UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXIII CICLO DEL DOTTORATO DI RICERCA IN FISICA

Measurement of the Zγγ bosons production cross section and search for new physics in proton proton collisions at $\sqrt{s}$ = 13 TeV with the CMS detector at the CERN LHC

Settore scientifico-disciplinare: **FIS/01 FISICA SPERIMENTALE**

DOTTORANDO / A
FEDERICO VAZZOLER

COORDINATORE
PROF. FRANCESCO LONGO

SUPERVISORE DI TESI
PROF. GIUSEPPE DELLA RICCA

ANNO ACCADEMICO 2019/2020

University of Trieste
Department of Physics

Graduate program in Physics

Measurement of the Zγγ bosons production cross section and search for new physics in proton proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the CMS detector at the CERN LHC

XXXIII Cycle
SSD: FIS/01

Author:

Federico Vazzoler

ID number: SM2300236

Supervisor:

Prof. Giuseppe Della Ricca

Coordinator:

Prof. Francesco Longo

Academic Year 2019/2020

Universita degli studi di Trieste
Dipartimento di Fisica

Programma di dottorato di Ricerca in Fisica

Misura della sezione d'urto di produzione e ricerca di nuova fisica in eventi $Z\gamma\gamma$ prodotti da collisioni protone protone a $\sqrt{s} = 13\,\text{TeV}$ misurate dal detector CMS presso l'acceleratore LHC

XXXIII Ciclo
SSD: FIS/01

Autore:

Federico Vazzoler

Matricola: SM2300236

Supervisore:

Prof. Giuseppe Della Ricca

Coordinatore:

Prof. Francesco Longo

Anno accademico 2019/2020

Questa tesi è dedicata a chi mi è stato vicino in questi anni.
A mamma e papà.
A Caterina, ai miei amici ed ai miei parenti, vicini e lontani.
Specialmente alla nonna Gigia ed alla nonna Nori.
Ancor più specialmente al nonno Giordano ed al nonno Bepi.
È incredibile come, pur attraverso gli anni, abbiamo fatto esperienze così simili.
Ad maiora.

Dicembre 2020

Ricordati dunque che, se credi che le cose che sono per natura in uno stato di schiavitù siano libere e che le cose che ti sono estranee siano tue, sarai ostacolato nell'agire, ti troverai in uno stato di tristezza e di inquietudine, e rimprovererai dio e gli uomini. Se al contrario pensi che sia tuo solo ciò che è tuo, e che ciò che ti è estraneo - come in effetti è - ti sia estraneo, nessuno potrà più esercitare alcuna costrizione su di te, nessuno potrà più ostacolarti, non muoverai più rimproveri a nessuno, non accuserai più nessuno, non farai più nulla contro la tua volontà, nessuno ti danneggerà, non avrai più nemici, perchè non subirai più alcun danno.

Epitteto, Manuale 3 p.143

Introduction

The Standard Model of particle physics is one of the most predictive and precise theories conceived by mankind. Its predictions have been proven, along decades of several tests, to accurately describe the vast majority of the experimental observations. The recent discovery of a boson with mass at about 125 GeV, showing properties compatible with the Standard Model predicted Higgs boson, had finally provided the last missing piece of the model itself.

Despite its outstanding success, there is a growing set of phenomena for which the descriptions of the Standard Model are insufficient or even missing. To overcome these shortcomings, "new physics" models have been developed over the years, whose validity can only be tested through experiments. A general way to parametrise new physics phenomena is represented by the Effective Field Theory framework, which extends the Standard Model by considering new interactions between its constituent particles. A specific class of these interactions, labelled as anomalous Quartic Gauge Couplings, is probed in this work. Another goal of this analysis is to perform high-precision studies of the Standard Model itself, by investigating the validity of its predictions in the multiple gauge vector boson production channel.

Processes involving the interaction of multiple gauge vector bosons are very rare and neutral coupling between the Z boson and one or more photons are forbidden, at tree level, in the Standard Model. Possible anomalous properties of the Z boson have been studied, before the CERN Large Hadron Collider era, at the Large Electron-Positron collider and the Tevatron, while triboson processes become accessible at the Large Hadron Collider only recently. Both the CMS and ATLAS Collaborations have studied the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ channel, at a centre of mass energy of 7 and 8 TeV, by exploiting the Run 1 available luminosity of about $20\,fb^{-1}$. The measured cross sections were found compatible with the Standard Model prediction and limits on anomalous Quartic Gauge Couplings were placed.

This work represents the first attempt to measure the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ cross section at a centre of mass energy of 13 TeV by making use of about $137\,fb^{-1}$ of luminosity collected by the Compact Muon Solenoid experiment at the Large Hadron Collider. The measurement presented in this analysis makes it possible to probe the electroweak sector of the Standard Model in a very high energy regime and also benefits from an increased statistics, up to 10 times the one accessible before.

Observed events as a function of the diphoton transverse momentum $p_T^{\gamma\gamma}$ variable are used to measure the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section and to extract limits on the anomalous Quartic Gauge Couplings. The signal region of this analysis is defined by a pair of same flavour opposite sign leptons (electrons or muons) produced

alongside two isolated photons.

The measured $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section is found to be consistent with the Standard Model predictions and results are also interpreted in terms of physics beyond the Standard Model. No significant excess over the expected numbers of events is observed and exclusion limits on various anomalous Quartic Gauge Couplings parameters are set.

This work is organised as follows. Chapter 1 is devoted to a brief theoretical introduction to the Standard Model of particle physics. The main theoretical topics of interests for this analysis are covered, focusing on the electroweak theory and on the Effective Field Theory framework used to describe new physics phenomena.

The physics basis of Standard Model perturbative calculations, the main features and the use of general purpose Monte Carlo programs for the simulation of the signal and background processes, expected to contribute to the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section, are presented in Chapter 2.

The phenomenology of the $pp \to Z\gamma\gamma \to l^+l^-$ production at hadron colliders, the state of the art theoretical calculation of its cross sections and the anomalous Quartic Gauge Couplings theory are reviewed in Chapter 3.

Chapter 4 gives a brief overview of the Large Hadron Collider, the world largest and most powerful accelerator complex ever built, with a particular focus on the Compact Muon Solenoid experiment, one of the main experiments along the accelerator ring.

The sub-detectors of the Compact Muon Solenoid experiment are sensitive to different stable particles, such as electrons, muons, photons, neutral and charged hadrons. A combination of several kinematic variables is used to select, with the highest efficiency and purity, the building blocks of the $l^+l^-\gamma\gamma$ events. Chapter 5 is devoted to the description of the algorithms used for the offline particle reconstruction and identification, together with the object selection criteria exploited in this work.

The core of the analysis is presented in the Chapter 6. The data and Monte Carlo samples used for signal and background contributions modelling are here defined, together with the object based corrections. The main background source comes from $Z + \text{jets}$ and $Z\gamma + \text{jets}$ events where hadronic jets are misidentified as prompt photons. A data-driven approach, used to estimate this background contribution, is presented and discussed in detail. Other small residual background contributions are estimated from Monte Carlo simulations. The systematic uncertainties related to the estimate of the background contributions are described as well.

The measurement of the fiducial cross section in the electron, muon and combined lepton channels is presented in Chapter 7. A discussion of the statistical methods employed to obtain the final results, together with a comparison with the theoretical prediction, is given.

In Chapter 8 the $l^+l^-\gamma\gamma$ events are used to investigate the presence of new physics. For these studies, events are selected in the same phase space defined for the fiducial cross section measurement. No evidence of new physics phenomena is obtained and limits on the anomalous Quartic Gauge Couplings are set.
The work presented in this thesis is currently under internal review by the CMS Collaboration. The author of this thesis has provided fundamental contributions in all the major aspect of the analysis.

Contents

1 | Introduction to high energy physics

Elementary particles are the smallest constituents of all the ordinary matter. Nevertheless, the notion of "elementary" is subject to change with the pursuit of the scientific progress, as can be seen in Table 1.1. Particles that were considered as "elementary" in the past are now known to be composite states of smaller constituents, and nothing prevents our current knowledge in this field to be expanded or completely revamped in the future.

Nowadays, the framework used to catalogue all the elementary particles is represented by the Standard Model (SM). It can describe, with great precision, the interactions between particles via the fundamental forces of nature. The SM predictive power, enforced by an incredible number of experimental confirmations, makes it one of the most successful scientific theories developed by mankind.

This chapter is organised as follows. In Section 1.1 the SM of particle physics is presented while in Section 1.2 special attention is given to the gauge symmetries of the model and their implications in terms of particle interactions. The electroweak and quantum chromodynamics theories are introduced in Section 1.2.2 and Section 1.2.3, respectively. Finally, in Section 1.3 a summary of the experimental tests of the SM done so far is presented, outlining its current limitations and revising a possible framework to parametrise new physics processes beyond it.

This chapter follows the notation of reference [1] and it is further based on reference [2, 3].

1.1 The Standard Model of particle physics

The goal of fundamental high energy physics is to reduce all natural phenomena to a (small) set of common rules able to reproduce and predict the experimental observations. At microscopic level, forces between elementary particles are parametrised through three classes of fundamental interactions, the *electromagnetic*, the *weak* and the *strong* one. For all microscopic objects a fourth interaction, the *gravitational* force, plays a role but appears to be negligible for sub-nuclear physics. The SM of particle physics reflects our current knowledge of elementary constituents of matter at scales of about 10^{-15} m (1 fm) and below [4, 5]. Particle interactions are described within the framework of quantum mechanics and relativity, more precisely by a local relativistic quantum field theory (QFT). Each particle is considered to be point-like and associated with a specific field which undergoes suitable transformations under the Lorentz group.

Year	Discovery	Elementary particle
1869	Period system	Atom
1890	Electron	Atom, electron
1910	Radioactivity Alpha scattering	Atomic nucleus, electron
1932	Neutron Positron	Proton, neutron, electron ... and their antiparticles
1947	Pion, muon	The "particle zoo"
1956	Neutrino	
1967	Electroweak theory	
1968	Deep inelastic scattering	Quarks, leptons
1972	Quantum chromodynamics	
1974	c quark, τ lepton	
1977	b quark	
1979	Gluon	
1983	W, Z bosons	
1995	t quark	
2012	Higgs boson	Standard Model

Table 1.1: An historical outline of the keystone discoveries in high-energy physics. The concept of elementary changes through time as new particles/interactions are introduced in the model. With the discovery of the Higgs boson the SM can be considered as "complete".

According to the Heisenberg uncertainty principle, to study particles with a spatial dimension Δx one needs, as a suitable probe, a beam of particles with impulse $p \gtrsim \hbar/\Delta x$, where $\hbar$ is the reduced Plank constant ($\hbar = h/2\pi$). It comes straightforward that the smallest the spatial dimension Δx under study the greater the energy of the beam of particles is required. To probe the SM predictions at very small distances, high energy beams of particles are typically produced at the particle accelerators. An accelerator such as the CERN Large Hadron Collider (LHC) allows to study collisions between particles with a total centre of mass energy up to $2pc \lesssim 13\,\text{TeV}$, giving the possibility to probe particle interactions at a distance scale of $\Delta x \gtrsim 10^{-20}\,\text{m}$ (see also Section 4.2). From the experimental results obtained so far all the fundamental constituents of the SM do not show any appreciable internal structure.

The SM predicts the existence of two kinds of elementary particles, spin-$\frac{1}{2}$ *fermions* which obey the Fermi-Dirac statistics and the Pauli exclusion principle, and integer-spin *bosons* subjected to the Bose-Einstein statistics. In Figure 1.1 (left) a table of the elementary matter particle constituents of the SM is presented. All the known fundamental matter constituents are represented by twelve elementary fermions. There exist six types of *leptons*, which are not affected by the strong force and are seen in experiments as individual particles. Three of them have an electric charge and are labelled as the electron (e), the muon (μ) and the tau (τ) lepton. The other three leptons are the neutrinos, electrically neutral and extremely weakly interacting particles. They are labelled as the electron (ν_e), the muon (ν_μ) and the tau (ν_τ) neutrino. Similarly to the leptons, there exist six types of *quarks* all of them affected by the electromagnetic,

the weak and the strong interactions. Quarks are hidden from view but represent the fundamental constituents of several non-fundamental particles, such as protons and neutrons. Three of them have a fractional electric charge of 2/3 and are labelled as the up (u), charm (c) and top (t) quark, while the other three have a fractional electric charge of $-1/3$ and are labelled as the down (d), strange (s) and bottom (b) quark.

Quarks and leptons are arranged into three *families* with equal quantum numbers but different masses. The description of matter through different families evolved naturally from the Cabibbo's quark mixing theory into the Cabibbo-Kobayashi-Maskawa matrix [6, 7].

A fundamental prediction of the SM is that quarks can never appear singly in nature. Indeed there is no experimental evidence of isolated quarks so far. Nevertheless, there exists a large number of quark composite states, collectively called *hadrons*, which are arranged in different families such as *mesons*, bound states of two quarks ($q\bar{q}'$) and *baryons*, formed by three quarks ($qq'q''$).

All the matter particles listed above are accompanied by the respective anti-particles with opposite quantum numbers and the same couplings of their counterparts.

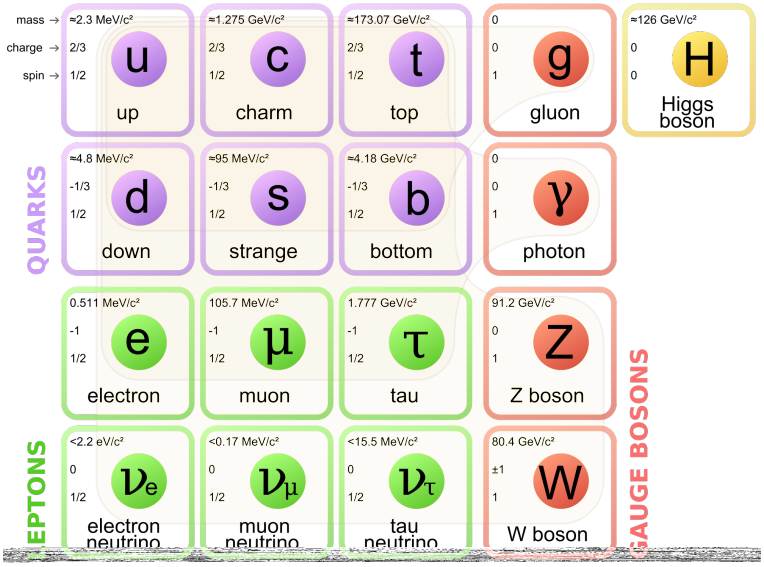


Figure 1.1: The Standard Model elementary particles content. Matter particles are indicated in the first three columns, gauge bosons in the fourth and the Higgs boson in the fifth. Each fermion has an antiparticle partner, with the same mass but different quantum numbers. In the top left corner the mass, charge and spin of the particle are reported. Figure adapted from reference [8].

The interactions between fermions are described in relativistic QFT via the exchange of additional particles labelled as *force carriers*. In the SM, the force carriers obey the Bose-Einstein statistics and are represented by twelve integer-spin bosons, as can be seen in Figure 1.1 (right). The photon (γ), the carrier of the electromagnetic force, is a massless spin-1 boson while the weak bosons (W and Z), are the massive mediators of the weak force. Eight gluons (g) are the massless carriers of the strong interaction. In addition to these spin-1 bosons there is the Higgs boson, a neutral massive scalar

has spin 1 and is massless (providing the symmetry is unbroken). These vector bosons, called *gauge bosons*, are the mediators of the interactions.

The principle of non-Abelian gauge invariance, which dictates the structure of the interactions between fermions and vector bosons as well as the vector boson self-interactions, is the generalisation of the Abelian gauge symmetry found in QED. It is implemented in the SM through the Yang-Mills theory [9]. The starting point to construct a non-Abelian gauge invariant theory is the general Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$, which is assumed to be invariant under a D dimensional continuous group of transformations

$$\phi'(x) = U(\theta^A)\phi(x) \quad (A = 1, 2, \ldots D),$$

where

$$U(\theta^A) = e^{ig\sum_A \theta^A T^A} \sim 1 + ig\sum_A \theta^A T^A + \ldots.$$

The symbols θ^A are numerical parameters of the transformation, g is called the *coupling constant* and T^A are the generators of the group. The T^A satisfy the commutation relation

$$[T^A, T^B] = iC_{ABC}T^C,$$

where the group structure constants C_{ABC} are non-vanishing for a non-Abelian theory. If the gauge transformation is local, so that it depends on the space-time coordinates x_μ via $U(\theta^A) \to U(\theta^A(x_\mu))$, the Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$ will be no longer invariant under the transformation (since in general $\partial(U\phi) \neq U(\partial\phi)$). Gauge invariance can then be restored if the ordinary derivative ∂_μ is replaced by the covariant derivative

$$D_\mu = \partial_\mu + ig\mathbf{V}_\mu, \tag{1.1}$$

where $\mathbf{V}_\mu = \sum_A V_\mu^A T^A$ and V_μ^A are a set of D gauge vector fields, in one-to-one correspondence with the group generators, which transform according to $\mathbf{V}'_\mu = U\mathbf{V}_\mu U^{-1} - \frac{1}{ig}(\partial_\mu U)U^{-1}$. The gauge invariance of $\mathcal{L}(\phi, D_\mu\phi)$ is re-established at the expense of introducing additional external fields $\mathbf{V}_\mu$, which were not present from the beginning in the theory. In analogy with QED, the gauge-invariant kinetic term for the gauge fields V_μ^A takes the form $F^{\mu\nu,\ A}F^A_{\mu\nu}$ and the tensor $F^A_{\mu\nu}$ is defined as

$$F^A_{\mu\nu} = \partial_\mu V^A_\nu - \partial_\nu V^A_\mu - gC_{ABC}V^B_\mu V^C_\nu. \tag{1.2}$$

The complete Yang-Mills Lagrangian can be finally written as

$$\mathcal{L}_{\mathrm{YM}} = -\frac{1}{4}\sum_A F^A_{\mu\nu}F^{A\mu\nu} + \mathcal{L}(\phi, D\phi), \tag{1.3}$$

where it has to be noticed that the kinetic energy term is an operator of dimension $d = 4$, and no mass term for the gauge boson of the form $m^2 V_\mu V^\mu$ is allowed by gauge invariance.

The gauge symmetry group $SU(3)\times SU(2)\times U(1)$ is not the only symmetry embedded in the SM theory. The invariance of the theory under space-time translations, rotations and boosts (Poincare symmetry) leads to conservations of energy, momentum and angular momentum. The discrete symmetries parity (P), where an inversion of coordinates takes

place, time-reversal (T), where $t \to -t$ and charge conjugation (C), that transforms a particle into the corresponding antiparticle, are involved differently. While strong and electromagnetic forces conserve each symmetry individually, the electroweak symmetry group $SU(2) \times U(1)$ breaks both the P and C symmetries by acting differently on the left and right-handed fields. The T and the combined CP symmetries are broken also by the Yukawa couplings between the fermions and the Higgs boson. The combination CPT is instead conserved in every known interaction, ensuring that the mass and lifetimes of all particles is equal to that of the corresponding antiparticles. The SM Lagrangian is also accidentally invariant under a global common phase transformation of all lepton fields. The associated charges are the baryon number $B = \pm 1/3$ for quarks and antiquarks, and the lepton number $L = \pm 1$ for leptons and anti-leptons.

1.2.2 Electroweak theory

The EW theory [10–12] is the unified description of the QED and weak nuclear forces. The unification arises at a fixed energy scale, the *electroweak scale*, of around 100 GeV. The EW theory describes the interactions between matter particles, leptons and quarks, and weak gauge bosons ($W^{\pm}$, Z) or photons (γ), as dictated by the $SU(2) \times U(1)$ gauge symmetry. In the following the EW Lagrangian will be split in two parts,

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{Higgs}}, \tag{1.4}$$

by separating the term containing the Higgs field, which will be described in Section 1.2.2.1, from the one involving only gauge bosons and fermions.

The gauge Lagrangian $\mathcal{L}_{\text{gauge}}$ can be written, according to the prescription of Equation 1.3, as

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}\sum_{A=1}^{3} F^{A}_{\mu\nu}F^{A\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \bar{\psi}_L i\gamma^{\mu}D_{\mu}\psi_L + \bar{\psi}_R i\gamma^{\mu}D_{\mu}\psi_R, \tag{1.5}$$

where the gauge antisymmetric tensors $B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}$ and $F^{A}_{\mu\nu} = \partial_{\mu}W^{A}_{\nu} - \partial_{\nu}W^{A}_{\mu} - g\epsilon_{ABC}W^{B}_{\mu}W^{C}_{\nu}$ are built from the gauge field B_{μ} associated with the $U(1)$ symmetry generator, and W^{A}_{μ} which correspond to the three $SU(2)$ symmetry generators. The action of covariant derivative D_{μ} on the fermion fields $\psi_{L,R}$ is explicitly given by

$$D_{\mu}\psi_{L,R} = \left[\partial_{\mu} + ig\sum_{A=1}^{3} t^{A}_{L,R}W^{A}_{\mu} + ig'\frac{1}{2}Y_{L,R}B_{\mu}\right]\psi_{L,R}, \tag{1.6}$$

where the values g and g' are the coupling constants of the three gauge fields to the fermion fields and t^{A} correspond to the Pauli matrices. It has to be noticed that the standard EW theory is "chiral", in the sense that the $SU(2) \times U(1)$ symmetry transformation acts differently on left-handed (ψ_L) and right-handed (ψ_R) fermion fields. In the approximation of massless fermions, chirality reduces to the helicity which corresponds to the projection of the particle spin on the particle momentum vector. The left and right-handed fermion fields are obtained from the general Dirac field ψ by defining the projectors

$$\psi_L = \frac{1-\gamma_5}{2}\psi \text{ and } \psi_R = \frac{1+\gamma_5}{2}\psi.$$

The $SU(2)$ part of the electroweak symmetry group is referred to as the weak isospin group, with associated quantum numbers I and I_3. Left-handed fields have $I = 1/2$ and form doublets which are subject to the weak isospin transformation $U = \exp(i\alpha^a\sigma^a/2)$ (as an example, the electron doublet is $(\nu_\mathrm{e}, \mathrm{e})_L$), while right-handed fields have $I = 0$ and form singlets invariant under the weak isospin transformation (such as $\mathrm{e}_R, \mathrm{u}_R, \mathrm{d}_R$). For each doublet or singlet, the $U(1)$ part corresponds to a multiplication by a phase factor $\exp(i\alpha Y/2)$ and the quantum number associated to the transformation Y is the weak hyper-charge. The electromagnetic gauge group appears as a subgroup of the electroweak symmetry group obtained by combining a hyper-charge transformation with a particular isospin transformation. The electric charge of a particle is indeed given by the Gell-Mann-Nishijima formula as

$$Q = I_3 + \frac{Y}{2}. \tag{1.7}$$

The assignment of quantum numbers to the lepton and quark fields is given in Table 1.2.

	$\nu_{\mathrm{e}\,L}$	e_L	e_R	u_L	d_L	u_R	d_R
	$\nu_{\mu\,L}$	μ_L	μ_R	c_L	s_L	c_R	s_R
	$\nu_{\tau\,L}$	τ_L	τ_R	t_L	b_L	t_R	b_R
I_3	$+\frac{1}{2}$	$-\frac{1}{2}$	0	$+\frac{1}{2}$	$-\frac{1}{2}$	0	0
Y	-1	-1	-2	$+\frac{1}{3}$	$+\frac{1}{3}$	$+\frac{4}{3}$	$-\frac{2}{3}$
Q	0	-1	-1	$+\frac{2}{3}$	$-\frac{1}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$

Table 1.2: The weak isospin and hyper-charge of the left-handed and right-handed leptons and quarks groups. The rows correspond to the different generations.

The mass eigenstates of the charged and neutral boson fields are defined as

$$\mathrm{W}^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \text{ and } \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix},$$

where θ_w is the electroweak mixing angle. The photon field A_μ couples identically to the ψ_L and ψ_R fields, with a strength equal to the electric charge $e = g\sin\theta_w = g'\cos\theta_w$. The $\mathrm{W}^\pm$ bosons couple to the weak isospin and therefore only to left-handed fermions. Finally the Z boson couples to both left-handed and right-handed fermions.

1.2.2.1 Spontaneous symmetry breaking and the Higgs field

A gauge invariant way to recover the fermions and bosons masses proceeds via a spontaneously breaking of the $SU(2) \times U(1)$ symmetry. In the SM the electroweak symmetry breaking (EWSB) is achieved through the Brout-Englert-Higgs mechanism [13]. In 2012, the CMS and ATLAS Collaborations jointly announced the discovery of a new particle with a mass of about 125 GeV and Higgs-boson-like properties [14, 15].

