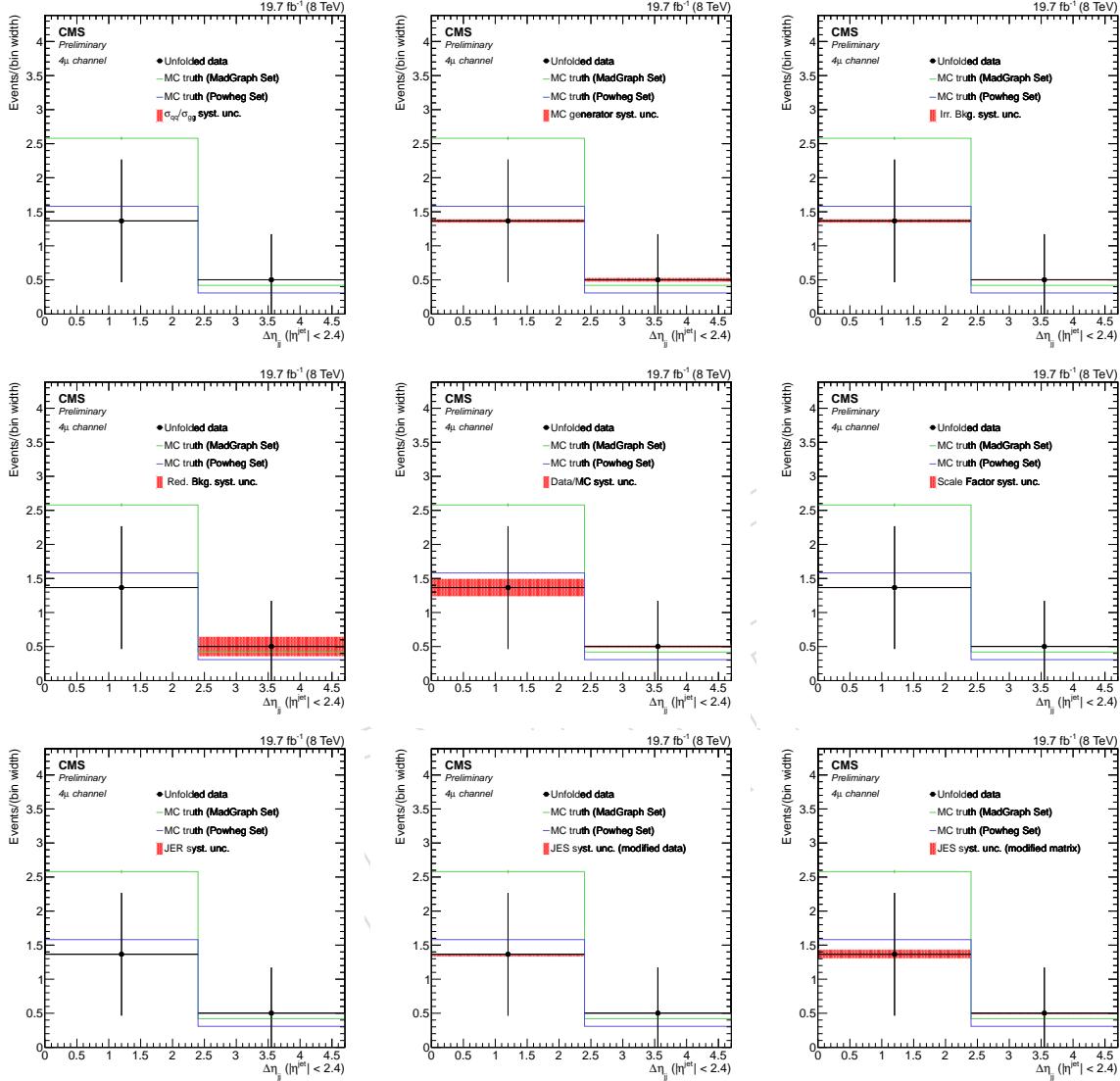


Figure 93: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\Delta\eta_{jj}$ (with $|\eta^{\text{jet}}| < 2.4$), for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

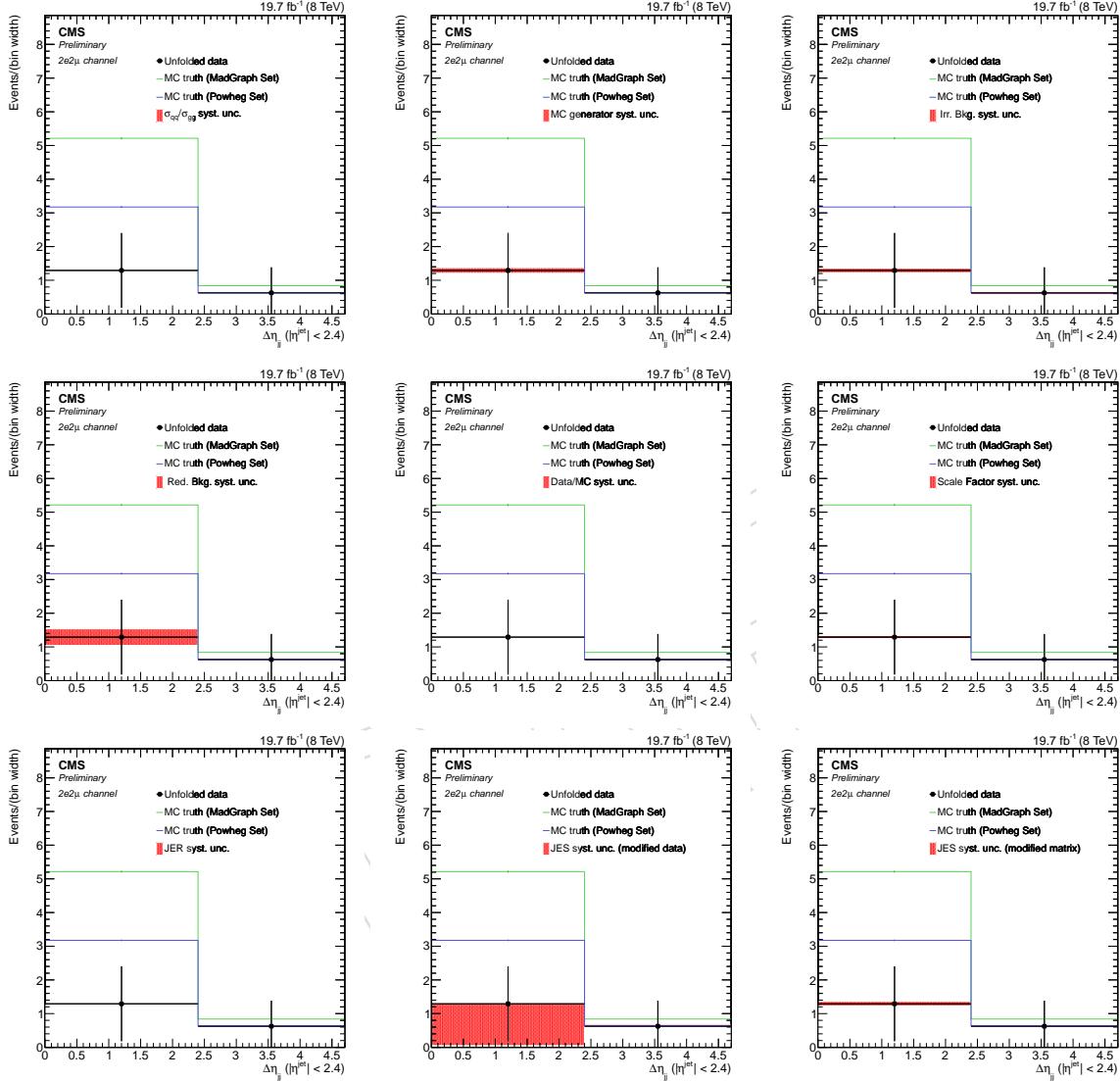


Figure 94: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\Delta\eta_{jj}$ (with $|\eta^{jet}| < 2.4$), for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq}/σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/true ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

