

Università degli Studi di Torino



SCUOLA DI DOTTORATO IN SCIENZE DELLA NATURA E
TECNOLOGIE INNOVATIVE

Indirizzo in Fisica ed Astrofisica

Tesi di Dottorato

**Measurement of the ZZ cross-section in the
fully leptonic decay channel in association
with jets, and studies on vector boson
scattering with the CMS Experiment**

Coordinatore

Prof. Mauro Gallio

Tutore

Prof. Nicola Amapane

Candidata

Linda Finco

Ciclo XXVIII

Date of dissertation

January 12th 2015

Candidate

Linda Finco

Tutor

Prof. Nicola Amapane

Università degli Studi di Torino

Degree Board

Prof. Cristiana Peroni

Università degli Studi di Torino

Prof. Marco Zanetti

Università degli Studi di Padova

Prof. Andrea Lavagno

Politecnico di Torino

Ph.D. Coordinator

Prof. Mauro Gallio

Università degli Studi di Torino

Contents

Introduction	vi
1 The Standard Model of Fundamental Interactions	1
1.1 The Standard Model of Elementary Particles	1
1.2 The Electroweak Theory	3
1.3 The Higgs Mechanism	5
1.3.1 Vector Boson Masses and Couplings	7
1.3.2 Fermion Masses and Couplings	8
1.4 Boson Interaction Vertices	9
1.5 Vector Boson Scattering	11
1.6 Anomalous Quartic Gauge Couplings	15
1.6.1 Higher Dimension Operators	15
1.6.2 Unitarity Bounds	16
1.7 Beyond the Standard Model	18
1.7.1 Dark Matter and Dark Energy	18
1.7.2 Matter-Antimatter Asymmetry	18
1.7.3 Neutrino Masses	19
1.7.4 The Hierarchy Problem	19
2 The CMS Experiment at LHC	21
2.1 The Large Hadron Collider	21
2.1.1 The LHC Design	21
2.1.2 LHC Operations during Run I	22
2.1.3 Run II and Future Perspectives	24
2.2 The Compact Muon Solenoid Experiment	25
2.2.1 Definition of Kinematic Variables	26
2.2.2 The Tracker	27
2.2.3 The Electromagnetic Calorimeter	28
2.2.4 The Hadron Calorimeter	28
2.2.5 The Magnet	30
2.2.6 The Muon System	31
2.3 The CMS Trigger System	32
2.3.1 The Level-1 Trigger	32
2.3.2 The High Level Trigger	33
2.4 The CMS Results in Run I	33

3	Study of Massive Vector Boson Pair Production	35
3.1	Massive Vector Boson Pair Production at the LHC	35
3.1.1	WW Production	36
3.1.2	WZ Production	36
3.1.3	ZZ Production	36
3.2	ZZ $\rightarrow 4\ell$ Processes	37
3.3	Leading and Next-to-Leading Order Cross-Sections	40
3.3.1	A Simple Example	41
3.3.2	Monte Carlo Generators	43
3.4	Current Status of Calculations	44
3.4.1	QCD and EW Corrections at NLO	44
3.4.2	Corrections at Higher Orders	50
4	Physics Objects Reconstruction	53
4.1	The Particle-Flow Algorithm	53
4.1.1	Track Reconstruction	53
4.1.2	Calorimeter Clustering Reconstruction	55
4.1.3	The Linking Algorithm	55
4.2	Leptons	55
4.2.1	Electron Reconstruction and Identification	56
4.2.2	Muon Reconstruction and Identification	56
4.2.3	Lepton Impact Parameter and Isolation	57
4.2.4	Leptons Efficiency	58
4.3	Photons	59
4.3.1	Photon Reconstruction and Identification	59
4.3.2	Photon Isolation	60
4.3.3	Final State Radiation	60
4.4	Jets	60
4.4.1	The Anti- k_T Algorithm	60
4.4.2	Jet Energy Corrections	61
4.4.3	Jet Energy Resolution	62
4.4.4	Jet Reconstruction	63
5	Analysis of ZZ Events	65
5.1	Data Sets	65
5.1.1	Experimental Data	65
5.1.2	Simulated Samples	66
5.2	Event Selection and Kinematics	68
5.2.1	Best Candidate Selection	68
5.2.2	Signal Efficiency and Acceptance	69
5.3	Background Estimation	75
5.3.1	Irreducible Background	75
5.3.2	Reducible Background	76
5.4	Inclusive ZZ Cross-Section Measurement	78
5.5	Systematic Uncertainties	83

6	The Differential ZZ Cross-Section Measurement	85
6.1	Data Unfolding	85
6.1.1	Unfolding Procedure	86
6.1.2	Bin-to-Bin Migration	86
6.1.3	Building the Response Matrix	91
6.1.4	Choice of the Unfolding Algorithm	91
6.1.5	Unfolding Validation Tests on Monte Carlo Samples	104
6.2	Unfolded Distributions	106
6.3	Differential Cross-Section Measurement	119
6.4	Systematic Uncertainties	136
6.5	Interpretation of the Results	139
7	Vector Boson Scattering and Anomalous Quartic Interactions	141
7.1	Signal Signature	141
7.2	Analyses at Run I	142
7.2.1	Z-Production in Association with Forward/Backward Jets . .	143
7.2.2	W-Production in Association with Forward/Backward Jets .	146
7.2.3	$\gamma\gamma$ -Production of W^+W^- pairs	146
7.2.4	Vector Boson Scattering in the $W^\pm W^\pm + 2$ jets Channel . . .	148
7.3	ZZ + 2 jets Channel	152
7.4	Future Perspectives	153
7.4.1	ZZ + 2 jets Channel	155
7.4.2	WZ + 2 jets Channel	156
7.4.3	$W^\pm W^\pm + 2$ jets Channel	157
7.4.4	$Z\gamma\gamma$ Channel	158
	Conclusions	159
	A Unfolding Validation Tests	161
	B Systematic Uncertainty Contributions	179
	Bibliography	191

Introduction

The Standard Model (SM) of Particle Physics is, up to now, one of the best tested physics theories of modern Physics. Developed in the early 1970s, it has successfully explained several experimental results and precisely predicted a wide variety of phenomena. This theory illustrates how the twelve elementary particles and three of the four fundamental forces are related to each other and predicts the existence of the Higgs scalar boson, which represents the key to the origin of particles mass.

The *Higgs hunting* has been a central purpose of the physics program at the Large Hadron Collider (LHC), the world's major particle accelerator. On the 4th of July 2012, the ATLAS and CMS Collaborations, two of the four main experiments at the LHC, announced the discovery of a new particle with mass around 125 GeV, which has been then confirmed to be consistent with the SM Higgs boson. This discovery represents the greatest result achieved by the LHC during its first period of data-taking and has led to the Nobel Prize in Physics in 2013, awarded for the particle's theoretical prediction.

Even if the discovery of the Higgs boson is an important step which confirms the validity of the Standard Model, it is not conclusive, because this theory, despite its effectiveness at describing the phenomena within its domain, is nevertheless incomplete. There are several questions without a satisfactory answer within the SM framework that still need to be resolved, and shading light on them is the central goal of the experiments at the LHC. After the discovery of the Higgs boson, one of the most urgent experimental tasks is to test if, according to the Standard Model, this particle is the only responsible for the electroweak symmetry breaking (EWSB) mechanism. If this is not the case, other hypotheses must be investigated. The Compact Muon Solenoid (CMS) plays a leading role in checking the Standard Model validity and probing new phenomena beyond it.

One of the most important processes to test the Standard Model description of electroweak and strong interactions is the study of the electroweak vector boson pair production. Any deviation from the expectations in kinematic distributions of vector boson pairs or their decay products could provide first evidence for effects from New Physics. The production cross-section of vector boson pairs in association with one or more jets is also very interesting: by comparing results with predictions of perturbative quantum chromodynamics, the precision of these corrections can be tested. Moreover, vector boson pair production processes enter as irreducible backgrounds in many Higgs boson analyses and New Physics searches and it is thus extremely important to measure them with high precision.

The work presented in this thesis concerns the study of ZZ vector boson pairs produced in association with jets. In its fully leptonic decay channel, ZZ has a cleaner signature with respect to the other massive vector boson pair production processes (such as W^+W^- and $W^\pm Z$) because, thanks to the four leptons in the final state, its kinematics is entirely reconstructible. Furthermore, $ZZ \rightarrow 4\ell$ is the only irreducible background of the Higgs boson decaying into four leptons ($H \rightarrow ZZ \rightarrow 4\ell$), the golden channel for Higgs boson searches. Moreover, ZZ +jets processes contribute to the background of several Higgs production mechanisms, such as in association with a vector boson, very rare phenomena which will become more important at higher luminosity, in the second data-taking period of LHC. For all these reasons, a proper evaluation of Z boson pair production processes in association with jets is mandatory.

This thesis is organized as follows. After a brief overview of the Standard Model, which describes the theoretical framework of this analysis, in Chapter 1 the attention is focused on the EWSB mechanism and the vector boson scattering, the key process to probe its effectiveness. In Chapter 2, after an introduction to the LHC and its operation during the Run I, the CMS detector is described. The overall structure of CMS, consisting of several cylindrical layers closed at both ends by detector disks, allows an excellent lepton reconstruction and particle identification to be achieved. Thanks to these features, in its first operation period CMS has obtained several important results.

In order to cover the broad scope of SM measurements and search strategies performed at LHC, precise predictions are needed for a wide variety of processes. In Chapter 3, the state-of-the-art calculations on vector boson pair production is reported, to show the current status of theoretical expectations. Results obtained by both ATLAS and CMS Collaborations are also summarized and compared to the predictions.

The measurement of ZZ vector boson pairs is based on the identification and reconstruction of leptons (electrons and muons), their decay products. Moreover, in order to measure the dependency to the jet multiplicity, the definition of jets is also a key element of this analysis. The description of the features of these physics objects is reported in Chapter 4. The selection of signal ZZ events decaying in the leptonic mode $ZZ \rightarrow \ell\ell\ell'\ell' (\ell, \ell' = e, \mu)$ is then presented, by analyzing the data sample collected during the Run I at a center-of-mass energy of 8 TeV, and the procedure applied to extract the inclusive cross-section is illustrated (Chapter 5). Later, the measurement is performed differentially, as a function of several variables related to jets (Chapter 6). Distributions are first corrected for the finite resolution and limited acceptance of the detector through an unfolding procedure, and the cross-section is then computed.

Finally, the study of ZZ + jets event production is exploited in order to perform a preliminary analysis on vector boson scattering. This process has indeed a very similar final state compared to the ZZ + jets production, being characterized by a pair of Z bosons created in association with two high- p_T jets. In Chapter 7, the main features of vector boson scattering processes at LHC are described and the ZZ + 2 jets channel is explored, projecting my work towards future searches.

Chapter 1

The Standard Model of Fundamental Interactions

To our knowledge, everything in the universe is made from a few basic building blocks called fundamental particles, governed by four fundamental forces: the gravitational, the electromagnetic, the weak and the strong force. The Standard Model (SM) describes how these particles and three of the forces (the gravity is not yet included) interact to each other. This important theory was developed in the early 1970s and predicts a wide variety of phenomena, which has been almost completely verified through many experiments. Thanks to that, the Standard Model of elementary particles is a very successful and well-tested physics theory.

In this Chapter the theory of the Standard Model is briefly reviewed: after an introduction of the theoretical framework and on the Electroweak Theory, the Higgs Mechanism is described. The attention is then focused on massive vector bosons, since their pair production is the main topic of this thesis, in order to illustrate how they interact to each other and with the Higgs boson. Among the vector boson interactions, the scattering process of two massive vector bosons is particularly interesting, being a crucial test of the SM validity, and it is thus analyzed. Finally, the main still unsolved questions of the SM are outlined, in order to provide a complete picture of the theoretical framework.

1.1 The Standard Model of Elementary Particles

The SM describes the matter as composed by twelve elementary particles, the *fermions*, all having half-integer spin. Fermions can be divided into two main groups, *leptons* and *quarks*, organized in three families, together with their anti-particles (Table 1.1). Leptons can just interact by electroweak forces, while quarks are subject to both strong and electroweak interactions. Moreover, quarks do not exist as free states, but only as constituents of *hadrons*, a class of particles such as protons and neutrons.

In the SM, the interactions between elementary particles are mediated by *bosons*, integer-spin particles which are carriers of the fundamental forces. The main characteristics of bosons and of the corresponding interactions are summarized in Table

1.2. The gravitational interaction is not taken into account, as it is not relevant at the typical mass and distance scales of particle physics.

Table 1.1: Classification of the three families of fundamental fermions.

Fermions	1 st fam.	2 nd fam.	3 rd fam.	Charge	Interactions
Quarks	u	c	t	$+\frac{2}{3}$	All
	d	s	b	$-\frac{1}{3}$	
Leptons	e	μ	τ	-1	Weak, E.M.
	ν_e	ν_μ	ν_τ	0	Weak

Table 1.2: Properties of the three fundamental forces (gravitational interaction is not taken into account).

	Electromagnetic	Weak	Strong
Quantum	Photon (γ)	$W^\pm, Z$	Gluons
Mass [GeV/c^2]	0	80, 90	0
Coupling constant	$\alpha(Q^2 = 0) \approx \frac{1}{137}$	$\frac{G_F}{(\hbar c)^3} \approx 1.2 \cdot 10^{-5} \text{ GeV}^{-2}$	$\alpha_s(m_Z) \approx 0.1$
Range [cm]	∞	10^{-16}	10^{-13}

This phenomenology arises from a mathematical formalism according to which the SM is a perturbatively renormalizable Quantum Field Theory (QFT) based on the *local gauge symmetry* of its Lagrangian. By the Noether's theorem, a conservation law corresponds to each of these local invariances, explaining why they are so important. Therefore the SM is a local gauge quantum field theory describing three of the four fundamental interactions: electromagnetic, weak and strong interaction. It is based on the symmetry group

$$SU(3)_C \otimes SU(2)_I \otimes U(1)_Y,$$

the direct product of $SU(3)_C$, the color symmetry group upon which Quantum Chromo Dynamics (QCD) is built, the gauge groups of weak isospin, $SU(2)_I$, and hypercharge, $U(1)_Y$. Electromagnetic and weak interactions are unified in the *electroweak* gauge group $SU(2)_I \otimes U(1)_Y$, upon which the Glashow-Weinberg-Salam (GWS) model is built.

Despite this symmetry predicts with precision and accuracy the phenomenology of particle interactions, it is broken by the mass terms of the Lagrangian. A necessary ingredient of the SM is therefore a *symmetry breaking* mechanism that allows to introduce the mass terms in a local gauge invariant Lagrangian.

1.2 The Electroweak Theory

From a historical point of view, the starting point for the study of electroweak interactions is Fermi's theory of muon decay [1], which is based on an effective four-fermion Lagrangian:

$$\mathcal{L} = -\frac{4G_F}{\sqrt{2}} \bar{\nu}_\mu \gamma^\alpha \frac{1-\gamma_5}{2} \mu \bar{e} \gamma_\alpha \frac{1-\gamma_5}{2} \nu_e, \quad (1.1)$$

where G_F is the Fermi coupling constant reported in Table 1.2. Equation 1.1 represents a *point-like* interaction, with only one vertex and without any intermediate boson exchanged. It is usually referred to as $V - A$ interaction, being formed by a vectorial and an axial component. The term $\frac{1}{2}(1 - \gamma_5)$ that appears in it is the negative helicity projector. Only the negative helicity (*left-handed*) component of fermions takes part to this interaction, corresponding to the projection of spin opposite to particle's motion.

Fermi's Lagrangian is not renormalizable and it results in a non-unitary scattering matrix. Both problems of renormalizability and unitarity are overcome requiring the weak interaction Lagrangian to be invariant under local transformations generated by the elements of a Lie group (*gauge transformations*). The resulting Lagrangian must reduce to Equation 1.1 in the low energy limit.

A gauge theory for weak interactions is conceived as an extension of the theory of electromagnetic interaction, the Quantum Electro-Dynamics (QED), which is based on the gauge group $U(1)_{EM}$, associated to the conserved quantum number Q (electric charge). In this case, the condition of local invariance under the $U(1)_{EM}$ group leads to the existence of a massless vector boson, the *photon*.

A theory reproducing both the electromagnetic and weak interaction phenomenology is achieved by extending the gauge symmetry to the group $SU(2)_I \otimes U(1)_Y$, following the GWS model. In this sense, the weak and electromagnetic interactions are said to be unified. The generators of $SU(2)_I$ is the weak isospin operator and the generator of $U(1)_Y$ is the weak hypercharge Y operator. The corresponding quantum numbers satisfy the Gell-Mann–Nishijima formula

$$Q = I_3 + \frac{Y}{2},$$

where I_3 is the third component of the weak isospin. Fermions can be divided in doublets of negative-helicity (*left-handed*) particles and singlets of positive-helicity (*right-handed*) particles, as follows:

$$L_L = \begin{pmatrix} \nu_{\ell,L} \\ \ell_L \end{pmatrix}, \ell_R, Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, u_R, d_R, \quad (1.2)$$

where $\ell = e, \mu, \tau$, $u = u, c, t$ and $d = d, s, b$. Neutrinos have no right component, as their mass is taken as null. In Table 1.3, I_3 , Y and Q quantum numbers of all fermions are reported. As well as for QED, the requirement of local gauge invariance with respect to the $SU(2)_I \otimes U(1)_Y$ group introduces now four massless vector fields (gauge fields), $W_\mu^{1,2,3}$ and B_μ , which couple to fermions with two

Table 1.3: Isospin (I_3), hypercharge (Y) and electric charge (Q) of all fermions.

	I_3	Y	Q
$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$
u_R, d_R	$0, 0$	$\frac{4}{3}, -\frac{2}{3}$	$+\frac{2}{3}, -\frac{1}{3}$
$\begin{pmatrix} \nu_{\ell,L} \\ \ell_L \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$
ℓ_R	0	-2	-1

different coupling constants, g and g' . Note that B_μ does not represent the photon field, because it arises from the $U(1)_Y$ group of hypercharge, instead of $U(1)_{EM}$ group of electric charge. The gauge-invariant Lagrangian for fermion fields can be written as follows:

$$\mathcal{L} = \bar{\Psi}_L \gamma^\mu \left(i\partial_\mu + g t_a W_\mu^a - \frac{1}{2} g' Y B_\mu \right) \Psi_L + \bar{\Psi}_R \gamma^\mu \left(i\partial_\mu - \frac{1}{2} g' Y B_\mu \right) \Psi_R, \quad (1.3)$$

where

$$\Psi_L = \begin{pmatrix} \Psi_L^1 \\ \Psi_L^2 \end{pmatrix}$$

and where Ψ_L and Ψ_R are summed over all the possibilities in Equation 1.2. $W_\mu^{1,2,3}$ and B_μ do not represent physical fields, which are given instead by linear combinations of the four mentioned fields. The charged bosons W^+ and W^- correspond to

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp i W_\mu^2 \right), \quad (1.4)$$

while the neutral bosons γ and Z correspond to

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \quad (1.5)$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W, \quad (1.6)$$

obtained by mixing the neutral fields W_μ^3 and B_μ with a rotation defined by the Weinberg angle θ_W . In terms of the fields reported in Equations 1.4 to 1.6, the interaction term between gauge fields and fermions, taken from the Lagrangian in Equation 1.3, becomes

$$\mathcal{L}_{int} = \frac{1}{2\sqrt{2}} g (J_\alpha^+ W^{(+)\alpha}) + \frac{1}{2} \sqrt{g'^2 + g^2} J_\alpha^Z Z^\alpha - e J_\alpha^{EM} A^\alpha, \quad (1.7)$$

where J^{EM} is the electromagnetic current connected to the photon field, while J^+ , J^- and J^Z are the three weak isospin currents. It is found that

$$J_\alpha^Z = J_\alpha^3 - 2\sin^2\theta_W \cdot J_\alpha^{EM}.$$

A^α can then be identified with the photon field and, requiring the coupling terms to be equal, one obtains

$$g \sin\theta_W = g' \cos\theta_W = e, \quad (1.8)$$

which represents the electroweak unification (e is the electron charge). The GWS model thus predicts the existence of two charged gauge fields, which only couple to left-handed fermions, and two neutral gauge fields, which interact with both left- and right-handed components.

1.3 The Higgs Mechanism

In order to correctly reproduce the phenomenology of weak interactions, both fermion and gauge boson fields must acquire mass, in agreement with experimental results. Up to this point, however, all particles are considered massless: in the electroweak Lagrangian a mass term for the gauge bosons would violate gauge invariance, which is needed to ensure the renormalizability of the theory. Explicit mass terms for fermions, instead, would not violate gauge invariance, but in the GWS model the Lagrangian is also required to preserve the invariance under chirality transformations, and this is achieved only with massless fermions. Masses are thus introduced with the *Higgs mechanism* [2], which allows fermions and $W^\pm$, Z bosons to be massive, while keeping the photon massless. Such mechanism is accomplished by means of a doublet of complex scalar fields,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix}, \quad (1.9)$$

which is introduced in the electroweak Lagrangian within the term

$$\mathcal{L}_{EWSB} = (D^\mu \phi)^\dagger (D_\mu \phi) + V(\phi^\dagger \phi), \quad (1.10)$$

where $D_\mu = \partial_\mu - ig t_a W_\mu^a + \frac{i}{2} g' Y B_\mu$ is the covariant derivative. The Lagrangian in Equation 1.10 is invariant under $SU(2)_I \otimes U(1)_Y$ transformations, since the kinetic part is written in terms of covariant derivatives and the potential V only depends on the product $\phi^\dagger \phi$. The ϕ field is characterized by the quantum numbers reported in Table 1.4 Writing the potential term as follows (see also Figure 1.1 for a graphical representation)

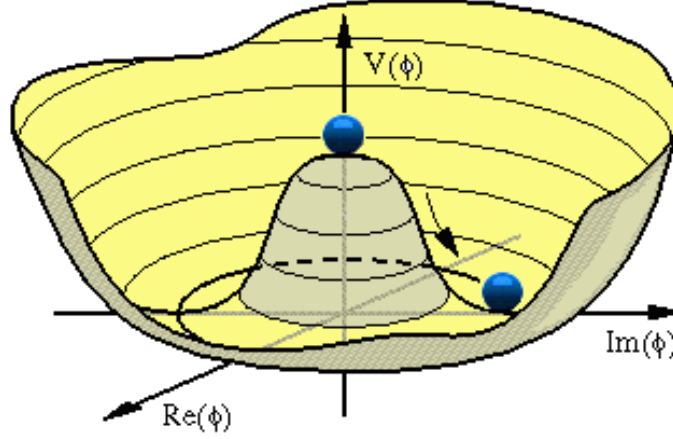
$$V(\phi^\dagger \phi) = -\mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2, \quad (1.11)$$

with $\mu^2 < 0$ and $\lambda > 0$, it results to have a minimum for

$$\phi^\dagger \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \quad (1.12)$$

Table 1.4: Isospin (I_3), hypercharge (Y) and electric charge (Q) of the Higgs boson.

	I_3	Y	Q
$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

**Figure 1.1:** Shape of the Higgs potential.

This minimum is not found for a single value of ϕ , but for a manifold of non-zero values. The choice of (ϕ^+, ϕ^0) corresponding to the ground state, i.e. the lowest energy state or *vacuum*, is arbitrary, and the chosen point is not invariant under rotations in the (ϕ^+, ϕ^0) plane: this is referred to as *spontaneous symmetry breaking*. It is therefore chosen arbitrarily a minimum along the ϕ^0 axis:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v^2 = -\frac{\mu^2}{\lambda}. \quad (1.13)$$

The ϕ field can thus be rewritten in a generic gauge, in terms of its vacuum expectation value:

$$\phi = \frac{1}{\sqrt{2}} e^{i\phi^a t_a} \begin{pmatrix} 0 \\ H + v \end{pmatrix}, \quad a = 1, 2, 3,$$

where the three fields ϕ^a and the fourth $\phi^4 = H + v$ are called *Goldstone fields*. Being scalar and massless, they introduce four new degrees of freedom, in addition to the six degrees due to the transverse polarizations of the massless vector bosons $W^\pm$ and Z . The unitary gauge is fixed by the transformation

$$\phi' = e^{-\frac{i}{v}\phi^a t_a} \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ H + v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \phi^4 \end{pmatrix}.$$

The remaining field, the *Higgs field*, has now a zero expectation value.

Rewriting the Lagrangian in 1.10 with the ϕ field in the unitary gauge, $\mathcal{L}_{EWSB}$

results from the sum of three terms:

$$\mathcal{L}_{EWSB} = \mathcal{L}_H + \mathcal{L}_{HW} + \mathcal{L}_{HZ}, \quad (1.14)$$

where the three terms can be written as follows, using the approximation $V \sim \mu^2 H^2 + \text{const}$ and neglecting higher order terms:

$$\begin{aligned} \mathcal{L}_H &= \frac{1}{2} \partial_\alpha H \partial^\alpha H + \mu^2 H^2 \\ \mathcal{L}_{HW} &= \frac{1}{4} v^2 g^2 W_\alpha W^{+\alpha} + \frac{1}{2} v g^2 H W_\alpha W^{+\alpha} \end{aligned} \quad (1.15)$$

$$\begin{aligned} &= m_W^2 W_\alpha W^{+\alpha} + g_{HW} H W_\alpha W^{+\alpha} \\ \mathcal{L}_{HZ} &= \frac{1}{8} v^2 (g^2 + g'^2) Z_\alpha Z^\alpha + \frac{1}{4} v (g^2 + g'^2) H Z_\alpha Z^\alpha \\ &= \frac{1}{2} m_Z^2 Z_\alpha Z^\alpha + \frac{1}{2} g_{HZ} H Z_\alpha Z^\alpha \end{aligned} \quad (1.16)$$

Equations 1.15 and 1.16 now contain mass terms for fields $W^\pm$ and Z : each of the three gauge bosons has acquired mass and an additional degree of freedom, corresponding to the longitudinal polarization. At the same time, three of the four Goldstone bosons have disappeared from the Lagrangian $\mathcal{L}_{EWSB}$, thus preserving the total number of degrees of freedom: the degrees related to the missing Goldstone bosons have become the longitudinal degrees of the vector bosons. Only the H scalar field is still present and has acquired mass itself: it is the Higgs field.

Summarizing, the Higgs mechanism is used to introduce the weak boson masses, without explicitly breaking the gauge invariance, thus preserving the renormalizability of the theory. When a symmetry is *spontaneously* broken it is not properly eliminated: it is rather *hidden* by the choice of the ground state. It can be shown that the minimum of the Higgs field is still invariant under the $U(1)_{EM}$ group. Hence, the electromagnetic symmetry is unbroken and the photon does not couple to the Higgs boson and remains massless.

1.3.1 Vector Boson Masses and Couplings

Equations 1.15 and 1.16 show that the masses of vector bosons $W^\pm$ and Z are related to the parameter v , characteristic of the electroweak symmetry breaking, and to the electroweak coupling constants:

$$\begin{cases} m_W = \frac{1}{2} v g \\ m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2} \end{cases} \rightarrow \frac{m_W}{m_Z} = \frac{g}{\sqrt{g^2 + g'^2}} = \cos \theta_W. \quad (1.17)$$

Also the couplings of vector bosons to the Higgs can be obtained from Equations 1.15 and 1.16, and are found to depend on the square of m_W and m_Z :

$$g_{HW} = \frac{1}{2} v g^2 = \frac{2}{v} m_W^2 \quad (1.18)$$

$$g_{HZ} = \frac{1}{2}v(g^2 + g'^2) = \frac{2}{v}m_Z^2. \quad (1.19)$$

A relation between the decay ratios of the Higgs boson to a W pair and to a Z pair can be derived from Equations 1.18 and 1.19:

$$\frac{BR(H \rightarrow W^+W^-)}{BR(H \rightarrow ZZ)} = \left(\frac{g_{HW}}{\frac{1}{2}g_{HZ}} \right)^2 = 4 \left(\frac{m_W^2}{m_Z^2} \right)^2 \simeq 2.4.$$

Finally, the EWSB energy scale can be determined from the relation between the v parameter and the Fermi constant G_F :

$$v = \left(\frac{1}{\sqrt{2}G_F} \right)^{\frac{1}{2}} \simeq 246 \text{ GeV}. \quad (1.20)$$

1.3.2 Fermion Masses and Couplings

The Higgs mechanism is also used to generate the fermion masses, by introducing in the SM Lagrangian an $SU(2)_I \otimes U(1)_Y$ invariant term, called *Yukawa term*, which represents the interaction between the Higgs and the fermion fields. Since ϕ is an isodoublet, while the fermions are divided in left-handed doublet and right-handed singlet, the Yukawa terms (one for each fermion generation) must have the following expression for leptons:

$$\mathcal{L}_\ell = -G_{H\ell} \cdot \bar{L}_\ell \phi \ell_R + \bar{\ell}_R \phi^\dagger L_\ell.$$

In the unitary gauge, the first component of ϕ is zero, therefore a mass term will arise from the Yukawa Lagrangian only for the second component of L_ℓ : this correctly reproduces the fact that neutrino is (approximately) massless.

$$\mathcal{L}_\ell = -\frac{G_{H\ell}}{\sqrt{2}}v\bar{\ell}\ell - \frac{G_{H\ell}}{\sqrt{2}}H\bar{\ell}\ell. \quad (1.21)$$

As far as the quark fields are concerned, the *down* quarks (d, s, b) are treated in the same way as leptons; *up* quarks (u, c, t), instead, must couple to the charge-conjugate of ϕ

$$\phi^c = -i\tau_2\phi^* = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^3 - i\phi^4 \\ -\phi^1 + i\phi^2 \end{pmatrix},$$

which becomes in the unitary gauge

$$\phi^c = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta + v \\ 0 \end{pmatrix}.$$

Therefore, the Yukawa Lagrangian is

$$\mathcal{L}_Y = -G_{H\ell}\bar{L}_L\phi\ell_R - G_{Hd}\bar{Q}_L\phi d_R - G_{Hu}\bar{Q}_L\phi^c u_R + h.c. \quad (1.22)$$

From Equation 1.21, the mass of a fermion (apart from neutrinos) and its coupling constant to the Higgs boson are found to be

$$m_f = \frac{G_{Hf}}{\sqrt{2}} v \quad (1.23)$$

$$g_{Hf} = \frac{G_{Hf}}{\sqrt{2}} = \frac{m_f}{v}. \quad (1.24)$$

Being G_{Hf} free parameters, the mass of the fermions cannot be predicted by the theory.

1.4 Boson Interaction Vertices

As previously illustrated, in the Standard Model electroweak gauge bosons are introduced to preserve the local $SU(2)_I \times U(1)_Y$ symmetry. The Lagrangian describing the gauge fields kinetics, which is invariant by construction, is of the form

$$\mathcal{L}_{gauge} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} \quad (1.25)$$

with

$$\begin{cases} B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \\ W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon_{bc}^a W_\mu^b W_\nu^c, \end{cases}$$

from which the physical fields W_μ^+ , W_μ^- , Z_μ and A_μ can be obtained by

$$\begin{cases} W_\mu^1 = \frac{1}{\sqrt{2}}(W_\mu^+ + W_\mu^-) \\ W_\mu^2 = \frac{1}{\sqrt{2}}(W_\mu^+ - W_\mu^-) \\ W_\mu^3 = \sin\theta_W A_\mu + \cos\theta_W Z_\mu \\ B_\mu = \cos\theta_W A_\mu + \sin\theta_W Z_\mu. \end{cases}$$

These relations represent the inverse of Equations 1.4, 1.5 and 1.6.

$\mathcal{L}_{gauge}$ can be expanded in terms of number of interacting fields, obtaining

$$\mathcal{L}_{gauge} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} = \mathcal{L}^{(2)} + \mathcal{L}^{(3)} + \mathcal{L}^{(4)},$$

where

$$\mathcal{L}^{(2)} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - \frac{1}{2}W_{\mu\nu}^+ W_-^{\mu\nu} \quad (1.26)$$

$$\begin{aligned} \mathcal{L}^{(3)} = & ig \sin\theta_W (W_{\mu\nu}^+ W_-^\mu A^\nu - W_{\mu\nu}^- W_+^\mu A^\nu + F_{\mu\nu} W_+^\mu W_-^\nu) + \\ & + ig \cos\theta_W (W_{\mu\nu}^+ W_-^\mu Z^\nu - W_{\mu\nu}^- W_+^\mu Z^\nu + Z_{\mu\nu} W_+^\mu W_-^\nu) \end{aligned} \quad (1.27)$$

$$\mathcal{L}^{(4)} = -\frac{g^2}{2}(2g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} - g^{\mu\sigma}g^{\nu\rho})[W_\mu^+W_\nu^-(A_\rho A_\sigma \sin^2 \theta_W + Z_\rho Z_\sigma \cos^2 \theta_W + 2A_\rho Z_\sigma \sin \theta_W \cos \theta_W) - \frac{1}{2}W_\mu^+W_\nu^+W_\rho^-W_\sigma^-]. \quad (1.28)$$

Equation 1.27 describes triples $WW\gamma$ and WWZ gauge bosons interactions, called triple gauge couplings or TGC, while Equation 1.28 represents quartic interactions, such as $WWWW$, $WWZZ$, $WWZ\gamma$ and $WW\gamma\gamma$, called quartic gauge couplings or QGC. These interactions can be illustrated in a graphical way by Feynman diagrams, in which particles are described by lines and their interactions by vertices. Diagrams representing TGC and QGC interactions are reported in Figures 1.2 and 1.3.

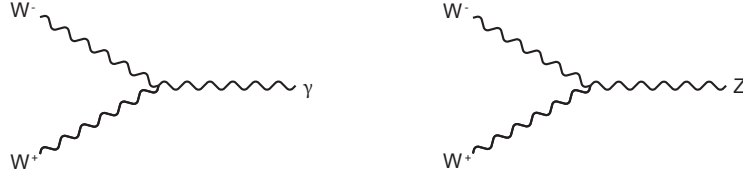


Figure 1.2: Standard Model gauge bosons triple couplings.

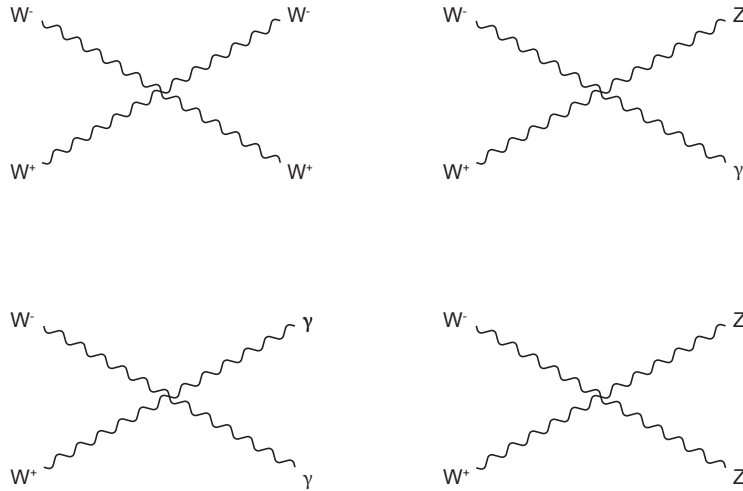


Figure 1.3: Standard Model gauge bosons quartic couplings.

Furthermore, as explained in Section 1.3, through the EWSB mechanism W and Z bosons couple to the Higgs (Equations 1.15 and 1.16), by becoming massive particles and acquiring the longitudinal polarization degree. Vertices describing these

interactions are reported in Figure 1.4.

The SM predicts also the Higgs boson self-interaction: by substituting $\phi(x) = \frac{1}{\sqrt{2}}(\begin{smallmatrix} 0 \\ v+H \end{smallmatrix})$ in the Higgs potential (Equation 1.11), one obtains

$$V(\phi) = -\frac{\mu^4}{4\lambda} + \mu^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4,$$

in which clearly the H^2 , H^3 and H^4 terms emerge, describing respectively the Higgs boson self-interaction and the triple and quartic self-couplings.

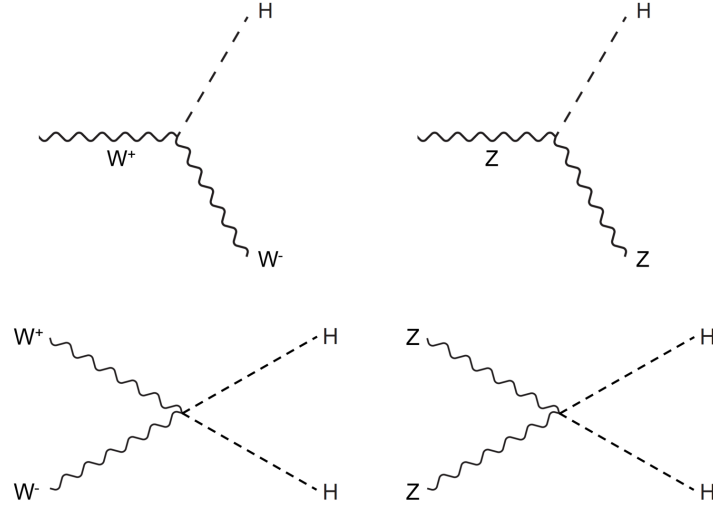


Figure 1.4: Standard Model gauge bosons couplings to Higgs particle.

Figure 1.5 summarizes all the described interactions, which constitute the theoretical structure of the SM Lagrangian: interacting particles are connected by lines, showing the couplings allowed by the theory. For W gauge boson, gluons, and Higgs boson the self-interactions are also predicted and they are represented by small loops. All these vertices are typical features of the SM and almost all of them have been investigated, nevertheless some still have to be tested and precisely measured. Both triple and quartic gauge couplings and the interaction with a Higgs particle are involved in the production mechanisms of vector boson pairs, in particular through Vector Boson Scattering (Section 1.5). Vector boson pair production is the main topic of this thesis and it is analyzed in Chapter 3, by describing the theoretical aspects and calculations on which this work is based. Moreover, quartic gauge couplings can be further explored in order to search for new physics, as explained in Section 1.6. In Chapter 7 anomalous quartic interactions are analyzed experimentally together with Vector Boson Scattering searches at the LHC.

1.5 Vector Boson Scattering

The production processes involving the scattering of two massive vector bosons (VBS), $VV \rightarrow VV$ with $V = W$ or Z , are of particular interest: they provide useful

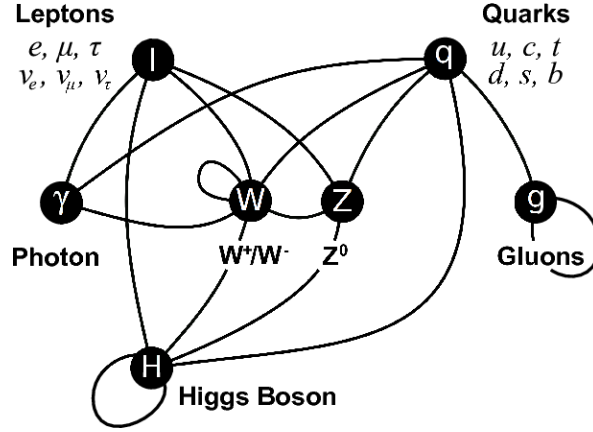


Figure 1.5: Couplings predicted by the Standard Model.

information on the validity of the SM and, in particular, they could be the final test of the nature of the Higgs boson. As described in previous paragraphs, the set of vector boson interactions that can produce di-boson pairs is determined by the gauge structure of the SM. These scattering interactions receive contribution through a variety of mechanisms, including both TGC and QGC diagrams by the exchange of an additional boson (W , Z or γ) and the interaction with a Higgs particle. In the former case, the involved diagrams are reported in Figure 1.6 (left). Each diagram corresponds to a scattering amplitude $\mathcal{M}$, which is related to the probability that the process occurs. Scattering amplitude of vector boson pair are found to be:

$$\mathcal{M}(W^+W^- \rightarrow W^+W^-) \sim \frac{1}{v^2}(s+t)$$

$$\mathcal{M}(W^+W^+ \rightarrow W^+W^+) \sim -\frac{1}{v^2}(t+u)$$

$$\mathcal{M}(ZZ \rightarrow W^+W^-) \sim \mathcal{M}(W^+W^- \rightarrow ZZ) \sim -\frac{1}{v^2}s$$

$$\mathcal{M}(ZW^+ \rightarrow ZW^+) \sim -\frac{1}{v^2}t.$$

Variables s , t and u are the so-called *Mandelstam variables*, encoding the energy, momentum, and angles of particles in a scattering process:

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2$$

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2$$

$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2,$$

in which p_1 and p_2 are the four-momenta of the incoming particles and p_3 and p_4 are the four-momenta of the outgoing particles (s is also known as the square of the center-of-mass energy).

The scattering amplitudes of longitudinally polarized vector bosons thus increase as a function of the center-of-mass energy $\sqrt{s}$ and violate unitarity at 1 – 1.5 TeV. Since unitarity is essentially the statement of conservation of total probability, it cannot be violated in Nature and some phenomena must intervene. In the SM

framework these divergences are totally canceled by the diagrams in which a Higgs boson is exchanged (Figure 1.6 right). In case of transversally polarized bosons, the amplitudes of these diagrams are negligible at high invariant mass: through the EWSB mechanism vector bosons acquire the longitudinal polarization and it is thus natural that the Higgs couples mainly to this polarization degree, while the transversally polarized bosons are less sensitive to the Higgs interactions. The amplitudes of longitudinally polarized vector bosons involving the Higgs exchange are:

$$\begin{aligned}\mathcal{M}(W^+W^- \rightarrow W^+W^-) &\sim -\frac{1}{v^2} \left(\frac{s^2}{s^2 - M_H^2} + \frac{t^2}{t^2 - M_H^2} \right) \\ \mathcal{M}(W^+W^+ \rightarrow W^+W^+) &\sim -\frac{1}{v^2} \left(\frac{t^2}{t^2 - M_H^2} + \frac{u^2}{u^2 - M_H^2} \right) \\ \mathcal{M}(ZZ \rightarrow W^+W^-) &\sim \mathcal{M}(W^+W^- \rightarrow ZZ) \sim -\frac{1}{v^2} \left(\frac{s^2}{s^2 - M_H^2} \right) \\ \mathcal{M}(ZW^+ \rightarrow ZW^+) &\sim -\frac{1}{v^2} \left(\frac{t^2}{t^2 - M_H^2} \right)\end{aligned}$$

and the total amplitudes become thus:

$$\begin{aligned}\mathcal{M}(W^+W^- \rightarrow W^+W^-) &\sim \frac{1}{v^2} \left(s + t - \frac{s^2}{s^2 - M_H^2} - \frac{t^2}{t^2 - M_H^2} \right) \sim -\frac{M_H^2}{v^2} \left(\frac{s}{s - M_H^2} + \frac{t}{t - M_H^2} \right) \\ \mathcal{M}(W^+W^+ \rightarrow W^+W^+) &\sim \frac{1}{v^2} \left(t + u - \frac{t^2}{t^2 - M_H^2} - \frac{u^2}{u^2 - M_H^2} \right) \sim -\frac{M_H^2}{v^2} \left(\frac{t}{t - M_H^2} + \frac{u}{u - M_H^2} \right) \\ \mathcal{M}(ZZ \rightarrow W^+W^-) &\sim \mathcal{M}(W^+W^- \rightarrow ZZ) \sim \frac{1}{v^2} \left(s - \frac{s^2}{s^2 - M_H^2} \right) \sim -\frac{M_H^2}{v^2} \left(\frac{s}{s - M_H^2} \right) \\ \mathcal{M}(ZW^+ \rightarrow ZW^+) &\sim -\frac{1}{v^2} \left(t - \frac{t^2}{t^2 - M_H^2} \right) \sim -\frac{M_H^2}{v^2} \left(\frac{t}{t - M_H^2} \right).\end{aligned}$$

Since they are constant as a function of the vector bosons invariant mass, the unitarity is therefore restored.

It should be noted that the only channel without divergences is $ZZ \rightarrow ZZ$: in this case just diagrams with a Higgs exchange in the s , t and u channel are involved and the amplitude shows a constant behavior as a function of the center-of-mass energy:

$$\begin{aligned}\mathcal{M}(ZZ \rightarrow ZZ) &\sim -\frac{1}{v^2} \left(\frac{s^2}{s^2 - M_H^2} + \frac{t^2}{t^2 - M_H^2} + \frac{u^2}{u^2 - M_H^2} \right) \\ \lim_{s,t,u \rightarrow +\infty} \mathcal{M}(ZZ \rightarrow ZZ) &= -\frac{1}{v^2} (s + t + u) = \frac{4M_Z^2}{v^2}.\end{aligned}$$

The study of VBS processes is thus particularly important in order to probe the nature of the EWSB mechanism and check if the unitarization is completely restored by it. The Higgs boson provides a plausible explanation for the mechanism that unitarizes these processes, but however many other physics scenarios are able to predict enhancements in VBS, either from additional resonances or if the observed SM-like Higgs boson only partially unitarizes this amplitude. Understanding which theory is correct is crucial and an in-depth analysis is reported in Chapter 7.

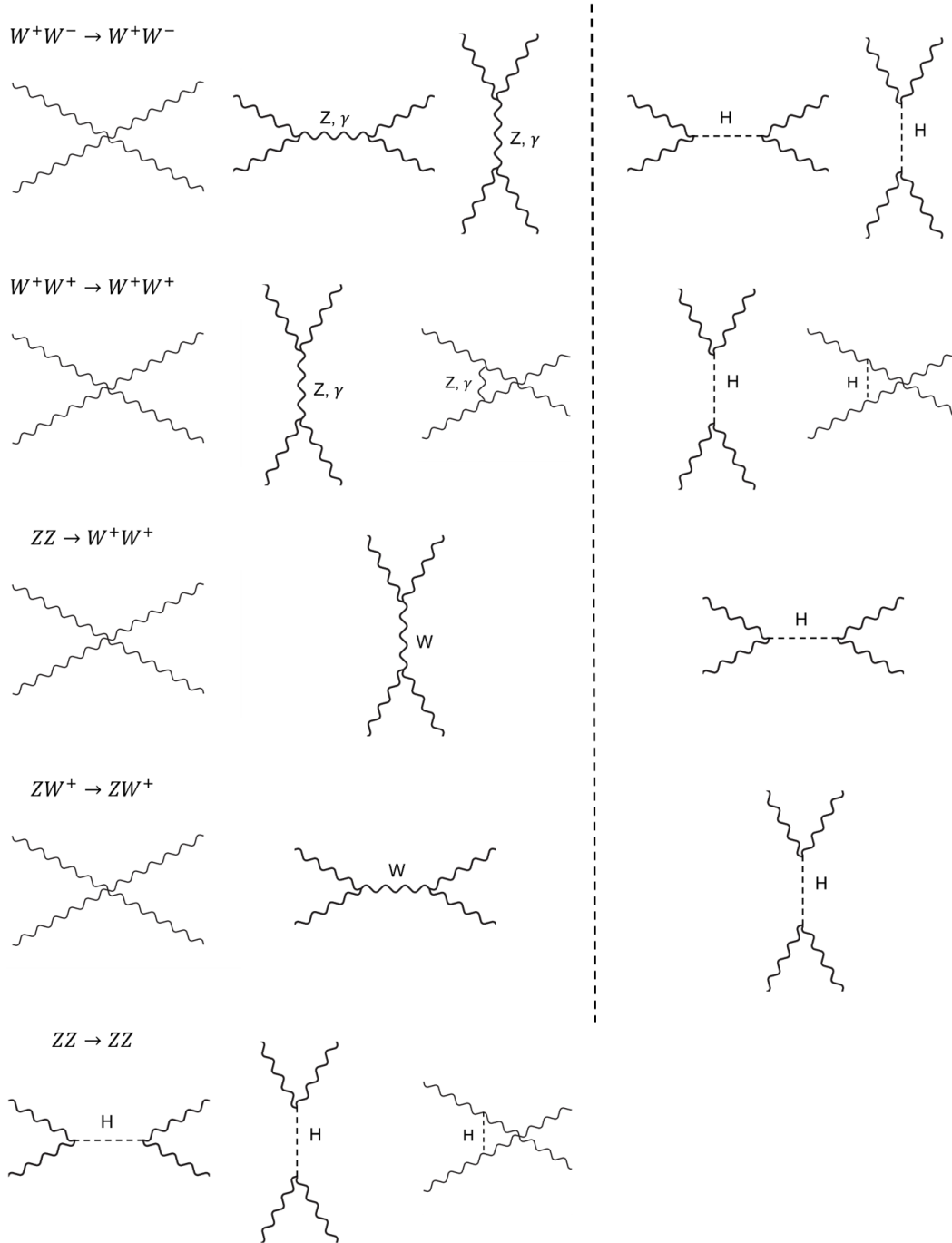


Figure 1.6: Vector boson scattering diagrams for each channel, without (left) and with (right) Higgs exchange.

1.6 Anomalous Quartic Gauge Couplings

The non-abelian nature of the electroweak sector predicts the self interaction of electroweak gauge bosons (see Section 1.4) in the form of triple and quartic gauge boson couplings (TGC and QGC). Besides the study of these couplings offers an important test of the SM, QGCs are additionally connected to the electroweak symmetry breaking sector, together with the Higgs boson, to ensure unitarity at high energy for scattering processes. QGCs therefore offer a new window on the electroweak symmetry breaking mechanism and represent a sector for observing new physics. Triboson production and VBS are directly sensitive to QGCs and their in-depth analysis provides complementary ways to study them. If any deviation from the SM predictions is observed, independent tests of the triple and quartic gauge boson couplings can give important information on the type of new physics responsible for the departures from the SM. For example, the exchange of heavy bosons can generate a tree level contribution to four gauge boson couplings, while its effect in the triple gauge vertex would only appear at one-loop level, and consequently be suppressed with respect to the quartic one.

1.6.1 Higher Dimension Operators

New physics can be searched using essentially two strategies [3]: by looking for new phenomena directly, or by looking for novel interactions of the known particles of the SM. Focusing on the latter method, one can formulate a more general theory using a model-independent approach which allows to search for new physics without committing a particular extension of the SM. The new extended theory should have the following characteristics:

- it should satisfy the S-matrix axioms of unitarity, analyticity, etc;
- it should respect the symmetries of the SM, namely Lorentz invariance and $SU(3)_C \times SU(2)_I \times U(1)_Y$ gauge symmetry;
- it should be possible to recover the SM in an appropriate limit;
- it should be general enough to capture any physics beyond the SM.

The unique way to incorporate all these features is via an Effective quantum Field Theory (EFT).

In the SM, all operators in the Lagrangian are restricted to be of mass dimension four or less. In order to extend the theory, operators of higher dimension can be added. By dimensional analysis, these operators have coefficients of inverse power of mass and hence are suppressed if this mass is large compared with the experimentally accessible energies. The Lagrangian of the effective quantum field theory of the SM can be written as

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{c_i}{\Lambda_i^2} \mathcal{O}_i + \dots,$$

where $\mathcal{O}_i$ are the higher dimension operators, while Λ is the mass scale that characterizes the coefficients of these operators and it can be regarded as the scale of new physics. This scale must be large compared with the energy reached experimentally and could be anywhere, from about a TeV up to the Plank scale. The coefficients c_i are dimensionless and parameterize the strength with which the new physics couples to the SM particles. It can be noticed that in the limit $\Lambda \rightarrow \infty$ the SM is recovered.

By dimension analysis, the dimension-six operators are expected to be dominant. There are many dimension-six operators, but only few of them affect physical processes and their coefficients can be measured or constrained by experiments. Electroweak boson pair production involves the coupling of electroweak vector bosons to fermions, that is generally constrained by other processes, and to each other. Focusing on the latter case, there are just three independent dimension-six operators that affect electroweak vector boson self interactions assuming conjugation (C) and parity (P) conservation. They are chosen as

$$\mathcal{O}_{WWW} = \text{Tr}[W_{\mu\nu}W^{\nu\rho}W_{\rho}^{\mu}]$$

$$\mathcal{O}_{\phi W} = (D^{\mu}\phi)^{\dagger}W^{\mu\nu}(D_{\nu}\phi)$$

$$\mathcal{O}_{\phi B} = (D^{\mu}\phi)^{\dagger}B^{\mu\nu}(D_{\nu}\phi)$$

where ϕ is the Higgs doublet field, D^{μ} is the covariant derivative and $W^{\mu\nu}$ and $B^{\mu\nu}$ are defined in Equation 1.4. These operators are constrained only by electroweak vector pair production.

If C and/or P violation are allowed, two additional operators have to be considered and are

$$\mathcal{O}_{\tilde{W}WW} = \text{Tr}[\tilde{W}_{\mu\nu}W^{\nu\rho}W_{\rho}^{\mu}]$$

$$\mathcal{O}_{\tilde{W}} = (D^{\mu}\phi)^{\dagger}\tilde{W}^{\mu\nu}(D_{\nu}\phi).$$

Together these five operators parametrize the leading effect of physics beyond the SM on the electroweak vector boson self interaction.

Dimension-six operators are the lowest order operators which can be built using the effective field theory approach. They modify the four gauge boson vertices, but affect either the two or three gauge boson couplings as well. Consequently they can be better searched for by looking into those effects. The lowest dimension operators that modify the quartic boson interactions but do not exhibit two or three weak gauge boson vertices are dimension-eight. There are three classes of parity-conserving dimension-eight operators:

- Operators containing just $D_{\mu}\phi$, without derivatives of the gauge fields. They are only two:

$$\mathcal{O}_{S,0} = [(D_{\mu}\phi)^{\dagger}D_{\nu}\phi] \times [(D^{\mu}\phi)^{\dagger}D^{\nu}\phi]$$

$$\mathcal{O}_{S,1} = [(D_{\mu}\phi)^{\dagger}D_{\nu}\phi] \times [(D^{\mu}\phi)^{\dagger}D^{\nu}\phi]$$

- Operators containing $D_{\mu}\phi$ and derivatives of the gauge fields, for example:

$$\mathcal{O}_{M,0} = \text{tr}[(W_{\mu\nu}W^{\mu\nu}) \times [(D_{\beta}\phi)^{\dagger}D^{\beta}\phi]]$$

- Operators containing only the gauge fields, for example:

$$\mathcal{O}_{T,1} = \text{tr}[(W_{\alpha\nu}W^{\mu\beta}) \times \text{tr}[(W_{\mu\beta}W^{\alpha\nu})].$$

Given that dimension-eight operators modify only four boson vertices, they are largely investigated in the search for new physics at particle colliders.

1.6.2 Unitarity Bounds

Since unitarity must be preserved in Nature, to ensure the conservation of total probability, bounds are typically set in order to prevent amplitudes and cross-section from increas-

ing too much with energy. In the EFT approach these unitarity bounds are irrelevant. Dimension-six(eight) operators yield terms in the amplitude that grow as $s/\Lambda^{2(4)}$ and will eventually violate the unitarity bound at high energy. However, when that happens, the effective field theory has become useless and should be discarded. Once s/Λ^2 is of order unity, there is no justification for ignoring the yet higher dimension operators, since they are not suppressed. The effective field theory, by construction, is a low energy theory and has no ambitions of being applicable at arbitrarily high energy. The theory must respect the unitarity bound in the region where there is data. Since the data necessarily respect the unitarity bound, an effective field theory that fits data will automatically respect the bound.

In order to better understand the experimental results summarized in Chapter 7, in Figure 1.7 the W^+W^- cross-section obtained by adding the operator $\mathcal{O}_{WWW}$ to the SM is reported, with a coefficient $c_{WWW}/\Lambda^2 = (400 \text{ GeV})^{-2}$. As it can be noticed, the cross-section deviates from the SM at large invariant mass, as expected. The pink-dashed line represents the sum of the SM cross-section, the interference between SM and the dimension-six contribution and the square of the dimension-six contribution. In this particular case, it can be shown that interference term is order v^2/Λ^2 and does not grow with energy. The square of the dimension-six contribution is order s^2/Λ^4 and leads to the growth of the cross-section at high invariant mass. If experimental data fits the non-standard curve, it can be concluded that $c_{WWW}/\Lambda^2 \sim (400 \text{ GeV})^{-2}$ (but c_{WWW} and Λ^2 cannot be separated) and Λ must be greater than the invariant mass of the largest data point. In addition, Λ must be less than the energy at which the curve violates unitarity, thus one can bound Λ from above and from below. However a measurement of $c_{WWW}/\Lambda^2 \sim (400 \text{ GeV})^{-2}$ does not indicate that $\Lambda \sim 400 \text{ GeV}$, because the coefficient c_{WWW} is not constrained to be unity: it can be less or much greater and some constraints are set by the theory.

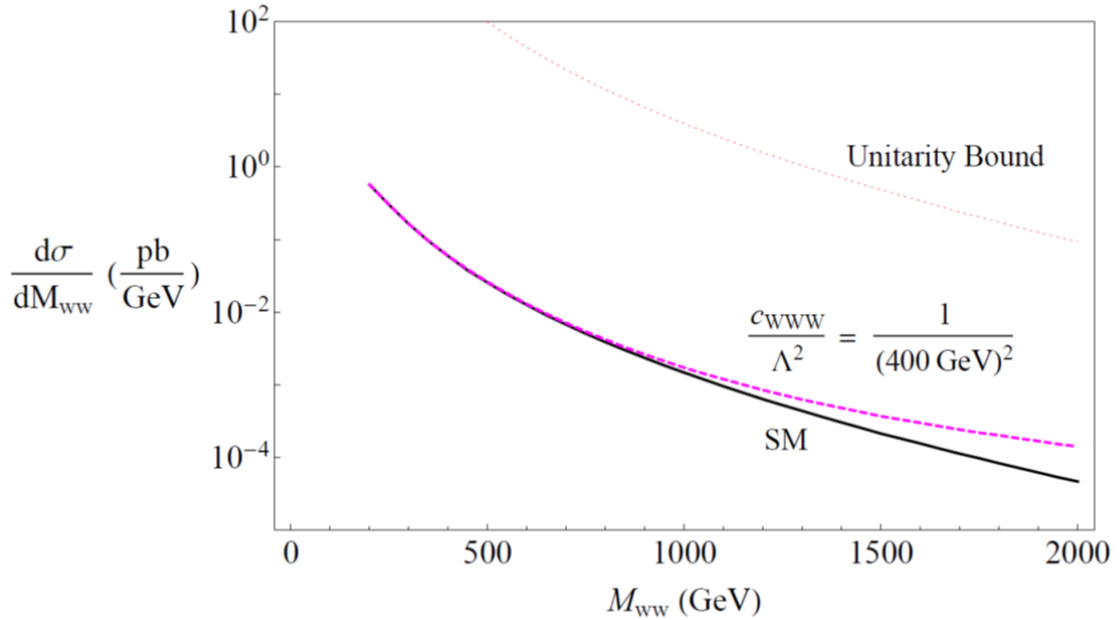


Figure 1.7: Differential cross-section versus invariant mass for the process $pp \rightarrow W^+W^-$ at 14 TeV. Both the SM cross-section and the cross-section including the dimension-six operator $\mathcal{O}_{WWW}$ are shown. The unitarity bound on the cross-section is also reported [3].

1.7 Beyond the Standard Model

Although the SM is a very successful theory in describing particle physics, several experimental observations and theoretical issues show that it is incomplete and at some higher energy it must be embedded into a wider new theory. Phenomena such as the existence of dark matter or the matter-antimatter asymmetry of the universe are not explained by the SM, even if experimentally observed. In parallel theoretical inconsistencies and features artificially added by hand, like for example the so-called *hierarchy problem*, are believed to be unnatural. All these issues motivate the community of particle physicists to look for a more complete theory which is able to solve the internal problems of the SM and to give a satisfactory answer to the open questions in particle physics.

In the following the most important unsolved issues of the SM are briefly summarized, together with the theoretical attempts designed to solve them.

1.7.1 Dark Matter and Dark Energy

According to cosmological observations, the SM explains about 5% of the energy present in the universe, while the rest is formed of *dark matter* ($\sim 27\%$) and *dark energy* ($\sim 68\%$). The dark matter is a non-baryonic, not electromagnetically interacting kind of particle matter, which interacts with the ordinary matter only through the gravitational and weak forces. The existence of dark matter is confirmed by various observations of the stellar objects, like the measurements of the gravitational lensing of galaxy clusters. Dark matter is studied through both direct searches, in cosmic rays by underground experiments, and searches at hadron colliders, where according to the Supersymmetry model it would be produced in pairs of neutral particles (the neutralinos), often called Weakly Interacting Massive Particles or WIMPS. Despite the existence of several Beyond Standard Model (BSM) theories describing this phenomenon, to date no evidence of dark matter has been found in any experiment.

The dark energy is an unknown form of energy which permeates space and could explain the observations indicating the universe's expansion at an accelerating rate. Several possibilities have been formulated for the dark energy: it could be a property of space, a new dynamic fluid, or a new theory of gravity. However, in order to decide among them, more data are needed.

1.7.2 Matter-Antimatter Asymmetry

Another open issue in cosmology is the matter-antimatter asymmetry of the universe. Despite at the time of Big Bang an equal amount of matter and antimatter should have been produced, the universe consists predominantly of matter, while antiparticles appear only in cosmic rays and cosmic phenomena or when produced artificially by collisions at the accelerators. This asymmetry can only be understood in theories that violate baryon number and CP (particle-antiparticle) symmetry, and in which the universe evolves far from thermal equilibrium during part of its early life. The SM predicts processes which proceed via CP-violation interactions, but they are too small to generate the observed amount of asymmetry. In BSM theories other possible CP-violating mechanisms are taken into consideration.

1.7.3 Neutrino Masses

The SM doesn't provide for a Higgs boson coupling to neutrinos, describing them as massless particles. However, in 1998 the Super Kamiokande Experiment observed the oscillation of neutrino flavors ($\nu_e \rightarrow \nu_\mu$), that is possible only if neutrinos are massive. By combining results from various cosmology and oscillating experiments, it has been discovered that neutrino masses are different for the three lepton families and are of the order of a few eV. Mass terms for neutrinos can be added by hand to the theory and several extensions of the SM explain this phenomenon, like for example the Majorana/Dirac representation, the study of right-handed extended symmetry group and the so-called *see-saw* mechanism.

1.7.4 The Hierarchy Problem

The *hierarchy problem* is an important aspect of the SM and is one of the big reasons why we expect physics beyond it. There's many different ways to describe this problem, each emphasizing a different feature of it. First of all a large discrepancy of the order of 10^{32} is observed between electroweak force and gravity. This translates into a huge mass difference between the W and Z boson, describing the scale of the electroweak interactions ($\sim 10^2$ GeV), and the Plank mass, defining the scale beyond which gravity must be described by quantum mechanics ($\sim 10^{18}$ GeV). The W and Z bosons acquire mass through the Higgs mechanism, as explained in Section 1.3. Despite the Higgs mass is measured to be about 125 GeV, quantum mechanics corrections due to interactions with virtual particles would lead to a very larger value of it. This is another way of seeing the *hierarchy problem*, which occurs when the fundamental value of some physical quantity, like for example the Higgs mass, is much different from its effective value, i.e. the one measured in an experiment, because of quantum corrections to it (*renormalization*). Typically the renormalized value is close to its fundamental value, but in some cases it appears that there has been an incredible fine-tuning cancellation between the radiative corrections and the bare quantity.

Although the fine-tuning of the Higgs field would not be itself an inconsistency of the SM, since in this strict context the Higgs mass cannot even be calculated, theoretical physicists are looking for new models providing a mechanism to avoid such large corrections. One proposed solution is Supersymmetry, an elegant theory involving superpartner particles. These objects introduce huge contributions to the Higgs mass-squared parameter but with opposite sign with respect to the radiative corrections, removing the divergences.

Extensive searches have been done in the first phase of data taking of LHC, showing no evidence of Supersymmetry in a wide range of parameters and masses of the superpartner particles.

Chapter 2

The CMS Experiment at LHC

The Large Hadron Collider (LHC) is the largest particle accelerator ever built, colliding beams of protons and heavy ions at an unprecedented energy. Four experiments are located in correspondence to the four collision points: ATLAS, ALICE, LHCb and CMS. The latter is a multi-purpose detector designed to study a wide range of physics processes. In particular, it was built to search for the Higgs boson and to provide evidence of new physics at the TeV scale.

The present Chapter gives a short description of the LHC and focuses on the Compact Muon Solenoid (CMS) experiment, one of the four main detectors installed on the LHC beam-line. The LHC design is described, together with its successful operation during Run I and the scheduled future plans. Later, the CMS experiment hardware setup, the reconstruction software and the main trigger features are outlined. Finally, a brief summary of the most important results achieved by CMS during the first period of data-taking is reported.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [4] is the world's biggest particle accelerator, designed to collide opposing particle beams of either protons and lead ions. It has been built by the European Organization for Nuclear Research (CERN) in order to find an answer to the most fundamental questions of physics, like understanding the electroweak symmetry breaking for which the Higgs mechanism is presumed to be responsible. It is an unprecedented machine in terms of energy, luminosity, size, complexity of experiments, cost and involvement of human resources.

2.1.1 The LHC Design

The collider is placed in a circular tunnel 27 km long, situated at a depth of about 100 m underground, on the French-Swiss border. This tunnel was built between 1983 and 1988 for the Large Electron-Positron Collider (LEP) and it contains two adjacent and parallel beam pipes, where proton (or ion) beams travel in opposite directions around the ring. The LHC is designed to collide two beams of protons, each one with a nominal energy of 7 TeV, for a total center-of-mass energy of 14 TeV. Moreover, for the ion physics program, shorter running periods with lead nuclei collisions at 2.76 TeV/NN are performed.

Since collisions occur between particles of the same charge, the two acceleration cavities are



Figure 2.1: View of the CERN site: the LHC tunnel, with the four experiments located at each interaction point.

required to have different magnetic configurations. The beams are kept on their circular path by means of 1232 dipole magnets along the tunnel and are focused by 392 quadrupole magnets. The magnets of the LHC are superconducting and are kept at an operational temperature of 1.9 K using about 96 tonnes of liquid helium. The field in the magnet varies between 0.53 – 8.3 T during the acceleration of the protons from 450 GeV to 7 TeV. The two beams cross in four different points in correspondence to the four main experimental detectors: ATLAS, ALICE, LHCb and CMS (Figure 2.1).

One of the most important characteristic of a collider is its luminosity $\mathcal{L}$, crucial indicator of performance for an accelerator, that ties event rate to the cross-section (σ) of a process:

$$\frac{dN_i}{dt} = \sigma_i \mathcal{L}(t)$$

(where the i subscript indicates a general process). The luminosity is proportional to the number of particles colliding within a defined amount of time and, since discoveries in particle physics rely on statistics, the greater the number of collisions, the more chances physicists have to see a particle or process that they have not seen before. Assuming a Gaussian beam shape, the luminosity at LHC depends only on the accelerating machine parameters and it can be written as:

$$\mathcal{L} = f \frac{n_b N_1 N_2}{4\pi\sigma_x\sigma_y}$$

where f is the revolution frequency of the n_b proton bunches per beam, N_1 and N_2 are the number of protons in the colliding bunches, σ_x and σ_y are the beam profiles in horizontal and vertical directions at the interaction point.

The designed peak of instantaneous luminosity for ATLAS and CMS is $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The total integrated luminosity delivered to CMS in 2011 was 6.1 fb^{-1} , while 23.3 fb^{-1} were provided in 2012 (see Figure 2.2).

The main parameters of the LHC are summarized in table 2.1.

2.1.2 LHC Operations during Run I

The first data-taking run of the LHC started in Fall 2010, being the origin schedule delayed by about one year because of an accident in 2008, and ended in Spring 2013. During its first three years, the LHC ran at an initial collision energy of 7 TeV, 3.5 TeV per beam, rising

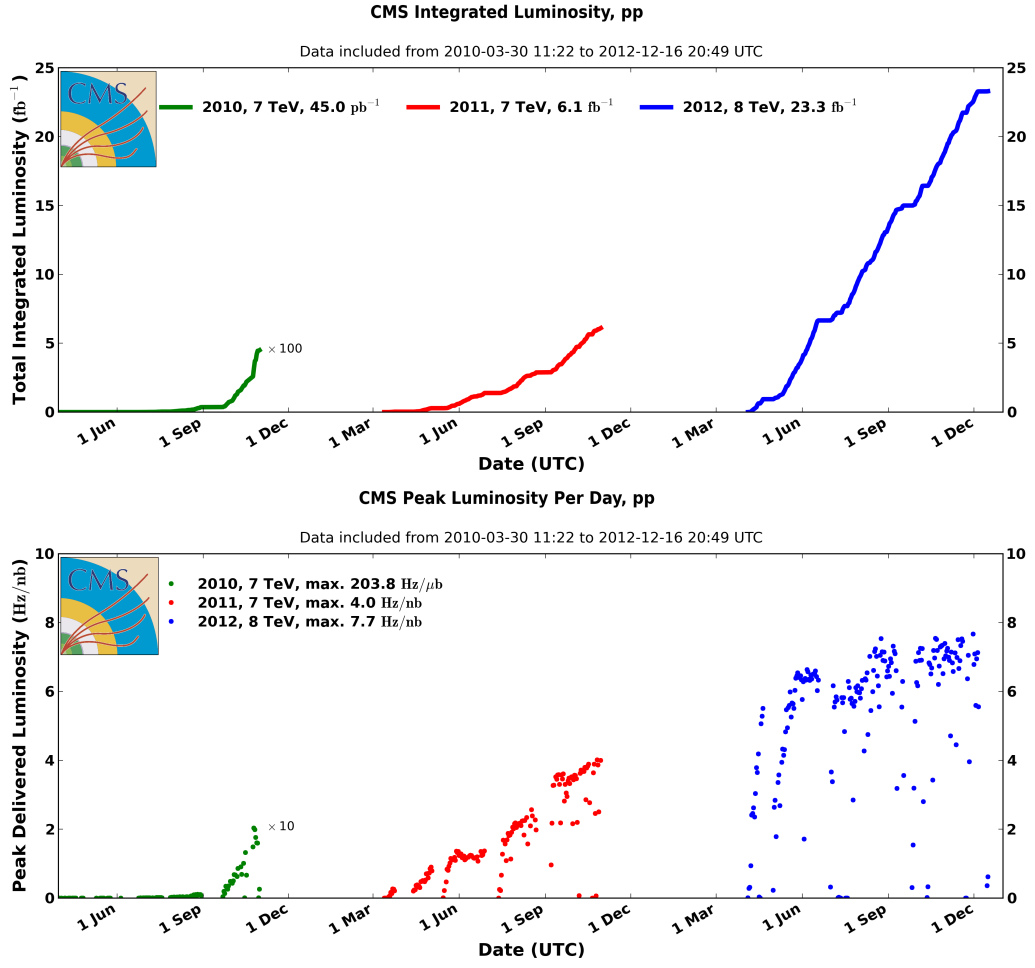


Figure 2.2: Cumulative luminosity (top) and peak luminosity (bottom) delivered to the CMS experiment as a function of time, shown for 2010 (green), 2011 (red) and 2012 (blue) data-taking [5].

Table 2.1: The LHC design parameters for $p - p$ and $Pb - Pb$ collisions.

Parameter	$p - p$	$Pb - Pb$
Circumference [km]	26.659	
Beam radius at interaction point [μm]	15	
Dipole peak field [T]	8.3	
Design center-of-mass energy [TeV]	14	1148
Design Luminosity [$\text{cm}^{-2} \text{s}^{-1}$]	10^{34}	$2 \cdot 10^{27}$
Luminosity lifetime [h]	10	4.2
Number of particles per bunch	$1.1 \cdot 10^{11}$	$\sim 8 \cdot 10^7$
Number of bunches	2808	608
Bunch length [mm]	53	75
Time between collisions [ns]	24.95	$124.75 \cdot 10^3$
Bunch crossing rate [MHz]	40.08	0.008

to 8 TeV from 2012. The beams consisted of trains of several proton bunches separated by 150 ns in 2010, 75 ns in the beginning of 2011 and 50 ns in 2011 and 2012. In November 2010 LHC delivered the first lead-ion collision run, at the end of which it was shutdown and restarted on March 2011. In April of the same year, LHC became the world's highest luminosity accelerator, with a peak luminosity of $4.67 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, by beating the Tevatron's record of $4 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$. One year later, the beam-energy was increased up to 4 TeV in order to gain both in luminosity and in Higgs production cross-section and this was crucial for the discovery of a new resonance with mass around 125 GeV, announced on the 4th of July by the ATLAS and CMS collaborations. On the 13th of February 2013 the LHC's first run officially ended, and the machine was shut down for planned upgrades.

2.1.3 Run II and Future Perspectives

After a long technical stop of more than two years, in June 2015 the LHC has started its second three-year data-taking period, the so-called *Run II*. Over the last 16 months, the LHC has been through a major program of maintenance and upgrading, in order to prepare the machine for running at its design energy, almost double the energy of Run I. On the 20th of May 2015, LHC reached the center-of-mass energy of 13 TeV, opening up a new window on the sub-nuclear universe. Thanks to the improved performances, in the next years it will be possible to search for new physics as well as verify theories that could not be tested before. The analysis of the first data has already started and it is expected to lead to new important results in several fields. The schedule of the Run II and the luminosity collected so far by the CMS experiment are reported in Figures 2.3 and 2.4.

To extend its discovery potential, the LHC will need a major upgrade around 2020 and this necessity has given rise to the High-Luminosity (HL) LHC project. The HL-LHC will increase the luminosity by a factor of 10, delivering 10 times more collisions than the LHC would do over the same period of time. Upgrading the LHC will be a challenging procedure and relies on several breakthrough technologies currently under development, such as cutting-edge 12 Tesla superconducting quadrupole magnets built using a superconducting compound of niobium and tin, which will strongly focus the beam to increase the probability of collisions occurring. The HL-LHC will therefore provide more accurate measurements of fundamental particles and enable physicists to observe rare processes that occur below the current sensitivity level of the LHC. With this upgrade, the LHC will continue to push the limits of human knowledge, enabling physicists to explore beyond the Standard Model and the electroweak symmetry breaking mechanism.

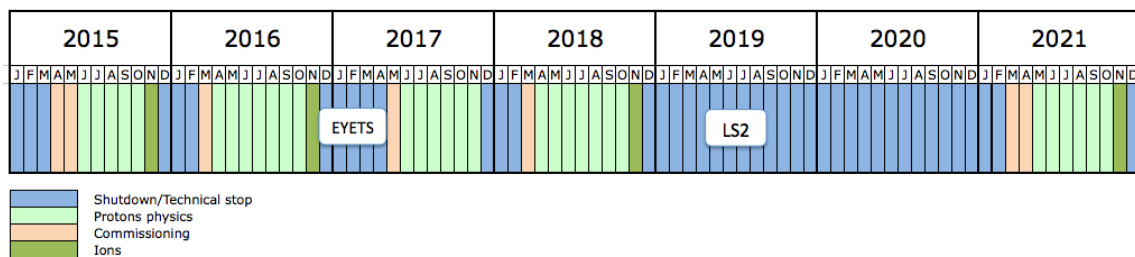


Figure 2.3: The medium-term LHC schedule including Run II from 2015 to 2018 [6].

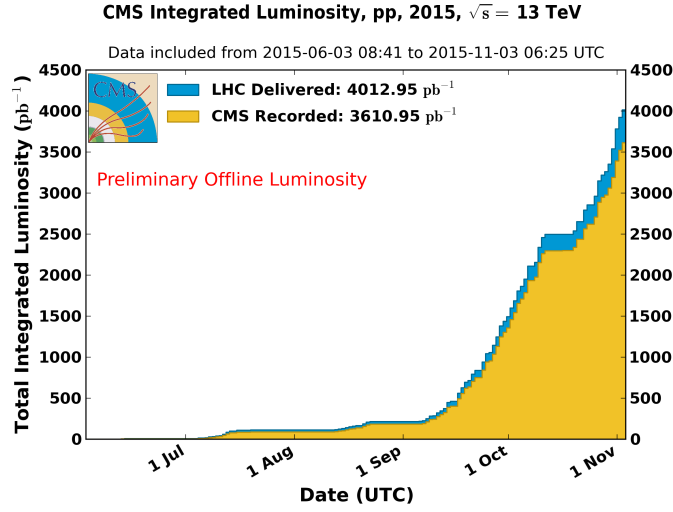


Figure 2.4: Cumulative off-line luminosity versus day delivered to (blue), and recorded by (orange) CMS during stable beams and for pp collisions at 13 TeV center-of-mass energy in 2015 [5].

2.2 The Compact Muon Solenoid Experiment

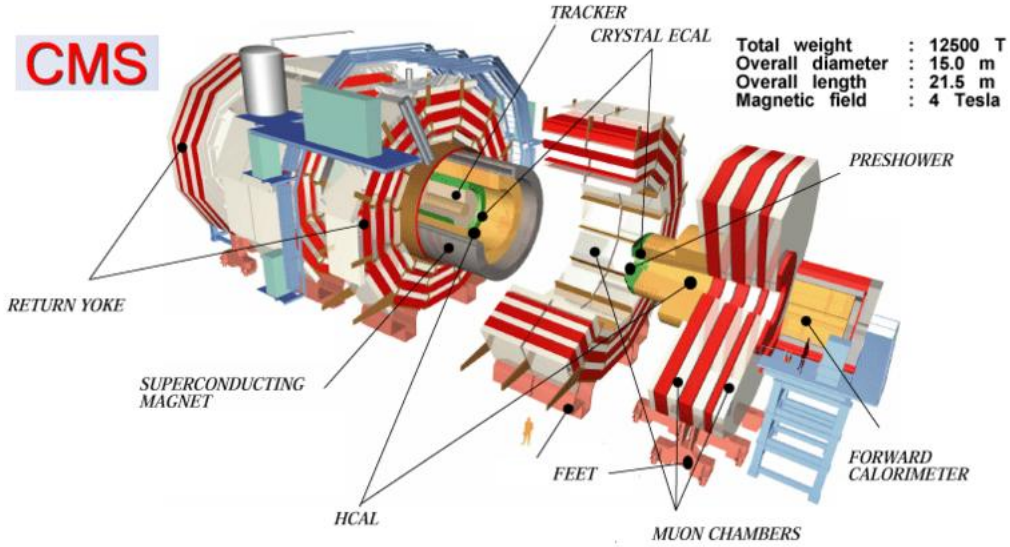


Figure 2.5: Schematic picture of CMS experiment at the LHC.

CMS [7] is one of the four main experiments at LHC. Its primary goals are the study of the electroweak symmetry breaking mechanism linked to the Higgs mechanism, the search for physics Beyond the Standard Model (BSM) and the precision measurements of already known physics processes. In order to fulfill these goals, excellent lepton reconstruction and particle identification are required. A particular magnetic field configuration is necessary to provide a large bending power and to have thus good resolution in measuring charge particles of high momentum. The main feature of the CMS detector is the 3.8 T superconducting solenoid, which allows a compact design with a strong magnetic field, besides a

robust and redundant muon system, a fully active scintillating crystals-based electromagnetic calorimeter and a high-quality tracking system.

The overall structure of CMS consists of several cylindrical layers coaxial to the beam axis (the *barrel layers*), closed at both ends by detector disks orthogonal to the beam direction (the *endcaps*), to ensure optimal hermeticity. The dimensions of the overall detector are fairly remarkable: the total length is 21.5 m, the diameter 15 m and the weight about 12500 tons. A schematic view of the CMS detector is shown in Figure 2.5.

2.2.1 Definition of Kinematic Variables

The CMS experiment uses a cylindrical coordinate system, centered at the interaction point. The direction of the z -axis is chosen along the beam and it is referred as *longitudinal*. The x -axis is horizontal and pointing towards the center of the LHC ring and the y -axis is vertical and pointing upwards. The $x - y$ plane, orthogonal to the beam line, is called *transverse plane*. According to these definitions, the momentum of a particle can be divided in two components: the *longitudinal momentum* p_z and the *transverse momentum* p_T , defined as

$$p_T = \sqrt{p_x^2 + p_y^2}.$$

The *rapidity* of a particle of energy E is defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}$$

and for high-energy particles it can be approximated by the *pseudorapidity*

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right),$$

which only depends on the polar angle θ of the particle momentum, measured with respect to the z -axis. The azimuthal angle, measured in the $x - y$ plane from the x -axis, is called ϕ (Figure 2.6).