The Higgs Lagrangian is defined, according to the $SU(2) \times U(1)$ gauge symmetry and the requirement of renormalizability, as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi) - \mathcal{L}_{\text{Yukawa}} \tag{1.8}$$

where ϕ is a complex scalar field with $Y = 1$ and weak isospin $T = 1/2$

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{1.9}$$

The potential $V(\phi^\dagger \phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2$ is referred as the Higgs potential. As can be seen in Figure 1.2, if the mass μ^2 and the self-coupling λ terms are positive then the potential $V(\phi^\dagger \phi)$ is also positive and its minimum will be located at $\langle 0|\phi|0\rangle = 0$. Conversely, if μ^2 in negative but $\lambda > 0$ is positive the potential has a whole orbit of

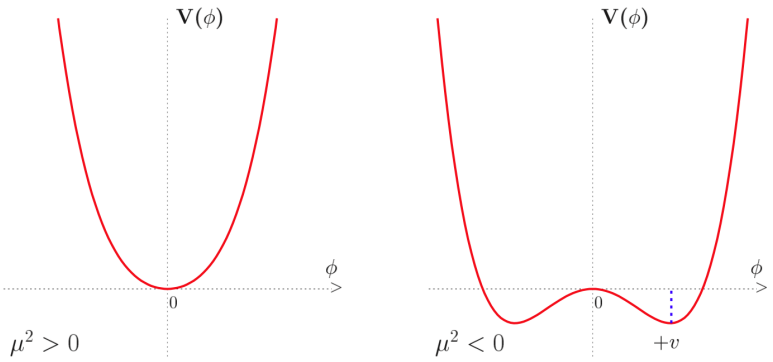


Figure 1.2: The potential of a scalar field ϕ, when $\lambda > 0$, in the case $\mu^2 > 0$ (left) or $\mu^2 < 0$ (right). Figure adapted from reference [16].

minima located at a radius v, the vacuum expectation value (vev) of the scalar field ϕ, given by

$$\phi^\dagger \phi = v^2 = -\frac{\mu^2}{\lambda}. \tag{1.10}$$

By choosing the following vacuum configuration

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.11}$$

the $SU(2) \times U(1)$ symmetry is spontaneously broken such that the EW Lagrangian is still invariant under the symmetry transformation but the potential expanded around the minimum is not.

The coupling of the Higgs boson to the EW gauge bosons is responsible for their masses and is given by the $(D_\mu \phi)^\dagger (D^\mu \phi)$ terms of Equation 1.8

$$D_\mu \phi = \left[\partial_\mu + ig \sum_{A=1}^{3} t^A W_\mu^A + ig' \frac{Y}{2} B_\mu \right] \phi. \tag{1.12}$$

By inserting Equation 1.11 into Equation 1.12 the following mass terms are obtained

$$M_{\mathrm{W}} = \frac{1}{2}vg, \quad M_{\mathrm{Z}} = \frac{1}{2}v\sqrt{g^2 + g'^2}, \quad M_{\gamma} = 0.$$

The mass of the W and Z bosons are well determined experimentally and their most updated values correspond to $M_{\mathrm{W}} = 80.379(12)\,\mathrm{GeV}$ [17] and $M_{\mathrm{Z}} = 91.1876(21)\,\mathrm{GeV}$ [18]. It has to be noticed that, as a consequence of Equation 1.7, the scalar field does not carry electric charge thus the photon remains massless. There is no experimental evidence, so far, for the photon to be massive.

The coupling of the Higgs boson to the fermions, described in Equation 1.8 by the $\mathcal{L}_{\mathrm{Yukawa}}$ term, is responsible for their masses

$$\mathcal{L}_{\mathrm{Yukawa}} = \bar{\psi}_L \Gamma \psi_T \phi - \bar{\psi}_R \Gamma^{\dagger} \psi_L \phi^{\dagger}. \tag{1.13}$$

In Equation 1.13 the quantities Γ are matrices that make the Yukawa couplings invariant under the Lorentz and gauge groups. By inserting Equation 1.11 into Equation 1.13 the mass term M is obtained as $M = \bar{\psi}_L \mathcal{M} \psi_R + \bar{\psi}_R \mathcal{M}^{\dagger} \psi_L$, where $\mathcal{M} = \Lambda \cdot v$.

Finally, the self-coupling of the Higgs boson is responsible for its mass. This can be seen by expanding the scalar field ϕ around the vacuum

$$\phi(x) = \begin{pmatrix} \phi^{+}(x) \\ \phi^{0}(x) \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ v + (H/\sqrt{2}) \end{pmatrix}. \tag{1.14}$$

By inserting the Higgs potential of Equation 1.8 into Equation 1.14, given the minimum condition of Equation 1.10, one obtains $M_{\mathrm{H}} = \sqrt{2\lambda v^2}$.

1.2.3 Quantum chromodynamics

The theory describing the interactions between quarks and gluons was developed since the 1950s and ended up with the QCD, the gauge theory of strong interactions. As for the EW case, the QCD Lagrangian is derived from the Yang-Mills theory by introducing a $SU(3)$ symmetry and an associated charge called *colour*. The choice of $SU(3)$ as the gauge symmetry is the result of several considerations. Only quark bound states of the form $\mathrm{q}\overline{\mathrm{q}}'$ but not qq' are observed in nature thus the group should be able to distinguish a quark from an antiquark, by allowing quark complex representations. No free colour charge has been observed so far, that is the group must include a completely antisymmetric colour singlet made up of three quarks $\mathrm{qq}'\mathrm{q}''$. Finally, the choice of $N_C = 3$ colour charges is well confirmed from experiments such as the measurement of the total rate for the hadronic production in $\mathrm{e}^{+}\mathrm{e}^{-}$ annihilations.

The QCD Lagrangian can be written as

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4}\sum_{A=1}^{8} F^{A\mu\nu} F^{A}_{\mu\nu} + \sum_{j=1}^{n_f} \bar{q}_j (i\not{D} - m_j) q_j, \tag{1.15}$$

where q_j are the quark fields (n_f are the different flavours) with mass m_j. The covariant derivative D_μ is defined as

$$D_\mu = \partial_\mu + i e_s \mathbf{g}_\mu, \tag{1.16}$$

where $\mathbf{g}_\mu = \sum_A t^A g_\mu^A$. The g_μ^A are the gluon fields and t^A the $SU(3)$ group generators, which obey the commutation relation $[t^A, t^B] = iC_{ABC}t^C$. The gluon dynamics is given by

$$F_{\mu\nu}^A = \partial_\mu g_\nu^A - \partial_\nu g_\mu^A - e_s C_{ABC} g_\mu^B g_\nu^C . \tag{1.17}$$

It has to be noticed that, according to Equation 1.17, the QCD theory includes triple and quartic gauge couplings between the gluons. Finally, it is customary to redefine the gauge coupling constant e_s, in analogy with the QED, as

$$\alpha_s = \frac{e_s^2}{4\pi}. \tag{1.18}$$

1.2.3.1 Quark confinement and asymptotic freedom

The QCD Lagrangian of Equation 1.15 has a simple structure but a very rich dynamical content. It can describe the complex spectrum of hadrons and it implies the striking properties of *confinement* and *asymptotic freedom*. Confinement is the property that no colour charged particles can be observed but only colour singlet ones. A complete understanding of this mechanism is particularly difficult, due to the non-perturbative regime in which it appears, but a great improvement has been obtained over time from so-called "lattice" simulations of QCD. It turns out that the potential between a quark and an antiquark has a Coulomb part at short distances r and a linearly rising term at long ones

$$V_{\mathrm{q}\overline{\mathrm{q}}} \sim \frac{\alpha_s(r)}{r} + \sigma \cdot r. \tag{1.19}$$

The $\sigma \cdot r$ term is the one that makes it energetically impossible to separate a $\mathrm{q}\overline{\mathrm{q}}'$ pair. If the pair is created at one point in space-time and then the quarks start to move away from each other, it becomes energetically favourable to create additional quark pairs between the two leading charges, thus screening any net colour charge. Confinement is fundamental to explain why the strong force has a very limited range (approximately the proton radius) despite the gluons being massless.

In Equation 1.19 the strong coupling constant $\alpha_s(r)$ is dependent on the distance r, which can also be thought as the Fourier-conjugate of the momentum transfer. In perturbation theory, the calculation of the interaction amplitudes will contain, at a certain point, particle "loops": the integration over the loop particle momenta generate infinities because, in principle, the particle momentum is not bounded from above. These divergences, usually called *ultraviolet divergences*, are tackled with the *regularisation* procedure, that is an unphysical energy scale μ_R is introduced in the theory as an effective cut-off. In a renormalisable theory any dependence on the cut-off is "reabsorbed" into the bare parameters of the theory by modifying the masses, field strengths and coupling constants. The renormalisation procedure is necessary since the observed results cannot depend on any unphysical cut-off. After renormalisation is achieved, the physical observable is no longer constant but depends on the momentum transfer Q^2 (alternatively the distance r) at which the observable is probed. This dependency can be specifically parametrised by employing the renormalisation group equation (RGE). The RGE of the strong coupling constant α_s is

$$\mu^2 \frac{\partial \alpha_s(\mu^2)}{\partial \mu^2} = -\alpha_s \left[\beta_0 \frac{\alpha_s}{4\pi} + \beta_1 \left(\frac{\alpha_s}{4\pi} \right)^2 + \mathcal{O}(\alpha_s^3) \right], \tag{1.20}$$

where β_i are constants related to the gauge group. Solving Equation 1.20 at leading order, assuming α_s has been determined at a certain energy scale Q_0^2, gives

$$\alpha_s(Q^2) = \frac{\alpha_s(Q_0^2)}{1 + \alpha_s(Q_0^2)\beta_0 \ln\left(\frac{Q^2}{Q_0^2}\right)},$$

and for high momentum transfer the coupling constant α_s becomes small as $\beta_0 > 0$. This important result, which leads to the QCD fundamental property of *asymptotic freedom* [19,20], implies that quarks and gluons can be considered as free, non-interacting particles, as long as the process of interest occur at large Q^2.

1.3 Summary, limitations and extensions of the Standard Model

In Section 1.1, Section 1.2.2 and Section 1.2.3 the particle content and the underlying theory of the SM were presented. Ideally, any interaction between fundamental particles can be described by the SM Lagrangian

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{EW}} + \mathcal{L}_{\mathrm{QCD}} + \mathcal{L}_{\mathrm{Higgs}}. \tag{1.21}$$

In Equation 1.21 several parameters, such as the coupling constant or the particle masses, are free and have to be determined through experiments. Over the last decades, the predictions of the SM has been validated by several experiments (see also Table 1.1). Consistency tests of the SM were performed both by the LEP and SLC Collaborations [21, 22] and are nowadays in progress at the LHC [23].

Although the SM gives a highly successful description of the physics governing the fundamental particle interactions, it still leaves some open questions and strong arguments are calling for its extension. The numerous shortcomings of the SM are one of the main motivation for the construction of new collider facilities, able to probe at larger and larger energies the theory. The most striking deficits of the SM are briefly mentioned in the following:

- the hierarchy problem is the question why the scale of EW symmetry breaking (or alternatively the Higgs boson mass) is $\mathcal{O}(100\,\mathrm{GeV})$. This is a fundamental problem since, in principle, this scale should receive corrections of the order of the largest scale relevant in the SM (like the Plank scale or the grand unification scale). This discrepancy can either be explained by a fine tuning of the perturbative expansion contributions to the Higgs mass or by introducing a new gauge symmetry in the model (a so-called "supersymmetry");

- the large number of parameters of the SM is an aesthetic and conceptual problem. In a fundamental theory, one would expect motivations for the values taken by certain parameters but, in the SM, the values of more than 30 parameters (masses, mixing angles, couplings) must be put "by hand" (through experimental measurements). In addition, the values of some of them are rather puzzling from a phenomenological point of view (as an example the mass range of fermions starts from the MeV energy scale for electrons and extends to more than 170 GeV for the top quark);

- the particle character of the neutrino is still open. Its neutral electric charge makes it special among other fermions since it may be its antiparticle ("Majorana" instead of "Dirac" neutrino);
- the quantisation of electric charge can only be explained if magnetic monopoles exist (which, so far, have not be observed);
- the SM does not predicts any charge or mass unification at some large energy scale;
- from a cosmological point of view, there is not a renormalizable quantum theory of gravity, nor an observation of a "graviton" (the boson associated with this force field) or a candidate able to explain the origin of the observed dark matter/energy content of the Universe.

1.3.1 The Effective Field Theory framework

A possible extension of the SM, which allows to parametrise in a general way possible contribution of new physics (NP) beyond the SM (BSM), is presented in this section. Two complementary approaches could be used for BSM searches: on one hand, the *direct* approach aims to find evidence of NP via the production of new particles in high energy collisions, while, on the other hand, the *indirect* approach searches for novel interactions between the know SM particles. The latter is commonly pursued by developing a so-called "Effective Field Theory" (EFT) expansion of the SM. It turns out that there exists a natural way to extend the SM, by adding additional unknown interactions, in such a way that the $SU(3) \times SU(2) \times U(1)$ gauge symmetry is respected. The EFT framework seeks to parametrise any NP contributing to the known particle interactions by following a largely model-independent approach. Its main advantages are the power to search for NP without restricting to a particular theoretical model and, in case of no discovery, the ability to quantify the accuracy with which the BSM physics can be excluded. To build a model-independent extension of the $\mathcal{L}_{\text{SM}}$, the following guiding principles are used [24]:

- it should satisfy the basic requirements for a QFT of particle interactions, such as unitarity and analyticity;
- it should not spoil the $SU(3) \times SU(2) \times U(1)$ gauge symmetry. Moreover, the SM predictions should be restored in an appropriate limit;
- it should simultaneously be sufficient generic, as to include any possible NP contribution, but also sufficiently precise to give guidance where to look for new effects;
- the perturbative expansion corrections should be calculable in the new theory.

In Equation 1.21 all operators are restricted to be of mass dimension $d \leq 4$, thus a natural way to extend the theory is by inserting operators of dimension higher than that. It can be proven that all these new higher dimension operators must be accompanied by terms proportional to inverse powers of a mass such that the dimensional properties of the Lagrangian density are maintained. The mass scale Λ characterise the coefficients

of the higher dimensional operators, can take any value from few TeV to very high energies, and it is assumed to be far beyond the experimentally accessible energies. This allows a perturbative expansion of the SM Lagrangian as an effective theory at low energies, while the expansion becomes meaningless when approaching the scale at which NP occurs. A general EFT expansion of Equation 1.21 can be written as

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{f_i^{(6)}}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j \frac{f_j^{(8)}}{\Lambda^4}\mathcal{O}_j^{(8)} + \cdots , \tag{1.22}$$

where only the dimension-six and dimension-eight operators were explicitly included. The dimensionless Wilson coefficients $f_a^{(b)}$ parametrise the coupling strength of NP to the SM particles. In the EFT framework the particle content of the theory, that is the fields from which the operators $\mathcal{O}_i$ are constructed from, must be fully specified. There is no evidence for new particles at the energy scale available from particle accelerators so it is reasonable to assume the $\mathcal{O}_i$ operators to be built using only the SM fields. Since the Wilson coefficients are fixed by the complete high energy theory, any extension of the SM can be parametrised by Equation 1.22, where the coefficients of the operators $(f_a^{(b)}/\Lambda^{b-4})$ are kept as free parameters to be constrained by experiments. From Equation 1.22 it can also be seen that the $\mathcal{L}_{\text{SM}}$ is re-obtained in the limit $\Lambda \to \infty$.

As further explained in Section 3.2 the EFT framework can be used to parametrise BSM interactions which can arise in the $\text{pp} \to \text{Z}\gamma\gamma$ channel.

2 | Perturbative calculations and Monte Carlo tools

In this chapter the physics basis, the main features and the use of general-purpose Monte Carlo (MC) programs for the simulation of the signal and background processes expected to contribute to the $pp \to Z\gamma\gamma$ production cross section are presented. The key aspect of the event generators is that, by simulating the final state topology of each process, they can provide a valuable description of the experimental observables under study. It is by comparing observed and simulated data that the theoretical models, developed to describe reality, can be tested.

In Section 2.1 the basic features of a LHC proton proton collision are described. At high energy scales the partons are considered as free particles which eventually undergo hard scattering processes, whose matrix elements can be computed in perturbation theory, as described in Section 2.2. At very low energy scale, of the order of 1 GeV, the perturbative calculations cannot be used any more and an effective approach is used to model the parton confinement. The non-perturbative QCD methods used to parametrise the partons hadronisation process in the generated events are described in Section 2.3. The evolution from the hard to the soft scattering regime is modelled by a perturbative QCD approach, which implies the production of many additional particles from both initial state and final state hard scattering partons via initial (final) state radiation processes. The approaches used to model the so-called "showers of partons" are presented in Section 2.2.2. To obtain a complete picture of the interaction, predictions obtained from the parton shower algorithms and the matrix element calculations are matched together. The procedures developed to properly combine the hard scattering amplitude calculations to the parton showers are described in Section 2.2.4. To be compared with observed data, simulated events need to include the interactions of the generated particles with the detector. A brief overview of the CMS detector simulation is given in Section 2.4.

The simulation of high energy particle collisions requires calculations which often involve integrations over large and variable dimension phase spaces. Those calculations are particularly suitable to be performed by computer simulations. Moreover, the evolution from the hard to the soft scattering regime can be modelled as a Markov process and can be handled effectively with a MC approach. The same applies for what concerns the parton hadronisation process. Finally, the factorised nature of the problem allows to treat each regime separately and to eventually improve the precision of each calculation once more sophisticated models become available.

2.1 Overview of an LHC collision

A precise understanding of the properties of the complicated final states that could arise from high energy proton proton collisions is an extremely challenging theoretical and computational problem. The main issue involves dealing with the calculations of the QCD interactions between initial and final state partons: despite being a perturbative theory at short distances (high energies), phenomena occurring at lower-energy scales are intrinsically non-perturbative. The transition from the perturbative regime, the one which characterises the scattering between two partons, to the non-perturbative one, the one which regulates the final state hadrons, needs to be modelled with care. These multiple regimes cannot be treated all at once. Instead, processes are *factorised* such that each of them is treated differently according to the scale of the momentum transfer involved. The build-up of each event structure can be then divided into several basic phases, illustrated in Figure 2.1, which are separately simulated:

- at the highest energy scale the *hard process* of interest occurs, which is characterised by a large momentum transfer between the constituent partons of the colliding particle beams. The outcome of each collision is usually represented by a small number of energetic outgoing partons, leptons or gauge bosons. The matrix element of the hard process can be computed in perturbative QFT. It has to be noticed that when a short-lived resonance, such as the Z boson, is produced its decay is included in the process itself, thus describing any spin correlation of the decay products;

- each collision necessarily implies the presence of accelerated colour and electromagnetic charges, which could produce bremsstrahlung radiation. Radiation emissions associated with the incoming particles are usually referred to as *initial state radiation* (ISR); alternatively, emissions from the final state particles are labelled as *final state radiation* (FSR). Both ISR and FSR are modelled by the *parton shower* algorithm. The distinction between the hard process and the ISR and FSR could be ambiguous, especially when dealing with higher-order computations and great care is needed when the predictions obtained at different energy scales are merged;

- only the $pp \to Z\gamma\gamma$ hard processes are of interest for this analysis, while the vast majority of collisions are indeed soft and originate from additional interactions between the hadron remnants. The *multi-parton interactions* (MPI) fall in the non-perturbative QCD regime, meaning that dedicated models with tunable parameters are needed to describe the data;

- finally, as soon as partons created from the previous steps move away from each other, the QCD interaction between partons becomes relevant. The structure of the confinement field is not currently described from first principle and a dedicated model, tuned on data, is used to effectively predict the production of additional quark-antiquark pairs which results into colour neutral final states. The algorithm used to simulate hadrons production is called *hadronisation*. It also includes the modelling of any subsequent unstable hadron decay into long-lived particles.

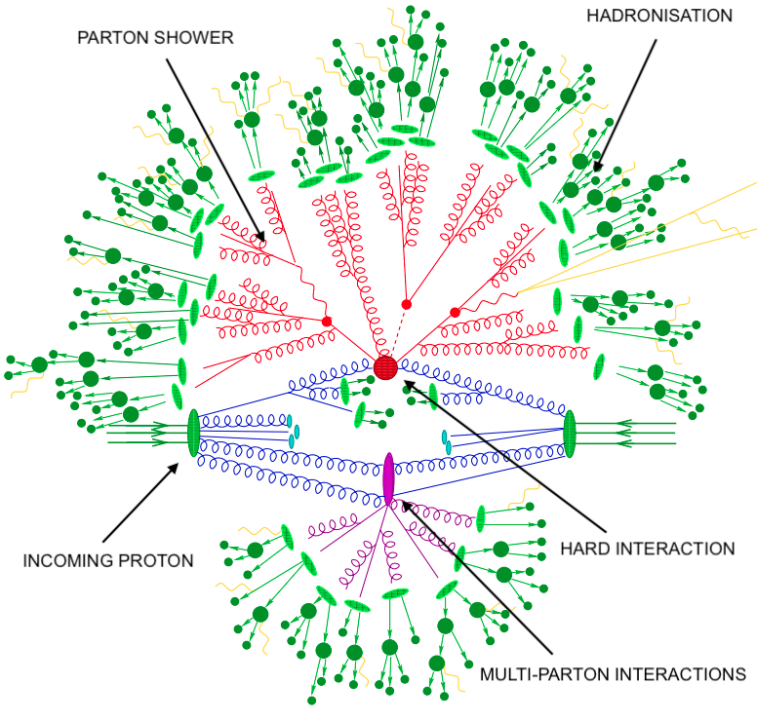


Figure 2.1: Schematic illustration of an high energy proton proton collision in the LHC. Figure adapted from reference [25].

The factorisation approach allows these steps to be considered, to some extent, individually and sequentially. Dedicated sets of rules are then exploited, by the MC method, to iteratively construct more and more complex states. The aim is to provide the most realistic possible description of the process of interest.

2.1.1 Collinear and soft emission and jet clustering

Phenomena which involve the QCD potential are described through interactions between quarks and gluons. However, neither quarks nor gluons are experimentally visible. Instead, right after its production each parton undergoes fragmentation, by emitting additional soft and collinear partons, and ends up in a collimated spray of energetic particles referred to as *jet*. The hadronisation process (see Section 2.3) is such that the jet structure of an event is preserved. Final states dominated by jets of hadrons are, indeed, an experimental evidence in high-momentum transfer hadronic events.

Despite representing a predominant feature of QCD interactions, jets are not fundamental objects and need to be properly defined. A *jet algorithm* is a set of rules regulating how partons are clustered together and how the momenta of the single partons contribute to the energy of the whole jet. As further explained in Section 2.2.4, one important application of the jet object is to match the hard process fixed order calculations with the parton shower evolution of the outgoing hard partons. This can be done only if the jet algorithm, used to cluster the partons, is well defined at each

order of perturbation theory. This is true only for jet algorithms that are collinear and infrared safe (IRC). This property states that, for any partonic configuration, if one replaces a parton with a collinear set of particles sharing the same momentum or adds any number of infinitely soft partons, emitted in every direction, the number of hard jets found in the event should remain unchanged. It can be proved that, once the IRC requirement is satisfied by any jet algorithm used, jets cross sections are finite at any perturbative order [26].

Particles in data and MC samples, used in this work, are clustered by using the IRC safe anti-k_T algorithm. The clustering procedure is done by assessing the distance between the particles i and j, which is defined according to

$$d_{ij} = \min(p_{T,i}^{2s}, p_{T,j}^{2s}) \frac{\Delta R_{ij}^2}{R^2}, \tag{2.1}$$

where $p_{T,i}$ is the transverse momentum of particle i, R_{ij} the angular distance between the particles i and j and R a fixed parameter which defines the "radius" of each jets. It can be proven that, if s is chosen to be equal to -1, the clustering procedure tends to involve hard rather than soft particles [27], meaning that jets are grown aggregating soft particles around hard seeds.