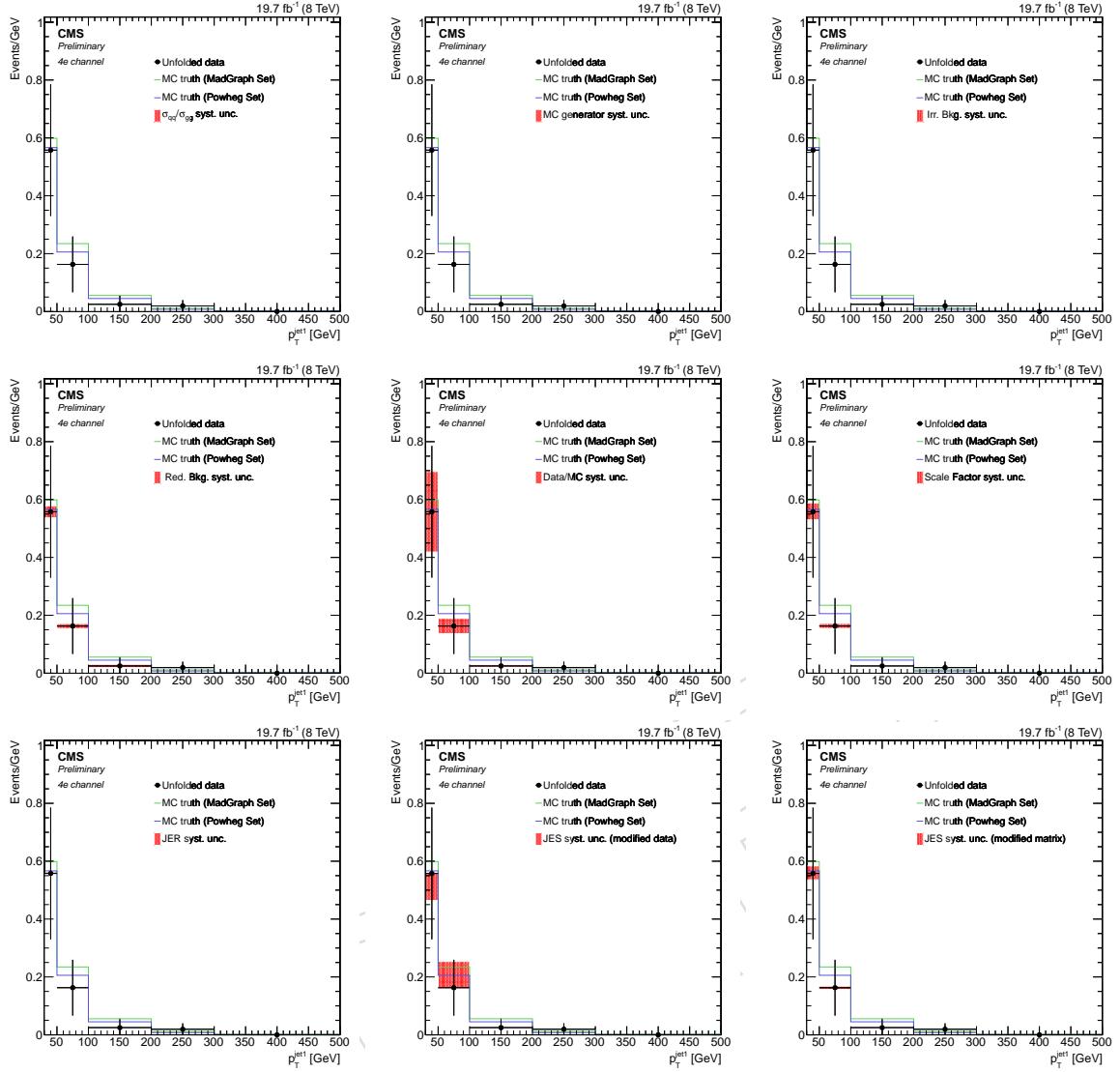


Figure 95: Effects of the different sources of systematic uncertainty on the unfolded distributions of p_T^{jet1} , for the 4e final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

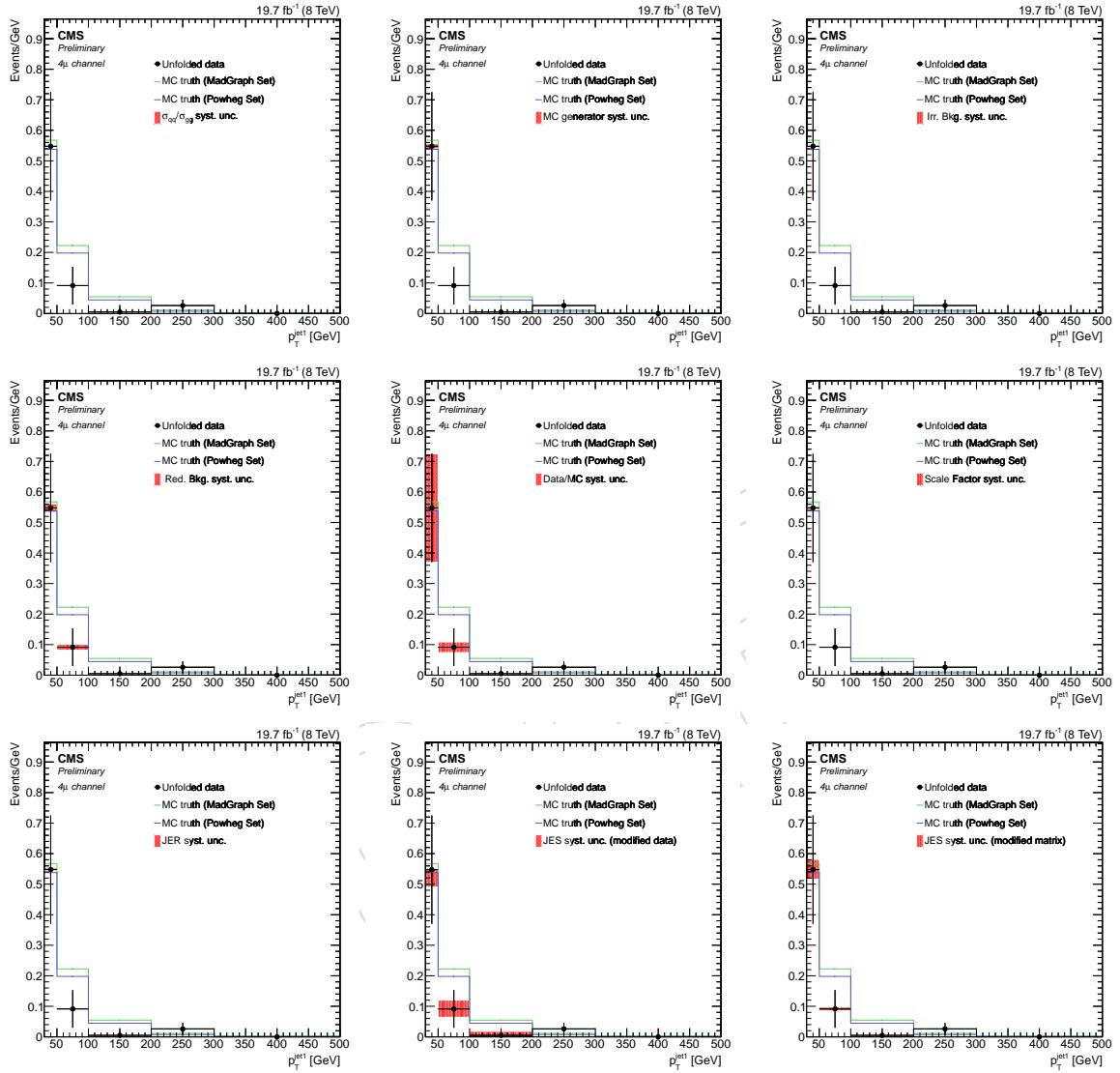


Figure 96: Effects of the different sources of systematic uncertainty on the unfolded distributions of p_T^{jet1} , for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