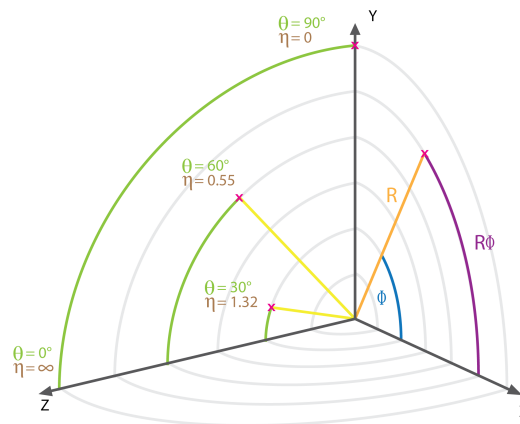


Figure 2.6: Pseudorapidity η and azimuthal angle ϕ used to track particles inside CMS.

In the following, every component of the CMS detector is briefly presented, focusing on the performances and on their impact on physics analysis.

2.2.2 The Tracker

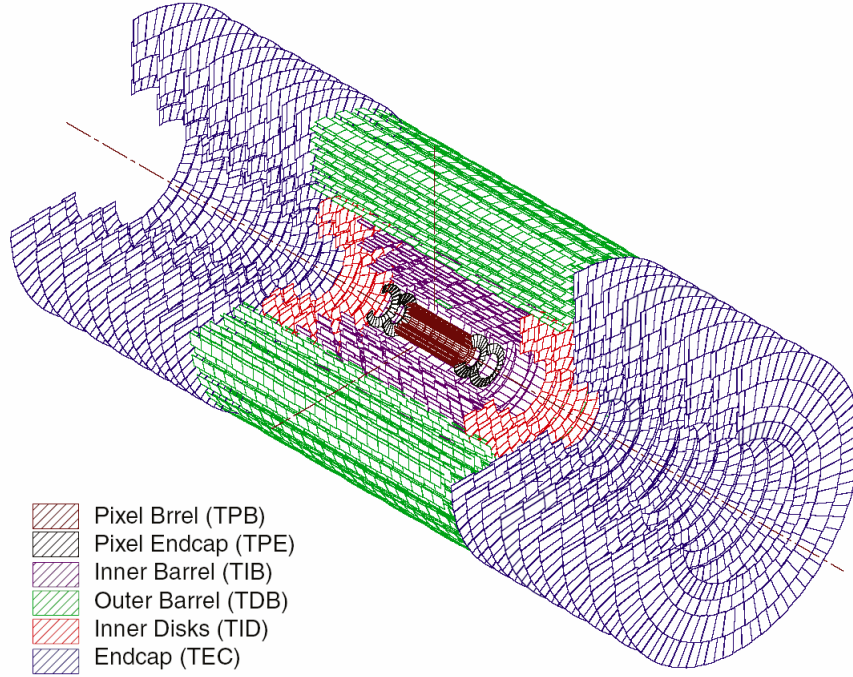


Figure 2.7: View of the silicon tracker, including the pixel detector [7].

The Tracker (Figure 2.7) is the CMS innermost sub-detector and the closest to the interaction point. Its main purpose is the reconstruction of charged tracks and vertices, together with the measurement of charged particles p_T , that are important tools for identifying signal events and rejecting background.

It extends in the region of $|\eta| < 2.5$, $r < 120$ cm, $|z| < 270$ cm and it is completely based on semiconductor detectors made of silicon, that cover a total surface of 210 m^2 .

Because of its tasks, the major physics requirements for the CMS Tracker are an efficient reconstruction of isolated lepton tracks (in all the pseudorapidity coverage is close to 100%), a good lepton momentum resolution (for $|\eta| < 2$, $\sigma(p_T)/p_T < 4\%$ for single muons in a large range of p_T) and a good reconstruction of interaction vertices.

For optimal pattern recognition, the Tracker is designed to fulfill two basic properties: low cell occupancy and large hit redundancy. The low occupancy is obtained by high granularity detectors, mainly those closest to the interaction point (because of higher particle fluxes), and a fast primary charge collection, obtained using thin detectors and reducing the silicon bulks. The redundancy is guaranteed by the overall design which provides many measured points per track (on average 12-14, see Figure 2.8). All these characteristics allow for a high tracking efficiency and a low rate of fake tracks, i.e. reconstructed tracks that do not correspond to any real track.

Moreover, one of the major constraints of the Tracker design is to reduce the amount of material crossed by particles, to minimize electron bremsstrahlung and hadronic interactions. The inner tracking system of the CMS detector has a concentric layout. Pixel detectors are

placed closest to the interaction point ($r \leq 10$ cm), because of the high particle density (see Figure 2.7). They consist of three barrel layers and two endcap disks for each side. In the region $20 < r < 55$ cm the particle flux decreases enough to allow the use of silicon microstrips (made of 4 barrel layers and 3 disks at each side), while in the outermost region of the tracker ($r > 55$ cm) larger pitch microstrips are placed (consisting of 6 barrel layers and 9 disks for each endcap).

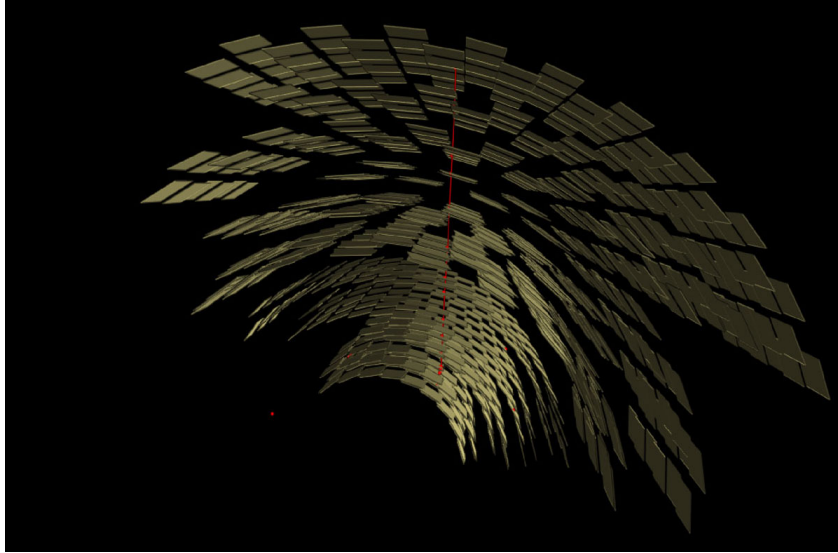


Figure 2.8: Example of a particle track into the tracker.

2.2.3 The Electromagnetic Calorimeter

The primary aim of the Electromagnetic Calorimeter (ECAL) is the identification and measurement of the energy of electrons and photons. The physics process that dictates the strictest requirements on its performance is the low mass Higgs decay into two photons $H \rightarrow \gamma\gamma$, one of the leading channels for the study of the Higgs properties. The goal is 1% resolution on the di-photon invariant mass and the natural choice to reach this goal is a homogeneous calorimeter.

ECAL is made of 74848 lead tungstate ($PbWO_4$) crystals, chosen because of their excellent energy resolution. It is a transparent material characterized by a high density (8.28 g/cm^3), a short radiation length ($X_0 = 0.89 \text{ cm}$) and a small Molière radius (2.2 cm). These features allow a very compact shape and a fine granularity, necessary because of the high particle density produced at the LHC. Moreover, these crystals have a trapezoidal shape and they are characterized by a very short scintillation decay time, which permits to collect about 80% of the light within 25 ns, so that they can be used at the crossing rate of 40 MHz.

The ECAL barrel covers the central rapidity region ($|\eta| < 1.48$) and the two ECAL endcaps extend the coverage up to $|\eta| = 3$ (Figure 2.9).

2.2.4 The Hadron Calorimeter

The Hadron Calorimeter (HCAL) surrounds ECAL and its design is strongly influenced by the choice of the magnet parameters and by the fact that most of it is placed inside the

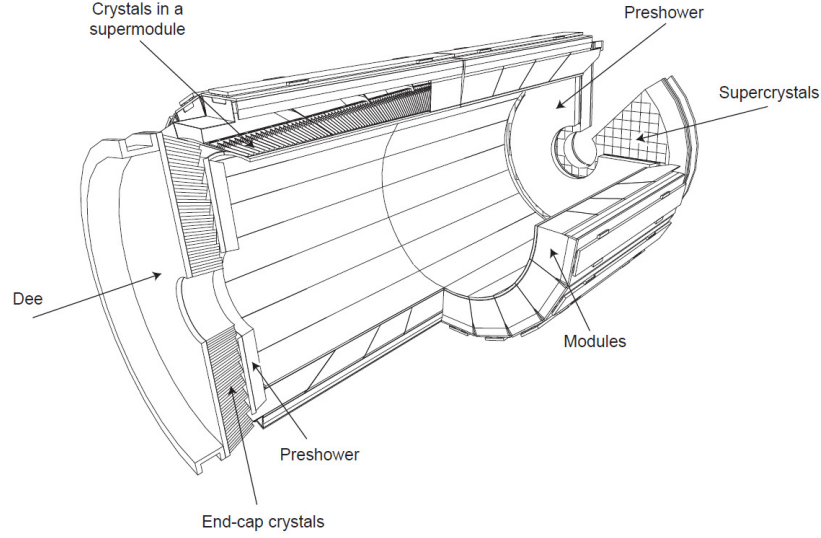


Figure 2.9: Layout of the CMS Electromagnetic Calorimeter showing the arrangement of crystal modules, supermodules and endcaps, with the preshower in front [7].

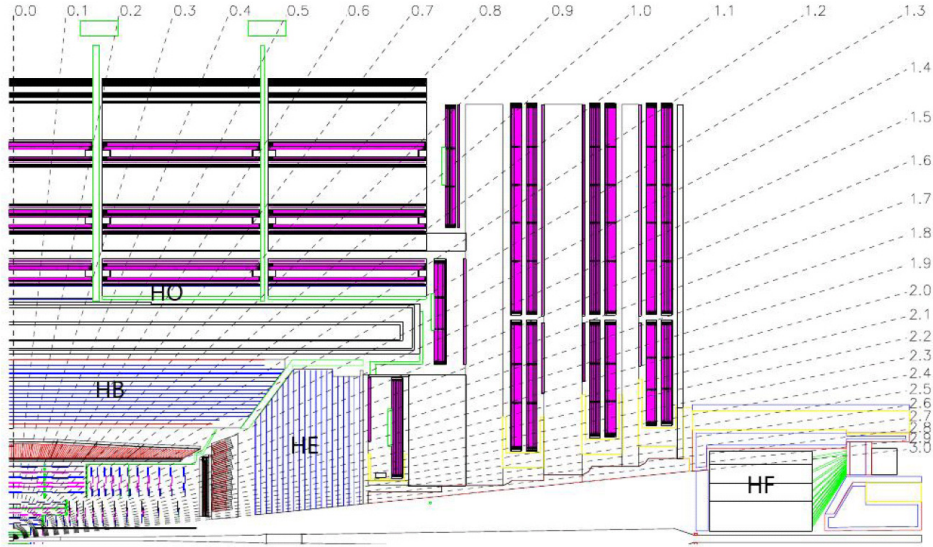


Figure 2.10: Longitudinal view of one quarter of the detector in the $r - z$ plane, showing the positions of the HCAL parts: Hadron Barrel (HB), Hadron Outer (HO), Hadron Endcap (HE) and Hadron Forward (HF) [7].

magnet coil. This detector plays an essential role in the identification and measurement of hadrons by estimating the energy and the direction of jets. It is also used for neutrino detection by measuring the missing transverse energy of the event. For this reason, one of the main design requirements is a high hermeticity. In particular, the HCAL angular coverage must include the very forward region, since the identification of forward jets is very important for the rejection of many backgrounds and the evidence of BSM signatures. The HCAL is a sampling calorimeter, that means it finds particles position, energy and arrival time using alternating layers of *absorber* and *scintillator* materials. In particular, brass

was chosen as absorber material, thanks to its non-magnetic behavior and its quite short interaction length ($\lambda_I \sim 151$ mm). To maximize the amount of absorber before the magnet, the space devoted to the active medium is minimized. The active part is made of scintillator tiles coupled with wavelength shifting fibers and clear fibers carrying the light to the read-out system.

The Hadronic Calorimeter can be divided in four parts, that permit a good segmentation, a moderate energy resolution and a full angular coverage (Figure 2.10). The *Barrel Hadronic calorimeter* (HB) surrounds the electromagnetic calorimeter and covers the central pseudorapidity region up to $|\eta| = 1.3$. The endcap regions are covered up to $|\eta| = 3$ by the two *Endcap Hadron calorimeters* (HE). The HB and HE are located inside the solenoid magnet. To satisfy the hermeticity requirements, *Forward Hadronic calorimeters* (HF) are placed outside the magnet yoke, 11 m far from the interaction point, extending the pseudorapidity coverage up to $|\eta| = 5$. Finally, an array of scintillators located outside the magnet, which is referred to as the *Outer Hadronic calorimeter* (HO), is used to improve the central shower containment.

2.2.5 The Magnet

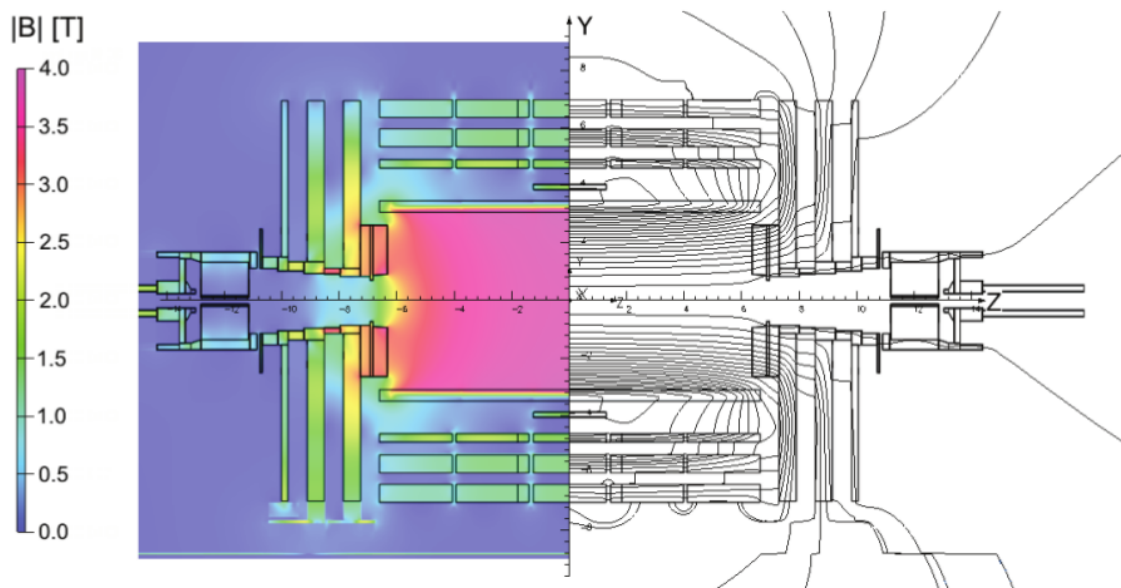


Figure 2.11: Values of the magnetic field $|B|$ (left) and field lines (right) predicted on a longitudinal section of the CMS detector [8].

In order to achieve a good momentum resolution for momenta up to 1 TeV, CMS needs a strong solenoidal magnetic field. The CMS Magnet is the central device around which the experiment is built and its dimensions limited the size of the total apparatus.

Its purpose is to bend the paths of particles emerging from high-energy collisions. The higher particle momentum is, the less its trajectory is curved by the magnetic field. A higher strength field, combined with high-precision position measurement in the tracker and muon system, gives accurate measurement of momentum.

The CMS Magnet is a 13 m long superconducting solenoid, the largest ever built. It is able to generate a uniform magnetic field of 4 T in the inner region, storing about 2.5 GJ of energy (Figure 2.11). It operates with a superconducting technology at a temperature of 4 K, ensured by a helium cooling system. The Magnet is then contained in a huge vacuum cylinder, which isolates it from the external environment. Outside, an iron structure constitutes the iron yoke, necessary to bridge the return magnetic flux, which otherwise would get lost, disturbing the surrounding environment.

2.2.6 The Muon System

Because muons can penetrate several meters of iron without being stopped, they are detected by muon chambers, placed in the outermost part of the experiment. The Muon Detection System is therefore set outside the magnetic coil and it has multiple tasks: triggering on muons, identifying them and assisting the tracker in measuring their momentum. After being detected by the inner tracker, muons are detected again in the Muon System. This redundancy allows a sensible improvement in the resolution of high-momentum particles, for which spatial resolution dominates, while for lower momenta (up to p_T values of 200 GeV) the inner tracker is sufficient. The trajectory beyond the return yoke can be then extrapolated back to the beam-line, thanks to the compensation of the bend before and after the coil. The minimum value of the muon transverse momentum required to reach the system is 3 GeV.

The spectrometer is composed by three independent subsystems, i.e. three different types of gaseous particle detectors: *Drift Tube* (DT) *Chambers* and *Cathode Strip Chambers* (CSC) are used in the regions of $|\eta| < 1.2$ and $0.9 < |\eta| < 2.4$ respectively, and are complemented by a system of *Resistive Plate Chambers* (RPC) covering the range of $|\eta| < 1.6$. The detector can thus be divided into three regions, referred to as *barrel* ($|\eta| < 0.9$), *overlap* ($0.9 < |\eta| < 1.2$) and *endcap* ($|\eta| > 1.2$). The reason for these different technologies lies in the different particle rates and occupancies, both higher in the endcaps, and in the intensity of the stray magnetic field, which is lower in the barrel. Thanks to the fast response they provide, all the muon detectors are also used within the first level of the trigger system.

In the barrel, the expected occupancy is low, allowing for the use of Drift Tubes as detection element. In this region the detector stations are interleaved with the iron plates of the yoke. The chambers segmentation follows that of the iron yoke, consisting of five wheels along the z -axis, each one divided into 12 azimuthal sectors. As one can see in Figure 2.12, in each sector the chambers are organized in four stations (labeled *MB1*, *MB2*, *MB3* and *MB4*). There are 12 chambers in each of the 3 inner layers; in the 4th layer the top and bottom sectors host 2 chambers each, thus leading to a total of 14 chambers per wheel in this outermost layer. The DTs provide a precise track measurement in the bending plane; their maximum drift time is about 400 ns, with a time resolution of about 1 ns. Each chamber has a resolution of about 100 μm in $r\phi$ and 1 mrad in ϕ . Each DT chamber has 1 or 2 RPCs coupled to it before installation, depending on the station. RPCs are a double-gap bakelite chambers operating in avalanche mode, with gap width of 2 mm.

In each of the endcaps, the Cathode Strip Chambers are arranged in 4 disks perpendicular to the beam and in concentric rings, 3 rings in the innermost station and 2 in the others. Each CSC has a trapezoidal shape and is made of 6 gas gaps (7 layers) with planes of cathode strips in the radial directions and anode wires almost perpendicular to the strips. Most CSCs are overlapped in ϕ in order to avoid gaps in acceptance. The ionization of a charged particle passing through the planes causes the charge to be collected on the anode wire and the image charge on the cathode strips, thus allowing to get (r, z, ϕ) hits in each layer.

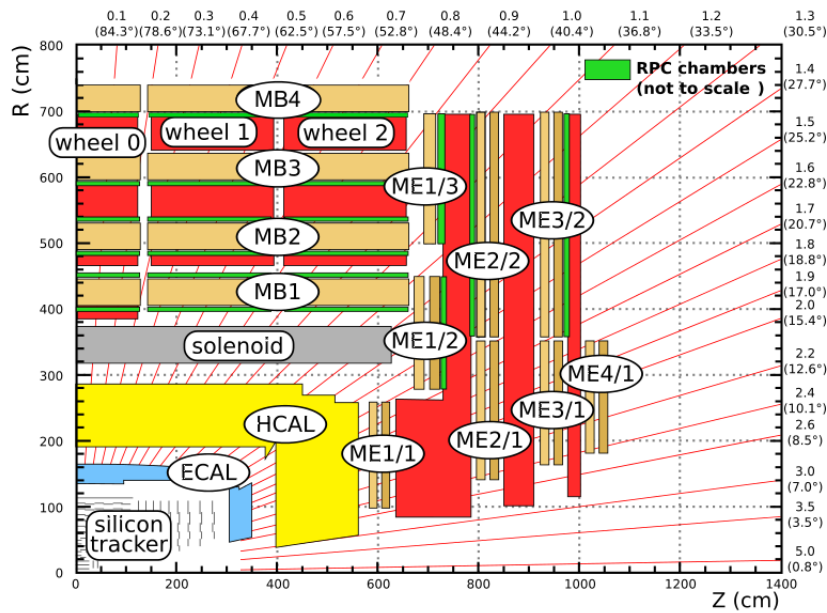


Figure 2.12: An $r - z$ cross-section of a quadrant of the CMS detector. The locations of the various muon stations and the steel disks (red areas) are shown: four DTs stations are labeled MB (*muon barrel*) and CSCs are labeled ME (*muon endcap*). RPCs (in green) are both the barrel and the endcaps of CMS [9].

2.3 The CMS Trigger System

At the instantaneous luminosity expected in proton-proton collisions at LHC, the interaction rate (~ 40 MHz) produces 10^9 interactions/second, but only a small part of them can be written to archival media. Such a high luminosity causes several interactions overlap in the same bunch crossing, effect known as *pile-up*, as well as the overlap in the detector of signals from different bunch crossings, due to the limited speed of detector response and read-out. There are therefore several technical difficulties in handling, storing and processing this huge amount of data and a selection of events written to permanent storage is thus necessary. As a consequence, one of the most important and difficult aspects of the experiment is the Trigger, the real-time selection and recording of useful events.

The data size per event is ~ 1 MB and just a rate of 300 Hz can be stored for off-line analysis. In order to fulfill this data reduction, the CMS Trigger is divided into two levels: the *Level-1 Trigger* (L1), mainly rested on hardware, and the *High Level Trigger* (HLT), software-based. The following sections briefly describe the design of the L1-Trigger and the general structure of the HLT.

2.3.1 The Level-1 Trigger

The Level-1 Trigger is based on custom-made hardware and its task is to take an accept-reject decision for each bunch crossing, every 25 ns, reducing the rate of selected events down to 100 kHz. The Level-1 decision must be fast and it is therefore taken within a fixed time of $3.2 \mu\text{s}$. This time includes also the transit of data from front-end electronics to Level-1 Trigger processors and back and the time available for calculations can thus be no more than $1 \mu\text{s}$.

Because of the bunch crossing high rate and slowness of the tracker algorithms, the Level-1

Trigger uses only coarsely segmented data from calorimeters and muon detectors, while all the high-resolution data is stored in pipelines of processing elements, waiting for the trigger decision. If the L1 accepts the event, the data are moved to be processed by the High Level Trigger.

The calorimeter and the muon sub-detectors provide objects called *trigger primitives* in the form of local energy deposits in calorimeter trigger towers (identifying electron, photon and jet candidates) and track segments or hits in muon chambers (identifying muon candidates). These objects are then ranked and stored and they form the basis for trigger decisions taken by the final L1 stage, the *Global L1 Trigger*, according to programmable algorithms. Finally, the *Trigger Control System* determines if the sub-detectors are ready to read out the event and if the Data Acquisition (DAQ) system is ready to receive it.

2.3.2 The High Level Trigger

The High Level Trigger accepts all events that pass the L1 Trigger filter, using detector data transferred from the pipelines to the front-end read-out buffers. The reconstruction and selection in the HLT take place in steps and uninteresting events are rejected as quickly as possible. The basic idea of the HLT software is to reconstruct those parts of each physics object that can be used for selection, while minimizing the overall CPU usage. The first selection step in the HLT process, called *Level-2 Trigger*, has the full Level-1 rate as input and uses only information from the calorimeter and muon detectors. On the other hand, *Level-3 Trigger*, the second step of the selection, includes the reconstruction of full tracks in the tracker. In order to minimize the CPU time required by the HLT, a key feature of the algorithms is to reconstruct the information in the CMS detector only partially, taking the accept-reject decision on the reconstruction of quantities in a limited region of the detector. The HLT software system reduces the event rate from 100 kHz to 300 Hz.

2.4 The CMS Results in Run I

During the first three years of the LHC operation, the CMS Experiment has been working extremely well, and more than 90% of delivered data has been used for physics analysis. Several results have been achieved, advancing our knowledge of the fundamental particles and their interactions.

Thanks to the good understanding of the detector and the accurate theoretical predictions, it has been possible to measure with very high precision many processes predicted by the Standard Model. As reported in Figure 2.13, several cross-sections have been evaluated, finding excellent agreement with next-to-leading (NLO) or next-to-NLO theoretical expectations (see Chapter 3).

On July 2012 CMS, together with the ATLAS collaboration, has reached its greatest success: the discovery of the Higgs boson, first observed in the di-photon and ZZ decay channels and then confirmed by measurements performed in other modes [10]. The Higgs boson mass has been obtained combining the CMS and ATLAS results and it is already known with an impressive precision ($m_H = 125.09 \pm 0.21 \pm 0.11$ GeV [11]). The ZZ invariant mass distribution with the Higgs signal over the background is reported in Figure 2.14 (left).

In the hard sector, no discrepancies with predictions have been observed in jet distributions over many orders of magnitude [12] and the strong-force coupling constant α_s has been measured up to so-far unproved scales (~ 2 TeV [13]). The top mass, another fundamental parameter of the Standard Model related to the electroweak vacuum stability, has been measured from a kinematic fit of $t\bar{t}$ pair decaying into a lepton plus at least four jets,

Up to now, no evidence of super-partners has been observed, pushing their masses increasingly away from the electroweak scale. Finally, all searches of new resonances or particles connected to beyond Standard Model scenarios have been unsuccessful so far, making the upcoming Run II necessary in order to solve the fundamental physics problems that still remain.

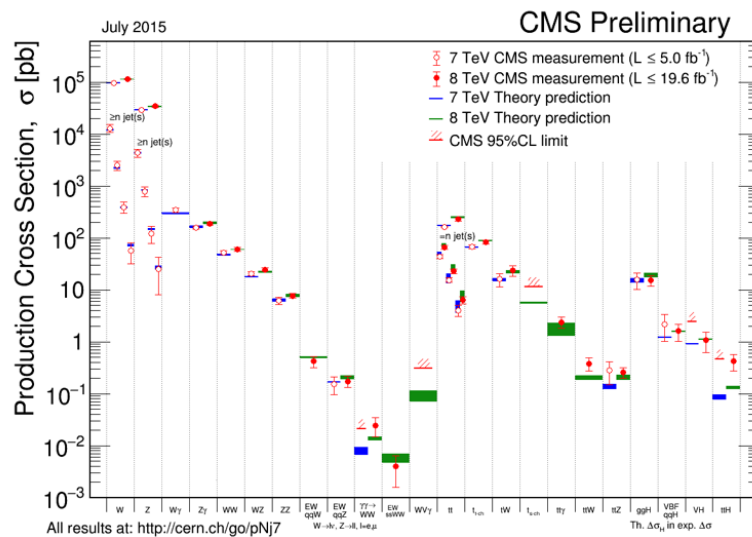


Figure 2.13: Summary plot of CMS cross-section measurements [16].

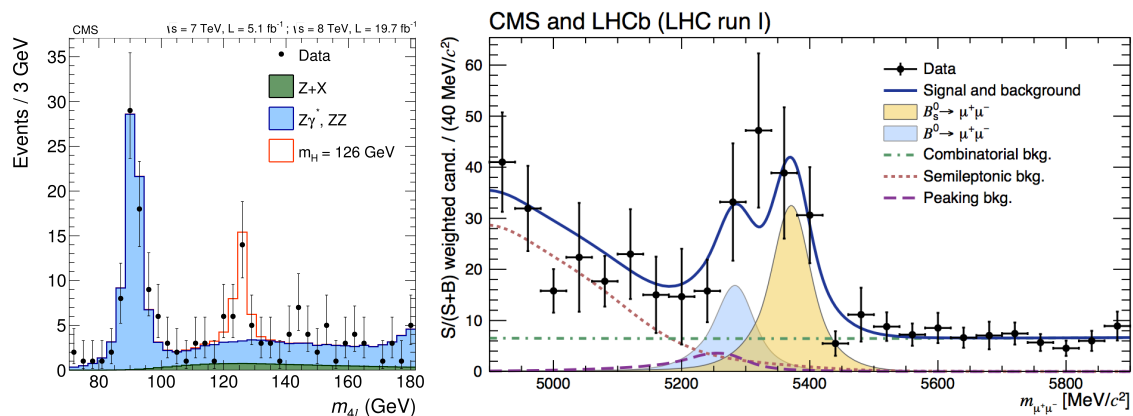


Figure 2.14: Distribution of the four-lepton reconstructed mass for the mass region $70 < m_{4\ell} < 180$ GeV (left) [17]. Measured invariant dimuon mass in the B_s^0, B^0 decay range (right) [15].

Chapter 3

Study of Massive Vector Boson Pair Production

Massive vector boson pair production is among the most important electroweak processes at hadron colliders, being relevant for physics analyses within and beyond the Standard Model. Through its investigation, detailed studies of the gauge symmetry structure of electroweak interactions and of the mechanism of electroweak symmetry breaking can be carried out. Moreover, these processes could provide a first evidence for new physics, probing deviations from Standard Model expectations through the measurement of production rates and kinematic distributions. Vector boson pair production is also an important background in the study of the Higgs boson: although the Higgs resonance is well below the WW and ZZ threshold, the off-shell WW and ZZ backgrounds are crucial both in the extraction of the Higgs signal and in the measurement of its properties. All these features lead to the requirement of precise experimental measurements and rigorous theoretical understanding of massive vector boson pair production.

In this Chapter, the analyses on vector boson pair production carried out by ATLAS and CMS Collaborations are summarized. The ZZ production is studied carefully, being the central subject of this thesis, to introduce the work presented in the next Chapters. In order to perform precise SM measurements, detailed theoretical predictions are needed: after an introduction on the definition of leading and next-to-leading order cross-sections, the current status of calculations is then summarized and the main theoretical results on massive vector boson pair production are finally reported.

3.1 Massive Vector Boson Pair Production at the LHC

A profound understanding of vector boson pair production processes at the LHC is needed for various reasons. First of all, such processes contribute as an important irreducible background to the Standard Model Higgs production at moderate energies. Second, they could also provide deeper insight into the physics of the weak interaction at the high energy frontier, possibly even allowing for the discovery of beyond Standard Model physics. Weak bosons can decay into jets, charged leptons or neutrinos and hence produce the same signatures as new colored particles, new electroweak gauge bosons or dark matter candidates. At the LHC experiments ZZ , $W^\pm Z$ and W^+W^- production processes can be investigated in a plethora of final states, that sometimes link together the different VV production mechanisms. The experimental results reviewed here use pp collisions at the LHC at a

center-of-mass energy $\sqrt{s} = 7$ and 8 TeV, with an integrated luminosity of about 5 and 20 fb⁻¹.

3.1.1 WW Production

WW production processes are measured in the $WW \rightarrow 2\ell 2\nu$ and $WW \rightarrow \ell\nu 2j$ (together with WZ) final states. In the former case, the signal is characterized by two opposite charged isolated leptons (electrons or muons) accompanied by significant amount of missing energy. Background processes include top production, suppressed by applying an additional veto jet; W + jets contributions, estimated from data; Drell-Yan processes, evaluated from Monte Carlo samples normalized to data; di-boson processes, such as WZ and ZZ production, suppressed by third lepton veto; $W\gamma$ production, in which photon converts, reduced by rejecting electrons consistent with a photon conversion.

$WW + WZ \rightarrow \ell\nu 2j$ events are selected with one well-identified and isolated lepton (muon or electron), large missing energy, and exactly two jets with high transverse momentum. This semileptonic final state could be more convenient than the purely leptonic process, thanks to the W and Z bosons larger branching ratio to quarks decay than leptons decay. However this advantage is partially nullified by the huge W + jets background, which has a very large cross-section. The other main backgrounds are Z + jets, top production and QCD multi-jet processes. These contributions are reduced by additional jet and lepton vetoes and cuts on the W transverse mass.

3.1.2 WZ Production

The WZ production measurement is performed by the analysis of $WZ \rightarrow 3\ell\nu$ events. The signal is thus characterized by two opposite charge and same flavor leptons to pair into an on-shell Z boson, typically within a mass window between 10 – 20 GeV around the nominal mass of the Z boson, an additional isolated lepton with high momentum, and missing transverse energy greater than few ten of GeV, associated with the escaping neutrino. The most important background for this channel is the production of Z + fake lepton, including fake or non-isolated leptons from jets or photons. $ZZ \rightarrow 4\ell$ processes also contribute to the background when one of the four leptons is not detected, but this component is quite small.

Furthermore, the WZ production is also investigated in the case in which the Z boson decays hadronically into a $b\bar{b}$ pair, while the W decays leptonically into $\ell\nu$. Here the main backgrounds are the production of the W with light flavor jets, $t\bar{t}$ and single-top events. These contributions are reduced using the b -tagging technique and additional jet activity veto.

3.1.3 ZZ Production

The ZZ production cross-section is measured in the following final states: $ZZ \rightarrow \ell\ell\ell'\ell'$ (with $\ell = e, \mu$ and $\ell' = e, \mu, \tau$), $ZZ \rightarrow 2\ell 2\nu$ and in the hadronic decay $ZZ \rightarrow 2\ell(\nu)b\bar{b}$. For $ZZ \rightarrow \ell\ell\ell'\ell'$ signal is characterized by four charged leptons. Z candidate is formed using a pair of well-identified and isolated leptons of the same flavor and opposite charge with invariant mass consistent with the nominal mass of Z boson. The major background contributions arise from the production of Z and WZ in association with jets and from $t\bar{t}$, where a jet or a non-isolated lepton is misidentified as a lepton (see Chapter 5). On the other hand, the $ZZ \rightarrow 2\ell 2\nu$ final state is characterized by a pair of opposite charge, isolated electrons

or muons, with invariant mass close to nominal mass of Z boson (additional leptons are vetoed) and large missing transverse energy, due to the neutrino pair which escapes direct detection. In order to minimize backgrounds coming from top quarks, events are vetoed if a b -tagged jet is found. Finally, in the hadronic decay case the event selection is based on the reconstruction of one of the two Z bosons in its leptonic decay modes and the other in its decay into two b -tagged jets. As in the WZ case, to suppress the Z + light flavor jets component b -tagging is used, while an additional jet activity veto is employed to reduce $t\bar{t}$ and single top background events.

In Tables 3.1 and 3.2 the total cross-sections of WW, WZ and ZZ processes, measured by ATLAS and CMS with $\sqrt{s} = 7$ and 8 TeV, are presented. They are consistent with the Standard Model cross-sections, taking into account the uncertainties. Figures 3.1 and 3.2 show the relative production cross-section for various electroweak boson pairs for both the experiments.

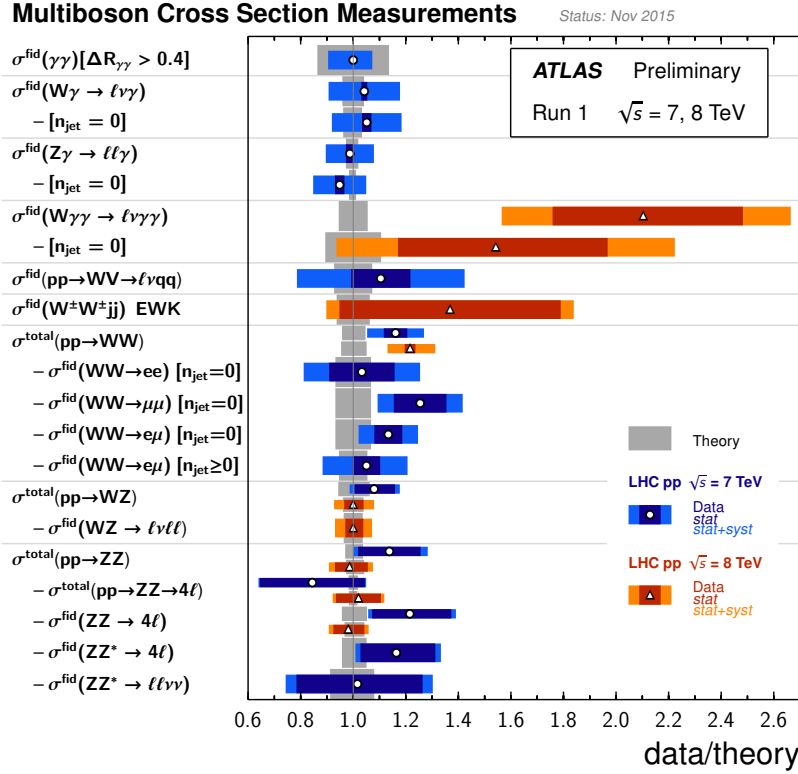


Figure 3.1: Summary plot for the measured multi-boson cross-sections and their ratio with theory predictions for ATLAS [18].

3.2 $ZZ \rightarrow 4\ell$ Processes

The study of ZZ boson pair production is the central topic of the presented thesis. As already pointed out, the production of vector boson pairs is a crucial process for physics

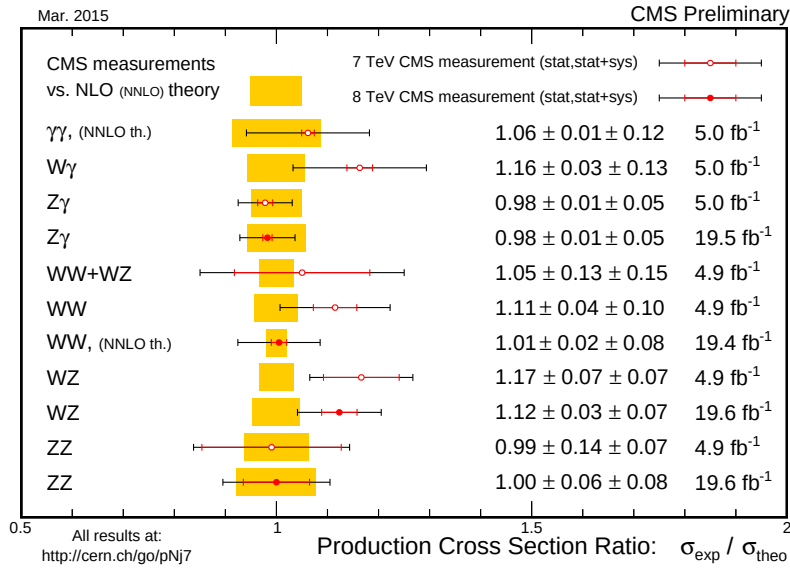


Figure 3.2: Summary plot for the ratio between measured multi-boson cross-sections and theory predictions for CMS [16].

studies within and beyond the Standard Model. In particular the production of Z boson pairs plays an important role as the principal background to the Higgs boson signal. Moreover, neutral ZZZ and $ZZ\gamma$ couplings are forbidden in the Standard Model, thus any deviation of the observed strength from the Standard Model prediction, manifested as a change in the cross-section or as a deviation in the kinematic distributions of the vector boson pair, would indicate new physics. For these reasons measuring this process with high precision is a crucial task. In parallel, great effort has been made during the last years to push its theory predictions to a new level, by extensively studying QCD and EW corrections, as reported in Section 3.4.

In the Standard Model, ZZ production proceeds primarily via the t - and u -channel $q\bar{q}$ scattering diagrams, with additional contributions from gluon gluon *box* diagrams (Figure 3.3). Although the production of Z pairs is much smaller than the other di-boson cross-sections, this process has the advantage to present a very clean signature, due to the four leptons in the final state.

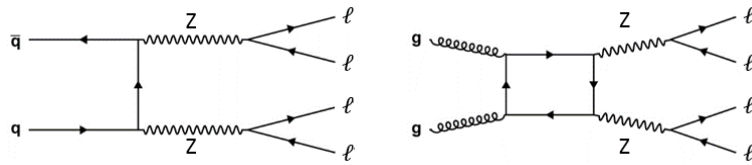


Figure 3.3: Lowest order diagrams for the $q\bar{q} \rightarrow ZZ$ process (left) and for the $gg \rightarrow ZZ$ process (right).

Various measurements of ZZ production have been performed in the past, as briefly summarized in Subsection 3.1.3. The CMS Collaboration measured the cross-section 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 [34, 33]. The measured total

Table 3.1: Observed and expected cross-sections for vector boson pair production at $\sqrt{s} = 7$ TeV.

Process	Exp.	$\int \mathcal{L} dt$	σ_{tot}	(stat.)	(syst.)	(lumi.)	σ_{tot}^{pred} [pb]	Ref.
$WW \rightarrow 2\ell 2\nu$	ATLAS	4.6 fb^{-1}	51.9	2.0	3.9	2.0	$44.7^{+2.1}_{-1.9}$	[19]
$WW \rightarrow 2\ell 2\nu$	CMS	4.92 fb^{-1}	52.4	2.0	4.5	1.2	$47.0^{+2.0}_{-2.0}$	[20]
$WZ \rightarrow 3\ell 1\nu$	ATLAS	4.6 fb^{-1}	19.0	1.4	0.9	0.4	$17.6^{+1.1}_{-1.0}$	[21]
$WZ \rightarrow 3\ell 1\nu$	CMS	4.9 fb^{-1}	20.8	1.3	1.1	0.5	$17.8^{+0.7}_{-0.5}$	[22]
$WZ, WW \rightarrow \ell\nu 2j$	ATLAS	4.6 fb^{-1}	68	7	19	—	$61.1^{+2.2}_{-2.2}$	[23]
$WZ, WW \rightarrow \ell\nu 2j$	CMS	5.0 fb^{-1}	68.9	8.7	9.7	1.5	$65.6^{+2.2}_{-2.2}$	[24]
$ZZ \rightarrow 4\ell + 2\ell 2\nu$	ATLAS	4.6 fb^{-1}	6.7	0.7	0.4	0.3	$5.89^{+0.22}_{-0.18}$	[25]
$ZZ \rightarrow 4\ell$	CMS	5.0 fb^{-1}	6.24	0.86	0.41	0.14	$6.3^{+0.4}_{-0.4}$	[26]
$ZZ \rightarrow 2\ell 2\nu$	CMS	5.1 fb^{-1}	5.1	1.5	1.4	0.2	$6.3^{+0.4}_{-0.4}$	[27]

cross-sections are:

$$\sigma(pp \rightarrow ZZ) = 6.24^{+0.86}_{-0.80} (stat.)^{+0.41}_{-0.32} (syst.) \pm 0.14 (lumi.) \text{ pb (7 TeV)}$$

$$\sigma(pp \rightarrow ZZ) = 7.7 \pm 0.5 (stat.)^{+0.5}_{-0.4} (syst.) \pm 0.4 (theo.) \pm 0.2 (lumi.) \text{ pb (8 TeV)},$$

for both Z bosons in the mass range $60 < m_Z < 120$ GeV. The ATLAS Collaboration measured a total cross-section of

$$\sigma(pp \rightarrow ZZ) = 6.7 \pm 0.7 (stat.)^{+0.4}_{-0.3} (syst.) \pm 0.3 (lumi.) \text{ pb}$$

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

$$\sigma(pp \rightarrow ZZ) = 7.1^{+0.5}_{-0.4} (stat.) \pm 0.3 (syst.) \pm 0.2 (lumi.) \text{ pb},$$

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 [35, 36]. All measurements are in agreement with the corresponding SM predictions.

The analysis presented in this thesis extends the measurement performed in [33], using the final data reprocessing, with state-of-the-art calibrations, alignment and reconstruction techniques, and more precise Monte Carlo generators, to event quantities that have jets as basic observable. The main goal is to measure the differential cross-section according to the number of jets accompanying the two Z bosons, as well as the differential cross-section as a function of many other variables related to jets, like the angular separation of the two energy-leading jets ($\Delta\eta_{jj}$), the invariant mass (m_{jj}) of the di-jet system and their transverse momentum and pseudorapidity.

Very useful information can be extracted from these differential cross-sections. First of all,

Table 3.2: Observed and expected cross-sections for vector boson pair production at $\sqrt{s} = 8$ TeV.

Process	Exp.	$\int \mathcal{L} dt$	σ_{tot}	(stat.)	(syst.)	(lumi.)	σ_{tot}^{pred} [pb]	Ref.
$WW \rightarrow 2\ell 2\nu$	ATLAS	20.3 fb^{-1}	71.4	1.2	5.0	2.2	$59.8^{+1.3}_{-1.1}$	[28]
$WW \rightarrow 2\ell 2\nu$	CMS	19.4 fb^{-1}	60.1	0.9	4.5	1.6	$59.8^{+1.3}_{-1.1}$	[29]
$WZ \rightarrow 3\ell 1\nu$	ATLAS	13 fb^{-1}	20.3	0.8	1.2	0.7	$20.3^{+0.8}_{-0.8}$	[30]
$WZ \rightarrow 3\ell 1\nu$	CMS	19.6 fb^{-1}	24.6	0.8	1.1	1.1	$21.9^{+1.2}_{-0.9}$	[22]
$WZ \rightarrow \ell \nu b \bar{b}$	CMS	18.9 fb^{-1}	30.7	9.3	8.2	1.0	$22.3^{+1.1}_{-1.1}$	[31]
$ZZ \rightarrow 4\ell$	ATLAS	20.3 fb^{-1}	7.1	0.5	0.3	0.2	$7.2^{+0.3}_{-0.2}$	[32]
$ZZ \rightarrow 4\ell$	CMS	19.6 fb^{-1}	7.7	0.5	0.6	0.2	$7.7^{+0.6}_{-0.6}$	[33]
$ZZ \rightarrow 2\ell 2\nu$	CMS	19.6 fb^{-1}	7.2	0.8	1.9	0.2	$7.7^{+0.6}_{-0.6}$	[27]
$ZZ \rightarrow 2\ell(2\nu)b\bar{b}$	CMS	18.9 fb^{-1}	6.5	1.7	1.2	0.2	$7.7^{+0.4}_{-0.4}$	[31]

the cross-section dependency to the jet multiplicity has never been measured so far and can show how well the QCD correction to the ZZ production is understood: through this measurement, the NLO QCD calculations implemented in Monte Carlo samples used in the analysis can be tested, and an estimate of the size of higher-order contributions can be evaluated. The ZZ production is also sensitive to physics beyond the Standard Model: the measurement in the $\Delta\eta_{jj}$ variable will be instrumental for the vector boson scattering analysis in this channel, the main process to probe the electroweak symmetry breaking mechanism and thus the role of the Higgs Boson in the electroweak theory (see Chapter 7). Moreover, the study of m_{jj} will be also the basis for future multi-boson final state searches, in which an additional electroweak boson decays hadronically. The di-jet invariant mass is also a key variable for the investigation of anomalous quartic gauge couplings (aQGC), as explained in Section 1.6.

3.3 Leading and Next-to-Leading Order Cross-Sections

In order to cover the broad scope of SM measurements and search strategies performed at LHC, precise predictions are needed for a wide variety of processes. In the following, an introduction on the definition of leading and next-to-leading order cross-section is reported together with a simple example, in order to explain how very precise theoretical expectations can be obtained to describe experimental data.

The Lorentz invariant scattering amplitude $\mathcal{M}$ is a central element in theoretical particle physics, since allows to calculate the probability for relativistic scattering processes, by connecting initial state particles with well defined momenta to other final state particles, often different from the original ones. Each contribution to $\mathcal{M}$ can be represented in a graphical way by a Feynman diagram, in which particles are described by lines and create or annihilate in vertices. Simple rules are defined in order to extract from the diagram the complex number contributing to the amplitude: each vertex is associated to a vertex factor, propor-

tional to a dimensionless charge or coupling constant g , which represents the strength of the interaction between the particles and the force carrier. This coupling constant is proportional to $\sqrt{\alpha}$, the fine structure constant of the interaction (Table 1.2). In principle, all

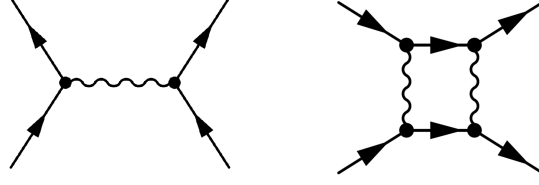


Figure 3.4: Example of *tree-level* scattering diagram (left) and diagram with loop (right). According to the Feynman rules, each vertex is proportional to $\sqrt{\alpha}$.

the infinite contributions of $\mathcal{M}$ have to be calculated and summed up in order to get the total amplitude of a process. However, in certain cases it is possible to add just the simplest diagram(s), making the calculation easier. The simplest *tree level* scattering diagram (Figure 3.4, left) has two vertices and it thus contains two factors of g . The diagrams with loops (Figure 3.4, right), on the other hand, present four vertices and hence four factors of g . In the electromagnetic interactions, the coupling constant $\alpha_{EM} = \frac{g_{EW}^2}{4\pi} \sim \frac{1}{137} \ll 1$ and therefore the more complicated diagrams with more vertices will contribute much less to the amplitude than the simplest ones, since they contain higher powers of α_{EM} . A good approximation to $\mathcal{M}$ can usually be obtained from the sum of the amplitudes for the *Leading Order* (LO) diagrams, those with the smallest power of α_{EM} that make non-zero contributions. Analyzing these terms and regarding the other smaller terms as negligible gives the main behavior of a process, while its true behavior is expected to be only a small deviation away from this. The process of truncating the sum of diagrams is a form of perturbation theory. The coupling constant of the weak force is larger ($\alpha_W \sim \frac{1}{29}$) than α_{EM} , but it appears however weak thanks to its very heavy carriers (the W and Z bosons). On the contrary, the strong force is so-called because is characterized by a coupling constant close to 1, which is about a hundred times larger than the electromagnetic constant. Hence, the contribution to the amplitude from the higher order diagrams, which are proportional to g_s^n with $n > 2$, is compatible with the LO contribution. Each diagram is equally important and thus, having all the terms of $\mathcal{M}$ the same order, the series becomes divergent. Nevertheless, α_s is found to be small in high energy or short distance interactions, conditions reproduced at hadron colliders, allowing perturbation theory to be applied. If the coupling constant is small enough, the infinite number of terms can be approximated accurately by a much more manageable number of contributions, and up to now this results as the most precise approach to test Quantum-Chromodynamics (QCD).

Calculating the contribution of Next-to-Leading Order (NLO) diagrams is therefore extremely important, especially for strong interactions. They allow more accurate estimates of signal and background processes, useful in the search for new physics, and more confidence in predicting cross-sections for precision measurements. A lot of techniques have been developed for this purpose and many programs are available for making NLO calculations, such as generic programs like MCFM [37], which has been employed in this analysis.

3.3.1 A Simple Example

In order to explain how very precise theoretical predictions can be made to describe experimental data, a simple example is reported [38]. At an electron-positron collider, such as

LEP, a common interaction is the production of two jets, which can be represented by a single Feynman diagram at lowest order, reported in Figure 3.5. The spin- and color-

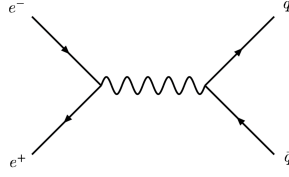


Figure 3.5: Feynman diagram representing the lowest-order production of two jets at an electron-positron collider.

summed squared matrix element of this process, i.e. the squared scattering amplitude, is obtained assuming all the particles are massless and is:

$$\sum |\mathcal{M}|^2 \propto \left(\frac{(p_1 \cdot q_1)^2 + (p_2 \cdot q_2)^2}{(p_1 \cdot p_2)^2} \right), \quad (3.1)$$

where p_i and q_i are the momenta of the colliding electrons and jets produced in the interaction:

$$e^-(p_1) + e^+(p_2) \rightarrow q(q_1) + \bar{q}(q_2).$$

To calculate the effect of NLO QCD corrections, additional contributions proportional to α_s must be evaluated. Two classes of diagrams have to be considered:

- **Real radiation diagrams:** this class corresponds to simple additional gluon emission and it is represented by the two diagrams reported in Figure 3.6. Each additional vertex is proportional to the strong coupling constant g_s , resulting in a contribution of $\alpha_s \propto g_s^2$ when the squared matrix element is evaluated.
- **Virtual corrections:** these diagrams involve the emission of a gluon from a quark and re-absorption on the same or a different quark-line and are shown in Figure 3.7. They contribute as interference with the lowest order diagram.

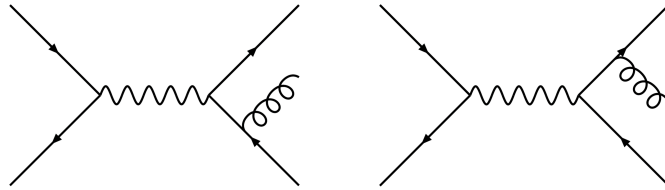


Figure 3.6: Additional gluon emission diagrams contributing to the NLO QCD corrections of the di-jet production.

The NLO cross-section is obtained by summing the real and virtual contribution together. The real contribution is simpler to be calculated, since it is just the lowest order matrix element for $e^+e^- \rightarrow q\bar{q}g$. According to the Feynman rules, the intermediate quark propagator before the gluon emission contributes a factor of (considering $q^2 = k^2 = 0$)

$$\frac{1}{2q \cdot k} = \frac{1}{2E_q E_g (1 - \cos\theta_{qg})},$$

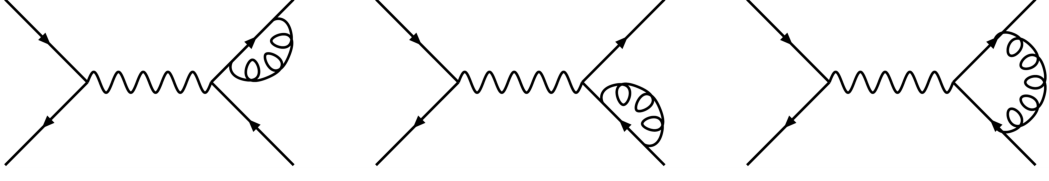


Figure 3.7: Virtual gluon emission diagrams contributing to the NLO QCD corrections of the di-jet production.

where q and E_q are the momentum and energy of the quark (or anti-quark), k and E_g are the gluon momentum and energy and θ_{qg} is the angle between (anti-)quark and gluon. This propagator can vanish in two cases:

- the gluon and the quark are collinear ($\theta_{qg} \rightarrow 0$);
- the gluon is soft ($E_g \rightarrow 0$).

These two problems are called *infrared (IR) singularities* and arise from the fact that particles have been considered massless: adding a mass to the gluon or putting the quarks slightly off-shell would turn these singularities into a logarithmic divergence, more easily manageable, but would complicate the algebra making calculations intractable.

The most common trick to overcome this complication is to use the *dimensional regularization* (DR) technique [39], by introducing the regularization parameter ϵ and expanding the cross-section in power of this deviation. This procedure allows to control all the singularities, which now manifest themselves as factor of $\frac{1}{\epsilon}$. By integrating the matrix element, one obtains:

$$\sigma_{real} \sim \frac{C_F \alpha_s}{2\pi} \left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + \frac{19}{2} + \mathcal{O}(\epsilon) \right) \cdot \sigma_{LO}, \quad (3.2)$$

where $C_F = \frac{4}{3}$ is the color factor.

Soft and collinear singularities generate also from the virtual contributions. Moreover, as the gluon momentum tends to infinite, power counting makes some terms look logarithmically divergent ($\sim \int \frac{dy}{y}$) and this effect is known as *ultraviolet (UV) divergence*. The dimensional regularization can be applied also in the virtual case and leads to a similar expression:

$$\sigma_{virt} \sim \frac{C_F \alpha_s}{2\pi} \left(-\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \mathcal{O}(\epsilon) \right) \cdot \sigma_{LO}. \quad (3.3)$$

Thus, by adding the real and virtual parts and taking $\epsilon \rightarrow 0$, the next-to-leading order cross-section is:

$$\sigma_{NLO} = \left(1 + \frac{\alpha_s}{\pi} \right) \cdot \sigma_{LO}. \quad (3.4)$$

This correction agrees well with very precise data from experiments.

3.3.2 Monte Carlo Generators

In the previous paragraph, a NLO calculation has been analyzed quite easily. Unfortunately, such integrals can only be performed in very simple cases. The complexity of the matrix elements and the phase space for increasing particle multiplicity make the calculations more difficult. Moreover, when experimental cuts are enforced, integrals become even harder. For these reasons, flexible tools have been developed to perform the NLO calculation numerically and in a general manner, so that any desired experimental cuts can

be applied. Although the procedure is more complex than the example reported above, the basic idea of a NLO Monte Carlo is the same: calculate the virtual and real contributions of the process and add them together. Since, as already pointed out, each contribution is separately divergent, the solution is to render the integrations finite in some way and this is made possible, for example, by the factorization properties of QCD matrix elements in soft and collinear limits.

At hadron collider, the situation becomes even more complicated: the two colliding quarks can radiate gluons in the initial state, making the redefinition of the Parton Density Functions (PDF) at NLO necessary, and other additional processes can enter, making the cross-section sensitive to the gluon PDF. A lot of effort has been employed in order to address these challenging issues and provide NLO corrections for a very large number of processes at LHC.

3.4 Current Status of Calculations

Currently, NLO available calculations included in Monte Carlo simulations cover a wide range of physics processes, starting with two particles annihilation to a maximum of five final state objects. The theoretical efforts for a precise prediction of vector boson pair production in the Standard Model began more than 20 years ago, with the first NLO QCD calculations with stable vectors [40, 41, 42, 43, 44, 45]. Since then, great improvements have been achieved, obtaining NLO QCD and electroweak corrections available for all di-boson processes, including correlations and off-shell effects [46, 47, 48].

Recently, new approaches in subtraction methods and algorithms have lead to further progresses in theoretical calculations, allowing the evaluation of next-to-NLO (NNLO) corrections in almost all di-boson processes. However, their implementation in Monte Carlo generators, useful for experimental analysis, has just started and it is not available for all cases.

3.4.1 QCD and EW Corrections at NLO

The main mechanism to produce two on-shell massive gauge bosons at the LHC is via quark anti-quark annihilations, as shown in Figure 3.8 (top). For the case of W^+W^- and ZZ , a contribution comes also from the one-loop gluon fusion processes, which are formally NNLO diagrams, as reported in Figure 3.8 (bottom). Moreover, the $\gamma\gamma \rightarrow W^+W^-$ process occurs at tree level and includes the $\gamma\gamma W^+W^-$ coupling: it is very small at the total cross-section level, but increases with the WW invariant mass and must thus be taken into account (Figure 3.8, center).

The NLO QCD contribution includes the virtual loop corrections with one gluon in the loops and the real emission corrections with one additional parton in the final state. The real emission contributions can be classified into two groups: gluon-quark radiated processes ($q\bar{q} \rightarrow VV'g$), where the gluon is radiated from an initial (anti-)quark, and gluon-quark induced processes ($qg \rightarrow VV'q'$), which are related to gluon-quark radiated processes via crossing-symmetry. Both the virtual and real corrections are separately IR divergent. In order to overcome these divergences, the dimensional regularization method is used.

The NLO EW corrections are also divided into similar sub-contributions as in the QCD evaluation (with photons instead of gluons), but in this case loop corrections contain also UV divergences. These divergences are regularized using the DR method and by the renormalization of the relevant EW parameters, namely M_W , M_Z and the fine-structure constant

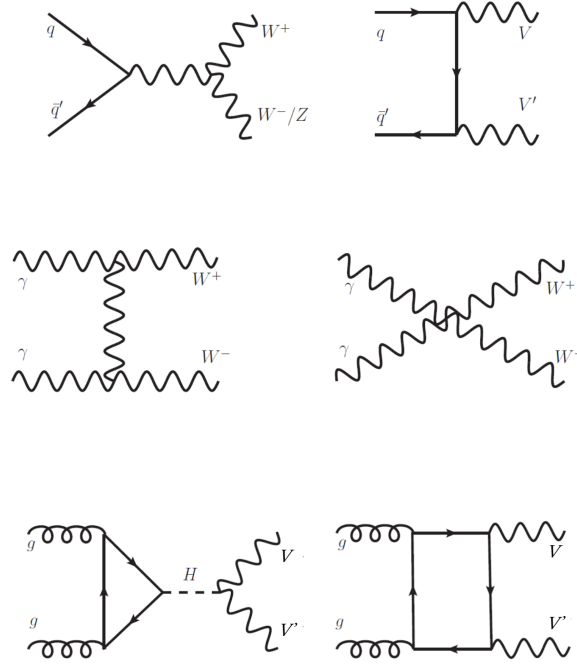


Figure 3.8: Representative tree-level diagrams for VV' production processes via quark anti-quark annihilation (top), $\gamma\gamma \rightarrow W^+W^-$ (center) and gluon-gluon fusion (bottom) processes.

α . Real and virtual corrections are combined together, using the *mass regularization* (MR) method, which introduces a common mass regulator for the light fermions and a fictitious photon mass to regularize IR divergences. For the same purpose, the DR scheme is also employed.

In the following, the full next-to-leading order predictions for the production cross-sections and distributions of on-shell massive gauge boson pair production (W^+W^- , $W^\pm Z$ and ZZ) in the Standard Model is reported [48]. The invariant mass distribution of the gauge boson pair is displayed at $\sqrt{s} = 14$ TeV, designed center-of-mass energy of the LHC.

W^+W^- Productions

In Figure 3.9 (top-left) the differential distribution of the invariant mass M_{WW} of the W boson pair is shown, calculated at LO and NLO including QCD and EW corrections. Separately, the gg contribution is also reported. Moreover, in Figure 3.9 the M_{WW} distribution of the QCD (top-right) and EW (bottom) corrections, relative to the LO $q\bar{q}$ results, is displayed for the various sub-corrections. It can be noticed that the gluon-quark induced (in blue) and the gluon-quark radiated (in green) processes compensate each other and the full QCD corrections are dominated by the virtual corrections (Figure 3.9, top-right). The gluon fusion mechanism amounts to $\sim +5\%$. Turning to EW corrections (Figure 3.9, bottom), it can be observed that the virtual EW contributions (represented by the dotted red curve) are compensated by the effect of photon-quark induced processes (displayed in dotted-dashed blue). The total EW corrections to the $q\bar{q}$ sub-processes drop thus very fast from $+2\%$ at the threshold to small variations around zero. The $\gamma\gamma$ sub-process is not negligible and increases up to $+6\%$ at about 700 GeV.

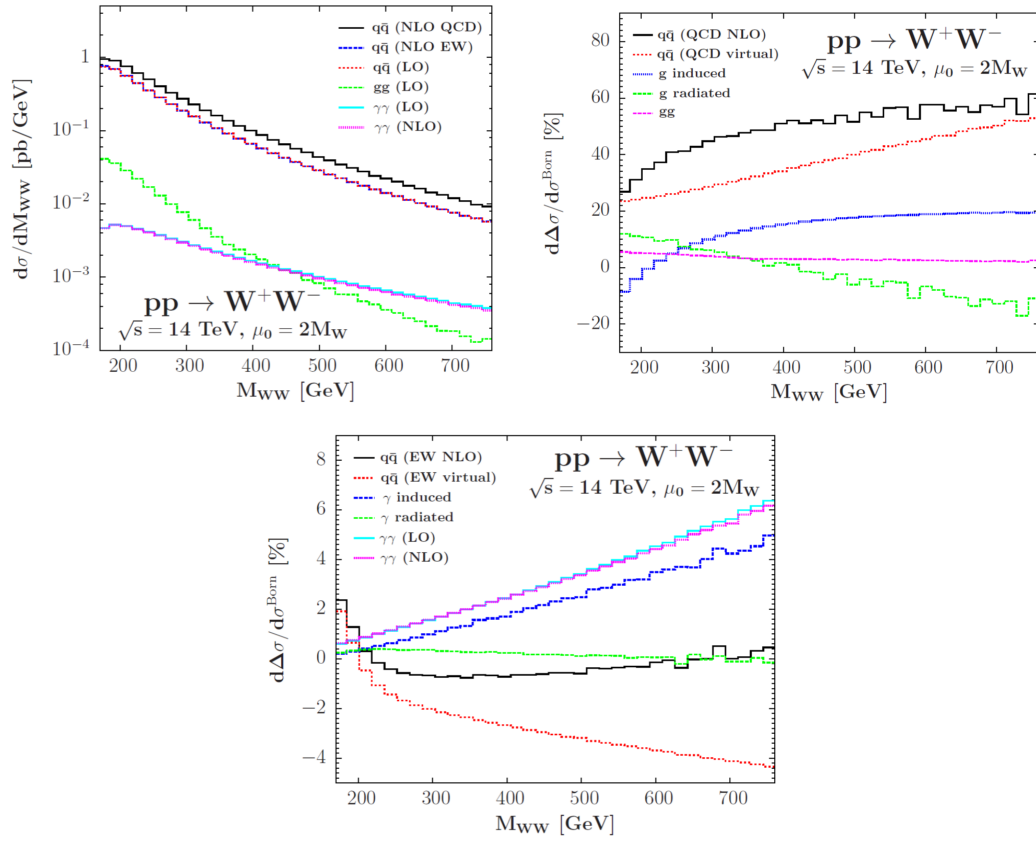


Figure 3.9: Invariant mass M_{WW} distribution of the $pp \rightarrow WW$ cross-section at the LHC (top-left) and its corrections of the NLO QCD (top-right) and EW (bottom) [48].

$W^\pm Z$ Production

The invariant mass distributions in W^+Z and W^-Z channels are similar and their behavior is displayed in Figure 3.10. The QCD corrections (Figure 3.10, middle) are of the order of +60% (with some fluctuations up to +70% near the threshold). Furthermore, EW corrections are very small (Figure 3.10, bottom), $\sim +1\%$ in W^+Z channel for $M_{WZ} \geq 500$ GeV (and close to zero just after the threshold) and $\sim +2\%$ in W^-Z at $M_{WZ} \sim 700$ GeV. A compensation of the virtual part (in dotted red) can be observed, giving a decreasing virtual corrections down to -4% in both channels, to be added to the positive correction coming from the photon-quark induced processes (dotted dashed blue) up to $\sim +4\%$ in W^+Z channel and $\sim +5\%$ in W^-Z channel.

ZZ Production

The trend of the M_{ZZ} distribution of the QCD corrections, reported in Figure 3.11, is very similar to the one seen for the WW production: the gluon-quark induced (in blue) and the gluon-quark radiated (in green) processes compensate each other and the full QCD corrections are dominated by the virtual corrections (Figure 3.11, top-right). The gluon fusion mechanism amounts to $\sim +10\%$. As regards EW corrections, it can be observed that the total corrections are dominated by the virtual EW contribution and drop to -8% at $M_{ZZ} = 700$ GeV (Figure 3.11, bottom). Comparing to the $W^\pm Z$ cases reported above, the virtual correction is significantly more negative. This suggests that there are less cancellations between negative and positive corrections in the ZZ case.

Going beyond final states with two particles, NLO QCD corrections have been computed for the production of three weak vector bosons, the production of a weak boson in association with up to three jets and weak boson pair production in vector boson fusion.

The production of weak boson pairs with one additional jet at NLO is also of particular interest for the presented analysis. This process is interesting in its own right, due to the enhanced jet activity, particularly at the LHC. It also provides the real-virtual contribution to the NNLO QCD corrections to weak boson pair production. The production of W -boson pairs with an additional jet has been calculated at NLO without [49] and with [50, 51] decays. Given the central subject of this analysis, ZZ events produced in association with jets, the $\mathcal{O}(\alpha_s)$ corrections to Z boson pair production with an additional jet are now summarized [52]. The sub-processes which contribute to the $ZZ + \text{jet}$ production at hadron colliders are:

$$q\bar{q} \rightarrow ZZg, \quad qg \rightarrow ZZq, \quad \bar{q}g \rightarrow ZZ\bar{q},$$

where q is an up- or down-type quark and g is a gluon. Representative LO diagrams for the first sub-process are shown in Figure 3.12. The produced weak bosons are identical for ZZ production and lead to a significant amount of Feynman diagrams, making the calculations very difficult. At $\mathcal{O}(\alpha_s)$ the most complicated loop topologies are graphs derived from the tree-level diagrams via virtual gluon exchange and box graphs derived by closing the quark line in the tree-level diagrams and attaching a $gq\bar{q}$ current (representative one-loop graphs for the process $q\bar{q} \rightarrow ZZg$ are displayed in Figure 3.13). On the other hand, real corrections channel for $ZZ + \text{jet}$ production are related to the emission of an additional gluon or quark. While all virtual corrections channels are already active at LO, new real corrections channels open up at NLO, namely the $gg, qq', q\bar{q}'$ ($q' \neq q$) and $\bar{q}\bar{q}'$ channels. In Figure 3.14, LO and NLO predictions for $ZZ + \text{jet}$ production cross-section at the LHC ($\sqrt{s} = 14$ TeV) are reported. It can be noticed that the shape of the cross-section variation is qualitatively unchanged when going from LO to NLO, because the new channels becoming active at

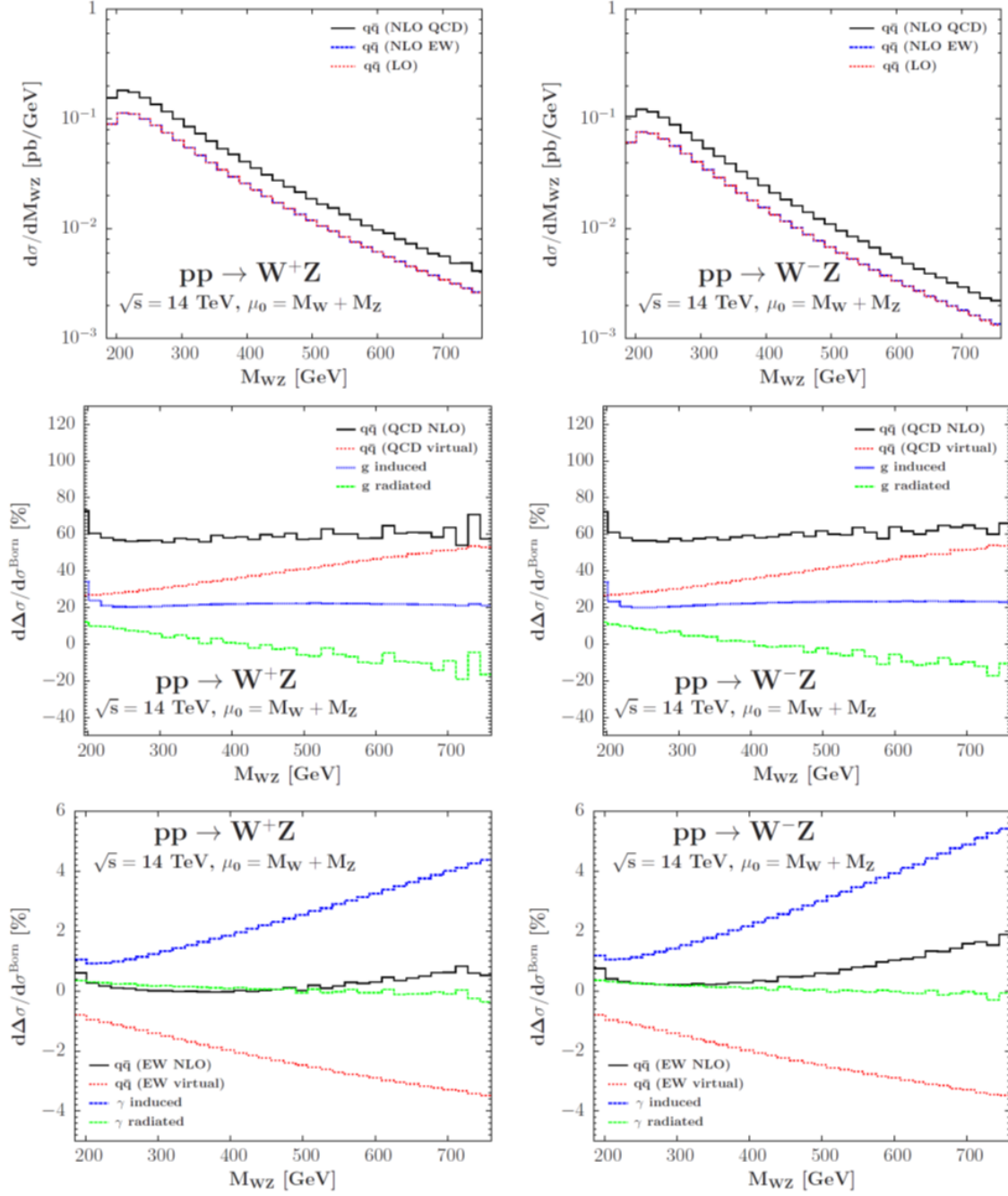


Figure 3.10: Invariant mass M_{WZ} distribution of the $pp \rightarrow W^\pm Z$ cross section at the LHC (top) and its corrections of the NLO QCD (center) and EW (bottom) [48].

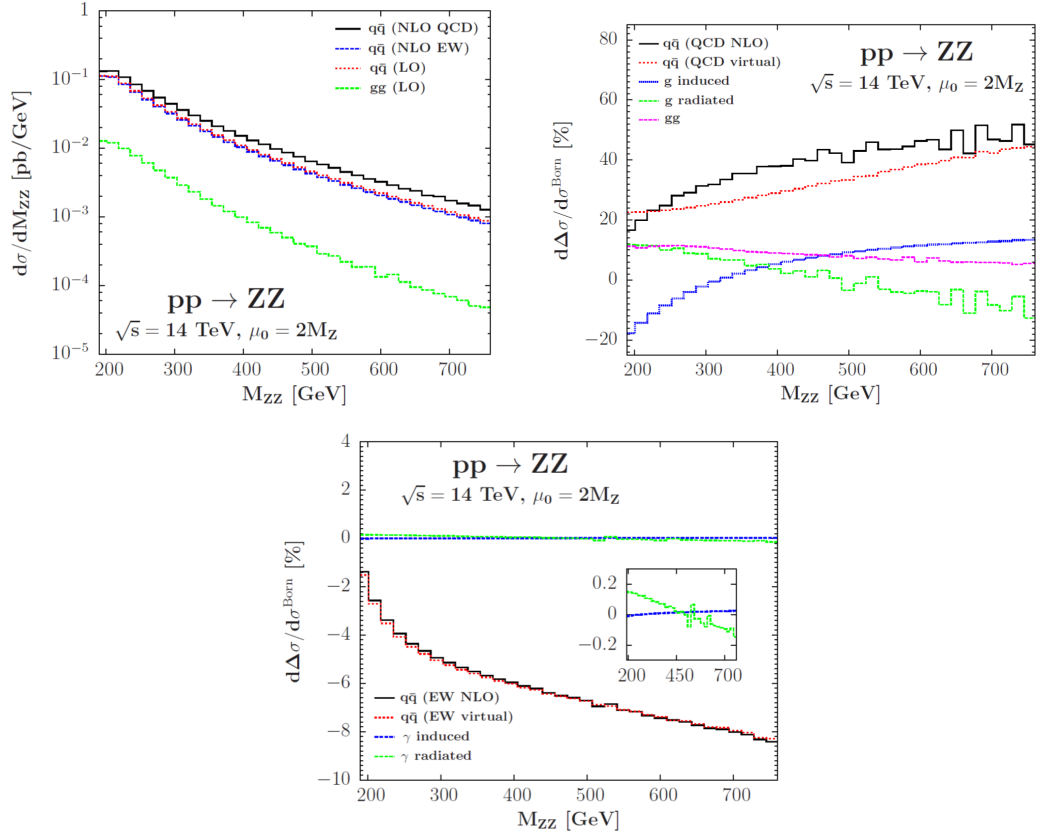


Figure 3.11: Invariant mass M_{ZZ} distribution of the $pp \rightarrow ZZ$ cross section at the LHC (top-left) and its corrections of the NLO QCD (top-right) and EW (bottom) [48].

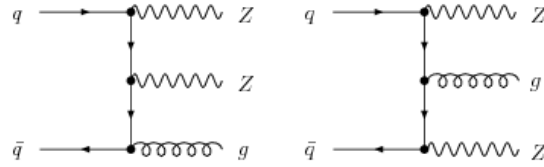


Figure 3.12: Representative LO graphs for the partonic process $q\bar{q} \rightarrow ZZg$ [52].

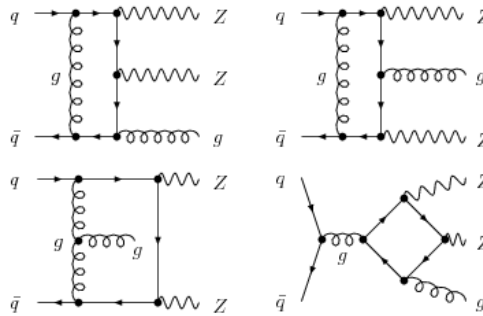


Figure 3.13: Representative one-loop graphs for the partonic process $q\bar{q} \rightarrow ZZg$ [52].

NLO have a sizable impact at the LHC energies, due to parton density being probed in different x regions. The K -factor, i.e. the ratio between the NLO contributions and the LO cross-section, does not vary significantly, ranging from 0.8 to ~ 1.5 . In red an exclusive NLO cross-section by vetoing 2 jet events with a second hardest jet with $p_T > 50$ GeV is also reported.

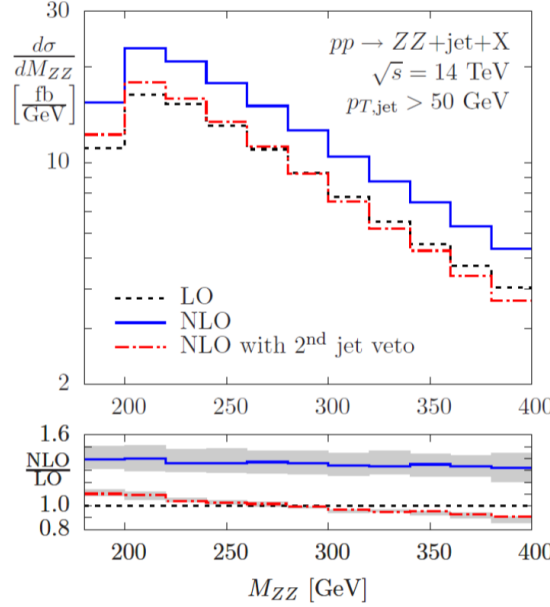


Figure 3.14: ZZ invariant mass distribution for $ZZ + \text{jet}$ production at the LHC. The differential K -factor is also shown [52].

3.4.2 Corrections at Higher Orders

In some cases, the NLO accuracy is not sufficient and more precise corrections are needed. It is already possible to extend the calculation beyond the NLO in few processes, requiring the resummation of all the Feynman diagrams at the α_s^2 order, the Next-to-Next-to-Leading Order (NNLO). From the computational point of view, it is extremely challenging to compute the exponential growth of all the diagrams involving more and more loops. The NNLO QCD computation of VV' production [53] requires the evaluation of the tree-level scattering amplitudes with two additional (unresolved) partons, of the one-loop amplitudes with one additional parton, and of the one-loop-squared and two-loop corrections to the Born sub-process $q\bar{q} \rightarrow VV'$. Up to now, the bottleneck for the NNLO corrections has been the knowledge of the two-loop amplitudes. Although now all the relevant amplitudes have been calculated, the computation of the NNLO corrections is still a non-trivial task, mainly because of the IR singularities which prevent a straightforward application of numerical techniques. In order to handle and cancel these divergences at NNLO, the q_T -subtraction formalism is used [54]. In the following, the results of the NNLO QCD corrections to the inclusive production of W and Z boson pairs are summarized [55, 56], in order to show the differences with respect to the NLO predictions.