2.2 Factorisation formula for QCD cross sections

The process of interest for this analysis, the associated production of a massive vector boson (Z) and photons, involves large momentum transfers between the colliding protons. The proton is not a fundamental particle but a composite state made of three valence quarks, as well as sea-quarks and gluons. At the energies reached by the proton beams circulating into the LHC, the incoming partons could be considered as *asymptotically free* and their hard scattering processes can be described by perturbative QCD. The cross section σ for a generic hard scattering process $ab \to n$ between the partons a and b is computed by applying the *collinear factorisation principle* [28] as

$$\sigma = \sum_{a,b} \int_0^1 \mathrm{dx_a dx_b} \int \mathrm{f_a^{h_1}(x_a, \mu_F^2) f_b^{h_2}(x_b, \mu_F^2) d\hat{\sigma}^{ab\to n}(x_a x_b s, \mu_F^2)}, \tag{2.2}$$

where s is the hadronic centre of mass energy, x_i is the fraction of the total proton momentum carried by the parton i, $f_i^{h_i}(x_i, \mu_F)$ is the parton distribution function (PDF) of the parton, which depends on the momentum fraction x_i and on the factorisation scale μ_F (see Section 2.2.1) and $\hat{\sigma}^{ab\to n}$ denotes the parton level hard scattering cross section for the production of the final state n through the initial partons a and b (see Section 2.2.2).

2.2.1 Parton Distribution Functions in event generators

The proton PDFs play a central role in the event generation not only for the computation of the hard processes but also for the parton shower simulation. An appropriate choice of the PDFs to be used when generating the events is important since it would influence both the event cross section and shapes of the kinematical observables.

In Equation 2.2, the function $f_i^{h_i}(x_i, \mu_F)$ can be seen as the probability to find a parton i with a defined momentum fraction x_i when the proton is probed at a energy scale μ_F. This is only a leading-order approximation since collinear and soft emissions below μ_F are considered "unresolved" and included into the PDFs themselves. One can define the total momentum splitting between the proton constituents, by invoking energy conservation, as

$$\sum_i \int_0^1 x f_i(x, \mu_F^2) = 1.$$

The PDFs cannot be calculated explicitly from any QCD first principle but can be considered, with good approximations, as universal functions. This means that they can be derived by independent measurements, usually represented by deep inelastic scattering experiments between a probing lepton and proton beam. By controlling the energy of the incident lepton, the number of the outgoing particles can be related to the abundance and momentum fraction of the proton constituents. Such measurements define the PDFs values for a fixed value of μ_F but, fortunately, the evolution of the PDFs as a function of μ_F can be predicted in perturbative QCD via the DGLAP equation [29–31]

$$\mu_F \frac{\partial f_a(x, \mu_F^2)}{\partial \mu_F^2} = \sum_b \frac{\alpha_s(\mu_F^2)}{2\pi} \int_x^1 P_{ab}(z) f_b\left(\frac{x}{z}, \mu_F^2\right) \frac{\mathrm{dz}}{z}, \tag{2.3}$$

where $P_{ab}(z)$ is the Atarelli-Parisi splitting function which describes the probability for a parton a to split into two $a \to bc$. The longitudinal momentum of the resulting particle is $p_b = zp_a$ and the additional parton c is absorbed into the proton sea-quarks.

Contributions to the PDFs are measured at several μ_F values with experiments sensitive to different quark flavours. The datasets obtained are combined and fitted with a suitable functional form, considering the DGLAP evolution Equation 2.3, to obtain an effective parametrisation of the PDFs, referred to as *PDF set*. Uncertainties related to the fitting procedure and the specific choice of the input datasets are provided as variations with respect to the nominal PDF set. There exists a large number of different PDF sets, collected in the LHAPDF common library [32], which are differentiated by the specific choice of the initial datasets and by the implemented fitting procedures.

In this work, simulated events compared with data collected during the 2016 (2017 and 2018) year of data-taking make use of the neural networks NNPDF3.0 [33] (NNPDF3.1 [34]) set of PDFs. In Figure 2.2 the scale dependence of the NNPDF3.1 set is shown: the main differences between the NNPDF3.1 and NNPDF3.0 release consist in the inclusion of new LHC data, with high-precision PDF-sensitive measurements, and of all the recent progress in QCD calculations.

2.2.2 Perturbative expansion and matrix element calculation

Having defined the PDFs, the remaining missing piece to the calculation of the hard process cross section is represented by the computation of the parton-level cross section $\hat{\sigma}^{ab \to n}$. This quantity depends on the allowed momenta phase space of the final state particles Φ_n, as well as on the factorisation scale μ_F and the renormalisation scale μ_R. The latter is introduced to cure the ultraviolet divergences which could appear during computation. The fully differential parton-level cross section can be computed as the

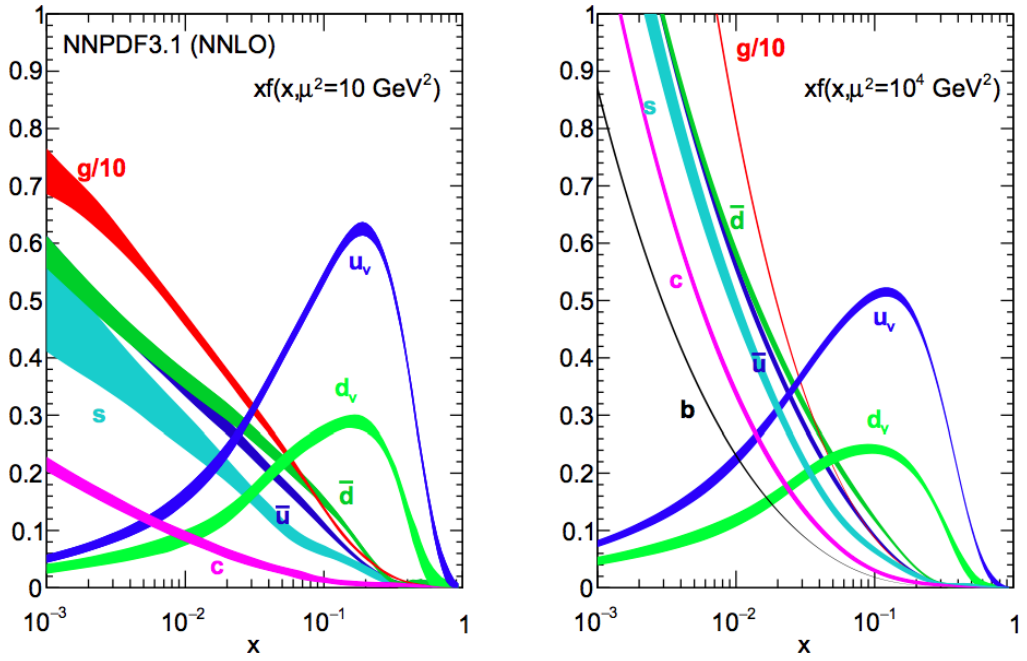


Figure 2.2: The scale dependence of the NNPDF3.1 set of PDFs evaluated at $\mu_F^2 = 10\,\mathrm{GeV}^2$ (left) and $\mu_F^2 = 10^4\,\mathrm{GeV}^2$. Figure adapted from reference [34].

product of the corresponding matrix element squared $|M_{ab\to n}|^2$ (which is averaged over the spin and colour degree of freedom of the initial state partons) and the parton flux $1/(2x_a x_b s)$ following

$$\mathrm{d}\hat{\sigma}^{\mathrm{ab\to n}} = \frac{1}{2\mathrm{x_a x_b s}}|\mathrm{M_{ab\to n}}|^2(\Phi_\mathrm{n}, \mu_\mathrm{F}, \mu_\mathrm{R})\mathrm{d}\Phi_\mathrm{n}. \tag{2.4}$$

According to the perturbative QCD prescriptions, the parton-level cross section can be expanded in series with respect to the coupling constant of the theory as

$$\mathrm{d}\hat{\sigma}^{\mathrm{ab\to n}} = \mathrm{d}\hat{\sigma}_0^{\mathrm{ab\to n}}\alpha_\mathrm{s}^0 + \mathrm{d}\hat{\sigma}_1^{\mathrm{ab\to n}}\alpha_\mathrm{s}^1 + \ldots, \tag{2.5}$$

where the expansion holds as long as the strong coupling constant α_s is sufficiently small. The perturbative expansion of Equation 2.5 can be absorbed into the definition of the matrix element squared $|M_{ab\to n}|^2$ according to

$$M_{ab\to n} = \sum_i F_{ab\to n}^{(i)},$$

where the $F_{ab\to n}^{(i)}$s are the *Feynman diagrams* describing the process of interest. In the Feynman formalism each perturbative contribution to the final amplitude of a quantum transition is represented by a specific diagram. The coupling constants are included into the QCD (EW) interaction vertices which are proportional to $\sqrt{\alpha_s}$ ($\sqrt{\alpha_{EW}}$). A calculation at a given perturbative order includes the Feynman diagrams such as the

product of all the vertices connecting initial final states is equal or lower than the order of the desired expansion. Diagrams representing the leading-order (LO) term of the perturbative expansion are those without any loop contribution, while the next-to-leading-order (NLO) corrections include extra particles in loops or additional legs. An example of LO and NLO QCD contributions to the pp $\rightarrow$ Z$\gamma\gamma$ production cross section is given in Figure 2.3. These extra emitted particles make the combination with subsequent steps of the event generation not trivial and introduce infrared divergences which must cancel between the various terms of the series expansion. An automated procedure to generate Feynman diagrams at NLO, to derive the corresponding amplitudes and to match predictions at different scales has been made available recently through the MADGRAPH5_aMC@NLO [35] program, which is used extensively for the MC simulations used in this work.

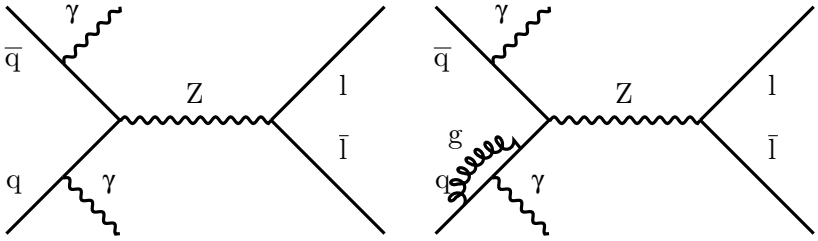


Figure 2.3: Example of LO (left) and NLO QCD contribution to the pp $\rightarrow$ Z$\gamma\gamma$ cross section. The LO diagram, also referred to as "tree-level", does not contain any loops and the strength of the interaction is proportional to α_{EW}^4 while for the NLO loop diagram, interfering with the tree-level one, gives a contribution proportional to $\alpha_{EW}^4\alpha_s$

2.2.3 Parton showering

Since particles entering the hard scattering process and those leaving it are typically QCD partons, they can radiate gluons and photons by reducing their energy. Similarly, outgoing leptons generated in the hard subprocess can radiate photons. Both radiated gluons and photons can produce additional quarks and lepton pairs, thus generating showers of outgoing particles. A crucial observation is that, despite soft emissions can occur with an energy spectrum down to arbitrarily low frequencies so that the total number of soft emitted particles is ill-defined, the number of the observable particles above a given energy threshold is finite. The fact that the bremsstrahlung emission from different particles involved in the same scattering process is coherent in one important property of QED. One manifestation of this effect is that, when a high energy photon produces an e^+e^- pair in the field of a nucleus, the charged particles only behave independently in a forward cone with an opening angle given by their separation, while at larger angles they can be treated as a coherent pair. There is a corresponding effect also in QCD thus both QCD and QED emission processes can be handled by a common algorithm by simulating successive soft or collinear emissions of massless particles. The cross section for a process which involves a parton emission diverges if the emitted particle is collinear (that is emitted with an angle $\theta \rightarrow 0$) or soft (that is its energy going to zero, independently of the emission angle θ), but it turns out that these divergences

do not prevent the computation of a finite probability for a showering process to occur. The cross section of a hard process that produces a certain number of partons of any flavour i can be factorised, in analogy with Equation 2.3, as

$$\mathrm{d}\sigma \approx \sigma_0 \sum_{partons,\ i} \frac{\alpha_s}{2\pi} \frac{\mathrm{d}\theta^2}{\theta^2} \mathrm{d}z P_{ij}(z,\phi)\mathrm{d}\phi, \tag{2.6}$$

where j is a generic parton emitted with a momentum fraction z. As can be seen from Equation 2.6, the $\mathrm{d}\sigma_{\mathrm{n+1}}$ cross section is proportional to the corresponding $\mathrm{d}\sigma_{\mathrm{n}}$ one and the rest of the expression represents the (differential) probability for a parton emission. The $P_{ij}(z,\phi)$ splitting functions are the non-spin averaged version of the Altarelli-Parisi splitting functions of Equation 2.3. The evident collinear divergence in Equation 2.6 is not an issue if one observes that no physical measurements can distinguish two exactly collinear partons from a single one carrying the same total momentum. A resolution criterion is introduced such that only observable topologies can be generated: two particles are considered as "resolvable" if their relative transverse momentum is larger than some cut off energy Q_0, under which the parton hadronisation process becomes important and must be treated with a different approach (see Section 2.3). It can be shown that this criterion corrects for both the collinear and soft divergences by allowing to evaluate a finite total-resolvable emission probability.

Equation 2.6 is a general expression, valid for any hard process accompanied by a collinear splitting, which gives the inclusive emission distribution of all partons emissions. However, to generate one collinear splitting and then treat the obtained final state as a new hard process it is convenient to consider separately the contribution of every emission process. This is done by introducing an ordering variable, such as the virtuality of the internal parton line q^2. The emission processes can then be simulated with a step-wise Markov chain, generally referred to as *parton shower algorithm* (PS), by choosing on a probabilistic basis to add one or more partons (photons) to the final state. The probability distribution of each branching, $\Delta_i(Q^2, q^2)$, is given by the probability that no branching with virtualities greater than q^2 has occurred given the initial available maximum possible transverse momentum of the emitting parton Q^2. A particular value of Δ_i is when q^2 is chosen to be equal to Q_0^2. In such case $\Delta_i(Q^2, Q_0^2)$ is referred to *Sudakov form factor*, and gives the probability for a parton to evolve from the hard process till the hadronisation scale cut off without any branching. The inclusion of the Sudakov probability term in the PS algorithm ensures that the total probability for a parton to branch never exceed unity. The showering process is formulated both as a forward evolution, for the outgoing partons, and a backward evolution of the incoming partons, thus modelling both FSR and ISR processes. There exists different programs able to handle the parton shower processes and the one more extensively used in this analysis is the PYTHIA8 [36] one.

2.2.4 Matrix element calculations and parton showers

Combining the hard process matrix element (ME) calculation with the parton shower evolution of the initial and final state partons is particularly desirable and a lot of effort in the recent years has been dedicated to this task. Fixed order ME calculations are particularly suitable to simulate separated hard partons processes but do not guarantee

high precision when trying to describe collinear or soft emission. The PS algorithm, on the other side, is specifically designed to populate the soft and collinear region, but becomes inefficient when a hard process is simulated. A good description of soft and collinear multi-parton states in an event is fundamental to guarantee subsequent hadronisation models to properly work.

Matching ME and PS prediction is not a trivial task and to add a parton shower to a hard process, simulated with a matrix element generator, simply does not work. Among all the technical consideration, one major problem involves the avoidance of "double counting" some regions of the phase space or, conversely, "undercounting" others. The problem lies in the fact that one identical final state configuration can arise in different ways: final states characterised by a hard emission within a soft scattering process are indistinguishable from final states characterised by a soft emission within a hard scattering process. For LO ME predictions the problem is circumvented by simply limiting the energy Q^2 at which the shower starts to be lower than the typical energy scale at which partons are generated by the ME event generator. This procedure is commonly referred to as *shower matching*, and several alternative algorithms exist to combine LO ME predictions to the PS evolution. The LO MC samples considered in this analysis make use of the MLM approach [37] to match the MADGRAPH5_aMC@NLO ME calculation with the PYTHIA8 PS predictions. Matching ME predictions at NLO is far more ambitious than the LO case. As before, hard emissions need to be generated according to the ME prescriptions while soft and collinear emissions should fall into the PS description. Nevertheless, both real and virtual contributions to the Feynman diagrams need to be matched to the PS, so that cross sections are accurate to NLO and NLO results are obtained for all the observables when expanded in powers of α_s. Among the different approaches developed to accomplish this task, the MC@NLO [38] one has been worked out for a large number of different processes and it is thus extensively used to generate the vast majority of the MC samples used in this analysis. The MC@NLO approach exploits analytical calculations of how the first branching in a shower, starting from a n-body topology, would populate the $n+1$ phase space to correct the ME predictions event by event. The $n+1$ analytical PS prediction is subtracted from the ME $n+1$ real emission while living any virtual correction unaltered. The kinematic of the hard process is thus modified accordingly. Finally, the standard PS evolution is applied to each corrected ME contribution. The MADGRAPH5_aMC@NLO NLO samples used in this work exploit the FxFx approach [38] to match NLO ME element prediction to the PYTHIA8 PS evolution.

2.3 Hadronisation, multi-parton interactions and pileup

Since no free quarks or gluons can be experimentally seen, the PS evolution of the final state partons must be bounded by a lower limit value of the parton virtuality. Indeed, when each final state parton is evolved down to an energy scale $Q_0^2 \sim 1\,\text{GeV}$ the strong coupling constant α_s is large enough to prevent any further perturbative approach. At this energy scale partons end to act as free particles and start to form composite colourless states of two or three quarks. There exists yet no theoretical models able to describe the *hadronisation process* from first principles but several effective models, able to describe with great precision the experimentally observed hadronic final

states, were developed over the years. Two of them, the *cluster model* [39] and the *string model* [40], are mainly used in the general-purpose MC generators.

The cluster model, implemented in the SHERPA [41] program, is based on the fundamental observation that the colour structure of a parton shower, in the proximity of the hadronisation scale Q_0^2, is such that colour singlet combinations of the partons can be formed according to a universal mass distribution. The latter is independent of the hard energy scale of the process Q^2 [42]. At the end of the PS shower evolution, all gluons are split into $q\overline{q}$ pairs and neighbouring quarks are combined into colour singlets, known as "clusters", by keeping track of the colour flow in the event. Clusters are then decayed into hadrons: while low energy clusters will originate a single hadron, high energetic clusters are further split into lower energy ones before decay.

The string model exploits a rather different approach based on lattice QCD studies. In this manifestly not Lorentz covariant framework quarks are assumed to be static and the energy stored by the QCD field between a colour and an anti-colour charge increases linearly with the spatial separation between the particles. This peculiar behaviour of the QCD potential is completely different from that of the (usual) QED ones, and it is deeply connected to the fact that the massless QCD quanta can interact via three gluon vertices (see Section 1.2.3). The quark linear confinement principle is the starting point for the string model interpretation of the hadronisation process. As schematically presented in Figure 2.4 (left), the QCD field is parametrised as a colour flux tube, connecting each $q\overline{q}$ pair, which is stretched as the partons move apart from each other. To obtain a Lorentz covariant causal description of the rising potential, the colour flux

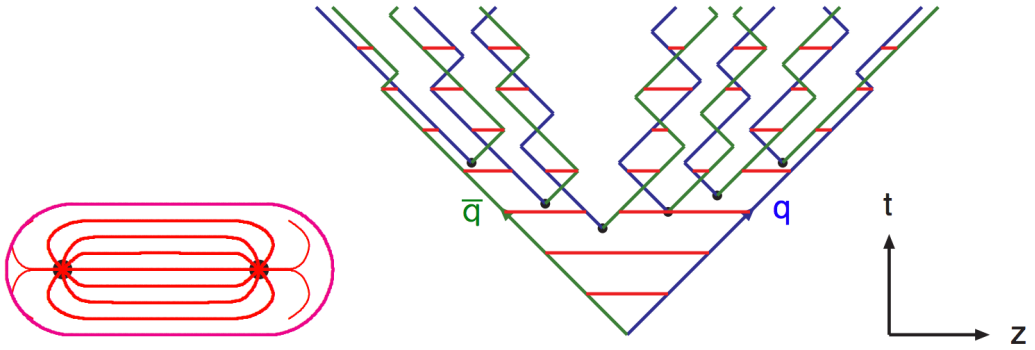


Figure 2.4: Sketches of the Lund string model. On the left, the QCD potential flux tube between two partons is depicted while, on the right, the motion and break-up of a string system in different $q\overline{q}$ pairs is presented. Figure adapted from reference [43]

tube is mathematically modelled as one-dimensional string (Figure 2.4 right), where local kinks represent emitted gluons. The string constant, that is the amount of energy stored by the colour potential between the two partons, is found to be $\kappa \sim 1\,\text{GeV/fm}$, a value consistent with results obtained from hadron spectroscopy measurements. As the system evolves in spacetime, the energy stored by the string linearly increases until it may eventually break producing additional quark pairs. In the *Lund string model* [44] used by PYTHIA8 the break-up process is assumed to occur until only on-shell hadrons remain. The final state hadrons correspond to segments of the original string connecting

the initial $q\bar{q}$ pair.

Besides hadrons coming from the hard process, an event is also populated by secondary soft interactions of the remaining partons of the proton beams which need to be included to obtain a realistic description of the observed final state. Those secondary interactions are collectively referred to as *underlying events* and can be divided into multiple parton interactions (MPI) and remnants of the interacting protons. The MPI are primarily soft interactions and do not generally end up in identifiable jets. In any case, they impact the total scattered energy and could cause colour exchange between the hard process particles affecting the hadronisation stage. This is also the case for the soft elastic scattering processes between the beam remnants. In this work, the effect of the underlying events on the MC simulations is modelled using PYTHIA8 with specific parameters tuned by the CMS experiments, referred to as CUETP8M1 [45] and CP5 [46] tunes.

Additional proton proton interactions occurring in the same event, known as *pileup events* are included in the simulated samples as well. A large dataset of minimum bias events, composed by elastic and low energy proton proton scattering events, has been simulated with PYTHIA8. Events were then sampled randomly, following a Poissonian distribution with mean the expected number of pileup events for each LHC collision, and added to the hard scattering events (see also Section 6.3.1).

2.4 Detector simulation

To compare predictions to data, the interaction of simulated particles with the CMS sub-detectors is modelled with the GEANT4 software [47]. A detailed version of the CMS geometry, shown in Figure 2.5, is implemented in the software to properly include any active and inactive volume. The GEANT4 suite covers several physics processes, including electromagnetic, hadronic and nuclear interactions, over a comprehensive energy range from the eV to the TeV scale. Monte Carlo methods are used to track the particle trajectories inside the different sub-detectors, to model their decays and to simulate the released energy deposits and hits. The digitised information from the simulated sub-detectors is then fed into the same reconstruction algorithms used to process data. The performances of the CMS detector simulation are constantly tuned and refined, thus accounting for any modification in the active and inactive material characteristics, but residual differences between collected and simulated data are unavoidable. They arise due to small inaccuracy when modelling either the detector geometry or response. These difference are usually small and corrected by comparing data and simulation results for know quantities, as it is further explained in Section 6.3.3.

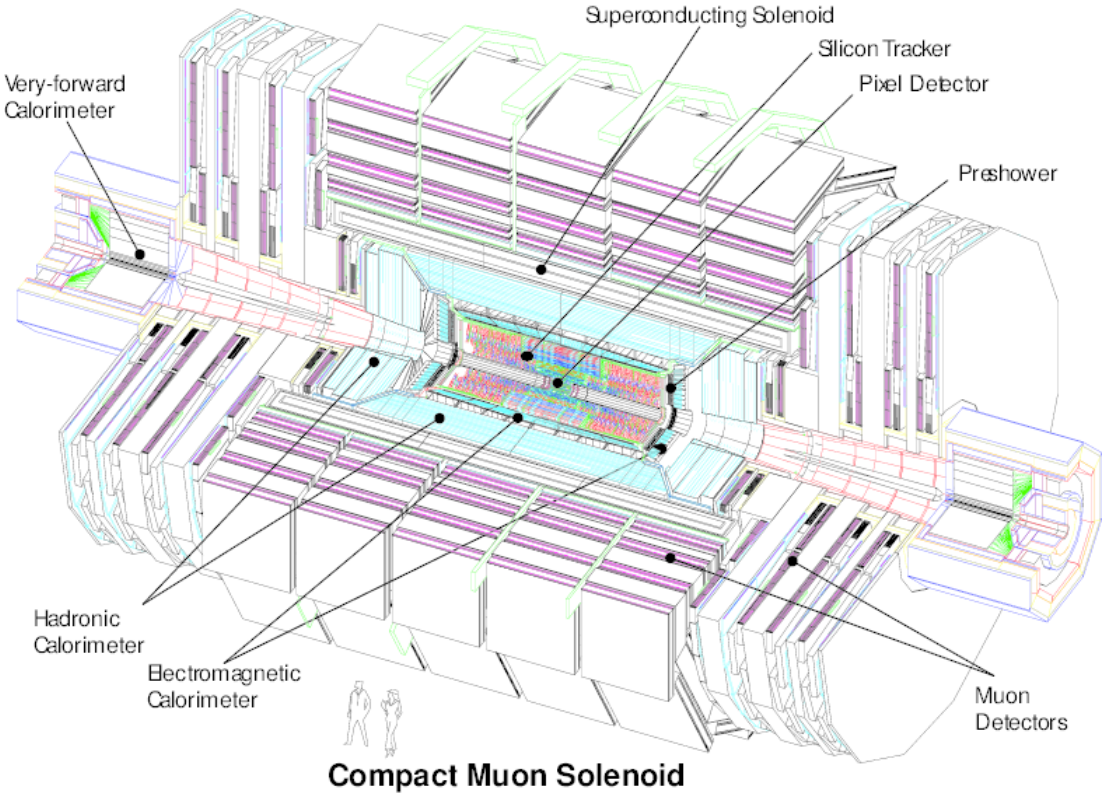


Figure 2.5: Schematic illustration of the Compact Muon Solenoid detector geometry. Figure adapted from reference [48].