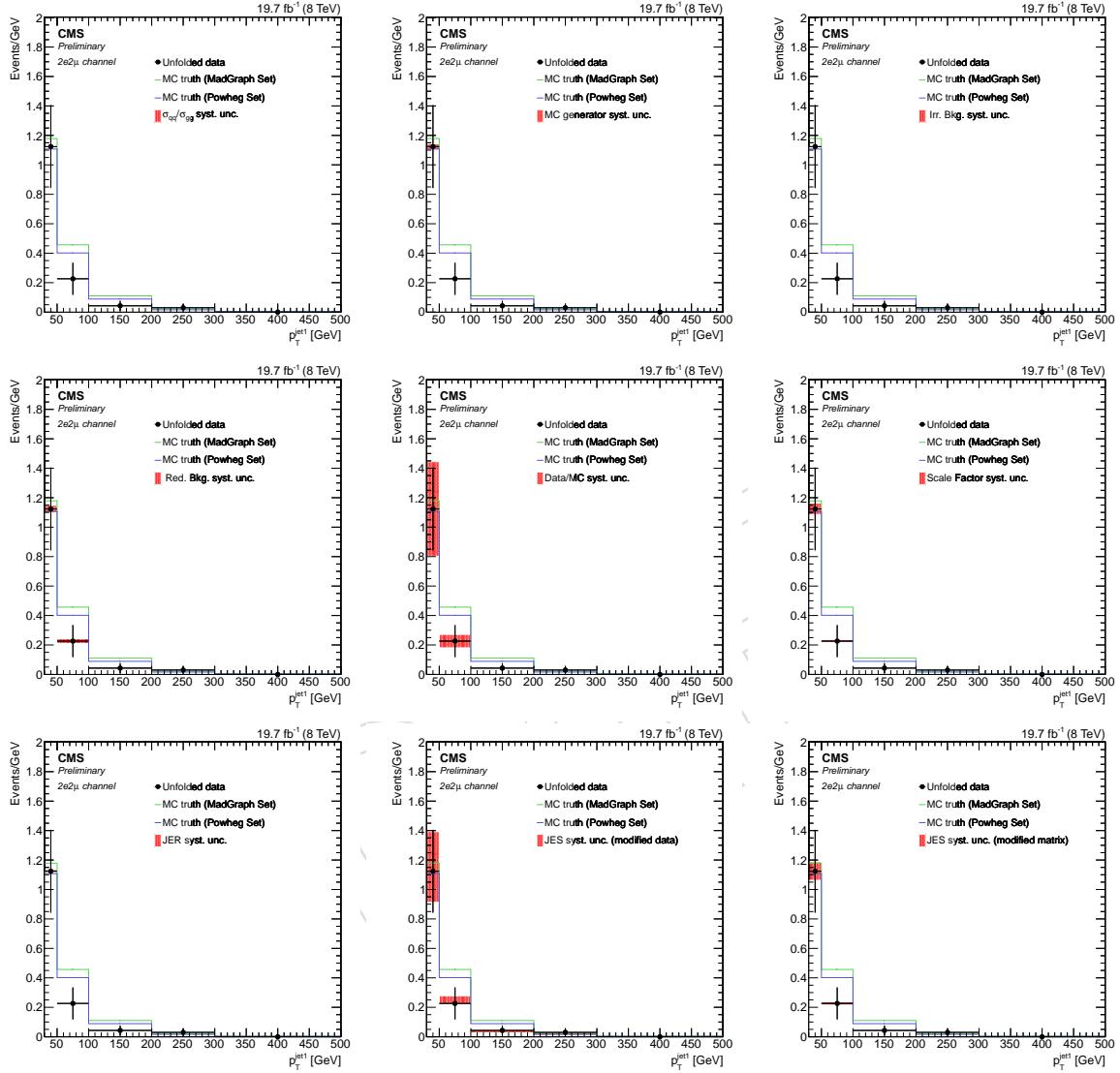


Figure 97: Effects of the different sources of systematic uncertainty on the unfolded distributions of p_T^{jet1} , for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

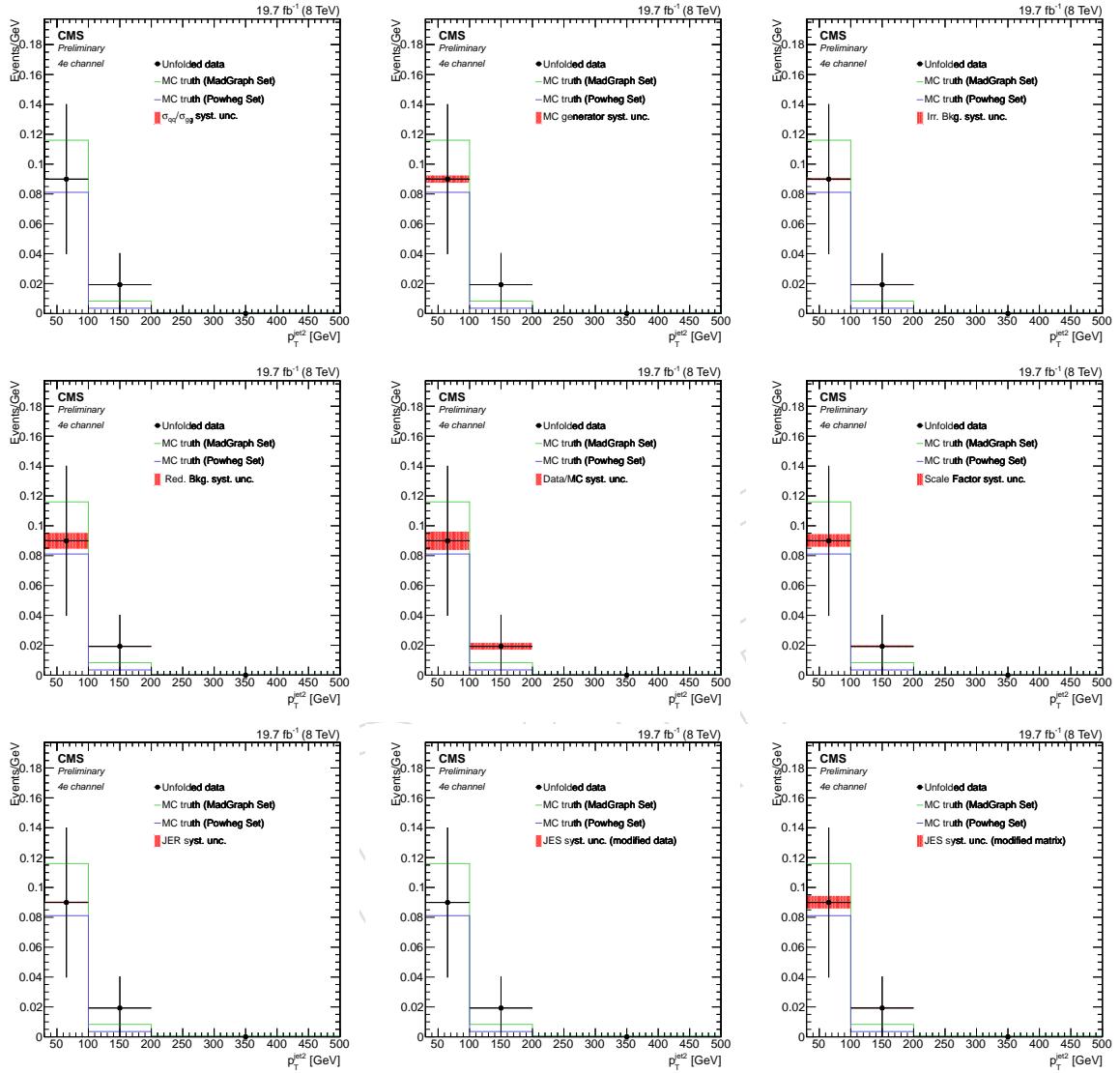


Figure 98: Effects of the different sources of systematic uncertainty on the unfolded distributions of p_T^{jet2} , for the $4e$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