W^+W^- Production

Figure 3.15 (left) shows the LO, NLO and NNLO predictions for inclusive W^+W^- production at LHC, with $\sqrt{s}$ ranging from 7 to 14 TeV. The higher-order QCD corrections to W^+W^- production include partonic channels with b -quarks in the final state, which lead to a subtle interplay between W^+W^- and top production processes. It can be demonstrated that this huge top contamination can be subtracted without significant loss of theoretical precision. From the plotted distribution it can be noticed that at 7 (14) TeV the LO predictions receive a positive NLO shift of 53 (58)%, and the NNLO corrections induce a further enhancement of 9 (12)%, with an estimated 3% residual uncertainty from missing contributions beyond NNLO. The NLO result supplemented with the loop-induced gluon fusion contribution (“NLO+gg”) is also shown for comparison. Gluon fusion amounts to about 35% of the total NNLO contribution. The lower panel reports the NNLO and NLO+gg predictions normalized to the NLO result. Comparing the theoretical predictions to CMS and ATLAS results [19, 20, 28, 29], it can be seen that the inclusion of the newly computed NNLO corrections provides an excellent description of recent measurements of the W^+W^- cross-section at 7 TeV and diminishes the significance of an observed excess at 8 TeV.

ZZ Production

The LO, NLO and NNLO cross-sections for the ZZ production is reported in Figure 3.15 (right) as a function of $\sqrt{s}$, ranging from 7 to 14 TeV. It can be observed that the NLO corrections increase the LO result by $\sim 45\%$, while the effect of NNLO corrections with respect to the NLO calculations ranges from 11% at $\sqrt{s} = 7$ TeV to 17% at $\sqrt{s} = 14$ TeV. The gluon fusion contribution provides $\sim 60\%$ of the full NNLO corrections. The theoretical predictions can be compared to the ATLAS and CMS measurements [25, 26, 32, 33], carried out at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV, which are also displayed in the plot. As it can be seen, the experimental uncertainties are still relatively large. Even if the ATLAS and CMS results look compatible with both the NLO and NNLO predictions, the comparison between the theoretical expectations must be interpreted with care: the LHC experiments obtain their ZZ production cross-section without including contributions from far off-shell Z bosons and theoretical calculations don’t consider EW corrections, which are expected to provide a negative contribution to the inclusive cross-section.

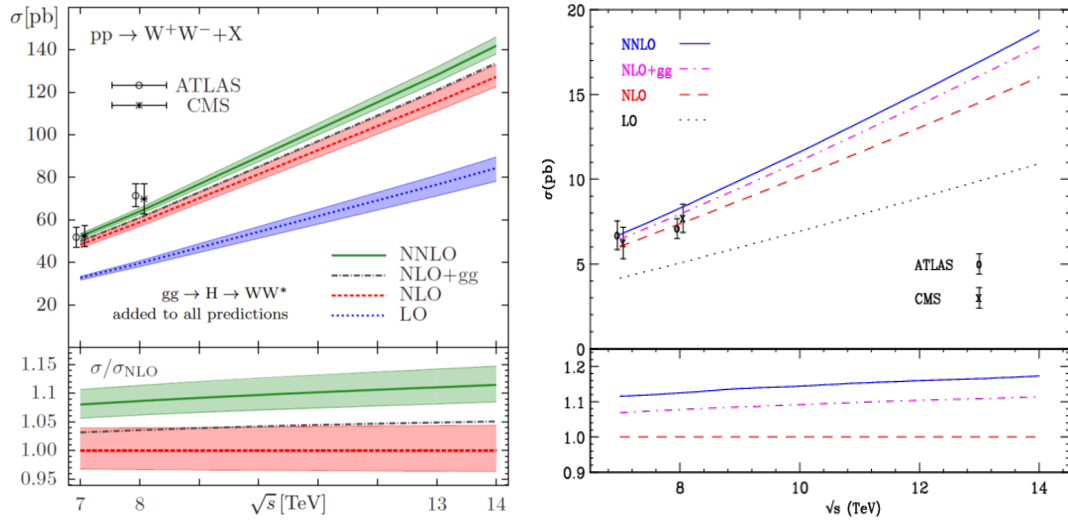


Figure 3.15: The on-shell W^+W^- (left) and ZZ (right) cross-sections at LO (dots), NLO (dashes), NLO+gg (dot dashes) and NNLO (solid) are compared to ATLAS and CMS measurements. In the lower panel NNLO and NLO+gg results are normalized to NLO predictions. In the left-plot the $gg \rightarrow H \rightarrow WW^*$ is also included and the bands describe scale variations [55, 56].

Chapter 4

Physics Objects Reconstruction

In order to identify Z boson pairs produced in association with jets, a detailed knowledge of the physics objects characterizing these processes is mandatory. The reconstruction of the $ZZ \rightarrow 4\ell$ events requires high performance lepton reconstruction, identification and isolation as well as excellent lepton energy-momentum measurements. Moreover, being jets a central element of the analysis, their precise reconstruction and correct energy-momentum evaluation are also crucial aspects.

This Chapter is dedicated to the electrons, muons, photons and jets reconstruction with the CMS Experiment. In the first part, a description of the particle-flow technique, the most used and performing algorithm for object reconstruction in CMS, is reported. In the second part, the reconstruction and identification criteria for leptons, photons and jets are presented.

4.1 The Particle-Flow Algorithm

A complete reconstruction of the individual particles emerging from each collision event is obtained via a Particle-Flow (PF) technique [57]. The particle-flow event reconstruction aims at reconstructing and identifying all stable particles in the event, i.e. electrons, muons, photons, charged hadrons and neutral hadrons, with an accurate combination of all CMS sub-detectors towards an optimal determination of their direction, energy and type (Figure 4.1). This list of individual particles is then used, as if it came from a MC event generator, to build jets, to determine the missing transverse energy E_{miss}^T (which gives an estimate of the direction and energy of the neutrinos and other invisible particles), to reconstruct and identify taus from their decay products, to quantify charged lepton isolation with respect to other particles. The CMS PF algorithm relies on an efficient and pure track reconstruction, on a clustering algorithm able to disentangle overlapping showers within the calorimeters, and on an efficient link procedure to connect together the deposits of each particle in the sub-detectors.

4.1.1 Track Reconstruction

The CMS tracking system provides a precise measurement of the charged particle direction at the production vertex. The trajectory is defined as the sum of all the hits inside the silicon and pixel detectors that represent the path followed by the particle coming from the pp collision. Particle tracks are reconstructed adopting an iterative strategy. The hits

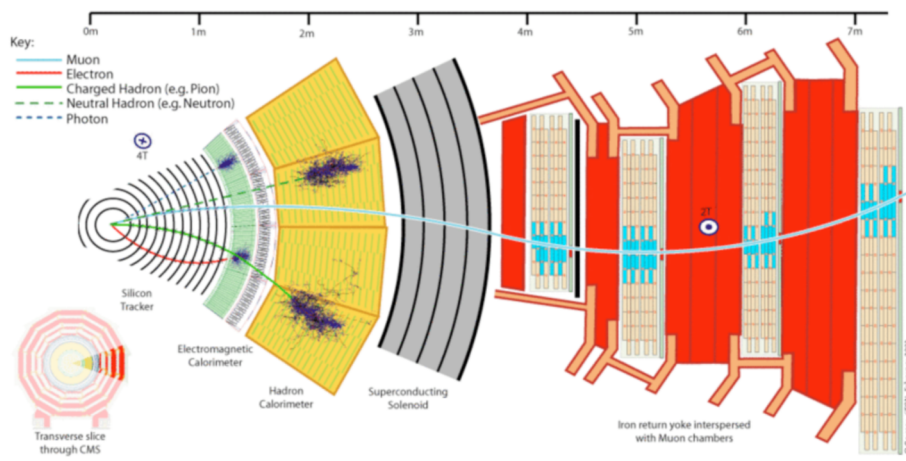


Figure 4.1: Transverse slice view of the CMS sub-detectors energy measurement for different particles. The particle-flow algorithm combines all the information given by each detector to build the physics object.

collected by the tracking system are selected if they satisfy very tight criteria. These hits unambiguously assigned to the tracks are then removed from the collection and an iterative fit technique is applied: the statistical compatibility between the remaining hits and the track is estimated (taking into account the energy loss in the material and the magnetic field) and track selecting criteria are progressively loosened. The fit procedure finally extracts the first estimate of the particle transverse momentum and pseudorapidity. After this selection, the tracking fake rate is estimated to be at the order of a per cent, while the tracking efficiency is greater than 90% as it can be seen in Figure 4.2.

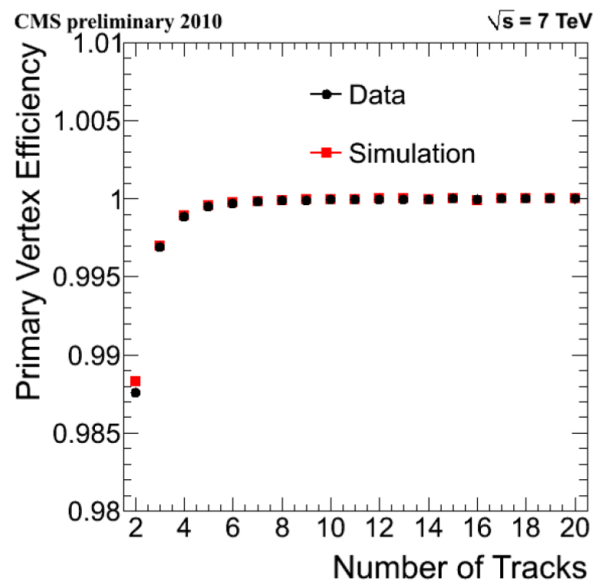


Figure 4.2: Tracking efficiency as a function of the number of tracks measured in CMS and compared to simulation (7 TeV) [58].

4.1.2 Calorimeter Clustering Reconstruction

The calorimetric information is indispensable for the measurement and identification of electromagnetic particles (electrons and photons) through their energy loss in matter as well as charged and neutral hadrons. The algorithm is based on the clustering of the energy deposits released by the different interacting particles inside the ECAL and HCAL calorimeters and in the pre-shower detector (PS) in the endcaps. It aims to detect and measure the energy and direction of stable neutral particles (photons and neutral hadrons), separate these neutral particles from energy deposits from charged hadrons, reconstruct and identify electrons and all accompanying Bremsstrahlung photons, and help the energy measurement of charged hadrons for which the track parameters were not determined accurately (low-quality, or high- p_T tracks). The algorithm is performed separately in each sub-detector (ECAL barrel, ECAL endcap, HCAL barrel, HCAL endcap, PS first layer and PS second layer). Clusters are constructed if they have a minimum energy threshold and in this case they are labeled as *cluster seeds*. Clusters are then grown from the seeds by aggregating cells with at least one side in common with a cell already in the cluster and with an energy that is found to be 2σ away from the electronics noise in the calorimeters. These new objects are called *particle-flow seeds* and are used as one of the inputs to the linking algorithm in order to build the different particles.

4.1.3 The Linking Algorithm

A given particle is, in general, expected to give rise to several particle-flow elements in the various CMS sub-detectors: one charged-particle track, and/or several calorimeter clusters, and/or one muon track. These elements are connected to each other by a linking algorithm in order to fully reconstruct each single particle, avoiding any possible double counting from different detectors. The algorithm is performed for each pair of elements and leads to new reconstructed objects called *blocks*. Thanks to the granularity of the CMS detectors, blocks typically contain only one, two or three elements, and constitute simple inputs for the particle reconstruction and identification algorithm. A link between a charged particle track and a calorimeter cluster proceeds by extrapolating the track from its last measured hit in the tracker to the two layers of the PS, the ECAL and the HCAL. The track is linked to any given cluster if the extrapolated position in the corresponding calorimeter is within the cluster boundaries. Similarly, a link between two calorimeter clusters (either between an HCAL and an ECAL cluster, or between an ECAL and a PS cluster) is established when the cluster position in the more granular calorimeter (PS or ECAL) is within the cluster envelope in the less granular calorimeter (ECAL or HCAL). Finally, a link between a charged particle track in the tracker and a muon track in the muon system is established (and is called a *global muon*) when a global fit between the two tracks returns an acceptable χ^2 . The reconstruction and identification of a set of particles from each block of elements is finally performed by the particle-flow algorithm.

4.2 Leptons

The reconstruction of a ZZ boson pair imposes high performance lepton reconstruction, identification, isolation and excellent lepton energy-momentum measurements. The identification of isolated leptons emerging from the event primary vertex cuts drastically QCD-induced sources of misidentified leptons or non-prompt leptons coming from hadron decays (*fake leptons*). The precise momentum measurement leads to a precise measurement

of $m_{4\ell}$, one of the most interesting observables for the Higgs boson search, in which non-resonant $pp \rightarrow ZZ$ processes represent the main background.

4.2.1 Electron Reconstruction and Identification

The electron reconstruction in CMS [59] combines the information of ECAL and tracker measurements. Electron candidates are reconstructed from clusters of energy deposits in ECAL, which are then matched to tracks reconstructed in the silicon tracker. The energy deposit in ECAL is measured in arrays of energy clusters in the ϕ direction (called *super-clusters*), which collect Bremsstrahlung photons emitted in the tracker and part of the collinear final state radiation (FSR). Super-clusters are also used to identify hits in the track layers in order to start the reconstruction of electron tracks. This procedure is complemented by a tracker-driven approach allowing to improve the reconstruction efficiency at low p_T , in particular in the transition between the ECAL barrel (EB) and the ECAL endcap (EE). The track-cluster matching is performed either *outside-in* from energy cluster measurements, or *inside-out* from track reconstruction. Trajectories in the tracker volume are reconstructed using a dedicated modeling of the electron energy loss due to Bremsstrahlung and fitted with a Gaussian Sum Filter algorithm. In order to preserve the highest possible efficiency and at the same time remove part of the background, electron candidates are selected by applying loose requirements on the track–super-cluster matching. Additional cuts are also employed to reject electrons coming from photon conversion inside the tracking system.

The electron momentum is extrapolated combining the track momentum estimate and the super-cluster energy (corrected for several effects) by using a multivariate regression function. Its precision is dominated by the tracker for $p_T < 15$ GeV, while at higher momenta the ECAL measurement is predominant.

For the physics analysis, electron candidates are required to have a transverse momentum p_T^e larger than 7 GeV and a reconstructed geometrical acceptance $|\eta^e| < 2.5$. Integrated over the acceptance, the reconstruction efficiency for basic electron objects steeply rises to reach 90% at $p_T^e = 10$ GeV and then more slowly to reach a plateau of 95% for $p_T^e > 30$ GeV.

Electron identification relies on a multivariate technique that combines observables sensitive to the amount of Bremsstrahlung along the electron trajectory, the geometrical and momentum matching between the electron trajectory and associated clusters, as well as shower-shape observables. The multivariate identification is trained using a Higgs boson Monte Carlo sample for the signal and a $W + 1$ jet data sample for background and it is optimized using a $Z + 1$ jet data sample.

4.2.2 Muon Reconstruction and Identification

The robust detection of muons over the full acceptance of the CMS detector in a condition of very high rate and background is obtained with a muon system, that allows an efficient and pure identification of muons, and with the inner tracker, that provides a very precise measurement of their properties (as seen in Chapter 2).

Muons are required to have $p_T^\mu > 5$ GeV and $|\eta^\mu| < 2.4$. Their tracks are first reconstructed independently in the inner tracker (*tracker tracks*) and in the muon system (*standalone-muon tracks*) and then the information are combined, using two different approaches. The matching between the inner and outer tracks is initiated either *outside-in* (*Global Muon reconstruction*), starting from a track in the muon system, or *inside-out* (*Tracker Muon reconstruction*),

starting from a track in the silicon tracker. The second method is more efficient than the *Global Muon reconstruction* at low momenta ($p_T \leq 5$ GeV), because it requires only a single muon segment in the muon system, whereas *Global Muon reconstruction* is designed to have high efficiency for muons penetrating through more than one muon station and typically requires segments in at least two muon stations.

A given physics analysis can achieve the desired balance between identification efficiency and purity by applying a selection based on various muon identification variables. For this analysis the *particle-flow Muon selection* is chosen. To identify particle-flow muons, a selection is performed on all the muon candidates reconstructed by applying minimal requirements on the track components in the muon system and takes into account a matching with small energy deposits in the calorimeters.

The resolution for the muons used in the analysis is dominated by the multiple Coulomb scattering in the detector material and is about 2% in the barrel and 6% in the endcaps.

4.2.3 Lepton Impact Parameter and Isolation

The electron or muon pairs from Z decays should originate from the primary vertex. This is ensured by requiring that the significance of the impact parameter to the event vertex, called SIP_{3D} , satisfies $|SIP_{3D} = \frac{IP}{\sigma_{IP}}| < 4$ for each lepton. The IP is the lepton impact parameter in three dimensions, i.e. the lepton distance at the point of closest approach with respect to the primary interaction vertex, and σ_{IP} is the associated uncertainty.

Lepton candidates originating from Z decays are required to be well isolated from other particles, unlike leptons coming from other processes, which are typically close to jets. The isolation of individual leptons is measured relative to their transverse momentum p_T^ℓ , by summing over charged and neutral particles in a cone $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ around the lepton direction at the interaction vertex:

$$R_{iso}^\ell \equiv \frac{\left(\sum p_T^{charged} + \text{MAX} \left[0, \sum p_T^{neutral} + \sum p_T^\gamma - p_T^{PU} \right] \right)}{p_T^\ell}.$$

The $\sum p_T^{charged}$ is the scalar sum of the transverse momenta of charged hadrons originating from the primary vertex, without consider the lepton momentum itself. The primary vertex is chosen as the vertex with the highest sum of p_T^2 of its constituent tracks. The $\sum p_T^{neutral}$ and $\sum p_T^\gamma$ are the scalar sums of the transverse momenta for neutral hadrons and photons, respectively. The latter excludes photons that are candidates for final-state radiation (FSR) from the lepton (see Subsection 4.3.3).

Isolation variables are among the most pile-up sensitive variables in this analysis. Pile-up causes the mean energy deposited in the detector to increase, leading to a shift in the isolation distribution for higher values. Thus, the efficiency of a cut on isolation variables strongly depends on pile-up conditions. In order to have a pile-up robust analysis, the isolation variable has to be corrected using the term p_T^{PU} , representing the contribution of neutral hadrons coming from pile-up interactions. For electrons it is defined as

$$p_T^{PU} \equiv \rho \times A_{eff},$$

where ρ is the median of the energy-density distribution around *jets* (any PF jet in the event having $p_T^{jet} > 3$ GeV). A_{eff} is the *effective area* defined as the geometrical area of the isolation

cone corrected for the different PU depositions as a function of the electron pseudorapidity. Effective areas A_{eff} increase with the η of the lepton and this is mainly due to increasing pile-up at large pseudorapidity and with the fact that the geometrical isolation cone gets truncated at the end of the tracker acceptance.

For muons the contribution to the isolation from pile-up interactions is

$$p_T^{PU} \equiv \sum_i 0.5 \times p_T^{PU,i},$$

where the sum is extended to all the charged hadrons originating from PU vertices and 0.5 is to correct for the different fraction of charged and neutral hadrons in the isolation cone. The electrons or muons are considered isolated in this analysis if $R_{Iso} < 0.4$.

4.2.4 Leptons Efficiency

The efficiencies for reconstruction, identification and selection (SIP_{3D} and isolation) for electrons and muons is measured with a data-driven method based on a selection of events of inclusive single Z production. The *tag-and-probe technique* [60] combines the requirements of a vertex from a pair of basic objects (e.g. tracks for muons or clusters of calorimetry cells for electrons) with a tight lepton selection applied on one leg (the *tag*), in order to ensure sufficient purity. The other leg (the *probe*) is used to measure the efficiency of a given reconstruction algorithm or identification criterion. Since no selection is applied on the probe, its efficiency is unbiased with respect to this analysis selection criteria (Figure 4.3). The efficiency is defined as the ratio of the number of probes passing a certain cut to the

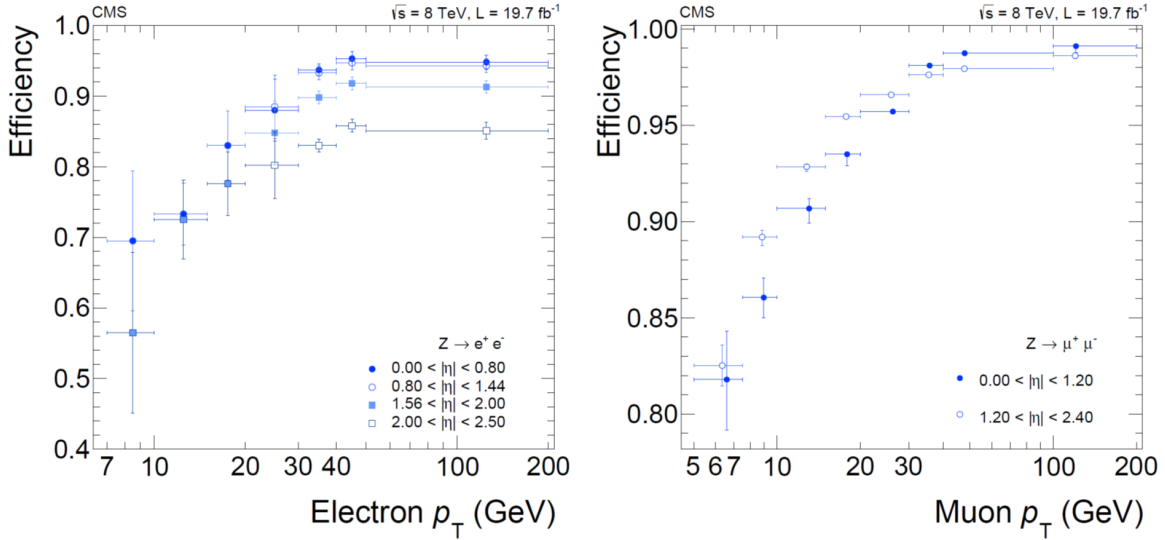


Figure 4.3: Efficiency for the reconstruction, identification and selection of electrons (left) and muons (right) as a function of the lepton p_T , for the 8 TeV data. [61, 62].

total number of probes. The yields are computed by means of a fit to the mass spectrum around the resonance.

By using appropriate definitions for probes, the overall efficiency per lepton is factorized in a series of terms, that can be measured independently:

$$\epsilon = \epsilon_{RECO} \times \epsilon_{ID|RECO} \times \epsilon_{SIP_{3D}|ID} \times \epsilon_{ISO|SIP_{3D}},$$

where each term represents the efficiency for the probe to pass a certain selection or reconstruction step, given that it passes the criteria for the previous one:

- ϵ_{RECO} is the efficiency to reconstruct a track in the inner detector and in the ECAL or muon system;
- $\epsilon_{ID|RECO}$ is the efficiency of the particle-flow reconstruction and identification for leptons that have been successfully reconstructed in the previous step;
- $\epsilon_{SIP_{3D}|ID}$ is the efficiency of the impact parameter requirement, for leptons passing the identification;
- $\epsilon_{ISO|SIP_{3D}}$ is the efficiency of the isolation requirement, for leptons passing all previous selection criteria.

By applying the method to both data and simulation one can derive data to simulation scale factors, used to either correct the signal efficiency in the simulation or to provide systematic uncertainties. In general data and MC distributions agree at the level of 3% or better, with the exception of electron in the range $p_T < 15$ GeV. The muon scale factors consider also the tracking efficiency, which is always almost 1, except for the highest PU data-taking periods. Scale factors and their associated uncertainties are the same as in [17].

4.3 Photons

A Z decay into a lepton pair can be accompanied by final state radiation (FSR), $Z \rightarrow \ell^+ \ell^- \gamma$. If the photon transverse momentum p_T^γ is required to exceed 2 GeV, about 8% (15%) of the decays into muons (electrons) are affected. As the photon emission is most often collinear with one of the leptons, electron measured energies automatically include the energy of a large fraction of the emitted photons in the associated electromagnetic super-cluster. On the other hand, muon measured momenta do not include the emitted photons, absorbed by the calorimeter. Final state radiation is therefore expected to degrade the Z mass resolution when measured with the sole muon pair and in turn degrade the ZZ di-boson mass resolution when measured with the four-lepton momenta, especially in the 4μ and in the $2e2\mu$ final states and, to a lesser extent, in the $4e$ final state. It is also expected to reduce the efficiency of the lepton isolation cut when the emitted photon is in the lepton isolation cone.

Both an excellent di-boson mass resolution and a large selection efficiency are essential ingredients in view of the small production cross section in the three channels. Therefore it is necessary to recover the FSR photons with large efficiency and purity, to remove the energy of the recovered photons from the lepton isolation cones and to measure the mass of the ZZ candidate from the momenta of the leptons and the recovered photons. In addition to being collinear with the leptons, final state radiation also tends to favor low energy photon emission.

4.3.1 Photon Reconstruction and Identification

Photons are reconstructed and identified using the particle-flow reconstruction with a specific clustering algorithm. The determination of the photons energy and direction is monitored in the data with $\pi^0 \rightarrow \gamma\gamma$ decays and is shown to be accurate, reliable and in agreement with the predictions from simulation.

4.3.2 Photon Isolation

The photon isolation R_{Iso}^γ is defined in a similar way as for leptons (Subsection 4.2.3), except for the PU neutral contribution, which is not removed here. R_{Iso}^γ is determined from charged and neutral hadrons and from photons identified by the particle-flow reconstruction in a cone of size $\Delta R = 0.3$ around the candidate photon direction. The absolute photon isolation is defined as the sum of the transverse momenta of all these iso-deposits. To discriminate against photons produced in pile-up interactions, an additional isolation deposit is defined. It corresponds to the sum of charged particles from the vertices other than the primary vertex. The isolation observable corrected for pile-up is then divided by the photon transverse momentum p_T^γ . Isolated photons must satisfy $R_{Iso} < 1$.

4.3.3 Final State Radiation

Photons within $|\eta^\gamma| < 2.4$ are possible FSR candidates. To be accepted as FSR, a reconstructed photon must either have $p_T^\gamma > 2$ GeV and be found within a conical distance $\Delta R < 0.07$ from a selected lepton candidate, or have $p_T^\gamma > 4$ GeV and be found isolated within the conical distance of $0.07 < \Delta R < 0.5$ around a selected lepton candidate. In case there are multiple FSR photon candidates associated to a Z boson, only the highest- p_T one is considered, if it has $p_T > 4$ GeV, otherwise the one closest to any Z daughter lepton is chosen. In this analysis the FSR photons are kept if $|m_{\ell\ell\gamma} - m_Z| < |m_{\ell\ell} - m_Z|$ and $m_{\ell\ell\gamma} < 100$ GeV and, in the following, the presence of the recovered photon in the $\ell\ell$ kinematics is implicit.

The performance of the FSR selection algorithm has been measured using MC simulation samples and the rate was verified with single Z data events. The photons within the acceptance for the FSR selection are measured with an efficiency of 50% and with a mean purity of 80%. FSR photons are selected in 5% of single Z events with muon pairs, and 0.5% of single Z events with electron pairs. A gain of 3% (2%, 1%) in efficiency is expected for the selection of $ZZ \rightarrow 4\mu$ ($2e2\mu$, $4e$) events in this analysis.

4.4 Jets

In high-energy processes such as proton-proton collisions, quarks and gluons cannot exist as free particles, since they carry out a net color charge that must be conserved between the initial and the final state of the interaction (*color confinement*). If produced in the collision, they thus combine with quarks and anti-quarks spontaneously created from the vacuum to form a stable configuration of uncolored particles along the direction of the initial parton. The tight cone of particles created by this *hadronization* is called a *jet* and it is the experimental signature of quarks and gluons originated in hard collisions. A jet is formed of different kinds of particles, which are experimentally detectable thanks to their interaction with matter: light charged mesons (such as K , π), photons from π^0 decays, baryons (protons, neutrons...), leptons and neutrinos originating from the semi-leptonic decay of c and b quarks. The whole reconstruction of a jet can be achieved by many different algorithms, among which the anti- k_T algorithm is the most performing one.

4.4.1 The Anti- k_T Algorithm

Jet clustering algorithms are important tools for analyzing data from hadronic collisions. They are essential to investigate events involving quarks and gluons in the final state, since

they compute an estimate of energy and transverse momentum of the original parton, by analyzing all the particles within the jet cone. The anti- k_T algorithm [63] is based on the calculation of the distance between two particles i and j :

$$d_{ij} = \min(1/k_T^2(i), 1/k_T^2(j)) \cdot \frac{\Delta_{ij}^2}{R^2},$$

in which

- $k_T = |k|\sin\theta$ is the transverse momentum of particles;
- $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ is the distance between the two particles in the (η, ϕ) plane;
- R is a radius parameter characterizing the procedure (typically in the CMS jet reconstruction $R = 0.5$).

Another important parameter is the distance between the particle i and the beam, $d_{iB} = 1/k_T^2(i)$. The clustering algorithm proceeds by identifying the smallest of the two distances d_{ij} and d_{iB} : if the minimum is d_{ij} , it recombines together particles i and j by summing their momenta, while if it is d_{iB} , it labels the particle i as “jet” and removes it from the list of candidates. The distances are recalculated and the procedure is repeated until only jets are left. The functionality of the anti- k_T algorithm can be understood by considering an event with a few well-separated particles with high transverse momenta ($k_T(1), k_T(2), \dots$) and many soft particles. The $d_{1i} = \min(1/k_T^2(1), 1/k_T^2(i)) \cdot \frac{\Delta_{1i}^2}{R^2}$ between a hard particle 1 and a soft particle i is exclusively determined by the transverse momentum of the hard particle and the Δ_{1i} separation. The d_{ij} between similarly separated soft particles will instead be much larger, because of their smaller transverse momentum. Therefore soft particles will tend to cluster with hard ones long before they cluster among themselves. If a hard particle has no hard neighbors within a distance $2R$, then it will simply accumulate all the soft particles within a circle of radius R . With the anti- k_T algorithm, the hard jets are thus all circular with a radius R . Softer jets instead could also have more complex shapes. Finally, the anti- k_T algorithm takes into account for the infrared and collinear limits, ensuring reliable prediction at all orders in QCD. In Figure 4.4, a simulated jet reconstructed with the anti- k_T algorithm is shown.

4.4.2 Jet Energy Corrections

The calorimeter response to particles is not linear and therefore it is not straightforward to translate the measured jet energy to the true particle or parton energy. Several effects contribute to bias the measured jet energy, such as pile-up interactions, detector noise, and detector response non-uniformities in η and non-linearities in p_T . In order to obtain a uniform response of the energy with respect to these quantities, corrections must be applied. The CMS strategy is to factorize the contributions due to different physical effects. Following this approach, each *level* of correction scales the jet four-momentum depending on various jet related quantities (such as p_T , η , flavor, etc...). Moreover, the levels of correction are applied sequentially and with fixed order, meaning that the output of each step is taken as the input to the next one. The jet energy corrections are:

- **L1 - Pile-Up**: the energy coming from pile-up events is removed, taking into account the effect of overlapping collisions;

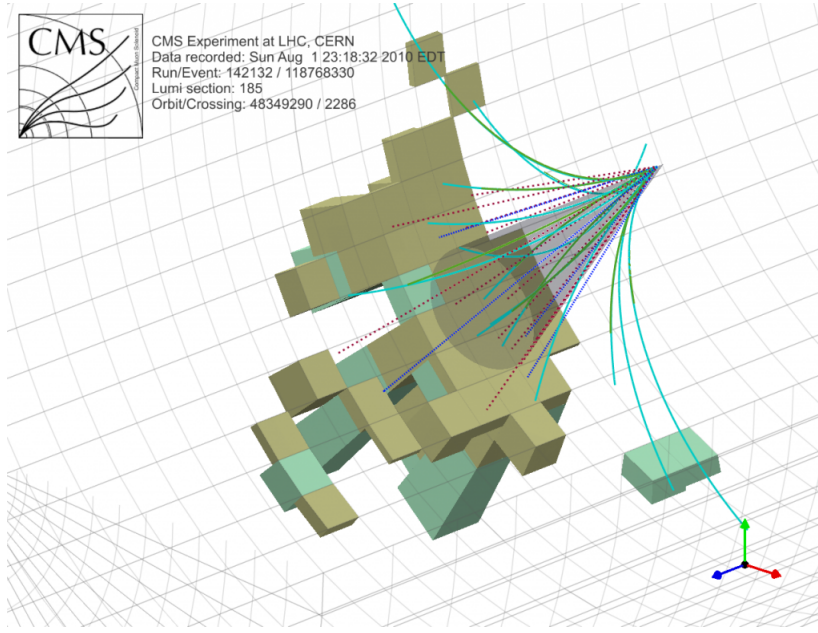


Figure 4.4: A graphical representation of a simulated jet clustered with the anti- k_T algorithm in the CMS Experiment, and measured as a release of energy in the hadronic calorimeter.

- **L2 - Relative:** a flat jet response is returned as a function of the jet pseudorapidity;
- **L3 - Absolute:** the non-uniformity of the jet response with respect to the jet transverse momentum is corrected;
- **L4 - Electromagnetic Energy Fraction:** a uniform jet response with respect to the electromagnetic energy fraction is obtained;
- **L5 - Jet Flavor:** the jet flavor dependence is corrected;
- **L7 - Parton:** the jet p_T is corrected by bringing it back to the originating parton p_T (on average);
- **L2L3 - Residuals:** small residual calibrations are applied to correct small differences between data and MC.

The first three items are considered mandatory for the reconstruction of a jet and the corresponding corrections are applied in this analysis, while the others are not. The factorized approach consists in the sequential application of the three levels L1+L2+L3 in order to correct the jet p_T . The correction values are derived from Monte Carlo simulation and then applied also on the data. The last correction is used only for data reconstruction, if significant discrepancies are found with respect to the MC.

4.4.3 Jet Energy Resolution

The energy resolution is an important characteristic of a reconstructed jet, since it represents the precision by which the jet energy is measured. It thus provides the systematic uncertainty associated to the reconstruction with respect to the generator-level jet. 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 [64], 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 4.1). 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 [64].

Table 4.1: Scaling factors applied on MC jets to described the data resolution.

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

4.4.4 Jet Reconstruction

In CMS jets are reconstructed using the anti- k_T clustering algorithm, with a distance parameter in the rapidity-azimuth plane of $R = 0.5$, by summing the four-momenta of individual PF particles according to the FASTJET package of reference [65]. The PF reconstruction is much more precise than other approaches, like calorimetric reconstruction, especially in the low p_T regions as Figures 4.5 shows. In order to assure a good reconstruction effi-

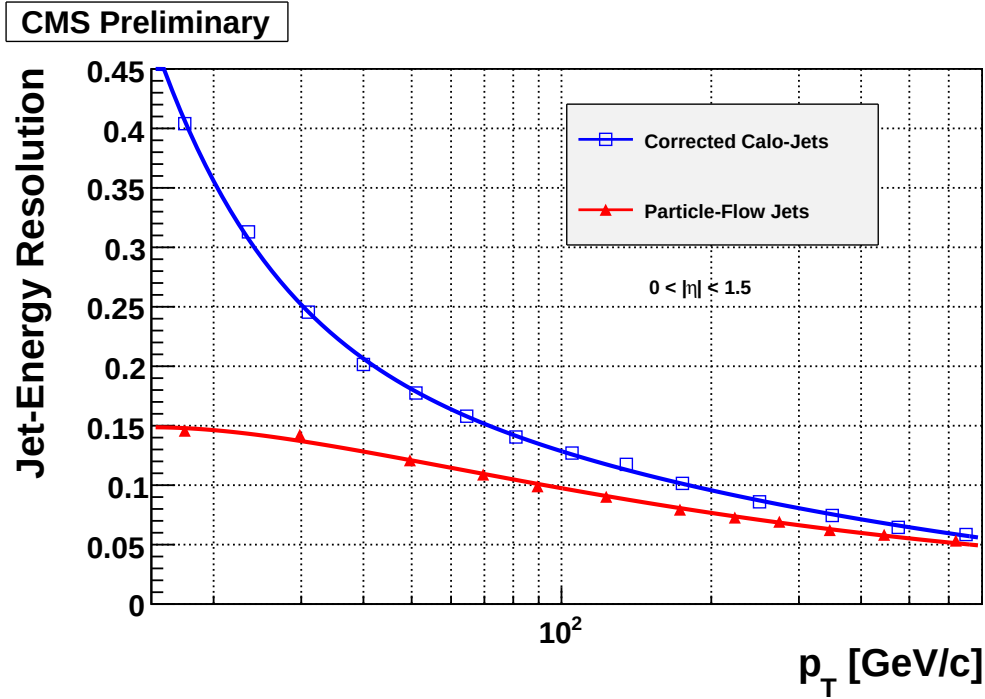


Figure 4.5: Jet-energy resolutions as a function of p_T for particle-flow jets (red curve) and for jets from calorimetric reconstruction (blue curve) in the barrel. [57]

ciency, identification quality criteria are imposed on the jets, based on the energy fraction of the charged, electromagnetic, and neutral hadronic components, and requiring at least one charged particle in the jet (Table 4.2). Once reconstructed, the jets overlapping (within $\Delta R = 0.5$) with any of the two leptons coming from the decay of the Z boson are removed from the jet collection. In order to remove jets coming from pile-up interactions, only PF candidates compatible with the primary vertex are considered in the jet reconstruction, if they are charged particles within the tracker acceptance. Otherwise, a multivariate technique is used. Moreover, jets are required to have $p_T > 30$ GeV, to further reduce the pile-up contamination as well as large uncertainty on the energy measurement, and $|\eta| < 4.7$, to ensure a good quality of the tracking information.

Table 4.2: Recommended cuts for the identification of PF jets at 8 TeV [66].

PF Jet ID	Loose	Tight
Neutral Hadron Fraction	< 0.99	< 0.90
Neutral EM Fraction	< 0.99	< 0.90
Number of Constituents	> 1	> 1
Muon Fraction	< 0.8	< 0.8
In addition for $-2.4 \leq \eta \leq 2.4$		
Charged Hadron Fraction	> 0	> 0
Charged Multiplicity	> 0	> 0
Charged EM Fraction	< 0.99	< 0.90

Chapter 5

Analysis of ZZ Events

The aim of this work is to perform the measurement of the inclusive and differential cross-sections of the $pp \rightarrow ZZ$ process, through the study of the 4ℓ decay channel. The analysis must preserve the highest possible reconstruction efficiency for the ZZ signal, while eliminating the contributions from the reducible and instrumental backgrounds. This is achieved relying on simple cuts for the lepton identification and isolation, kinematic selection, and specific requirements, as well as data-driven methods, for the background suppression.

In this Chapter, the analysis with the full statistics of the Run I of the LHC is presented, by explaining the experimental procedure and the strategy adopted in order to identify ZZ signal events. The employed data sets are reported, the event selection and signal efficiency are described and the techniques used to evaluate the background contribution are defined. Finally, the procedure applied to extract the inclusive measurement results is illustrated.

5.1 Data Sets

5.1.1 Experimental Data

The data samples used in this analysis have been collected by the CMS detector during 2012 and correspond to an integrated luminosity of 19.7 fb^{-1} at the center-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$. The integrated luminosity is measured using data from the hadron forward system and the pixel detector [67]. The uncertainty in the integrated luminosity measurement is 2.6%. A standard selection is applied, which requires high quality data with a good functioning of the sub-detectors (see Chapter 4).

As in the $H \rightarrow 4\ell$ and the previous ZZ analyses [17, 33], this study depends on primary data sets selected by the combination of different High Level Triggers (HLT). These include *DoubleMu* and *DoubleEle* triggers, which require two muons or electrons, the cross trigger *MuEG*, demanding one muon and one electron, and finally *TripleEle* triggers, requiring three electrons. The 4μ candidates are selected from the *DoubleMu* data sets, the $4e$ from the *DoubleEle* and the *TripleEle* data sets, while the $2e2\mu$ candidates come from the logical OR of all the data sets. The minimal transverse momenta (p_T) of the first and second lepton are 17 and 8 GeV, respectively, for the double lepton triggers, while they are 15, 8 and 5 GeV for the triple electron trigger. The trigger efficiency for ZZ events within the acceptance of this analysis is greater than 98%. The data samples processed in this analysis are summarized in table 5.1, along with the trigger selection operated on them.

Events are selected on-line from the presence of a pair of electrons or muons, or a triplet of

Table 5.1: List of data samples used in the analysis: 22Jan2013ReReco reprocessing of DoubleMu, DoubleEle, MuEG. The total integrated luminosity is 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

electrons.

5.1.2 Simulated Samples

Signal samples describing ZZ processes, as well as samples for a large variety of electroweak and QCD-induced SM background processes, have been obtained using detailed Monte Carlo (MC) simulations, based on the state-of-the-art theoretical calculations. Generated events are processed through the full GEANT 4 [68] CMS detector simulation and a trigger emulation. They are then reconstructed employing the same algorithms used for the data. A realistic misalignment scenario, based on the knowledge of positions of different elements of the muon and inner tracker systems, is used as input for the reconstruction. In this section, a description of the MC samples that have been employed for the event selection optimization prior to the direct investigation of experimental data is presented. These samples are further used in the analysis for the evaluation of signal efficiency and acceptance corrections (Subsection 5.2.2), the extraction of reducible and irreducible backgrounds (Section 5.3), the estimation of systematic uncertainties (Sections 5.5 and 6.4), the extraction of the unfolding response matrices (Section 6.1) and the comparison with the measurements (Sections 5.4 and 6.3).

The full set of MC simulation samples used in this analysis is reported in table 5.2.

The MadGraph 5.1 Monte Carlo event generator [69] is used to generate the signal $qq/qg \rightarrow ZZ \rightarrow 4\ell$ and $gg \rightarrow ZZ \rightarrow 4\ell$ (tree-level only) processes at LO, separately for 0, 1 and 2 jets. Because of this, no virtual diagrams are included in the cross-section calculation. Thus, in order to use the most accurate estimation possible, the MadGraph sample is scaled according to the theoretical ZZ cross-section calculated with MCFM at NLO [37]. For comparison signal $qq/qg \rightarrow ZZ$ samples generated with Powheg at NLO for 0 jet [70, 71] are also used. Since these samples don't contain events with 2 jets (a part from jets coming from the parton shower), the MadGraph sample is expected to better describe variables related to jets.

The signal $gg \rightarrow ZZ$ (box diagrams) processes are generated at LO with MCFM. Vector boson scattering processes with two Z bosons produced in association with 2 forward and backward jets are simulated using Phantom [72]. Other di-boson and tri-boson processes

Table 5.2: List of signal and background samples used in the analysis.

Process	Cross Sections [pb]	Sample
Signal Samples		
$ZZ \rightarrow 4l$ (MadGraph)	0.071 (no τ)	ZZJetsTo4L_TuneZstar_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/ADDSIM/PAT_CFG_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
$g\bar{g} \rightarrow ZZ \rightarrow 4\mu$ (MCFM)	0.000592340	GlUGluTo4mu_SMHContInInterf_M-125p6_8TeV-MCFM67-pythia6/Summer12_DR53X-PU_S10_START53_V19-v1
$g\bar{g} \rightarrow ZZ \rightarrow 4e$ (MCFM)	0.000592000	GlUGluTo4e_SMHContInInterf_M-125p6_8TeV-MCFM67-pythia6/Summer12_DR53X-PU_S10_START53_V19-v1
$g\bar{g} \rightarrow ZZ \rightarrow 2e2\mu$ (MCFM)	0.00118409	GlUGluTo2e2mu_SMHContInInterf_M-125p6_8TeV-MCFM67-pythia6/Summer12_DR53X-PU_S10_START53_V19-v1
$ZZ \rightarrow 4\mu + 2$ jets (Phantom)	$2.49 \cdot 10^{-4}$	ZZTo4muJJ_SMHContInInterf_M-125p6_8TeV-phantom-pythia6/Summer12_DR53X-PU_S10_START53_V19-v1
$ZZ \rightarrow 4e + 2$ jets (Phantom)	$2.48 \cdot 10^{-4}$	ZZTo4eJJ_SMHContInInterf_M-125p6_8TeV-phantom-pythia6/Summer12_DR53X-PU_S10_START53_V19-v1
$ZZ \rightarrow 2e2\mu + 2$ jets (Phantom)	$5.43 \cdot 10^{-4}$	ZZTo2e2muJJ_SMHContInInterf_M-125p6_8TeV-phantom-pythia6/Summer12_DR53X-PU_S10_START53_V19-v1
$WZZ +$ jets (MadGraph)	0.01968	WZZto4starJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZZ +$ jets (MadGraph)	0.005527	ZZZto4starJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$t\bar{t} + H$ (Pythia)	0.0001081681	TTbarH_HToZZTo4L_M-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	WWZto4starJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
Reducible Background Samples		
Z (MadGraph)	3503.71	DYJetsToLL_M-50_TuneZStar_8TeV-madgraph-tarball/Summer12_DR53X-PU_S10_START53_V7A-v1
WZ (MadGraph)	1.057	WZJetsTo3LJL_M-50_TuneZStar_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
WWW (MadGraph)	0.08058	WWWJets_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
$t\bar{t} \rightarrow 2l2\nu2B$ (Powheg)	23.64	TTTo2L2nu2B_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$t\bar{t} + W +$ jets (MadGraph)	0.232	TTWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1

(WZ, Zg, ZZZ, WZZ, WWZ) and the Z + jets samples are generated at LO with MadGraph, like events from $t\bar{t}$ production.

The general multi-purpose Monte Carlo event generator PYTHIA [73] is used for several processes including QCD multijet production and in particular it serves for the showering, hadronization, decays, and for adding the underlying event.

The default set of parton distribution functions (PDF) used for LO generators is CTEQ6L [74], whereas CT10 [75] is used for NLO generators.

The simulation includes pile-up effects, by matching the number of interactions per LHC beam-crossing observed in the data. The average number of measured pile-up interactions for the 8 TeV data set is 21.

5.2 Event Selection and Kinematics

The event selection is the same defined in the Higgs boson analysis [17] and relies on a simple sequence of cuts for the lepton identification and isolation, kinematic selection and specific background suppression requirements. The kinematic cuts allow for a drastic reduction of instrumental background rates in a transverse momentum and pseudorapidity acceptance of $p_T^e > 7$ GeV and $|\eta^e| < 2.5$ for electrons and $p_T^\mu > 5$ GeV and $|\eta^\mu| < 2.4$ for muons, as described in Chapter 4. The analysis is designed for the selection of Z boson pairs in association with jets.

5.2.1 Best Candidate Selection

The selection of the best ZZ candidate in the event is based on the reconstruction of four well-isolated leptons compatible with the primary vertex and the recovery of FSR described in Section 4.3. In the first step of the selection flow, a Z boson candidate is formed with pairs of leptons of opposite charge and same flavor. Among all such pairs, the one with mass closest to the nominal Z boson mass is retained and denoted Z_1 . Using the remaining leptons, additional $\ell^+\ell^-$ pairs are formed in order to build the second Z boson candidate, called Z_2 . If more than one Z_2 candidate is available, the one with the highest scalar sum of leptons' p_T is retained. Among the four leptons of the ZZ candidate, at least one is required to have $p_T > 20$ GeV and another one to have $p_T > 10$ GeV. These thresholds ensure that leptons are on the *plateau* of the trigger efficiency. Moreover, any opposite sign lepton pair is required to have $m_{\ell^+\ell^-} > 4$ GeV, independently from the flavor, in order to further reduce the contamination of events with a Z and a converted photon or events with low mass resonances that could survive due to wrong lepton pairing (such as J/ψ , Φ ...). It is found that, when the four leptons coming from the ZZ decay are within the detector acceptance, they are correctly selected in almost 100% of the cases.

Differently from the Higgs analysis, both the Z bosons are required to satisfy $60 < m_{Z_{1,2}} < 120$ GeV, because they must be on-shell.

In order to further suppress the reducible backgrounds (Subsection 5.3.2) from $Zb\bar{b}$, $Zc\bar{c}$, $t\bar{t}$ and the remaining WZ + jet(s), supplementary cuts are added to the selection, focusing on the rejection of leptons coming from the decay of b quarks. Such leptons are usually accompanied by hadronic products from the fragmentation and decay processes initiated in the b -quark jets. Moreover, because of the long lifetime of b -hadrons, they have a large impact parameter with respect to the primary vertex. Thus, lepton isolation and lepton impact parameter measurements can be used to obtain a powerful background rejection. Each lepton is therefore required to have $R_{Iso} < 0.4$ and $|SIP_{3D} = \frac{IP}{\sigma_{IP}}| < 4$, where SIP_{3D} is

the significance of the impact parameter (IP) to the event vertex (Subsection 4.2.3).

5.2.2 Signal Efficiency and Acceptance

The region in which the measurement is performed (called *fiducial region*) is defined by requiring the invariant mass of each Z boson between 60 and 120 GeV. The selection efficiency for the signal (ϵ) is estimated in the simulation separately for each final state and for each MC sample. It is defined as:

$$\epsilon = \frac{N_{fid}^{sel}}{N_{fid}(p_T^\ell, \eta^\ell)},$$

where

- N_{fid}^{sel} is the number of reconstructed signal events selected by the analysis;
- $N_{fid}(p_T^\ell, \eta^\ell)$ is the number of signal events generated in the fiducial region and within the detector acceptance (i.e. at least one of the four leptons with $p_T > 20$ GeV and another one with $p_T > 10$ GeV, $p_T^e > 7$ GeV and $|\eta^e| < 2.5$, $p_T^\mu > 5$ GeV and $|\eta^\mu| < 2.4$).

The efficiency has been studied as a function of several variables and results are reported in Figures 5.1 - 5.6 (on the left), showing an almost flat behavior in all the distributions. Moreover, plots reflect the higher efficiency in reconstructing muons with respect to electrons in CMS at this p_T -regime.

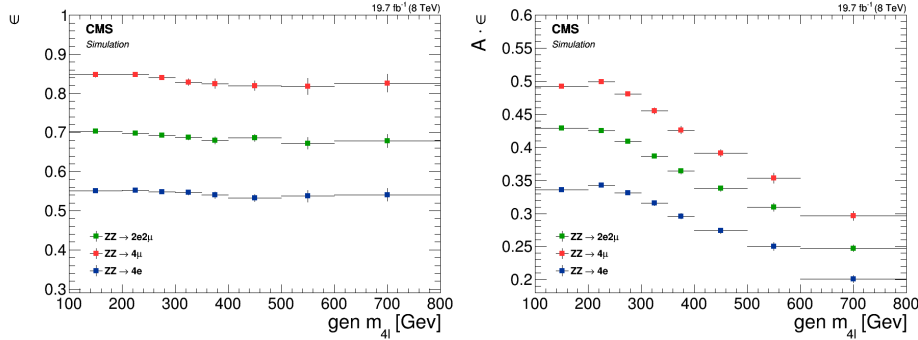


Figure 5.1: Selection efficiency (left) and acceptance times efficiency (right) distributions as a function of the ZZ mass, considering all simulated signal samples (Powheg+MCFM+Phantom+oth.). The three final states, i.e. $4e$ (blue), 4μ (red) and $2e2\mu$ (green), are overlaid.

The acceptance (A) is defined as:

$$A = \frac{N_{fid}(p_T^\ell, \eta^\ell)}{N_{fid}},$$

where N_{fid} is the total number of signal events generated in the fiducial region. The acceptance assesses the phase-space reduction due to the requirements on p_T and η for the final state leptons at generation level.

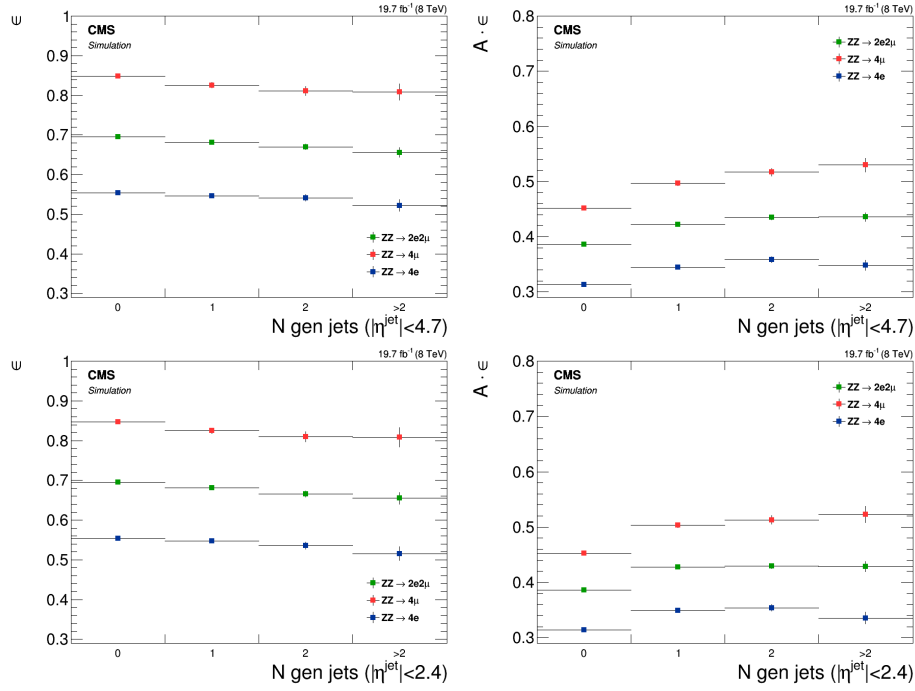


Figure 5.2: Selection efficiency (left) and acceptance times efficiency (right) distributions as a function of the number of jets (top) and central jets with $|\eta^{\text{jet}}| < 2.4$ (bottom), considering all simulated signal samples (MadGraph+MCFM+Phantom+oth.). The three final states, i.e. $4e$ (blue), 4μ (red) and $2e2\mu$ (green), are overlaid.

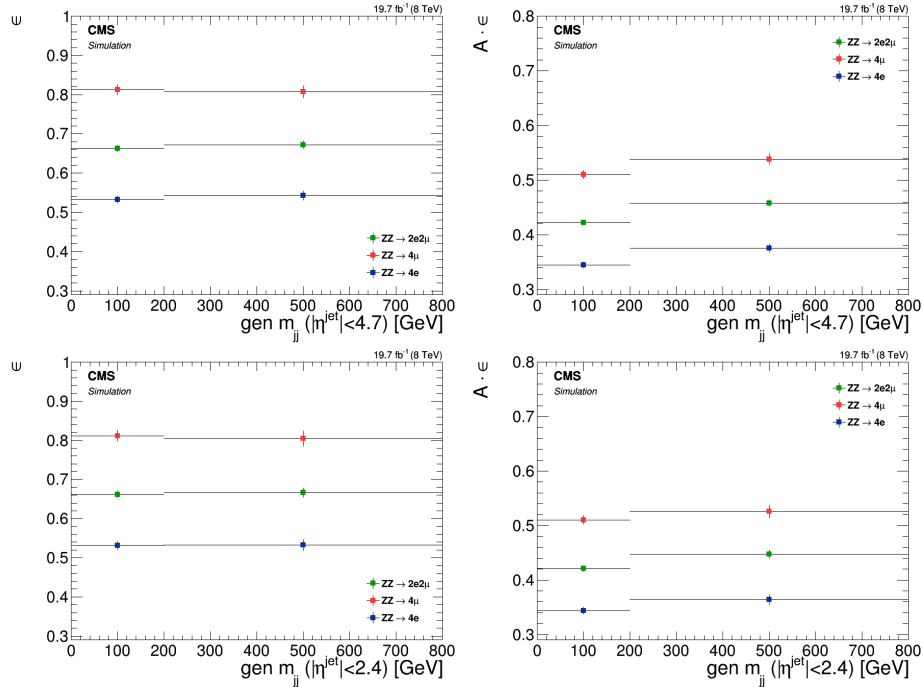


Figure 5.3: Selection efficiency (left) and acceptance times efficiency (right) distributions as a function of the mass of the two jets (top) and central jets (bottom) with the highest p_T , considering all simulated signal samples (MadGraph+MCFM+Phantom+oth.). The three final states, i.e. $4e$ (blue), 4μ (red) and $2e2\mu$ (green), are overlaid.

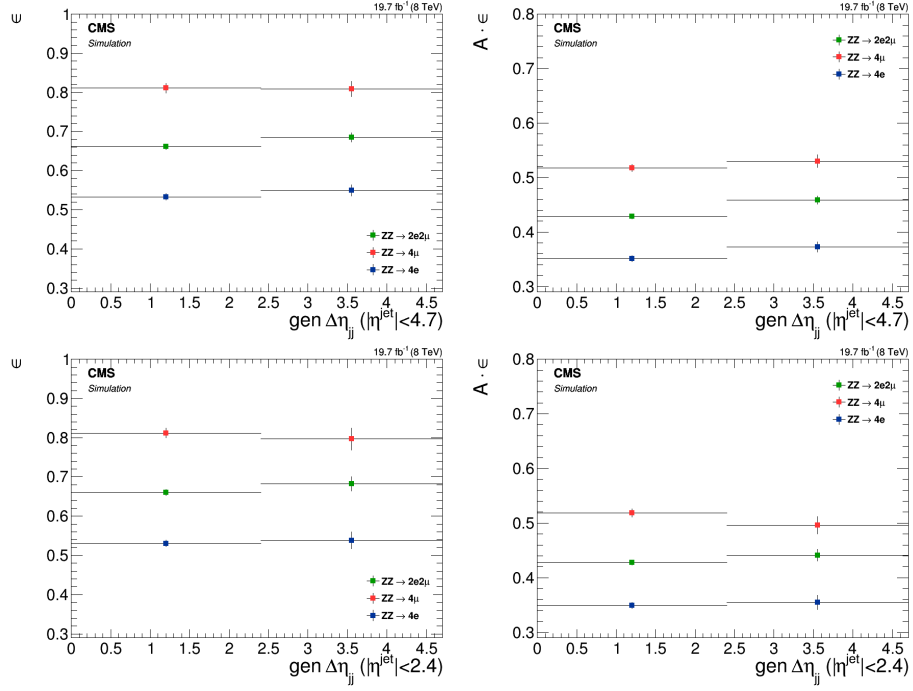


Figure 5.4: Selection efficiency (left) and acceptance times efficiency (right) distributions as a function of the separation in pseudorapidity of the two jets (top) and central jets (bottom) with the highest p_T , considering all simulated signal samples (MadGraph+MCFM+Phantom+oth.). The three final states, i.e. $4e$ (blue), 4μ (red) and $2e2\mu$ (green), are overlaid.

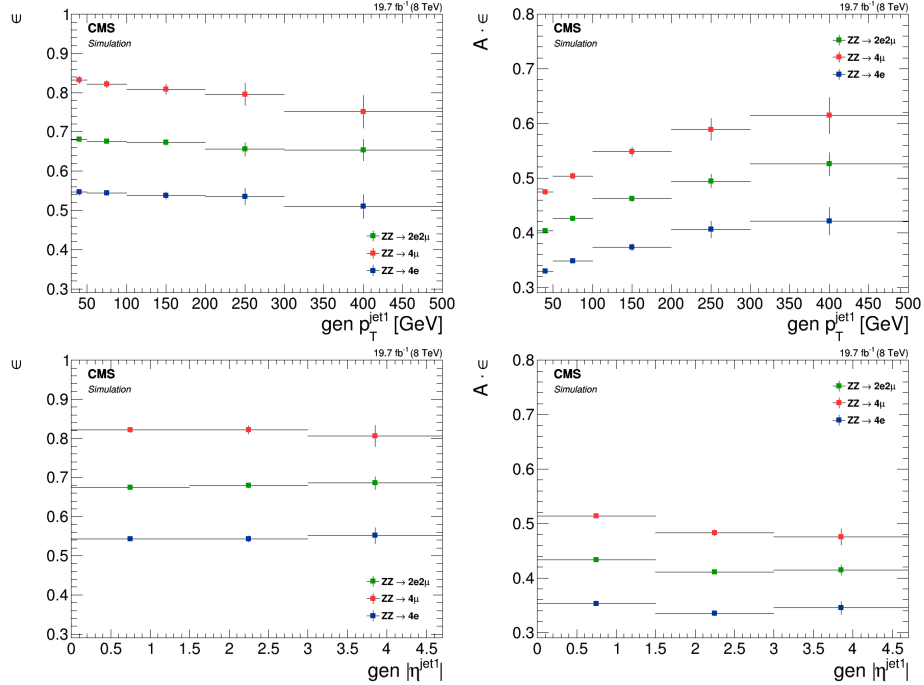


Figure 5.5: Selection efficiency (left) and acceptance times efficiency (right) distributions as a function of transverse momentum (top) and pseudorapidity of the jet with highest p_T , considering all simulated signal samples (MadGraph+MCFM+Phantom+oth.). The three final states, i.e. $4e$ (blue), 4μ (red) and $2e2\mu$ (green), are overlaid.

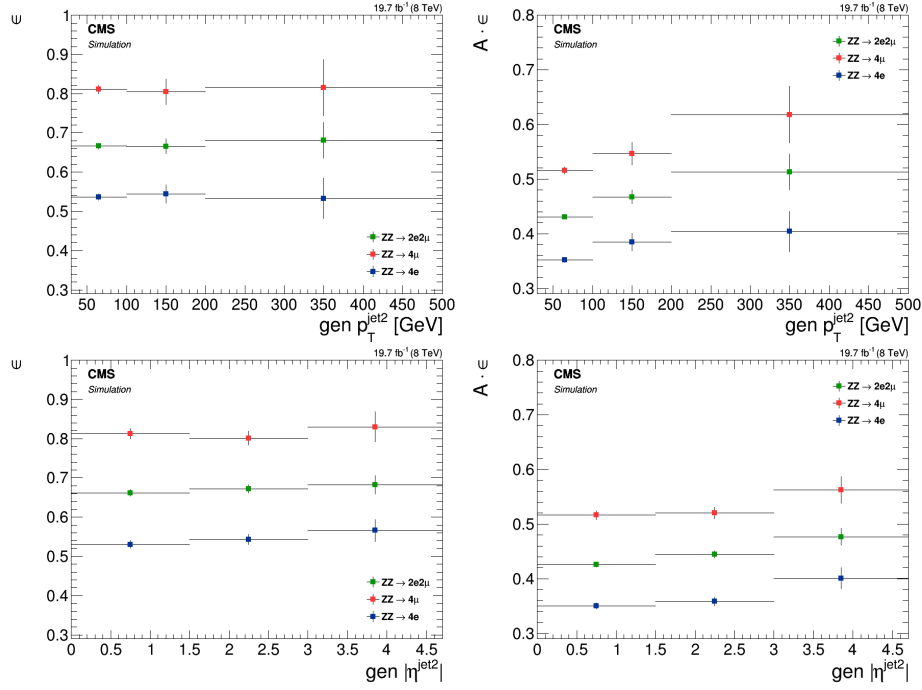


Figure 5.6: Selection efficiency (left) and acceptance times efficiency (right) distributions as a function of transverse momentum (top) and pseudorapidity of the jet with the second highest p_T , considering all simulated signal samples (MadGraph+MCFM+Phantom+oth.). The three final states, i.e. $4e$ (blue), 4μ (red) and $2e2\mu$ (green), are overlaid.

To measure the inclusive cross-section (Section 5.4) the product of acceptance and efficiency ($A \cdot \epsilon$) is essential. This product is evaluated from simulation and defined as the ratio between the number of signal events in the fiducial region which pass the analysis selection at reconstruction level and the number of signal events generated in that fiducial region, namely

$$A \cdot \epsilon = \frac{N_{fid}^{sel}}{N_{fid}}.$$

The requirements on p_T and η for the particles and the efficiency of the selection reduce the full possible phase space of the $ZZ \rightarrow 4\ell$ measurement by a factor within a range of 0.32 - 0.47 for the $4e$, 4μ and $2e2\mu$ (depending on the final state) with respect to all events generated in the mass window $60 < m_{Z_1}, m_{Z_2} < 120$ GeV. The $A \cdot \epsilon$ product estimated for each final state and for the different samples is listed in Table 5.4, while total values are reported in Table 5.3. In Figures 5.1 - 5.6 (on the right) the $A \cdot \epsilon$ behavior is shown as a function of several variables. The decreasing trend of the $m_{4\ell}$ dependency (Figure 5.1, right) has been further investigated, discovering that the number of events outside the η -acceptance grows with increasing lepton transverse momentum and, as a consequence, as a function of $m_{4\ell}$ (Figure 5.7). The number of events outside the p_T -acceptance is instead constant (Figure 5.8).

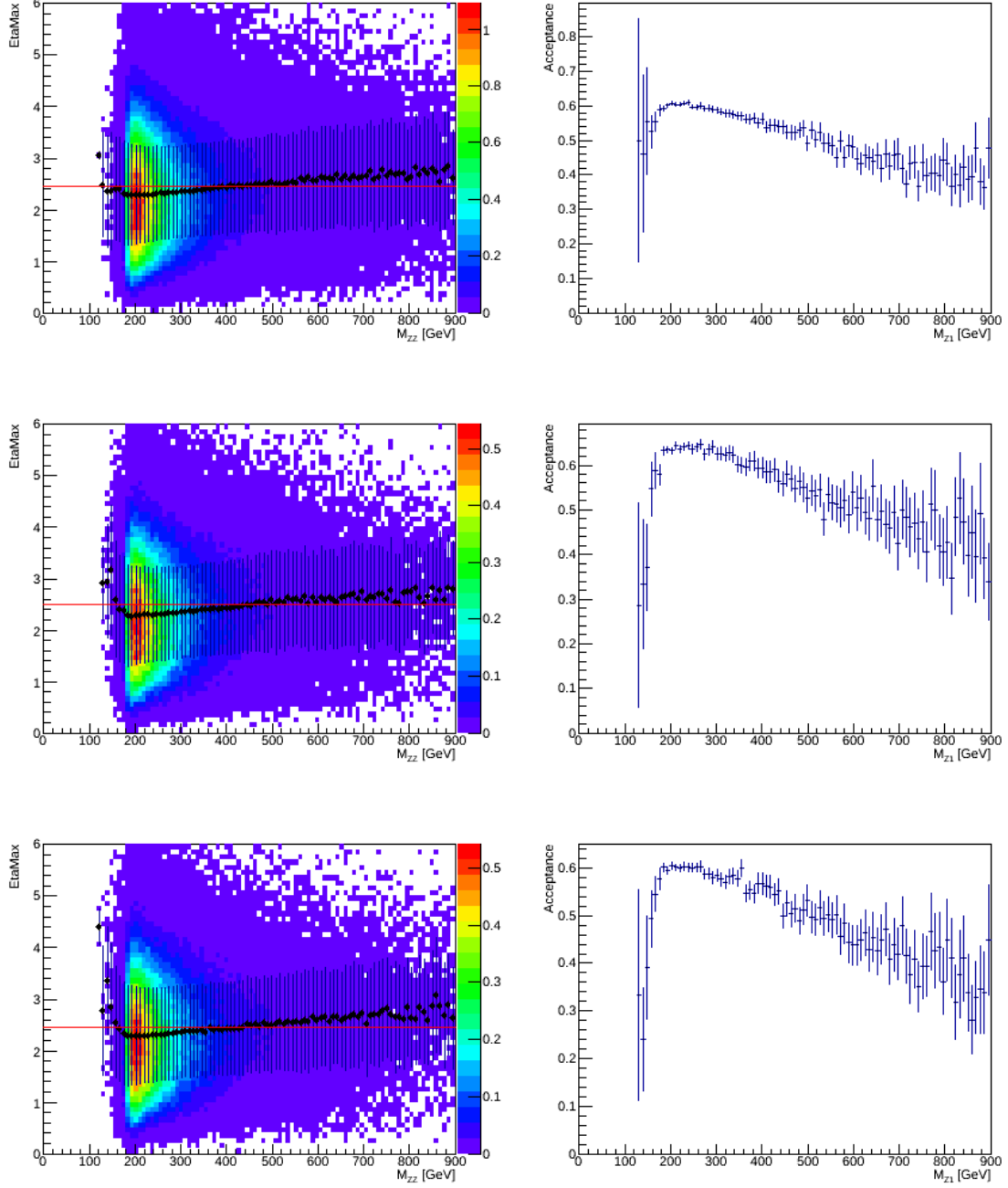


Figure 5.7: *Thermal map* (left) showing the density of yields (in arbitrary units) as a function of the mass of the four-lepton system (x axis) and the maximum lepton η -value (y axis). Points represent the mean value of each slice, while the red line represents the cut-value $|\eta^\ell| = 2.4$. Fraction of events generated in the fiducial region passing $|\eta^\ell| < 2.4$ as a function of m_{Z_1} (right).

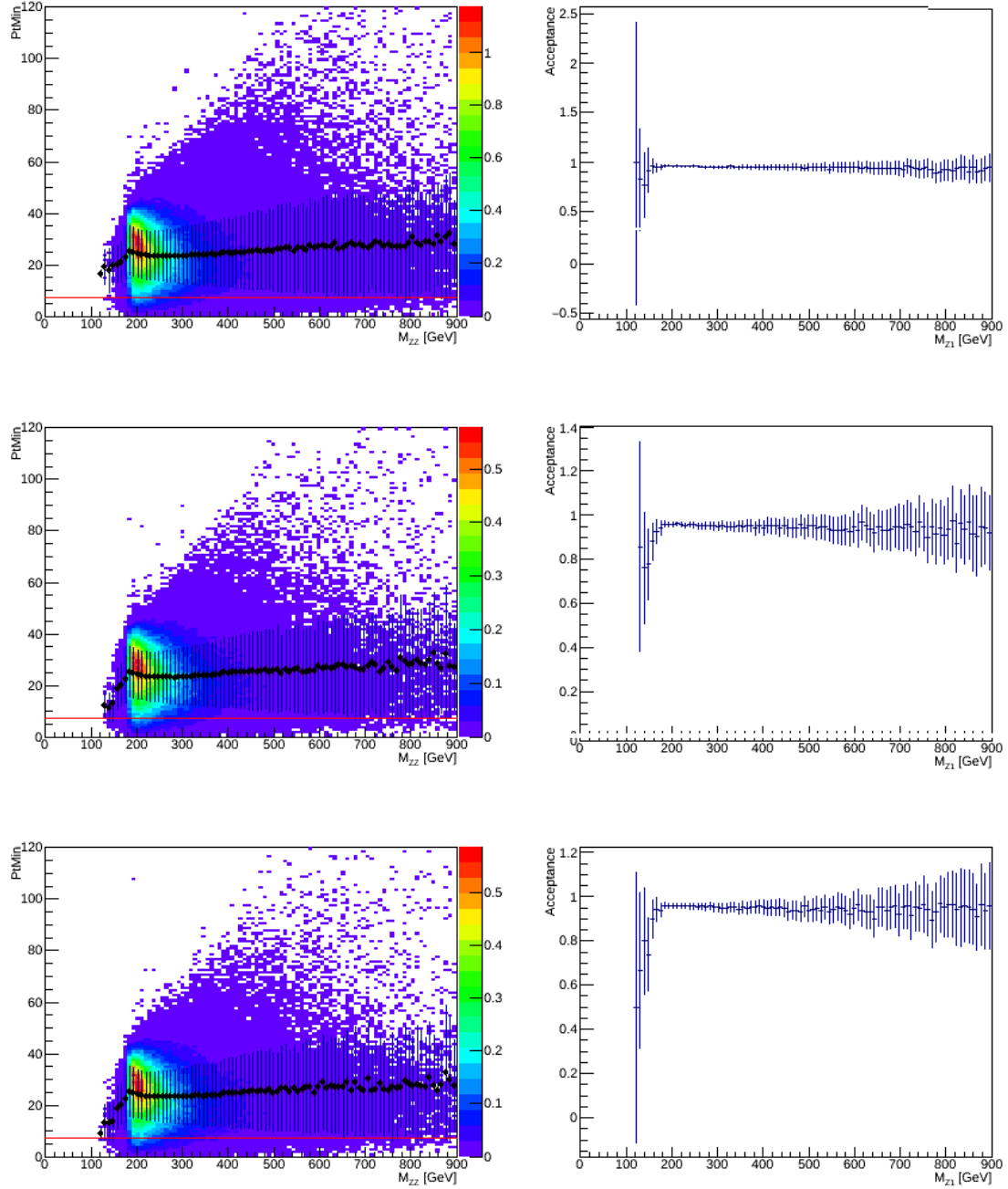


Figure 5.8: Thermal map (left) showing the density of yields (in arbitrary units) as a function of the mass of the four-lepton system (x axis) and the minimum lepton p_T -value (y axis). Points represent the mean value of each slice, while the red line represents the cut-value $p_T^\ell = 7$ GeV. Fraction of events generated in the fiducial region passing $|p_T^\ell| > 7$ GeV as a function of m_{Z1} (right).

Table 5.3: Selection efficiency and $A \cdot \epsilon$ product for the three final states used in the $pp \rightarrow ZZ$ cross-section measurement.

Final State	ϵ	$A \cdot \epsilon$
4μ	84%	47%
$4e$	55%	32%
$2e2\mu$	69%	40%

5.3 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. Two main background components can be defined: an irreducible background from genuine four high- p_T isolated leptons processes and a reducible background from processes with less than four leptons, but with jets which are misidentified as leptons.

5.3.1 Irreducible Background

The irreducible background contribution consists of processes which contain four-lepton events from non ZZ di-boson production, such as $t\bar{t}Z$, WWZ and $t\bar{t}WW$ processes. They have an event topology and kinematics very similar to those of signal events, from which they cannot be separated. However, this contribution is very small and it is estimated using MC samples (Figure 5.9).

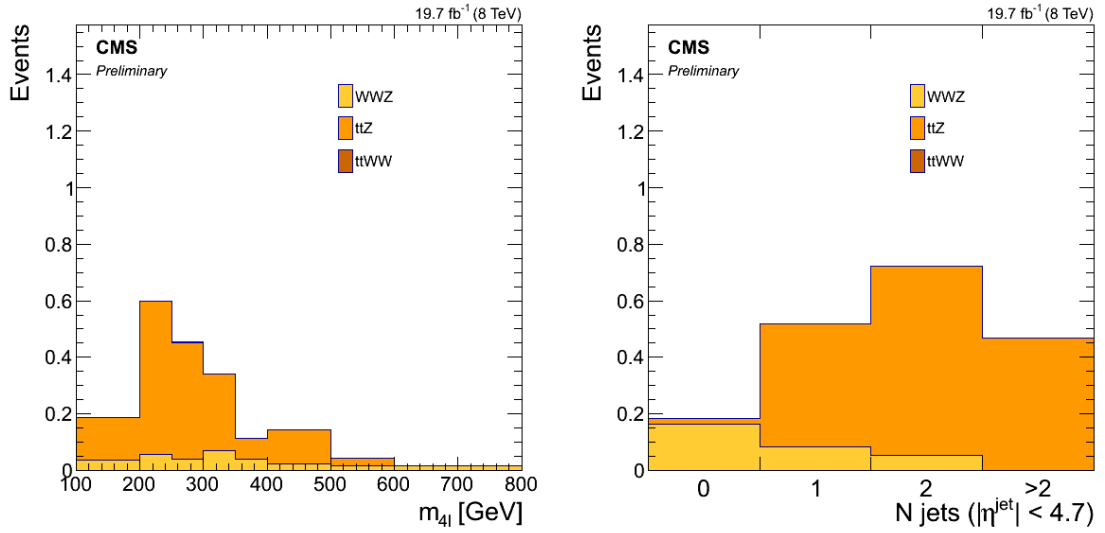
**Figure 5.9:** Number of events of the irreducible background component in the signal region as a function of the reconstructed invariant mass of the four-lepton system (left) and number of jets produced in the event (right).

Table 5.4: 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	—	—
$ZZ \rightarrow 4\mu + 2 \text{ jets}$ (Phantom)	—	—	0.633	—	—	0.65
$ZZ \rightarrow 4e + 2 \text{ jets}$ (Phantom)	—	0.440	—	0.45	—	—
$ZZ \rightarrow 2e2\mu + 2 \text{ jets}$ (Phantom)	1.08	—	—	0.56	—	—
$ZZZ + \text{jets}$ (MadGraph)	0.284	0.121	0.192	0.42	0.35	0.54
$WZZ + \text{jets}$ (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

5.3.2 Reducible Background

The main background contribution arises mostly from the production of a Z boson in association with jets, as well as from $t\bar{t}$, WZ and WWW + jets processes and 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 [17]. The low statistics of MC event yields after the event selection does not allow this contribution to be estimated directly from simulated samples.

The typical procedure to evaluate background from data consists of choosing a wide background *control region* (CR) outside the signal phase space. A control region is a kinematic region where the signal and the background are well discriminated, since signal is removed by the CR cuts and background can thus be characterized. The control region gets populated by relaxing the event selection and verifying that the event rates change according to the expectation from simulation. If a specific background contribution has to be determined, the corresponding control region must be chosen carefully since any of the other reducible backgrounds might rapidly become dominant if the event selection is relaxed, thus making the extrapolation to the signal phase space difficult.

The data-driven method chosen to evaluate the reducible background contribution is based on two control samples, obtained as subsets of four-lepton events, by demanding a Z_1 candidate reconstructed with the same requirements as the Z_1 built in the signal region, and two additional leptons, with opposite charge 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. Demanding that leptons fail some criteria is necessary in order to eliminate signal events in the control region. However, these leptons can be seen as *loose* leptons, i.e. leptons passing a loose set of cuts which could also survive to the full tighter selection. The $3P1F$ control sample is used to estimate the contribution in the signal region of events with three genuine leptons, such as WZ + 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 + jets and $t\bar{t}$. As well as with WZ events that intrinsically have three prompt leptons, the $3P1F$ control sample is also populated with the type of events contained in the $2P2F$ control region, albeit with different relative proportions, when one of the two loose leptons passes the full selection criteria. This contamination is evaluated as:

$$N_{3P1F}^{2P2F} = \sum \left(\frac{f_i}{1-f_i} + \frac{f_j}{1-f_j} \right) N_{2P2F}, \quad (5.1)$$

where

- N_{2P2F} is the number of events in the $2P2F$ control region;
- f_i is the fake rate of a failing lepton in the $2P2F$ sample.

The *fake rate* is the lepton misidentification probability $f(p_T^\ell, \eta^\ell)$ to extrapolate the background yields from the control region to the signal region. This probability is defined as the fraction of non-signal leptons passing the lepton selection of the analysis. The fake rate is estimated in a sample composed by $Z_1 + 1\ell_{loose}$ events, where in addition to the opposite sign, same flavor pair which forms the Z_1 , exactly one additional lepton is reconstructed fulfilling all the selection requirements, but for the identification and the isolation cuts. The lepton misidentification probability ranges from 1% to 15% and has a mild dependence on the pseudorapidities for the electrons.

The total number of reducible background events in the signal region is computed as the sum of the $3P1F$ control sample contribution (to which N_{3P1F}^{2P2F} is subtracted) and the $2P2F$ control region contribution. It is given by the following formula:

$$N_{exp}^{red\ bkg} = \sum_i^{N_{3P1F}} p_i - \sum_i^{N_{2P2F}} p_{i,1} p_{i,2}, \quad (5.2)$$

where:

- N_{3P1F} is the number of events in the $3P1F$ control region;
- $p_i = \frac{f_i}{1-f_i}$;
- $p_{i,1(2)} = \frac{f_{i,1(2)}}{1-f_{i,1(2)}}$;
- f_i is the fake rates of the failing lepton in the $3P1F$ control sample;
- $f_{i,1(2)}$ is the fake rates of the first (second) failing lepton in the $2P2F$ control sample.

In Figure 5.10 the contribution of the reducible background is reported, estimated both from MC samples and by using the *fake-rate* method. As already stressed, the lack of statistics makes the data-driven estimate necessary.