3 | Phenomenology of the Zγγ production at hadron colliders

In this chapter, the phenomenology of the Zγγ production at hadron colliders will be reviewed. Because the electroweak (EW) gauge bosons carry a weak charge, they are allowed to self interact. The Standard Model (SM) predicts interaction vertices which contain three or four EW gauge bosons. These interactions contribute to the inclusive production of diboson and triboson final states, the latter being interesting from an experimental point of view because of low background and high detection efficiency. Due to their kinematic characteristics, multi-boson final states represent exceptional experimental probes with which the SM gauge symmetry can be tested.

The study of the quartic gauge coupling (QGC) interactions, which leads to the Zγγ triboson final state, is the main focus of this work. In Section 3.1 the production mechanism of the $l^+l^-\gamma\gamma$ final state, through the reaction $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$, will be reviewed. The study of multiple gauge boson production, due to its connection to non-Abelian gauge theories and EW symmetry breaking, is also important for beyond the Standard Model (BSM) searches. The presence of new physics (NP) phenomena could manifest through departures of the measured $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ cross section with respect to the one predicted by the SM. Any deviation from the EW theory can be parametrised via the anomalous quartic gauge coupling (aQGC) formalism presented in Section 3.2. A brief overview of the most recent experimental results concerning the $pp \to Z\gamma\gamma$ production cross section and tests of the aQGC is given in Section 3.3.

3.1 The Zγγ production at the LHC

There are several theoretical motivations for studying multi-boson events at the LHC. Searches for NP signatures often involve the production of multiple particles in the final state, consisting of jets, leptons, photons and missing transverse energy. It comes straightforward that coherent searches of NP require precise knowledge of the SM background processes. Indeed, the $pp \to VV'V''$ process, where the vector bosons ($V = W$, Z or γ) decay leptonically, represents one of the possible multi-lepton backgrounds for BSM searches. Besides, triboson productions are also important as signal processes in their own right, being a new and independent way to test QGC and aQGC [49–51].

The production of a Z boson in association with two photons is one of the many gauge self-interactions predicted by the EW theory. In Section 3.1.1, the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$

process will be analysed from a theoretical point of view, while the state of the art of the theoretical calculation of its production cross section will be presented in Section 3.1.2.

3.1.1 EW gauge boson self-interactions

In quantum electrodynamic no interactions between the gauge bosons (the photons) are allowed. This is reflected by the structure of the gauge-invariant kinetic term of the theory. This is not the case of the more general EW theory where, as can be seen in Equation 1.5, gauge bosons self-interactions are explicitly allowed by the $g\epsilon_{ABC}W^B_\mu W^C_\nu$ term of the gauge boson dynamic evolution $F^A_{\mu\nu}$. Both triple and quartic gauge couplings are allowed by the theory. The triple gauge couplings (TGC) are proportional to the coupling constant g while the QGC to $g^2\epsilon_{ABC}W^BW^C\epsilon_{DEF}W^EW^F$. The case of the QGCs is of particular interest for this work and the allowed vertices can be explicitly written as

$$g_{\mathrm{WWWW}} = g^2,\ g_{\mathrm{WW}\gamma\gamma} = -\mathrm{e}^2,\ g_{\mathrm{WW}\gamma\mathrm{Z}} = -\mathrm{e}g\cos\theta_w,\ g_{\mathrm{WWZZ}} = -g^2\cos^2\theta_w.$$

It has to be noticed that, for both the TGC and QGC, direct couplings of the photon and the Z boson, such as the ZZγγ vertex, are explicitly forbidden, at leading-order (LO), by the SM. This is a consequence of the fact that both the photon and the Z boson are electrically neutral. The ZZγγ final state could arise, in proton proton collisions, as a result of initial-state radiated photons (ISR), emitted by the colliding quarks, of final-state radiated photons (FSR), emitted by the leptons from the Z boson decay, or by a mixture of the two processes (ISR + FSR). Some of the LO elementary processes contributing to the $\mathrm{pp} \to \mathrm{Z}\gamma\gamma \to \mathrm{l}^+\mathrm{l}^-\gamma\gamma$ channel are shown in Figure 3.1. In the same figure, an example of a possible aQGC, involving the Z boson and photons, is shown. As can be seen, the $\mathrm{pp} \to \mathrm{Z}\gamma\gamma \to \mathrm{l}^+\mathrm{l}^-\gamma\gamma$ production cross section is proportional, to the quartic power of the EW coupling constant α_{EW}. Since α_{EW} is significantly less than 1, the process is naturally suppressed.

3.1.2 State of the art theoretical calculations

The complete next-to-leading-order (NLO) QCD theoretical calculation of the $\mathrm{pp} \to \mathrm{Z}\gamma\gamma \to \mathrm{l}^+\mathrm{l}^-\gamma\gamma$ production cross section, including off-shell effects, Higgs resonances and leptonic decays, was first computed in 2011 [52]. Contributions from virtual diagrams, such as the one in Figure 2.3 (right), and of real emission by the incoming quark lines occur at NLO QCD. To avoid infrared divergences real and virtual contributions are summed together. The NLO real corrections are obtained, from the LO diagrams in Figure 3.1, by attaching a gluon to a quark line in every possible way, while the NLO virtual corrections are obtained by inserting gluon lines between the incoming quarks.

The NLO QCD corrections are particularly large for the $\mathrm{pp} \to \mathrm{Z}\gamma\gamma \to \mathrm{l}^+\mathrm{l}^-\gamma\gamma$ process and the obtained differential K-factor, defined as

$$K = \frac{\mathrm{d}\sigma_{NLO}/\mathrm{d}p_T^\gamma}{\mathrm{d}\sigma_{LO}/\mathrm{d}p_T^\gamma}$$

where the theoretical cross sections $\sigma_{(N)LO}$ is calculated in a phase space similar to the one used in this work (see Section 6.2), corresponds to approximately 50%.

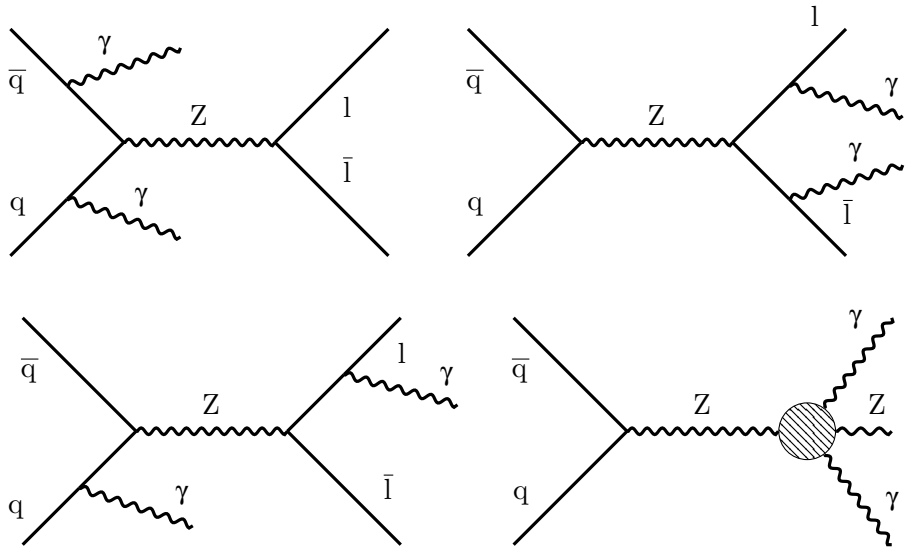


Figure 3.1: Feynman diagrams representing the EW tree level topologies contributing to the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ channel: ISR by the incoming quarks (top left); FSR by the outgoing leptons (top right); mixed ISR plus FSR channel (bottom left); anomalous quartic gauge coupling vertex (bottom right).

Moreover, the size of the NLO QCD correction exhibits a strong dependence on the phase space regions of the photon transverse momentum. For precise measurements of the process both the total cross section and all the kinematical distributions are needed at NLO: this an essential requirement because, unless the differential QCD corrections are properly included in experimental analyses, an excess in a particular region of the phase space, when comparing simulated events with data, could easily be misinterpreted as BSM physics [51]. A dedicated fully exclusive NLO parton Monte Carlo (MC) simulation of the $pp \to Z\gamma\gamma$ process is now available in the automated MC generator MADGRAPH5_aMC@NLO which is used for this work.

Reaching NLO QCD accuracy does not necessarily imply the need for NLO EW corrections, since at EW scale $\alpha_s \gg \alpha_{EW}^4$. Indeed, NLO EW corrections are available only for a limited number of multi-boson processes and become important only when the experimental uncertainty on the production cross section is reduced to the level of few per cent.

3.2 Anomalous gauge couplings through the Zγγ production process

The discovery of a particle compatible with the SM Higgs boson corroborates the validity, at the presently available energy scales, of the $SU(2) \times U(1)$ gauge symmetry group to be able to fully describe the EW interactions. Possible deviations from the SM in the multi-boson production are expected to manifest at a very high mass scale and can be parametrised by operators invariant under the $SU(2) \times U(1)$ symmetry

without introducing any additional source of electroweak symmetry breaking (EWSB). In this analysis, operators that lead to aQGCs arising from the Effective Field Theory framework described in Section 1.3.1 were considered. Dimension 8 operators are the first ones that give rise to quartic couplings among the neutral gauge bosons [53]. There are three classes of genuine aQGC dimension 8 operators [54]: operators containing only the Higgs kinetic term ($\mathcal{O}_S$), operators containing the Higgs kinetic term and two field strength tensors ($\mathcal{O}_M$) and those containing only field strength tensors ($\mathcal{O}_T$). The latter affect several multi-boson topologies, as can be seen in Table 3.1, and are the ones specifically probed in this work. The explicit form of the $\mathcal{O}_T$ operators is given in the following

$$\begin{aligned}
\mathcal{O}_{T,0} &= \mathrm{Tr}\left[\hat{W}_{\mu\nu}\hat{W}^{\mu\nu}\right] \times \mathrm{Tr}\left[\hat{W}_{\alpha\beta}\hat{W}^{\alpha\beta}\right] \\
\mathcal{O}_{T,1} &= \mathrm{Tr}\left[\hat{W}_{\alpha\nu}\hat{W}^{\mu\beta}\right] \times \mathrm{Tr}\left[\hat{W}_{\mu\beta}\hat{W}^{\alpha\nu}\right] \\
\mathcal{O}_{T,2} &= \mathrm{Tr}\left[\hat{W}_{\alpha\mu}\hat{W}^{\mu\beta}\right] \times \mathrm{Tr}\left[\hat{W}_{\beta\nu}\hat{W}^{\nu\alpha}\right] \\
\mathcal{O}_{T,5} &= \mathrm{Tr}\left[\hat{W}_{\mu\nu}\hat{W}^{\mu\nu}\right] \times B_{\alpha\beta}B^{\alpha\beta} \\
\mathcal{O}_{T,6} &= \mathrm{Tr}\left[\hat{W}_{\alpha\nu}\hat{W}^{\mu\beta}\right] \times B_{\mu\beta}B^{\alpha\nu} \\
\mathcal{O}_{T,7} &= \mathrm{Tr}\left[\hat{W}_{\alpha\mu}\hat{W}^{\mu\beta}\right] \times B_{\beta\nu}B^{\nu\alpha} \\
\mathcal{O}_{T,8} &= B_{\mu\nu}B^{\mu\nu}B_{\alpha\beta}B^{\alpha\beta} \\
\mathcal{O}_{T,9} &= B_{\alpha\mu}B^{\mu\beta}B_{\beta\nu}B^{\nu\alpha},
\end{aligned} \tag{3.1}$$

where $W^i_{\mu\nu}$ is the $SU(2)$ field strength, $B_{\mu\nu}$ the $U(1)$ one and all operators are charge conjugation and parity conserving. It has to be noticed that the two last operators, $\mathcal{O}_{T,8}$ and $\mathcal{O}_{T,9}$, give rise to aQGCs containing only the neutral electroweak gauge bosons. The $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ process is expected to be particularly sensitive to them.

Operators	WWWW	WWZZ	WWγZ	WWγγ	ZZZZ	ZZZγ	ZZγγ	Zγγγ	γγγγ
$\mathcal{O}_{T,0}, \mathcal{O}_{T,1}, \mathcal{O}_{T,2}$	✓	✓	✓	✓	✓	✓	✓	✓	✓
$\mathcal{O}_{T,5}, \mathcal{O}_{T,6}, \mathcal{O}_{T,7}$	✗	✓	✓	✓	✓	✓	✓	✓	✓
$\mathcal{O}_{T,8}, \mathcal{O}_{T,9}$	✗	✗	✗	✗	✓	✓	✓	✓	✓

Table 3.1: Dimension 8 operators containing only the EW field strength tensors and the quartic vertices they affect [24]. The ZZγγ quartic vertex is the one studied in this work.

3.3 Previous results in this field

The production of Zγγ triboson final state has been observed for the first time in e^+e^- collisions at LEP [55, 56]. The process was found to be well described by the SM predictions and limits were placed on aQGC parameters. During the LHC Run 1, both the CMS and ATLAS Collaborations have studied the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production using an 8 TeV data sample with an integrated luminosity of $20\,\mathrm{fb}^{-1}$ and provided the first cross-section measurement for this process, achieving a significance greater than 5σ [57, 58]. The observed production rate was found to be consistent with the theoretical

NLO QCD prediction and no evidence of a large BSM contribution was found. Limits on the aQGCs were placed by ATLAS using an exclusive fiducial cross section with high diphoton invariant mass requirements. No limits on the aQGCs were placed by the CMS experiment using this channel.

As can be seen in Table 3.1, the $\mathcal{O}_T$ operators studied in this work affect different multi-boson topologies. Figure 3.2 illustrates the constraints on the aQGC parameters, corresponding to the dimension 8 operators $\mathcal{O}_T$, obtained from other studies which use alternative production and decay channels. The most sensitive measurements are those performed in the WW and in the ZZ channels by the CMS Collaboration.

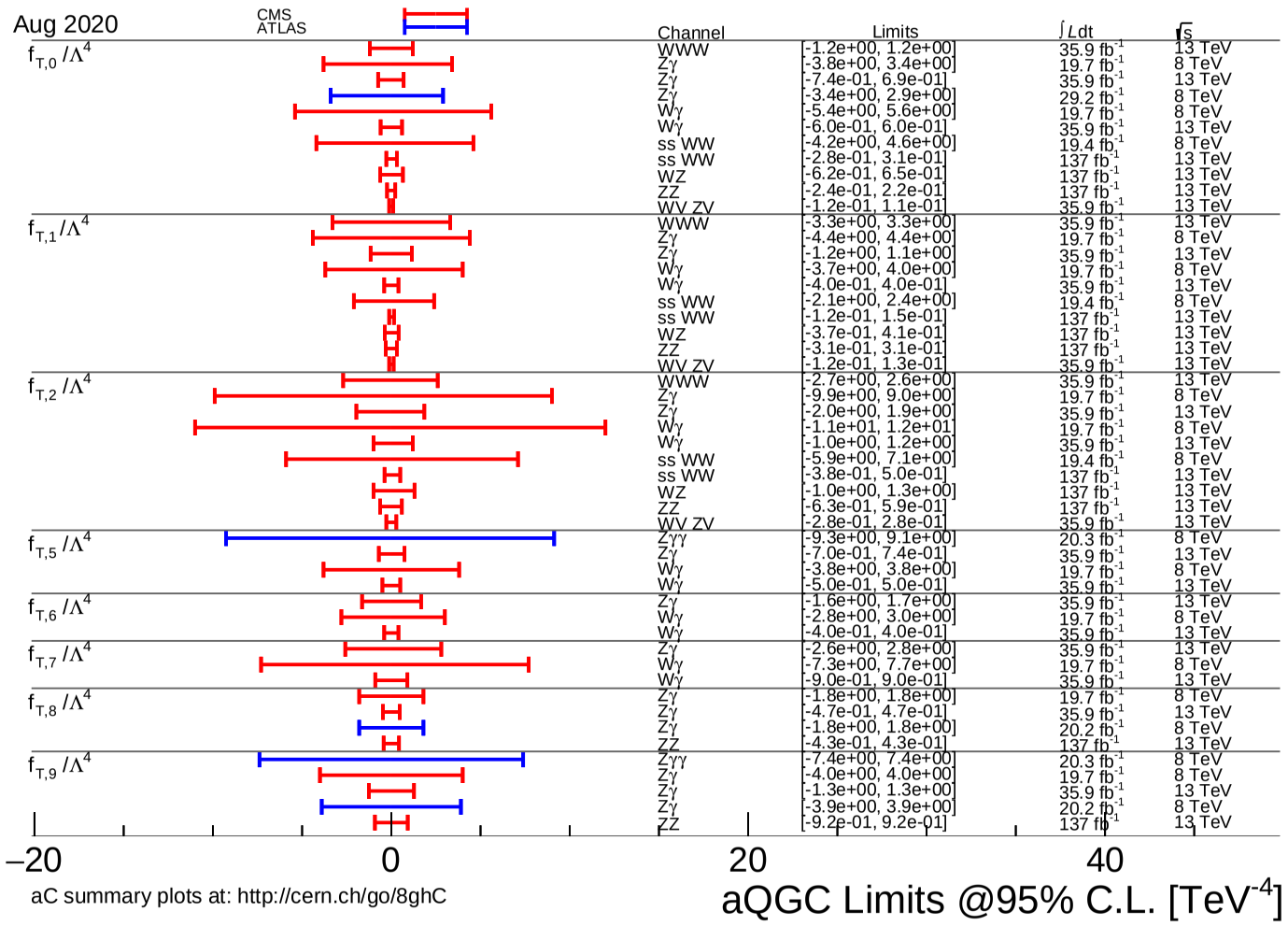


Figure 3.2: Constraints on the parameters $f_{T,i}/\Lambda^4$, corresponding to the dimension 8 operators $\mathcal{O}_T$, obtained by the CMS and ATLAS Collaborations during the Run 1 and 2 of the CERN LHC.

4 | The CMS detector at the LHC

In this work, events produced in high energy proton proton collisions at the world's largest and most powerful accelerator complex ever built, the Large Hadron Collider (LHC), are studied. The main features of the LHC are described in the following, with a particular focus on the Compact Muon Solenoid (CMS) experiment, one of the main experiments along the accelerator ring.

The chapter is organised as follow. In Section 4.1 a brief description of the European Laboratory for Particle Physics (CERN) is given. In Section 4.2 the CERN accelerator complex is described with a particular focus on its last step, the LHC. Finally, in Section 4.3 the CMS experiment it is described focusing on its relevant characteristics.

4.1 The European Laboratory for Particle Physics

The birth of the European Laboratory for Particle Physics was the result of the common effort of the European countries to stop the brain drain after the Second World War. Nowadays it is one of the biggest laboratories in the world by involving more than twenty countries and ten thousand cooperating people in particle physics research. CERN aims to offer a purely civil research facility by providing the environment and the fundamental tools necessary for particle physics research activity. It is also committed to give public access to every scientific result produced.

The organisation was founded in 1954 while the CERN Convention was signed later on in 1963. The construction of the first building, at the border between France and Switzerland, has begun in 1964. Five accelerators were built during the years: each one of them, after having accelerated the particles to the requested energy, injects the particles into the next one in the chain. As a result, from 1954 to 2018, the maximum centre of mass energy achievable at the CERN facilities has been raised from 50 MeV up to 13 TeV.

Since 1964 CERN has allowed nations around the world to communicate through the language of science. It played a central role in allowing mankind to answer many fundamental open questions in particle physics field, such as particles behaviour and their interactions.

4.2 The Large Hadron Collider accelerator

The Large Hadron Collider accelerator [59] is a superconducting accelerator and collider, the largest and most powerful hadron collider ever built. The project to build

a 14 TeV accelerator was approved in 1994 and the complex inaugurated on the 21st October 2009. It was installed by making full use of the preexisting tunnel and injection chains occupied by the Large Electron Positron Collider (LEP) [60]. New facilities and interaction points were added as well. The LHC tunnel has a circular shape with a total length of 27 km and it is located across the France-Swiss border, 45 m to 170 m below the surface.

The LHC design is based on some quite simple principles but employs the latest and cutting edge technology available. Being a particle particle collider it consists of two rings to accelerate and store two counter-rotating beams of particles. Due to lack of available space for two separate rings in the LEP cavern, LHC adopts a twin-bore magnet design [61], which is also one of the main distinctive features of the collider design. Superconducting magnet technology is employed to maintain the particle beam direction and collimation along the LHC circumference.

LHC is designed to collide beams of particles in four stations, called interaction points (IP). In each of them different detectors, able to collect and analyse the results of the beams collisions, are installed. CMS [62] and ATLAS [63] are general purpose detectors designed to cover the largest possible spectrum of energy allowed by LHC. Both of them were specifically designed for the Higgs boson detection but also to be sensitive to a wide range of possible new beyond the Standard Model phenomena. The idea to build two same purpose but differently designed detectors were found to be necessary to allow for a cross-check of any possible discovery, being CMS and ATLAS the only experiments available to study such a wide range of particles interactions at such high energies. The other two crossing points, previously occupied by the LEP experiment, host LHCb [64], a forward detector designed to study CP violations in the EW theory and the asymmetry of matter and anti-matter in our Universe, and ALICE [65], mainly designed to study collisions between heavy ions, in order to probe the condition of our Universe few seconds after the Big Bang.

The decision to build a proton collider after the LEP experience was mainly driven by the necessity to have a machine dedicated both to precision studies of the Standard Model and to the discovery of new physics. Indeed, the main advantages of a hadron collider with respect to a lepton are that:

- protons are heavier than electrons so they lose less energy by synchrotron radiation while circulating in the collider. Consequently, a higher energy is achievable in the centre of mass (CM) frame. LHC allows to test the widest mass range possible from few GeV to the TeV level;

- protons are composite so the CM energy is not fixed and a wide energy spectrum can be explored simultaneously in every collision. The additional production of many low-energy particles results in a complex experimental environment where clever detector design and particle reconstruction algorithm are needed.

4.2.1 Acceleration complex

LHC can deliver either bunches of protons, colliding at 13 TeV in the centre of mass frame, or heavy ions (typically lead nuclei) colliding at 2.76 TeV/nucleon in the centre of mass frame. LHC is only the last step of a succession of different sub-accelerators,

used in other CERN experiments in the past. The acceleration mechanism is divided into multiple stages and each machine is fed with a beam of particles which are boosted and injected in the next accelerator in the chain.

Protons are produced from hydrogen gas by stripping the electrons from the gas nuclei. This task is accomplished by applying an intense electric field through a device called *Duoplasmatron*. The proton beam, pulsed by the Duoplasmatron, is then fed into the first accelerator stage called *Linac 2*. This 36 m linear accelerator exploits radiofrequency cavities to charge cylindrical conductors placed along its length. The protons are pushed and pulled by the alternately charged cylindrical conductors into one direction until they have reached an energy of 50 MeV. The proton beam coming from the Linac 2 is then passed to the next accelerator, the *Proton Synchrotron Booster*, which consists of four superimposed rings for a total circumference of 157 m. Here the beam is divided into four bunches, each of them boosted up to 1.4 GeV and then recombined before being injected into the *Proton Synchrotron*. The Proton Synchrotron has a circumference of 628 m and it is able to deliver particles bunches interspersed by 25 ns (40 mHz) with an energy of 25 GeV to the *Super Proton Synchrotron*, a 7 km circular accelerator that raises the energy up to 450 GeV. The Super Proton Synchrotron is the last stage before the LHC main ring where two beams of particles, circulating in the opposite directions, are formed and then accelerated to the final energy of interest. A detailed description of the pre-injector complex for the LHC is given in reference [66] while a schematic representation of the whole CERN accelerator complex is shown in Figure 4.1.