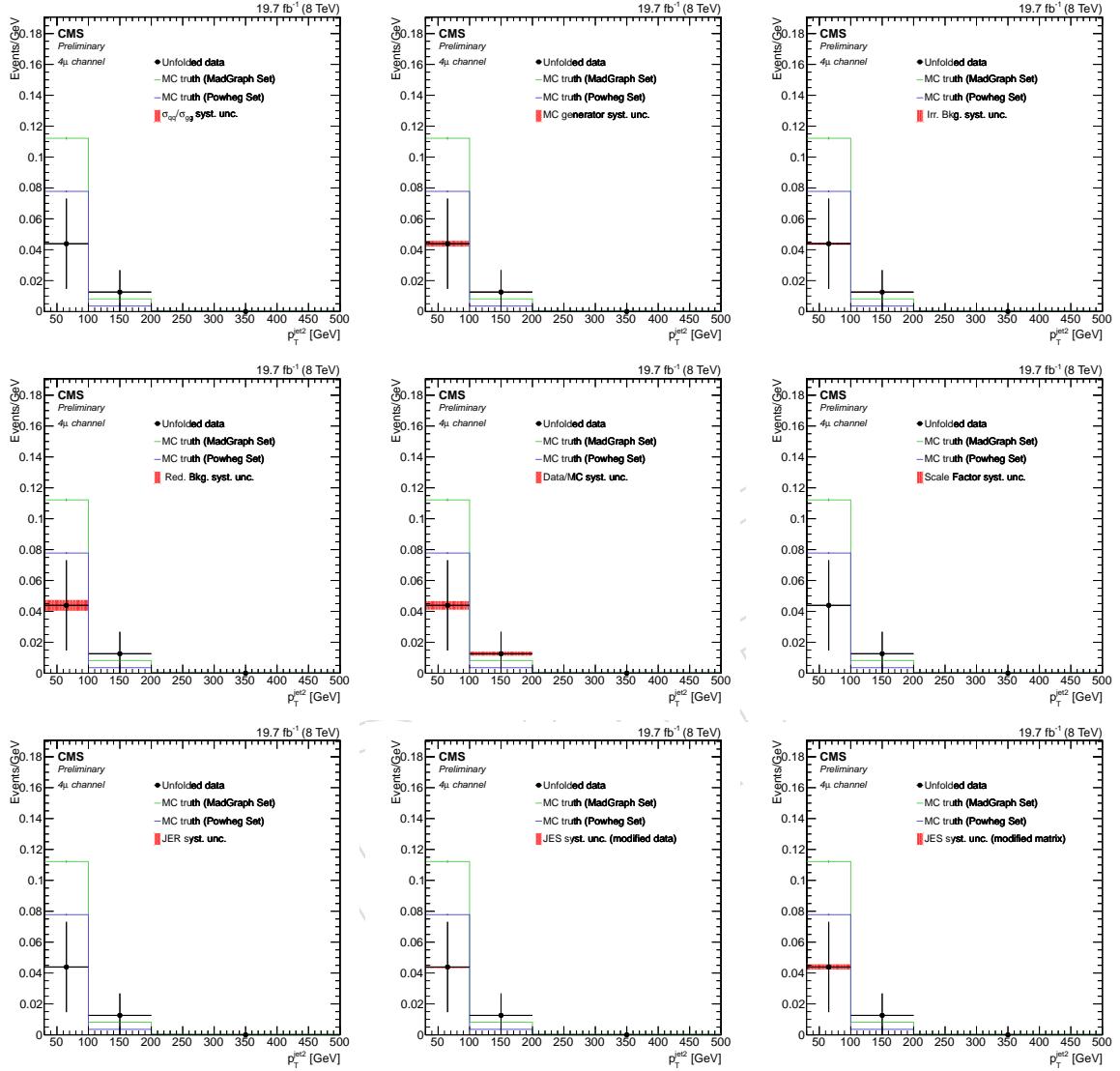


Figure 99: Effects of the different sources of systematic uncertainty on the unfolded distributions of $p_T^{\text{jet}2}$, for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

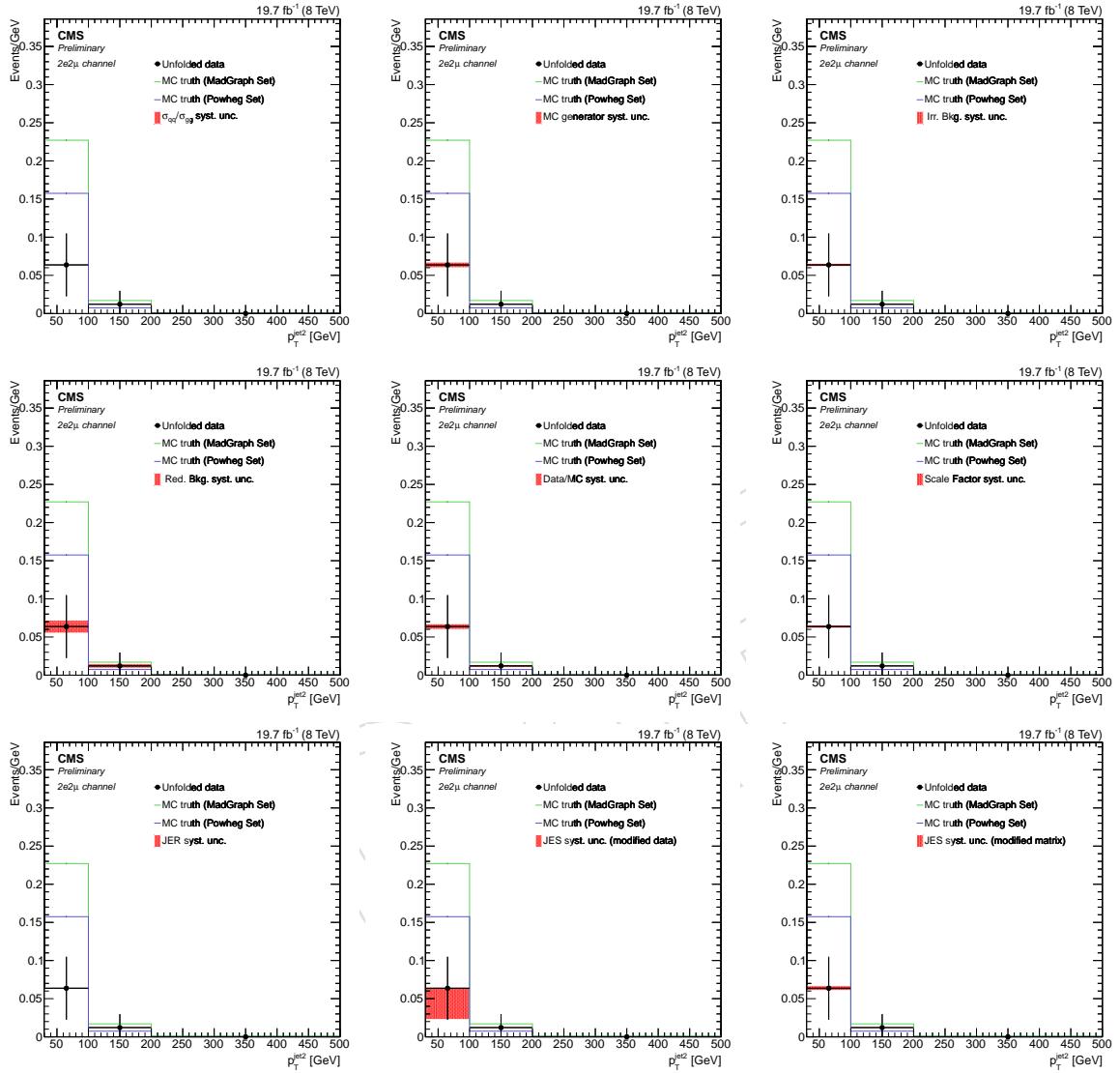


Figure 100: Effects of the different sources of systematic uncertainty on the unfolded distributions of p_T^{jet2} , for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/true ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