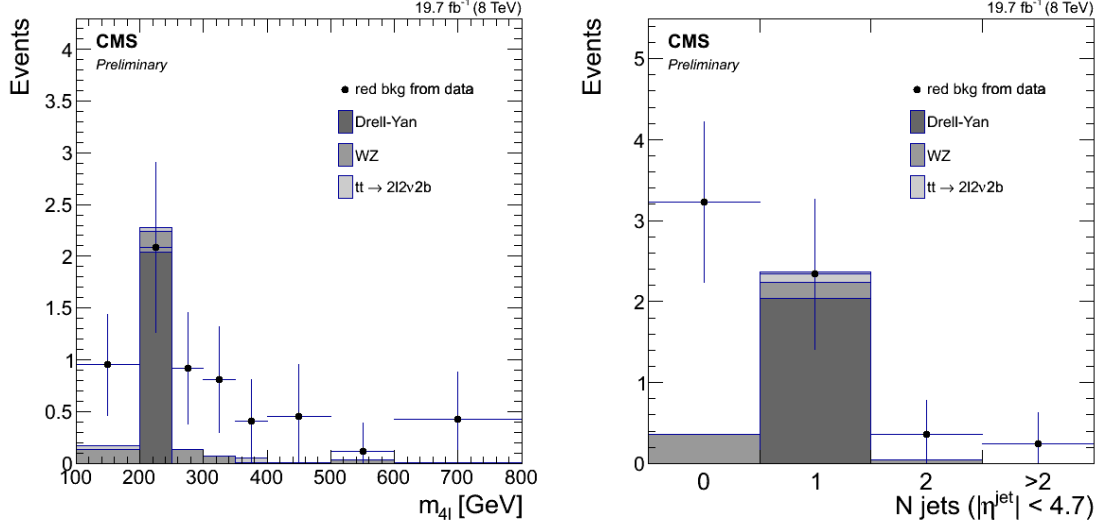


Figure 5.10: Reducible background component in the signal region as a function of the reconstructed invariant mass of the four-lepton system (left) and number of jets produced in the event (right). Points represent the data-driven estimate, the stacked histogram represents the Monte Carlo predictions, characterized by a very poor statistics.

5.4 Inclusive ZZ Cross-Section Measurement

After the event selection and background estimation, the inclusive $pp \rightarrow ZZ$ cross-section can be measured as:

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

where

- N_{data} is the total number of observed data events in each decay channel ($ZZ \rightarrow 4e$, $ZZ \rightarrow 4\mu$ and $ZZ \rightarrow 2e2\mu$);
- N_{bkg} is the total number of estimated background events in each decay channel;
- A is the p_T^ℓ and η^ℓ acceptance, evaluated with respect to generated events within the fiducial region ($60 < m_{Z_{1,2}} < 120$ GeV) as described in Subsection 5.2.2;
- ϵ is the signal efficiency, defined as the ratio between the reconstructed signal events passing the selection and the signal events generated within the geometrical acceptance in the fiducial region (Subsection 5.2.2);
- $\mathcal{L} = 19.7 \pm 0.51 \text{ fb}^{-1}$ is the total integrated luminosity;
- $\mathcal{B}(Z \rightarrow \ell^+ \ell^-) = 3.363 \pm 0.004$ (3.366 ± 0.007) is the branching fraction of a Z boson decaying into an electron (muon) pair [76].

The number of signal events used to evaluate the ZZ cross-section measurement is computed subtracting the irreducible and reducible background estimates from the selected data samples. Values are reported in Table 5.5.

In Figures 5.11, 5.12, 5.13 and 5.14 the distributions of some interesting variables related to lepton and jet quantities are reported, for both MC samples and data distributions. In all cases, $q\bar{q}(qg) \rightarrow ZZ \rightarrow 4\ell$ processes are described using the MadGraph sample, except for the $m_{4\ell}$ distribution for which Powheg samples are used.

The inclusive cross-section is obtained by dividing the distribution of the number of signal events as a function of $m_{4\ell}$ by the differential distribution of $A \cdot \epsilon$ computed using the Powheg set of samples and integrating the result. Compatible measurements are found by employing the number-of-jets variable and the MadGraph set of samples. Table 5.6 lists the inclusive cross-section obtained from each individual decay channel as well as the total cross-section based on the combination of all channels, computed as the weighted average of them. The total cross-section for a pair of Z bosons in the mass range $60 < m_Z < 120$ GeV is found to be:

$$\sigma_{pp \rightarrow ZZ} = 7.76 \pm 0.48 \text{ (stat.)}_{-0.29}^{+0.27} \text{ (syst.) pb.}$$

The measured cross-section agrees with the theoretical value of 7.5 ± 0.5 pb calculated with MCFM 6.6. In this calculation, the contribution from $q\bar{q} \rightarrow ZZ$ is obtained at NLO, while the smaller contribution (approximately 6%) from $gg \rightarrow ZZ$ is obtained at LO.

Table 5.5: The signal yield is obtained by subtracting from the observed yield the expected background events. Statistical uncertainties are reported.

Final state	Observed data	Irr. bkg	Red. bkg	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	280 ± 17

Table 5.6: 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.

Process	Total cross-section [pb]
$pp \rightarrow ZZ(4\mu)$	$7.46 \pm 0.86 \text{ (stat.)}_{-0.29}^{+0.27} \text{ (syst.)}$
$pp \rightarrow ZZ(4e)$	$7.32 \pm 1.02 \text{ (stat.)}_{-0.51}^{+0.48} \text{ (syst.)}$
$pp \rightarrow ZZ(2e2\mu)$	$8.22 \pm 0.70 \text{ (stat.)}_{-0.47}^{+0.43} \text{ (syst.)}$
$pp \rightarrow ZZ(4\ell)$	$7.76 \pm 0.48 \text{ (stat.)}_{-0.29}^{+0.27} \text{ (syst.)}$

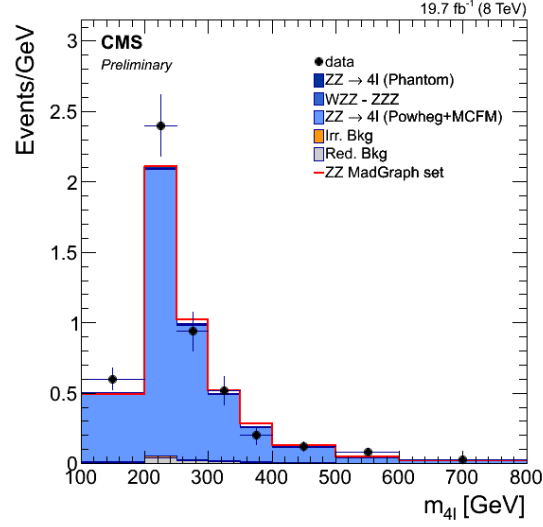


Figure 5.11: Different contributions to the distribution of the reconstructed four-lepton mass. Points represent the data and the stacked filled histogram represents the predictions for ZZ signal and background contributions using Powheg samples to describe $q\bar{q}(qg) \rightarrow ZZ \rightarrow 4\ell$ processes (while for the stacked histogram outlined in red the MadGraph simulation is used).

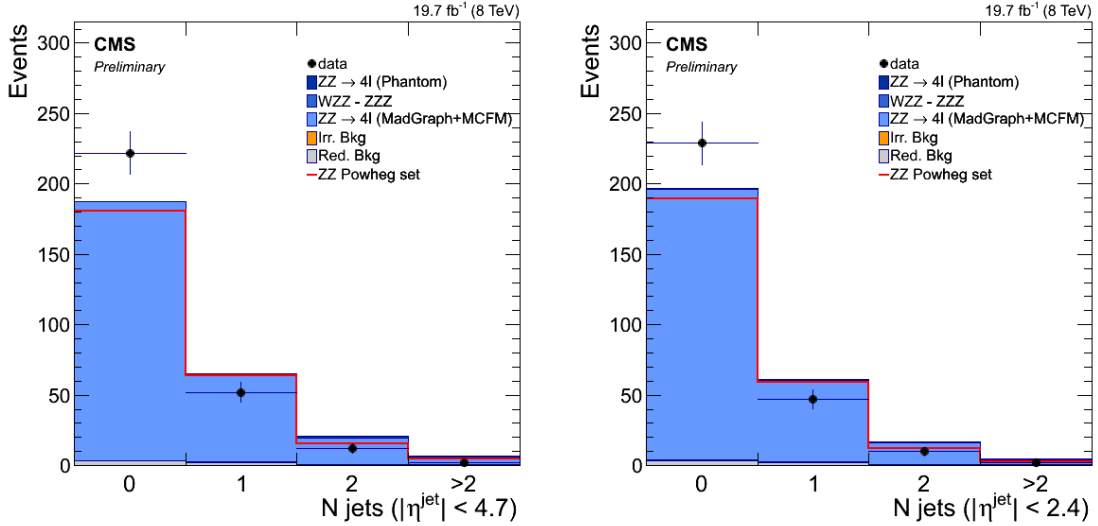


Figure 5.12: Different contributions to the distribution of the reconstructed number of jets produced in the event, with $\eta^{jet} < 4.7$ on the left and $\eta^{jet} < 2.4$ on the right. Points represent the data and the stacked filled histogram represents the predictions for ZZ signal and background 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).

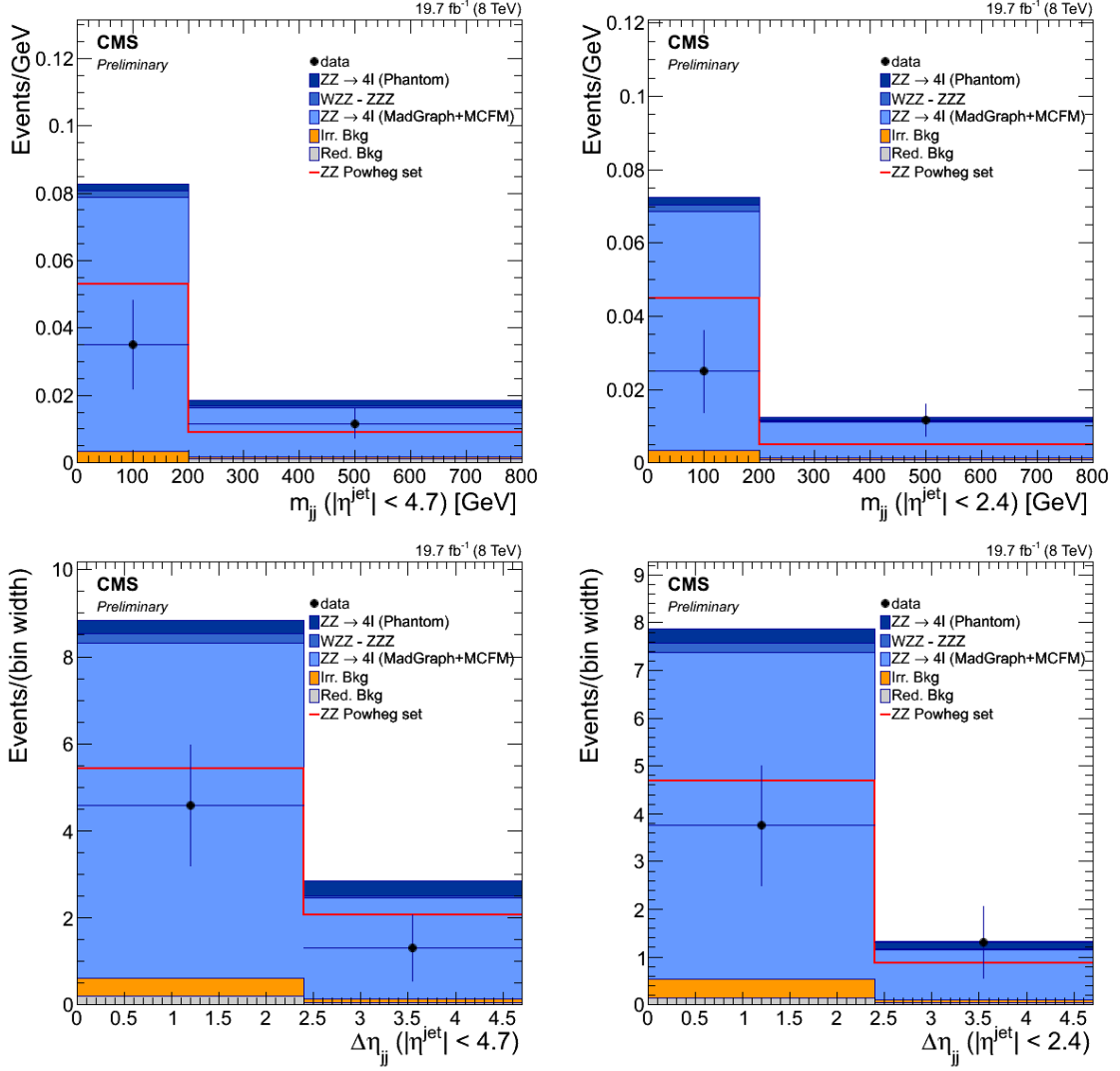


Figure 5.13: Different contributions to the distribution of the invariant mass of the two most energetic jets (top) and the pseudorapidity interval between them (bottom), with $\eta^{jet} < 4.7$ on the left and $\eta^{jet} < 2.4$ on the right. Points represent the data and the stacked filled histogram 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).

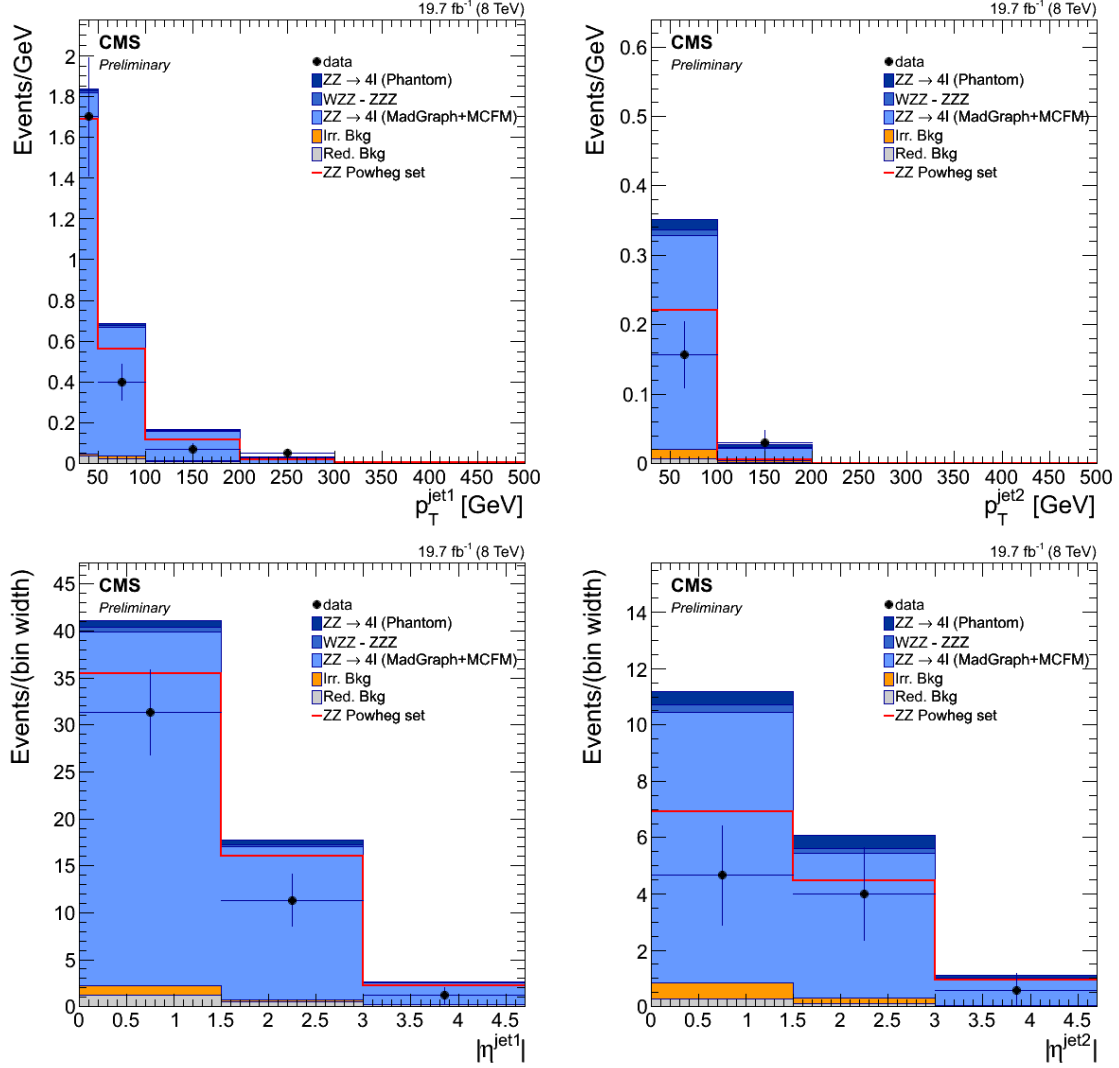


Figure 5.14: Different contributions to the distribution of the reconstructed transverse momentum (top) and pseudorapidity (bottom) of the leading (left) and sub-leading jet (right). Points represent the data and the stacked filled histogram represents the predictions for ZZ signal and background contributions 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).

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

- $60 < m_{Z_1}, m_{Z_2} < 120$ 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^ℓ and p_T^ℓ acceptance of the detector and are the same selection criteria demanded when leptons are reconstructed (Section 4.2), while the first requirement corresponds to the definition of the wider region used for the previous inclusive measurement.

The measurement of the $pp \rightarrow ZZ \rightarrow 4\ell$ cross-section is performed in this tight region, considering thus leptons within the η^ℓ and p_T^ℓ acceptance. The formula applied in this case is:

$$\sigma_{pp \rightarrow ZZ \rightarrow 4\ell} = \frac{N_{data} - N_{bkg}}{\mathcal{L} \cdot A_{tight} \cdot \epsilon},$$

where A_{tight} is the acceptance computed in the tight region, that by definition is equal to 1. Efficiency values are reported in Table 5.3, for the different final states. The term describing the branching ratio of the leptonic decay has been removed, since the formula above represents the cross-section of a Z boson pair decaying specifically into four leptons and no longer the process $pp \rightarrow ZZ$. The number of signal and background events used to evaluate the $pp \rightarrow ZZ \rightarrow 4\ell$ cross-section is the same employed for the inclusive measurement in the wider fiducial region and it is shown in Table 5.5. Table 5.7 contains the inclusive cross-section values estimated in the tight fiducial region. Contrary to the previous case, in which the *same* measurement ($\sigma_{pp \rightarrow ZZ}$) is obtained in each decay channel, here each final state describes a different process: $ZZ \rightarrow 2e2\mu$, $ZZ \rightarrow 4\mu$ or $ZZ \rightarrow 4e$. Indeed, in the $2e2\mu$ channel the cross-section is twice the measurement obtained in the other two final states. Therefore, instead of a weighted average, the 4ℓ cross-section is obtained as the sum of all final states. The total measured cross-section is in agreement with the theoretical expectation of 19.7 ± 1.6 fb obtained with MCFM 6.6 at NLO.

5.5 Systematic Uncertainties

Many sources of systematic uncertainty affect the inclusive cross-section measurement. The uncertainty on the LHC integrated luminosity of the data sample is 2.6% [67]. Theoretical uncertainties in the $ZZ \rightarrow \ell\ell\ell'\ell'$ acceptance are evaluated using MCFM and by varying the renormalization and factorization scales, up and down, by a factor of two with respect to the default values $\mu_R = \mu_F = m_Z$. Variations in the acceptance are 0.1% ($q\bar{q} \rightarrow ZZ$ at NLO) and 0.4% ($gg \rightarrow ZZ$), and can be neglected. Uncertainties related to the choice of the parton distribution functions (PDF) and the strong coupling constant α_s are evaluated following the PDF4LHC [77, 78, 75] prescriptions and using CT10, MSTW08, and NNPDF [79] PDF sets. They are found to be 1.2% in the wider region and 0.8% in the tighter region. Additional systematics affecting the estimated number of reducible backgrounds come from the limited amount of data in the background control regions, which propagates to

Table 5.7: The $ZZ \rightarrow 4\ell$ production cross-section as measured in each decay channel and for the sum of all channels in the tight fiducial region.

Process	Total cross-section [fb]
$ZZ \rightarrow 4\mu$	4.74 ± 0.54 (stat.) $^{+0.16}_{-0.18}$ (syst.)
$ZZ \rightarrow 4e$	4.99 ± 0.69 (stat.) $^{+0.32}_{-0.34}$ (syst.)
$ZZ \rightarrow 2e2\mu$	10.77 ± 0.90 (stat.) $^{+0.55}_{-0.58}$ (syst.)
$ZZ \rightarrow 4\ell$	20.50 ± 1.26 (stat.) $^{+0.79}_{-0.83}$ (syst.)

the background evaluation in the signal region, and from the uncertainty on the misidentification rate used to pass from the background control to the signal region. The reducible background uncertainties of $Z + \text{jets}$, $WZ + \text{jets}$ and $t\bar{t}$ processes vary between 30% and 70% on the background yield, which corresponds to a 2 – 3% of uncertainty on the cross-section measurement. The irreducible background uncertainty is smaller and varies between 13% and 26% on the yield (1 – 2% on the cross-section estimate). The uncertainty due to the selection efficiency of muon and electron measurement are also considered: the systematic uncertainties for trigger efficiency is estimated to be 1.5%, while lepton tracking, identification, impact parameter and isolation uncertainties are 1 – 5% for both muons and electrons. As in the Higgs analysis [17], the contributions coming from background uncertainties are estimated shifting data distributions up and down by their error, while for the selection efficiency two distributions of $N/(A \cdot \epsilon)$ are obtained by varying the value of ϵ by its uncertainty. Cross-sections are then computed using these new *up*- and *down*- distributions and the discrepancy with the central value is taken as the systematic uncertainty. Given that in the tighter fiducial region the four-lepton cross-section is obtained by summing up the measurements of the three different final states, correlations between the $4e/4\mu$ and $2e2\mu$ decay channels are taken into account.

Table 5.8 summarizes the sources of systematic uncertainty and their values.

Table 5.8: 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	2 – 3%
Irreducible background	1 – 2%
Acceptance for PDF	0.8 – 1.2%

Chapter 6

The Differential ZZ Cross-Section Measurement

After measuring the inclusive cross-section, useful information can be further extracted from the $pp \rightarrow ZZ$ process, by analyzing its behavior as a function of interesting variables. The measured distribution of physical observables is usually distorted by several effects, such as finite resolution and limited acceptance of the detector, and it is thus considered as a *noisy distortion* of the true one, i.e. the distribution one would observe under idealized conditions (ideal detector, no backgrounds...). In order to remove experimental effects and correct the observed measurement to the *particle level*, an unfolding procedure is needed. This procedure allows different experiments to achieve a direct comparison of the results with each other and with theoretical predictions.

In this Chapter the ZZ production cross-section is measured differentially as a function of the invariant mass of the four-lepton system, the number of jets produced in the event, the invariant mass of the two most energetic jets (m_{jj}), the difference of pseudorapidity 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 with $|\eta|^{jet} < 4.7$ and central jets with $|\eta|^{jet} < 2.4$ are considered. Reconstructed distributions are first unfolded, obtaining new distributions at particle level from which differential cross-sections are computed. The measurements are performed in the leptonic decay modes $ZZ \rightarrow \ell\ell\ell'\ell'$ ($\ell, \ell' = e, \mu$).

6.1 Data Unfolding

In order to recover the original *true* distribution from the *observed* one, smeared by the measurement, an unfolding procedure is needed. This procedure removes effects of finite resolutions, systematic biases and detection efficiency by using Monte Carlo simulation processes. The unfolding method is based on the so-called *response matrix*, derived using simulated Monte Carlo samples, which is a mapping between the true value (generated) of the observable and the reconstructed (measured) value, distorted by detector effects. Because of inevitable statistical uncertainties in the measured distribution, the exact solution that one would obtain by inverting the response matrix (if it exists) can lead to unacceptable results, wildly oscillating and useless. In order to avoid this unstable behavior, a *regularization condition* can be imposed, by requiring a smooth true distribution. The regularized unfolding methods investigated here, as implemented in the RooUnfold package

of C++ classes [80], are the *Singular Value Decomposition* (SVD) method [81] and the *iterative Bayesian* unfolding [82].

In the following, first the unfolding procedure is described in detail. The choice of the binning of the distributions is then discussed, in order to control migration from one bin to others, and the performance of the unfolding is validated through studies on Monte Carlo samples. Finally the unfolding procedure is applied on data distributions.

6.1.1 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 [81, 83]. 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.

The distribution N_{gen} of dimension m is then generated according to some idea of the physical process under study, and the detector simulation is performed. 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 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:

$$\sum_j A_{ij} N_{gen}^j = N_{sig}^i := N_{data}^i - N_{bkg}^i. \quad (6.1)$$

The linear system of equations can be solved using the exact inverse of the response matrix on measured histogram and obtaining a data distribution *at particle level*. However a direct inversion of the matrix usually leads to completely unacceptable wildly oscillating results. In order to overcome this problem, different regularized algorithms can be used, such as the SVD and iterative Bayesian methods, well described in [81] and [82].

6.1.2 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, evaluated using MC signal samples [83]. 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:

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

where *reco* refers to reconstructed events fulfilling the full selection requirements described in Section 5.2 and *gen* refers to generated events satisfying the phase space requirements. Without migration effects, purity and stability would be equal to one. The purity (stability) is sensitive to migrations into (out of) the bin. In order to keep bin-to-bin migrations acceptably small, the bin widths for each observable are optimized such that in each bin purity and stability are greater than about 70%. The binning selection for each observable is defined as:

- $m_{4\ell}$: $\{100, 200, 250, 300, 350, 400, 500, 600, \geq 800\}$ GeV;
- N_{jets} : $\{0, 1, 2, > 2\}$;
- m_{jj} : $\{0, 200, \geq 800\}$ GeV;
- $\Delta\eta_{jj}$: $\{0, 2.4, \geq 4.7\}$;
- p_T^{jet1} : $\{30, 50, 100, 200, 300, \geq 500\}$ GeV;
- p_T^{jet2} : $\{30, 100, 200, \geq 500\}$ GeV;
- η^{jet1}, η^{jet2} : $\{0, 1.5, 3, 4.7\}$.

Figures 6.1 - 6.6 show the purity and stability for the final binning definition. The Madgraph + MCFM + Phantom set of samples is used for all the variables, except $m_{4\ell}$ for which the Powheg + MCFM + Phantom set is employed, as explained in Subsection 6.1.3.

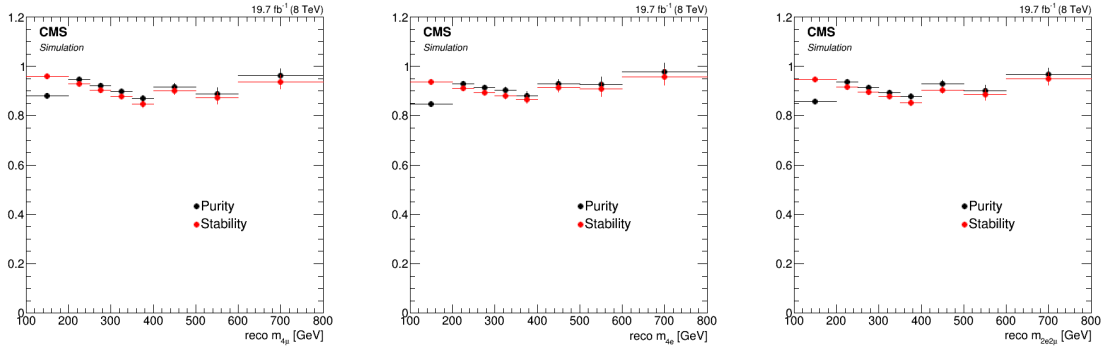


Figure 6.1: Purity and stability as a function of the four-lepton invariant mass, for the 4μ , $4e$ and $2e2\mu$ final states.

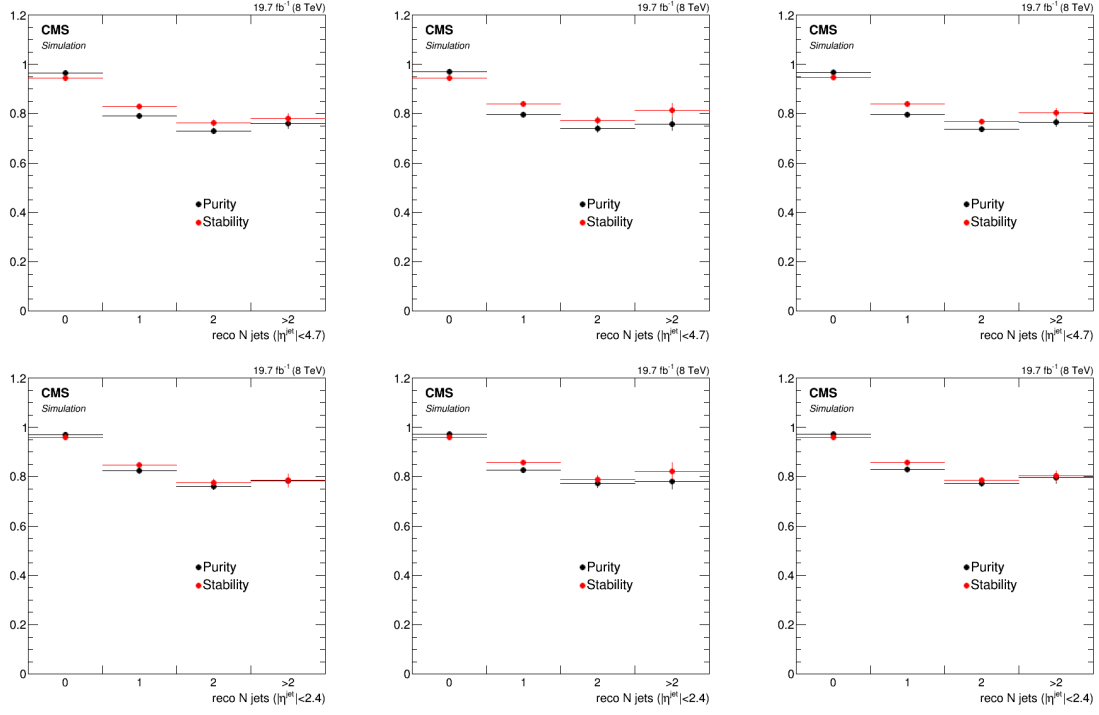


Figure 6.2: Purity and stability as a function of the number of jets (top) and central jets (bottom) in the event, for the 4μ , $4e$ and $2e2\mu$ final states.

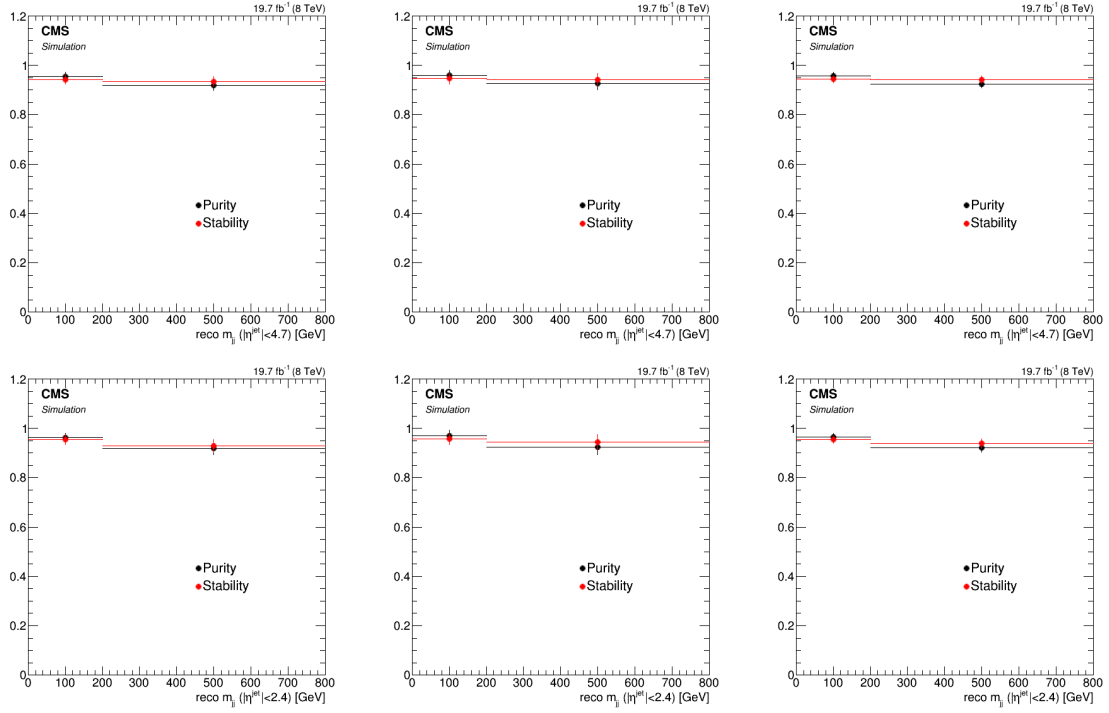


Figure 6.3: Purity and stability as a function of the invariant mass of the two p_T -leading jets (top) and central jets (bottom) in the event, for the 4μ , $4e$ and $2e2\mu$ final states.

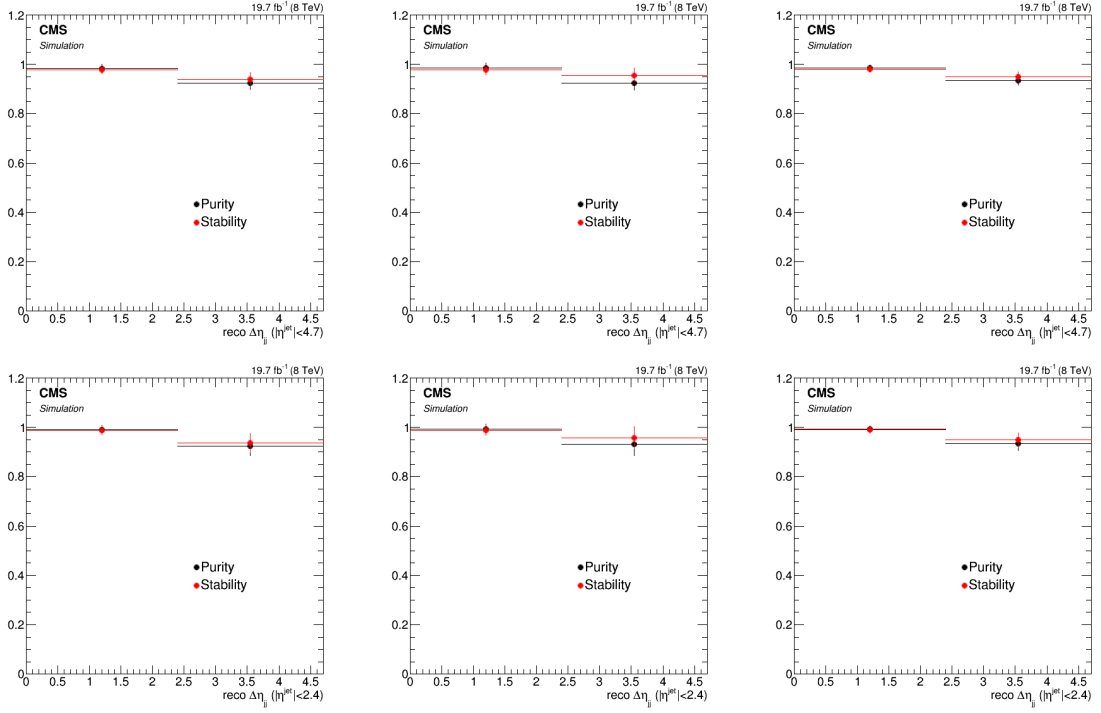


Figure 6.4: Purity and stability as a function of the $\Delta\eta$ between the two p_T -leading jets (top) and central jets (bottom) in the event, for the 4μ , $4e$ and $2e2\mu$ final states.

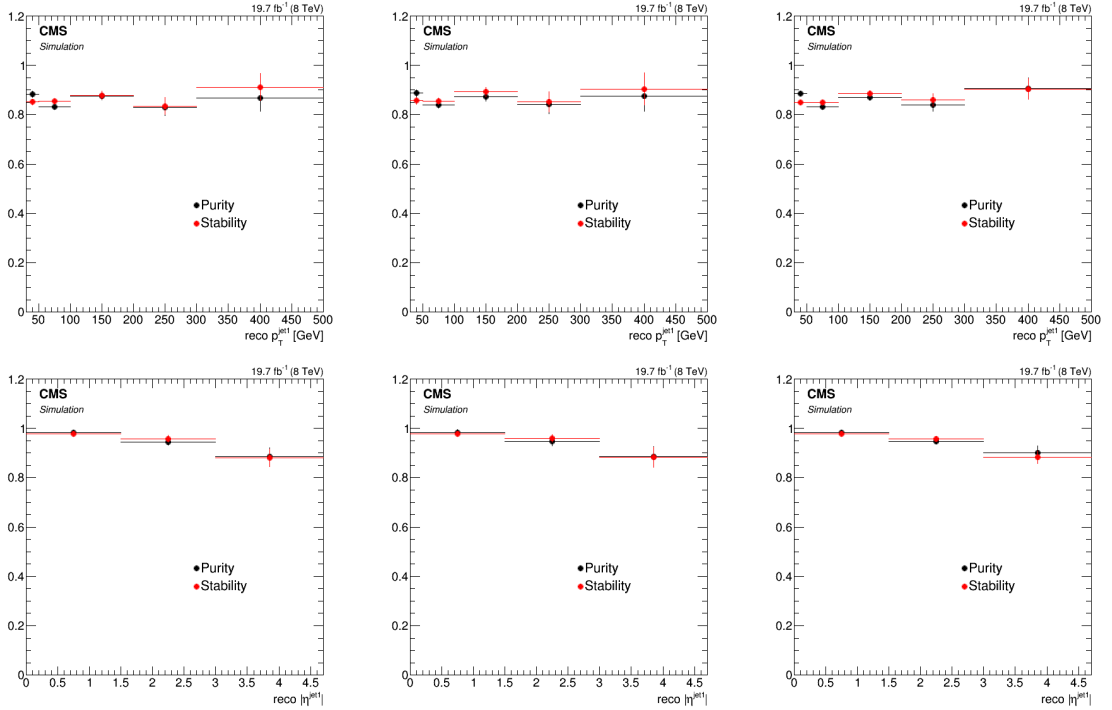


Figure 6.5: Purity and stability as a function of the p_T (top) and η (bottom) of the p_T -leading jet in the event, for the 4μ , $4e$ and $2e2\mu$ final states.

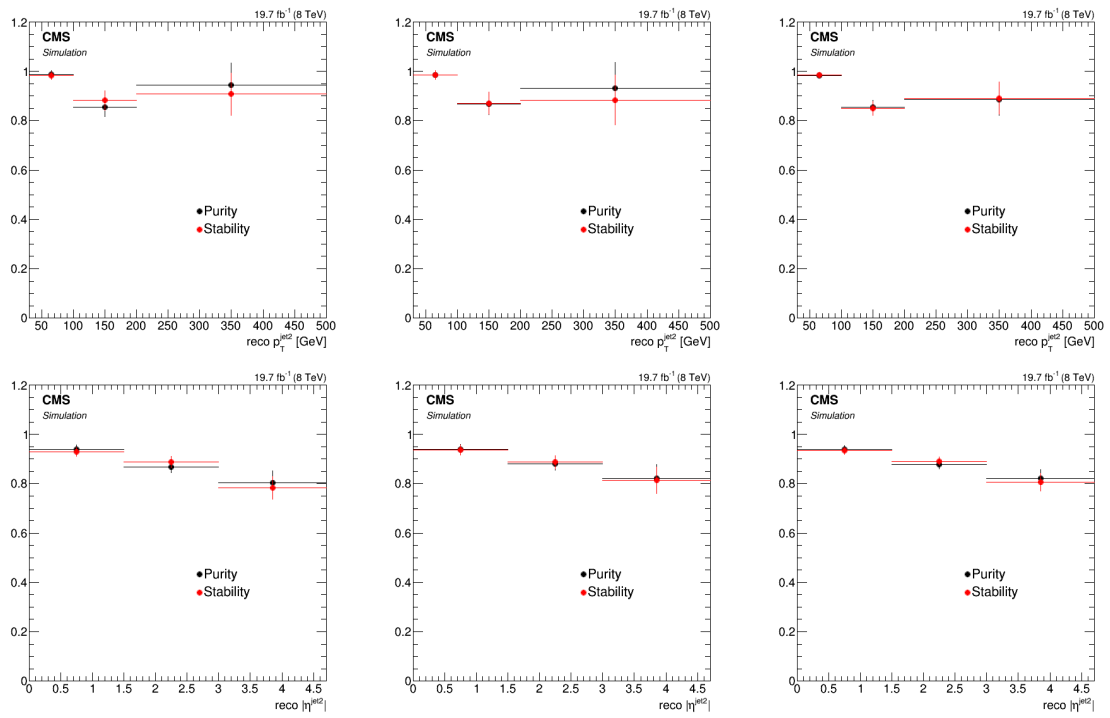


Figure 6.6: Purity and stability as a function of the p_T (top) and η (bottom) of the second p_T -leading jet in the event, for the 4μ , $4e$ and $2e2\mu$ final states.

6.1.3 Building the Response Matrix

Once the final binning is determined for all the analyzed variables, response matrices are built. As already introduced, the response matrix is the basic element on which the unfolding procedure is based, since it describes the bin migration between the reconstructed and true distribution of a chosen observable. It is obtained from Monte Carlo samples that contain signal-only events with pile-up reweighting, lepton and trigger efficiencies applied. Leptons are generated requiring $m_{\ell+\ell^-} > 4$ GeV in all samples but MadGraph, in which $m_{\ell+\ell^-} > 12$ GeV. Jets are generated with $p_T > 10$ GeV, but considered in the collection of generated jets only if their transverse momentum is $p_T > 30$ GeV. Leptons and jets are reconstructed following the same criteria reported in Chapter 4 and the ZZ best candidate is chosen as described in Subsection 5.2.1. At generation level events are taken into account if they enter the *wider* or *tighter* fiducial regions, defined in Section 5.4.

The response object needed for the data unfolding has three components, as illustrated in Figure 6.7:

- the *response matrix* itself, built using both events that are generated in the fiducial region and selected by the analysis cuts;
- the *generation vector*, made up of all events generated in the fiducial region, even if they are not reconstructed;
- the *reconstruction vector*, composed of all events passing the analysis selection, even if they are not generated in the fiducial region.

Thus, both generated events not measured because of detection or selection inefficiency and reconstructed events not generated as signal are considered, allowing to correct the measurement also for event selection efficiencies and not only for detector resolution effects. Given the different efficiencies, separate matrices are evaluated for the three different final states.

For each measured distribution, a response matrix is evaluated using two different sets of generators: the first one is composed of signal samples generated with MadGraph ($qq/qg/gg \rightarrow ZZ$), MCFM ($gg \rightarrow ZZ$) and Phantom ($qq \rightarrow ZZ + 2jets$). The second one has the Powheg sample ($qq \rightarrow ZZ$) instead of the MadGraph one. The MadGraph set is the reference set for the jet-related variables, while the Powheg one is used for check and comparison purposes. For the $m_{4\ell}$ distribution the role of the two sets of samples is switched.

Response matrices built using the reference set of samples are reported for all the analyzed variables in Figures 6.8 - 6.16 and are obtained in the tighter fiducial region (response matrices derived in the wider region are used in the analysis, but not reported).

6.1.4 Choice of the Unfolding Algorithm

As already illustrated in the previous Sections, the exact inverse of the response matrix in the linear system of Equation 6.1, if exists, can produce rapidly oscillating solutions, since small statistical fluctuations can cause large effects in the result. In order to avoid this meaningless behavior, the true distribution must be *smoothed* through a regularization procedure. Several regularized algorithms can be used to unfold distributions, such as the Singular Value Decomposition and the iterative Bayesian method: these techniques are based on regularization parameters which can be tuned to obtain results robust against numerical instabilities. In order to choose the proper algorithm, results obtained using the two methods are compared.

The SVD method depends on the choice of the regularization parameter (k_{reg}), which in the

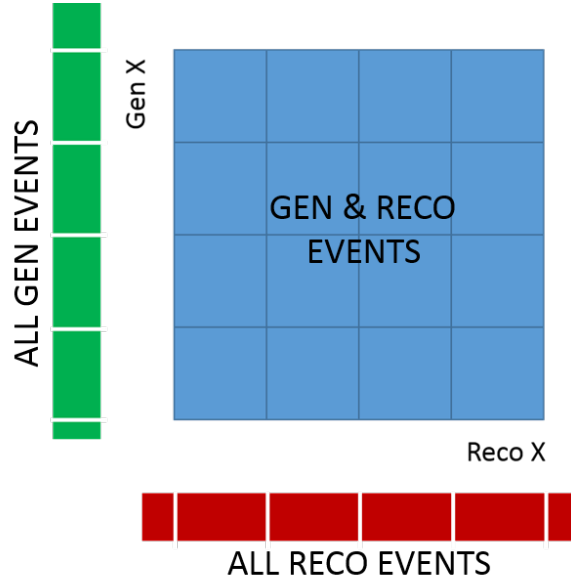


Figure 6.7: Graphic representation of the response object with its three components: the response matrix in blue, built using events generated in the fiducial region and selected by the analysis cuts; the generation vector in green, made of all events generated in the fiducial region; the reconstruction vector in red, composed of all events passing the analysis selection

range $[1, N_{bin}]$. The optimal value of k_{reg} is the one for which errors associated to the unfolding procedure are small when compared to the statistical ones. This parameter can be seen as a cut-off for quickly oscillating terms corresponding to data statistical fluctuations, as described in detail in [81]. Choosing a too small regularization parameter suppresses these spurious fluctuations but the result tends to be biased by the MC truth. On the other hand, a too large regularization parameter will decrease the MC dependence but give too large importance to data fluctuation which will be interpreted as real shape.

The D’Agostini method [82], based on the application of the Bayes theorem, uses an iterative approach and requires a choice for the maximum number of iterations to be done, that should not lead to different results. In this analysis, the iteration procedure is stopped as soon as the new unfolded distribution is compatible with the one obtained from the previous step. This is quantified by computing the $\chi^2/n.d.f.$ of the change in the unfolded distribution for every iteration step. The number of iterations in the Bayes unfolding procedure is set as the one when the next iteration is compatible with the current one by computing:

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

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 systematic uncertainty associated to b_n^i .

The unfolded $m_{4\ell}$ distributions are obtained using the SVD algorithm. In this case the regularization parameter is chosen to be $k_{reg} = 4$, the default value ($N_{bin}/2$), since as shown

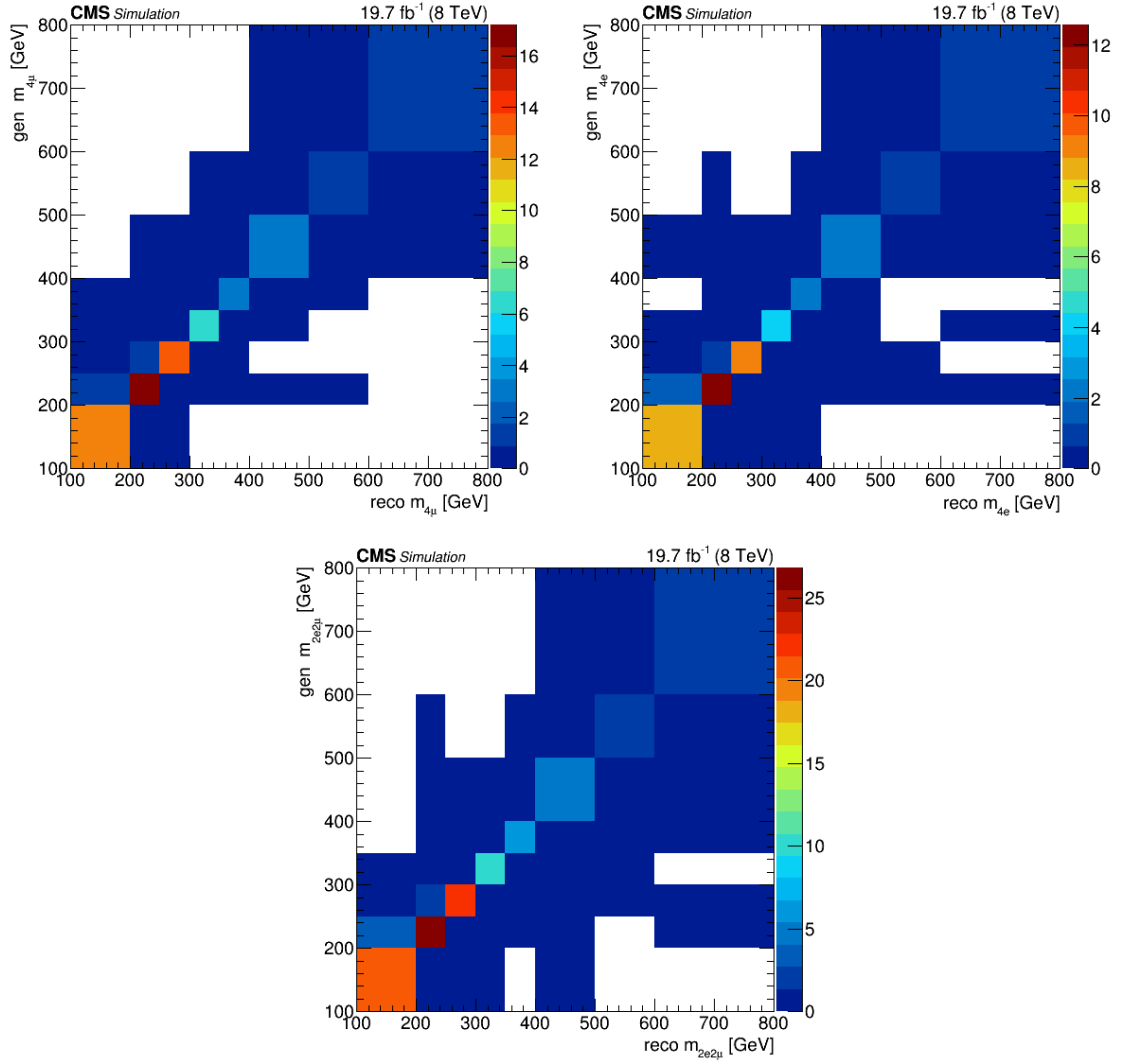


Figure 6.8: Response matrices for the $m_{4\ell}$ distribution, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the Powheg set of samples. The tight fiducial region is considered.

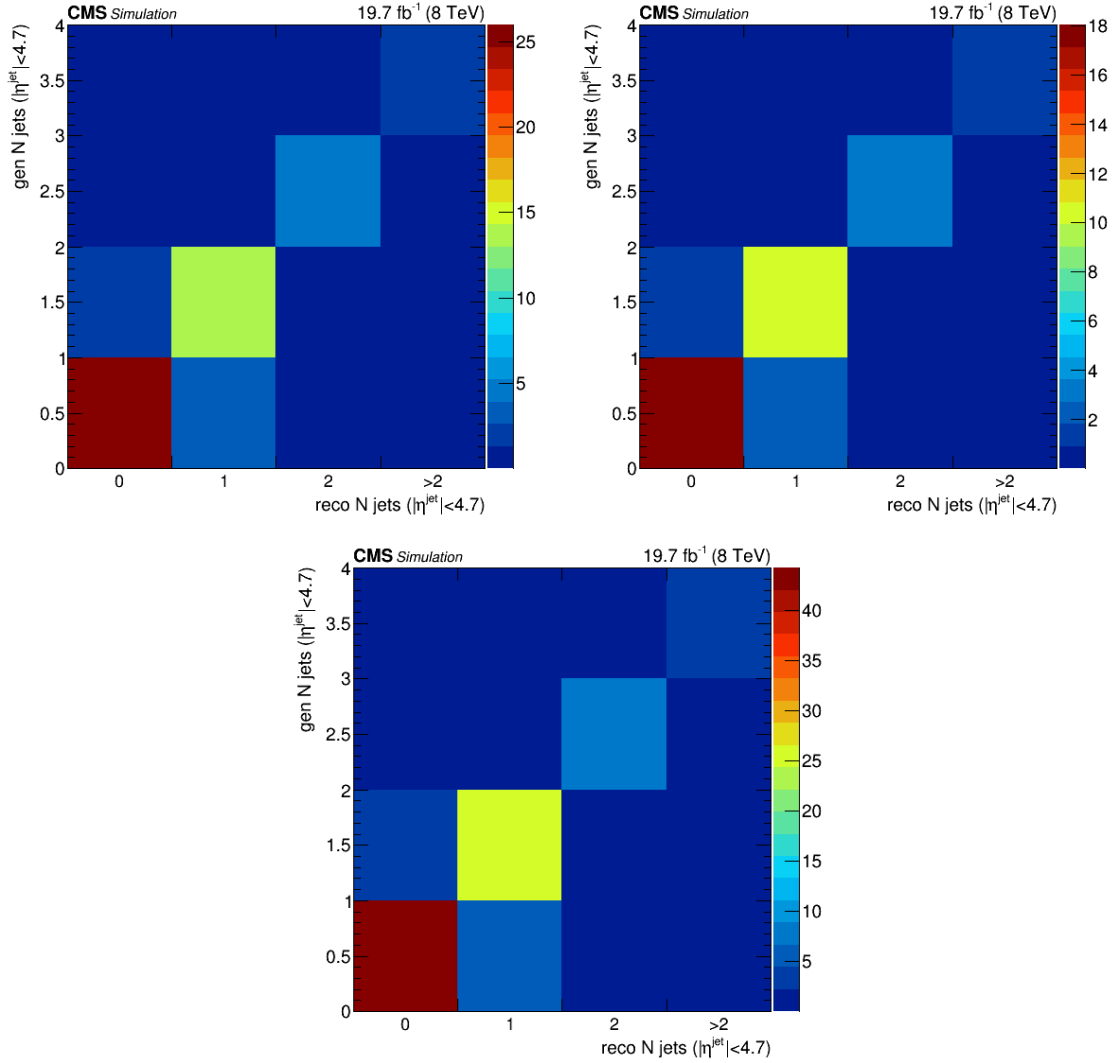


Figure 6.9: Response matrices for the distribution of the number of jets (with $|\eta^{jet}| < 4.7$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

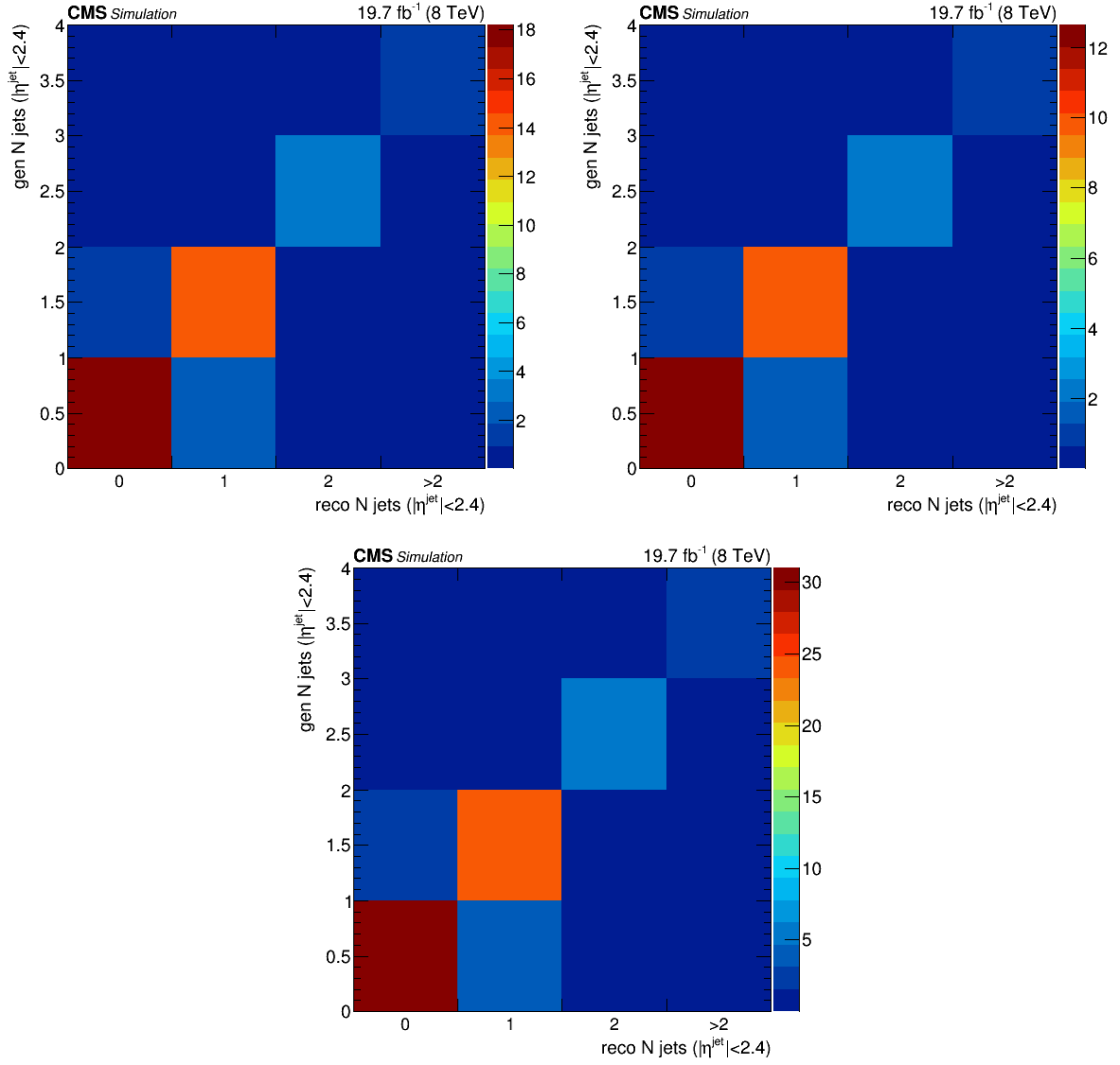


Figure 6.10: Response matrices for the distribution of the number of jets (with $|\eta^{jet}| < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

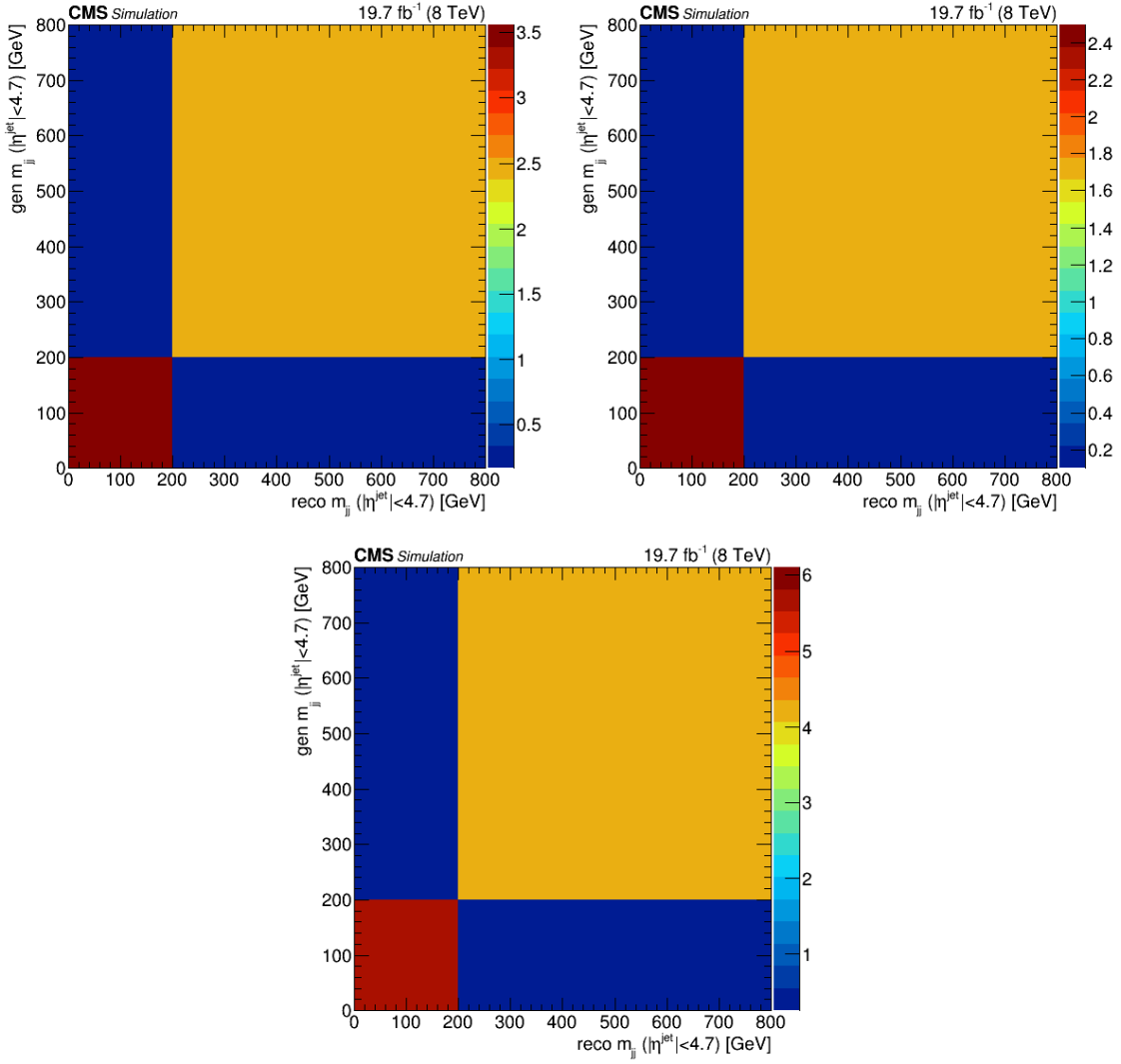


Figure 6.11: Response matrices for the m_{jj} distribution (with $|\eta^{jet}| < 4.7$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

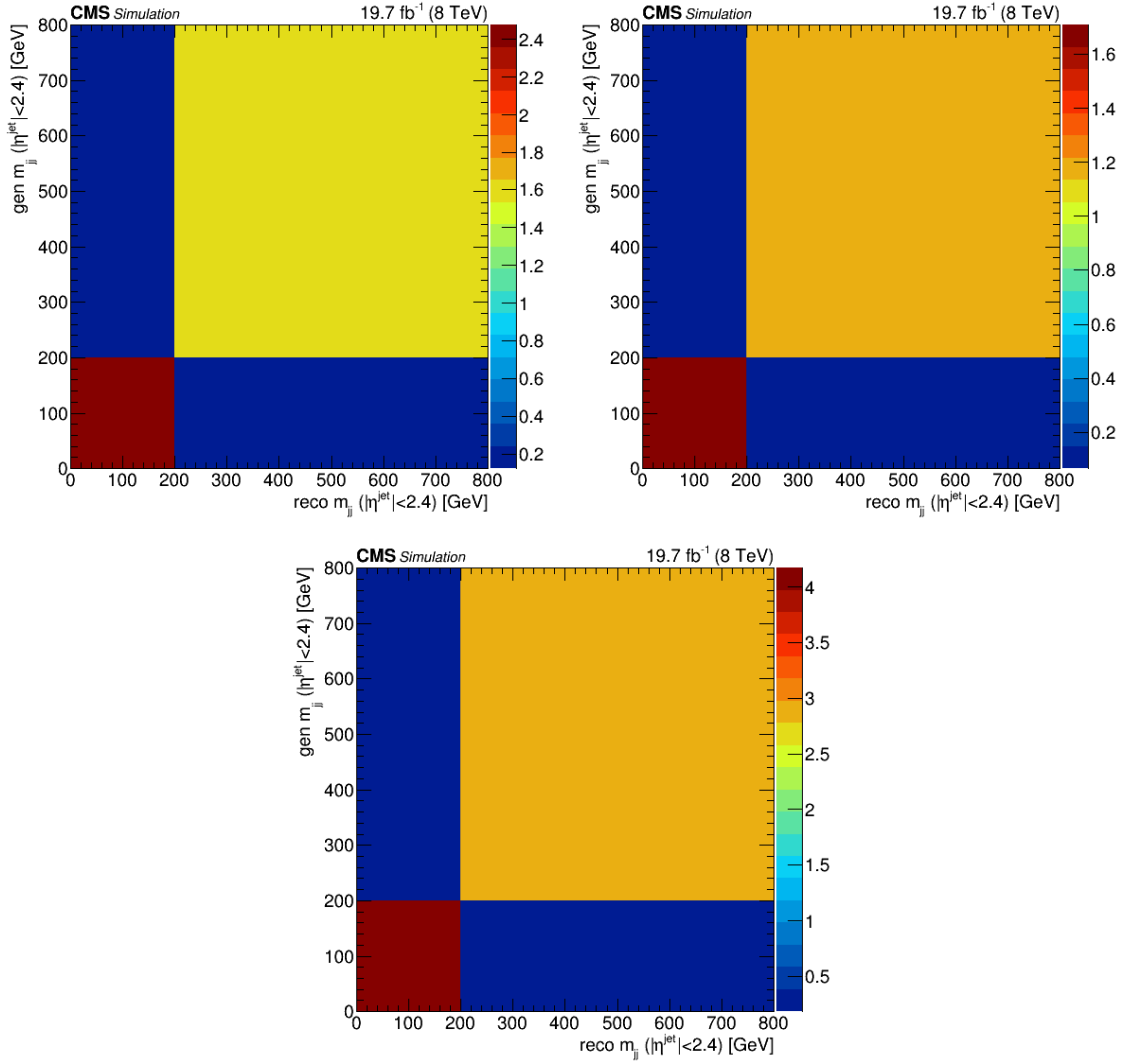


Figure 6.12: Response matrices for the m_{jj} distribution (with $|\eta^{jet}| < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

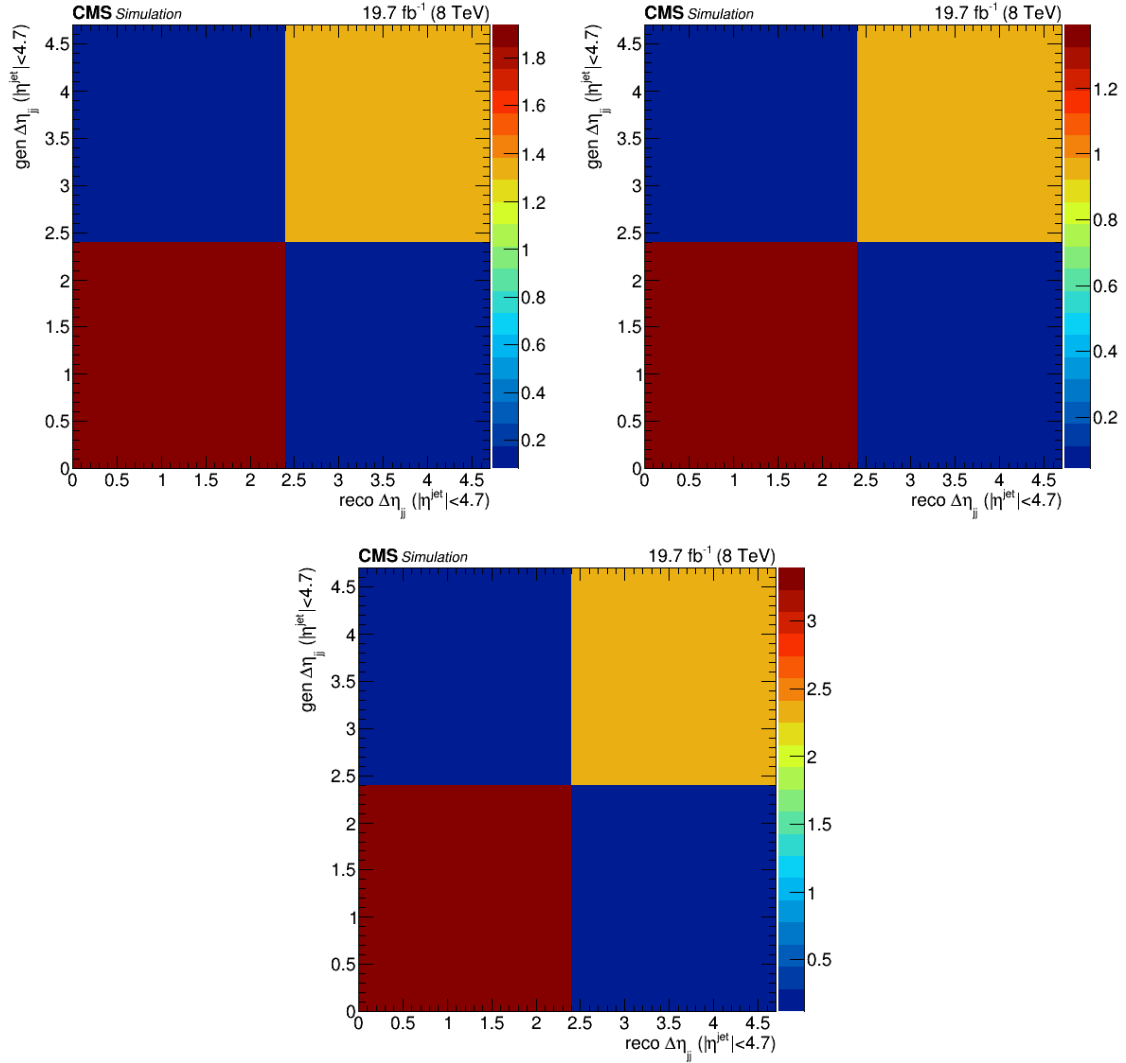


Figure 6.13: Response matrices for the $\Delta\eta_{jj}$ distribution (with $|\eta^{\text{jet}}| < 4.7$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

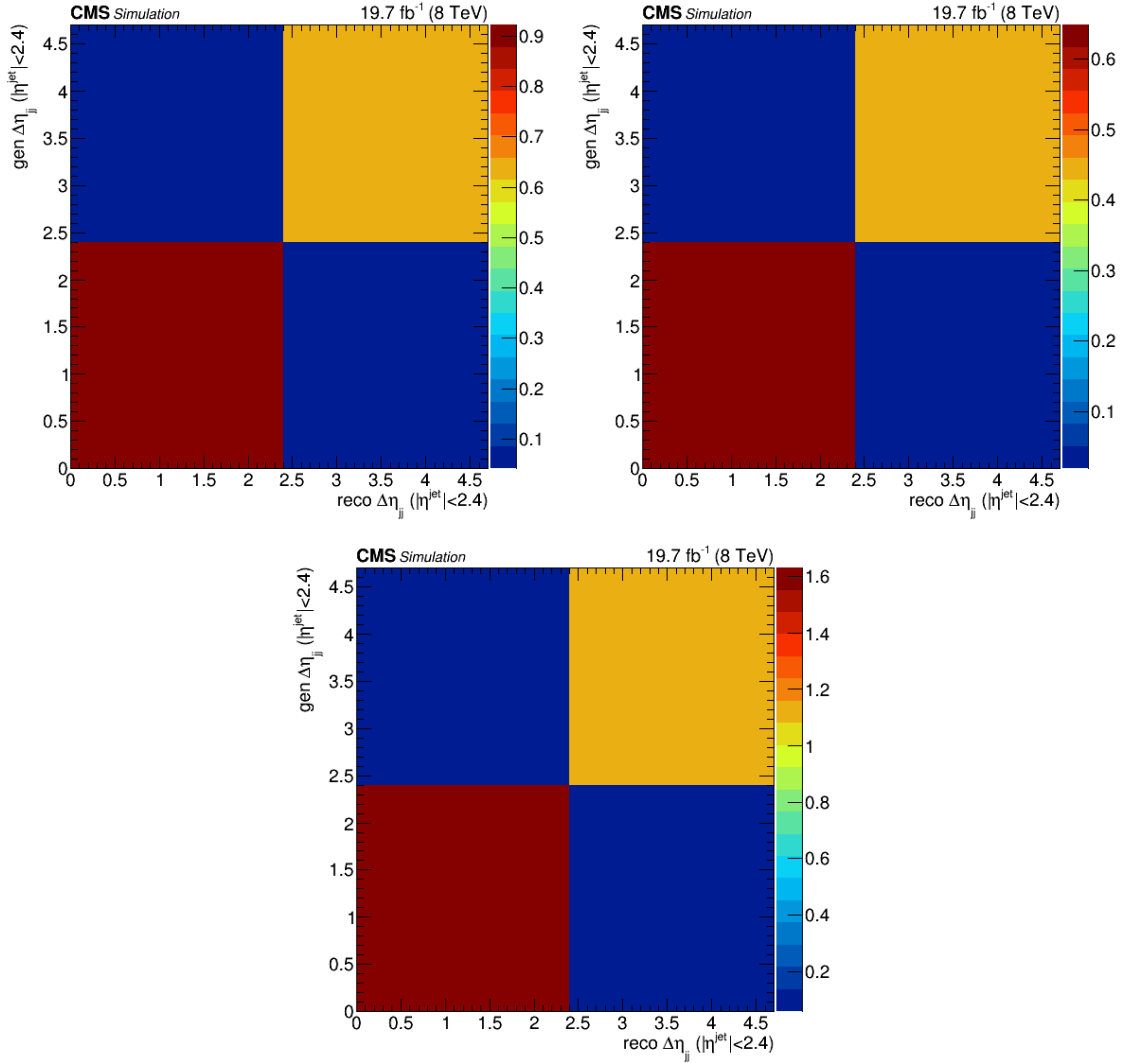


Figure 6.14: Response matrices for the $\Delta\eta_{jj}$ distribution (with $|\eta^{\text{jet}}| < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

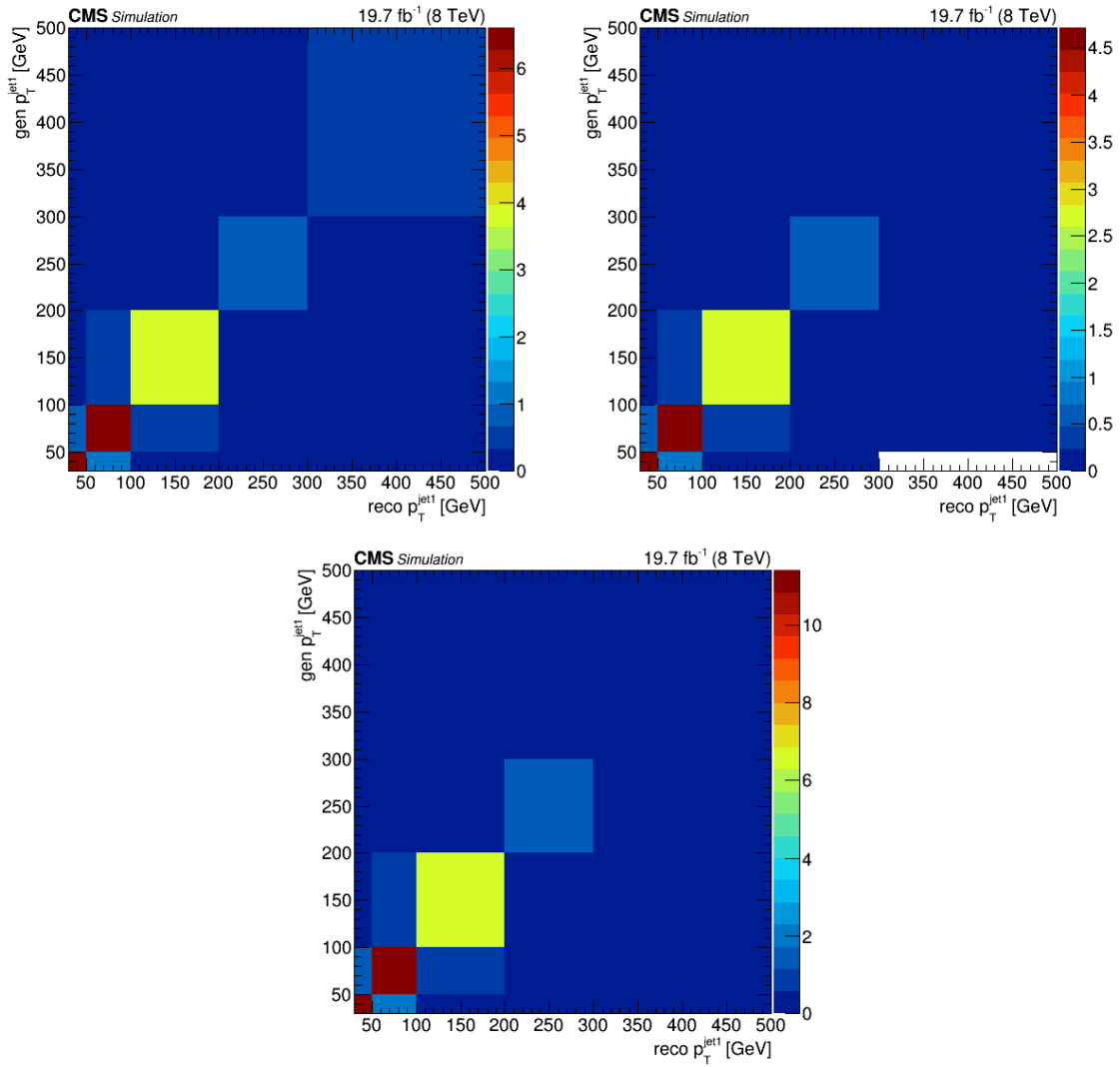


Figure 6.15: Response matrices for the p_T distribution of the p_T -leading jet, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

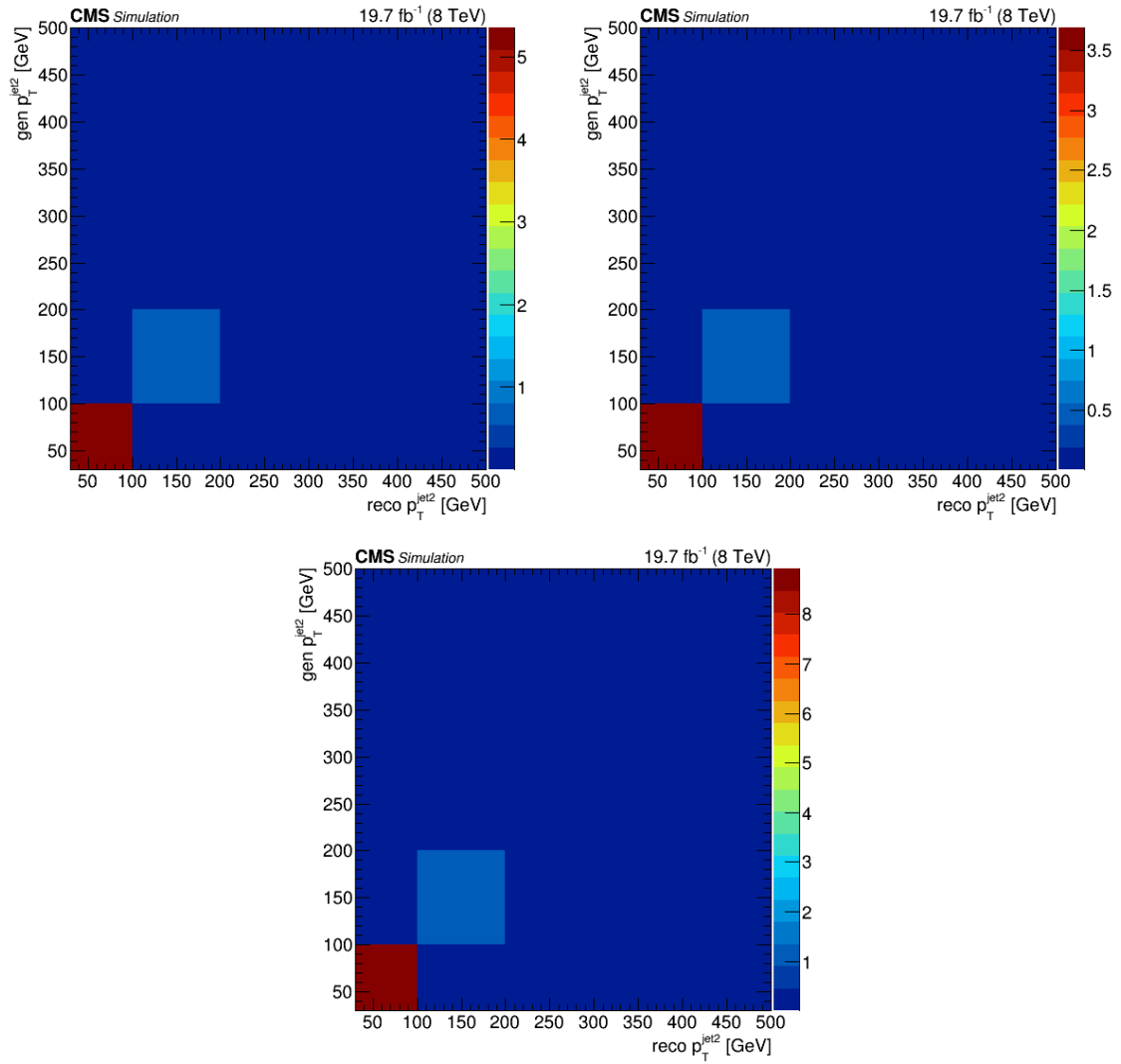


Figure 6.16: Response matrices for the p_T distribution of the sub- p_T -leading jet, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

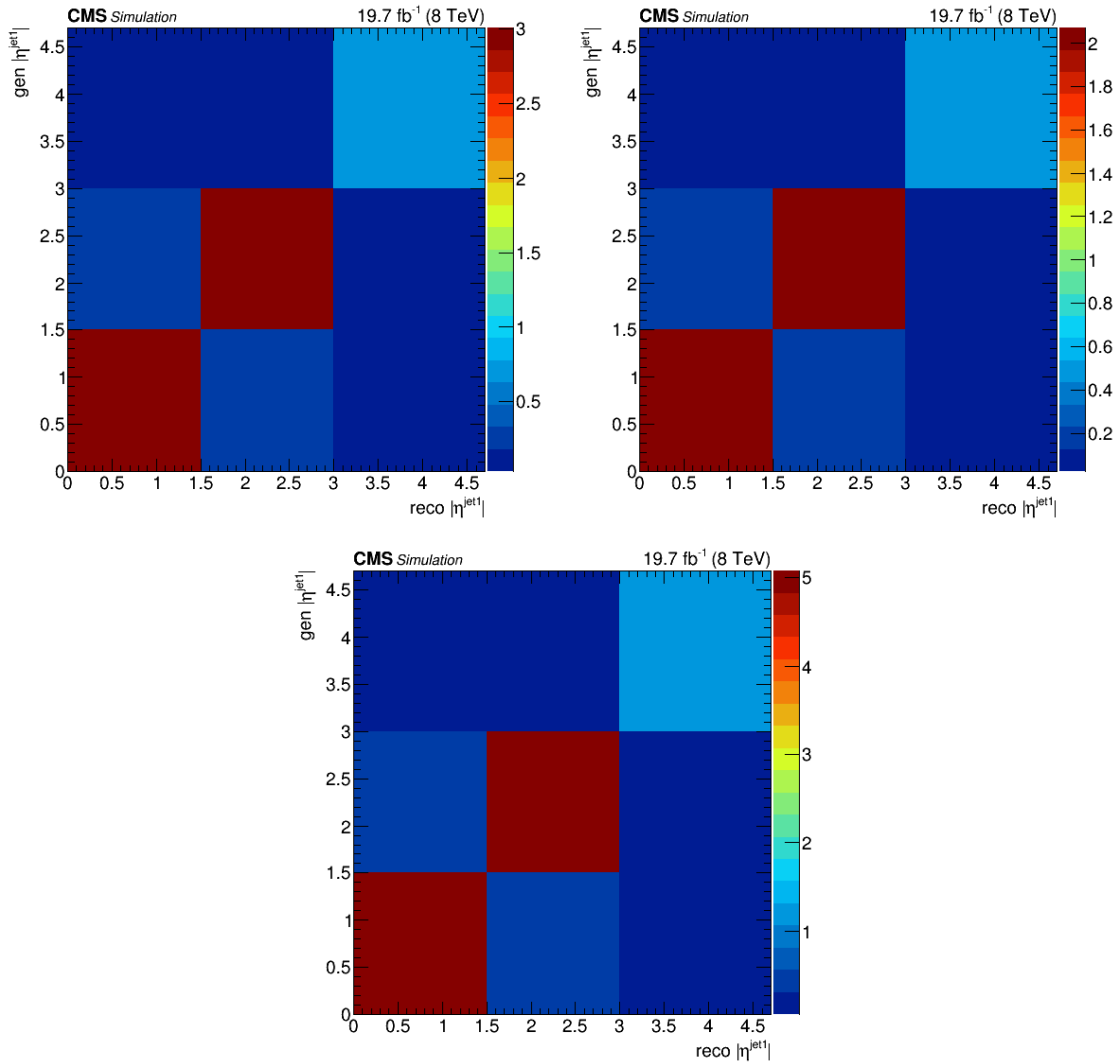


Figure 6.17: Response matrices for the η distribution of the p_T -leading jet, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

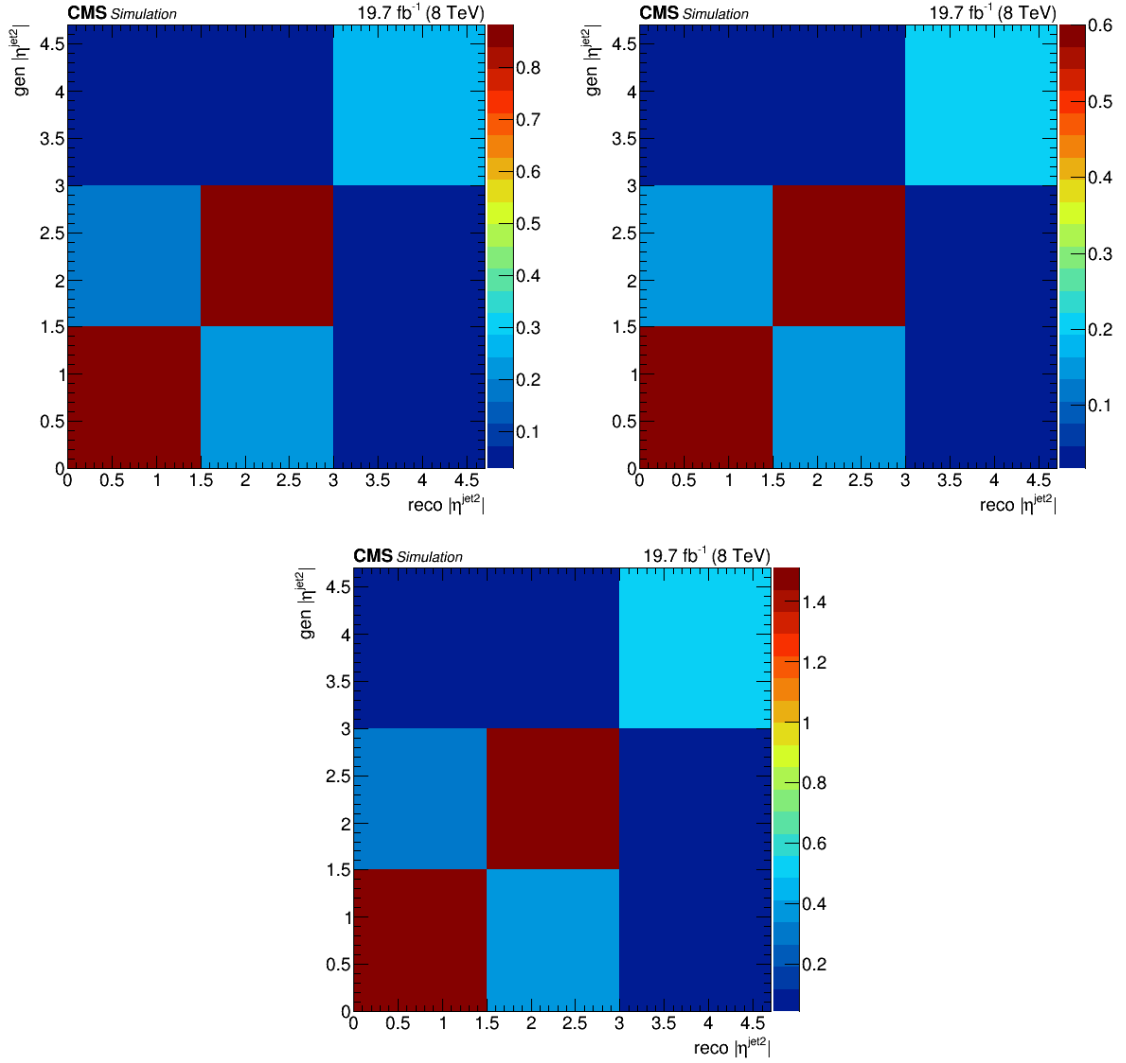


Figure 6.18: Response matrices for the η distribution of the sub- p_T -leading jet, according to the final state: 4 μ (top left), 4 e (top right), 2 $e2\mu$ (bottom). Matrices are obtained using the MadGraph set of samples. The tight fiducial region is considered.

in Figure 6.19 the distributions obtained using $k_{reg} = 2$ are clearly biased by the MC truth. On the other hand, all the other jet-related variables distributions are unfolded using the Bayesian algorithm, with two iterations. As shown in Figure 6.20, this choice is not biased by the MC truth distribution.