4.2.2 Structure overview and machine performances

The primary design considerations of a particle collider are the nature of the colliding particle, the centre of mass energy of the collision and the rate of collisions which is able to provide. Particle interactions, being quantum processes, are subjected to an intrinsic stochastic nature. For each process of interest, such as the $\mathrm{pp} \to \mathrm{Z}\gamma\gamma$ production (the topic of this work), the number of events per second generated by LHC can be calculated as

$$\frac{\mathrm{d}N_{\mathrm{Z}\gamma\gamma}}{\mathrm{d}t} = \mathcal{L}_{inst.} \cdot \sigma_{\mathrm{Z}\gamma\gamma}$$

where $\sigma_{\mathrm{Z}\gamma\gamma}$, the cross section of the process of interest, is related to the probability of producing a Z boson alongside two photons in a single proton proton collision, while $\mathcal{L}$, the machine instantaneous luminosity, measures the ability of a particle accelerator to produce the required number of interactions. If the process of interest is rare a high luminosity is crucial in order to be able to collect a decent amount of events to be analysed. One of the advantages of using proton proton beams instead of proton antiproton beams, as it was previously done at Tevatron [68], is the possibility to provide to the experiments unprecedented high luminosity beams. This is determined by the fact that protons, contrary to their antiparticles, are easier to be produced and stored in large quantities. The LHC is able to spread 2800 bunches of 10^{11} particles each along the main ring, separated by 25 ns and with a maximum energy of 6.5 TeV per beam. The instantaneous luminosity of the collider depends only on the beam characteristics

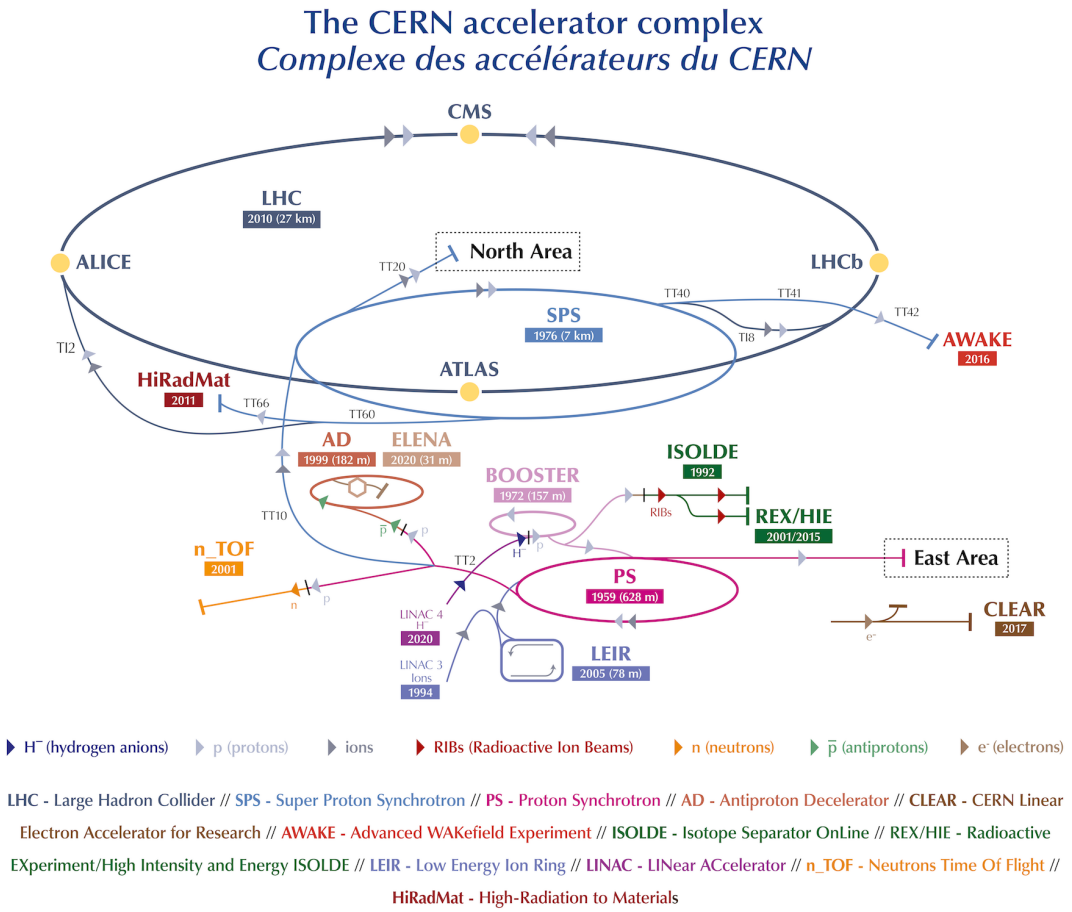


Figure 4.1: The CERN accelerators complex that delivers protons bunches to the LHC experiments. The LHC is the last ring (dark blue line) while smaller machines are used in a chain to help boost the particles to their final energies and to provide beams to a whole set of smaller experiments. Figure adapted from reference [67].

and can be written as

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} \cdot F \tag{4.1}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalised transverse beam emittance and β^* the beta function at the interaction point. Finally, F represents a geometric factor related to the beam crossing angle at the interaction point.

In order to span the widest mass region possible, the LHC increases the proton momentum to the limit of the current technology by using cyclic electrical fields. While the particles gain momentum high magnetic fields are used to maintain their trajectories along the synchrotron ring. The bunches of particles travel inside two parallel tubes, under strong vacuum. As can be seen in Figure 4.2, the tubes are arranged into a semi-circular shape composed of eight arcs called *octants*, in which are located the

bending dipole magnets, and eight long straight sections where particles are injected and Radio Frequency cavities (RF) or experiments are hosted.

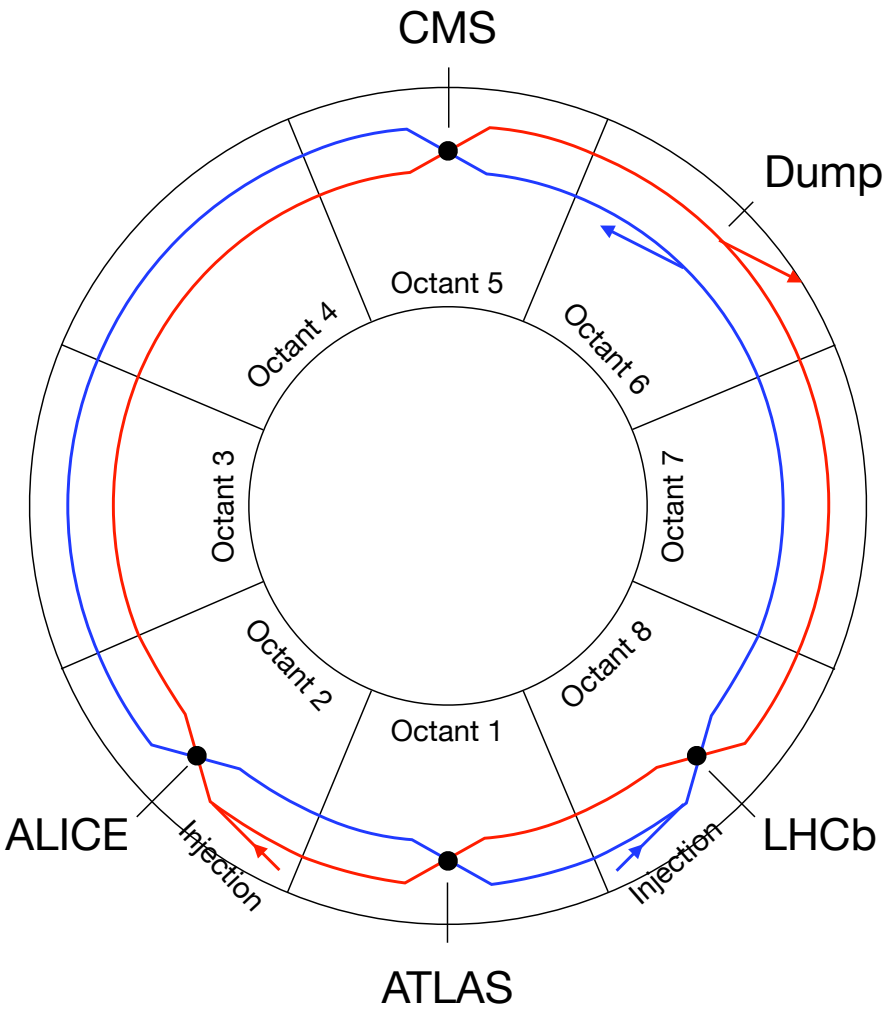


Figure 4.2: The structure of the LHC octants. Straight sections are use for inject/dump the beams. The positions of the experiments and other facilities are highlighted. Figure adapted from reference [59].

The LHC magnet system is at the core of the accelerator apparatus. More than 9600 superconducting magnets are distributed around the LHC circumference. Each octant contains 14.3 m long dipole magnets, which generate a magnetic field of 8.33 T to bend the particles trajectories along the circumference, plus hundreds of focusing and defocusing quadrupoles to maintain the beam orbit stable. A small amount of higher order multipoles gives additional minor corrections to the beam direction. Each superconducting magnet consists of Niobium-Titanium (NbTi) cables cooled at the unprecedented low temperature of 1.9 K. An important drawback in operating at such low temperature consists in a drastic reduction of the minimum energy deposition into

the magnet system able to trigger a magnet quench. This phenomenon is particularly dangerous being the NbTi superconducting only at very low temperatures. To prevent the magnet from burning an optimal heat dissipation and a tighter control of the cable movements are necessary. The twin-bore design adopted by LHC meets the need to have a stiff magnet structure by accommodating the winding for the two-beam channels in a common cold mass and cryostat. This peculiar design makes the magnet structure complicated, by coupling the dipoles both mechanically and magnetically, but was also motivated by the space limitation of the tunnel and the need to reduce the total costs.

The injected beam is captured and accelerated by sixteen superconducting RF cavities with a carrier electromagnetic wave of 400 MHz and a generated electric field of 5.5 MV/m. Longitudinal injection errors are damped by the same system while transverse injection errors are controlled by a separated one. All the RF accelerating cavities and beam control systems are located at Point 4, in the cavern previously occupied by the ALEPH experiment [69]. The RF accelerating system is able to provide to each proton beam 485 keV/revolution, and it takes about twenty minutes to reach the 7 TeV/beam energy. After the acceleration phase, the RF cavities are used only to supply the energy lost by synchrotron radiation which amounts to approximatively 7 keV/revolution. During the acceleration phase, the power delivered to each beam is relatively small but the one needed to control the beam direction is much larger. In order to reduce the overall power consumption, the two opposite circulating beams cannot be controlled by the same system. Instead, the beams separation is increased using dedicated dipole magnets so that each of them has its dedicated RF control system.

A dedicated vacuum system was developed in order to allow for the magnet system and the RF accelerating system to properly operate at superconducting temperatures. The LHC vacuum system consists of three subsystems: the insulation vacuum system for the cryomagnets, the insulation vacuum system for the helium distribution system and, finally, the vacuum system for the beam. Each subsystem is subdivided into subsections to provide easy access to each part, mainly for maintenance reasons. The subsections are divided using vacuum barriers and linked together by vacuum valves. Different vacuum performances are required for the three subsystems. The insulation and cryomagnets subsystems are required to provide a moderate vacuum, up to 10^{-6} mbar, while the requirements for the beam vacuum are much more stringent. Indeed, the stability of the protons beams against minor spurious collisions with other particles is achieved circulating the particles in a strong vacuum, up to 10^{-10} mbar.

4.3 The Compact Muon Solenoid experiment

The Compact Muon Solenoid experiment is one of the detectors that operate at the CERN LHC. It shares with the ATLAS experiment the same scientific goals, which after the Higgs boson discovery include precision study of the Higgs boson field as well as searches for physics beyond the Standard Model. The CMS and ATLAS experiments adopt different technical solutions, especially for what concerns the magnetic system design.

The CMS detector is installed about 100 m underground in the France countryside, close to the village of Cessy. Its prime design principle was the ability to shed light on the electroweak symmetry breaking mechanism, for which the Higgs boson was presumed to

be responsible. This was achieved in 2012 by the discovery and characterisation of a new scalar particle which was proven to be consistent with predicted SM Higgs boson [14].

Nowadays, the experimental studies of the Higgs boson mechanism and properties are of particular interest, being them able to probe the consistency of the SM at energy scales up to 1 TeV. This scale represents the energy frontier for high energy physics studies above which new phenomena, able to explain the SM pitfalls or to extend its predictions, could take place. The CMS detector was specifically designed to be able to investigate any hint of new symmetries, new particles or new forces at this unprecedented high energy scale.

The CMS detector requirements include the ability to operate in the harsh environment represented by the high centre of mass energy and luminosity delivered by the LHC. The main issue to cope with is the high event rate of particle collisions at which the detector is exposed. Indeed the short time between each bunch crossing and the great number of simultaneous interactions per beam crossing result in a huge rate of events that need to be reconstructed and selected in the shortest time possible. A dedicated trigger system was developed for such purpose and specifically design characteristics, such as high detector granularity and good time resolutions, were implemented in each CMS sub-detector to assure the lowest possible response time and occupancy. Another important issue is represented by the large flux of high energy particles coming from the interaction point which requires detectors and front end electronics to be particularly radiation-hard.

The CMS detector characteristics can be summarised as follows:

- a good muon and dimuon identification and momentum resolution over a wide range of momenta and angles plus the ability to determine unambiguously the charge of muons with momenta up to 1 TeV;
- a good identification and characterisation of charged particles with a short but measurable decay time by the tracking system. This includes efficient triggering and offline tagging of τ leptons and b-jets;
- a good electromagnetic energy resolution, including a diphoton and dielectron mass resolution up to 1% at 100 GeV, a wide geometric coverage, a good π^0 rejection and an efficient photon and lepton isolation from other charged and neutral particles especially at high luminosities;
- an acceptable hermetic geometric coverage and fine lateral segmentation to ensure good missing transverse energy measurements and dijet-mass resolution.

In the next sections, the CMS experiment design is presented. As implied by the experiment's name the main distinguishing features are the high-field superconducting solenoid and the muon detection system.

4.3.1 Structure overview

The CMS detector is a 21.6 m long cylinder with a diameter of 15 m and a weight of 12 500 t. It is designed to study proton proton collisions at a centre of mass energy of 14 TeV and luminosities up to $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The most impressive constructive element is represented by a huge solenoidal superconducting magnet that surrounds all the inner

detectors, with the only exception of the muon chamber on the outside region. The large superconducting bore hosts a silicon pixel and strip tracker, a lead-tungstate crystal Electromagnetic Calorimeter and a brass scintillator sampling Hadronic Calorimeter. The superconducting solenoid is enclosed into an iron yoke instrumented with the muon detector. A quasi-complete acceptance hermeticity is guaranteed by additional forward calorimeters.

The dodecagonal base prism structure of CMS is divided into two main regions: the barrel represents the central part of the prism while two endcaps, on both side, ensure the maximum tightness. The coordinate system used is a right-handed Cartesian one (x,y,z), with the origin fixed in the interaction point inside the detector. The z-axis defines the beam line direction, the y-axis points upward to the surface while the x-axis points to the centre of the LHC ring. Since the CMS detector has a cylindrical symmetry it is also useful to use the cylindrical coordinates (r,ϕ,θ) shown in Figure 4.3 to identify any point inside of it. The r-coordinate defines the distance from the z-axis,

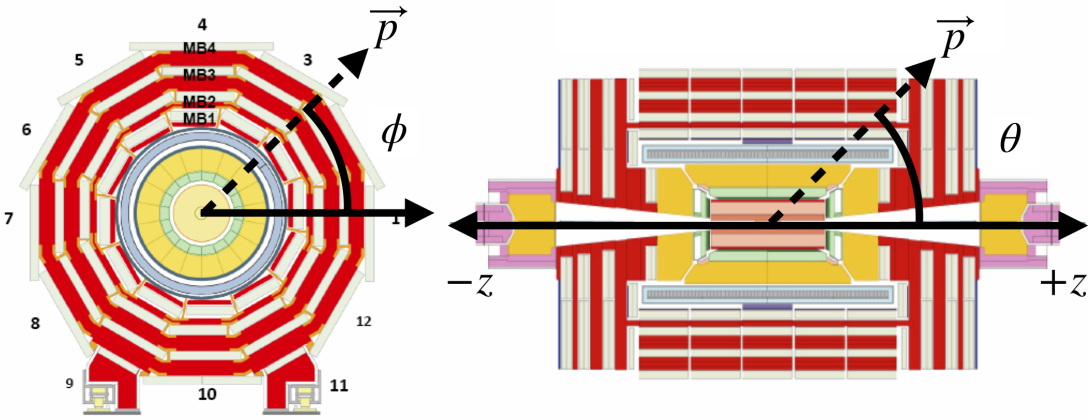


Figure 4.3: A schematic view of the polar coordinates used at CMS. Figure adapted from reference [70].

the ϕ-coordinate represents the azimuthal angle in the xy plane and θ is the angle between the particle three-momentum vector $\vec{p}$ and the positive direction of the beam axis. The momentum and energy transverse to the beam direction are denoted by p_T and E_T respectively. The transverse momentum can be computed from the p_x and p_y three-momentum vector coordinates as

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

The imbalance of energy measured in the transverse plane is denoted by E_T^{miss}. In experimental particle physics θ is often replaced by the pseudo-rapidity, defined as

$$\eta = -\ln\left[\tan\frac{\theta}{2}\right].$$

In the limit where a particle is travelling close to the speed of light (or when the mass of the particle is negligible), the pseudo-rapidity converges to the rapidity usually defined

in relativity. The main reasons to prefer the pseudo-rapidity over the polar angle θ are that the particle production is constant as a function of η and differences in rapidity are Lorentz invariant under boosts along the beam axis. This definition turns out to be extremely useful when dealing with hadron colliders where the colliding partons carry different longitudinal momentum fractions.

The CMS structure is shown, from inside to outside, in Figure 4.4 and summarised below:

1. the interaction point sits at the core of the apparatus. In this region interactions between bunches of particles take place. This work is based on data taken during the full LHC Run 2 with a maximum instantaneous luminosity of $21\,\mathrm{Hz\,nb^{-1}}$ achieved in 2018. The consequence of such high collision rate is reflected in a large number of superimposed interactions at the same time. The pileup affects the reconstruction of the particles tracks and is overcomed by providing a good time resolution and high granularity for all the sub-detectors;

2. the silicon tracker represents the inner part of CMS detector. It is an all-silicon detector which extends from the interaction point up to $|r| < 1.2\,\mathrm{m}$, covering a pseudo-rapidity range $|\eta| < 2.5$. Silicon pixel vertex detectors surrounded by silicon strip detectors able to reconstruct with high precision particles tracks and vertices are distributed over an active area of $215\,\mathrm{m}^2$;

3. the Electromagnetic Calorimeter (ECAL) extends over $1.2\,\mathrm{m} < r < 1.8\,\mathrm{m}$ by covering an angle $|\eta| < 3$. It is composed of scintillating crystals of lead-tungstate ($PbWO_4$) able to detect the energy and impact position of photons and electrons.

4. the Hadronic Calorimeter (HCAL) extends over $1.8\,\mathrm{m} < r < 2.9\,\mathrm{m}$ by covering the pseudorapidity range $|\eta| = 5$. Unlike ECAL, it is a sampling calorimeter composed of brass layers alternated with plastic scintillators able to detect charged and neutral hadrons;

5. the superconducting solenoidal magnet provides a 3.8 T internal constant magnetic field, along the beam direction, which is used to bend the charged particles trajectories in order to measure their momenta. The solenoid surrounds all the previous sections and extends from $2.9\,\mathrm{m} < r < 3.8\,\mathrm{m}$ over $|\eta| < 1.5$;

6. the muon system, embedded in the magnet return yoke, is placed in the outermost part of the experiment. It extends from $4\,\mathrm{m} < r < 7.4\,\mathrm{m}$ over $|\eta| < 2.4$, surrounded by a residual 1.8 T magnetic field. It is composed of drift tubes detectors in the barrel and cathode strip chambers detectors in the endcaps. Additionally, resistive plate chambers detectors are used in both regions. The muon system is able to reconstruct with high precision the tracks produced by muons.

4.3.2 Tracker

The tracker is the closest detector to the beams interaction point. It is specifically designed to provide the highest possible resolution measurement of the trajectories of charged particles with transverse momentum above 1 GeV in a pseudorapidity range

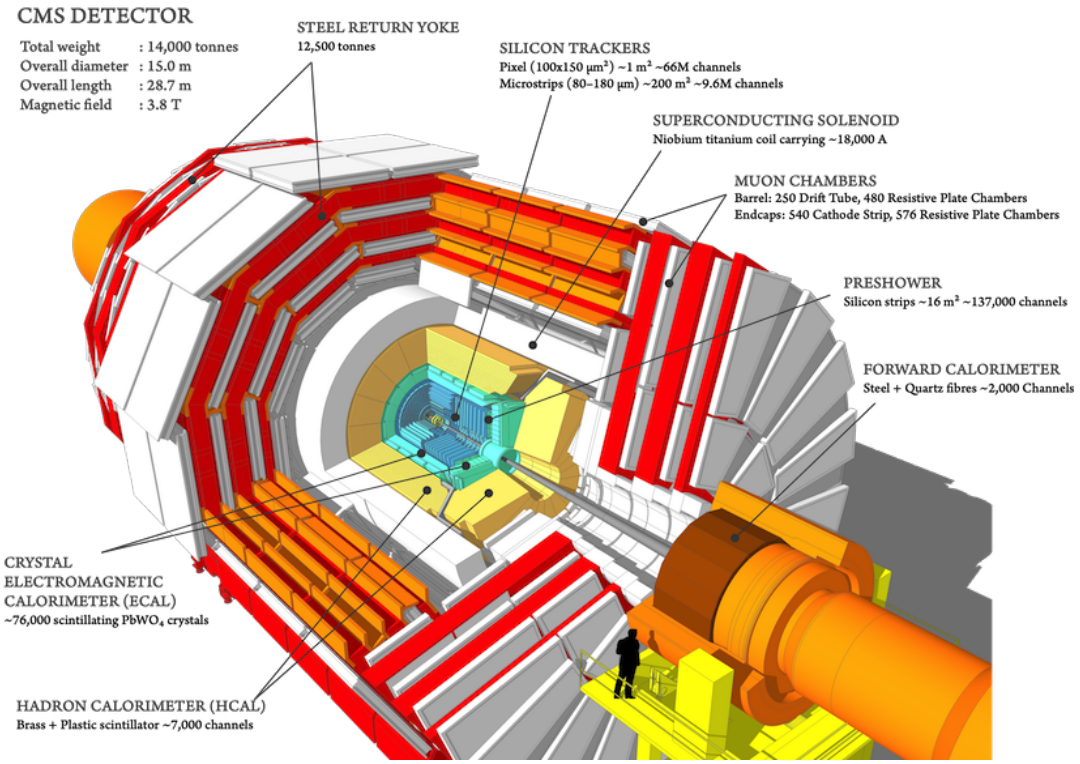


Figure 4.4: A schematic view of the structure of the CMS experiment. Figure adapted from reference [71].

up to $|\eta| = 2.5$. Furthermore, it is required to provide good discrimination between primary and secondary vertices in events containing short mean-life particles (such as heavy-flavour hadron ones).

The tracker plays a major role in the identification of electron and muons, along with ECAL and the muon system, and it is also fundamental to identify the tau leptons decays. The tracker information is also heavily exploited in the important task of the event triggering by providing valuable information to reduce the event rate originating from the LHC bunch crossing (see Section 4.3.7).