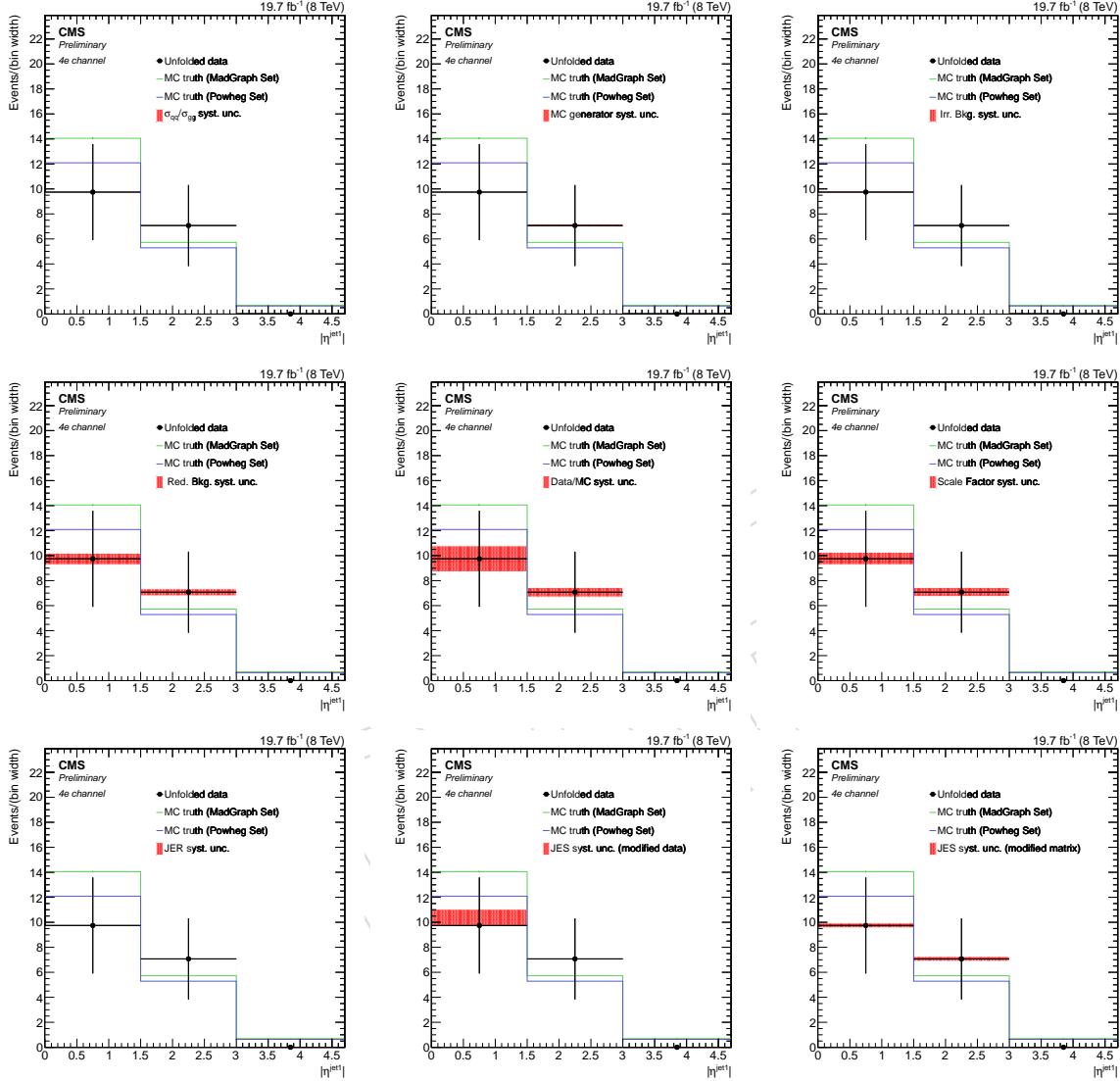


Figure 101: Effects of the different sources of systematic uncertainty on the unfolded distributions of η^{jet1} , for the $4e$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

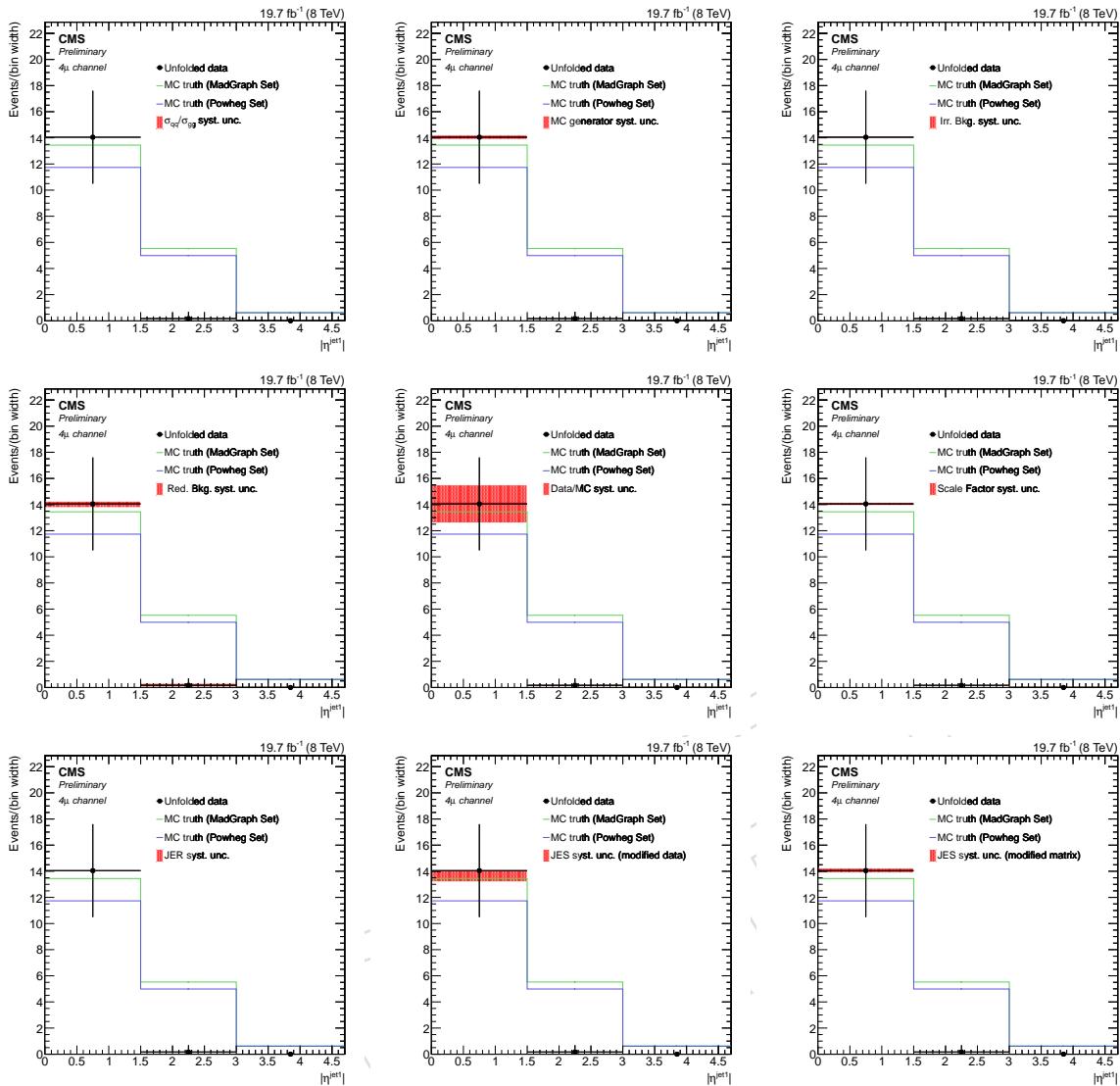


Figure 102: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\eta^{\text{jet}1}$, for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