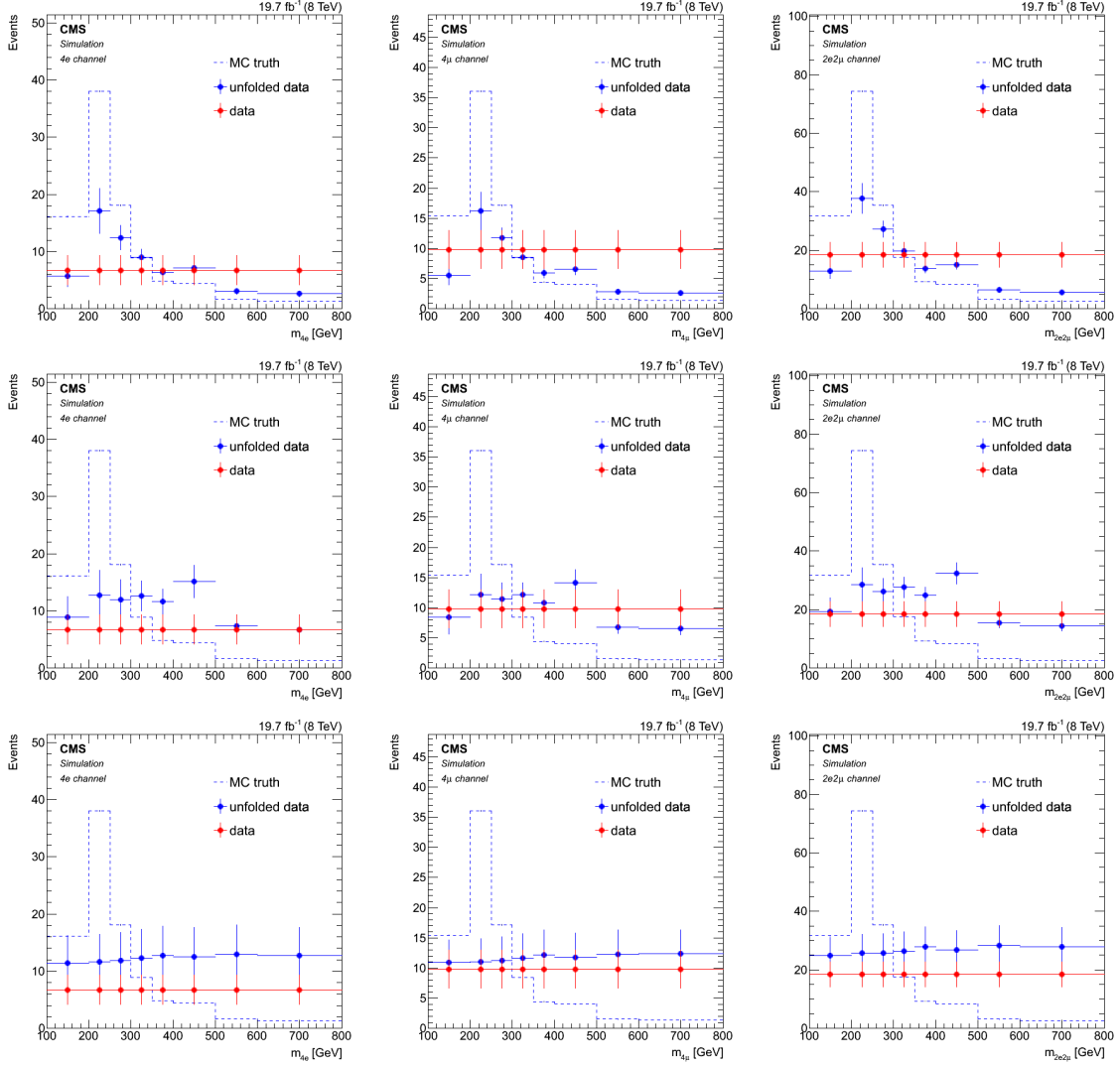


Figure 6.19: A flat distribution (in red) is unfolded using the SVD algorithm with $k_{reg} = 2$ (top) and $k_{reg} = 4$ (center) and using the D'Agostini method with 2 iterations (bottom). Results are reported as a function of the invariant mass of the four-lepton system and for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final state.

6.1.5 Unfolding Validation Tests on Monte Carlo Samples

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. In each distribution the ratio of unfolded over

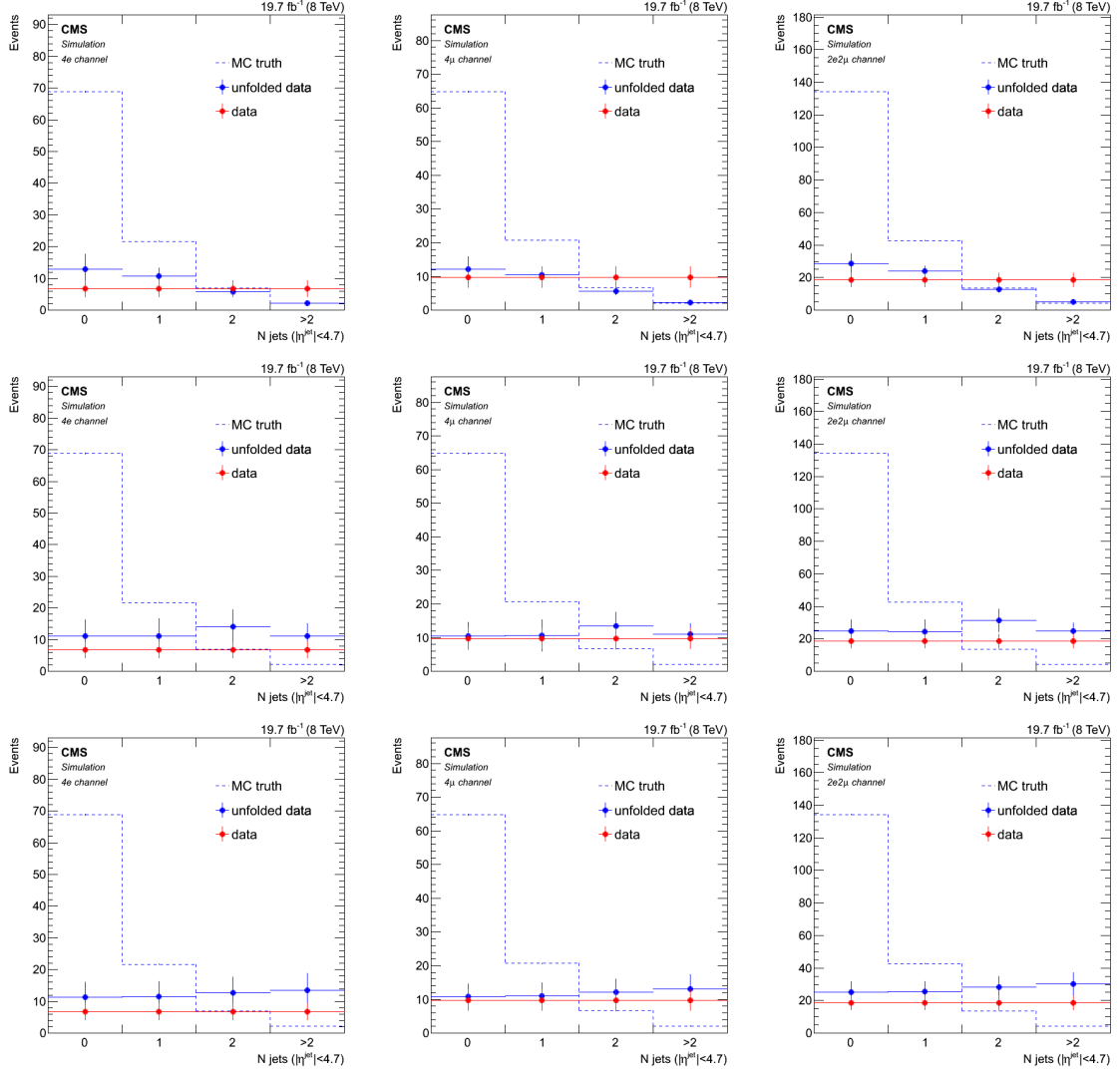


Figure 6.20: A flat distribution (in red) is unfolded using the SVD algorithm with $k_{reg} = 2$ (top) and $k_{reg} = 4$ (center) and using the D'Agostini method with 2 iterations (bottom). Results are reported as a function of the number of jets and for the $4e$ (left), 4μ (center) and $2e2\mu$ (right) final state.

generated output is as expected unity, demonstrating that the unfolding procedure does not introduce any bias.

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 employed 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. Closure tests produce a very good agreement between unfolded and truth distributions, showing no dependence on the particular Monte Carlo generator used for the unfolding, demonstrating the procedure effectiveness.

As an example, results of validation tests for the number-of-jets distribution in the $2e2\mu$ final state are reported in Figure 6.21. The whole set of plots, for all different variables and final states, is shown in Appendix A.

6.2 Unfolded Distributions

Distributions are first unfolded by estimating the number of events in the wider fiducial region. In order to minimize the model uncertainties due to unnecessary extrapolations of the measurement outside experimentally well-described phase space regions, distributions are then unfolded also in the tighter fiducial region, corresponding to the visible phase space defined by the kinematic and geometrical acceptance of leptons.

For the $m_{4\ell}$ 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 2 iterations. The unfolded distributions are presented in Figures 6.22 - 6.32, in which the content of each bin is divided by its width. The result of the unfolding procedure through the application of the response object on reconstructed distributions is presented by superimposing:

- the data distribution at reconstruction level (to which the background is subtracted);
- the MC distribution at reconstruction level, built using only signal samples;
- the MC truth distribution at generation level;
- the unfolded data distribution, at *particle level*, final result of the unfolding procedure.

Results are shown separately for the $4e$, 4μ , $2e2\mu$ final states and summed up in the 4ℓ distribution.

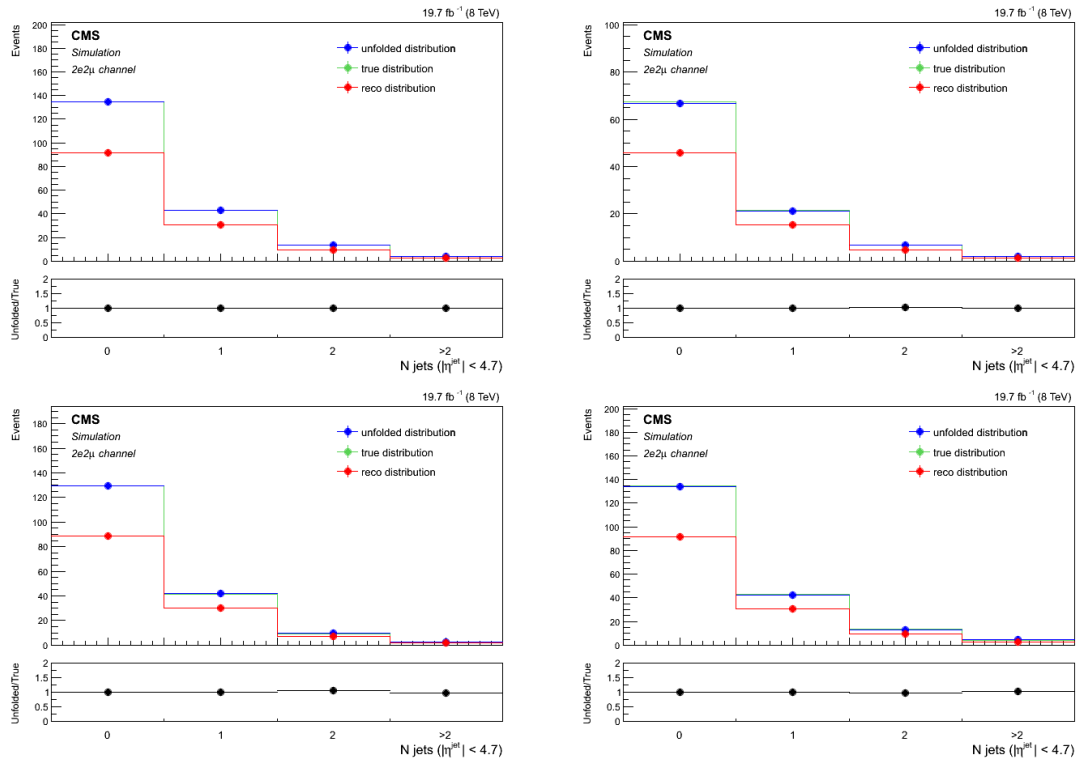


Figure 6.21: Unfolding validation tests of the number-of-jets distribution for the $2e2\mu$ final state: MadGraph matrix applied on MadGraph distribution using the full set (top left), MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set (top right), MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution using the full set. The tight fiducial region is considered.

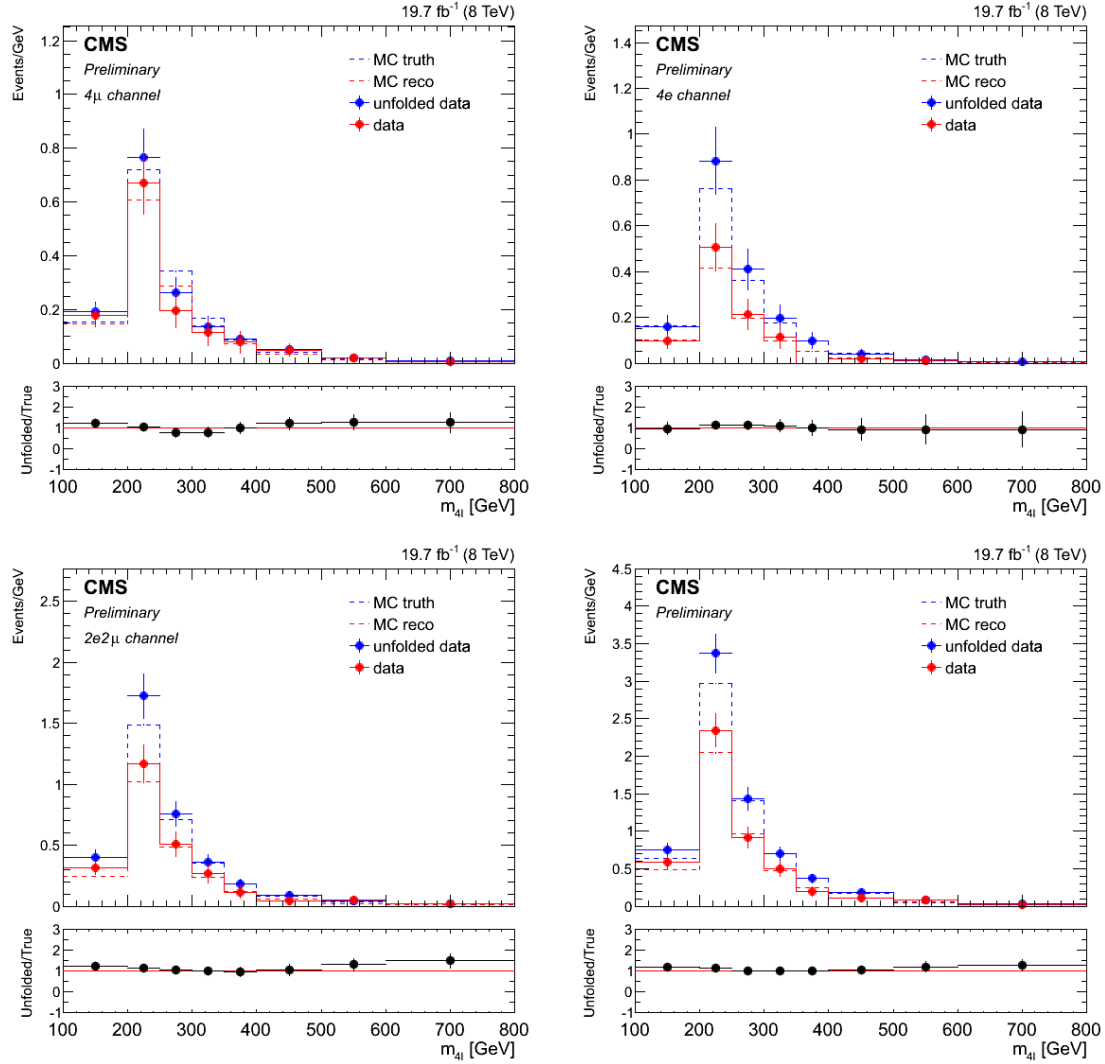


Figure 6.22: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) $m_{4\ell}$ distributions, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). The unfolded distributions are obtained in the tight fiducial region using the SVD algorithm, with $k_{reg} = 4$, and compared to predictions from the Powheg set of samples.

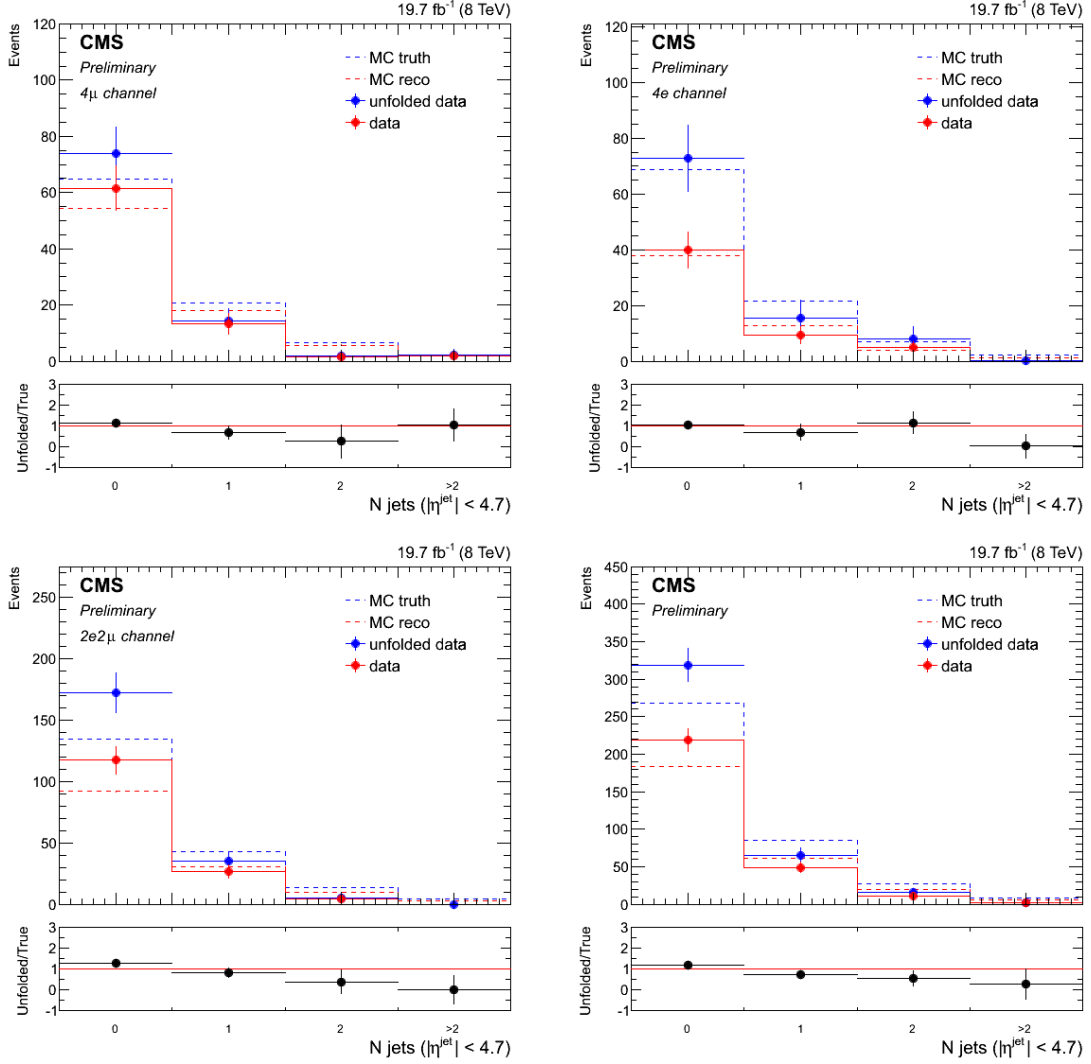


Figure 6.23: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) distributions of the number of jets with $|\eta^{\text{jet}}| < 4.7$, according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

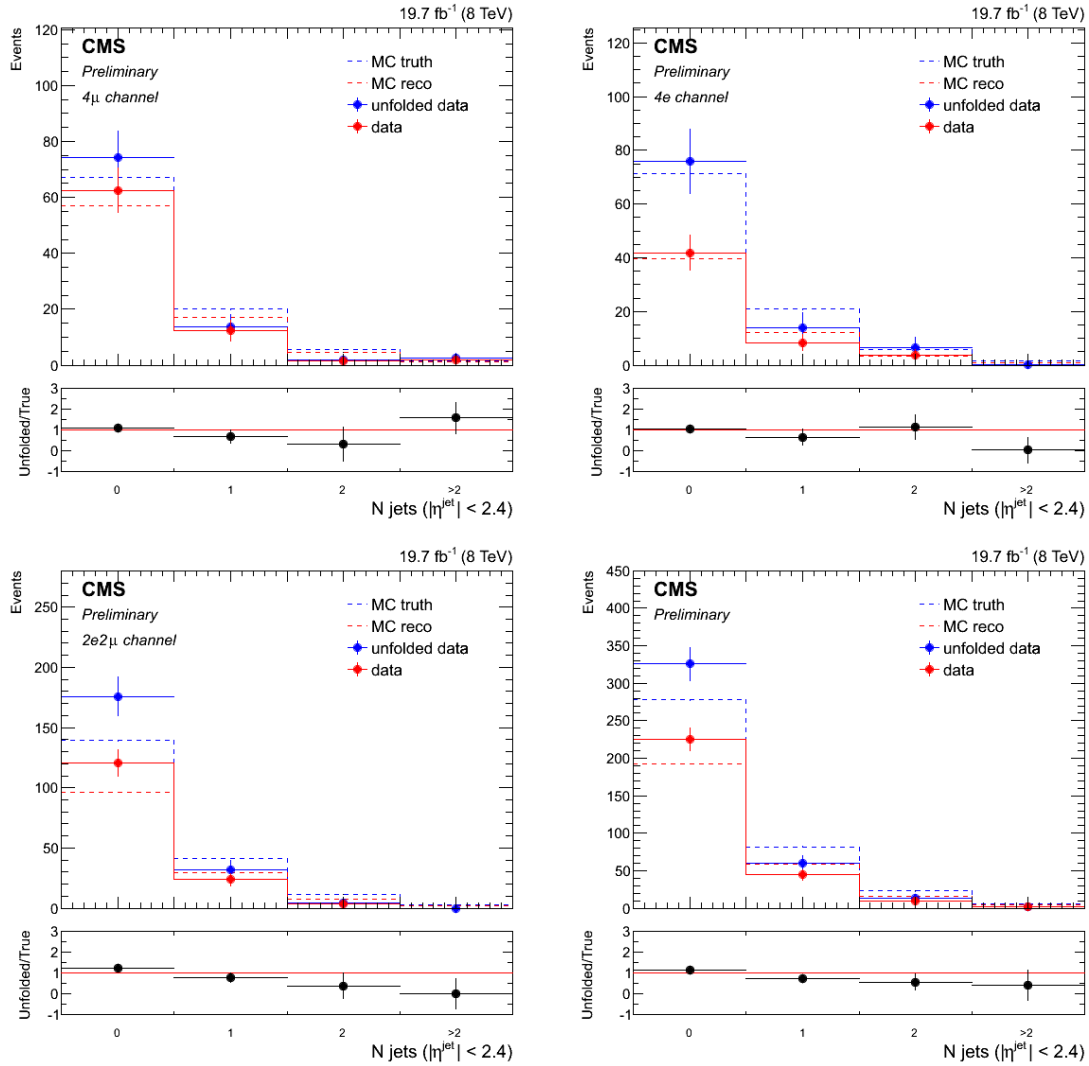


Figure 6.24: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) distributions of the number of jets with $|\eta^{\text{jet}}| < 2.4$, according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

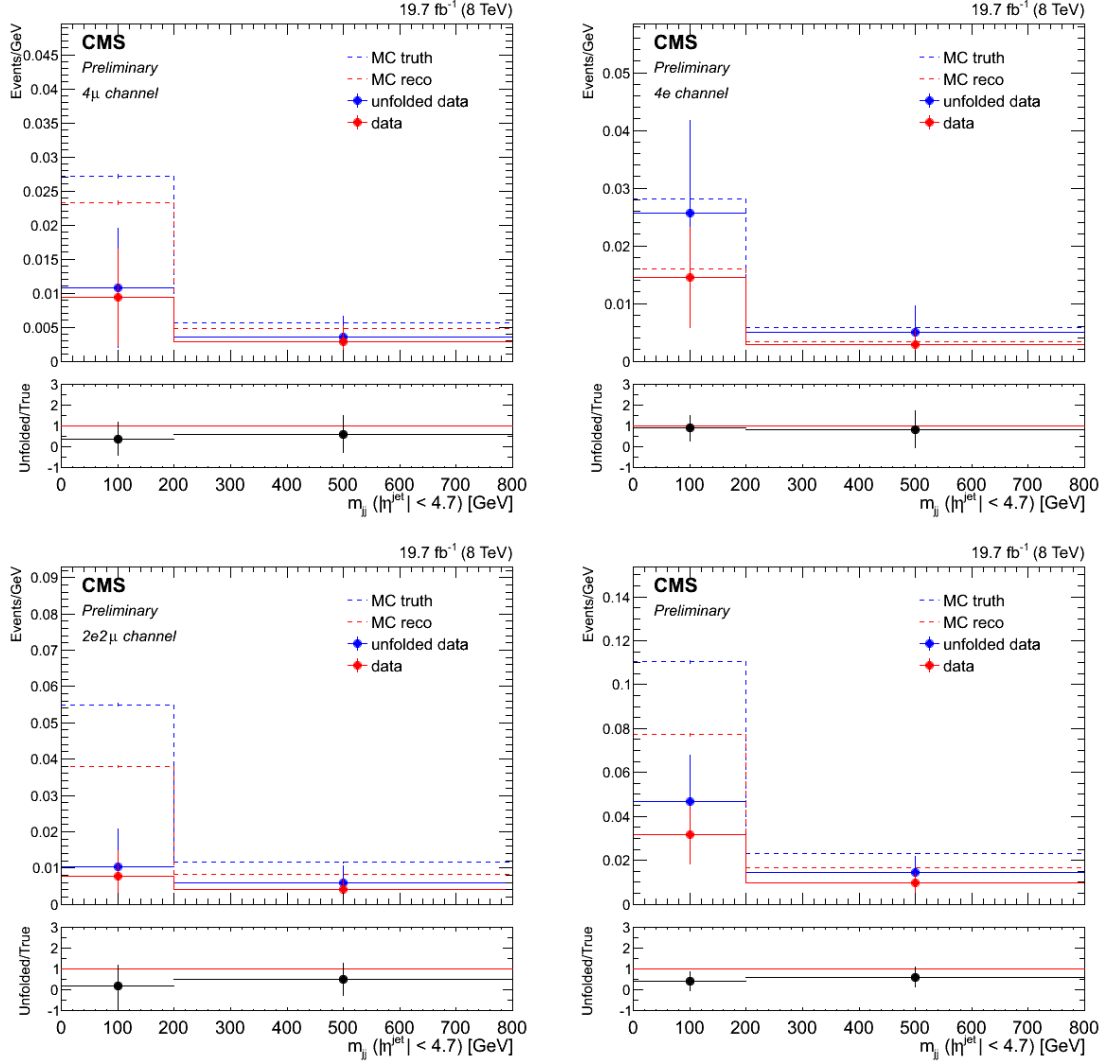


Figure 6.25: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) m_{jj} distributions (with $|\eta^{jet}| < 4.7$), according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

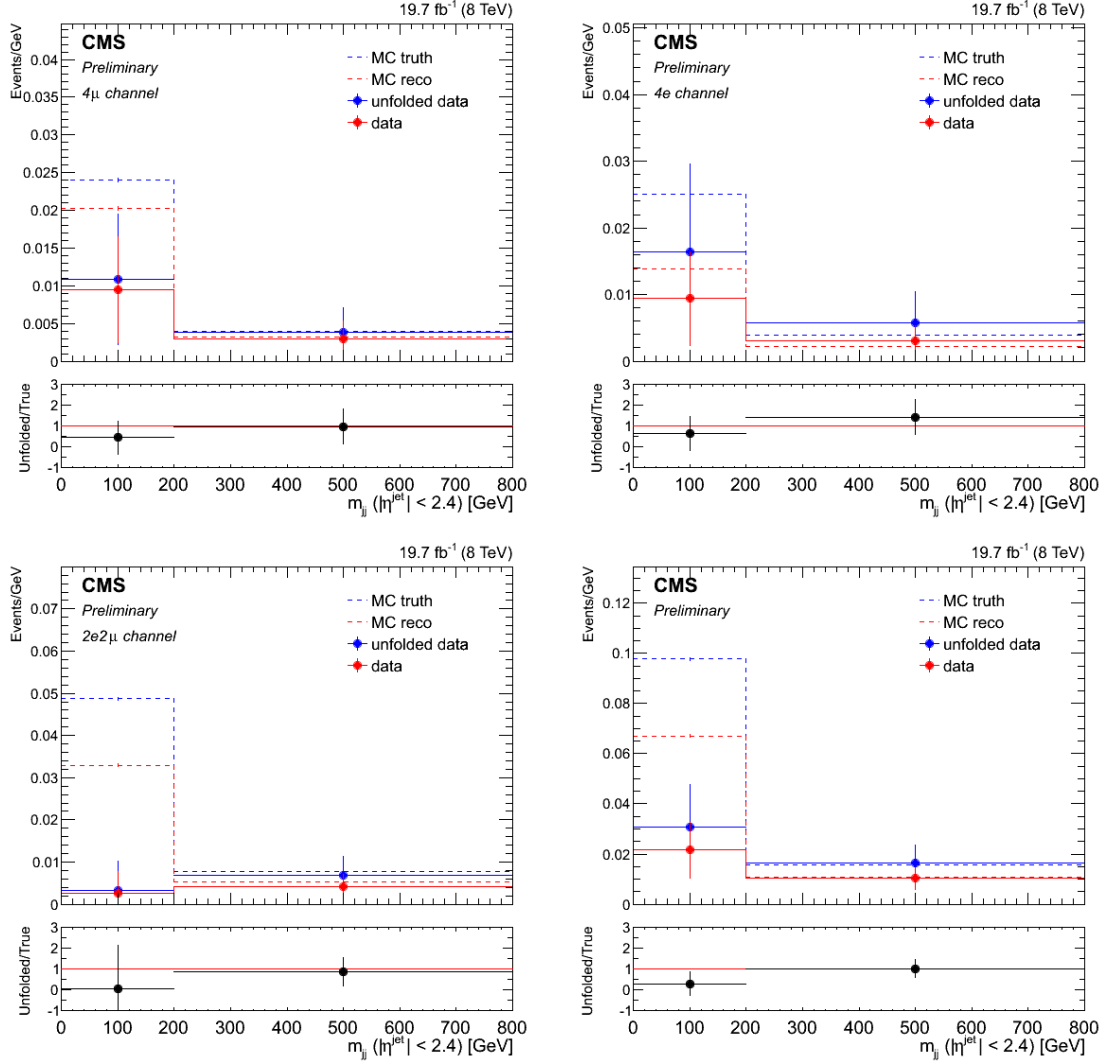


Figure 6.26: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) m_{jj} distributions (with $|\eta^{jet}| < 2.4$), according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

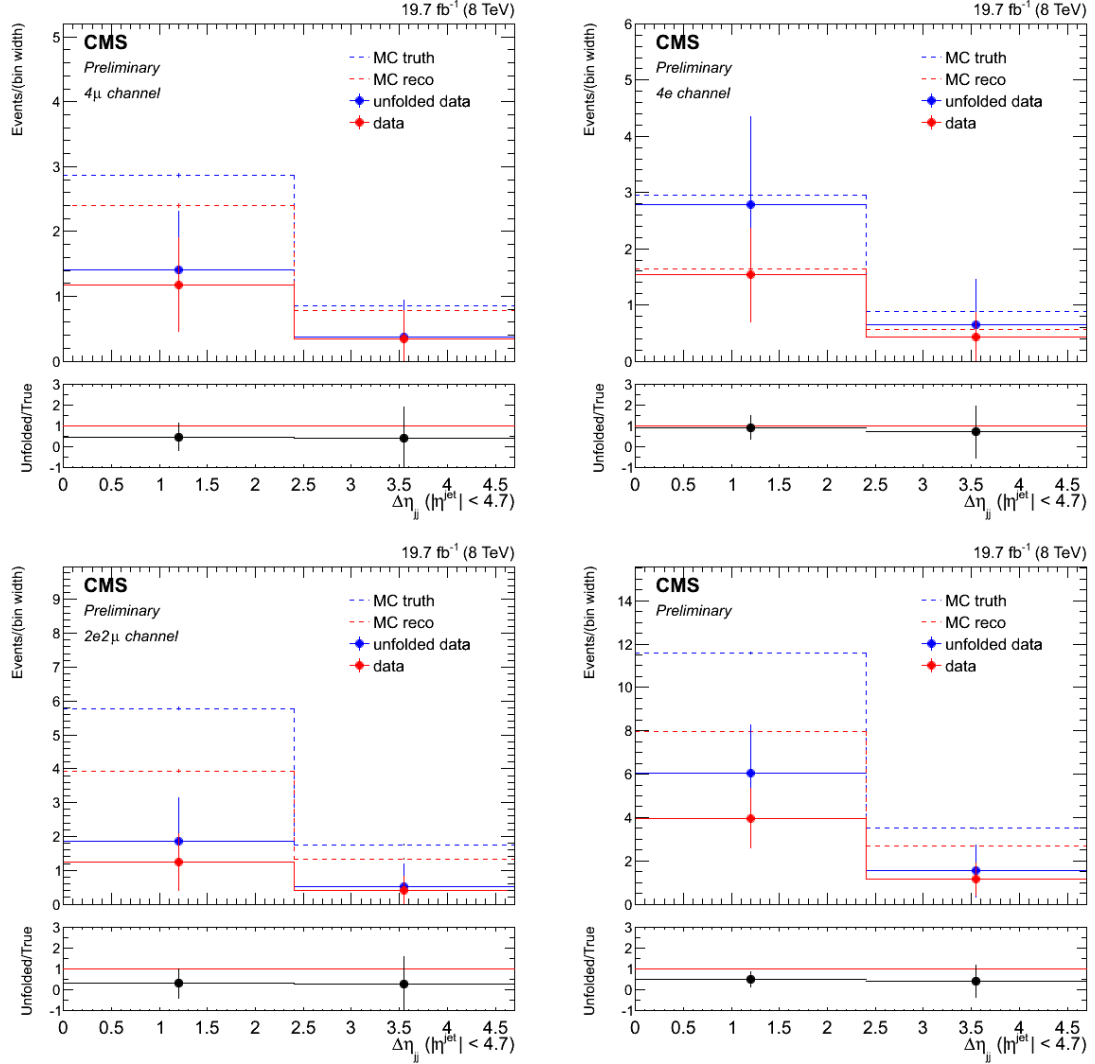


Figure 6.27: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) $\Delta\eta_{jj}$ distributions (with $|\eta^{jet}| < 4.7$), according to the final state: (top left) 4 μ , (top right) 4 e , (bottom) 2 $e2\mu$, (bottom right) 4 ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

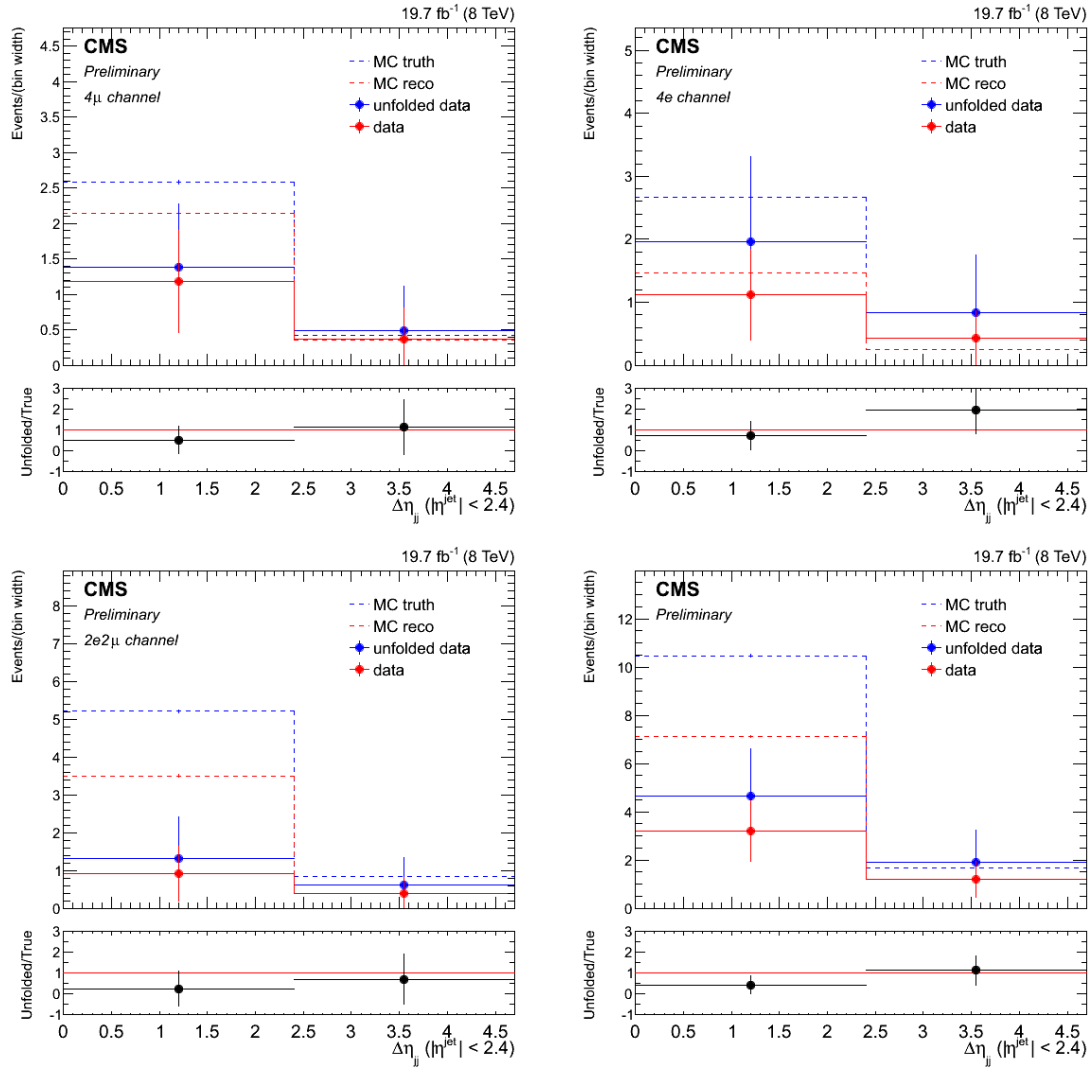


Figure 6.28: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) $\Delta\eta_{jj}$ distributions (with $|\eta^{jet}| < 2.4$), according to the final state: (top left) 4 μ , (top right) 4 e , (bottom) 2 e 2 μ , (bottom right) 4 ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

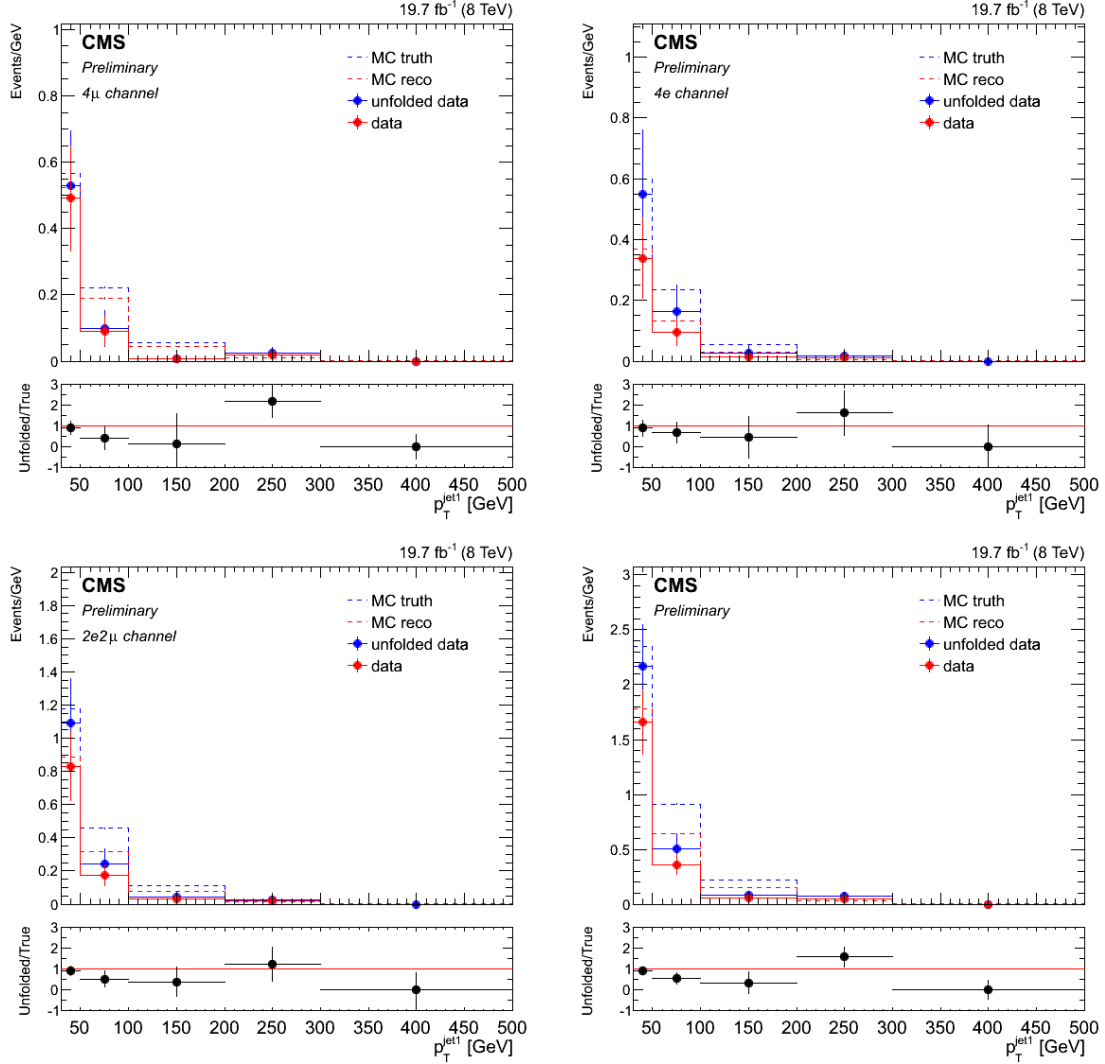


Figure 6.29: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) p_T distributions of the p_T -leading jet, according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

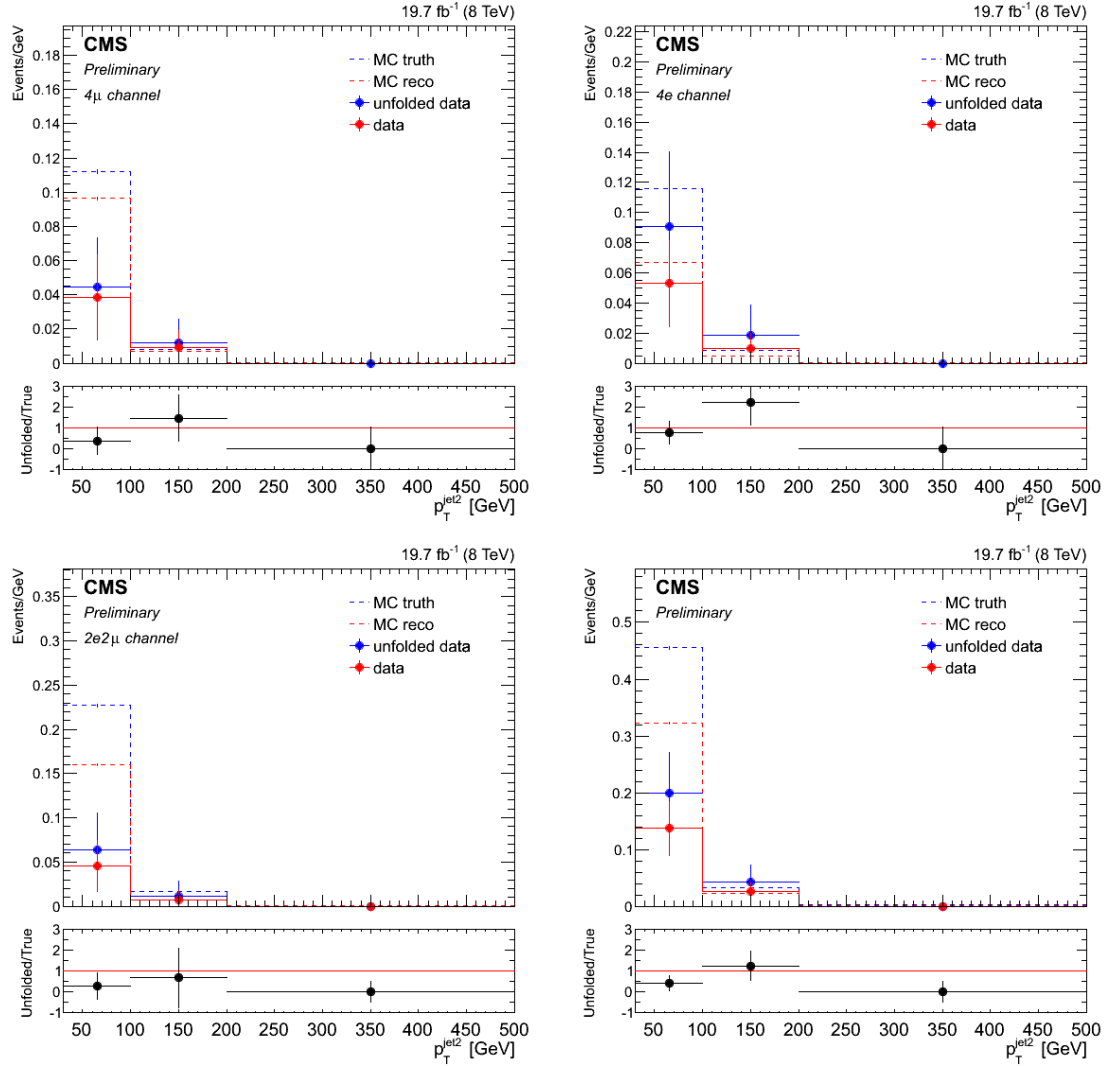


Figure 6.30: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) p_T distributions of the second p_T -leading jet, according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

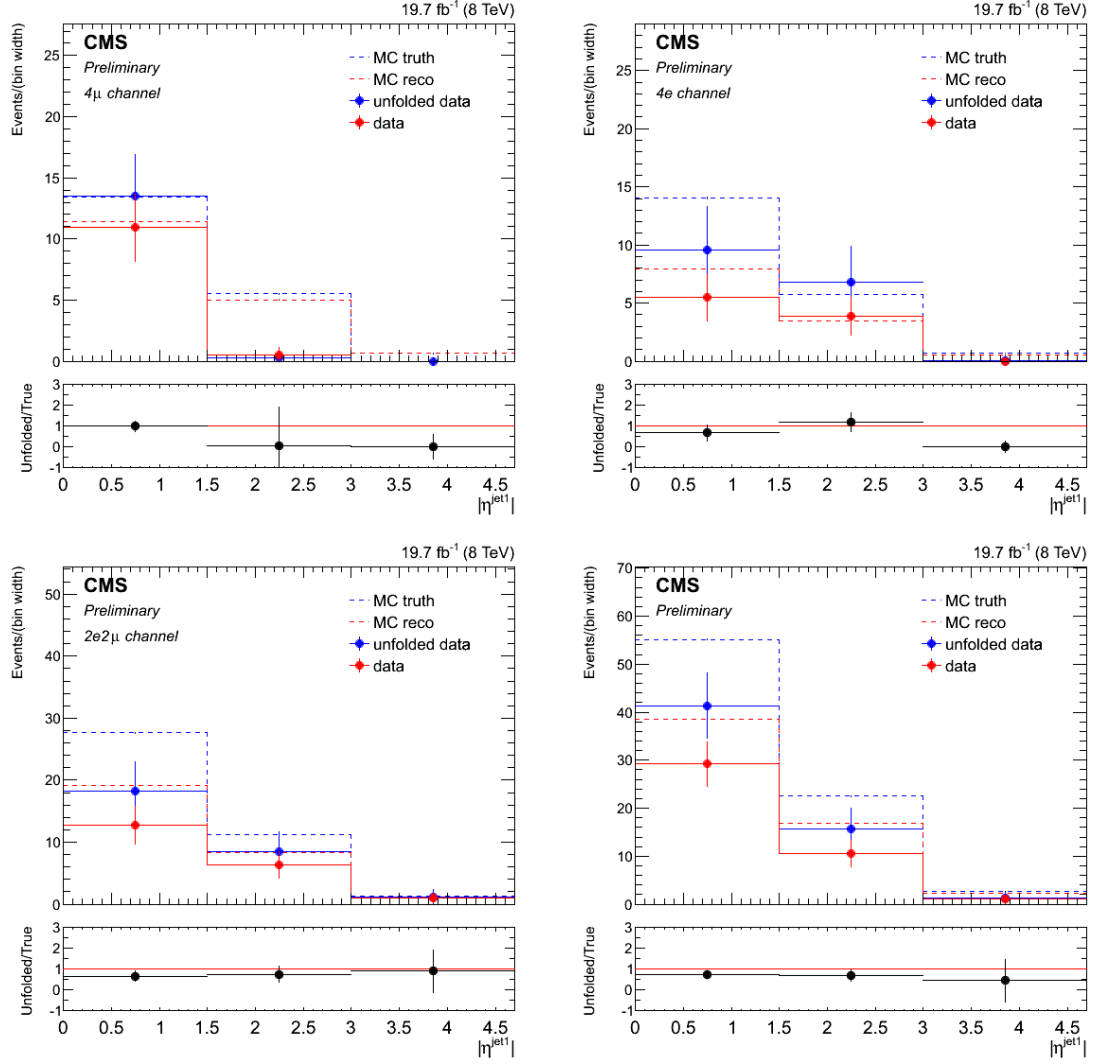


Figure 6.31: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) η distributions of the p_T -leading jet, according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

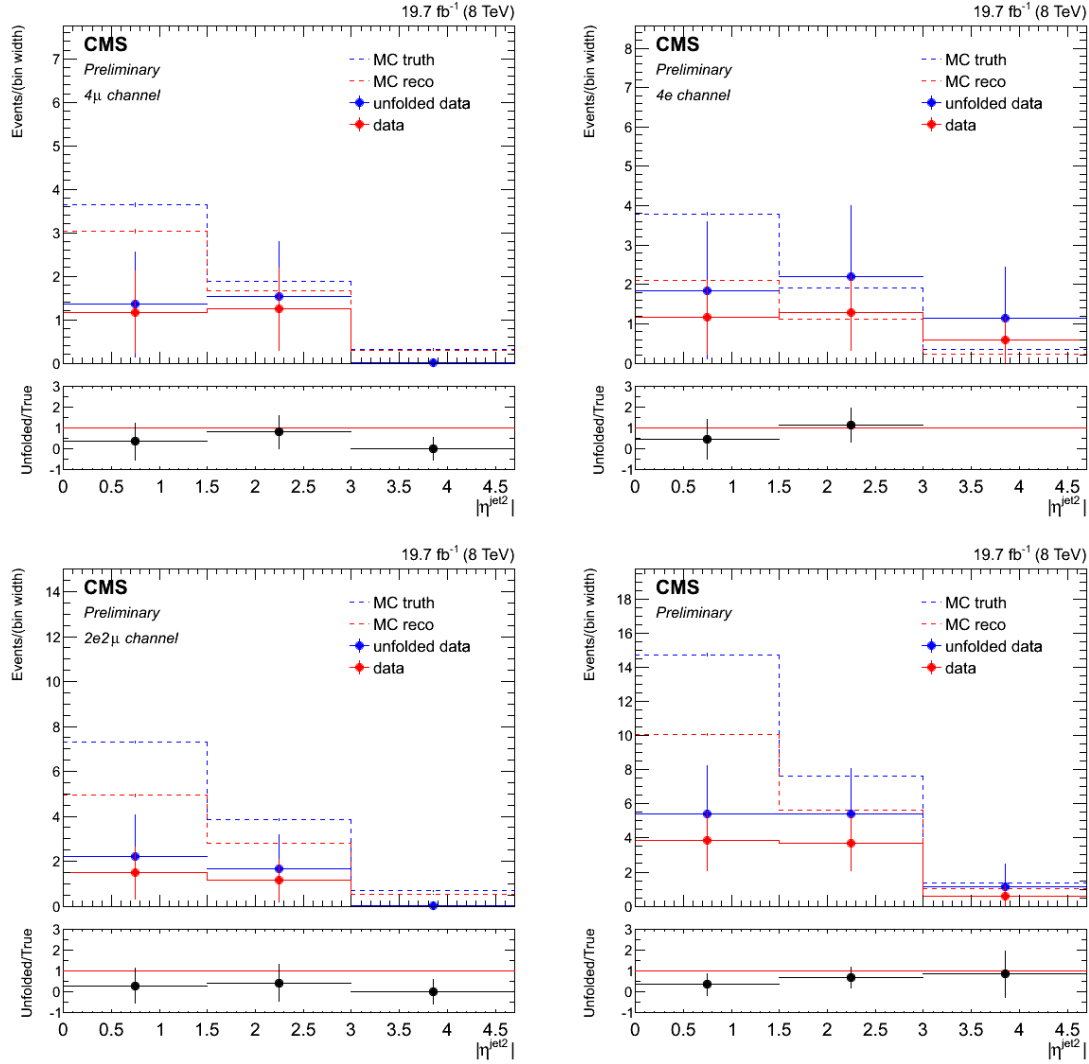


Figure 6.32: Data (red), unfolded data (blue), generated (dashed blue) and MC reconstructed (dashed red) η distributions of the second p_T -leading jet, according to the final state: (top left) 4μ , (top right) $4e$, (bottom) $2e2\mu$, (bottom right) 4ℓ . The unfolded distributions are obtained in the tight fiducial region using the D'Agostini algorithm, with 2 iterations, and compared to predictions from the MadGraph set of samples.

6.3 Differential Cross-Section Measurement

The cross-section in each bin of each observable is determined from the event yields subtracting the backgrounds. The same selection as in the inclusive cross-section measurement is applied to the reconstructed events. Each distribution is then corrected for event selection efficiencies and for detector resolution effects through an unfolding technique, in order to be compared with predictions from event generators. As reported in Section 5.4, two fiducial regions are taken into account:

- the *wider* fiducial region, in which only the requirement of two Z bosons with $60 < m_Z < 120$ GeV is demanded;
- the *tighter* fiducial region, in which conditions on lepton p_T and η are also required (at least one lepton with $p_T > 20$ GeV and another one with $p_T > 10$ GeV, $p_T^e > 7$ GeV and $|\eta^e| < 2.5$, $p_T^\mu > 5$ GeV and $|\eta^\mu| < 2.4$).

Two different sets of response matrices are built, according to the fiducial region, and used to correct reconstructed distributions to the *particle level*. The differential cross-section is then obtained by dividing the result of each bin to the luminosity, the proper branching ratio (if the measurement is performed in the wider fiducial region) and the bin width.

The distribution of the number of jets measured in both wider and tighter fiducial region is shown in Figure 6.33 and the corresponding cross-section values are reported in Tables 6.1, 6.2, 6.3 and 6.4, for all the different decay channels and for their combination. In the wider region, the 4ℓ combination is assessed by the weighted average of the $4e$, 4μ and $2e2\mu$ final states, computed bin-per-bin. In this case the measurement performed in each decay channel is the same and it is divided by the branching ratio. On the other hand, in the tighter region the computed 4ℓ combination is determined as the sum of the 4μ , $4e$ and $2e2\mu$ distributions.

In order to show the shape discrepancy between data distributions and MC predictions, measured and expected differential cross-sections are normalized to one, by dividing the result of each bin to their own integral. Since this is done in the tighter fiducial region, such integral corresponds to the inclusive fiducial cross-section in the experimentally visible phase-space defined by the kinematic and geometrical acceptance of leptons. The differential distributions normalized to the unity and extracted in the tight fiducial region are presented in Figures 6.34 - 6.44 for the 4μ , $4e$ and $2e2\mu$ decay channels and for their combination. The 4ℓ final state is obtained as the sum of the 4μ , $4e$ and $2e2\mu$ distributions, and is then normalized.

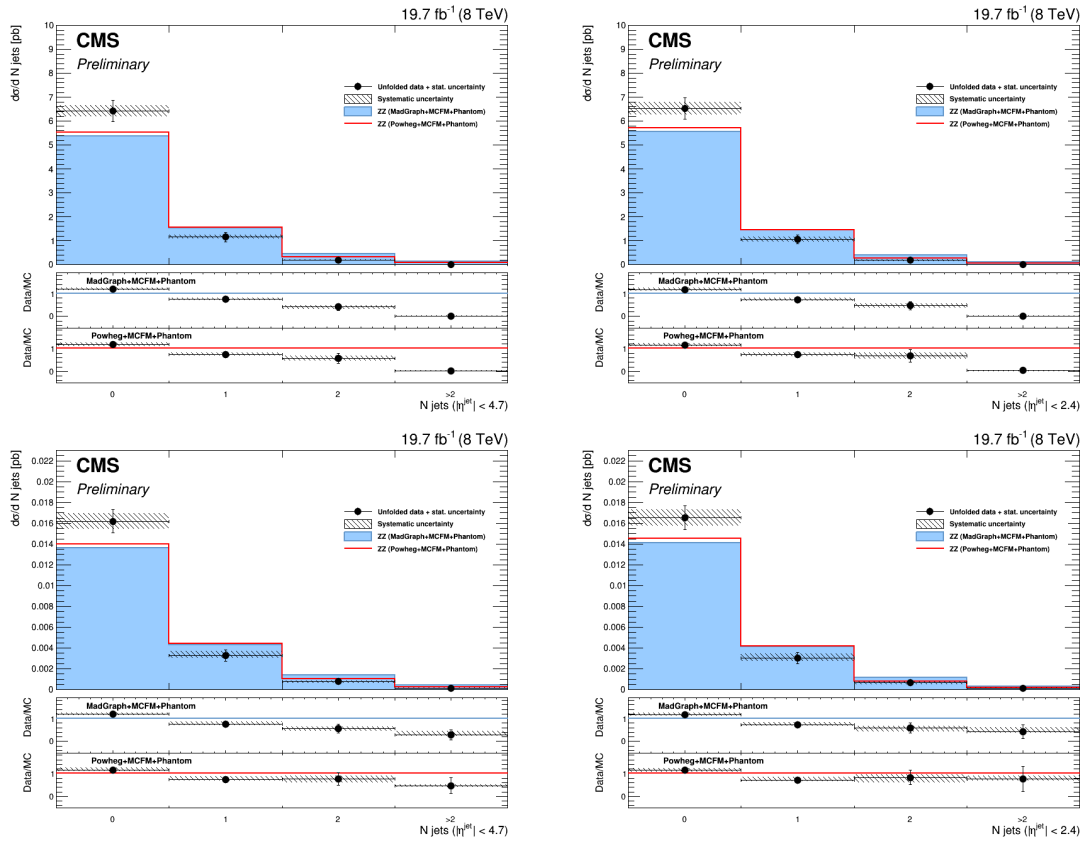


Figure 6.33: Differential cross-sections as a function of the number of jets (left) and central jets (right) in the 4ℓ final state. Cross-sections are extracted in both wider (top) and tighter (bottom) fiducial regions and are compared to predictions from the MadGraph and Powheg sets of samples.

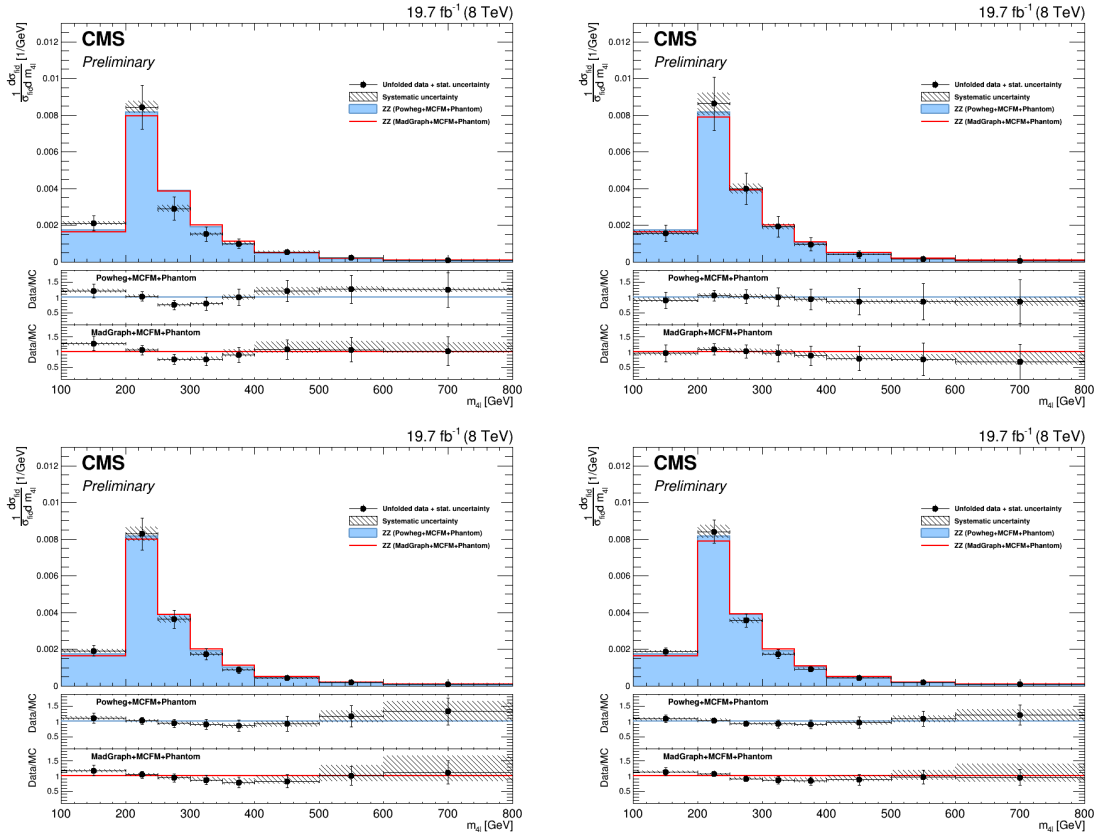


Figure 6.34: Normalized differential cross-sections as a function of the invariant mass of the four-lepton system, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the Powheg and MadGraph sets of samples.

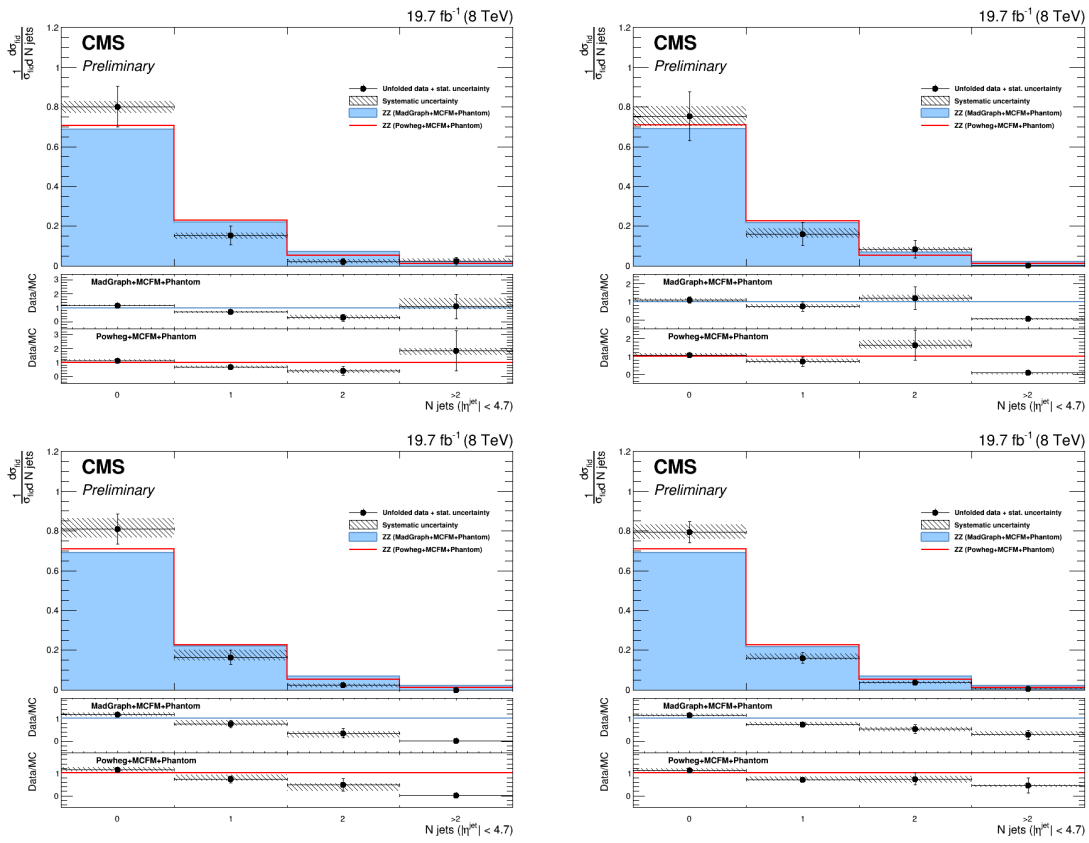


Figure 6.35: Normalized differential cross-sections as a function of the number of jets in the event (with $|\eta^{\text{jet}}| < 4.7$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2e2e$ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

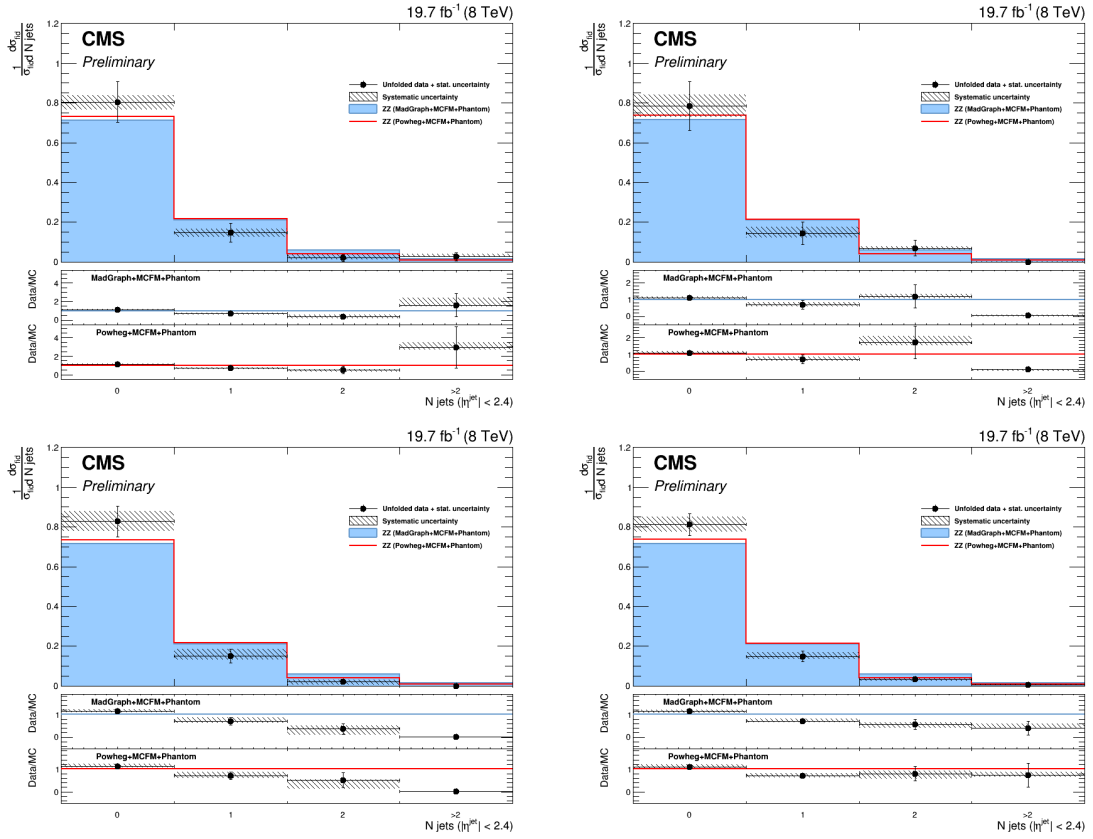


Figure 6.36: Normalized differential cross-sections as a function of the number of central jets in the event (with $|\eta^{\text{jet}}| < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

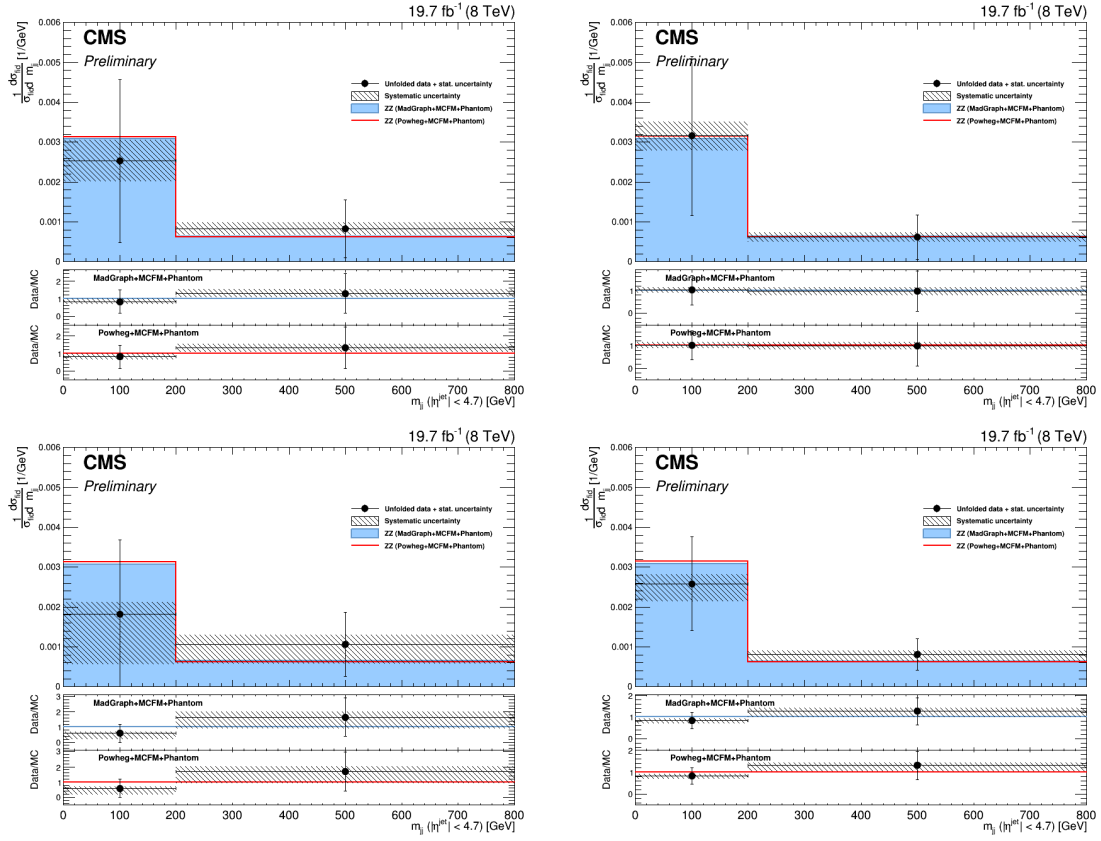


Figure 6.37: Normalized differential cross-sections as a function of the invariant mass of the p_T -leading jets in the event (with $|\eta^{jet}| < 4.7$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial prelin and compared to predictions from the MadGraph and Powheg sets of samples.