During LHC Run 2 the tracker experienced an average of 1000 tracks per bunch crossing (i.e. every 25 ns) from more than 20 quasi-simultaneous proton proton collisions. In order to operate in the LHC harsh environment the tracker was built with the following key characteristics:

- a low occupancy, needed to cope with the high hit rate density, is achieved by providing the highest possible granularity in (η, ϕ) thus by enhancing the resolution of the detector itself. A drawback of this design is the high power density of the detector electronics which thus need an efficient cooling system;
- a limited amount of material, needed to avoid compromising the trajectory measurements. Indeed an excessive amount of material gives rise to multiple scattering, bremsstrahlung, photon conversion or nuclear interactions that potentially hamper the accuracy of the hits points positions in the tracker sections;

- a high radiation hardness of the detector ($10^{15}\,n_{eq}cm^{-2}$), expected to sustain the radiation damage induced by the high particle flux at the LHC interaction region.

The tracker is a full silicon detector made by a cylinder 5.8 m long with an outer diameter of 2.5 m. To have maximum tightness it covers both the barrel and endcaps regions and it is optimised for collecting 12–14 hits per track. Its construction design can be divided in two sub-detectors, as can be seen in Figure 4.5:

1. the radial region between 4.4 cm and 10.2 cm is occupied by 1440 hybrid pixel detector modules arranged in three cylindrical layers, complemented by two disk of pixel modules on each endcap side. The pixel detector is essential for the high precision reconstruction of each charged particle tracks as well as for the reconstruction of primary and secondary vertices;

2. the radial region between 20 cm and 116 cm hosts 15 148 modules of the silicon strip detector arranged in ten barrel detector layers and complemented by twelve disks on each endcap side attached to a carbon fibre support.

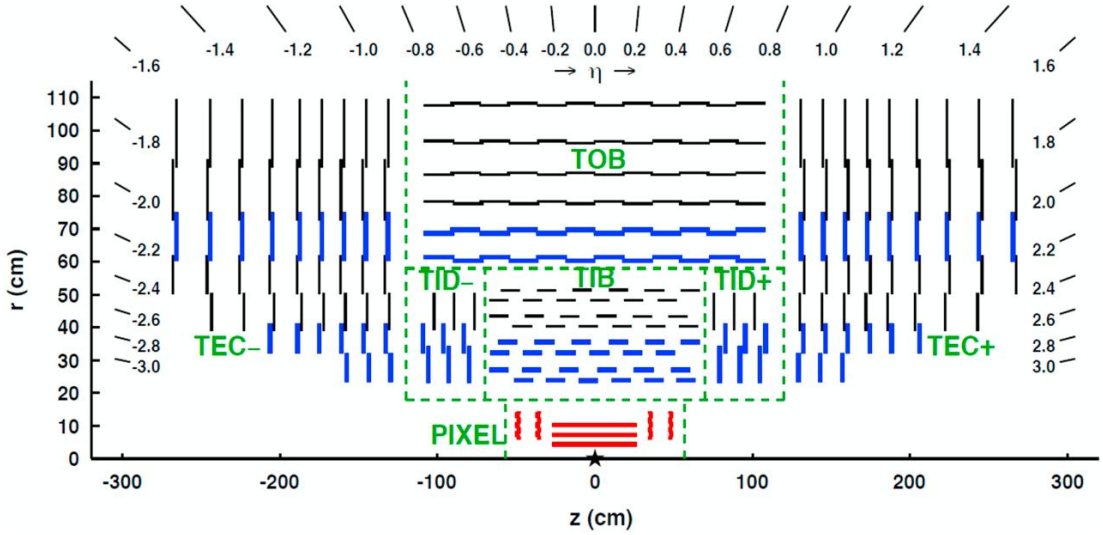


Figure 4.5: A schematic cross section through the tracker detector. Each line represent a detector module.

Once every trajectory is properly associated with the corresponding charged particle it is possible to reconstruct the particle momentum from the track curvature itself. The tracker performances allow for a momentum resolution for muons with transverse momentum up to 100 GeV and $|\eta| < 1.5$ of about 2% and a vertex association accurate to 10 µm.

4.3.3 Electromagnetic calorimeter

The Electromagnetic Calorimeter (ECAL) is a hermetic homogenous calorimeter designed to detect and measure with great precision the energy of electrons, positrons and photons originating from the LHC interaction point. The major driving criteria in

the ECAL design are the capability to detect the decay into two photons of the Higgs boson. The ECAL detector is then required to provide a fast signal response, a high granularity and a good energy resolution. It is also required to be radiation resistant to cope with the extremely harsh LHC environment. A detector with such characteristics is obtained exploiting the properties of high-density lead-tungstate ($PbWO_4$) crystals. The choice of the $PbWO_4$ as active detector material is motivated by its intrinsic characteristics, namely:

- a high density ($\rho = 8.3\,\mathrm{g/cm^3}$), short radiation length ($\mathrm{X_0} = 0.89\,\mathrm{cm}$) and reduced Moliere radius ($R_m = 2.2\,\mathrm{cm}$), which allow to build a compact and high granulated calorimeter;
- a short decay scintillation time which allows to collect about 80% of the emitted light in 25 ns thus between each LHC bunch crossing. This also allows for a fast particle energy reconstruction reducing any contamination from out of time vertices;
- an intrinsic radiation hardness which allows for the material to operate in the LHC environment for years with relatively modest deterioration in performances due to radiation damages.

Among the disadvantages of using the lead tungstate as active detector material the most important ones are related to the fact that:

- the light yield of the $PbWO_4$ is low and dedicated photo-multipliers are needed the enhance the signal. Avalanche photodiodes (APDs) are used as photodetectors in the barrel while vacuum photo-triodes (VPTs), able to better sustain radiation-induced damages, are used in the endcaps;
- the high flux of ionising particles could produce absorption bands inside the crystals by forming coloured centres. As a consequence, each crystal experience a wavelength-dependent loss of light transmission. This specific damage must be tracked and corrected for independently for each crystal since the radiation flux is highly $|\eta|$ dependent. To do so the optical transparency of the crystals is measured by regularly injecting laser light.

The layout of the calorimeter, showing the arrangement of crystals into the barrel and endcaps regions, is shown in Figure 4.6. The ECAL barrel is formed by 61 200 $PbWO_4$ crystals arranged in a cylindrical shape. It has an inner radius of 1.29 m, a length of 6.3 m and it is composed of 36 super-modules that cover a region of $|\eta| < 1.479$. Each super-module includes 20×85 crystals on the ϕ-η plane, sub-divided into four modules along the η direction. Every module is segmented into different sub modules which consist in 5×2 crystal arrays mounted on a glass fibre structure. In the barrel region the crystal are cut in a truncated pyramidal shape with a total length of 23 cm corresponding to a radiation length of $24.7\,\mathrm{X_0}$, a frontal and rear area of $22 \times 22\,\mathrm{mm^2}$ and $26 \times 26\,\mathrm{mm^2}$ and a granularity of $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ (1°). Crystals are grouped in 5×5 matrices called trigger towers. To avoid cracks aligning with particle trajectories crystal the axes are tilted by 3° with respect to the direction towards the interaction point. Following the η direction the barrel structure stops at $|\eta| = 1.479$ allowing some service

cables to exit the calorimeter area. The barrel region is complemented by 7324 crystals mounted in two endcaps covering a pseudorapidity angle of $1.479 < |\eta| < 3$. Each endcap consists of semicircular halves of aluminium instrumented with crystals shaped differently with respect to the barrel: they are 22 cm long with a frontal and rear area of $28.6 \times 28.6\,\mathrm{mm}^2$ and $30 \times 30\,\mathrm{mm}^2$. Crystals are arranged following a $x - y$ symmetry into 18 super-crystals modules each of them containing 5×5 elementary units.

A preshower detector, engineered with a different detector design, is placed in front of the endcaps crystals by covering a fiducial region of $1.653 < |\eta| < 2.6$. Its main function is to discriminate electromagnetic showers formed by prompt photons, that are isolated photons coming from the primary interaction vertex (i.e. from ISR/FSR), from photons coming from neutral pions decay. It also helps the identification and position determination of electron and photons by exploiting its high granularity. The preshower is a sampling calorimeter made by two layers of thick silicon strip detectors able to identify the particles shower profile and to measure their released energies. The silicon detectors are interspersed by disks of lead converter which triggers the electromagnetic shower of the incident particle.

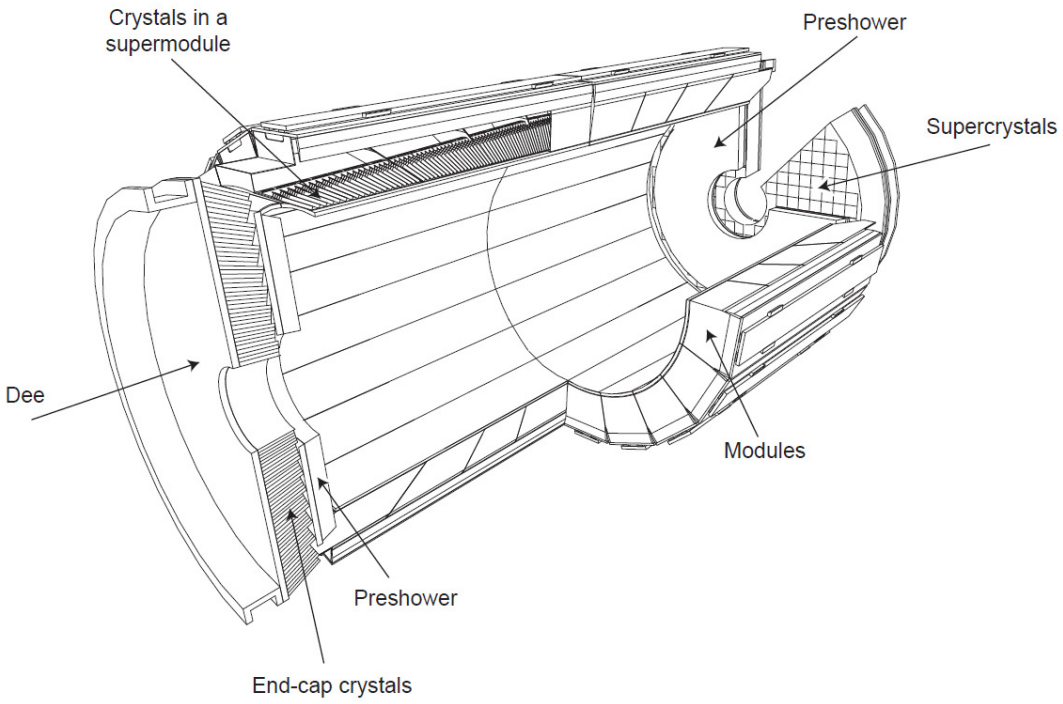


Figure 4.6: View of the main components of the ECAL detector apparatus. Figure adapted from reference [62].

For the range of energies 25 GeV–500 GeV the energy resolution of the ECAL can be parametrised according to [62] as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{4.2}$$

where E is the energy in GeV, a the stochastic term, b the noise term and c the constant term. The stochastic term is dominant at low energies and accounts for the

statistical fluctuations in the number of photoelectrons generated and collected. The noise term gives a contribution that varies with the pseudo-rapidity and the luminosity. It also includes the electronic and pileup noise contributions. Finally the constant term, which is dominant at high energies, describes different phenomena including the stability of the operating conditions, the presence of dead material and so on. The ECAL energy resolution has been measured in various test beams during the initial CMS commissioning [62] and was found to be

$$\frac{2.8\%\text{GeV}^{1/2}}{\sqrt{E}} \oplus \frac{12\%\text{GeV}}{E} \oplus 0.3\%.$$

These values are illustrative of the actual performances obtained during the LHC Run 2 data taking period at which this work is devoted.

4.3.4 Hadron calorimeter

The Hadron Calorimeter (HCAL) is specifically designed to study a wide variety of processes, including the energy and direction of hadronic jets, the reconstruction of the hadron decays and any signature of missing transverse energy in the events. It also measures the shower leakage from the rear of the ECAL, thus making complementary measurements with respect to it. In order to maximise its particle containment property it is built as a hermetic sample calorimeter. As can be seen in Figure 4.7, the calorimeter is divided into four modules, arranged both inside and outside the superconducting solenoid bore, which cover a pseudorapidity range up to $|\eta| < 5$.

The hadron barrel (HB) calorimeter sits between the outer part of the ECAL barrel ($r = 1.77\,\text{m}$) and the inner extent of the superconducting magnetic coil ($r = 2.95\,\text{m}$) covering a pseudorapidity range up to $|\eta| = 1.3$. The hadron calorimeter endcaps (HE) cover the solid angle between $|\eta| > 1.3$ and $|\eta| < 3$. This region contains about 34% of the particles produced by each proton proton collision final states. A sampling calorimeter design is used, for both HB and HE, by alternating absorbers with plastic scintillators as the active material. The readout is performed by hybrid photodiodes photodetector linked to the scintillators via wavelength-shifter optical fibres. The HB is divided into two half-barrels made of 36 identical azimuthal wedges. Each wedge is instrumented with flat steel and brass absorber plates aligned parallel to the beam axis. The wedges are arranged such as the final geometry contains no projective dead material for the full radial extension. The total absorber thickness ranges from 5.82 interaction lengths at $|\eta| = 0$ to 10.6 interaction lengths at $|\eta| = 1.3$. Since the HE is placed at the ends of the solenoidal magnetic field its absorber is made of brass, a non-magnetic material. The absorber is designed to minimise any crack between HB and HE. The HE is radially divided into 14 rings which accommodate the scintillators arranged into a trapezoidal shape. The outer calorimeter (HO), placed in the central pseudorapidity region $|\eta| < 3$ outside the superconducting solenoid, is used to improve the measurement of high energetic showers that can pass through EB and HB. The HO consists of several layers of plastic scintillators, with the same granularity of HB, while the superconducting solenoid acts as an additional absorber. The effective thickness is approximately 10 λ_I. The forward calorimeter (HF) is the last piece of HCAL. It is placed about 11.2 m from the collision point and guarantees the coverage of the system

up to $|\eta| = 5$. Its structure consists on two cylinders of iron absorbers and quartz fibres as the active material. The HF is optimised for very forward jets events expected in events.

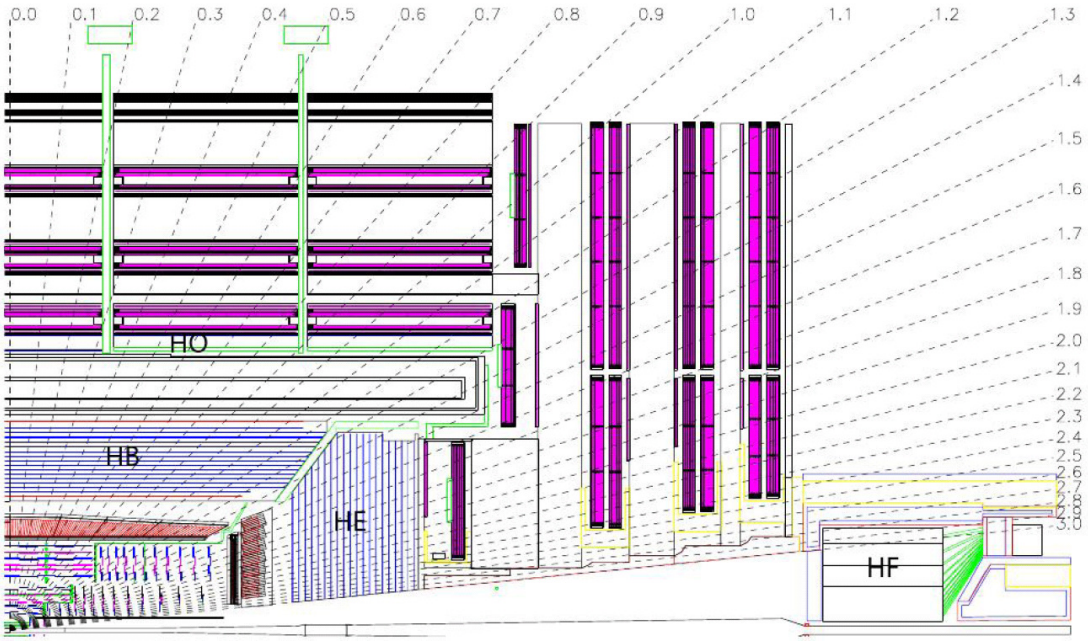


Figure 4.7: A longitudinal view of the HCAL. Figure adapted from reference [62].

The energy resolution as obtained from the commissioning test beam can be described as

$$\frac{\sigma_E}{E} = \frac{90\%\text{GeV}^{1/2}}{\sqrt{E}} \oplus 4.5\% \text{ (barrel/endcaps)}$$
$$\frac{\sigma_E}{E} = \frac{172\%\text{GeV}^{1/2}}{\sqrt{E}} \oplus 9\% \text{ (forward).}$$

4.3.5 Superconducting magnet

The superconducting magnet is one of the distinctive features of the whole CMS detector. It is designed to reach a maximum 4 T magnetic field in a region 6 m wide and 12.5 m long. It represents one of the biggest superconducting solenoid ever built. The solenoid bore contains the majority of the sub-detectors with the only exception of the muon chambers and part of the Hadronic Calorimeter. The high magnetic field needed to bend the particles trajectories is obtained by combining different magnets made by winding reinforced NbTi cables. The energy stored at the operating temperature of 4.6 K and full current amounts to 2.6 GJ. The magnetic flux is returned through a 10 000 t iron return yoke. The return yoke embeds the muon chambers which are surrounded by a residual magnetic field of 1.8 T. The radiation thickness of the magnet cold mass amount to 3.9 X_0. A general artistic view of the CMS superconducting solenoid is shown in Figure 4.8.

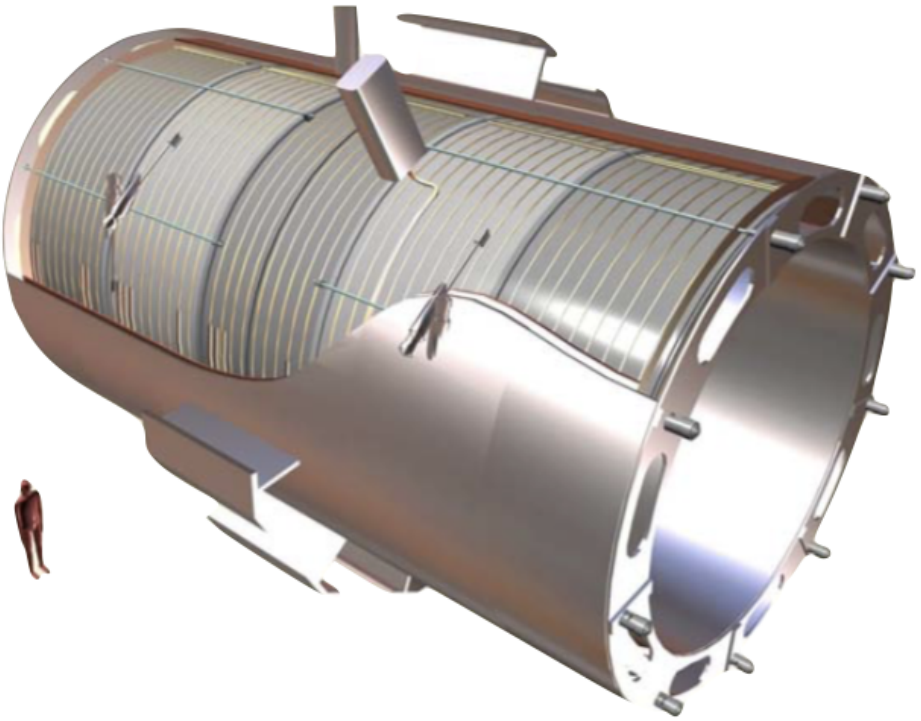


Figure 4.8: An artistic view of the CMS superconducting solenoid. Figure adapted from reference [62].

4.3.6 Muon system

The muon system is specifically designed to offer precise and robust measurements of muons with high transverse momentum. Its main functions consist on muon identification, momentum measurement and triggering (in combination with the tracker). These goals are achieved by exploiting the bending power of the superconducting magnet and its return yoke. The return yoke also serves as an effective hadron absorber thus enhancing the muon identification performances. The muon system guarantees a wide angular coverage and has the capability of reconstructing both charge and momentum of muons in the entire expected kinematic range. It is divided into three independent subsystems which make use of different types of gaseous particle detectors. The choice of using gaseous detector was mainly driven by the need to instrument a large area of about $25\,000\,\mathrm{m}^2$, thus the muon chambers had to be inexpensive, fast, reliable and robust. The muon system follows the cylindrical geometry of the rest of the detector and is composed of a barrel section closed by two endcap disks. Its lateral segmentation, including each independent subsystem, is shown in fig 4.9.

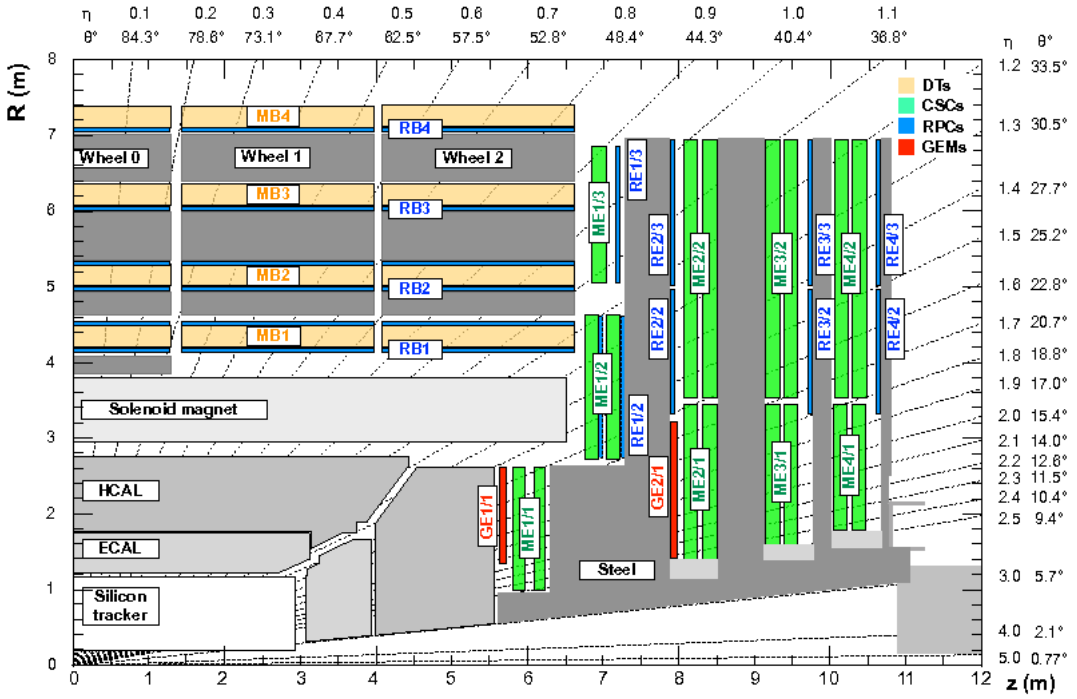


Figure 4.9: The structure of the muon detector system. Figure adapted from reference [72].

4.3.6.1 The Drift Tube system

The barrel region of the muon system is instrumented with drift tube (DT) chambers which cover a pseudorapidity range up to $|\eta| = 1.2$. Chambers with standard rectangular drift cells are installed in the barrel region of the muon system since the particle flux is low (thus the neutron-induced background is small), the 4 T magnetic field is homogeneous and contained in the magnet yoke and the muon rate is generally low. The chambers are arranged into five iron wheels each containing four concentric rings called stations. The first three stations contain eight chambers able to measure the muon coordinate in the transverse bending plane and four chambers dedicated to measure the muon z direction along the beam line. The fourth station is instrumented with eight chambers and it is not providing a z direction measurement. The DT chambers in each station are arranged in such a way to maximise the angular resolution while minimising any detector dead region.

The basic detector element of the chambers is a rectangular drift tube cell with a transverse size of $13 \times 42\,\text{mm}^2$ and a length of 2–4m corresponding to a drift time of about 380 ns. The volume is filled with a gas mixture made of 85% argon and 15% carbon dioxide. The average position resolution of the chambers is 170 µm both in the $r - \phi$ and $r - z$ plane.

4.3.6.2 The Cathode Strip Chambers system

The two endcap regions of the muon system are instrumented with cathode strip chamber (CSC) detectors, able to sustain the high muon and background particle rates and the non-homogeneous magnetic field that characterise this region. The CSCs are specifically designed to identify muons in the pseudorapidity range $0.9 < |\eta| < 2.4$ by providing a fast response time, a fine segmentation and a high radiation resistance. The chambers are distributed in four stations for each endcap, arranged perpendicularly to the beam line and separated by the flux return plates of the superconducting magnet system. The measurement of the particle trajectories in the transverse bending plane is performed by the radially oriented CSC's cathodes. The anodes are read out to provide complementary measurements of the muons pseudorapidity as well as beam-crossing time.