497 References

- 498 [1] CMS Collaboration, "Measurement of the ZZ production cross section and search for
499 anomalous couplings in $2\ell 2\ell'$ final states in pp collisions at $\sqrt{s} = 7\text{TeV}$ ", *Journal of High*
500 *Energy Physics* **2013** (2013), no. 1, doi:10.1007/JHEP01(2013)063.
- 501 [2] CMS Collaboration, "Measurement of the production cross section and constraints on
502 anomalous triple gauge couplings in four-lepton final states at $\sqrt{s} = 8\text{TeV}$ ", *Physics*
503 *Letters B* **740** (2015) 250–272,
504 doi:<http://dx.doi.org/10.1016/j.physletb.2014.11.059>.

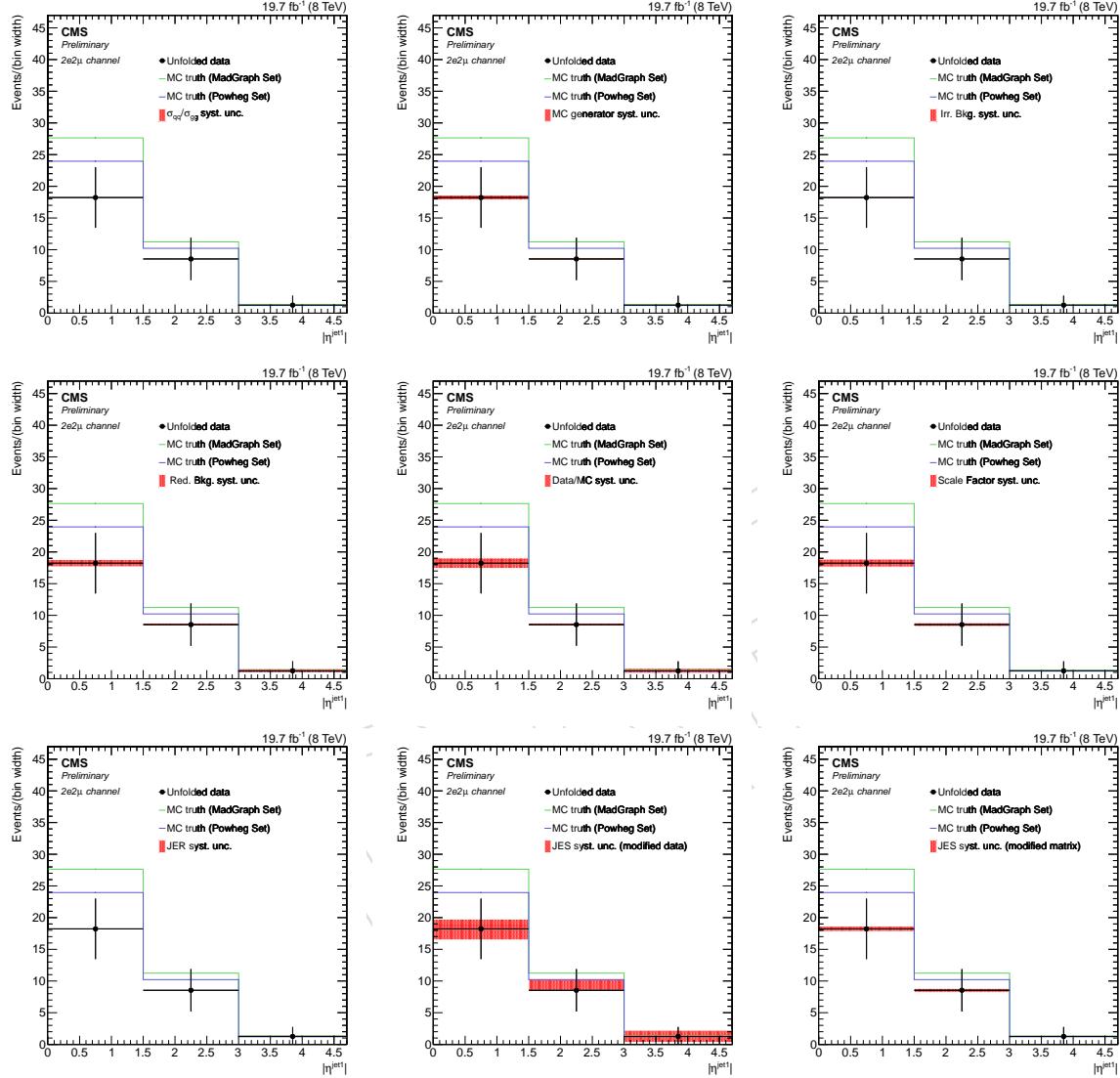


Figure 103: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\eta^{\text{jet}1}$, for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/true ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohwag sets of samples.