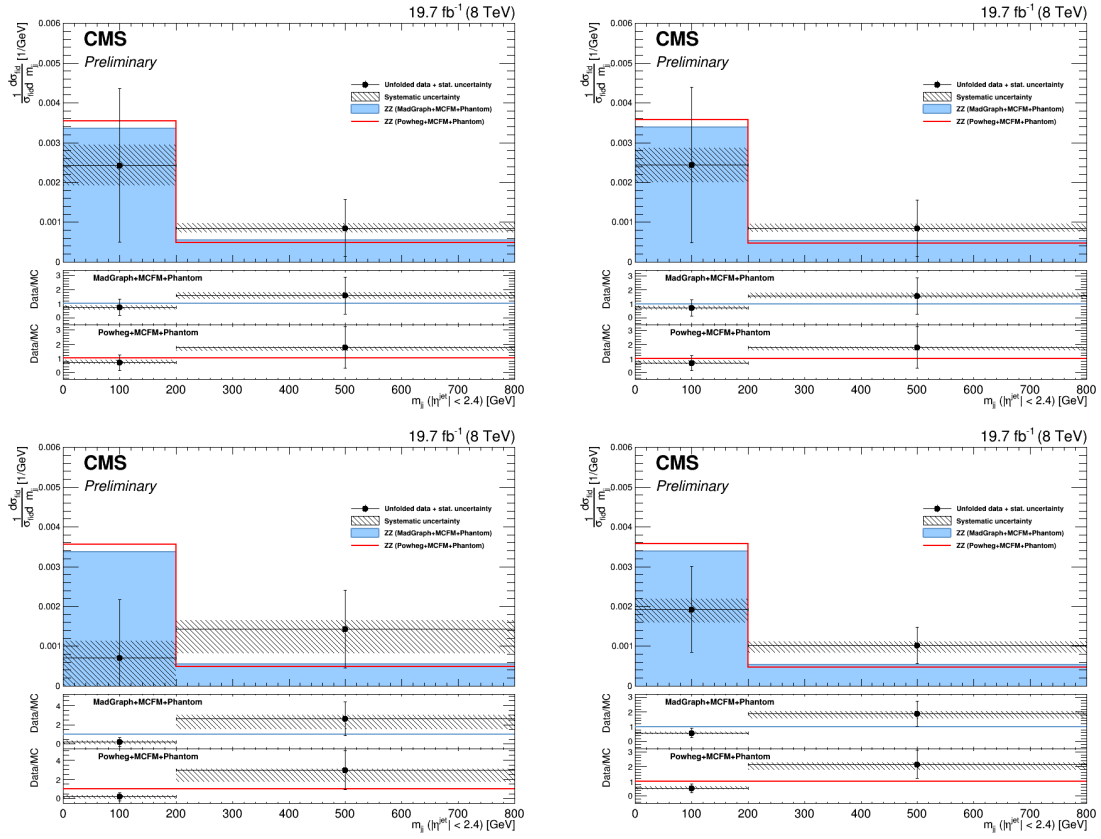


Figure 6.38: Normalized differential cross-sections as a function of the invariant mass of the p_T -leading central jets in the event (with $|\eta^{jet}| < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

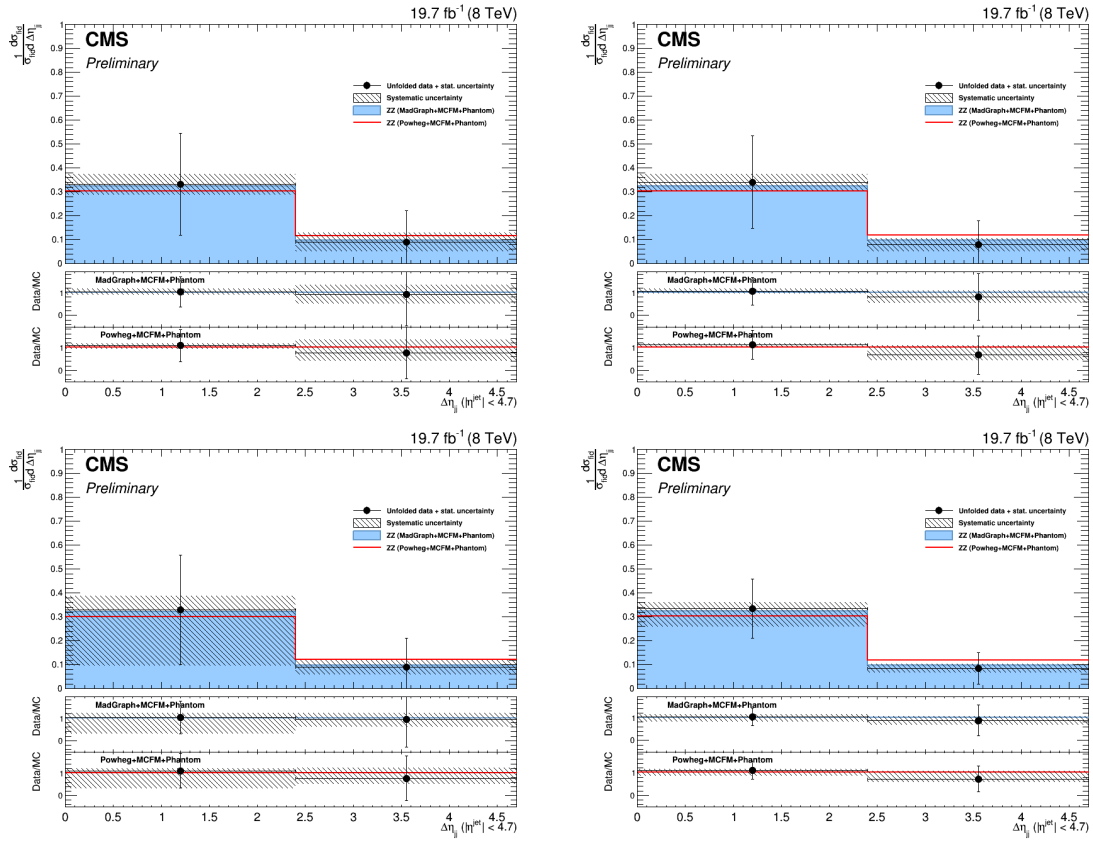


Figure 6.39: Normalized differential cross-sections as a function of the pseudorapidity interval between the p_T -leading jets in the event (with $|\eta^{jet}| < 4.7$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

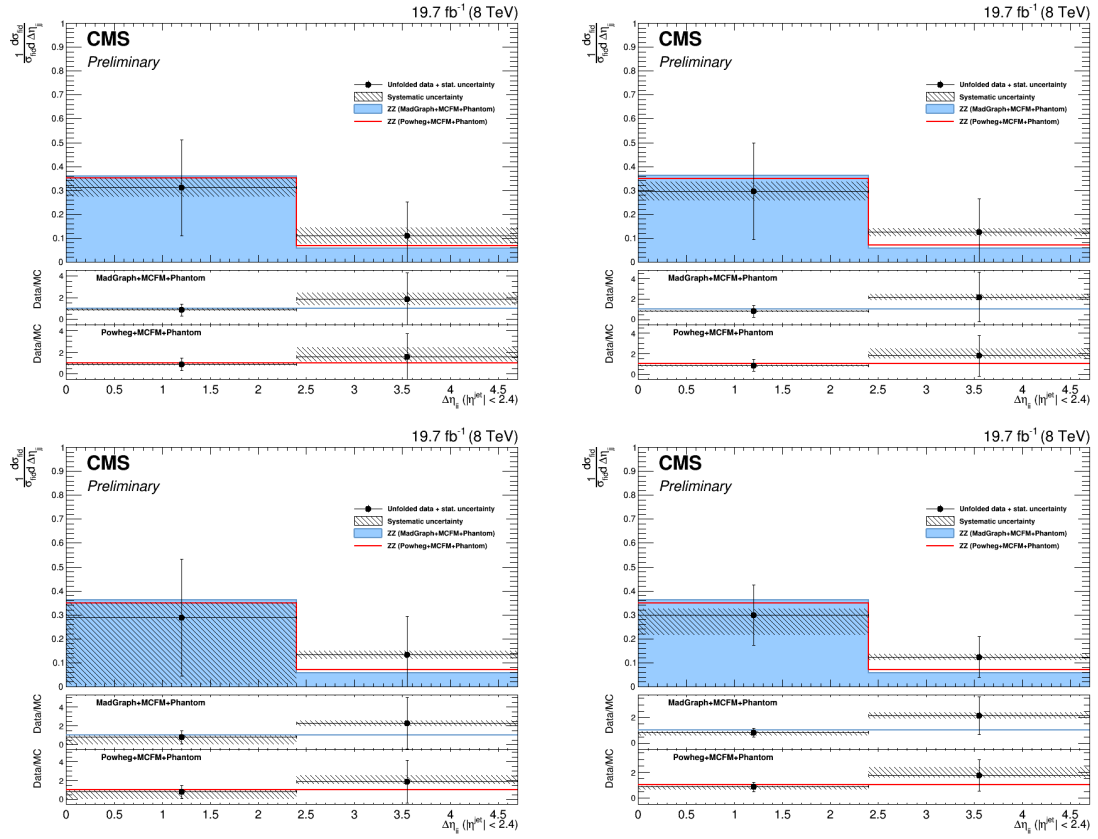


Figure 6.40: Normalized differential cross-sections as a function of the pseudorapidity interval between the p_T -leading central jets in the event (with $|\eta^{\text{jet}}| < 2.4$), according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

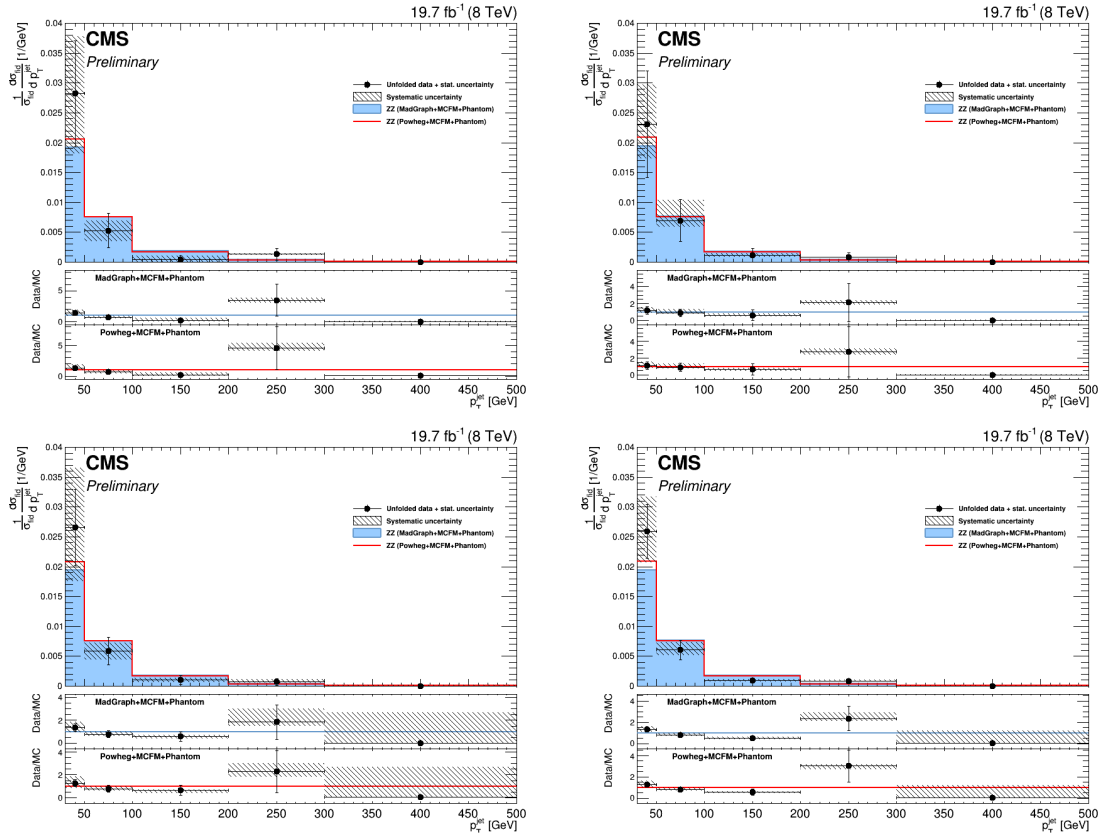


Figure 6.41: Normalized differential cross-sections as a function of the transverse momentum of the p_T -leading jet in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

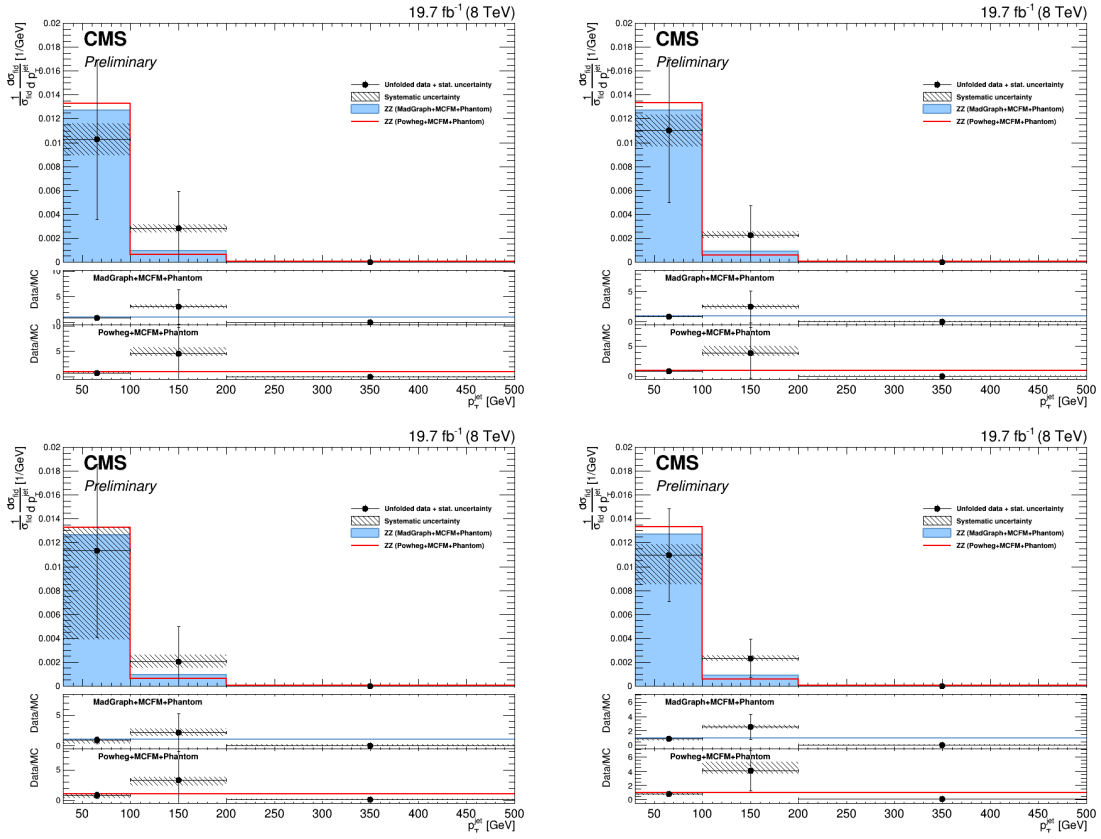


Figure 6.42: Normalized differential cross-sections as a function of the transverse momentum of the second p_T -leading jet in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

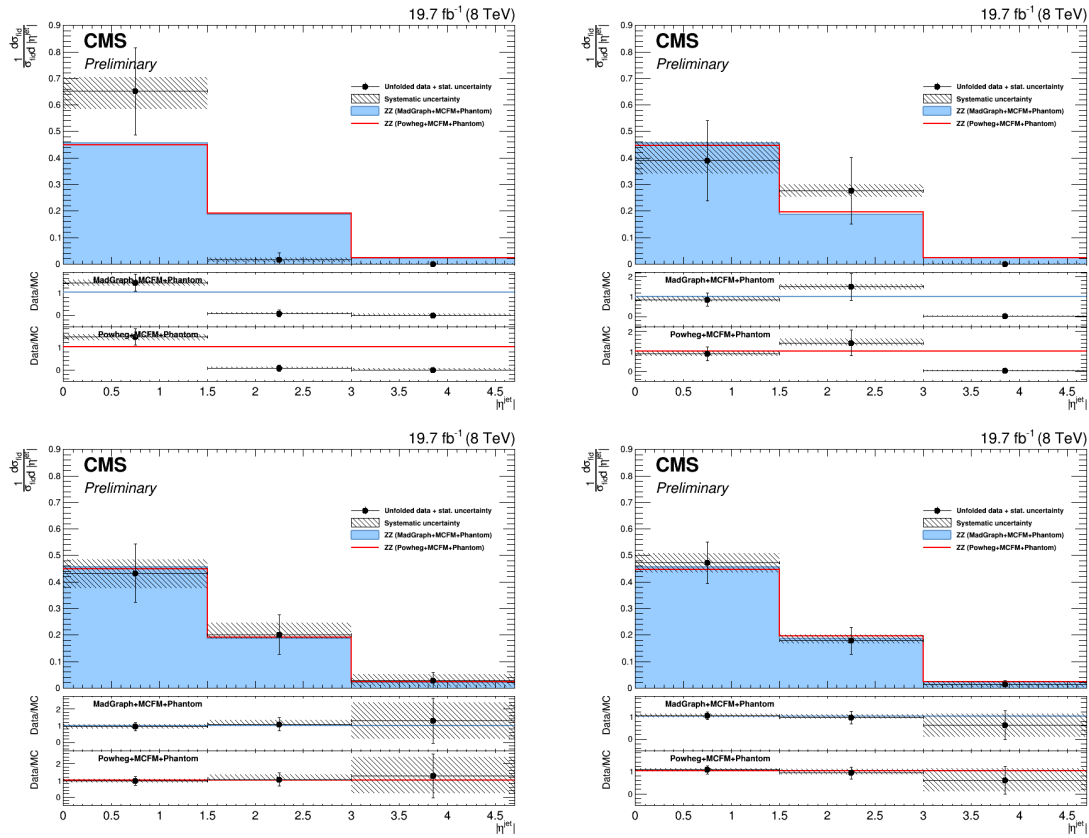


Figure 6.43: Normalized differential cross-sections as a function of the pseudorapidity of the p_T -leading jet in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

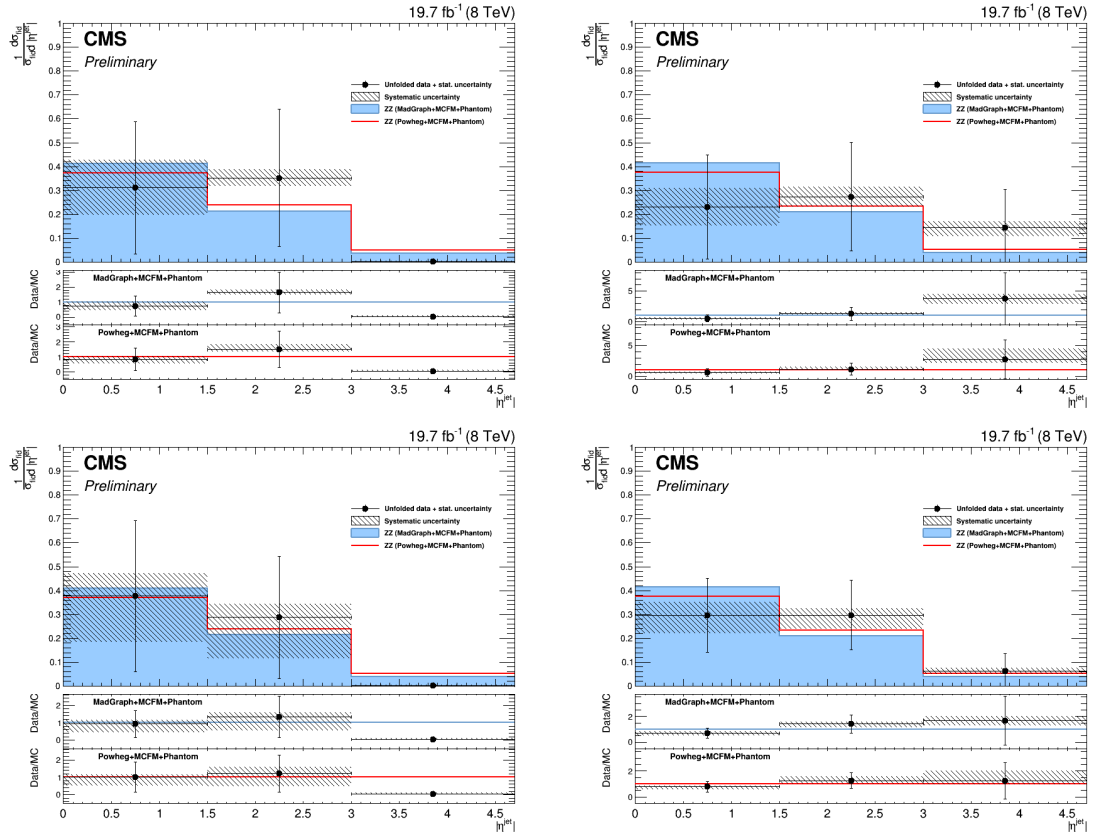


Figure 6.44: Normalized differential cross-sections as a function of the pseudorapidity of the second p_T -leading jet in the event, according to the final state: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), 4ℓ (bottom right). Cross-sections are extracted in the tight fiducial region and compared to predictions from the MadGraph and Powheg sets of samples.

Table 6.1: 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	6.20 ± 0.79 (stat.) $^{+0.23}_{-0.25}$ (syst.)
$pp \rightarrow ZZ(4e)$	0	5.78 ± 0.94 (stat.) $^{+0.39}_{-0.39}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	0	6.98 ± 0.65 (stat.) $^{+0.46}_{-0.40}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	0	6.43 ± 0.44 (stat.) $^{+0.23}_{-0.25}$ (syst.)
$pp \rightarrow ZZ(4\mu)$	1	1.06 ± 0.33 (stat.) $^{+0.09}_{-0.13}$ (syst.)
$pp \rightarrow ZZ(4e)$	1	1.11 ± 0.41 (stat.) $^{+0.18}_{-0.12}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	1	1.27 ± 0.28 (stat.) $^{+0.29}_{-0.12}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	1	1.15 ± 0.19 (stat.) $^{+0.08}_{-0.08}$ (syst.)
$pp \rightarrow ZZ(4\mu)$	2	0.13 ± 0.10 (stat.) $^{+0.09}_{-0.05}$ (syst.)
$pp \rightarrow ZZ(4e)$	2	0.55 ± 0.28 (stat.) $^{+0.07}_{-0.08}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	2	0.19 ± 0.10 (stat.) $^{+0.04}_{-0.10}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	2	0.19 ± 0.07 (stat.) $^{+0.03}_{-0.04}$ (syst.)
$pp \rightarrow ZZ(4\mu)$	> 2	0.16 ± 0.12 (stat.) $^{+0.09}_{-0.03}$ (syst.)
$pp \rightarrow ZZ(4e)$	> 2	0.0072 ± 0.0040 (stat.) $^{+0.0073}_{-0.0075}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	> 2	0.0023 ± 0.0015 (stat.) $^{+0.0023}_{-0.0028}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	> 2	0.0029 ± 0.0014 (stat.) $^{+0.0022}_{-0.0026}$ (syst.)

Table 6.2: 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	6.22 ± 0.80 (stat.) $^{+0.28}_{-0.30}$ (syst.)
$pp \rightarrow ZZ(4e)$	0	6.00 ± 0.95 (stat.) $^{+0.43}_{-0.43}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	0	7.11 ± 0.66 (stat.) $^{+0.48}_{-0.44}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	0	6.53 ± 0.45 (stat.) $^{+0.27}_{-0.27}$ (syst.)
$pp \rightarrow ZZ(4\mu)$	1	1.01 ± 0.32 (stat.) $^{+0.14}_{-0.16}$ (syst.)
$pp \rightarrow ZZ(4e)$	1	0.99 ± 0.39 (stat.) $^{+0.21}_{-0.16}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	1	1.15 ± 0.27 (stat.) $^{+0.26}_{-0.15}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	1	1.06 ± 0.18 (stat.) $^{+0.11}_{-0.10}$ (syst.)
$pp \rightarrow ZZ(4\mu)$	2	0.14 ± 0.11 (stat.) $^{+0.10}_{-0.06}$ (syst.)
$pp \rightarrow ZZ(4e)$	2	0.46 ± 0.26 (stat.) $^{+0.05}_{-0.05}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	2	0.16 ± 0.10 (stat.) $^{+0.06}_{-0.12}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	2	0.18 ± 0.07 (stat.) $^{+0.04}_{-0.04}$ (syst.)
$pp \rightarrow ZZ(4\mu)$	> 2	0.17 ± 0.13 (stat.) $^{+0.08}_{-0.01}$ (syst.)
$pp \rightarrow ZZ(4e)$	> 2	0.0045 ± 0.0028 (stat.) $^{+0.0061}_{-0.0047}$ (syst.)
$pp \rightarrow ZZ(2e2\mu)$	> 2	0.0017 ± 0.0012 (stat.) $^{+0.0017}_{-0.0021}$ (syst.)
$pp \rightarrow ZZ(4\ell)$	> 2	0.0021 ± 0.0011 (stat.) $^{+0.0016}_{-0.0019}$ (syst.)

Table 6.3: The $ZZ \rightarrow 4\ell$ production cross-section as measured in each decay channel and for the sum of all channels in the tight fiducial region, classified with respect to the number of jets.

Process	Number of jets ($ \eta^{jet} < 4.7$)	Total cross-section [fb]
$ZZ \rightarrow 4\mu$	0	3.74 ± 0.48 (stat.) $^{+0.14}_{-0.15}$ (syst.)
$ZZ \rightarrow 4e$	0	3.69 ± 0.60 (stat.) $^{+0.25}_{-0.24}$ (syst.)
$ZZ \rightarrow 2e2\mu$	0	8.74 ± 0.82 (stat.) $^{+0.56}_{-0.49}$ (syst.)
$ZZ \rightarrow 4\ell$	0	16.18 ± 1.12 (stat.) $^{+0.77}_{-0.72}$ (syst.)
$ZZ \rightarrow 4\mu$	1	0.72 ± 0.23 (stat.) $^{+0.06}_{-0.08}$ (syst.)
$ZZ \rightarrow 4e$	1	0.79 ± 0.29 (stat.) $^{+0.13}_{-0.09}$ (syst.)
$ZZ \rightarrow 2e2\mu$	1	1.78 ± 0.39 (stat.) $^{+0.40}_{-0.17}$ (syst.)
$ZZ \rightarrow 4\ell$	1	3.29 ± 0.54 (stat.) $^{+0.44}_{-0.23}$ (syst.)
$ZZ \rightarrow 4\mu$	2	0.09 ± 0.08 (stat.) $^{+0.06}_{-0.04}$ (syst.)
$ZZ \rightarrow 4e$	2	0.41 ± 0.21 (stat.) $^{+0.05}_{-0.05}$ (syst.)
$ZZ \rightarrow 2e2\mu$	2	0.27 ± 0.15 (stat.) $^{+0.05}_{-0.15}$ (syst.)
$ZZ \rightarrow 4\ell$	2	0.78 ± 0.27 (stat.) $^{+0.10}_{-0.17}$ (syst.)
$ZZ \rightarrow 4\mu$	> 2	0.12 ± 0.09 (stat.) $^{+0.06}_{-0.02}$ (syst.)
$ZZ \rightarrow 4e$	> 2	0.0054 ± 0.0030 (stat.) $^{+0.0055}_{-0.0057}$ (syst.)
$ZZ \rightarrow 2e2\mu$	> 2	0.0034 ± 0.0023 (stat.) $^{+0.0035}_{-0.0042}$ (syst.)
$ZZ \rightarrow 4\ell$	> 2	0.13 ± 0.09 (stat.) $^{+0.06}_{-0.02}$ (syst.)

Table 6.4: The $ZZ \rightarrow 4\ell$ production cross-section as measured in each decay channel and for the sum of all channels in the tight fiducial region, classified with respect to the number of central jets.

Process	Number of jets ($ \eta^{jet} < 2.4$)	Total cross-section [fb]
$ZZ \rightarrow 4\mu$	0	3.76 ± 0.48 (stat.) $^{+0.17}_{-0.18}$ (syst.)
$ZZ \rightarrow 4e$	0	3.84 ± 0.61 (stat.) $^{+0.28}_{-0.27}$ (syst.)
$ZZ \rightarrow 2e2\mu$	0	8.93 ± 0.82 (stat.) $^{+0.57}_{-0.52}$ (syst.)
$ZZ \rightarrow 4\ell$	0	16.53 ± 1.13 (stat.) $^{+0.80}_{-0.76}$ (syst.)
$ZZ \rightarrow 4\mu$	1	0.69 ± 0.22 (stat.) $^{+0.09}_{-0.11}$ (syst.)
$ZZ \rightarrow 4e$	1	0.71 ± 0.28 (stat.) $^{+0.15}_{-0.11}$ (syst.)
$ZZ \rightarrow 2e2\mu$	1	1.63 ± 0.38 (stat.) $^{+0.37}_{-0.21}$ (syst.)
$ZZ \rightarrow 4\ell$	1	3.04 ± 0.52 (stat.) $^{+0.42}_{-0.28}$ (syst.)
$ZZ \rightarrow 4\mu$	2	0.10 ± 0.08 (stat.) $^{+0.07}_{-0.04}$ (syst.)
$ZZ \rightarrow 4e$	2	0.34 ± 0.20 (stat.) $^{+0.04}_{-0.04}$ (syst.)
$ZZ \rightarrow 2e2\mu$	2	0.23 ± 0.15 (stat.) $^{+0.09}_{-0.17}$ (syst.)
$ZZ \rightarrow 4\ell$	2	0.67 ± 0.26 (stat.) $^{+0.12}_{-0.18}$ (syst.)
$ZZ \rightarrow 4\mu$	> 2	0.12 ± 0.09 (stat.) $^{+0.06}_{-0.01}$ (syst.)
$ZZ \rightarrow 4e$	> 2	0.0033 ± 0.0020 (stat.) $^{+0.0046}_{-0.0036}$ (syst.)
$ZZ \rightarrow 2e2\mu$	> 2	0.0025 ± 0.0018 (stat.) $^{+0.0025}_{-0.0032}$ (syst.)
$ZZ \rightarrow 4\ell$	> 2	0.13 ± 0.09 (stat.) $^{+0.06}_{-0.01}$ (syst.)

6.4 Systematic Uncertainties

To evaluate the effects of the different sources of systematic uncertainty on the unfolded distributions, the whole unfolding procedure is repeated for every contribution. The following systematic uncertainties are considered for the final result and their contributions are reported in Appendix B.

Unfolding

Uncertainties due to the unfolding procedure are computed by the RooUnfold toolkit and consist of two types, the *measurement uncertainties* and the *response matrix uncertainties*. The former come from the propagation of the statistical errors of the distribution that is unfolded through the unfolding matrix, in particular bin-to-bin correlations, and must be considered as statistical uncertainties. The latter are uncertainties on the response matrix elements due to the limited MC statistics and can be neglected.

σ_{qq} and σ_{gg} Ratio

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

Choice of the MC Generator

This uncertainty is computed comparing unfolded data distributions obtained applying the Madgraph + MCFM + Phantom or the Powheg + MCFM + Phantom response matrix and taking the difference between the two results. It varies between 1% and 10%.

Irreducible and Reducible Backgrounds

Uncertainties due to background estimate are evaluated by creating a new *data - background* distribution, in which the background is shifted up and down by its uncertainties, that is unfolded using standard matrices. This procedure is done separately for both reducible (1 – 16%) and irreducible (0.1 – 5.5%) contributions.

Unfolded Data over MC Truth Ratio

In order to take into account the discrepancy between data and MC at *particle level*, a systematic uncertainty is evaluated building up a new response matrix whose elements are weighted for the ratio between the unfolded data and the generator level information. The difference between the unfolded results obtained by the application of this weighted response matrix and the standard one is considered as a systematic error. This uncertainty

ranges from 1% to 29%, in the last bins where the statistics is very low, and it is thus one of the largest contributions.

Jet Energy Resolutions (JER)

As explained in Subsection 4.4.3, the simulated jet p_T needs to be smeared in order to describe data, which are characterized by a worse resolution. The smearing depends on a scaling factor c , whose values are reported in Table 4.1. The uncertainty on this measured factor has to be taken into account by scaling the simulated jets with two additional sets of values that correspond to varying the c -factor up and down by one standard deviation (Table 6.5). The effect is propagated accordingly by smearing the p_T of jets in the samples used to build the response matrices. This uncertainty is less than 3.5%.

Table 6.5: Scaling factors applied on MC jets to evaluate the systematic uncertainty on jet energy resolution corrections.

$ \eta $ -region	c_{down}	c_{up}
0.0 – 0.5	1.053	1.105
0.5 – 1.1	1.071	1.127
1.1 – 1.7	1.092	1.150
1.7 – 2.3	1.162	1.254
2.3 – 2.8	1.192	1.316
2.8 – 3.2	1.332	1.458
3.2 – 5.0	0.865	1.247

Jet Energy Scale (JES)

In Subsection 4.4.2 the different types of contributions needed to correct the measured jet energy to the true particle one are described. These corrections introduce systematic sources of uncertainty that are documented in Reference [84]. The corresponding uncertainty on the jet energy is transverse momentum and pseudorapidity dependent. It is computed 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 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.

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 of each lepton 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 contribution is

found to ranges between 1% and 6%.

As an example, in Figure 6.45 the individual contributions of all the systematic uncertainties described above is reported for the number-of-jets distribution.

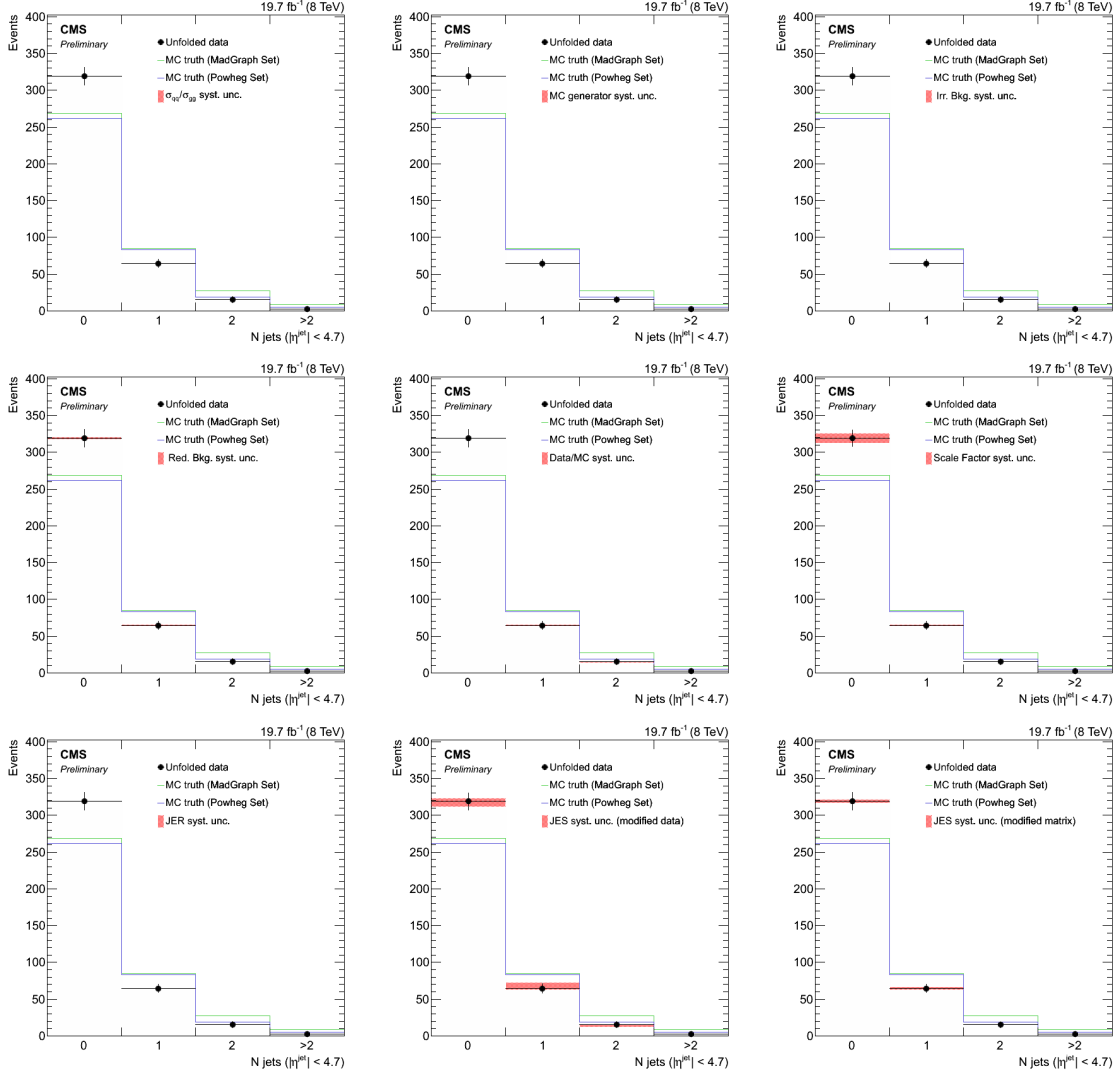


Figure 6.45: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of jets (with $|\eta^{jet}| < 4.7$), for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, 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. The tighter fiducial region is considered.

Besides the several sources of systematic uncertainties propagated through the unfolding procedure, additional contributions affect the overall measurement. The uncertainty on the LHC integrated luminosity is 2.6%. Uncertainties related to the choice of the parton distribution function and the strong coupling constant are evaluated to be 1.2% in the wider region and 0.8% in the tighter region. Finally, the uncertainty for the trigger efficiency is found to be 1.5%.

6.5 Interpretation of the Results

A discrepancy between data and MC predictions can be noticed already at reconstruction level and propagates through the unfolding procedure, affecting the final cross-section result. By analyzing the number-of-jets cross-section distribution (Figure 6.33), it can be observed that the number of events in the first bin with zero jets is significantly larger in data with respect to the predictions, while for the 1, 2 and > 2 bins, on the contrary, the reverse is true. This translates into a falling spectrum of the data/MC truth ratio for both sets of samples (MadGraph and Powheg), as reported below each plot. Thus, more events with at least one jet are expected than what we are actually measuring. This can be related to the distribution of the leading-jet transverse momentum (Figure 6.41). The jet- p_T presents a softer behavior in data with respect to predictions: jets measured from data tend to have a lower p_T -value and therefore, on average, they are less likely to pass the 30 GeV threshold, increasing the number of events with no jets.

The discrepancy between observed data and theoretic expectations is clear especially in distributions where more than one jet is required, as it can be seen in Figures 6.25 - 6.28, where the invariant mass and the pseudorapidity separation of the two p_T -leading jets are studied. For the reasons explained above, both MC sets of samples over-estimate the contribution measured from data and this behavior is preserved also in the unfolded distributions, confirming that no bias is introduced by the procedure. When distributions are normalized, the discrepancy between data and MC predictions almost disappears, underlining that the disagreement is probably a scale factor rather than a difference in describing shapes. However, the statistics of events with more than one jet is very low and does not allow precise measurements to be achieved.

As described in Subsection 5.1.2, the MadGraph Monte Carlo event generator used in this analysis generates the signal $qq/qg \rightarrow ZZ \rightarrow 4\ell$ and $gg \rightarrow ZZ \rightarrow 4\ell$ (tree-level only) processes at LO, separately for 0, 1 and 2 jets, and its total ZZ cross-section is then scaled at NLO using MCFM. On the other hand, Powheg samples, which describe $qq/qg \rightarrow ZZ$ processes, are generated at NLO for 0 jets. Thus, variables related to jet quantities suffer the consequences of these approximations that could be the origin of the observed discrepancy. In order to obtain more accurate theoretical predictions, a possible solution is to use the technique of p_T -resummation [85]. From a theoretical point of view, the jet- p_T threshold enhances renormalization logarithmic factors that are beyond the fixed-order calculation used by the generators employed in this analysis. The p_T of the jets produced in association with the ZZ system is strongly correlated with its transverse momentum, p_T^{ZZ} , because of the event kinematics. Thus, a more precise modeling of the p_T^{ZZ} is necessary. The p_T -resummation technique consists in re-weighting the predicted p_T^{ZZ} spectrum at Next-Next-To-Leading-Log (NNLL) order, by summing up the logarithmic terms and gaining a NLO precision for the additional jets. This technique has been already used in the measurement of WW cross-section at 8 TeV [86] and its implementation will be one of the future studies for the development of the analysis, given the available corrections also for the ZZ boson pair production [87].

Chapter 7

Vector Boson Scattering and Anomalous Quartic Interactions

Although very rare, the production of vector boson pairs includes processes involving the scattering of two massive vector bosons (VBS), $VV \rightarrow VV$ with $V = W$ or Z . As explained in Section 1.5, this class of phenomena is very interesting, being the key processes to probe the nature of the electroweak symmetry breaking (EWSB) mechanism and thus, as consequence, to test the Standard Model validity. For this reason, after the discovery of the Higgs boson, the study of VBS processes at hadron colliders is a very important issue to be addressed, and possibly the highest priority of particle physics today, which will open other interesting puzzles to be solved in the future.

Being characterized by a pair of vector bosons created in association with two high- p_T jets, VBS processes have a very similar final state compared to the $ZZ + \text{jets}$ event production. The work presented in the previous Chapters can thus be exploited in order to perform a preliminary study on VBS in the $ZZ + 2 \text{ jets}$ channel, starting exploring this phenomenon even at 8 TeV.

In the following, the main features of the VBS search at LHC are described, together with a summary on the recent results obtained by both ATLAS and CMS Collaborations using the data set collected during the Run I period. A possible signal region for the $ZZ + 2 \text{ jets}$ channel is then defined and projections for VBS analyses at higher luminosities are reported.

7.1 Signal Signature

At hadron colliders [88] VBS can be idealized as an interaction of gauge bosons radiated from initial-state quarks, yielding a final state with two bosons and two jets ($VV + 2 \text{ jets}$) in a purely electroweak (EW) process [89]. VBS diagrams are not separately gauge invariant and must thus be studied in conjugation with additional Feynman graphs, leading to the same $VV + 2 \text{ jets}$ final state. Two classes of physical processes give rise to $VV + 2 \text{ jets}$ final state [90]. The first class, which includes VBS contributions, involves exclusively weak-interactions at Born level (of order α_{EW}^4 without considering the boson decay) and is referred to as *electroweak production* (see for example Figure 7.9). The second class of processes involves both the strong and the electroweak interactions at Born level (of order $\alpha_S^2 \alpha_{EW}^2$) and is referred to as *strong production* (see for example Figure 7.10).

Processes via *electroweak production* are rarer with respect to strong $VV + 2 \text{ jets}$ events, but

are expected to carry a distinctive hallmark which can be explored experimentally: the two jets produced in association with the boson pairs are expected to have a very high energy. These jets, that originate from the fragmentation of the ongoing quarks in electroweak processes, are referred to as *tagging jets* and are characterized also by a large separation in (pseudo)rapidity ($\Delta\eta_{jj}$), which stems from the small angle scattering of the two initial partons. Because of both the topological configuration and the large transverse momentum of the colliding quarks, also the invariant mass of the two jets is expected to be large. These features yield thus the best separation power with respect to the events produced via strong interactions. Furthermore, the absence of color exchange in the hard scattering process leads to a rapidity gap in the central part of the detector. However the gap topology may be difficult to exploit due to the high level of pile-up interactions. Figure 7.1 summarizes the main features that characterize a generic VBS event.

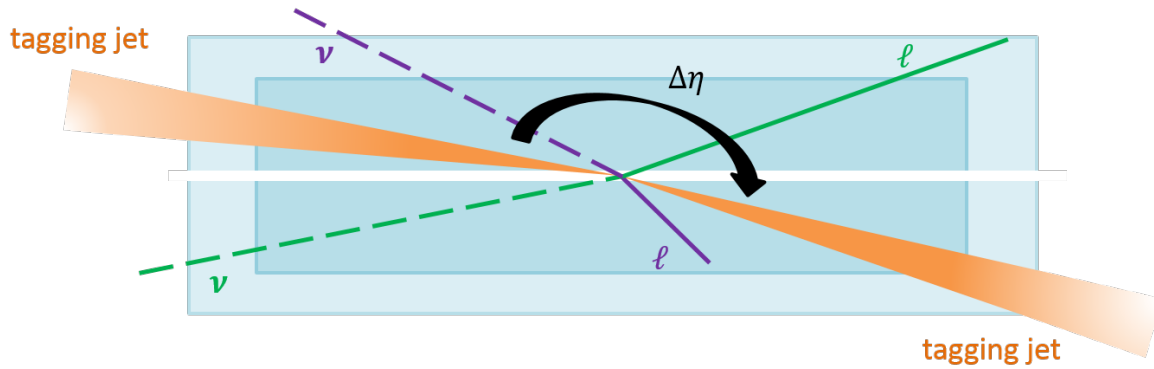


Figure 7.1: Generic vector boson scattering event, characterized by 2 jets of very high energy, well separated in pseudorapidity and with a large invariant mass. The decaying products of the vector boson pair (in this case WW) are expected to be produced centrally, in the rapidity gap between the jets.

7.2 Analyses at Run I

In order to deeply investigate vector boson scattering processes, an integrated luminosity of about 100 fb^{-1} at 14 TeV is needed, as evidenced by several theoretical studies and experimental researches carried out by both ATLAS and CMS Collaborations (Section 7.4). However, using data at 7 and 8 TeV collected during the Run I, it is possible to start exploring these phenomena through the study of different interactions:

- Mono-boson production via vector boson scattering, that allows triple gauge coupling investigation;
- Di-boson production via vector boson scattering, which is sensitive to triple and quartic gauge couplings;
- Di-boson inclusive (electroweak and strong) production, which is sensitive to triple gauge couplings;
- Multi-boson production, which is sensitive to quartic gauge couplings.

In the following, a summary of the published results of both ATLAS and CMS Collaborations is reported, in order to present the current state on this subject.

7.2.1 Z-Production in Association with Forward/Backward Jets

This analysis investigates the electroweak production of a Z boson in association with 2 jets, which is not truly a vector boson scattering process given that only a single boson (and not a pair of bosons) is produced. Nevertheless, it presents very interesting features: it has a comparable cross-section and a very similar topology to the vector boson fusion production of the Higgs boson, it is sensitive to new physics since involves the WWZ triple vertex and, most important in the context of this Chapter, the two jets produced in association with the Z boson have the same characteristics of the VBS tagging jets. This process can thus be exploited to refine forward jet selection and is very useful for vector boson scattering analyses.

In proton proton collisions at LHC the dominant source of production of a Z boson followed by a leptonic decay $Z \rightarrow \ell\ell$ in association with two jets is through mixed electroweak and strong processes of order $\mathcal{O}(\alpha_{EW}^2\alpha_S^2)$, commonly known as *Drell-Yan plus jets* processes. Pure electroweak productions of the $\ell\ell + 2$ jets final state, of order $\mathcal{O}(\alpha_{EW}^4)$, are rarer and include vector boson fusion (VBF) processes, Z-bremsstrahlung processes, and non-resonant processes (Figure 7.2). The VBF process is of particular interest because of the similarity to the VBF production of a Higgs boson and the sensitivity to anomalous WWZ triple gauge couplings.

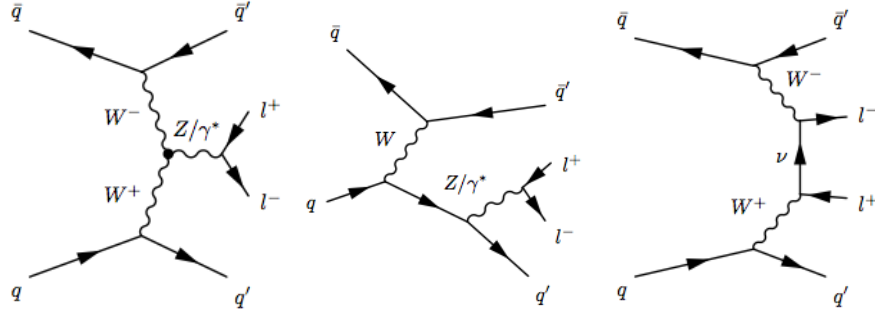


Figure 7.2: Representative diagrams for di-lepton production in association with two jets from pure electroweak processes. Vector boson fusion (left), Bremsstrahlung (center), and Non-resonant (right) productions [91, 92].

This analysis is performed by both ATLAS and CMS Collaborations using the data set collected at 8 TeV, with an integrated luminosity of about 20 fb^{-1} [91, 92]. The ATLAS Experiment presents a measurement of fiducial cross-sections and differential distributions of *inclusive* $Z + 2$ jets production, dominated by processes via strong interaction. These measurements are performed in five fiducial regions with different sensitivity to the electroweak component. The *search region* is chosen to optimize the expected significance when extracting the electroweak $Z + 2$ jets component and is the most selective region. It is defined requiring:

- a Z boson candidate, with exactly two oppositely charged and same flavor leptons with a di-lepton invariant mass of $81 < m_{\ell\ell} < 101 \text{ GeV}$;
- the transverse momentum of the di-lepton pair $p_T^{\ell\ell} > 20 \text{ GeV}$;
- at least two jets that satisfy $p_T^{j_1} > 55 \text{ GeV}$ and $p_T^{j_2} > 45 \text{ GeV}$, where j_1 and j_2 label the highest and second highest transverse momentum jet in the event;
- the invariant mass of the two leading jets $m_{jj} > 250 \text{ GeV}$;

- no additional jets with $p_T > 25$ GeV in the rapidity interval between the two leading jets;
- a low normalized transverse momentum balance between the two leptons and the two highest transverse momentum jets.

The inclusive cross-section containing both the electroweak and the strong production of $Z + 2$ jets events is measured in this region and is found to be

$$\sigma_{fid}^{search} = 1.10 \pm 0.01 (stat.) \pm 0.09 (syst.) \pm 0.03 (lumi.) \text{ pb.}$$

Inclusive $Z + 2$ jets differential distributions are measured as a function of several variables, such as m_{jj} , Δy_{jj} and other variables sensitive to the differences between electroweak and strong production. In Figure 7.3, the distributions of m_{jj} (left) and Δy_{jj} (right) in the search region are reported.

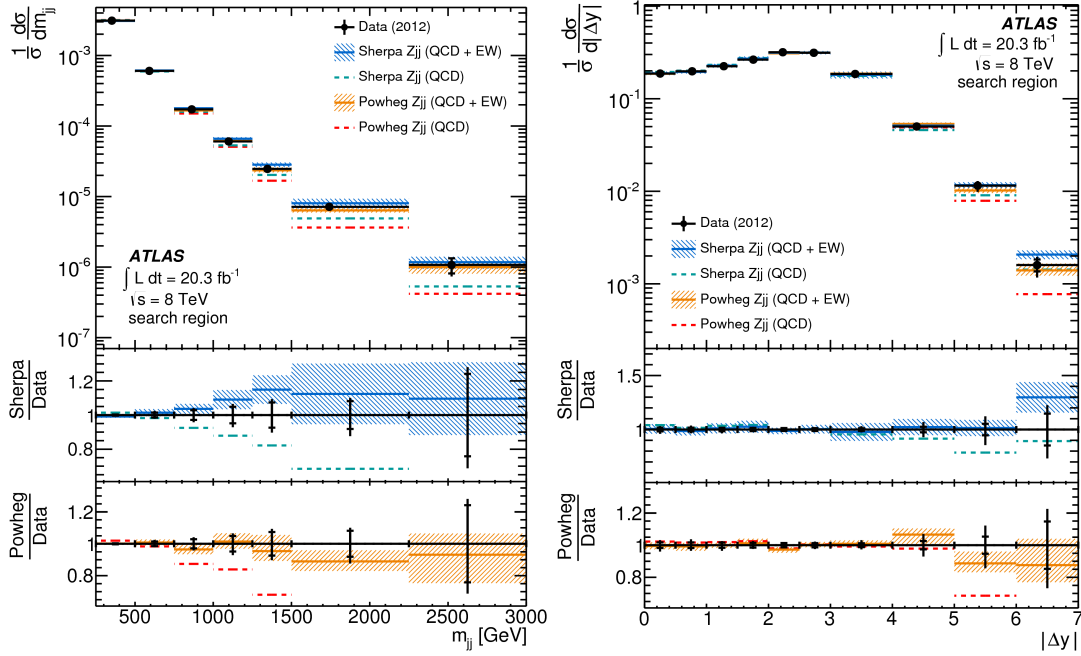


Figure 7.3: Unfolded normalized differential cross-section distribution as a function of the di-jet invariant mass (left) and the rapidity separation between the leading jets (right) in the search region. Distributions are measured by the ATLAS Experiment [91].

In addition to the measurement of inclusive $Z + 2$ jets production, the electroweak component is isolated by fitting the di-jet invariant mass reconstructed in the search region. The number of signal events produced via electroweak interaction is extracted and converted into a fiducial cross-section, obtaining:

$$\sigma_{EW}^{search} = 54.7 \pm 4.6 (stat.)_{-10.4}^{+9.8} (syst.) \pm 1.5 (lumi.) \text{ fb.}$$

The theoretical prediction from Powheg for the electroweak $Z + 2$ jets cross-section is $46.1 \pm 0.2 (stat.)_{-0.2}^{+0.3} (scale) \pm 0.8 (PDF) \pm 0.5 (model) \text{ fb}$, which is in good agreement with the data.

The CMS Collaboration performs a similar analysis. The electroweak production cross-section of a Z boson with two forward/backward jets is measured in di-electron and di-muon final states in the kinematic region $m_{\ell\ell} > 50$ GeV, $m_{jj} > 120$ GeV, transverse mo-

menta $p_T^j > 25$ GeV and pseudorapidity $|\eta_{jj}| < 5$. As the signal-to-background ratio for the cross-section measurement is small, two methods of signal extraction are employed to confirm and cross-check the presence of the signal. A multivariate analysis is performed building a discriminator based on simple variables which reflect mostly the kinematics of the tag jets and of the Z boson (Method I). In addition a new data-driven method is proposed (Method II): a $\gamma + 2$ jets sample is selected in data to estimate the main background due to the QCD production of the $Z + 2$ jets final state. The QCD production of $\gamma + 2$ jets is expected to resemble the production of QCD $Z + 2$ jets, yielding therefore a possible data-based model for the shapes of the kinematics of the tagging jets. In order to obtain a better equivalence with respect to the $\gamma + 2$ jets sample, the selection of the Z boson kinematics is tightened. A linear Fisher discriminator based on the di-jet kinematics is then defined. The distributions of the discriminators used in the two different analyses are reported in Figure 7.4.

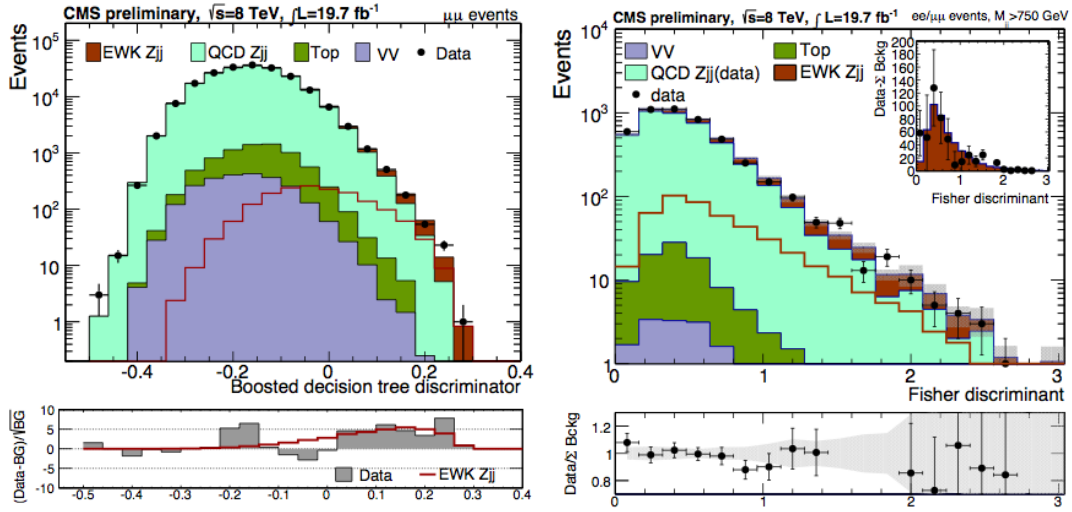


Figure 7.4: BDT shape discriminator used in Method I (left). Fisher shape discriminator used in Method II in the most sensitive region, $m_{jj} > 750$ GeV (right). Distributions are measured by the CMS Experiment [92].

The discriminator shapes are employed to extract the signal strength ($\mu = \sigma/\sigma_{th}$) from a fit to the data. The results obtained with each method are used to compute the cross-section corresponding to the generator-level phase space region and are combined to yield the final measurement:

$$\sigma_{EW} = 226 \pm 26 \text{ (stat.)} \pm 35 \text{ (syst.) fb.}$$

The result is in agreement with the theoretical cross-section at NLO which is 239 fb and the difference with respect to the measurement obtained by the ATLAS Experiment is due to the different kinematic region selected.

The properties of the hadronic activity of the selected events are also studied, defining a high purity electroweak region ($m_{jj} > 1250$ GeV). A significant suppression of the hadronic activity is expected for the signal, since the final state objects originate in pure electroweak interactions in contrast with the production of $Z + 2$ jets via QCD. A good agreement is found between data and different predictions.

7.2.2 W-Production in Association with Forward/Backward Jets

As the analysis described in the previous Subsection, the electroweak production of a W boson in association with two jets is not actually a vector boson scattering process, but it is characterized by the presence of two VBS-like tagging jets and thus can be used to improve jet selection in VBS-related studies. The three classes of pure electroweak diagrams for W boson production in association with two jets are shown in Figure 7.5 and mechanisms are the same as in 7.2.1.

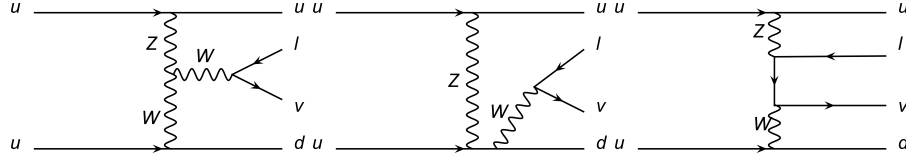


Figure 7.5: Representative diagrams for electroweak $\ell\nu + 2$ jets productions at the LHC: Vector boson fusion (left), Bremsstrahlung (center), and Non-resonant (right) productions [93].

The electroweak production of $W + 2$ jets events is studied by the CMS Collaboration, using 19.3 fb^{-1} of data recorded at $\sqrt{s} = 8 \text{ TeV}$ [93]. The cross-section of this process is measured requiring W bosons to be identified via their decay to muons or electrons and to be produced centrally with two jets well separated in rapidity. A Boosted Decision Tree (BDT) technique is used to distinguish between signal and background events. Given that the QCD $W + 2$ jets background normalization is known, the extraction of the signal yield is performed using an unbinned maximum likelihood fit to the m_{jj} distribution in the data. The normalization of the electroweak $W + 2$ jets contribution is a free parameter of the fit, while the shape parameters are fixed to the MC prediction. In Figure 7.6, the distribution of the m_{jj} variable for the muon and electron decay channels is reported.

The number of signal events is extracted from the fit and used to compute the fiducial electroweak $W + 2$ jets cross-section, calculated for W decaying to muons or electrons and for $p_T^{j1} > 60 \text{ GeV}$, $p_T^{j2} > 50 \text{ GeV}$, $|\Delta\eta_{jj}| < 4.7$ and $m_{jj} > 1000 \text{ GeV}$ at the true level. The combined result of muon and electron channels is:

$$\sigma_{EW}^{e+\mu} = 0.42 \pm 0.04 \text{ (stat.)} \pm 0.09 \text{ (syst.)} \pm 0.01 \text{ (lumi.) pb},$$

in agreement with SM prediction of $0.50 \pm 0.02 \text{ (scale)} \pm 0.02 \text{ (PDF)} \text{ pb}$ obtained by MadGraph generator interfaced with Pythia 6.

7.2.3 $\gamma\gamma$ -Production of W^+W^- pairs

The first $VV \rightarrow VV$ analysis carried out at the LHC is the study of exclusive two-photon production of W^+W^- in proton-proton collisions at $\sqrt{s} = 7$ [94] and 8 TeV [95]. This study is conducted by the CMS Collaboration using data collected in 2011 and 2012, with an integrated luminosity of 5.05 fb^{-1} (7 TeV) and 19.7 fb^{-1} (8 TeV).

Diagrams describing the under-study process are reported in Figure 7.7. In order to select it, the $\mu^\pm e^\mp$ final state is used, following the decay

$$pp \rightarrow p^{(*)} \gamma \gamma p^{(*)} \rightarrow p^{(*)} W^+ W^- p^{(*)} \rightarrow p^{(*)} \mu^\pm e^\mp p^{(*)}.$$

Since both very forward scattered protons escape detection, such a production process is characterized by a primary vertex formed of a $\mu^\pm e^\mp$ pair with no other tracks, with large transverse momentum $p_T(\mu^\pm e^\mp)$ and large invariant mass $m(\mu^\pm e^\mp)$. The choice of

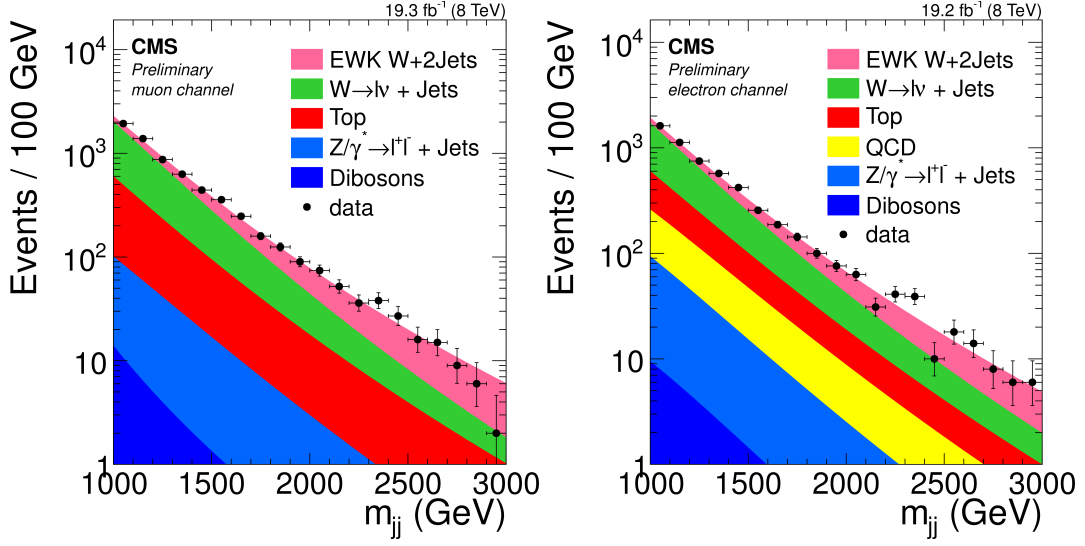


Figure 7.6: Distribution of m_{jj} for muon (left) and electron (right) channel measured with the CMS detector [93].

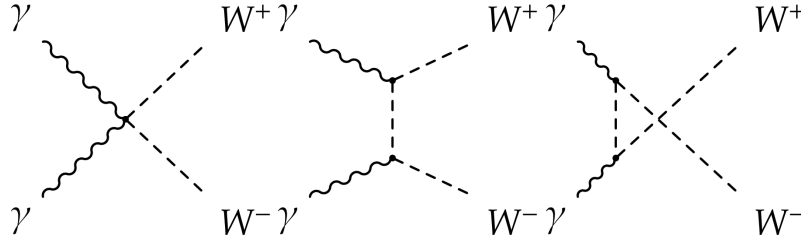


Figure 7.7: Quartic gauge coupling (left) and t - (center) and u -channel (right) W -boson exchange diagrams contributing to the $\gamma\gamma \rightarrow W^+W^-$ process at leading order in the SM [94].

this particular final state is due to the fact that in the case of W^+W^- pair decays to same-flavor $\mu^+\mu^-$ or e^+e^- final states, the backgrounds are more than an order of magnitude larger. The dominant backgrounds in the $\mu^\pm e^\mp$ final state, due to the inclusive production of W^+W^- and $\tau^+\tau^-$ pairs, are constrained using control regions.

Examining the SM signal region, defined by events with zero extra tracks on the $\mu^\pm e^\mp$ vertex and $p_T(\mu^\pm e^\mp) > 30$ GeV, two events passing all selection requirements are observed at 7 TeV, compared to the expectation of 2.2 ± 0.4 signal events and 0.84 ± 0.15 background events (see Figure 7.8). The observed result can be converted into a cross-section and upper limit for events with zero extra tracks within $|\eta| < 2.4$, using the expression $\sigma = N/(\epsilon \times A \times \mathcal{L})$, where N is the number of events observed and $\epsilon \times A$ is the efficiency times acceptance for a SM-like signal. Correcting for efficiency, acceptance and backgrounds, the best fit signal cross-section times branching fraction is:

$$\sigma(pp \rightarrow p^{(*)}W^+W^-p^{(*)} \rightarrow p^{(*)}\mu^\pm e^\mp p^{(*)}) = 2.2_{-2.0}^{+3.3} \text{ fb},$$

with a significance of $\sim 1\sigma$.

The observed upper limit is estimated using the Feldman-Cousins method [96] to be 2.6 times the expected SM yield at 95% C.L., that means a limit on the cross-section at 95% C.L. of

$$\sigma(pp \rightarrow p^*W^+W^-p^* \rightarrow p^{(*)}\mu^\pm e^\mp p^{(*)}) < 10.6 \text{ fb}^{-1}.$$

At 8 TeV, 13 events are observed in the signal region, with an expected background of 3.5 ± 0.5 events, consistently with SM predictions. In this case, the measured cross section is $\sigma = 12.3^{+5.5}_{-4.4}$ (stat.) fb, with a significance of 3.6σ .

Furthermore, this channel can be also sensitive to anomalous quartic gauge couplings of the gauge boson ($\gamma\gamma WW$ vertex) and it can be used to look for any deviation from the SM. In this perspective, in the analysis at 7 TeV a tighter cut on the transverse momentum of the lepton pair ($p_T(\mu^\pm e^\mp) > 100$ GeV) is added. Zero data events are observed in this region, which is consistent with the SM expectation of 0.14. A limit is set on the partial cross-section times branching fraction within the acceptance of $p_T(\mu, e) > 20$ GeV, $|\eta(\mu, e)| < 2.4$, $p_T(\mu^\pm e^\mp) > 100$ GeV at 95% C.L.:

$$\sigma(pp \rightarrow p^* W^+ W^- p^* \rightarrow p^{(*)} \mu^\pm e^\mp p^{(*)}) < 1.9 \text{ fb}^{-1}.$$

At 8 TeV several dimension-six and dimension-eight operators are studied. No significant deviations from the SM are observed and the resulting limits are interpreted in terms of improved constraints on aQGC.

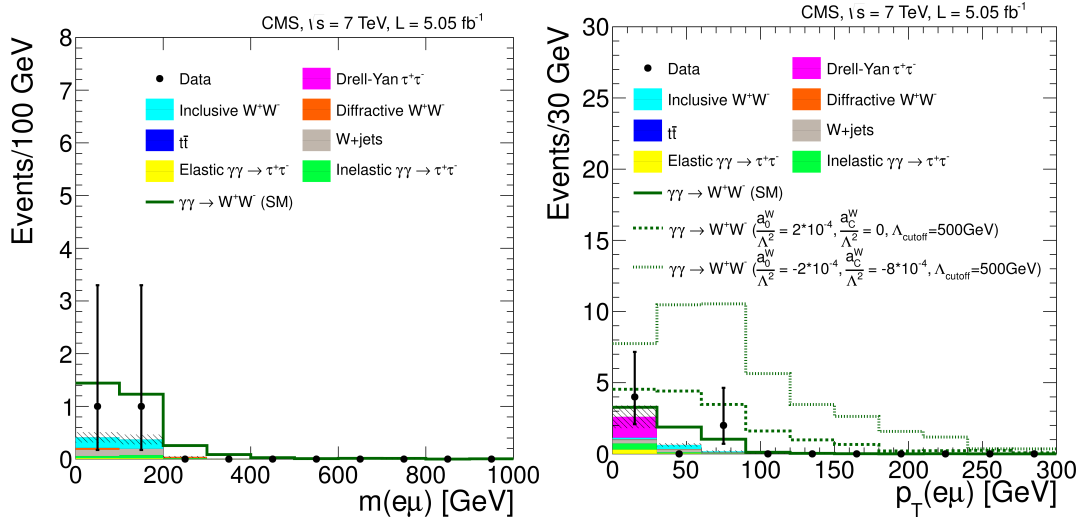


Figure 7.8: The $\mu^\pm e^\mp$ invariant mass (left) and transverse momentum (right) for events in the signal region with zero extra tracks on the $\mu^\pm e^\mp$ vertex at 7 TeV. In the mass distribution the cut $p_T(\mu^\pm e^\mp) > 30$ GeV is added. The backgrounds (solid histograms) are stacked with statistical uncertainties indicated by the shaded region, the signal (open histogram) is stacked on top of the backgrounds. Distributions are measured by the CMS Experiment [94].

7.2.4 Vector Boson Scattering in the $W^\pm W^\pm + 2$ jets Channel

The first evidence of $VV + 2$ jets electroweak production is reported in the study of vector boson scattering in events with two same-sign leptons and two jets. This analysis is carried out by both ATLAS and CMS Experiments [97, 98] and uses the data sample collected at 8 TeV with an integrated luminosity of about 20 fb^{-1} .

The aim of the analysis is to probe the electroweak production of same-sign W boson pair events. The strong production cross-section is reduced by the same sign requirement, making the experimental signature of same sign di-lepton events with two jets an ideal topology for VBS studies. Candidate events have exactly two high- p_T leptons of the same charge, two jets with large rapidity separation and di-jet mass, and moderate missing

transverse energy. The final states considered are $\mu^\pm\mu^\pm\nu_\mu\nu_\mu jj$, $e^\pm e^\pm\nu_e\nu_e jj$ and $\mu^\pm e^\pm\nu_\mu\nu_e jj$. Figures 7.9 and 7.10 show representative Feynman diagrams for the electroweak and QCD induced production.

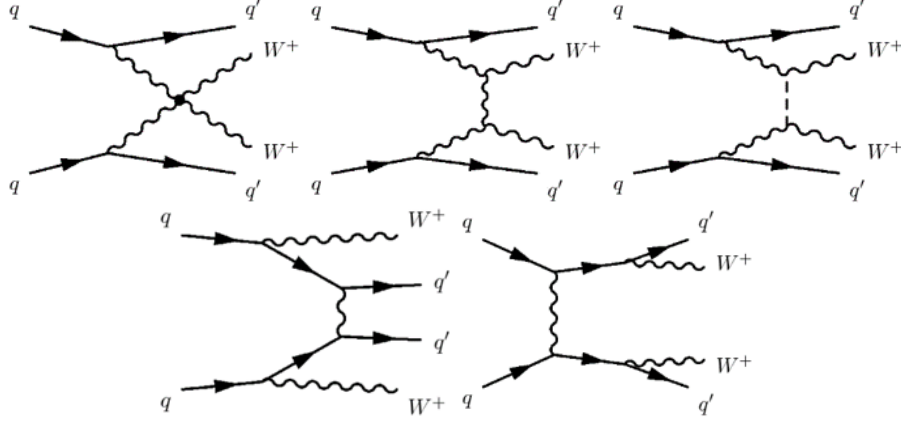


Figure 7.9: Representative Feynman diagrams for electroweak $W^+W^+ + 2$ jets production. The upper three diagrams represent VBS contributions, while the lower two diagrams represent non VBS contribution [98].

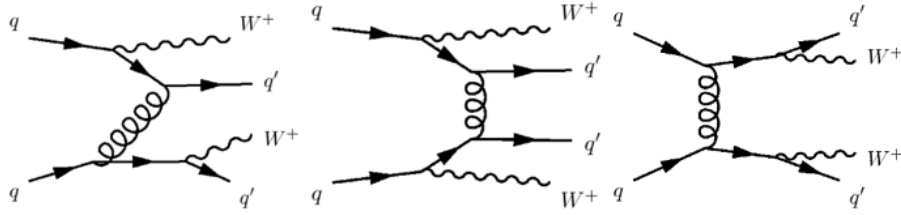


Figure 7.10: Representative Feynman diagrams for strong $W^+W^+ + 2$ jets production [98].

In the ATLAS analysis, two fiducial regions are defined: the *inclusive region*, to study the combination of electroweak and strong production mechanisms, and the *VBS region*, to investigate the electroweak interactions alone. The latter region is a subset of the inclusive one and requires the two jets with largest p_T to be separated in rapidity by $|\Delta y| > 2.4$. This enhances the purity of electroweak $W^\pm W^\pm + 2$ jets by removing most of the strong $W^\pm W^\pm + 2$ jets events, which are considered as a background in this region. The CMS analysis considers only the *VBS region* requiring $|\Delta\eta| > 2.5$, but extrapolates the $W^\pm W^\pm + 2$ jets production in an extended region.

Several SM processes enter the signal regions as irreducible physics processes or through instrumental effects. The main backgrounds are the irreducible *prompt-lepton background*, due to processes like $WZ/\gamma^* \rightarrow \ell^\pm \ell^\mp \ell^\pm \nu$ in which one lepton is outside of the experimental acceptance or does not satisfy the lepton identification criteria, the *conversion background*, in which oppositely charged prompt leptons are produced and one charge is misidentified, and other *non-prompt lepton background*, SM processes producing at least one non-prompt lepton from hadron decays.

Figure 7.11 shows the expected and observed m_{jj} distribution after all inclusive region selection criteria are applied, except $m_{jj} > 500$ GeV (left), and the $|\Delta y_{jj}|$ distribution after the inclusive region selection, measured by the ATLAS Collaboration. Results achieved by the CMS Experiment are presented in Figure 7.12, in which the m_{jj} (left) and the leading lepton p_T (right) distributions in the signal region are reported. All three di-lepton channels are summed in all figures.

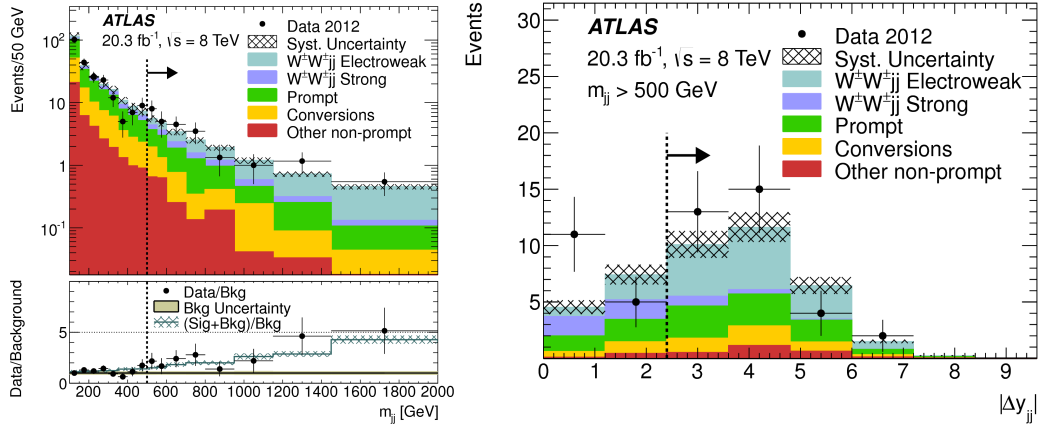


Figure 7.11: The m_{jj} distribution (left) for events passing the inclusive region selections except for the m_{jj} selection indicated by the dashed line. The $|\Delta y_{jj}|$ distribution (right) for events passing all inclusive region selections. The $|\Delta y_{jj}|$ selection is indicated by the dashed line. Distributions are measured by the ATLAS Experiment [97].

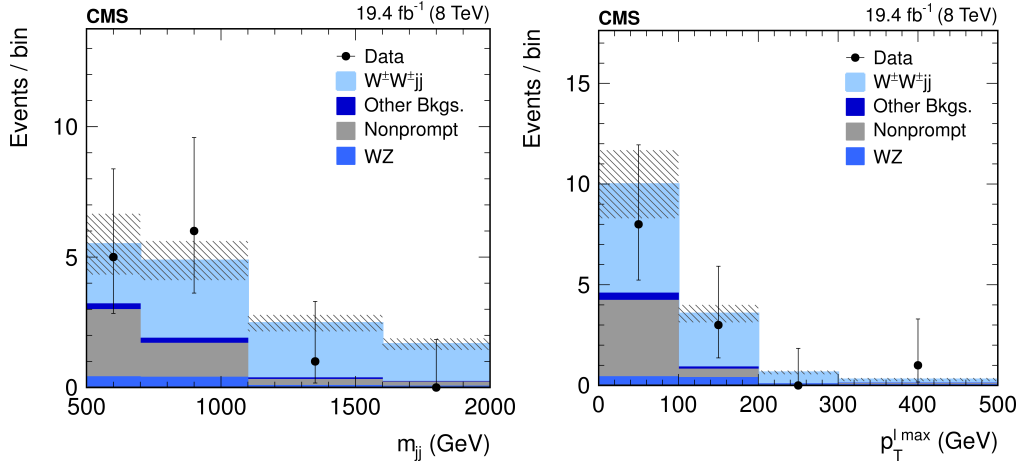


Figure 7.12: The distributions of m_{jj} (left) and leading lepton $p_T^{\ell, \max}$ in the signal region (right). The hatched bars include statistical and systematic uncertainties. Distributions are measured by the CMS Experiment [98].