The CSCs are multi-wire proportional chambers filled with a mixture of 40% argon and 50% carbon dioxide as active gas. A remaining 10% of carbon tetrafluoride acts as a quencher by providing a drift path shorter than the DTs. They are arranged into four superimposed disks called stations each formed by two rings divided into eighteen or thirty-six CSCs. Each CSC consists of six layers of sensitive wires with a resolution of 80–85µm in the $\eta \times \phi$ plane and a precision of 0.5 cm in the r direction. Each station provides a robust pattern recognition, able to discriminate between muon signal and non-muon background, and an efficient matching of hits with those recorded by the tracker.

Both the CSC and DT subsystems are intended to provide an efficient muon transverse momentum trigger system, independent from the rest of the detector. The combination of the muon system with the tracking system is of particular importance to achieve the high performances needed by the data triggering system.

4.3.6.3 The Resistive Plate Chamber system

Because of the possible uncertainties when measuring the bunch crossing time and discriminating the non-muon background, resistive plate chamber (RPC) detectors are placed both in the barrel and endcaps region. They are intended to provide robustness and redundancy to the muon system just described. These gaseous detectors provide a coarse spatial resolution but good time measurements by working in avalanche mode. They are specifically designed to complement the DTs and CSCs bunch crossing identification thus solving possible ambiguities in attempting to make tracks by the muon trigger system. The RPCs are made of four bakelite planes coated with graphite that acts as the electrode and two gas gaps of 2 mm. The central part of every chamber is equipped with insulated aluminium strips able to collect the signals generated by crossing particles. A total of six layers of RPCs are placed into the muon system.

4.3.7 Data triggering system

LHC provides proton proton or heavy ion collisions at an unprecedented high rate. Indeed the high particle density in each bunch combined with the limited separation time between each of them corresponds to an event rate of 40 MHz with an average of 20 simultaneous head-on particle collisions per event. Such high data stream is impossible

to handle with the currently available technology. The CMS experiment uses a dedicated trigger system which, by combining raw information from each sub-detector, is able to select possible physics events of interest among all the quasi-simultaneous collisions occurring.

The trigger system is divided into two main stages:

1. a Level-1 trigger (L1) system composed of custom-designed largely programmable electronics;

2. a High Level Trigger (HLT) system fully software based and implemented in a filter farm of about one thousand commercial processors.

The two combined systems, described in the following, are able to reduce the data stream rate to 1 kHz.

The L1 trigger is specifically designed to preselect events of interest in order to be further analysed by the slower and more precise HLT. For this reason, it uses coarsely segmented information coming from both the calorimeters and the muon systems. The input raw data are analysed by dedicated integrated circuit modules which guarantee high speed and radiation resistance while providing sufficient flexibility. Since the L1 trigger has to analyse every collision, the full event readout is pipelined in the front end electronics of each sub-detector during the decision process.

The L1 trigger path is subdivided into different logical components, namely the local, the regional and the global ones. The smallest bit of information used comes from the Trigger Primitive Generators (TPGs) which are built on localised energy deposits in the calorimeters called trigger towers and track segments or hits independently reconstructed by the muon system. The Regional Trigger (RT) then gathers the information coming from different TPGs by sorting and ranking them following a dedicated pattern logic. In this way trigger objects such as electron or muon candidates, localised in limited space regions of the detector, are formed. The rank assigned to each TPG entering the RT decision process is determined as a function of the energy (momentum) and quality of the processed information. The Global Calorimeter Trigger and the Global Muon Trigger gather the information from the relative RTs and, after determining which is the highest rank candidates, transfer them to the top entity of the L1 trigger system. The Global Trigger takes the final decision to reject or accept an event to be further scrutinised by the HLT.

A general view of the CMS L1 trigger system system is shown in Figure 4.10.

4.4 Luminosity collected during LHC Run 2

At the CMS experiment, the luminosity is measured combining the information coming from dedicated luminometers in the HF with information from the pixel detector and the DT system. The luminosity measurement system gathers the signals which come from these detectors and routes them to a high-frequency sampling dedicated electronics for the readout. Such system is able to determine the delivered luminosity once every few seconds, with a statistical precision at the per cent level. The luminosity information is then stored in a database for subsequent use in data analysis.

Precision measurement of the luminosity delivered by LHC to the CMS experiment is of particular interest for different reasons. The online luminosity measurements are

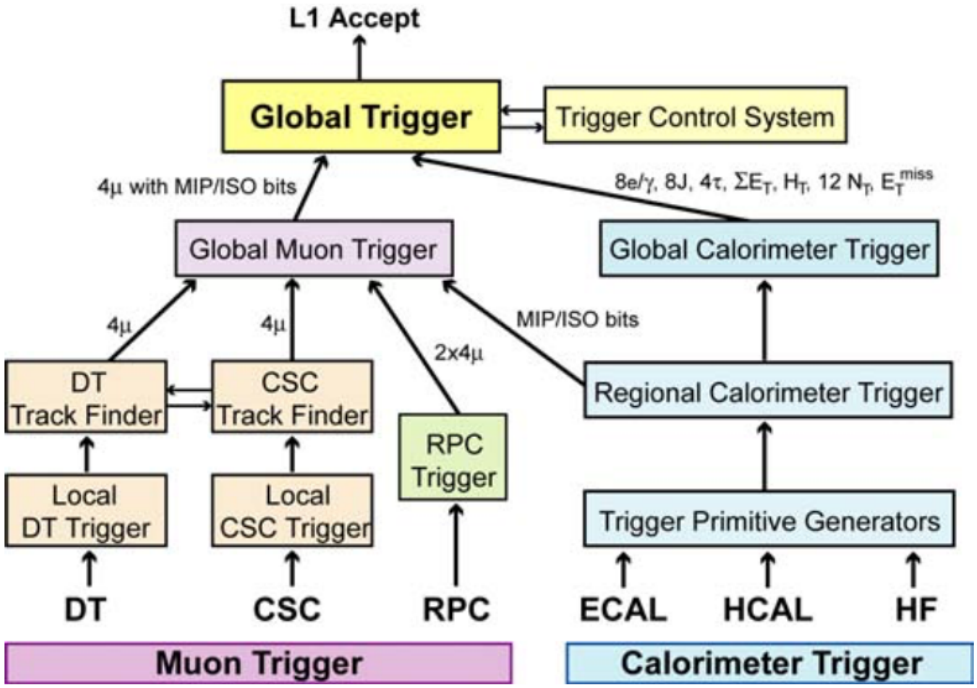


Figure 4.10: The structure of the L1 triggering system. Figure adapted from reference [62].

useful to constantly monitor the LHC performances and to measure the trigger rates. Offline, the recorded luminosity is a critical parameter used for the Monte Carlo events normalisation. Moreover, the luminosity is also needed to convert the observed rate of events in the physics channels of interest, which for this work are represented by the leptonic decay channels of the Z boson $Z\gamma\gamma \to ee\gamma\gamma$ and $Z\gamma\gamma \to \mu\mu\gamma\gamma$, to a measured cross-section.

The LHC luminosity is defined purely in terms of the beam characteristics, as can be seen in Equation 4.1. Nevertheless, the LHC beam properties are not necessarily constant during each data collection period. Indeed the number of particles per bunch is reduced after each beam crossing and the beams cross section is not constant during LHC operation due to the luminosity levelling procedure. The instantaneous luminosity thus cannot be measured using Equation 4.1 because the input parameters needed, if measured during a LHC run, will inevitably interfere with the beam conditions themselves. Instead, luminosity is measured by exploiting its linear relationship with the detector occupancy [73]. Each luminometer reads out the rate R of specific quantities measured by each detector (hits, tracks, energy clusters and so on). This rate is then related to the instantaneous luminosity by exploiting the following relation

$$R = \mathcal{L}_{inst} \cdot \sigma_{vis},$$

where σ_{vis} is the visible cross section of the system, measured in dedicated LHC runs

called *van der Meer scans*. During these special runs, the positions of the beams are varied in the transverse plane in a controlled manner, such that the occupancy gaussian profile of each detector can be measured. By exploiting these measurements with the actual occupancy at nominal LHC beam collision alignment the instantaneous luminosity can be derived. Separate values of R_i and $\sigma_{vis,i}$ are measured for each monitoring system. Their consistency is used for measurement validations and, in combination with the uncertainty in establishing the beam position during the van der Meer scans, to infer the systematic uncertainty of the measured luminosity.

The CMS experiment keeps track both of the delivered and recorded luminosity, the latter being the critical parameter used by physics analyses. The delivered and recorded luminosities could differ because, during the LHC runs, the CMS detector could be temporarily stale of not able to record data, either because the data acquisition chain is busy or any sub-detector system is temporarily unavailable. Each LHC run is then validated by sub-detector experts in order to log only the luminosity sections useful for data analyses. This work makes use of the full LHC Run 2 luminosity collected by the CMS experiment. The corresponding integrated luminosities for the three years of data taking periods are $35.9\,\mathrm{fb}^{-1}$ in 2016 with an overall uncertainty of 2.5% [73], $41.5\,\mathrm{fb}^{-1}$ in 2017 with an overall uncertainty of 2.3% [74] and $59.4\,\mathrm{fb}^{-1}$ in 2018 with an overall uncertainty of 2.5% [75] for a grand total of $136.7\,\mathrm{fb}^{-1}$.

5 | Physics objects reconstruction and object selection

The CMS sub-detectors, described in Section 4.3, are arranged following an "onion shape" around the beams interaction point, each layer ideally sensitive to different stable particles such as electrons, muons, photons and neutral and charged hadrons. Particles are then identified using the signals they leave when traversing the active material of the different sub-detectors.

Particles originating in the interaction point enter first a tracker where their trajectories (tracks) and their origins (vertices) are reconstructed by exploiting the signals left in the detector active layers (hits). Since the tracker is surrounded by a solenoidal magnetic field, the reconstructed track allows to determine both charge and momentum of the candidate particle. Particles with strong electromagnetic interaction, such as electron and photons, could dramatically interact with the tracker material and need to be measured as soon as possible to not degrade irretrievably their energy content. The ECAL is deputed to absorb the corresponding electromagnetic showers, which are detected as clusters of energy deposits in adjacent crystals. By looking at the cluster shape and position, it is possible to infer both the direction and the energy of the candidate electron or photon. The ECAL information is indeed fundamental for the identification of the photons which, differently from the electrons, do not leave any visible hit in the tracker. The ECAL volume is also interested by the initial showering process of charged and neutral hadrons which are then absorbed by the HCAL. Clusters of energy are formed in a similar way as what happens in the ECAL. The muons usually pass through the full detector material undisturbed and can be identified from the hits they produce in the muon system. Finally, the neutrinos do not interact at all with the detector and represent the missing transverse energy (E_T^{miss}) in the event.

A precise measurement of the E_T^{miss} is possible at the condition that the energy of all the other particles is measured with great precision. Indeed this is the main goal of the Particle Flow algorithm described in Section 5.1, which provides a precise identification and measurement of each particle in the event. Once the physics objects are identified their characteristics can be subsequently used for the object selection. Different criteria are exploited to select with the highest efficiency and purity the building blocks of the Zγγ events: electrons and muons, originating from the Z electroweak decay, photons, produced in the primary vertex of the interaction and jets, which are ubiquitous produced in proton proton collisions and play a central role in the background determination. The specific object selection criteria used in this work are presented in Section 5.2.

A schematic view of the specific particle interaction along the CMS detector is

presented in Figure 5.1

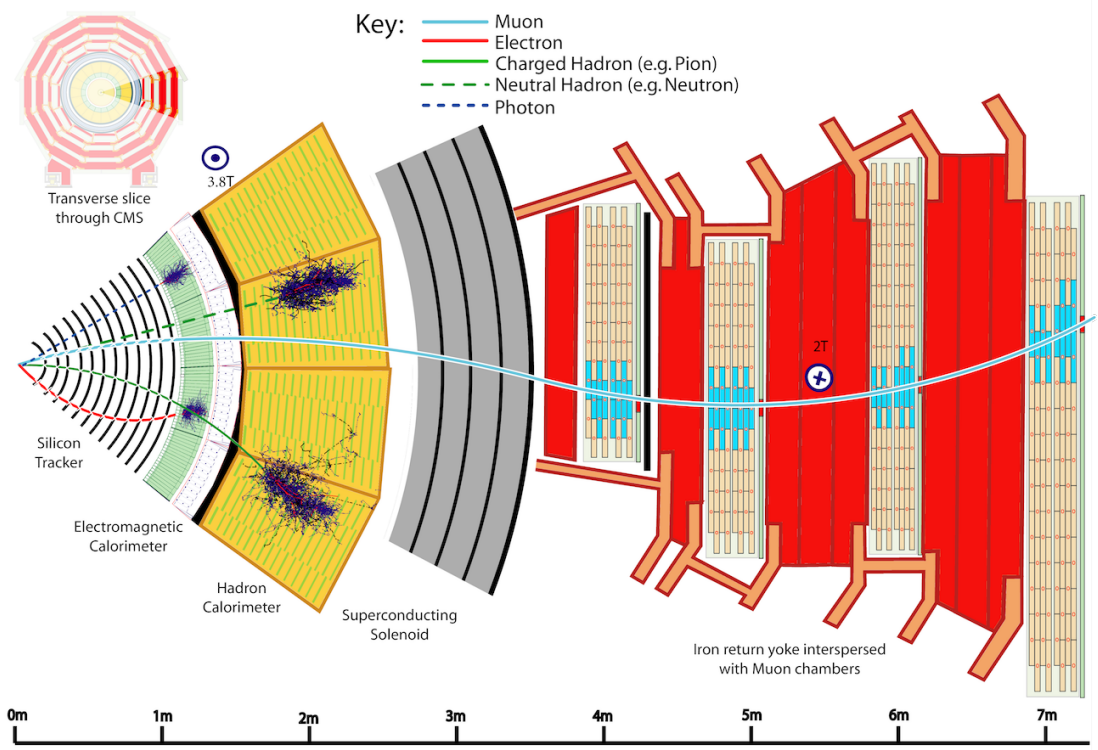


Figure 5.1: A sketch of different particle interactions processes from the beam interaction region along the whole CMS detector volume. Electrons are negatively charged while muons and charged pions are positively charged. Figure adapted from reference [76].

5.1 The particle flow algorithm

The interaction scheme of different kind of particles with the CMS detector elements, as shown in Figure 5.1, is certainly idealised and reality is much more subtle than this. Nevertheless, the basic principles for particle identification established at the beginning of the high energy colliders era can still be used as a guide to develop more and more sophisticated reconstruction algorithms. The development of such tools is a never-ending process, constellated by continuous refinements of the code. Indeed the Particle Flow (PF) algorithm, which is the reconstruction algorithm used by the majority of the CMS analyses including this work, was developed well before the LHC began producing proton proton collisions and it has been constantly updated for years. The techniques described in the following sections correspond to the ones exploited in the particle identification and reconstruction tasks by the CMS experiment.

Before the LHC era, the reconstruction of the physics objects was based on the signals collected separately by each sub-detector. Indeed, the energy of the jets can be reconstructed by exploiting the deposits in the electromagnetic and hadron calorimeters, without any attempt to separate their constituents. The electromagnetic calorimeter

is primarily dedicated to the reconstruction of electrons and photons while the muon identification primarily concerns the muon system.

An improved event description can be achieved by correlating the basic elements coming from each detector layer: instead of thinking of tracks and clusters of energy as separate elements, they can be combined for a more accurate energy, momentum and position measurement. This holistic approach is the founding principle of the CMS reconstruction algorithm. As suggested by its name, the Particle Flow algorithm is aimed to identify and follow each candidate particle in the event through its interactions with the different detector layers. The main goal of the PF algorithm is to reconstruct the relevant properties associated with each particle. It also aims to provide a precise particle identification.

The PF concept was developed to be used in electron-positron colliders and the first experiment which made extensive use of it was the ALEPH experiment at LEP. No attempts were made to implement the PF algorithm in experiments at hadron colliders priors to LHC for mainly two reasons:

- the PF reconstruction algorithm heavily relies on high granularity detector layers able to separate adequately the signals deposited from different particles. Detectors with a coarse granularity may induce signals to spatially merge, thus spoiling the resolution capabilities of the algorithm itself;
- differently from the lepton colliders, where each event can be reconstructed in a clean environment, the intricate final states arising in proton proton collisions is particularly challenging, especially from the computational point of view.

A detailed Monte Carlo simulation campaign that took place at the beginning of the LHC operations in 2009 had demonstrated the adequacy of the CMS high granularity detectors and accurate tracker system to be well suited for the PF algorithm to be used. Since 2010 the PF paradigm has provided accurate energy, momentum and spatial measurements, while also providing a way to quickly cross-calibrate the sub-detectors, validate their signal response and efficiently mask noisy detector channels.

The main benefits of the PF approach can be summarised as an improved determination of the energy and spatial direction of the charged hadrons obtained by combining the measurements in the tracker and the ones in the calorimeters. This a direct consequence of the outstanding CMS tracker resolution. Neutral hadrons and photons are identified as a calorimetric cluster of energy with no associated track. To better reconstruct the electron initial energy the ECAL clusters are linked to the corresponding track in the tracker with a momentum to energy ratio compatible with unity. Finally, muons are identified by a track in the inner tracker connected to the corresponding hits in the muon chambers.

In the following sections the basic detector elements, needed by the PF algorithm, will be described. The so-called linking algorithm, deputed to combine the signals from different detector layers, will be presented as well.

5.1.1 Tracks and vertices

The CMS tracking system aims to identify and measure particle tracks with the highest possible efficiency. Tracking in high energy particle experiments is primarily

targeting energetic particles which are associated with well-measured tracks. This approach, while assuring a high purity and a low misidentification rate, is not particularly efficient, in particular for reconstructing low momentum charged hadron tracks. As a consequence, each charged hadron not identified by the tracker system will inevitably be included into an energy cluster in the calorimeters system, thus being reconstructed as a neutral particle not associated with any track. This results in a largely degraded hadron energy resolution and a coarse reconstructed direction. In a typical LHC event two-thirds of each jet energy, on average, is carried by charged hadrons. Any tracking inefficiency will inevitably spoil a proper reconstruction of the jet energy fraction by overestimating the energy fraction associated with the jet neutral component and result in a degradation of both the jet energy and angular resolution. Increasing the track reconstruction efficiency while keeping under control the misidentification rate is therefore critical for the PF event reconstruction.

5.1.1.1 Charged particles tracking

The direction followed by charged particles transversing the CMS detector is reconstructed by combining the hits left in the pixel and strip layers of the tracker system. Tracks are reconstructed with a Combinatorial Track Finder (CTF) algorithm [76] which is based on a Kalman Filtering (KF) algorithm [77]. The CTF algorithm reconstructs a particle track in three stages: an initial projection, called track seed, is generated by using few hits in the tracker layers which are compatible with a charged particle trajectory. The track is then sequentially extrapolated outwards from the beam interaction point exploiting the KF algorithm which collects the hits in all the detector layers following the particle direction. In the last CTF step, the full candidate trajectory is fitted to determine the relevant charged particle parameters such as the track quality, its origin, transverse momentum and direction. The candidate track is discarded if it fails to meet the required quality criteria.

Increasing the CTF efficiency while keeping a low track misrecontructed rate is important for the PF event reconstruction to be sufficiently reliable. Two critical parameters for a track to be accepted by the CTF algorithm are relaxed to optimise the tracking efficiency. By accepting tracks with smaller transverse momenta the algorithm is able to recover all those candidate particles with a reduced probability to leave a signal in the calorimetric system. Additionally, by accepting tracks associated with fewer hits in the tracker layers the algorithm is able to better catch the candidate particles which, after strong interaction with the tracker material, have degraded their quality. These expedients to maximise the track reconstruction efficiency cause an exponential increase of the misreconstructed track rate. In order to avoid this, the CTF algorithm is applied in several subsequent interactions, each with a moderate efficiency but high purity requirements. At every step the misreconstruction rate is reduced by applying additional quality criteria on the track seed, the track fitted χ^2, its compatibility with the primary vertex and so on. In practice the iterative tracking algorithm initially searches for a high-quality track, that is one with a large number of hits in the tracker layers. When such a track is reconstructed no additional quality criteria are needed. Once reconstructed, the hits associated which the track are masked and the CTF algorithm applied once again to the remaining hits, which are considered to form new seeds and tracks. The main advantages of the interactive tracking approach applied to the CTF

algorithm are summarised in the following:

- the track reconstruction is faster when the reconstruction algorithm is split in multiple iterations than a single one;
- a considerable improvement in the efficiency reconstruction of tracks with transverse momenta as small as 200 MeV is observed;
- any nuclear interaction of a charged particle with the tracker material could degrade its track reconstruction, by resulting in kinks in the original particle direction or in the production of a considerable amount of secondary vertices. An iterative approach is able to better deal with these kinds of phenomena by subsequently applying different and stringent quality criteria to the reconstructed tracks.

5.1.1.2 Electrons tracking

When dealing with the electron track reconstruction two approaches can be followed. The first one targets energetic and well-isolated electron tracks and it is thus based on the ECAL measurements, without too much emphasis on the tracker capabilities. This out-in ECAL-based approach exploits the position of the energy clusters in the ECAL calorimeter to infer the hits left by the candidate electrons in the innermost tracker layers. As can be seen in Figure 5.2 there could be a significant amount of tracker material from the interaction region up to the ECAL radius. A consequence of the tracker thickness is that most of the electrons are likely to emit a sizeable amount of their energy via bremsstrahlung radiation while travelling through the tracker. The performances of the ECAL-based track reconstruction approach are then directly sensitive on how well the electromagnetic clusters gather all the energy radiated by the electron.

The ECAL-based track reconstruction approach fails when it has to deal with the reconstruction of electrons in jets or electrons with low transverse momentum. In the first case, the energy and thus the position of the energy deposits in the calorimeter can be contaminated by the energy deposits of the other particles belonging to the same jet. As a consequence, the back-propagation of the energy cluster position to the interaction vertex would be compatible, on average, with too many hits. The latter class of electrons, being at very low transverse momenta, are characterised by significantly bent tracks and their energy deposit in the calorimeter cannot be completely collected by the energy clustering algorithm. In both cases, tracks reconstructed with the ECAL-based track reconstruction approach will result in a high misreconstruction rate. To reconstruct tracks missed by the ECAL-based track reconstruction approach a tracker-based approach, very similar to the one implemented for the track reconstruction of generic charged hadrons, can be effectively used. In this method, the large probability for electrons to emit bremsstrahlung radiation is exploited to disentangle electron tracks from charged hadrons ones. By using the iterative tracking approach the non-radiating electrons can be reconstructed as efficiently as muons while the radiating ones are recovered in different steps by loosening the track quality criteria. The CTF algorithm is used to reconstruct non-radiating electron tracks, for which the ratio of the ECAL clustered energy to the momentum measured by the tracker is comparable to one. When energetic photons are radiated a Gaussian Sum Filter (GSF) algorithm [78] is employed

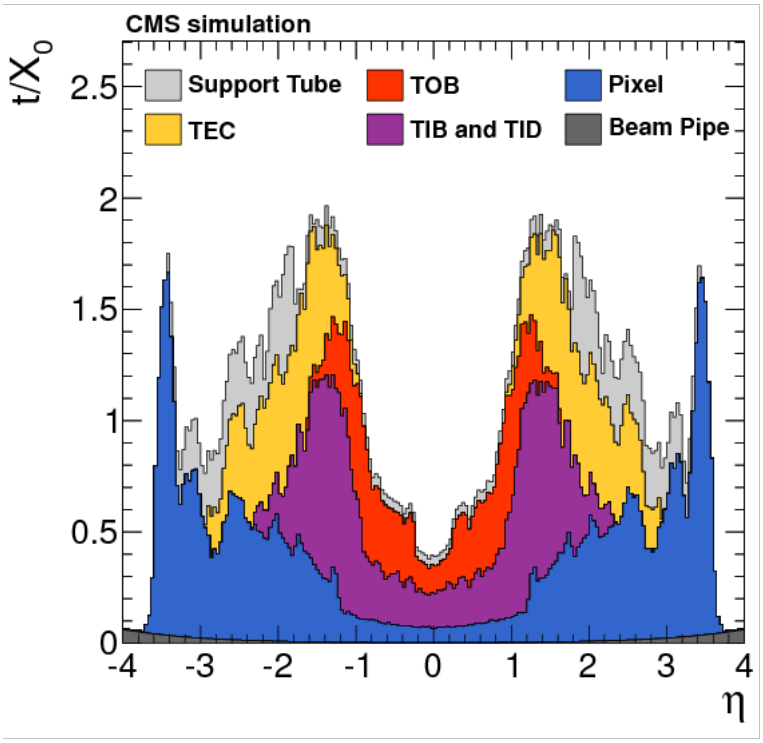


Figure 5.2: Total thickness of the inner tracker material expressed in units of radiation lengths, as a function of the pseudorapidity. Figure adapted from [76].

to better reconstruct the sharp kinks from sudden energy losses. While the KF exploited in the CFT algorithm accounts for the interaction of the candidate particle with the tracker material with a Gaussian smearing around each expected hit location, the GSF algorithm implements a sum of Gaussian terms to better reproduce the Bethe formula for the electron energy loss [79].