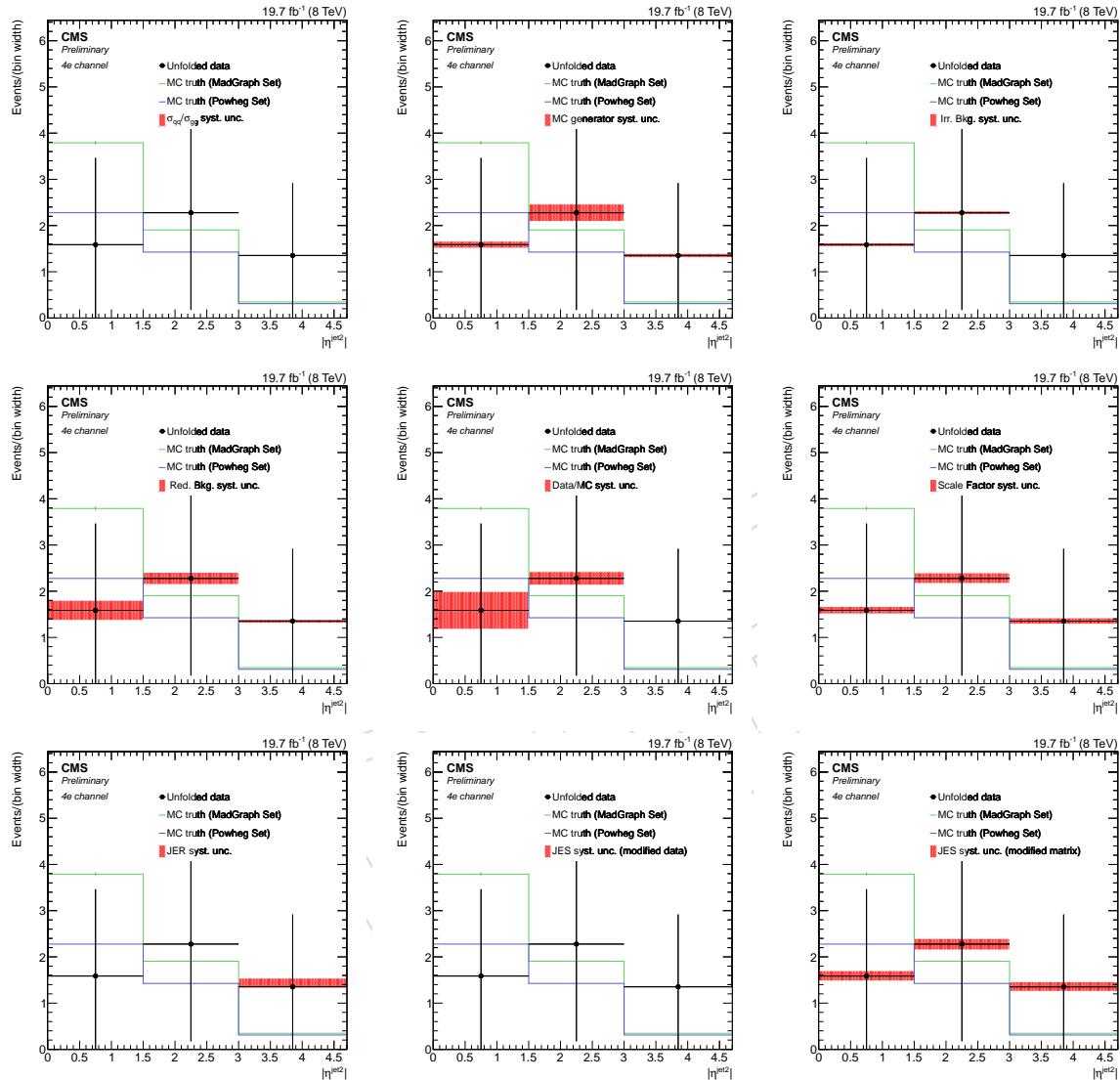


Figure 104: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\eta^{\text{jet}2}$, for the 4e final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

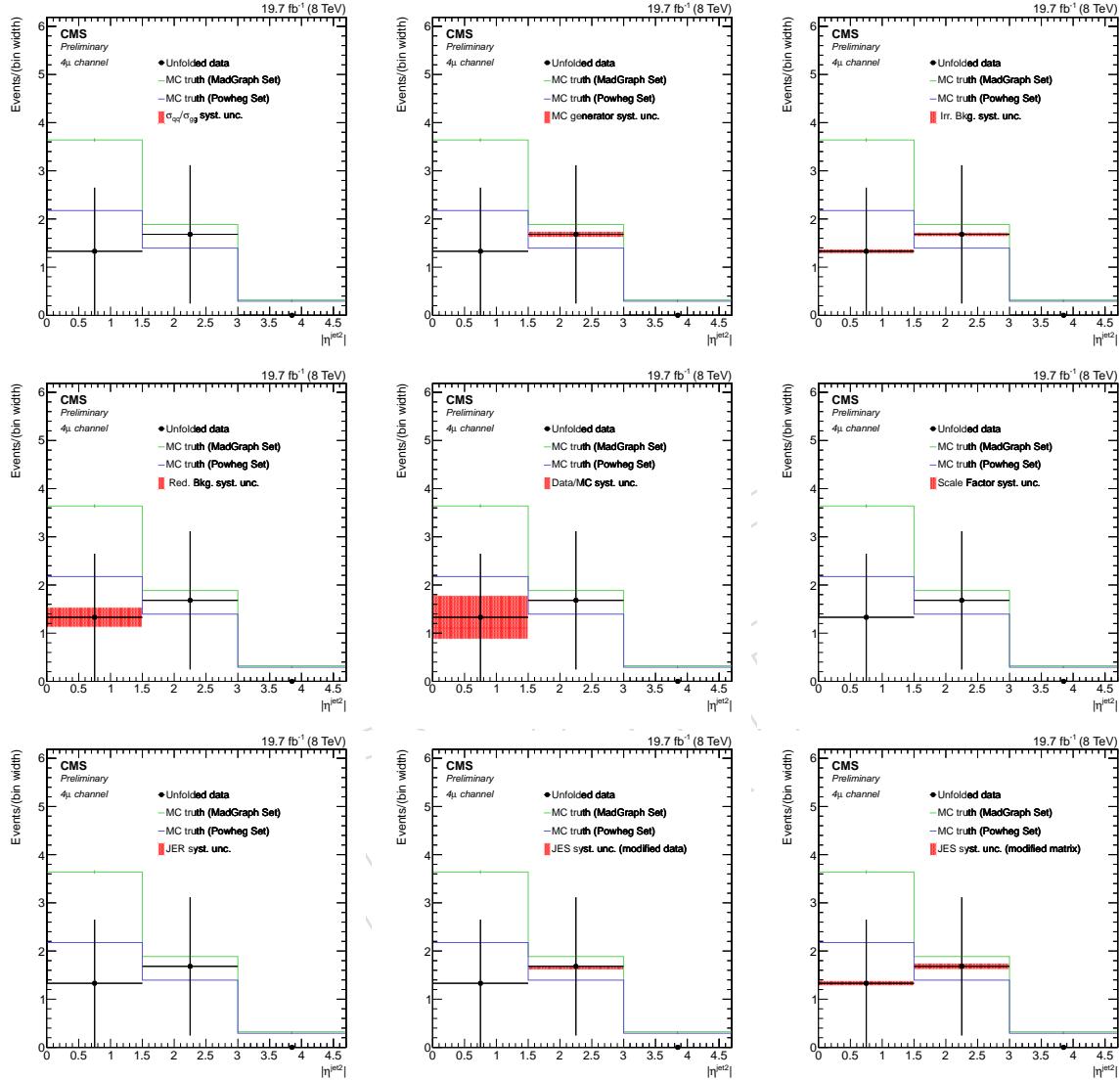


Figure 105: Effects of the different sources of systematic uncertainty on the unfolded distributions of $\eta^{\text{jet}2}$, for the 4μ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