The ATLAS Collaboration observes an excess of events, obtaining a significance over the background-only hypothesis of 4.5(3.6) standard deviations in the inclusive(VBS) region. This excess is interpreted as $W^\pm W^\pm + 2$ jets production and the fiducial cross-sections in the two regions are measured by combining the three decay channels in a likelihood function, finding:

$$\sigma_{incl}^{fid} = 2.1 \pm 0.5 \text{ (stat.)} \pm 0.3 \text{ (syst.) fb}$$

$$\sigma_{VBS}^{fid} = 1.3 \pm 0.4 \text{ (stat.)} \pm 0.2 \text{ (syst.) fb.}$$

The measured cross-sections are in agreement with the respective SM expectations of 1.52 ± 0.11 fb and 0.95 ± 0.06 fb. The cross-section measured by CMS in a wider fiducial region is

$$\sigma_{incl}^{fid} = 4.0^{+2.4}_{-2.0} \text{ (stat.)} {}^{+1.1}_{-1.0} \text{ (syst.) fb,}$$

with an expectation of 5.8 ± 1.2 fb.

Additional contributions to $W^\pm W^\pm + 2$ jets production can be expressed in a model-

independent way using higher-dimensional operators leading to anomalous quartic gauge boson couplings (aQGCs). For this measurement ATLAS uses the electroweak chiral Lagrangian approach, that can however be reinterpreted in terms of the EFT approach summarized in Section 1.6. The measured cross-section in the VBS fiducial region is used to set limits on aQGCs affecting vertices with four interacting W bosons and deviations from the SM are parameterized in terms of two parameters (α_4 and α_5). The one-dimensional projection at $\alpha_{4,5} = 0$ is respectively $0.14 < \alpha_4 < 0.16$ and $0.23 < \alpha_5 < 0.24$ compared to an expected $0.10 < \alpha_4 < 0.12$ and $0.18 < \alpha_5 < 0.20$ (Figure 7.13).

CMS uses instead the EFT approach. As explained in Section 1.6, there are several C and P conserving dimension eight effective operators modifying the quartic couplings between the weak gauge bosons. Figure 7.14 (left) shows the $m_{\ell\ell}$ distribution (which is found to be the most sensitive variable) for three values of $F_{T,0}/\Lambda^4$, a coefficient of one of the effective operators, together with the sum of all SM processes. The expected and observed two dimensional limits on the operator coefficients $F_{T,0}/\Lambda^4$ and $F_{T,1}/\Lambda^4$ are presented in Figure 7.14 (right).

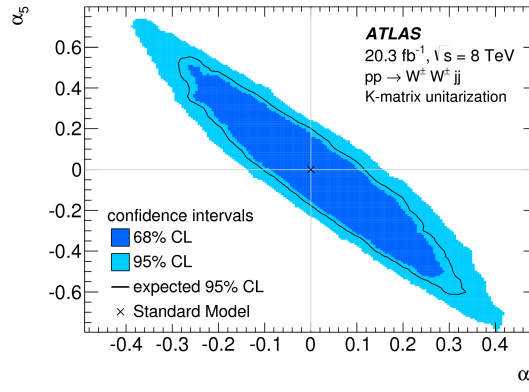


Figure 7.13: Limits on (α_4, α_5) . Points outside of the solid light ellipse are excluded by the data at 95% confidence level (CL). Points outside the inner dark ellipse are excluded at the 68% confidence level. The expected exclusion is given by the solid line. Distributions are measured by the ATLAS Experiment [97].

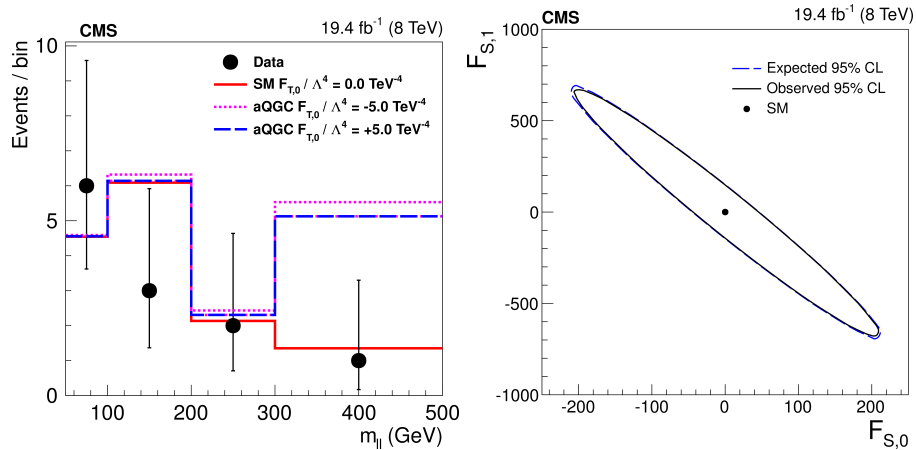


Figure 7.14: The $m_{\ell\ell}$ distributions (left) after full selection with all SM backgrounds and $F_{T,0}/\Lambda^4 = -5.0, 0$ (SM), and 5.0 TeV^{-4} ; the last bin includes overflow events. Observed and expected two-dimensional 95% C.L. limits (right) for $F_{T,0}/\Lambda^4$ and $F_{T,1}/\Lambda^4$. Distributions are measured by the CMS Experiment [98].

7.3 ZZ + 2 jets Channel

The production through vector boson scattering of a pair of Z bosons is very rare and up to now no analysis has been performed by the ATLAS and CMS Experiments using data collected in the Run I period. We will be able to fully investigate VBS interactions only with higher integrated luminosity, when it will be possible to observe these processes with a signal significance of at least 3σ in all the interesting channels, ZZ + 2 jets included. However, given the very similar final state with two Z bosons produced in association with jets, the work presented in Chapters 5 and 6 can be exploited to perform a preliminary study, starting exploring this channel even at 8 TeV.

In this Section, the electroweak production of a Z boson pair in association with 2 jets is investigated. Signal processes are simulated using Phantom MC generator, while the strong production of ZZ events, the main background of the analysis, is generated with MadGraph/Powheg ($qq/qg \rightarrow ZZ$) and MCFM ($gg \rightarrow ZZ$), using the same samples described in Subsection 5.1.2. Despite its low cross-section, the ZZ + 2 jets channel has a very clean signature, thanks to the four leptons in the final state. Leptons and jets are reconstructed following the same criteria reported in Chapter 4 and the ZZ best candidate is chosen as described in Subsection 5.2.1. A fiducial region is defined in order to enhance the purity of electroweak ZZ + 2 jets production and remove most of the strong events. This region is a subset of the tight region used to measure the cross-section of ZZ processes in Chapters 5 and 6. It is characterized by two Z bosons with mass between 60 and 120 GeV, at least one lepton with $p_T > 20$ GeV and another one with $p_T > 10$ GeV, electrons with $p_T^e > 7$ GeV and $|\eta^e| < 2.5$, muons with $p_T^\mu > 5$ GeV and $|\eta^\mu| < 2.4$.

In addition to that inclusive region, a set of cuts is defined exploiting the main features of VBS processes: the presence of two jets in the event with large invariant mass and well separated in pseudorapidity. The aim of the selection is to preserve the signal significance (roughly, the signal-over-background ratio) high, by eliminating the background contamination, while keeping the number of signal events as large as possible. In order to optimize the cuts on the di-jet invariant mass and pseudorapidity gap, Monte Carlo samples are used. Three different options on leading jets transverse momenta are considered:

- $p_T^{jet1} > 70$ GeV and $p_T^{jet2} > 50$ GeV;
- $p_T^{jet1} > 100$ GeV and $p_T^{jet2} > 70$ GeV;
- $p_T^{jet1} > 150$ GeV and $p_T^{jet2} > 100$ GeV;

where *jet1* and *jet2* are the leading- p_T and sub-leading- p_T jets. The second condition is found to be the most suitable. In Figure 7.15 signal and background contributions are reported as a function of the cut on m_{jj} , for four different values of $\Delta\eta_{jj}$ cut ($\Delta\eta_{jj} > 0$, $\Delta\eta_{jj} > 1.2$, $\Delta\eta_{jj} > 2.4$ and $\Delta\eta_{jj} > 4.2$). On the other hand, in Figure 7.16 signal and background contributions are shown as a function of the cut on $\Delta\eta_{jj}$, for four different cuts on m_{jj} ($m_{jj} > 0$ GeV, $m_{jj} > 300$ GeV, $m_{jj} > 600$ GeV and $m_{jj} > 900$ GeV). As it can be noticed, the tighter the thresholds, the more background is rejected. Nevertheless, a too stringent selection will result also in a loss of the large part of the signal. The fiducial region is finally defined requiring at least two jets with an invariant mass m_{jj} larger than 300 GeV and separated in pseudorapidity by $\Delta\eta_{jj} > 2.4$. Moreover, the magnitude of the missing transverse energy is demanded to be lower than 60 GeV, in order to reduce background events from *W* or top decays. Only one event passing all selection requirements is observed in the data, compared to a SM expectation of 0.14 signal events and 0.47 background events. Figure 7.17 (left) shows the expected and observed $m_{4\ell}$ distribution after all inclusive region selection

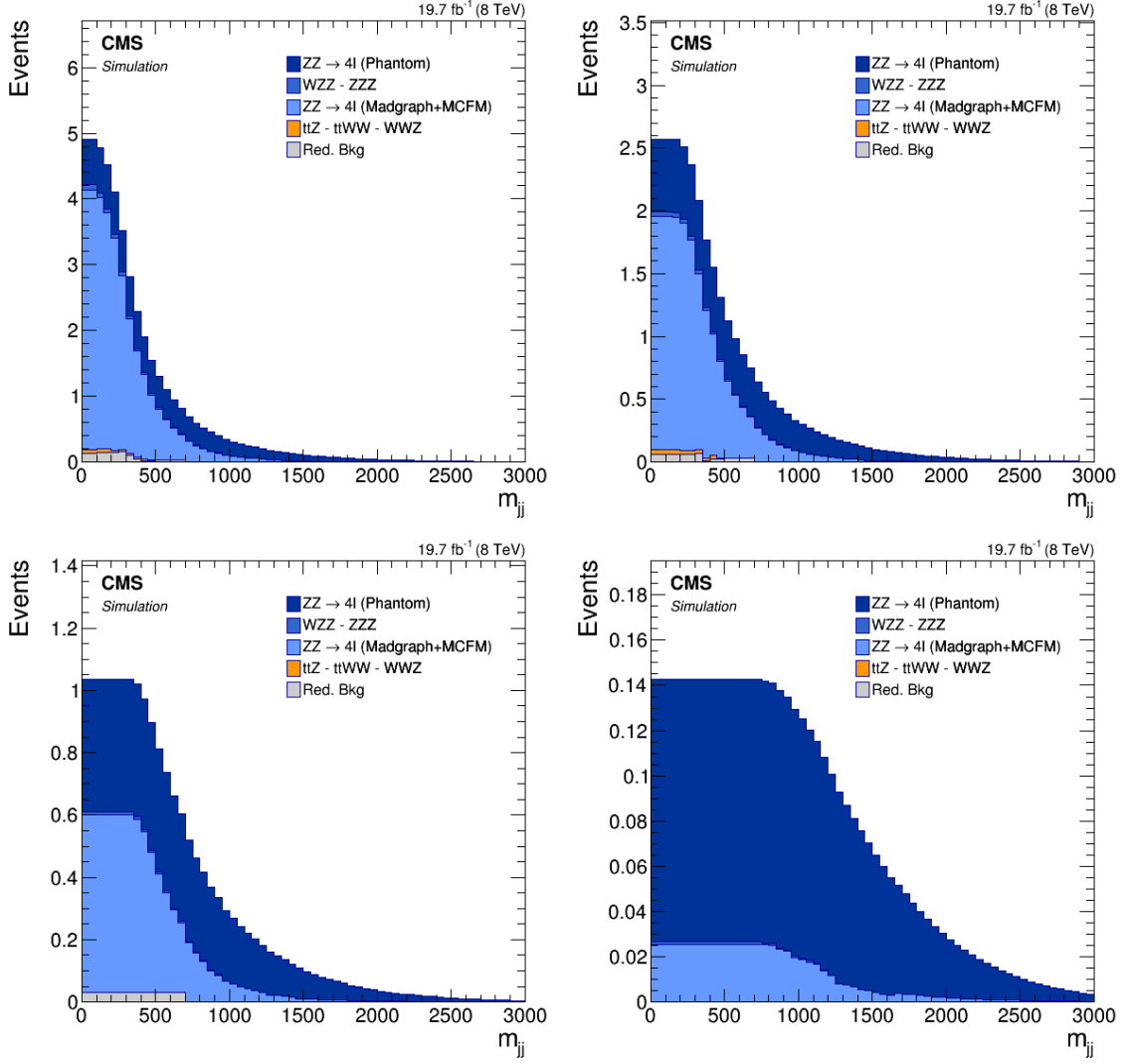


Figure 7.15: Signal (Phantom) and background contributions as a function of the cut on m_{jj} , for four different cuts on $\Delta\eta_{jj}$: $\Delta\eta_{jj} > 0$ (top-left), $\Delta\eta_{jj} > 1.2$ (top-right), $\Delta\eta_{jj} > 2.4$ (bottom-left) and $\Delta\eta_{jj} > 4.2$ (bottom-right).

criteria are applied, except $\Delta\eta > 2.4$. In this region, four events are observed and their kinematics is reported in table 7.1. Figure 7.17 (right) shows the m_{jj} distribution after the whole selection. All three final state channels are summed in all plots.

7.4 Future Perspectives

In the following, a summary of several studies at 14 TeV published by both ATLAS and CMS Collaborations is reported [99, 100]. Prospects for the observation of vector boson scattering and anomalous quartic gauge couplings are presented, including assessments on the discovery potential. Two different scenarios are taken into account:

- upgraded detector simulation based on the current one and integrated luminosity up to 300 fb^{-1} , reached at the end of the Run III period;
- integrated luminosity of 3000 fb^{-1} collected at a High-Luminosity LHC and using

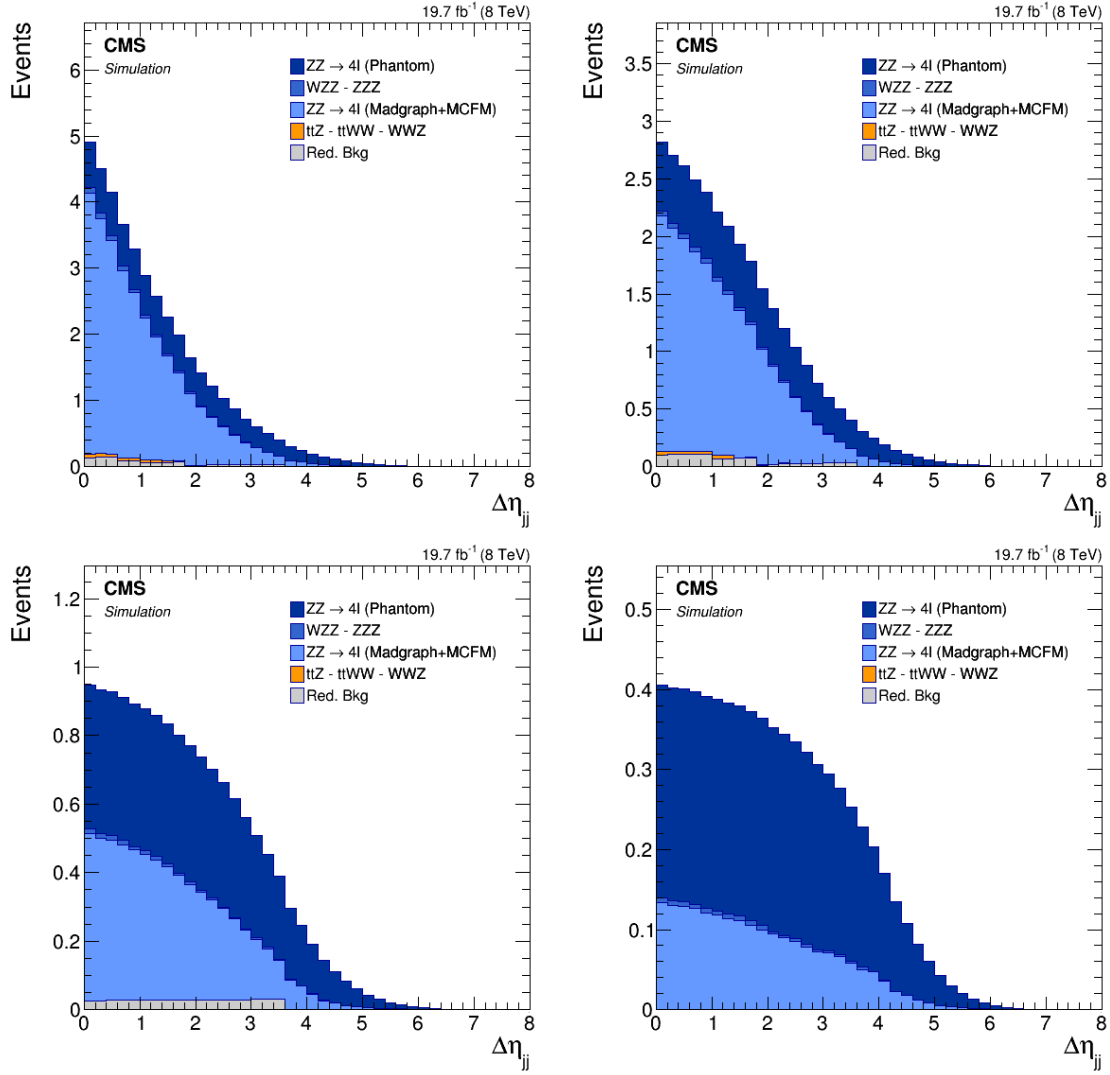


Figure 7.16: Signal (Phantom) and background contributions as a function of the cut on $\Delta\eta_{jj}$, for four different cuts on m_{jj} : $m_{jj} > 0$ GeV (top-left), $m_{jj} > 300$ GeV (top-right), $m_{jj} > 600$ GeV (bottom-left) and $m_{jj} > 900$ GeV (bottom-right).

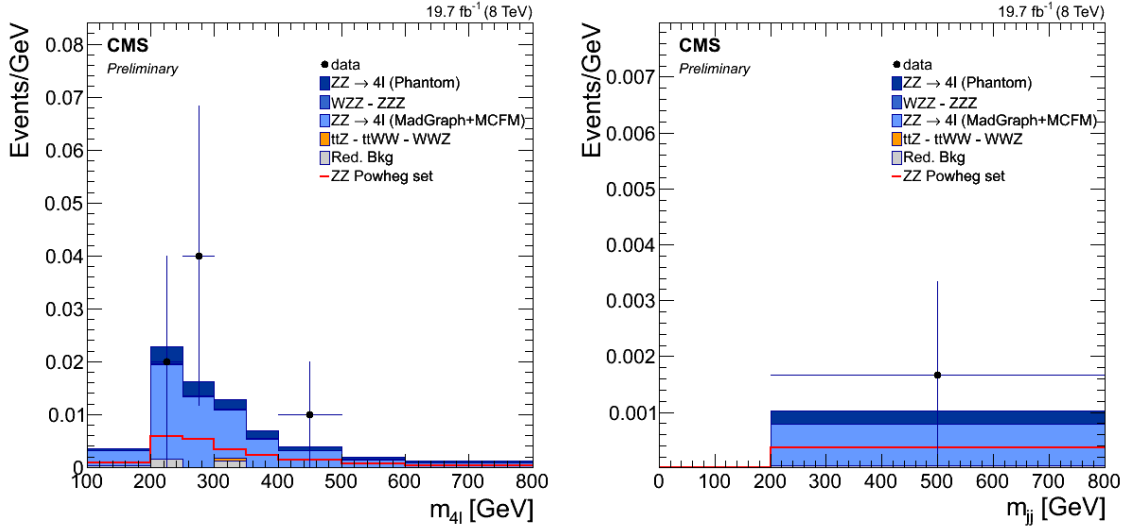
proposed upgraded detector configurations.

New physics effects are studied using an effective field theory approach for modeling anomalous contributions to quartic gauge couplings. The failure of direct searches to discover any new physics particles in the 7 and 8 TeV LHC runs indicates that in many classes of new physics theories new particles will have masses greater than 1 TeV. For sufficiently heavy new particles, not directly detectable, the observed effect would be an anomalous enhancement of the cross-section at higher scattering energies. Depending on the form of the new physics it may be represented by a large class of operators in effective field theory. In the analyses summarized below, dimension-six and dimension-eight operators sensitive to anomalous triple and quartic gauge couplings are investigated (Section 1.6).

Four studies of sensitivity to beyond-SM (BSM) scenarios are presented: VBS for ZZ production in the four-lepton channel, VBS for WZ production in the tri-lepton channel, VBS for WW production in the opposite sign di-lepton channel and $Z\gamma\gamma$ tri-boson production in the di-lepton plus di-photon channel.

Table 7.1: Kinematic values of events observed in the fiducial region without requiring $\Delta\eta_{jj} > 2.4$.

Event	$m_{4\ell}$ [GeV]	m_{jj} [GeV]	$\Delta\eta_{jj}$	p_T^{jet1} [GeV]	p_T^{jet2} [GeV]
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

**Figure 7.17:** The $m_{4\ell}$ distribution (left) for events passing the region selections except for the $\Delta\eta_{jj} > 2.4$ cut. The m_{jj} distribution (right) for events passing all the region selections. Points represent the data, the stacked fulfilled 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).

7.4.1 ZZ + 2 jets Channel

This channel suffers from a small cross-section, but provides a very clean final state, with two Z bosons decaying in four leptons. The kinematics of the outgoing di-boson system is fully reconstructible and this allows the cross-section measurement as a function of the center-of-mass energy, probing the energy dependence of the new physics. The ATLAS analysis [99] selects two pairs of opposite sign and same flavor leptons (electrons or muons) with a transverse momentum above 25 GeV. At least two jets with a transverse momentum above 50 GeV are required and the invariant mass of the two most highest- p_T jets must be above 1 TeV.

In order to determine the expected sensitivity to BSM ZZ contribution, the dimension-six operator $\mathcal{O}_{\phi W}$ is investigated through its coefficient $c_{\phi W}/\Lambda^2$. Figure 7.18 shows the reconstructed four-lepton invariant mass distribution and signal significance as a function of $c_{\phi W}/\Lambda^2$. In Table 7.2 the 5σ discovery potential is illustrated.

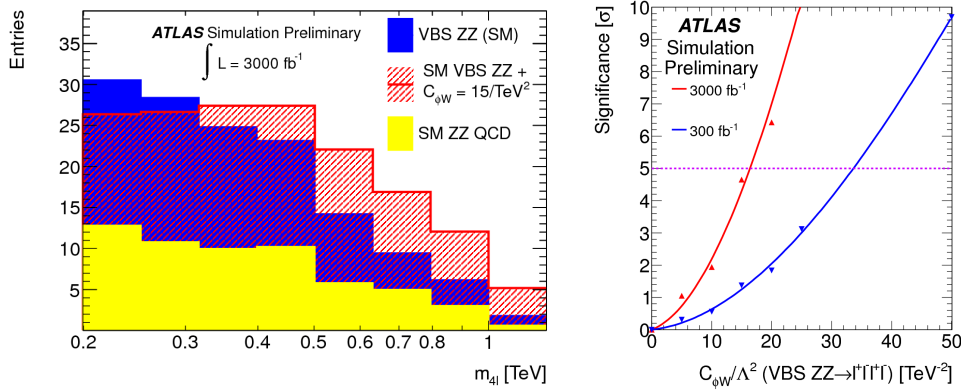


Figure 7.18: The reconstructed four-lepton mass ($m_{4\ell}$) spectrum after requiring $m_{jj} > 1 \text{ TeV}$ in the $pp \rightarrow ZZ + jj \rightarrow 4\ell + jj$ process (left). The corresponding signal significance as a function of $c_{\phi W}/\Lambda^2$ (right). Distributions are measured by the ATLAS Experiment [99].

7.4.2 WZ + 2 jets Channel

This channel presents a cross-section times branching ratio larger than the ZZ + 2 jets final state, but the full kinematics reconstruction of the process is more difficult since a neutrino is produced in the decay. Both ATLAS and CMS estimate up to which energy scale anomalous couplings could be observed at HL-LHC [99, 100]. The ATLAS(CMS) event selection requires exactly three high transverse momentum leptons (electrons or muons) with $p_T > 25(20) \text{ GeV}$. Two of them must be consistent with a Z boson. At least two jets with a transverse momentum above 50 GeV must be present and the di-jet invariant mass is required to be at least 1000(600) GeV. The CMS analysis additionally requires that the two jets are separated by at least 4.0 in pseudorapidity.

Both Collaborations obtain the expected sensitivity to BSM WZ contribution, by studying a dimension-eight operator through its coefficient f_{T1}/Λ^4 . Figure 7.19 shows the reconstructed WZ invariant mass distribution and signal significance as a function of f_{T1}/Λ^4 measured by ATLAS. Figure 7.20 presents the reconstructed WZ invariant mass distribution using 300 and 3000 fb^{-1} and assuming $f_{T1}/\Lambda^4 = 1.0 \text{ TeV}^{-4}$ obtained by CMS. In Table 7.2 the 5σ discovery potential is illustrated, for both ATLAS and CMS analyses.

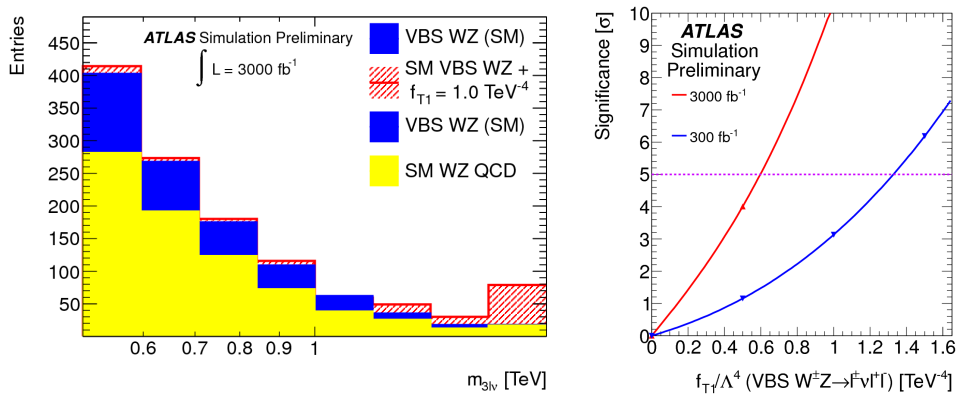


Figure 7.19: The reconstructed WZ mass spectrum, using the charged leptons and the neutrino solution after requiring $m_{jj} > 1 \text{ TeV}$, in the $pp \rightarrow WZ + jj \rightarrow 3\ell\nu + jj$ channel analysis (left). The corresponding signal significance as a function of f_{T1}/Λ^4 (right). Distributions are measured by the ATLAS Experiment [99].

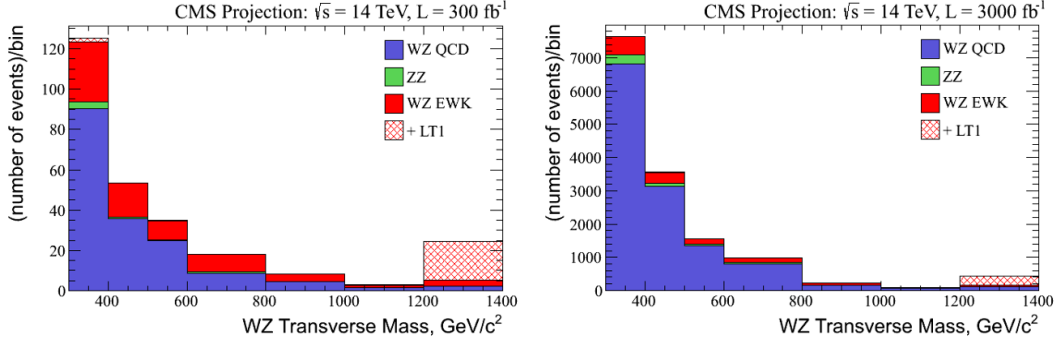


Figure 7.20: WZ transverse mass with 300 (left) and 3000 fb^{-1} (right). The additional contribution of an aQGC signal for the dimension-eight operator with $f_{T1}/\Lambda^4 = 1.0 \text{ TeV}^{-4}$ is illustrated by the red hatched histogram. Distributions are measured by the CMS Experiment [100].

7.4.3 $W^\pm W^\pm + 2 \text{ jets Channel}$

This channel has a cross-section times branching ratio larger with respect to the previous final states, but the second neutrino makes the reconstruction of the full kinematics of the weak boson scattering system impossible. As reported in Section 7.2, both ATLAS and CMS Experiments have analyzed the $W^\pm W^\pm + 2 \text{ jets}$ channel using the 8 TeV data set. However only the ATLAS Collaboration studies the sensitivity for this channel at HL-LHC. Exactly two leptons of the same charge, with a transverse momentum above 25 GeV are required in the selection, together with at least two jets with a transverse momentum above 50 GeV. The invariant mass of the two selected highest- p_T jets must exceed 1 TeV. The distribution of the mass of the two jets and two leptons is used to discriminate signal from background, since calculating the WW invariant mass is not possible when two neutrinos are present. In this case, a dimension-eight operator is chosen to parameterize new physics in terms of the magnitude of the coefficient f_{S0}/Λ^4 . Figure 7.21 shows the $m_{\ell\ell jj}$ distribution and signal significance as a function of f_{S0}/Λ^4 . In Table 7.2 the 5σ discovery potential is illustrated.

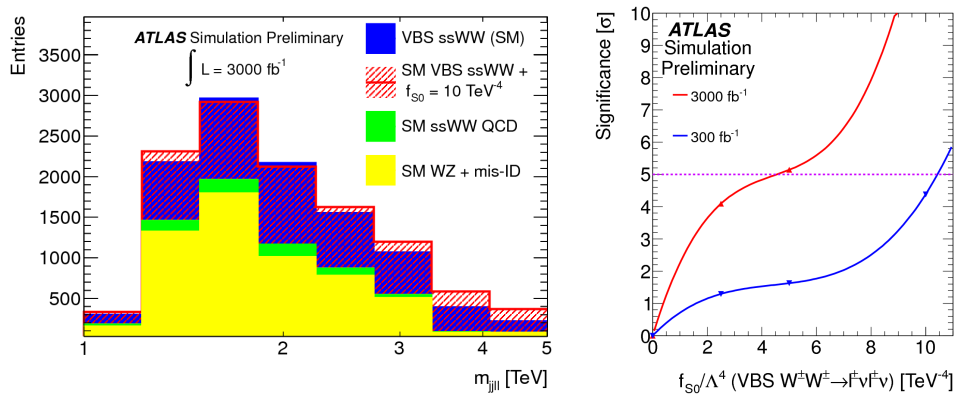


Figure 7.21: The reconstructed 4-body mass spectrum using the two leading leptons and jets in the $pp \rightarrow WW + jj \rightarrow 2\ell 2\nu + jj$ channel (left). The corresponding signal significance as a function of f_{S0}/Λ^4 (right). Distributions are measured by the ATLAS Experiment [99].

7.4.4 $Z\gamma\gamma$ Channel

The $Z\gamma\gamma$ mass spectrum at high mass is sensitive to BSM triboson contributions. The lepton-photon channel allows the final state to be fully reconstructed and thus the $Z\gamma\gamma$ invariant mass can be calculated. The sensitivity to new physics for this channel is studied by the ATLAS Experiment [99]. The selection requires a pair of opposite-sign and same-flavor leptons with p_T above 25 GeV and at least one lepton with p_T above 160 GeV. The same is required for the photon pair. Furthermore, photons and leptons must be well separated.

In this analysis, the new physics is parameterized using two dimension-eight operators and their coefficients f_{T8}/Λ^4 and f_{T9}/Λ^4 . Figure 7.22 shows the expected signal significance as a function of BSM physics parameters. In Table 7.2 the 5σ discovery potential is illustrated.

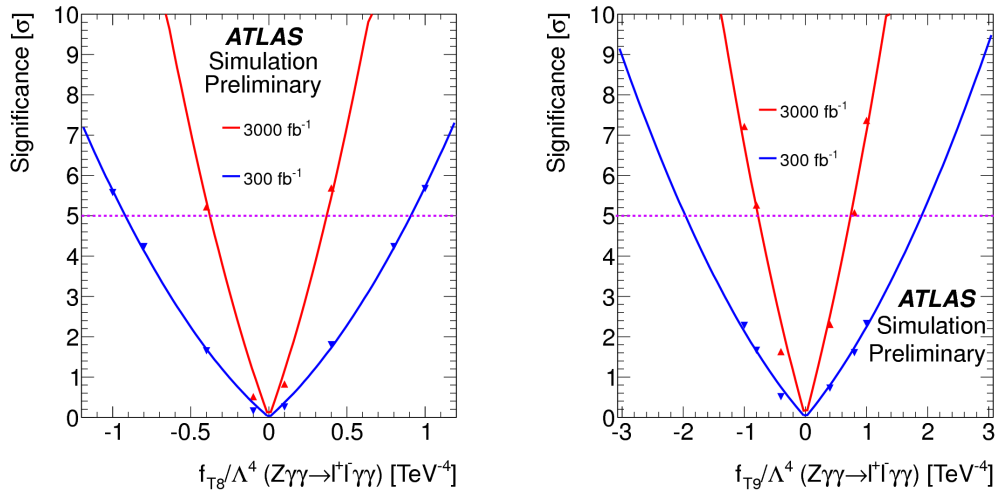


Figure 7.22: Signal significance as a function of f_{T8}/Λ^4 (left) and f_{T9}/Λ^4 (right) in the $Z\gamma\gamma$ channel analysis. Distributions are measured by the ATLAS Experiment [99].

Table 7.2: Projected 5σ discovery achieved for the vector boson scattering analyses at HL-LHC, with $\mathcal{L} = 3000 \text{ fb}^{-1}$.

Experiment	Final State	Parameter	Op. Dim.	5σ -significance
ATLAS	$\ell^+\ell^-\ell^+\ell^-jj$	$c_{\phi W}/\Lambda^2$	6	16 TeV^{-2}
CMS/ATLAS	$\ell^\pm\nu\ell^+\ell^-jj$	f_{T1}/Λ^4	8	$0.55/0.6 \text{ TeV}^{-4}$
ATLAS	$\ell^\pm\nu\ell^\pm\nu jj$	$f_{S,0}/\Lambda^4$	8	4.5 TeV^{-4}
ATLAS	$\ell^+\ell^-\gamma\gamma jj$	$f_{T8}/\Lambda^4 - f_{T9}/\Lambda^4$	8	$0.4 - 0.7 \text{ TeV}^{-4}$

Conclusions

In this thesis I have presented the study of the $pp \rightarrow ZZ$ process, carried out by analyzing proton-proton collision data recorded by the CMS detector during the 2012 run, corresponding to an integrated luminosity of 19.7 fb^{-1} . The total cross-section for the production of two Z bosons has been measured in the leptonic decay channels $ZZ \rightarrow 2\ell 2\ell'$, with $\ell, \ell' = e, \mu$. Signal events have been selected using simple sequential cuts based on lepton reconstruction, identification and isolation, as well as a set of kinematic conditions. The main background contributions have been estimated using Monte Carlo samples and data-driven techniques and found to be very small within the fiducial region in which the measurement is performed, i.e. $60 < m_{Z_{1,2}} < 120 \text{ GeV}$. In this region, the cross-section obtained combining the $4e$, 4μ and $2e2\mu$ final states is measured to be $\sigma_{pp \rightarrow ZZ} = 7.76 \pm 0.48 \text{ (stat.)}_{-0.29}^{+0.27} \text{ (syst.) pb}$, in perfect agreement with the theoretical value of $7.5 \pm 0.5 \text{ pb}$. The total cross-section is also evaluated in a tighter fiducial region, defined by the kinematic and geometrical acceptance of leptons, and is found to be $\sigma_{ZZ \rightarrow 4\ell} = 20.50 \pm 1.26 \text{ (stat.)}_{-0.83}^{+0.79} \text{ (syst.) fb}$, in agreement with the $19.7 \pm 1.6 \text{ fb}$ expected.

Besides the inclusive measurement, in this work I have also reported the study of differential cross-section as a function of several interesting variables, in particular related to the jets produced in association with the Z boson pair. The distributions of jet-multiplicity, invariant mass of the two p_T -leading jets, the separation between them and their transverse momentum and pseudorapidity have been analyzed. A slight discrepancy with respect to Monte Carlo samples is found and it is attributed to missing NNLO corrections, indicating the need to improve the MC simulation for future studies on $ZZ + \text{jets}$ processes.

These measurements have never been performed so far and carry very useful information that can be exploited for other related Standard Model processes or searches for new physics. Given the very similar final state with two Z bosons produced in association with jets, this analysis has thus been extended in order to start exploring vector boson scattering processes in the ZZ channel. A fiducial region has been defined to enhance the purity of electroweak $ZZ + 2 \text{ jets}$ events and remove the strong ZZ production. The observation of one event in this region agrees with the Standard Model expectation.

During the first three years of its operation, LHC has been working extremely well, providing collisions at a luminosity and energy never before achieved. A new high-energy frontier on the sub-nuclear sector has been opened, allowing precise studies on Standard Model processes to be obtained, and measurements reported in this work are an example of this. The discovery of a Higgs boson with a mass around 125 GeV announced by the CMS and ATLAS Collaborations represents an extremely important result, setting a milestone in the history of particle physics. Nevertheless, there is still much to be done and further questions have to be addressed by the experiments.

After a long technical stop of more than two years, in June 2015 the LHC has started its second data-taking period, reaching the center-of-mass energy of 13 TeV . New data are going to be collected and analyzed to further understand the electroweak symmetry breaking mechanism, possibly the highest priority of particle physics today. Vector boson scattering

processes are thus becoming even more important and the work reported in this thesis will be a starting point for future searches related to this subject. Moreover, other beyond Standard Model scenarios are going to be probed thanks to the improved sensitivity to signatures of new physics, possibly opening other interesting puzzles to be solved in the future.

Appendix A

Unfolding Validation Tests

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 A.1 to figure A.22 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.

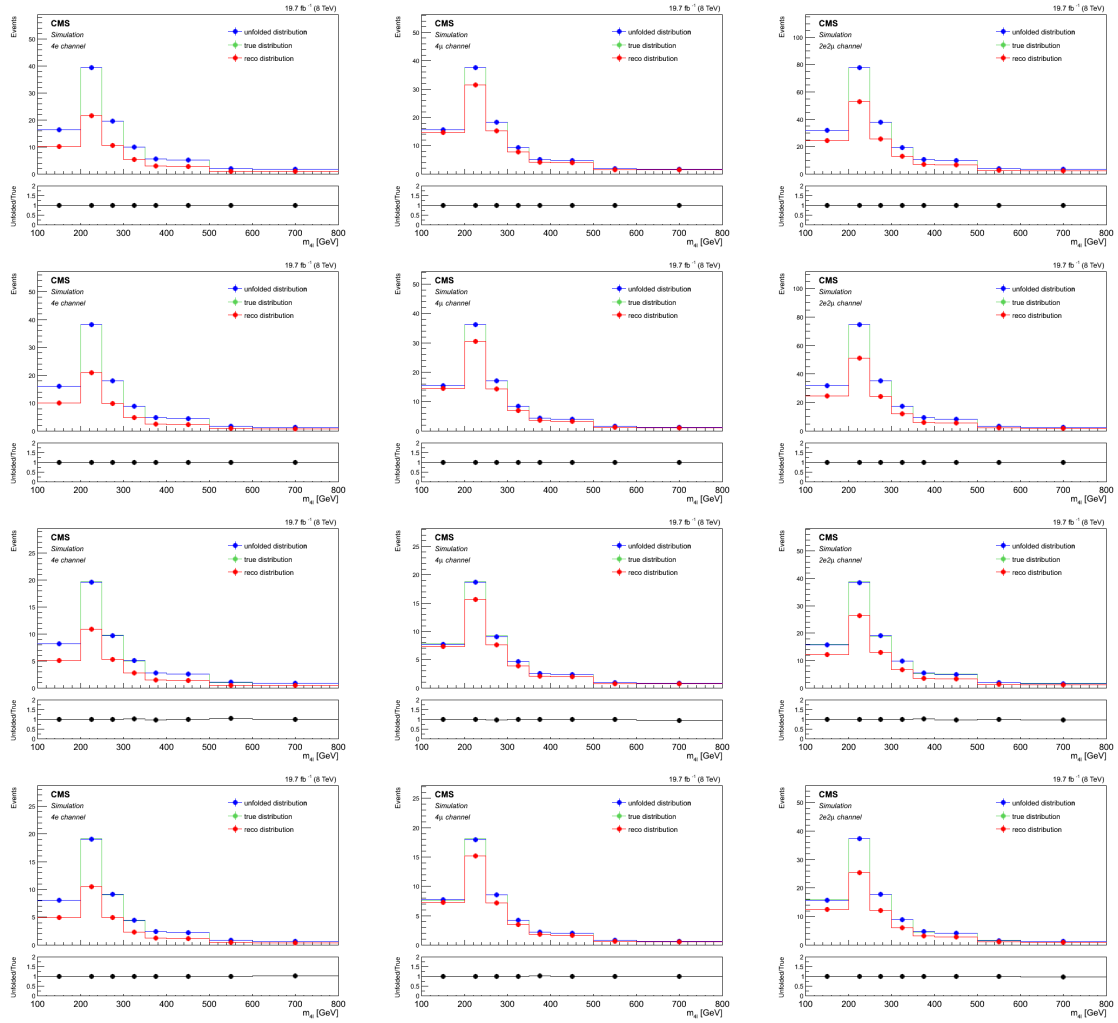


Figure A.1: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powhag matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powhag matrix applied on MadGraph distribution using half set. Results are reported as a function of the 4-lepton mass system for the $4e$, 4μ and $2e2\mu$ final states.

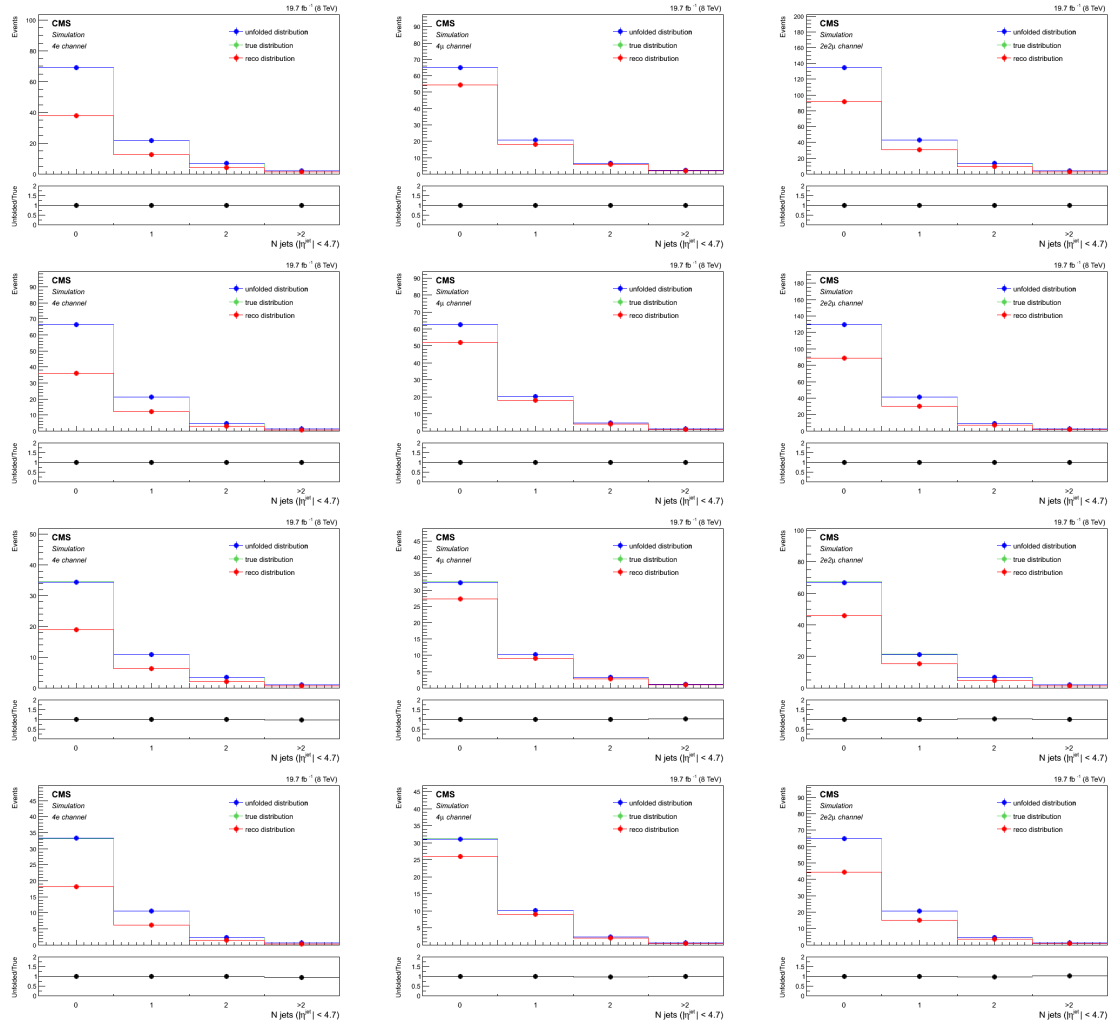


Figure A.2: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix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

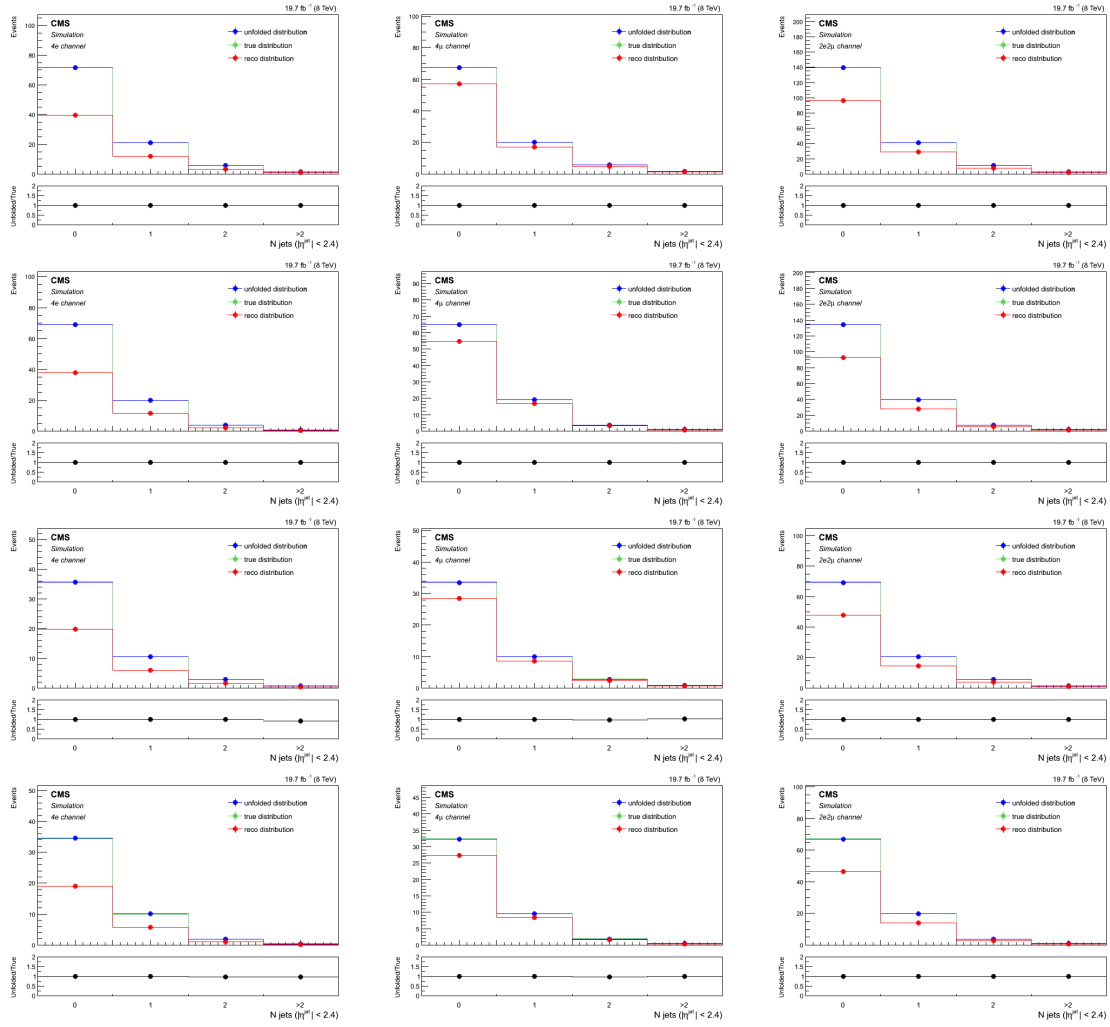


Figure A.3: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powhcg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powhcg matrix applied on MadGraph distribution using half set. Results are reported as a function of the number of central jets for the $4e$, 4μ and $2e2\mu$ final states.

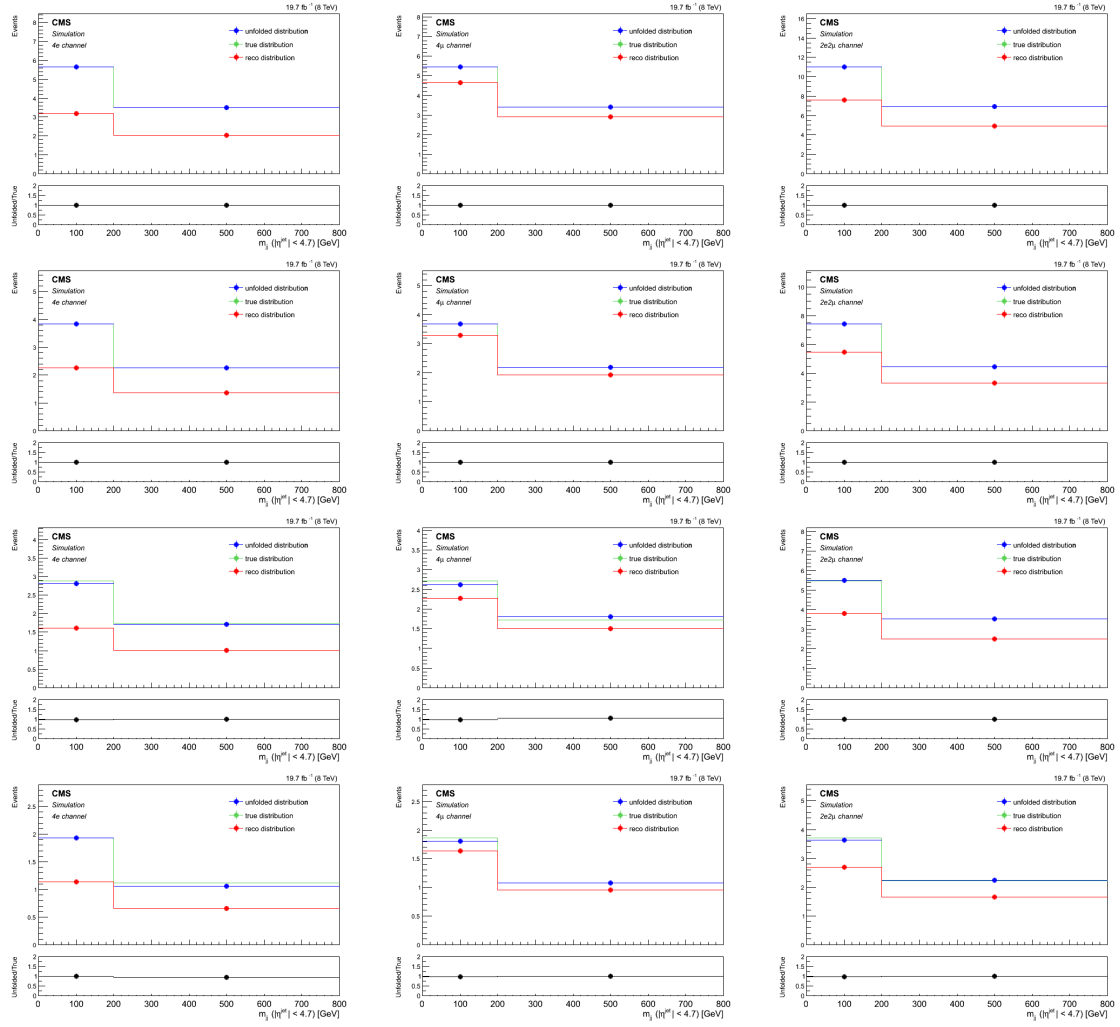


Figure A.4: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix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

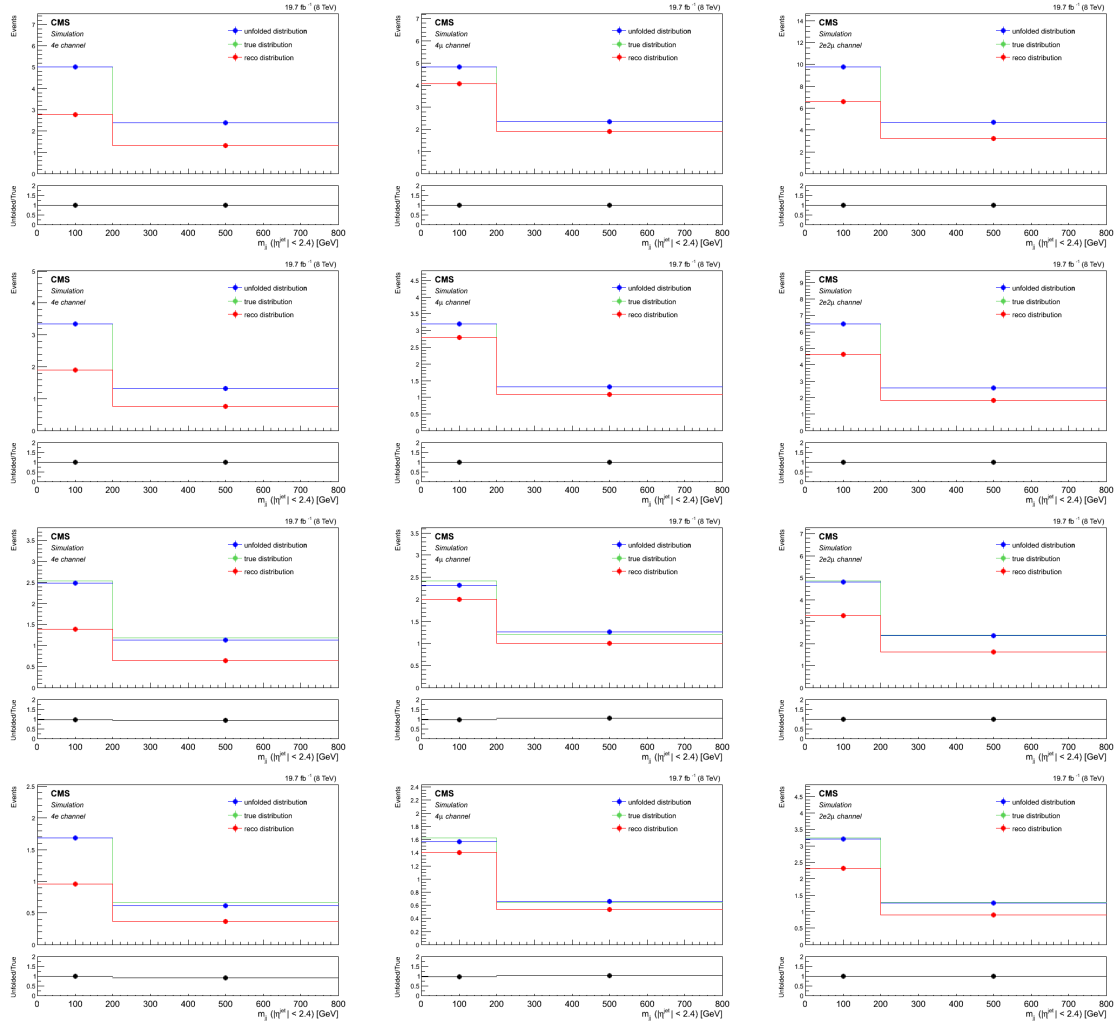


Figure A.5: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powhag matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powhag matrix applied on MadGraph distribution using half set. Results are reported as a function of m_{jj} (with $|\eta^{jet}| < 2.4$) for the $4e$, 4μ and $2e2\mu$ final states.

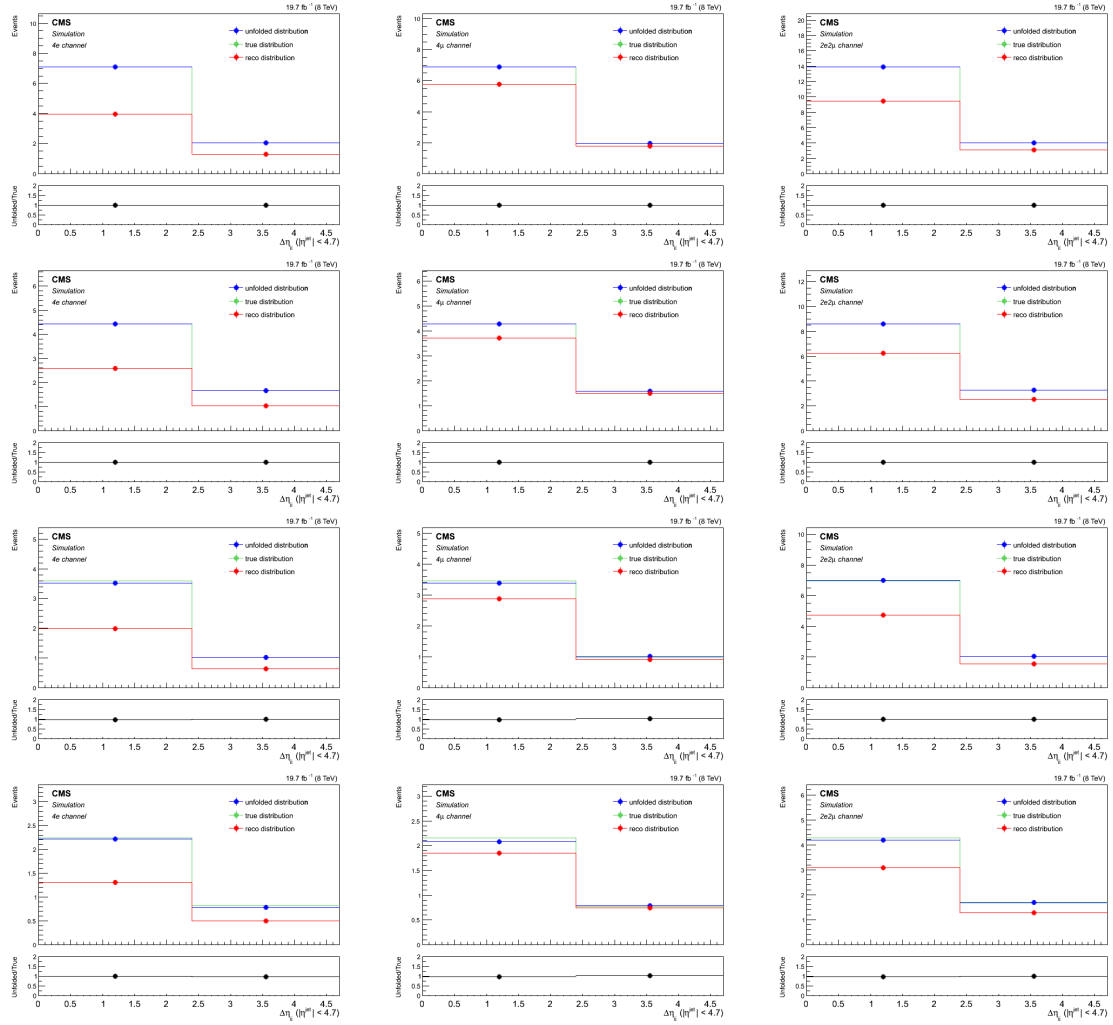


Figure A.6: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on MadGraph distribution using half set. Results are reported as a function of $\Delta\eta_{jj}$ for the $4e$, 4μ and $2e2\mu$ final states.

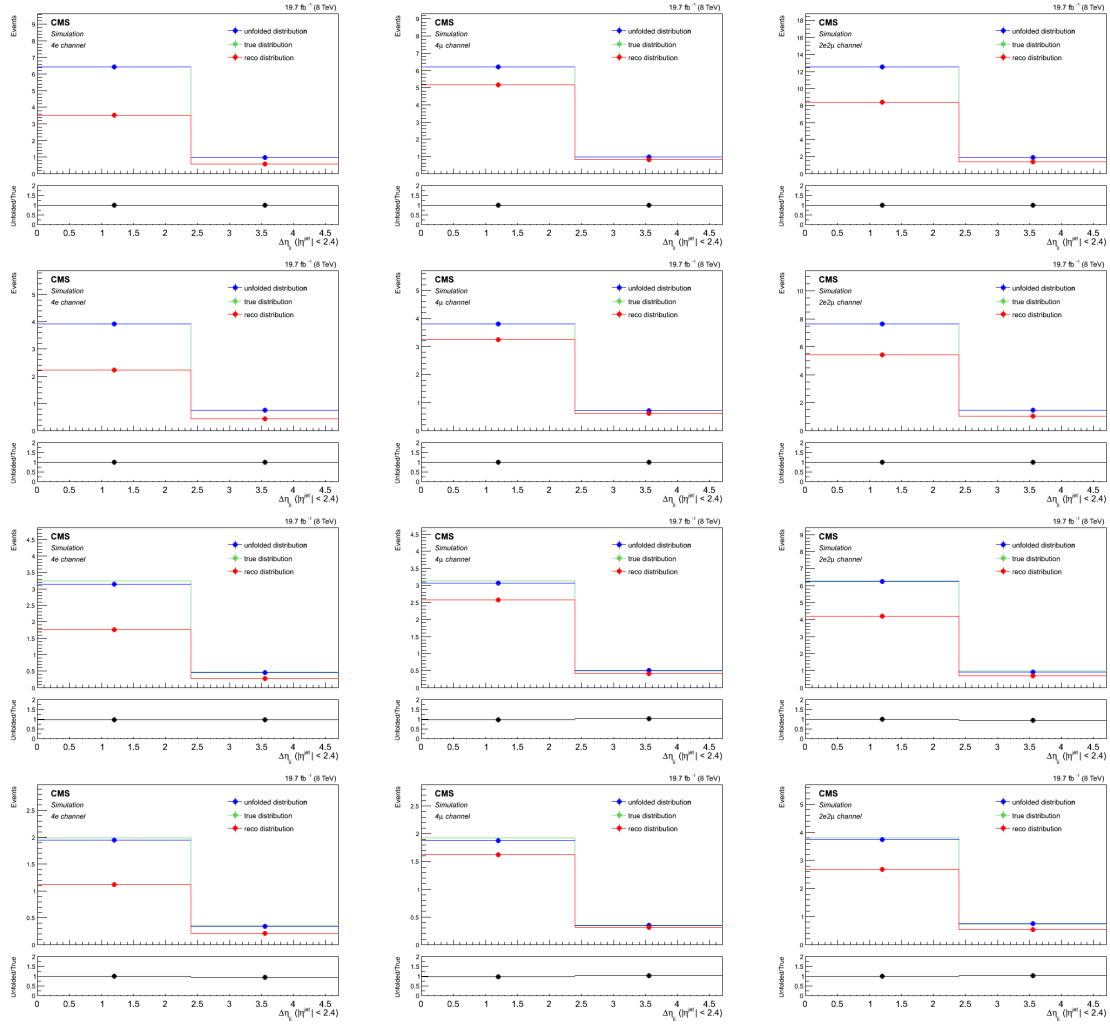


Figure A.7: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powhpeg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powhpeg matrix applied on MadGraph distribution using half set. Results are reported as a function of $\Delta\eta_{jj}$ (with $|\eta^{jet}| < 2.4$) for the $4e$, 4μ and $2e2\mu$ final states.

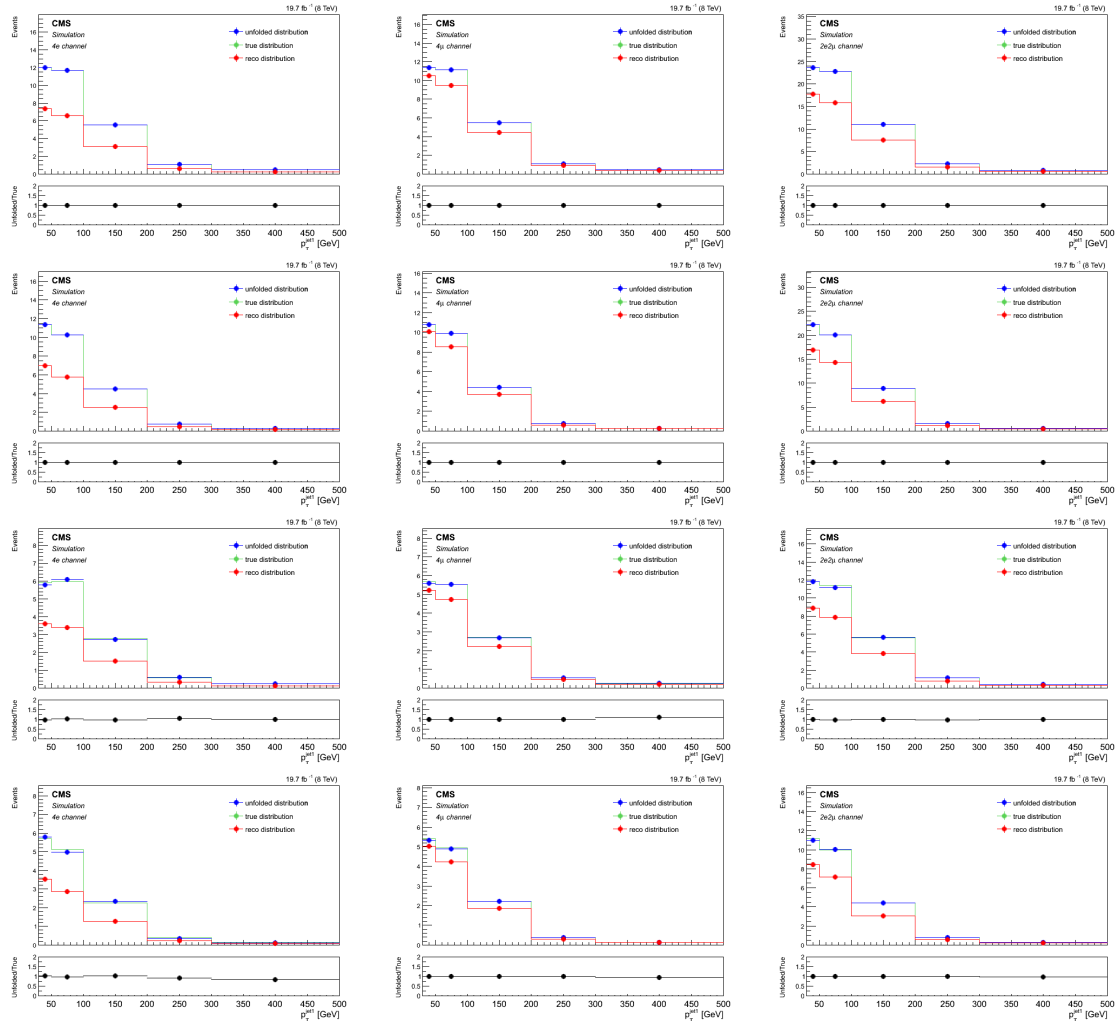


Figure A.8: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on MadGraph distribution using half set. Results are reported as a function of the p_T of the leading jet, for the 4e, 4μ and 2e2μ final states.

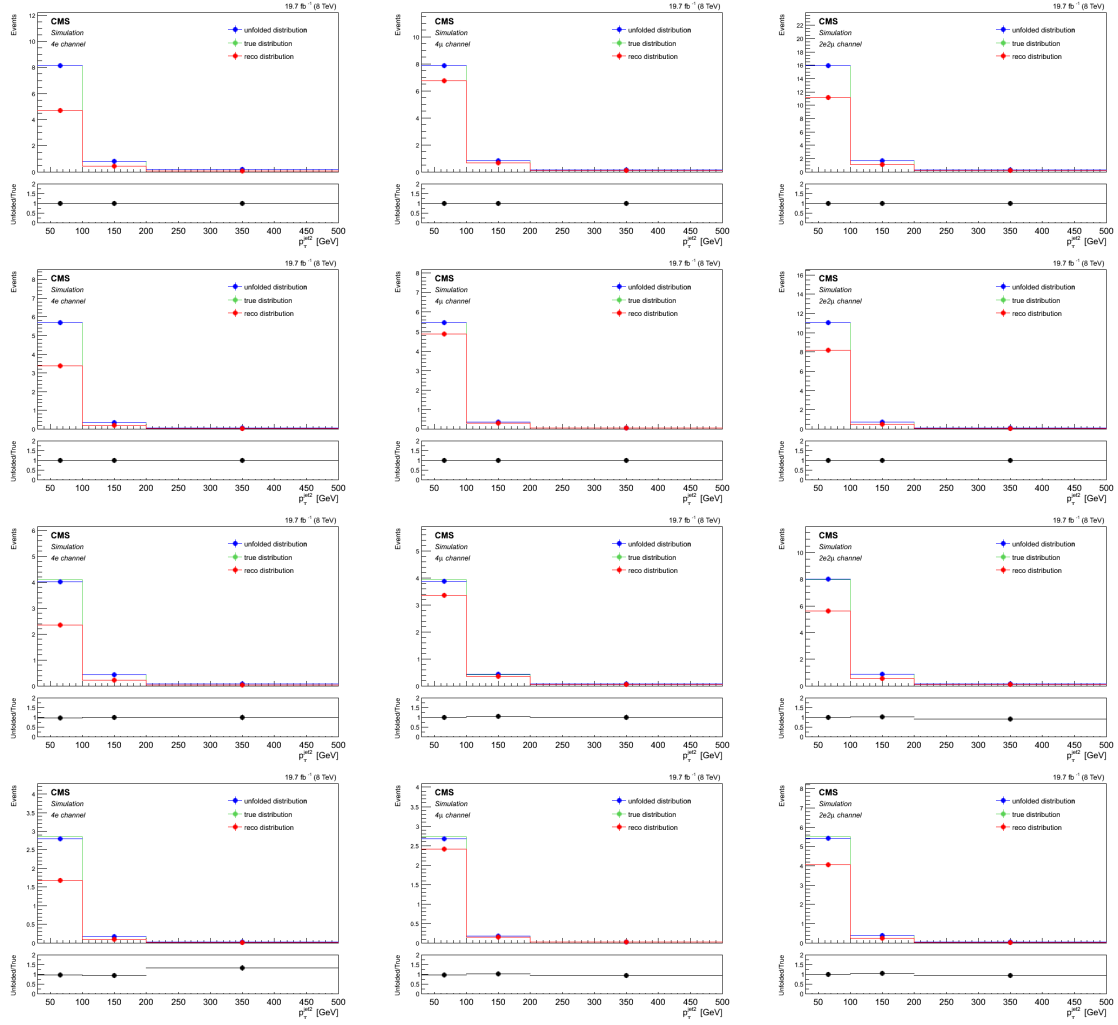


Figure A.9: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powhcg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powhcg matrix applied on MadGraph distribution using half set. Results are reported as a function of the p_T of the sub-leading jet, for the $4e$, 4μ and $2e2\mu$ final states.

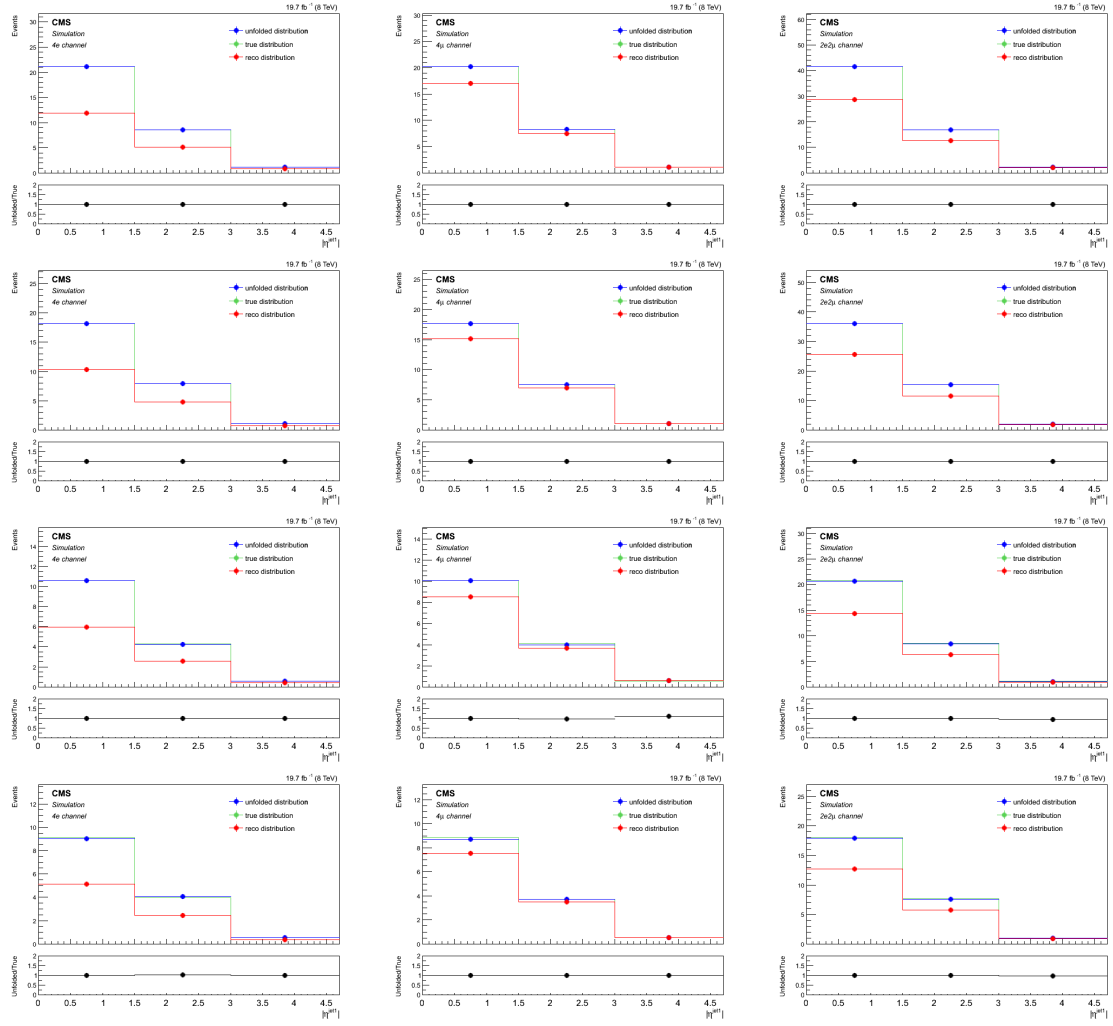


Figure A.10: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on MadGraph distribution using half set. Results are reported as a function of the η of the leading jet, for the $4e$, 4μ and $2e2\mu$ final states.

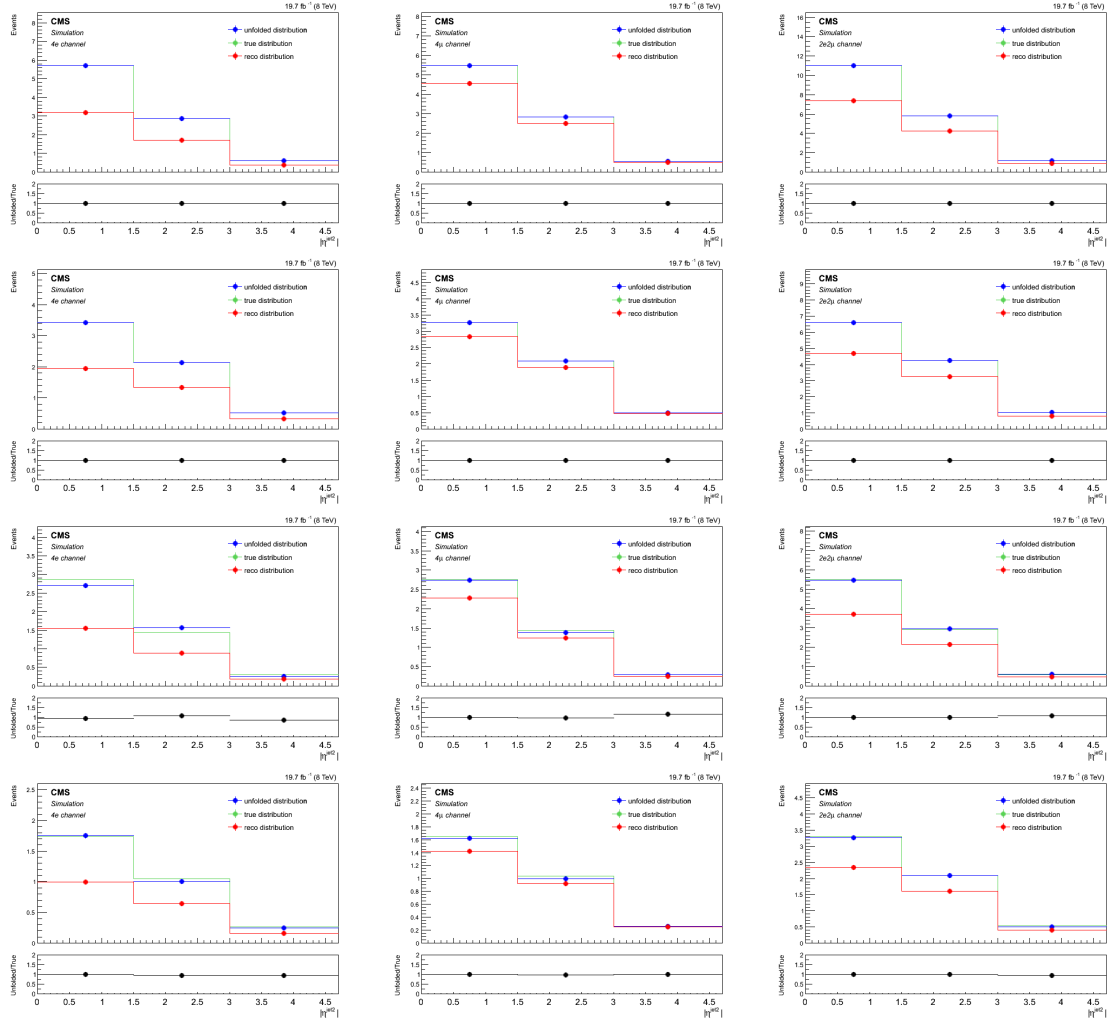


Figure A.11: Unfolding test. From top to bottom: MadGraph matrix applied on MadGraph distribution using the full set, Powheg matrix applied on MadGraph distribution using the full set, MadGraph matrix applied on MadGraph distribution using the two different halves of the total sample set, Powheg matrix applied on MadGraph distribution using half set. Results are reported as a function of the η of the sub-leading jet, for the $4e$, 4μ and $2e2\mu$ final states.

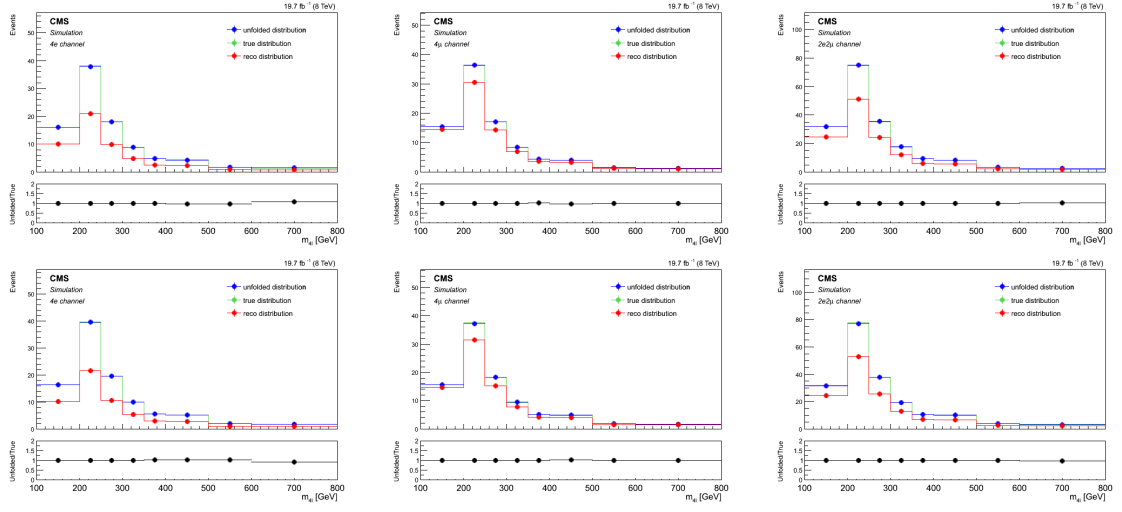


Figure A.12: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of the 4-lepton mass system for the 4e, 4μ and 2e2μ final states.

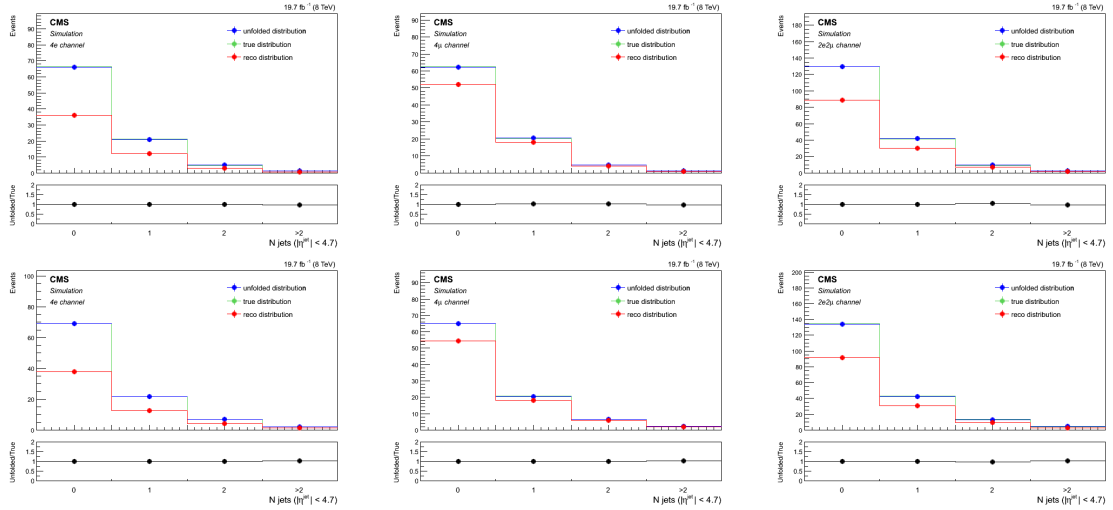


Figure A.13: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of the number of jets for the 4e, 4μ and 2e2μ final states.

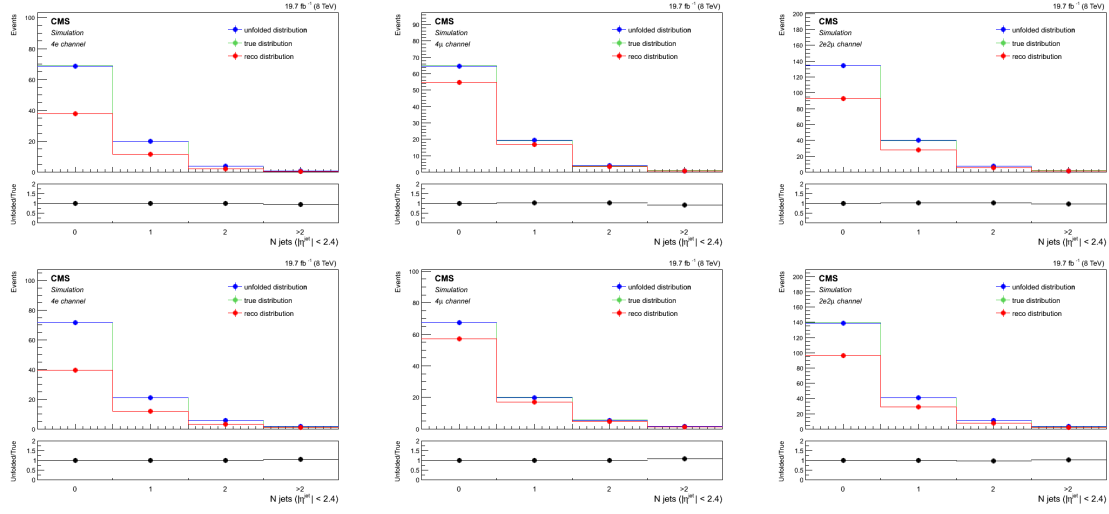


Figure A.14: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of the number of central jets for the 4e, 4μ and 2e2μ final states.

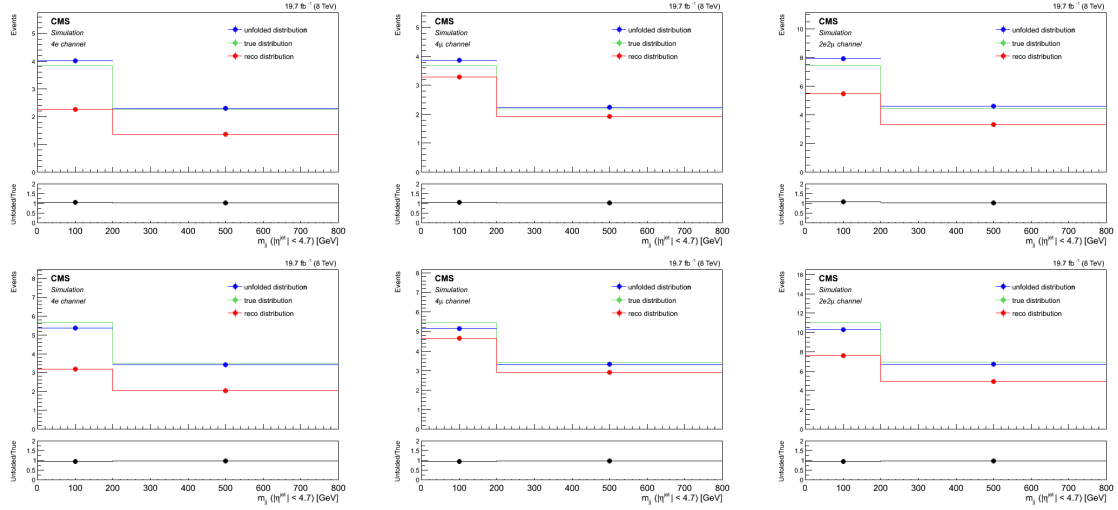


Figure A.15: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of m_{jj} for the 4e, 4μ and 2e2μ final states.

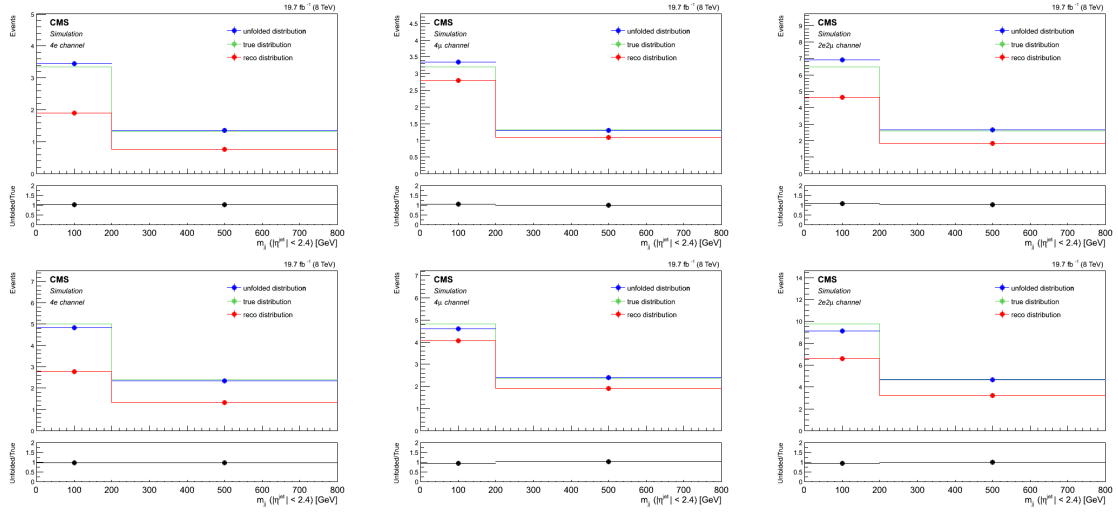


Figure A.16: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of m_{jj} (with $|\eta^{jet}| < 2.4$) for the 4e, 4μ and 2e2μ final states.

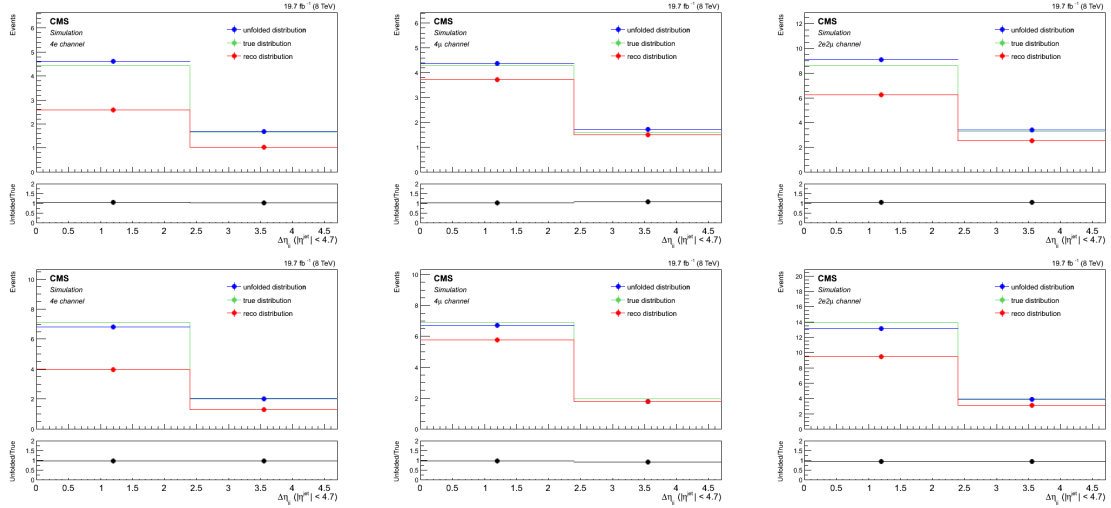


Figure A.17: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of $\Delta\eta_{jj}$ for the 4e, 4μ and 2e2μ final states.

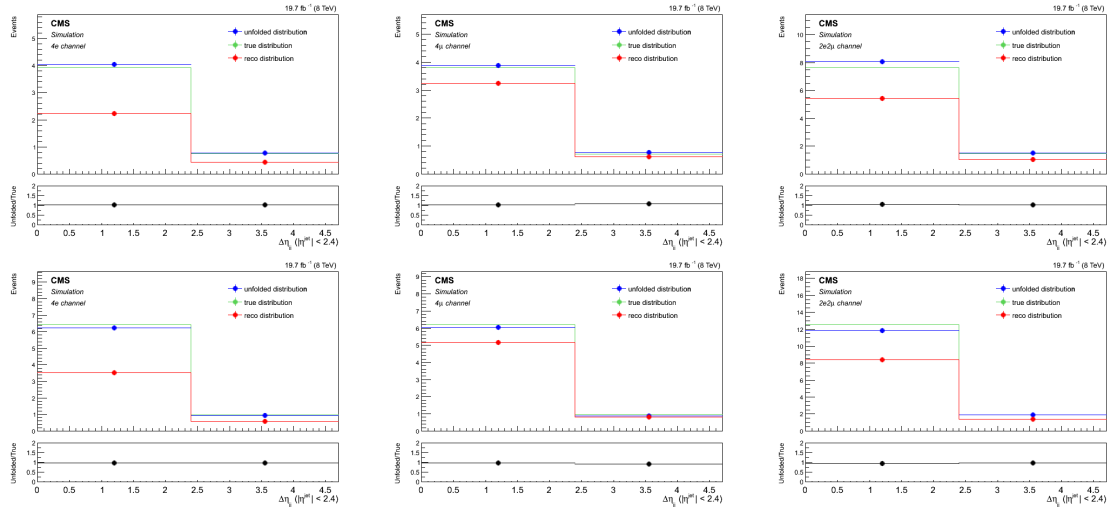


Figure A.18: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of $\Delta\eta_{jj}$ (with $|\eta^{jet}| < 2.4$) for the 4e, 4μ and 2e2μ final states.

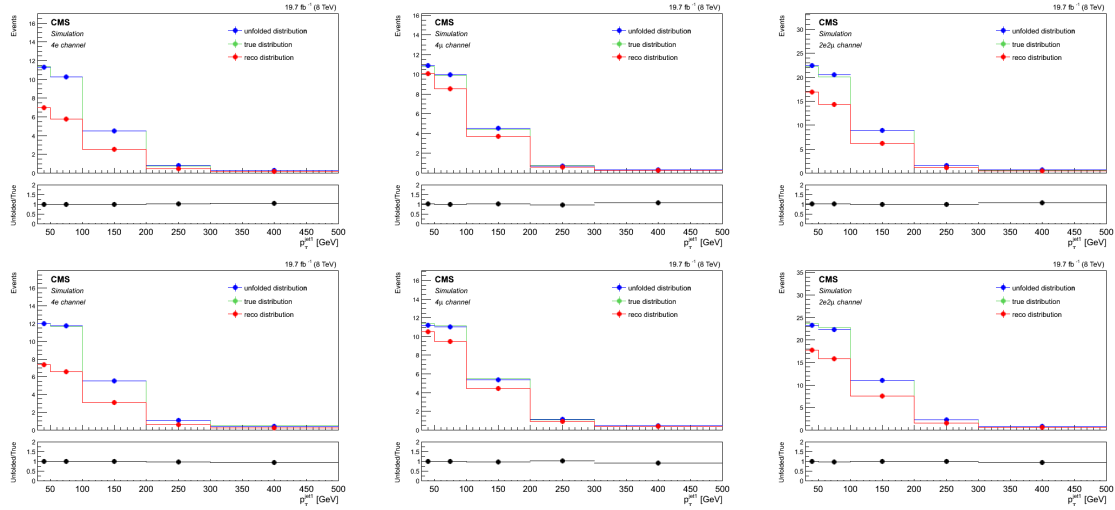


Figure A.19: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of the p_T of the leading jet, for the 4e, 4μ and 2e2μ final states.

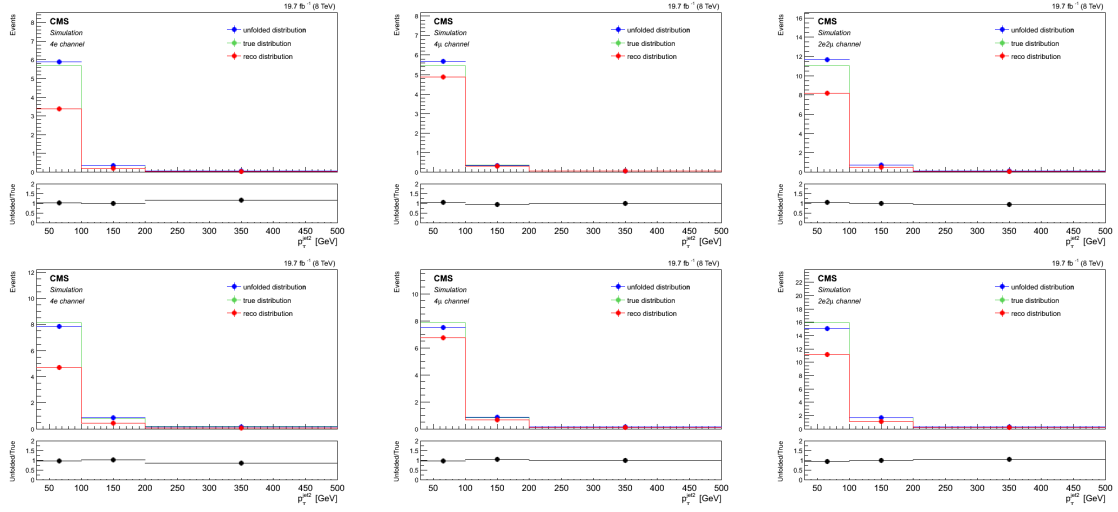


Figure A.20: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of the p_T of the sub-leading jet, for the $4e$, 4μ and $2e2\mu$ final states.

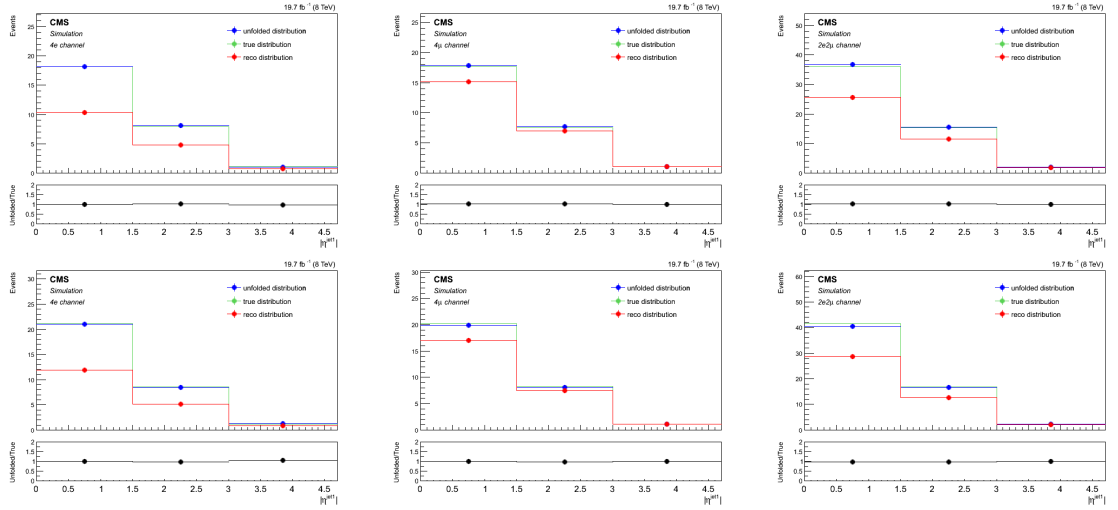


Figure A.21: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of the η of the leading jet, for the $4e$, 4μ and $2e2\mu$ final states.

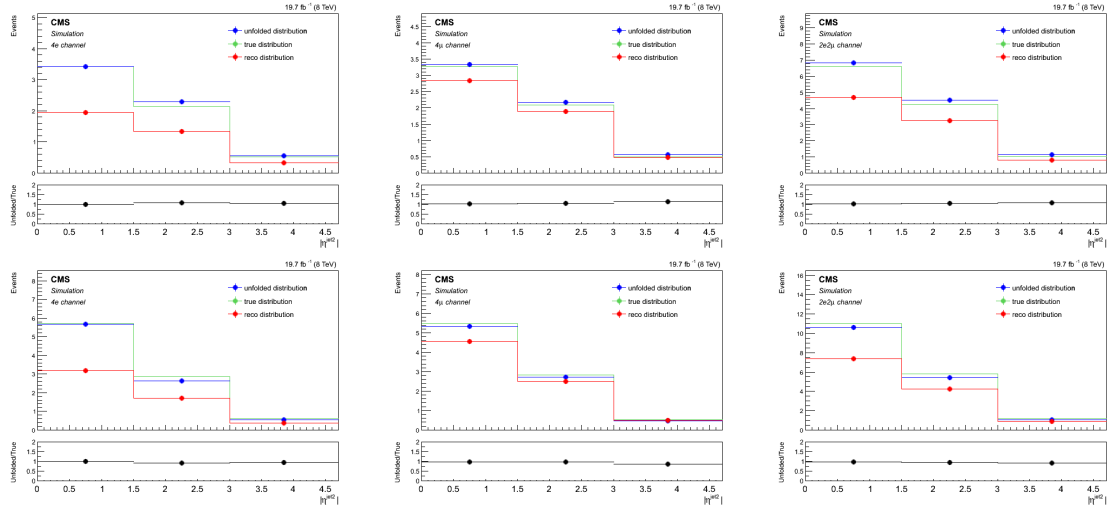


Figure A.22: Unfolding test. From top to bottom: MadGraph matrix applied on Powheg distribution using the full set, Powheg matrix applied on MadGraph distribution, using the full set, MadGraph matrix applied on MadGraph distribution, using the full set. Results are reported as a function of the η of the sub-leading jet, for the $4e$, 4μ and $2e2\mu$ final states.

Appendix B

Systematic Uncertainty Contributions

To evaluate the effects of the different sources of systematic uncertainty on the unfolded distributions, the whole unfolding procedure is repeated for every contribution. As reported in details in 6.4, the sources of systematic uncertainty considered for the final result are: σ_{qq}/σ_{gg} ratio, choice of the MC generator, irreducible and reducible backgrounds, unfolded data over MC truth ratio, lepton efficiency, jet energy resolution (JER), jet energy scale corrections on data distributions and jet energy scale corrections on MC distributions. The last three systematic uncertainties are evaluated only for jet-related variables. In the following, the contribution of each systematic uncertainty to the number-of-events distributions is reported for all the several variables analyzed in this thesis.

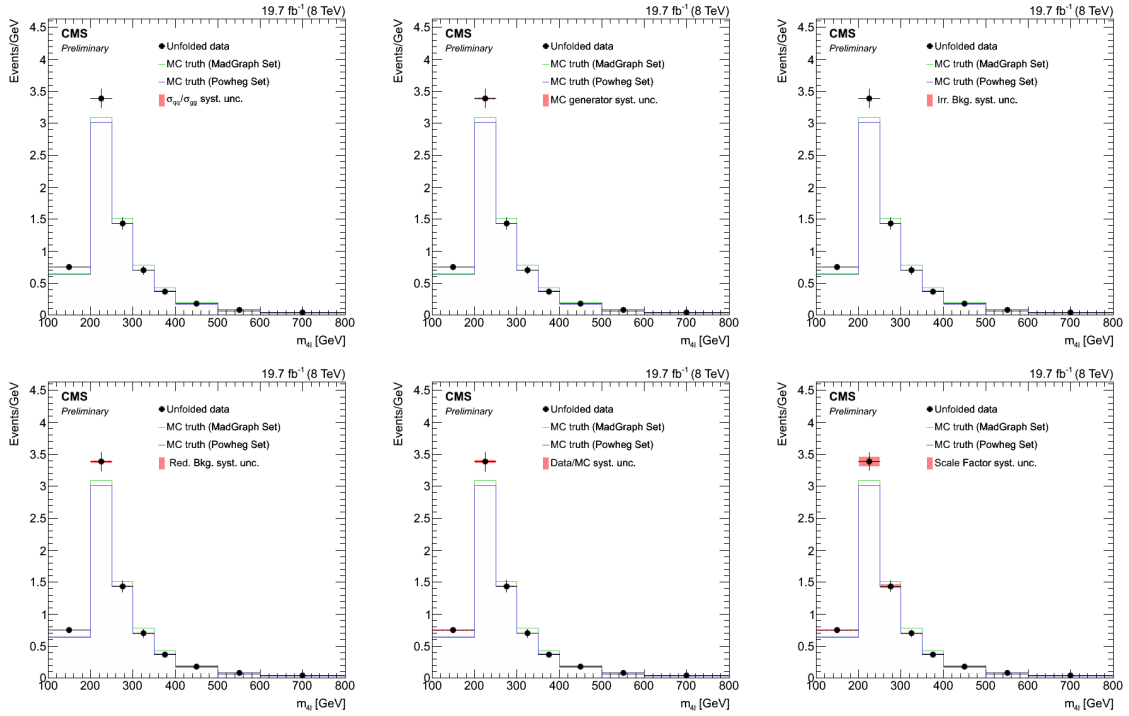


Figure B.1: Effects of the different sources of systematic uncertainty on the unfolded distributions of the mass of the 4-lepton system, for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth ratio and lepton efficiency. The systematic effect is superimposed on the nominal unfolded distribution, together with MC predictions from MadGraph and Powheg sets of samples.

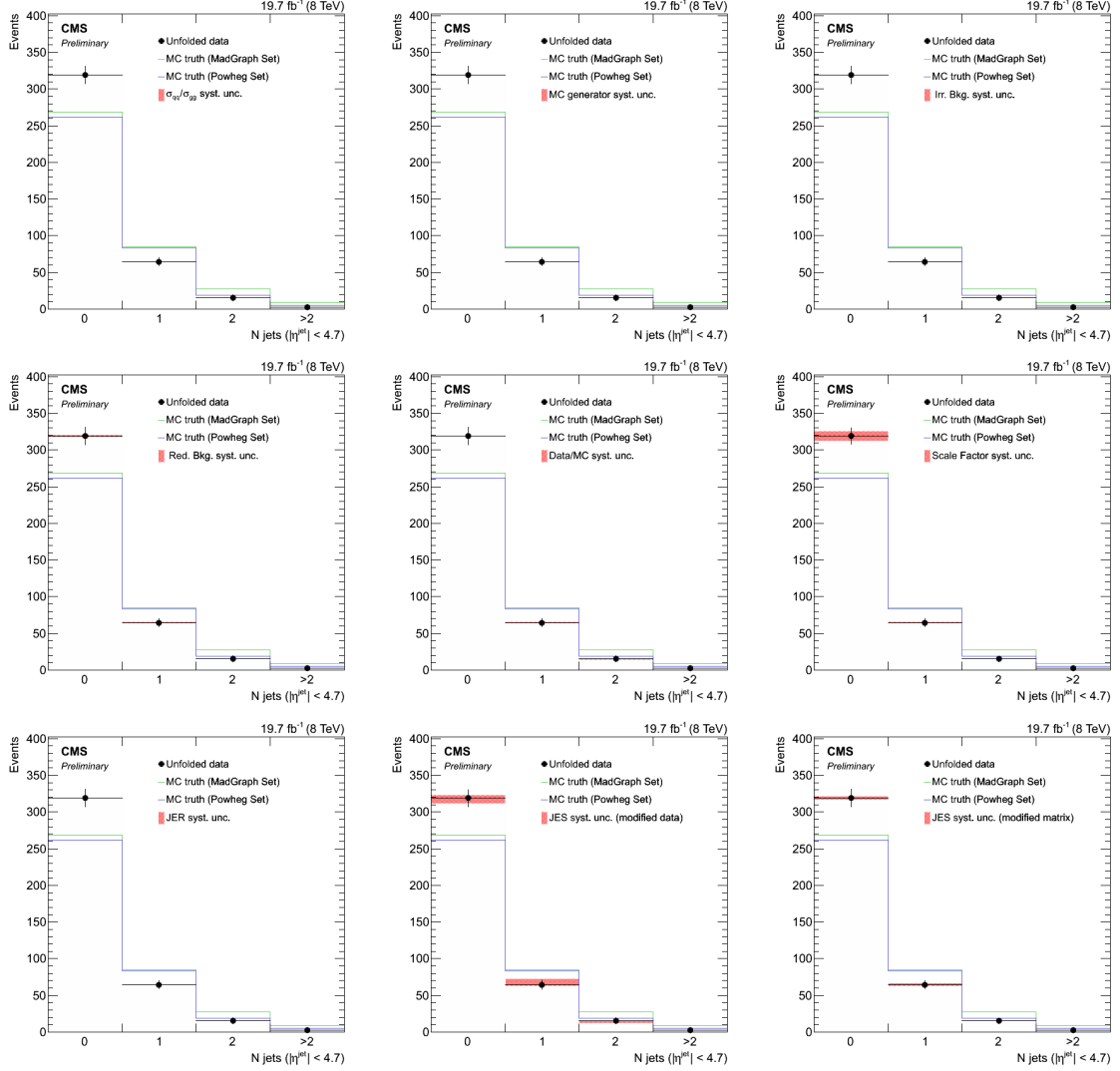


Figure B.2: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of jets (with $|\eta^{jet}| < 4.7$), for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, 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.

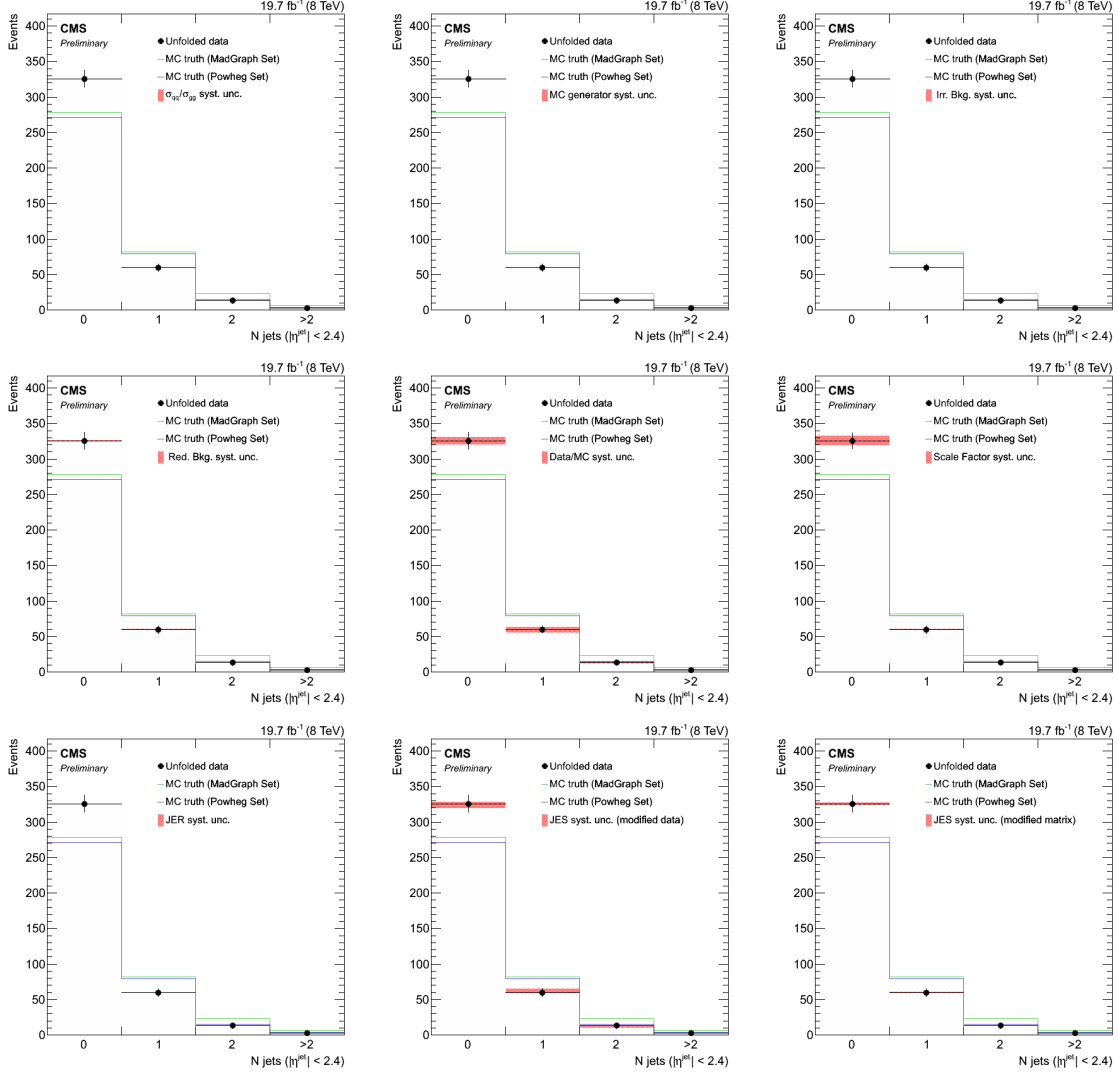


Figure B.3: Effects of the different sources of systematic uncertainty on the unfolded distributions of the number of central jets (with $|\eta^{jet}| < 2.4$), for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, 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.

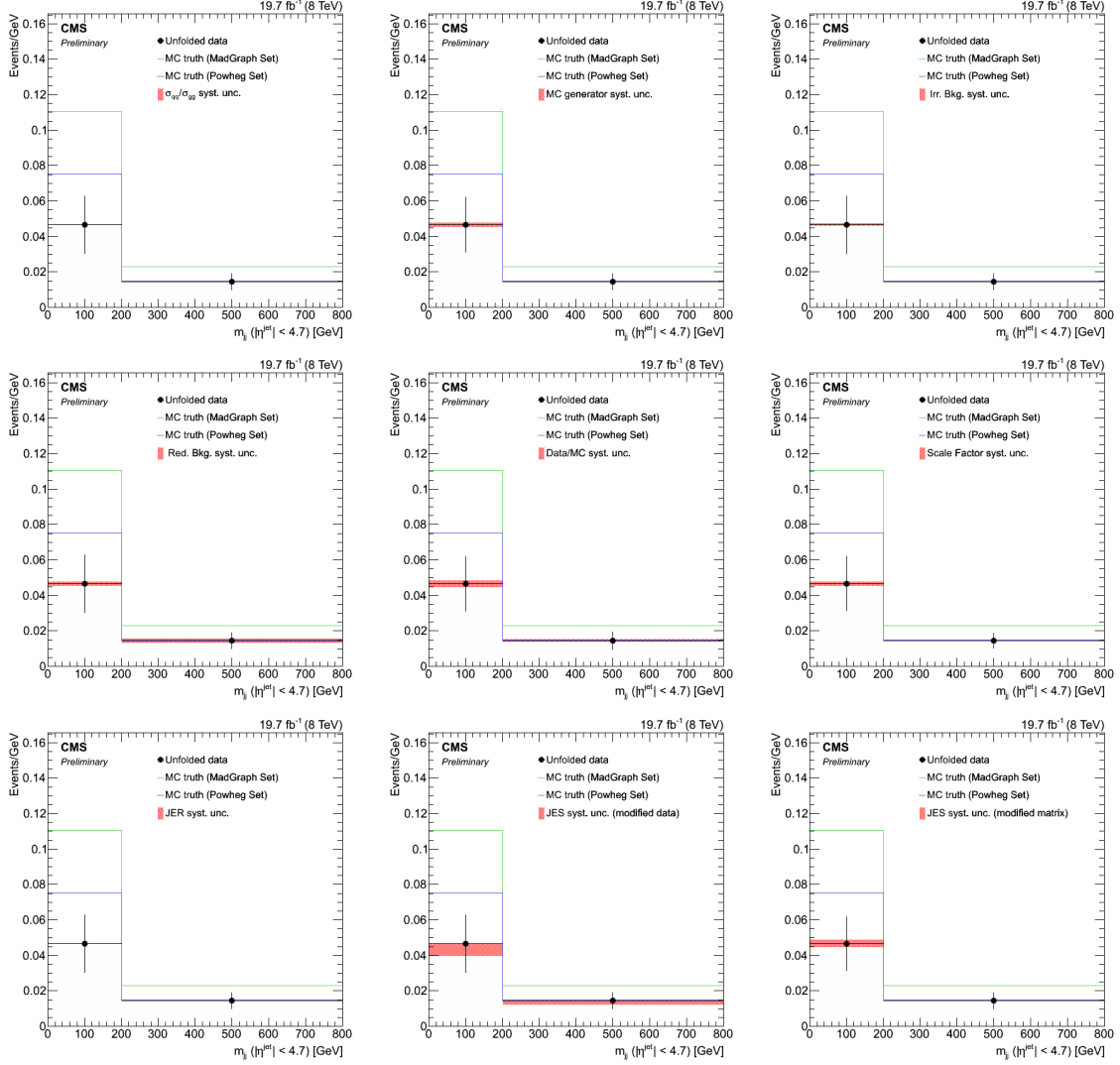


Figure B.4: Effects of the different sources of systematic uncertainty on the unfolded distributions of the invariant mass of the two p_T -leading jets (with $|\eta^{jet}| < 4.7$), for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, 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.

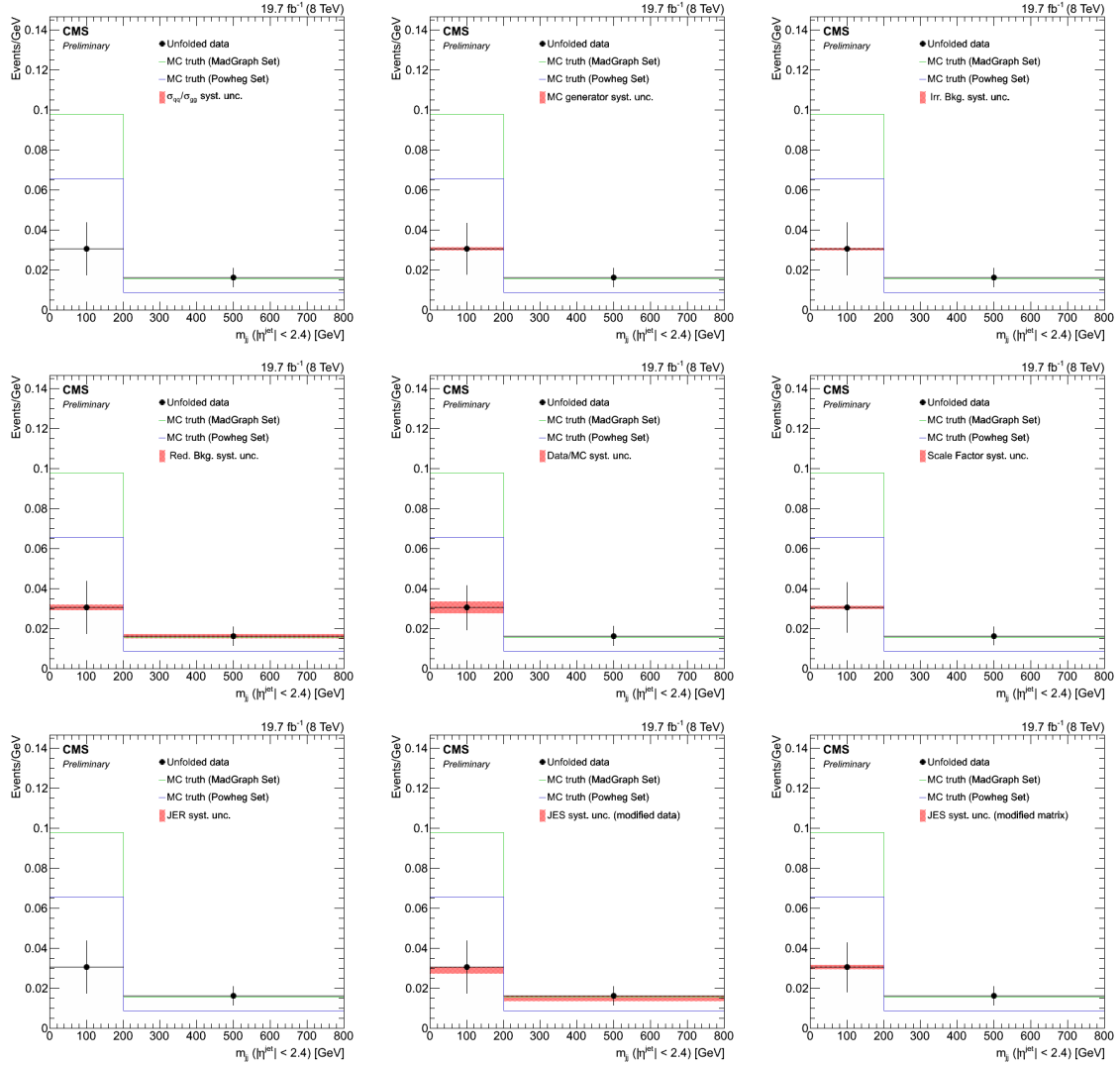


Figure B.5: Effects of the different sources of systematic uncertainty on the unfolded distributions of the invariant mass of the two p_T -leading jets (with $|\eta^{jet}| < 2.4$), for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, 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.

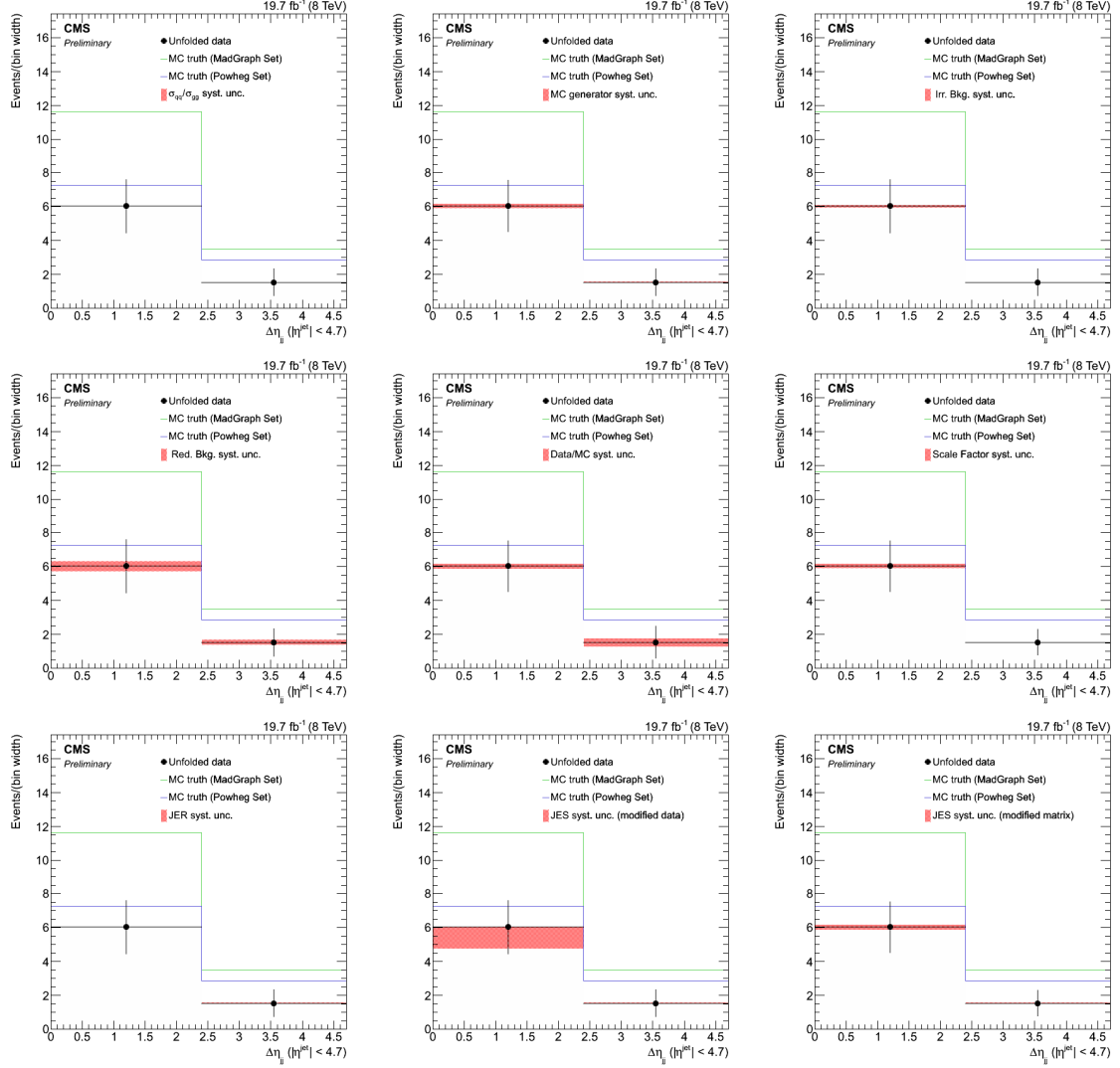


Figure B.6: Effects of the different sources of systematic uncertainty on the unfolded distributions of the $\Delta\eta_{jj}$ between the two p_T -leading jets (with $|\eta^{jet}| < 4.7$), for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible 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.

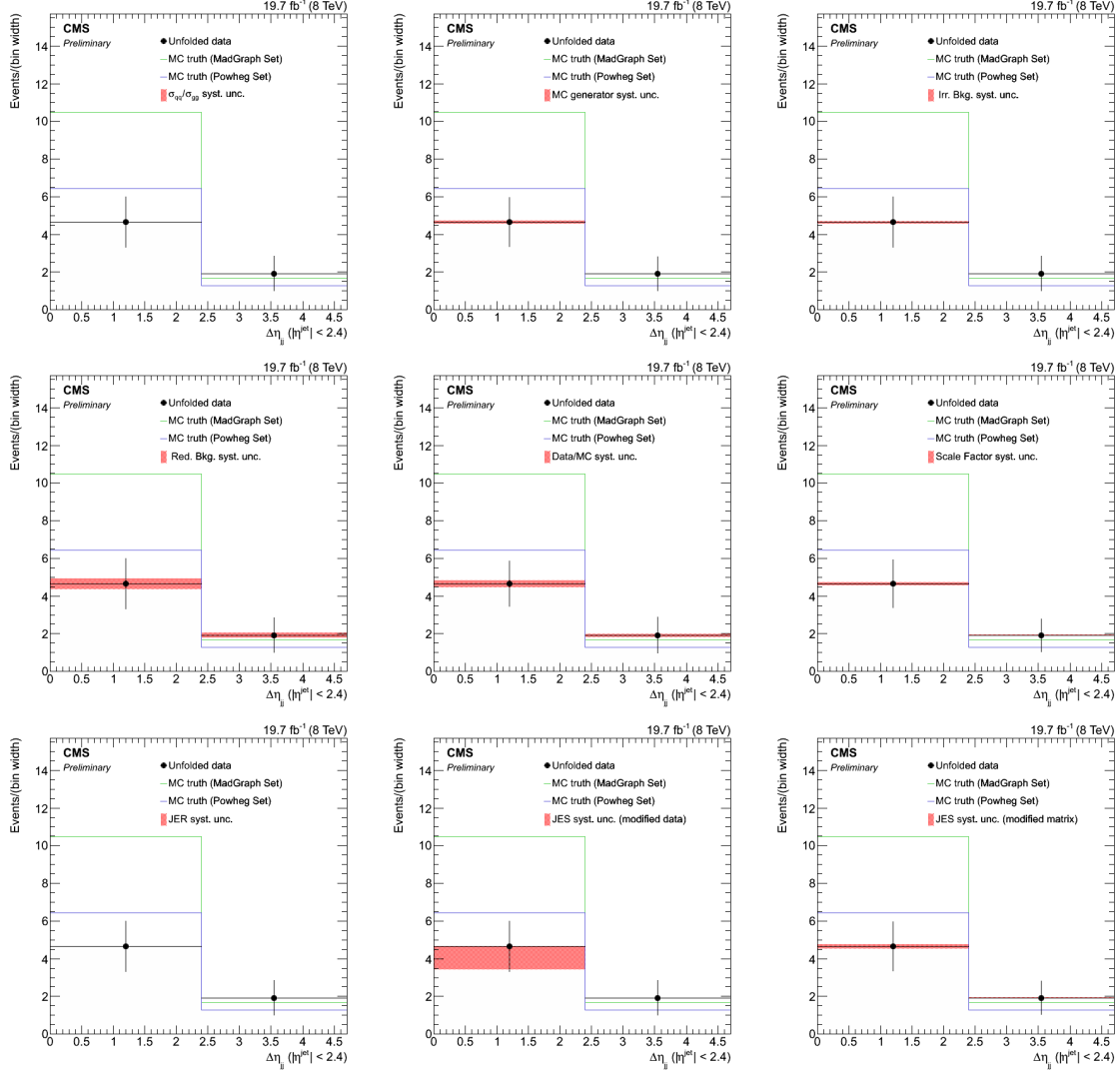


Figure B.7: Effects of the different sources of systematic uncertainty on the unfolded distributions of the $\Delta\eta_{jj}$ between the two p_T -leading jets (with $|\eta^{jet}| < 2.4$), for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible 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.

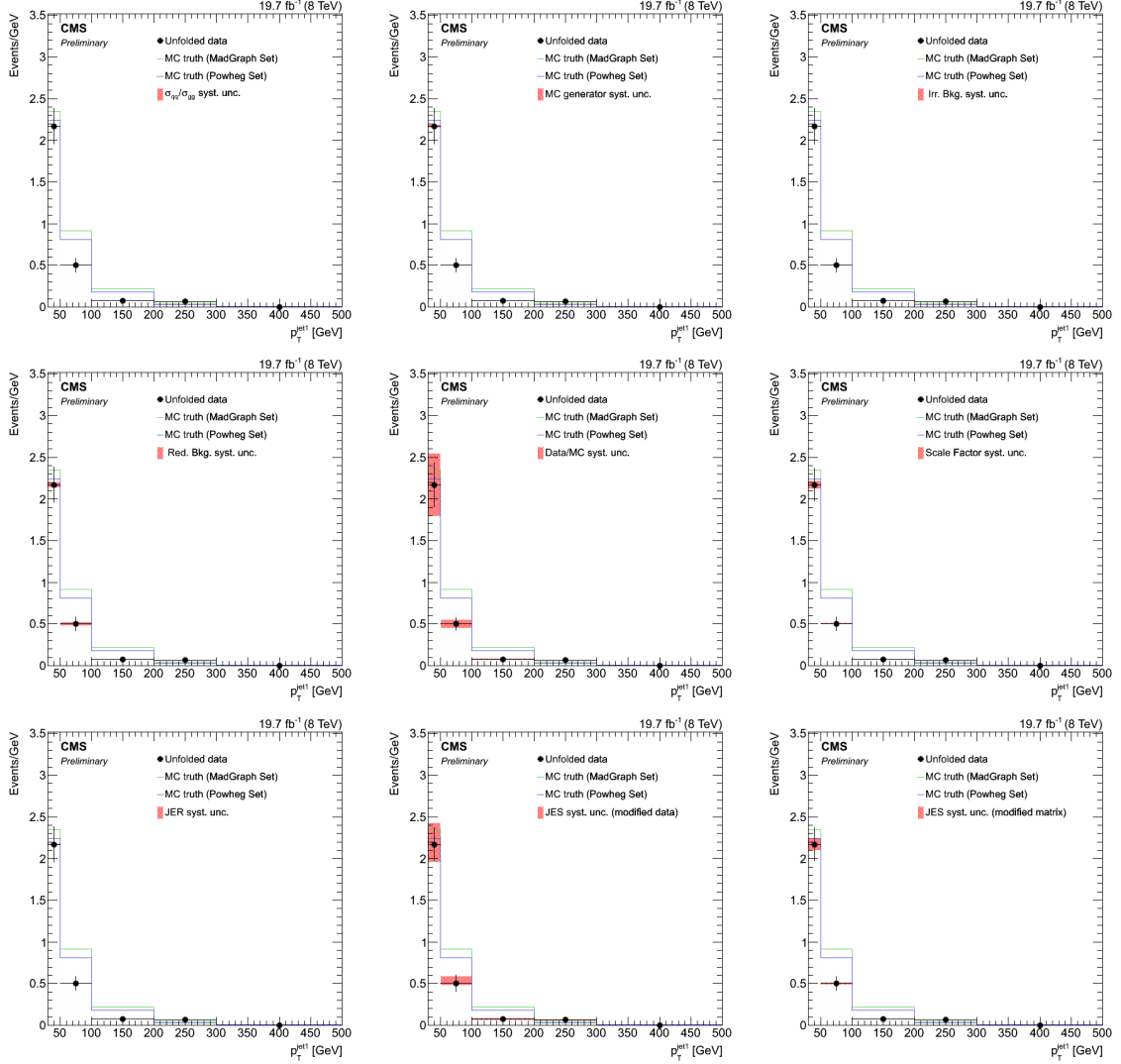


Figure B.8: Effects of the different sources of systematic uncertainty on the unfolded distributions of the leading jet transverse momentum, for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth 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.

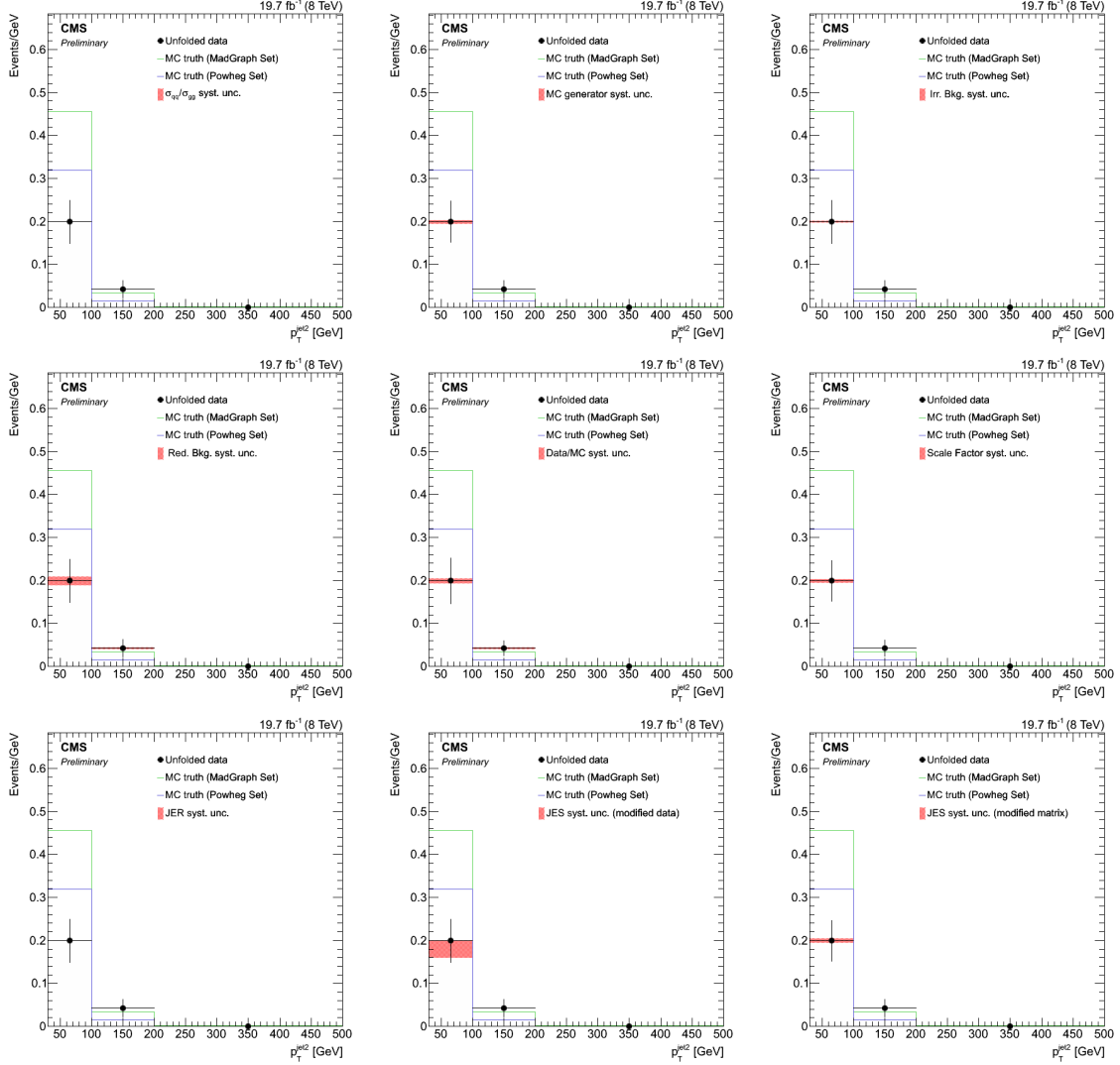


Figure B.9: Effects of the different sources of systematic uncertainty on the unfolded distributions of the sub-leading jet transverse momentum, for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth 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.

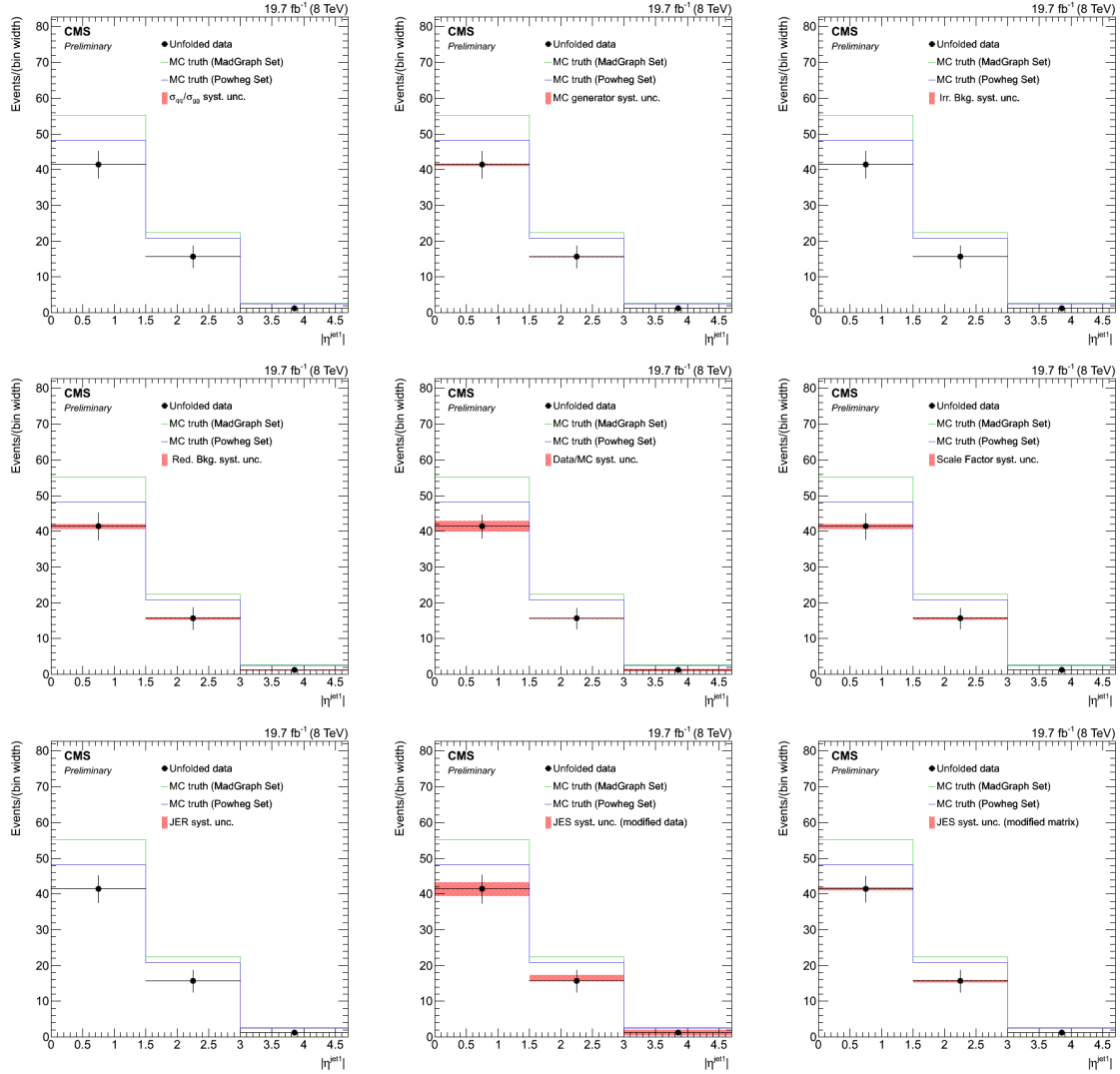


Figure B.10: Effects of the different sources of systematic uncertainty on the unfolded distributions of the leading jet pseudorapidity, for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth 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.

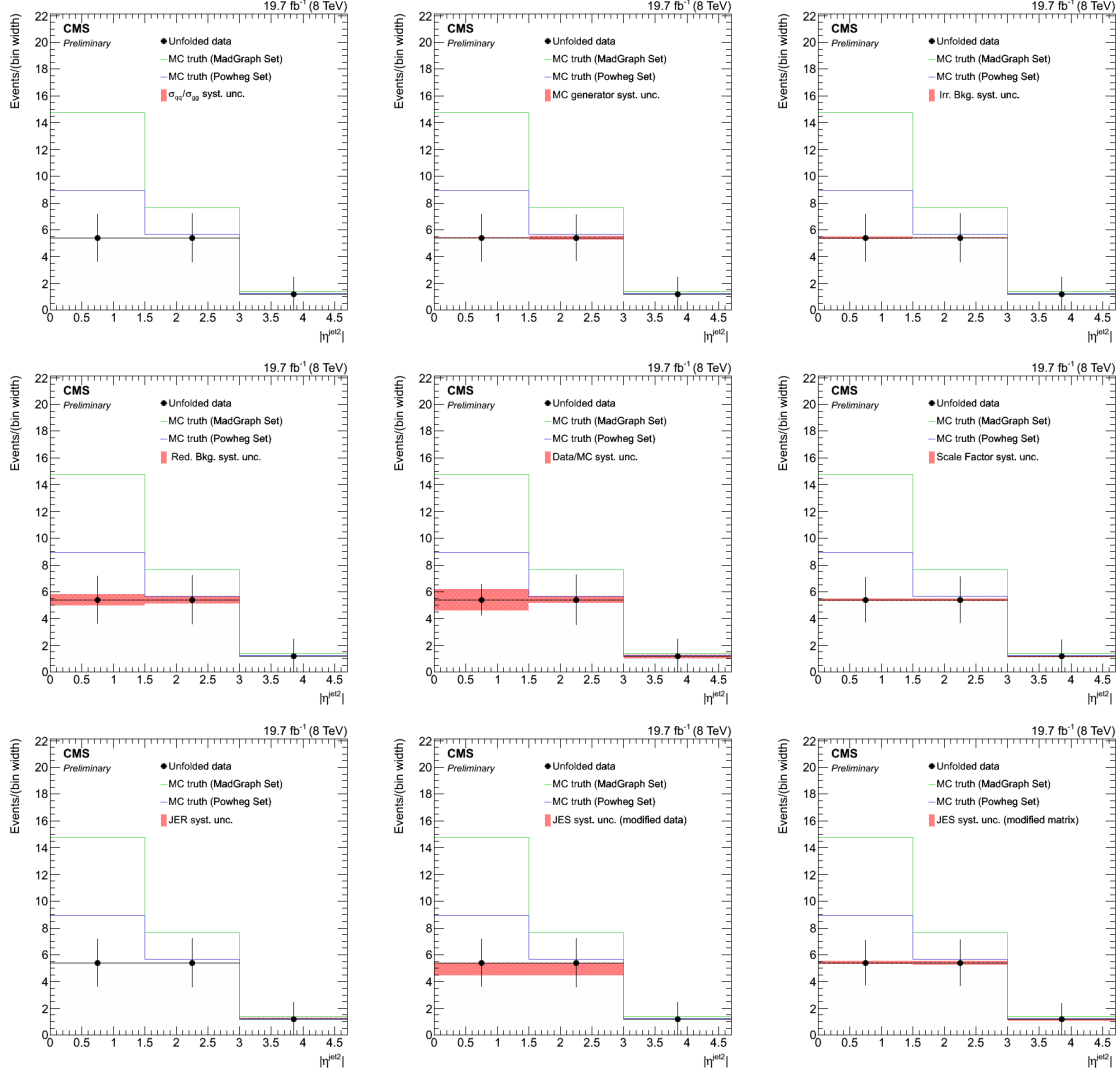


Figure B.11: Effects of the different sources of systematic uncertainty on the unfolded distributions of the sub-leading jet pseudorapidity, for the 4ℓ final state. From left to right, from top to bottom: σ_{qq} and σ_{gg} ratio, MC generator, irreducible background, reducible background, unfolded/truth 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.

Bibliography

- [1] E. Fermi. Trends to a theory of beta radiation. *Nuovo Cim.*, 11:1–19, 1934.
- [2] S. Dawson. Introduction to electroweak symmetry breaking. In *High energy physics and cosmology. Proceedings, Summer School, Trieste, Italy, June 29-July 17, 1998*, pages 1–83, 1998.
- [3] C. Degrande et al. Effective field theory: A modern approach to anomalous couplings. *Annals of Physics*, 335:21 – 32, 2013.
- [4] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN, Geneva, 2004.
- [5] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [6] <https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm>.
- [7] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008.
- [8] The CMS Collaboration. Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays. *Journal of Instrumentation*, 5(03).
- [9] The CMS Collaboration. Technical proposal for the upgrade of the CMS detector through 2020. Technical Report CERN-LHCC-2011-006, CERN, Geneva, June 2011.
- [10] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30 – 61, 2012.
- [11] The ATLAS and CMS Collaborations. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, May 2015.
- [12] The CMS Collaboration. Measurement of jet cross sections at low transverse momentum. Technical Report CMS-PAS-FSQ-12-031, CERN, Geneva, 2013.
- [13] The CMS Collaboration. Measurement of the inclusive 3-jet production differential cross section in proton–proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range. *The European Physical Journal C*, 75(5), 2015.
- [14] The CMS Collaboration. Measurement of the top-quark mass in $t\bar{t}$ events with lepton+jets final states in pp collisions at $\sqrt{s} = 8$ TeV. Technical Report CMS-PAS-TOP-14-001, CERN, Geneva, 2014.

-
- [15] The CMS and LHCb Collaborations. Observation of the rare $B_s^0 \rightarrow \mu^+ \mu^-$ decay from the combined analysis of CMS and LHCb data. *Nature*, 522, 2014.
- [16] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>.
- [17] The CMS Collaboration. Measurement of the properties of a Higgs boson in the four-lepton final state. *Phys. Rev. D*, 89:092007, 2014.
- [18] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults>.
- [19] The ATLAS Collaboration. Measurement of W^+W^- production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector and limits on anomalous WWZ and $WW\gamma$ couplings. *Phys. Rev. D*, 87:112001, Jun 2013.
- [20] The CMS Collaboration. Measurement of the W^+W^- cross section in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous $WW\gamma$ and WWZ couplings. *The European Physical Journal C*, 73(10), 2013.
- [21] The ATLAS Collaboration. Measurement of $W^\pm Z$ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *The European Physical Journal C*, 72(10), 2012.
- [22] Measurement of WZ production rate. Technical Report CMS-PAS-SMP-12-006, CERN, Geneva, 2013.
- [23] The ATLAS Collaboration. Measurement of the $WW + WZ$ cross section and limits on anomalous triple gauge couplings using final states with one lepton, missing transverse momentum, and two jets with the ATLAS detector at $\sqrt{s} = 7$ TeV. *Journal of High Energy Physics*, 2015(1), 2015.
- [24] The CMS Collaboration. Measurement of the sum of WW and WZ production with $W +$ dijet events in pp collisions at $\sqrt{s} = 7$ TeV. *The European Physical Journal C*, 73(2), 2013.
- [25] The ATLAS Collaboration. Measurement of ZZ production in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous ZZZ and $ZZ\gamma$ couplings with the ATLAS detector. *Journal of High Energy Physics*, 2013(3), 2013.
- [26] The CMS Collaboration. Measurement of the ZZ production cross section and search for anomalous couplings in $2\ell 2\ell'$ final states in pp collisions at $\sqrt{s} = 7$ TeV. *Journal of High Energy Physics*, 2013(1), 2013.
- [27] The CMS Collaboration. Measurements of the ZZ production cross sections in the $2\ell 2\nu$ channel in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV and combined constraints on triple gauge couplings. *The European Physical Journal C*, 75(10), 2015.
- [28] Measurement of the W^+W^- production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2014-033, CERN, Geneva, 2013.
- [29] The CMS Collaboration. Measurement of the W^+W^- cross section in pp collisions at $\sqrt{s} = 8$ TeV and limits on anomalous gauge couplings. 2015.

- [30] A Measurement of WZ Production in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2013-021, CERN, Geneva, Mar 2013.
- [31] The CMS Collaboration. Measurement of WZ and ZZ production in pp collisions at $\sqrt{s} = 8$ TeV in final states with b -tagged jets. *The European Physical Journal C*, 74(8), 2014.
- [32] Measurement of the total ZZ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV in 20 fb^{-1} with the ATLAS detector. Technical Report ATLAS-CONF-2013-020, CERN, Geneva, 2013.
- [33] The CMS Collaboration. Measurement of the production cross section and constraints on anomalous triple gauge couplings in four-lepton final states at $\sqrt{s} = 8$ TeV. *Physics Letters B*, 740:250–272, 2015.
- [34] The CMS Collaboration. Measurement of the ZZ production cross section and search for anomalous couplings in $2\ell 2\ell'$ final states in pp collisions at $\sqrt{s} = 7$ TeV. *Journal of High Energy Physics*, 2013(1), 2013.
- [35] The CDF Collaboration. Measurement of ZZ production in leptonic final states at $\sqrt{s}$ of 1.96 TeV at CDF. *Phys. Rev. Lett.*, 108:101801, Mar 2012.
- [36] The Do Collaboration. Measurement of the ZZ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 84, July 2011.
- [37] John M. Campbell and R.K. Ellis. MCFM for the Tevatron and the LHC. *Nuclear Physics B - Proceedings Supplements*, 205-206:10 – 15, 2010.
- [38] John Campbell. Next-to-Leading Order Tools for Colliders. 2004.
- [39] G. 't Hooft and M. Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44(1):189 – 213, 1972.
- [40] J. Ohnemus and J. F. Owens. Order- α_s calculation of hadronic ZZ production. *Phys. Rev. D*, 43:3626–3639, Jun 1991.
- [41] B. Mele, P. Nason, and G. Ridolfi. QCD radiative corrections to Z boson pair production in hadronic collisions. *Nuclear Physics B*, 357(2-3):409 – 438, 1991.
- [42] J. Ohnemus. Order- α_s calculation of hadronic W^+W^- production. *Phys. Rev. D*, 44:1403–1414, Sep 1991.
- [43] Stefano Frixione. A next-to-leading order calculation of the cross section for the production of WW pairs in hadronic collisions. *Nuclear Physics B*, 410(2):280 – 324, 1993.
- [44] J. Ohnemus. Order- α_s calculation of hadronic W^+Z production. *Phys. Rev. D*, 44:3477–3489, Dec 1991.
- [45] J. Ohnemus. Order- α_s calculations of hadronic $W^\pm\gamma$ and $Z\gamma$ production. *Phys. Rev. D*, 47:940–955, Feb 1993.
- [46] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *Journal of High Energy Physics*, 2011(7), 2011.

- [47] Anastasiya Bierweiler, Tobias Kasprzik, and JohannH. Kühn. Vector-boson pair production at the LHC to $\mathcal{O}(\alpha^3)$ accuracy. *Journal of High Energy Physics*, 2013(12), 2013.
- [48] Julien Baglio, Le Duc Ninh, and Marcus M. Weber. Massive gauge boson pair production at the LHC: A next-to-leading order story. *Phys. Rev. D*, 88:113005, Dec 2013.
- [49] S. Dittmaier, S. Kallweit, and P. Uwer. Next-to-leading-order QCD corrections to $WW + \text{jet}$ production at hadron colliders. *Phys. Rev. Lett.*, 100:062003, Feb 2008.
- [50] John M. Campbell, R. Keith Ellis, and Giulia Zanderighi. Next-to-leading order predictions for $WW + \text{jet}$ distributions at the LHC. *Journal of High Energy Physics*, 2007(12):056, 2007.
- [51] S. Dittmaier, S. Kallweit, and P. Uwer. NLO QCD corrections to including leptonic W-boson decays. *Nuclear Physics B*, 826(1–2):18 – 70, 2010.
- [52] T. Binoth, T. Gleisberg, S. Karg, N. Kauer, and G. Sanguinetti. NLO QCD corrections to production at hadron colliders. *Physics Letters B*, 683(2–3):154 – 159, 2010.
- [53] M. Grazzini. Vector-boson pair production at NNLO. *PoS*, LL2014:027, 2014.
- [54] Stefano Catani and Massimiliano Grazzini. Next-to-Next-to-Leading-Order Subtraction Formalism in Hadron Collisions and its Application to Higgs-Boson Production at the Large Hadron Collider. *Phys. Rev. Lett.*, 98:222002, May 2007.
- [55] T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, and L. Tancredi. W^+W^- Production at Hadron Colliders in Next to Next to Leading Order QCD. *Phys. Rev. Lett.*, 113(21):212001, 2014.
- [56] F. Cascioli, T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, L. Tancredi, and E. Weihs. ZZ production at hadron colliders in NNLO QCD. *Physics Letters B*, 735:311 – 313, 2014.
- [57] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, 2009.
- [58] Tracking and Primary Vertex Results in First 7 TeV Collisions. Technical Report CMS-PAS-TRK-10-005, CERN, 2010. Geneva, 2010.
- [59] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois. Electron reconstruction in CMS. *Eur. Phys. J.*, C49:1099–1116, 2007.
- [60] The CMS Collaboration. Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV. *Journal of High Energy Physics*, 2011(1), 2011.
- [61] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/EGMElectronsMoriond2013>.
- [62] Muon ID performance: low- p_T muon efficiencies. Technical Report CMS DP - 2014/020, CERN, Geneva, 2014.
- [63] G.P. Salam M. Cacciari and G. Soyez. The anti- k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.
- [64] Jet Energy Resolution: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution>.

-
- [65] G.P. Salam M. Cacciari and G. Soyez. Fastjet user manual. *The European Physical Journal C*, 72(3), 2012.
- [66] Jet Identification: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID>.
- [67] CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.
- [68] Rene Brun, Federico Carminati, and Simone Giani. GEANT Detector Description and Simulation Tool. 1994.
- [69] J. Alwall et al. Madgraph 5: going beyond. *Journal of High Energy Physics*, 2011, June 2011.
- [70] S. Frixione et al. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007(11):070, 2007.
- [71] S. Alioli et al. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6), 2010.
- [72] A. Ballestrero et al. PHANTOM: A Monte Carlo event generator for six parton final states at high energy colliders. *Computer Physics Communications*, 180(3):401–417, 2009.
- [73] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.
- [74] L Lai et al. Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions. *Physical Review D*, 82, 2010.
- [75] L Lai et al. New parton distributions for collider physics. *Physical Review D*, 82, 2010.
- [76] K.A. Olive and Particle Data Group. Review of Particle Physics. *Chinese Physics C*, 38(9):090001, 2014.
- [77] M. Botje et al. The PDF4LHC Working Group Interim Recommendations. 2011.
- [78] S. Alekhin et al. The PDF4LHC Working Group Interim Report. 2011.
- [79] R. D. Ball Ket al. Impact of heavy quark masses on parton distributions and LHC phenomenology. *Nuclear Physics B*, 849(2):296 – 363, 2011.
- [80] T. Adye. Unfolding algorithms and tests using RooUnfold. 2011.
- [81] A. Höcker and V. Kartvelishvili. SVD approach to data unfolding. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 372(3):469–481, 1996.
- [82] G. D’Agostini. A multidimensional unfolding method based on Bayes’ theorem. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 362(2-3):487–498, 1995.
- [83] Measurement of the W^+W^- cross section in pp collisions at $\sqrt{s} = 8$ TeV and limits on anomalous gauge couplings. Technical report, CERN, 2015.

- [84] Jet Energy Scale: <https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources>.
- [85] Daniel de Florian, Giancarlo Ferrera, Massimiliano Grazzini, and Damiano Tommasini. Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC. *Journal of High Energy Physics*, 2011(11), 2011.
- [86] The CMS Collaboration. Measurement of the W^+W^- cross section in pp collisions at $\sqrt{s} = 8$ TeV and limits on anomalous gauge couplings". *submitted to EPJC*, (CMS-SMP-14-016, CERN-PH-EP-2015-122), 2015.
- [87] D. Rathlev M. Wiesemann M. Grazzini, S. Kallweit. Transverse-momentum resummation for vector-boson pair production at NNLL+NNLO. (ZU-TH 20/15, MITP/15-051), 2015.
- [88] G. Aad et al. Evidence for electroweak production of $W^\pm W^\pm jj$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev. Lett.*, 113:141803, Oct 2014.
- [89] Robert N. Cahn, Stephen D. Ellis, Ronald Kleiss, and W. J. Stirling. Transverse-momentum signatures for heavy Higgs bosons. *Phys. Rev. D*, 35:1626–1632, Mar 1987.
- [90] Elena Accomando, Alessandro Ballestrero, Aissa Belhouari, and Ezio Maina. Isolating vector boson scattering at the CERN LHC: Gauge cancellations and the equivalent vector boson approximation versus complete calculations. *Phys. Rev. D*, 74:073010, Oct 2006.
- [91] The ATLAS Collaboration. Measurement of the electroweak production of dijets in association with a Z-boson and distributions sensitive to vector boson fusion in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Journal of High Energy Physics*, 2014(4), 2014.
- [92] The CMS Collaboration. Measurement of electroweak production of two jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$ TeV. *The European Physical Journal C*, 75(2), 2015.
- [93] Measurement of the cross section of the electroweak production of a W boson with two jets in pp collisions at $\sqrt{s} = 8$ TeV. Technical report, CERN.
- [94] The CMS Collaboration. Study of exclusive two-photon production of W^+W^- in pp collisions at $\sqrt{s} = 7$ TeV and constraints on anomalous quartic gauge couplings. *Journal of High Energy Physics*, 2013(7), 2013.
- [95] Evidence for exclusive gamma-gamma to W^+W^- production and constraints on Anomalous Quartic Gauge Couplings at $\sqrt{s} = 8$ TeV. Technical Report CMS-PAS-FSQ-13-008, CERN, Geneva, 2015.
- [96] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, Apr 1998.
- [97] G. Aad et al. Evidence for electroweak production of $W^\pm W^\pm jj$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev. Lett.*, 113:141803, Oct 2014.
- [98] V. Khachatryan et al. Study of vector boson scattering and search for new physics in events with two same-sign leptons and two jets. *Phys. Rev. Lett.*, 114:051801, Feb 2015.

-
- [99] Studies of Vector Boson Scattering And Triboson Production with an Upgraded ATLAS Detector at a High-Luminosity LHC. Technical Report ATLAS-PHYS-PUB-2013-006, CERN, Geneva, 2013.
 - [100] Vector Boson Scattering and Quartic Gauge Coupling Studies in WZ Production at 14 TeV. Technical Report CMS-PAS-FTR-13-006, CERN, Geneva, 2013.