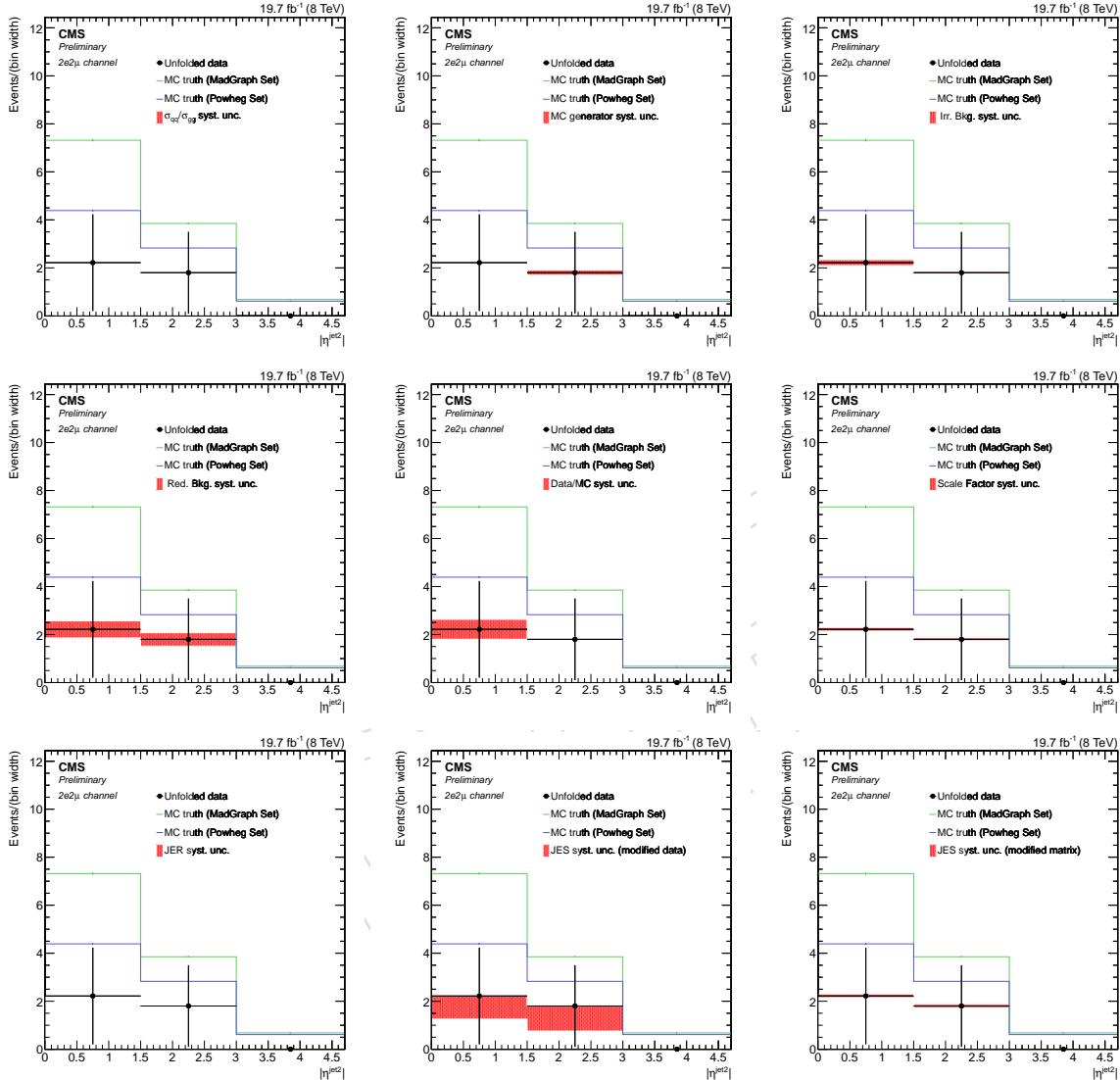


Figure 106: Effects of the different sources of systematic uncertainty on the unfolded distributions of η^{jet2} , for the $2e2\mu$ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/true ratio, lepton efficiency, JER, JES modifying data distribution, JES modifying the response matrix. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Pohweg sets of samples.

- [3] CDF Collaboration, "Measurement of ZZ Production in Leptonic Final States at \sqrt{s} of 1.96 TeV at CDF", *Phys. Rev. Lett.* **108** (Mar, 2012) 101801, doi:10.1103/PhysRevLett.108.101801.
- [4] D0 Collaboration, "Measurement of the ZZ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\text{TeV}$ ", *Phys. Rev. D* **84** (July, 2011) doi:10.1103/PhysRevD.84.011103.
- [5] CMS Collaboration, "Measurement of the properties of a Higgs boson in the four-lepton final state", *Phys. Rev. D* **89** (2014) 092007, doi:10.1103/PhysRevD.89.092007.
- [6] J. A. et al., "MadGraph 5: going beyond", *Journal of High Energy Physics* **2011** (June, 2011) doi:10.1007/JHEP06(2011)128.
- [7] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC", *Nuclear Physics B - Proceedings Supplements* **205-206** (2010) 10 – 15, doi:10.1016/j.nuclphysbps.2010.08.011. Loops and Legs in Quantum Field TheoryProceedings of the 10th DESY Workshop on Elementary Particle Theory.
- [8] S. F. et al., "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *Journal of High Energy Physics* **2007** (2007), no. 11, 070, doi:10.1088/1126-6708/2007/11/070.
- [9] S. A. et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *Journal of High Energy Physics* **2010** (2010), no. 6, doi:10.1007/JHEP06(2010)043.
- [10] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [11] A. B. et al., "PHANTOM: A Monte Carlo event generator for six parton final states at high energy colliders", *Computer Physics Communications* **180** (2009), no. 3, 401–417, doi:<http://dx.doi.org/10.1016/j.cpc.2008.10.005>.
- [12] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [13] L. L. et al., "Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions", *Physical Review D* **82** (2010) doi:<http://dx.doi.org/10.1103/PhysRevD.82.054021>.
- [14] L. L. et al., "New parton distributions for collider physics", *Physical Review D* **82** (2010) doi:<http://dx.doi.org/10.1103/PhysRevD.82.074024>.
- [15] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.
- [16] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", technical report, CERN, 2009.
- [17] CMS Collaboration, "Double Muon Trigger efficiency in 2012 data", technical report, CERN, 2014.
- [18] G. S. M. Cacciari and G. Soyez, "The anti- k_t jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008), no. 04, 063, doi:10.1088/1126-6708/2008/04/063.

- 545 [19] G. S. M. Cacciari and G. Soyez, “FastJet user manual”, *The European Physical Journal C* **72**
 546 (2012), no. 3, doi:10.1140/epjc/s10052-012-1896-2.
- 547 [20] J.-M. Group, “Jet Identification”,
- 548 [21] T. C. collaboration, “Determination of jet energy calibration and transverse momentum
 549 resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11, P11002.
- 550 [22] J.-M. Group, “Jet Energy Resolution”,
- 551 [23] M. B. et al., “The PDF4LHC Working Group Interim Recommendations”,
 552 arXiv:1101.0538.
- 553 [24] S. A. et al., “The PDF4LHC Working Group Interim Report”, arXiv:1101.0536.
- 554 [25] R. D. B. K. al., “Impact of heavy quark masses on parton distributions and LHC
 555 phenomenology”, *Nuclear Physics B* **849** (2011), no. 2, 296 – 363,
 556 doi:<http://dx.doi.org/10.1016/j.nuclphysb.2011.03.021>.
- 557 [26] K. Olive and P. D. Group, “Review of Particle Physics”, *Chinese Physics C* **38** (2014), no. 9,
 558 090001.
- 559 [27] T. Adye, “Unfolding algorithms and tests using RooUnfold”, arXiv:1105.1160v1.
- 560 [28] A. Höcker and V. Kartvelishvili, “SVD approach to data unfolding”, *Nuclear Instruments
 561 and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and
 562 Associated Equipment* **372** (1996), no. 3, 469–481,
 563 doi:[http://dx.doi.org/10.1016/0168-9002\(95\)01478-0](http://dx.doi.org/10.1016/0168-9002(95)01478-0).
- 564 [29] G. D’Agostini, “A multidimensional unfolding method based on Bayes’ theorem”,
 565 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers,
 566 Detectors and Associated Equipment* **362** (1995), no. 2-3, 487–498,
 567 doi:[http://dx.doi.org/10.1016/0168-9002\(95\)00274-X](http://dx.doi.org/10.1016/0168-9002(95)00274-X).
- 568 [30] CMS Collaboration, “Measurement of the W^+W_- cross section in pp collisions at $\sqrt{s} = 8$
 569 TeV and limits on anomalous gauge couplings”, technical report, CERN, 2015.
 570 arXiv:1507.03268.
- 571 [31] Atlas Collaboration, “Observation of a new particle in the search for the Standard Model
 572 Higgs boson with the ATLAS detector at the LHC”, *Physics Letters B* **716** (2012), no. 1, 1 –
 573 29, doi:<http://dx.doi.org/10.1016/j.physletb.2012.08.020>.
- 574 [32] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS
 575 experiment at the LHC”, *Physics Letters B* **716** (2012), no. 1, 30 – 61,
 576 doi:<http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
- 577 [33] M. Grazzini, S. Kallweit, and D. Rathlev, “ZZ production at the LHC: fiducial cross
 578 sections and distributions in NNLO QCD”, *Phys. Lett. B* **750** (2015) 407–410,
 579 doi:10.1016/j.physletb.2015.09.055.
- 580 [34] G. D. R. V. Candelise, F. Cossutti and C. L. Licata, “Associated production of Z boson and
 581 b-jets in pp collision at 8 TeV”, CMS Note 2013/165, CERN, 2014.
- 582 [35] T. C. Collaboration, “Measurement of the differential production cross section of Z bosons
 583 in association with jets in pp collisions at $\sqrt{s} = 8$ TeV”, CMS PAS 2013/007, CERN, 2014.