The electron seeds from the ECAL-based and the tracker-based approaches are merged in a unique collection. The impact of using these two combined methods to track electrons with respect to an ECAL-based only seeding is shown in Figure 5.3.

5.1.1.3 Primary vertex reconstruction

A typical LHC proton proton collision involves multiple parton interactions, thus a large number of tracks are likely to originate from a variety of vertices. A collection of vertices is identified for each hard scattering process. Tracks used for the vertex reconstruction have to be compatible with the beam spot and need to be well-reconstructed (a low χ^2 is required) [80]. The vertices are then built by a deterministic annealing algorithm [81], which associates to each track i in the event a probability $p_{i,k}$ to come from vertex k. The vertex coordinates are derived by a fit to the associated tracks which takes into account the likelihood of misassignment to the vertex.

The *primary vertex* (PV) is the vertex associated with the hard interaction of interest. It is chosen as the vertex that maximises $\sum_i p_{T,i}^2$, where the sum runs over all the tracks associated with the vertex itself.

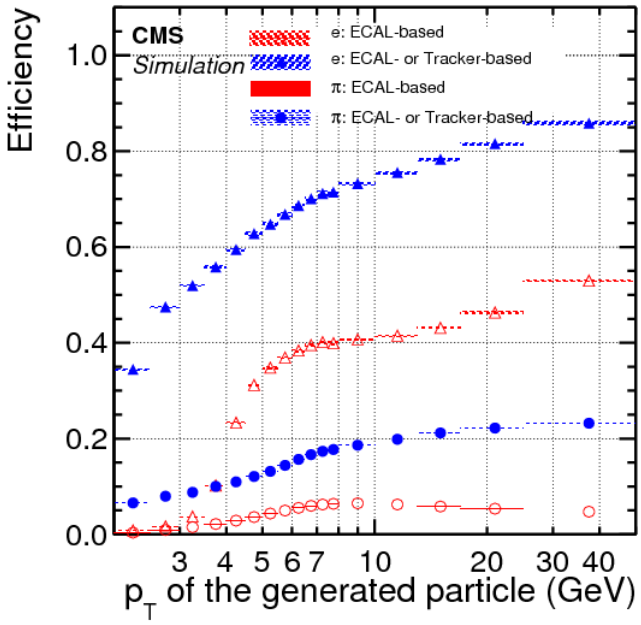


Figure 5.3: The electron seeding efficiency for electrons (triangles) and pions (circles) as a function of transverse momentum, from a simulated event sample enriched in b quark jets. Figure adapted from [76].

5.1.2 Muons tracking

Due to the efficiency of the tracking system (see Section 4.3.2) and of the muon spectrometers (see Section 4.3.6) muons can be identified over the full CMS detector acceptance with high precision and efficiency. High purity in the reconstructed tracks is granted by the considerable material which needs to be passed through by a particle to reach the muon chambers. The vast majority of particles produced in each beam collision are absorbed by the calorimeter systems and, except for neutrinos, in practice only muons can reach the muon system. Muons tracks are the easiest one to be reconstructed, thus the muon tracking is not specific to the PF reconstruction. Instead, it makes use of both the muon system and tracker to produce several collections of muon objects. The broader class of muon tracks is the so-called standalone-muon track: it is formed by fitting hits collected in all the muon systems (DT, CSC and RPC) compatible with an initial seed. If a standalone-muon track is matched to a track in the tracker the hits from the latter detector and the muon system are combined and fitted to form a global-muon track. The global-muon tracking procedure aims to improve the momentum resolution for high energetic muons.

Each track in the tracker can be extrapolated to the muon system. If at least one muon segment is matched to the extrapolated track the correspondent candidate is classified as a tracker-muon track. By construction, global-muon tracks are designated to offer high efficiency when reconstructing muons highly penetrating the muon system. Conversely, for muons with momenta below 10 GeV, the tracker-muon tracks show better efficiency performances due to the multiple scatterings occurring in the steel return yoke.

Global muons and tracker muons which shares the same inner track are merged into single candidates. About 99% of the muon produced in the CMS detector acceptance are reconstructed as global muons or tracker muons.

5.1.3 Calorimeters clusters

The CMS calorimeter systems are dedicated to the measurement of the energy and position of interacting electrons, photons, charged and neutral hadrons. The energy deposited from the interaction of either charged or neutral particles in the calorimeter material forms a shower characterised by a significant lateral extent. The main purpose of the calorimeters clustering algorithm is to efficiently collect nearly as much as possible of the deposited energy. This is fundamental in order to deliver a precise reconstruction of the energy and direction of stable particles such as photons and neutral hadrons. It also helps to separate the neutral particles from the charged hadron and charged lepton deposits. Finally, a precise reconstruction of the energy shower shape improves the energy measurement of charged hadrons associated with low-quality or high transverse momentum tracks.

The PF clustering algorithm collects the energy deposits left by particles interacting with the ECAL and HCAL calorimeters and the two ECAL preshower layers. In the forward region of the HCAL no clustering is performed and the energy information is directly given by the electromagnetic or hadronic components of each calorimeter cell. The algorithm is implemented separately for each sub-detector and for the barrel and endcap regions by exploiting the different characteristics of the detector layout and the energy shower shapes. The first step in the clustering process is to identify the so-called *cluster seeds*, that are cells with energy deposits larger than a given energy threshold and larger than any other neighbouring cell. The threshold energy E_{seed} values are optimised independently for each detector regions, they range from hundreds of MeV to over 1 GeV and are determined from dedicated studies. *Topological clusters* are grown from each seed by aggregating cells with at least one corner in common with a cell already included in the cluster and with energy above the cell threshold value $E_{threshold}$. Finally, an iterative expectation-maximisation algorithm is used to reconstruct the clusters within each topological cluster. The algorithm performs a maximum likelihood fit to the energy deposit positions assuming the energy deposits in each cell to be Gaussian distributed. After convergence, the position and energies of the Gaussian functions are assumed to correspond to the cluster parameters.

In the PF algorithm, the identification and reconstruction of neutral particles rely entirely on the reconstruction of their calorimeter energy clusters. Energy clusters not compatible with any reconstructed tracks are a clear signature of a neutral particle energy deposit. Things start to get worse when the energy deposits from neutral and charged particles overlap: in such cases, the identification and reconstruction of the neutral particles depend on the ability to separate the energy excess with respect to the sum of the associated charged particle momenta. In order to maximise the reconstruction efficiency of neutral particles while maintaining an acceptable misidentification rate, an accurate calibration of the calorimeter response to photons, electrons, charged and neutral hadrons is performed.

5.1.4 The linking algorithm

The *linking algorithm* is specifically designed to connect the PF elements, described in the previous sections, by testing virtually any pair of such elements for each event. To avoid to algorithm execution time to diverge, especially for highly populated events, the pairs of elements being tested are restricted to the ones occupying the same region in the detector (η, ϕ) plane. When combining the information from the sub-detectors different linking criteria are applied. More specifically, the linking algorithm is designed to establish links between the following PF elements:

- a track in the central tracker and a calorimeter cluster, which is determined by extrapolating the track from the last layer of the tracker to the calorimeters. The extrapolation is performed at a depth consistent to the expected maximum of a typical longitudinal electron shower profile in the ECAL and up to one interaction length in the HCAL. In order to collect all the bremsstrahlung photons emitted by electrons traversing the tracker material the energy deposits in the ECAL compatible with tangents line to the GSF tracks are linked together. The probability of a photon conversion for these bremsstrahlung photons is high and a dedicated conversion finder algorithm is developed to link any pair of tracks likely originating from a photon conversion process. The link distance is defined as the displacement between the calorimeter cluster and the extrapolated track in the detector (η, ϕ) plane. If more than one cluster is linked to the same track the one with the smallest distance is retained;

- a calorimeter cluster in the ECAL with a calorimeter cluster in the preshower or in the HCAL. A link is established when the position of the cluster belonging to the more granular calorimeter is found to be within the cluster belonging to the less segmented one;

- a charged particle track and one (or more) charged particle tracks originating from a common secondary vertex. This procedure aims to reconstruct tracks from nuclear interactions with the tracker material;

- finally a track in the tracker system and the information collected by the muon system, as explained in Section 5.1.2.

The outputs of the linking procedure are *PF blocks* whose common elements are associated by direct or indirect links. Due to the high granularity of the CMS sub-detectors (see Section 4.3) the vast majority of the PF blocks ideally contains one or very few candidate particles. For each PF block, the identification and reconstruction procedures follow a defined order. Once each particle is identified the corresponding PF elements (tracks and energy clusters) are removed from the list of the PF blocks. The first particles to be identified are muons. Isolated global muons are selected by considering the additional inner tracks and energy contribution around the muon direction. For a muon to be considered as isolated the sum of these contributions should not exceed a threshold value, depending to the muon transverse momentum. Non-isolated global muons are subjected to additional quality criteria in order to be properly identified. The muon momentum is chosen to be the one measured by the tracker below 200 GeV, while for energies above this threshold a combination of the

tracker and muon systems information is considered. Electrons and isolated photons are identified next. Being the basic properties between electrons and isolated photons similar, they can be reconstructed in the same PF step. In a PF block, an electron candidate is seeded by a GSF track, while a photon candidate is seeded from an ECAL cluster not linked with any GSF track. Several quality criteria on the shower shape characteristics and on the amount of energy radiated from the GSF tracks are applied to each reconstructed candidate. Once muons, electron and isolated photons are identified and removed from the PF blocks, the remaining PF elements are combined to form charged hadrons, neutral hadrons and non-isolated photons. The ECAL and HCAL clusters, reconstructed within the tracker acceptance, are identified as photons or neutral hadrons respectively if they are not linked to any tracks. Beyond the tracker coverage, charged and neutral hadron cannot be distinguished anymore. For this reason, an energy deposit in the ECAL linked to an energy deposit in the HCAL is assumed to originate from the same charged or neutral hadron shower while an energy deposit in the ECAL with no corresponding cluster in the HCAL is identified as a non-isolated photon Each of the HCAL cluster is then combined with one or several tracks. This procedure helps to disentangle the presence of neutral hadrons or photons quasi-collinear to the charged hadron tracks. Indeed if the calibrated calorimetric energy is in excess with respect to the sum of the track momenta the difference is interpreted as produced by neutral particles energy deposits. On the other side, if the calorimetric energy is compatible with the tracker measurement no neutral particles are identified. The charged hadron momentum is determined by combining the measurement in the tracker and the energy deposit in the calorimeters thus enhancing the energy resolution. Hadrons that have experienced nuclear interaction in the tracker material, thus producing secondary vertices, are identified and reconstructed in the last step of the PF reconstruction.

5.2 Object selection

The particles identified and reconstructed by the PF algorithm, described in the previous section, represent the objects pool from which the analyses can start. The PF object selection is looser than the requirements typically applied at analysis level and the optimal balance between efficiency and misidentification rate depends on the characteristics of the signal and background targeted by each particular analysis. The criteria used for the object selection are presented in the following sections. Section 5.2.2 and 5.2.1 describes the characteristics that need to be satisfied by the charged leptons. Section 5.2.3 is of central importance since it describes the quality criteria imposed to the photon candidates. The photon object selection criteria are chosen in such a way as to provide high efficiency while still offering high photon purity, thus keeping as much as possible under control the background contamination coming from hadronic jets misreconstructed as photons (see also Section 6.4.2). Finally, Section 5.2.4 describes the reconstruction of the hadronic jets and their characteristics.

5.2.1 Electrons

The electron selection aims to identify prompt isolated electrons likely originating from the $Z \rightarrow ee$ decay channel and to discard those coming from background sources,

such as photon conversions, jet misidentified as electrons or electrons from semileptonic decays of b and c quarks. Electrons are required to be reconstructed within the detector acceptance by imposing $|\eta| < 2.5$ and excluding the gap region between the ECAL barrel and endcaps $1.442 < |\eta| < 1.566$. Only candidate particles with a transverse momentum greater than $p_T > 15\,\text{GeV}$ are considered. On top of these kinematical cuts, a sequential selection applies additional requirements on several identification variables, including the combined PF isolation relative to the electron transverse momentum, and on the specific variables used to reject converted photons. The selection criteria an electron must satisfy are presented in the following while detailed information on their characteristics can be found in reference [82]. The selection requirements, summarised in Table 5.1, are categorised depending on the electron location in the ECAL calorimeter thus reflecting the different characteristics of the electron energy deposits in the ECAL barrel and endcaps. They correspond to the tight electron selection provided by the CMS experiment with only some slight modifications. The average selection efficiency is 70(70)% while the misidentification rate is 1(3)% for electrons reconstructed in the ECAL barrel (endcaps) [83].

Several requirements on the GSF tracks and the linked PF clusters in the ECAL and HCAL are applied. The track-cluster association criterion requires a geometrical matching between the GSF track and the ECAL cluster which is defined through the variables $|\Delta\eta| = |\eta_{SC} - \eta_{track}^{extrap.}|$ and $|\Delta\phi| = |\phi_{SC} - \phi_{track}^{extrap.}|$, respectively. The quantity $\Delta\eta$ is the difference between the position of the ECAL cluster energy-weighted in η and the track pseudorapidity extrapolated from the innermost track position and directed to the position of closest approach to the ECAL cluster. The quantity $\Delta\phi$ is analogous to $\Delta\eta$ but in the ϕ direction. The track of the candidate electron is required to be reconstructed as close as possible to the corresponding ECAL clusters by requiring, for candidate electrons reconstructed in the ECAL barrel (endcaps), $|\Delta\eta| < 0.00255(0.00501)$ and $|\Delta\phi| < 0.0220(0.0236)$. The lateral extension of the electromagnetic shower along the η direction is accounted by the shower width variable $\sigma_{i\eta i\eta}$ defined by

$$\sigma^2_{i\eta i\eta} = \frac{\sum_{5\times5} w_i(\eta_i - \bar{\eta})^2}{\sum_{5\times5} w_i}$$

where the sums run over a 5×5 matrix of crystals around the most energetic crystal of the cluster. The quantity $\sigma^2_{i\eta i\eta}$ is the weighted variance around the crystal mean in the η direction. The weights w_i depend logarithmically on the deposited energy in each crystal considered for the calculation of $\sigma_{i\eta i\eta}$ and are defined as $w_i = \max(0, w_0 + \ln(E_i/E_{5\times5}))$. Finally the positions of each energy deposit η_i are expressed in units of crystals thus accounting for the variable size gaps existing between the ECAL crystals. The shower distribution on the η direction is a parameter of particular importance since it has a considerable discriminating power between genuine electrons and hadronic jets signals. Since jets are typically composed of a collection of different particles, their shower distribution on the η direction tends to be broader than the one of a single electromagnetic shower, initiated by a genuine electron. Electrons are required to satisfy tight requirements on different variables in the ECAL barrel (endcap). The shape variable $\sigma_{i\eta i\eta}$ must be lower than 0.0104(0.0353). To further suppress spurious electrons from photon conversion or displaced vertices the electron candidate must be compatible with the primary vertex, thus the displacement variable d_{xy} is required to

be less than 0.05(0.10)mm while d_z must be less than 0.10(0.20)mm. While hadronic jets are likely to deposit their energy in both ECAL and HCAL, electrons are expected to be completely absorbed by the electromagnetic calorimeter. For this reason, the ratio H/E between the energy deposit measured in the HCAL calorimetric towers linked by the PF algorithm to a corresponding ECAL cluster is required to be small. Since a flat H/E cut may not be optimal for electrons with transverse momentum greater than 100 GeV a scaled cut is used which reflects the H/E variable dependence on pileup while being asymptotically tighter for higher electron energies and looser at lower electron energies thus better accounting for the HCAL noise. Candidate electron objects are selected by requiring their transverse momentum to be $p_T > 15\,\text{GeV}$. At this energy scale, the electron mass is negligible thus the electron energy measured by the ECAL calorimeter E_{SC} should be compatible with the electron momentum p_T^{track} measured by the tracker at point of closest approach to the vertex. This compatibility is ensured be requiring $|1/E_{SC} - 1/p_T^{track}|$ to be smaller than $0.159(0.0197)\text{GeV}^{-1}$. To ensure a well-reconstructed electron track the expected number of missing hits in the tracker layers is required to be lower than one both for ECAL barrel and endcaps electrons. To further reject electrons likely originating from photon conversion in the tracker material a dedicated photon conversion veto is required to be satisfied. To select only prompt electrons while rejecting those coming from misidentified objects selected electrons are expected to be isolated from any energy flow around their trajectories. The isolation is defined using the PF candidates reconstructed in a cone around the electron direction. In this way, when a generic particle inside a jet is misidentified by the PF algorithm as an electron candidate, it will likely be associated with larger isolation values with respect to genuine electrons from the primary vertex of the interaction. The electron PF isolation is defined as

$$I_{\text{e}} = \left[\sum p_T^{charged} + \max\left(0, \sum E_T^{h^0} + \sum E_T^{\gamma} - p_T^{PU}\right) \right]_{\Delta R<0.3}$$

where the sums run over the charged PF candidates associated with the PV, the neutral hadrons h^0 and photons within a cone of radius $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.3$ around the electron direction. The charged candidates are required to originate from the primary vertex of the interaction. I_{e} is strongly sensitive to any extra energy coming from pileup interaction: in order not to spoil the efficiency of the isolation variable, the contribution coming from neutral pileup processes p_T^{PU} must be subtracted from the total sum. This contribution is computed following the *jet area* approach [84] by assuming $p_T^{PU} = \rho A_{eff}$ where ρ is the transverse momentum per unit area defined as the median between the ratio of each jet transverse momentum to its area p_i^{jet}/A_i^{jet}. The effective area A_{eff} is the geometric area of the isolation cone in the (η, ϕ) plane with an η-dependent correction factor that accounts for any residual dependence of the isolation on the pileup. A relative cut is used used by scaling the isolation variable with the electron transverse momentum with a similar approach as for that used for the H/E variable.

5.2.2 Muons

The muon selection aims to identify prompt isolated muons, likely originating from the $\text{Z} \to \mu\mu$ decay channel. Muons are identified as global muons by the PF algorithm (see Section 5.1.2). They are required to be reconstructed within the detector

Variable	ECAL barrel	ECAL endcap
$\lvert\Delta\eta\rvert$	< 0.00255	< 0.00501
$\lvert\Delta\phi\rvert$	< 0.022	< 0.0236
$\sigma_{i\eta i\eta}$	< 0.0104	< 0.0353
d_{xy}	< 0.05	< 0.10
d_z	< 0.10	< 0.20
H/E	$< 0.026 + 1.15/E_{SC} + 0.0324\rho/E_{SC}$	$< 0.0188 + 2.06/E_{SC} + 0.183\rho/E_{SC}$
$\lvert 1/E_{SC} - 1/p_T^{track}\rvert$	< 0.159	0.0197
Max. number of missing hits	≤ 1	≤ 1
Pass conversion veto	Yes	Yes
$I_{\mathrm{e},\,rel}$	$< 0.0287 + 0.506/p_T$	$0.0445 + 0.963/p_T$

Table 5.1: The tight identification criteria used to select the electron objects used in this analysis.

acceptance by imposing $|\eta| < 2.4$ and only those with a transverse momentum greater than $p_T > 15\,\mathrm{GeV}$ are retained. With respect to the baseline muon selection of the PF algorithm, described in Section 5.1.4, additional quality requirements are put in place. These additional quality criteria aim to suppress non-prompt muons arising from hadrons misidentification and cosmic muon contamination while maintaining a reconstruction efficiency greater than 98% [85]. The final muon object has to pass the tight muon selection described in reference [86] with only some slight modifications. The selection requirements a muon must satisfy are detailed in the following and summarised in Table 5.2.

A tight muon is a global muon with a tracker track that uses hits from at least six layers of the inner tracker, including at least one hit in the pixel detector. The candidate must be reconstructed as a global muon. To suppress hadronic punch-through, that is hadron shower remnants that reach the muon system, and muons from decays in flight the global muon fit must have a χ^2/d.o.f smaller than 10. Additionally, at least one muon chamber hit must be included in the global muon track fit. To further suppress the punch-through as well as any accidental track-to-segment match in the tracking procedure the number of muon segments in the muon stations must be at least two. A tight muon must also be compatible with the primary vertex thus the distance of the track vertex from the primary vertex in the (x, y) plane d_{xy} is required to be less than 0.2 mm. The longitudinal distance in the z direction along the beam axis of the muon tracker vertex with respect to the primary vertex d_z must be less than 0.5 mm. This loose cut helps to further suppress cosmic muons, muons from decays in flight and spurious tracks originating from pileup events. The significance of the muon track vertex with respect to the primary vertex in the event is required to be lower than 4σ. Finally, corrections for bias in the measurement of the muons momenta are applied. Such bias can originate from different sources, including detector misalignment, software reconstruction biases and uncertainties in the magnetic field inside the detector [87].

To select only prompt muons while neglecting those likely to arise from weak decays within the jets muons are required to be isolated. The muon isolation is evaluated relative to its transverse momentum by summing up the energy contribution coming

from each PF element in a geometrical cone surrounding the muon direction. More explicitly, the isolation variable is defined as

$$I_{\mu,rel} = \frac{1}{p_T^{\mu}} \cdot \left[\sum p_T^{h_{PV}^{\pm}} + \max\left(0,\ \sum E_T^{h^0} + \sum E_T^{\gamma} - \frac{1}{2} \cdot \sum p_T^{h_{PU}^{\pm}}\right)\right]_{\Delta R<0.3}.$$

Here the transverse momentum of charged hadrons $h_{PV}^{\pm}$, neutral hadrons h^0 and photons, reconstructed by the PF algorithm within a cone of radius $R = 0.3$ around the muon direction and originating from the primary vertex are summed together. In principle, only particles coming from the primary vertex of the event should contribute to $I_{\mu,rel}$ but in practice only charged hadrons can be associated to the primary vertex $h_{PV}^{\pm}$ or to the pileup contribution $h_{PU}^{\pm}$ while this distinction is not reliable for neutral particles. The neutral component of the $I_{\mu,rel}$ thus needs to be corrected for by removing the contribution from pileup events. This is accomplished by computing the sum of charged hadron deposits originating from all the pileup vertices in the event $h_{PU}^{\pm}$ and by subtracting it from neutral and photon energy sum after being scaled down by a factor 0.5. This scale factor roughly represents the ratio of neutral particles to charged hadron production in inelastic proton proton collisions and it is estimated in simulation. The muons selected for this work are required to have a tight PF based relative isolation value of $I_{\mu,rel} < 0.15$ which corresponds to a selection efficiency of about 95% [86].

Variable	Cut
Number of tracker layers with hits	> 5
Reconstructed Global muon	Yes
Reconstructed PF muon	Yes
Global muon track fit χ^2/d.o.f.	< 10
Number of pixel hits	> 0
Muon chambers hits included in the fit	≥ 1
Number of matched stations	≥ 2
d_{xy}	< 0.2
d_z	< 0.5
IP significance	$< 4\sigma$

Table 5.2: The tight identification criteria used to select the muon objects used in this analysis.

5.2.3 Photons

The photon selection is of central importance for this analysis. It aims to select prompt photons, that are isolated photons coming from the primary interaction vertex (i.e. from ISR/FSR), in order to identify and reconstruct with great precision the Zγγ events. Photon candidates are required to be reconstructed within the detector acceptance by imposing $|\eta| < 2.5$ and excluding the gap region between the ECAL barrel and endcaps $1.442 < |\eta| < 1.566$. Only candidate particles with a transverse momentum greater than $p_T > 20\,\text{GeV}$ are considered.

In the transverse momentum range of interest, photons originating from neutral meson decays are highly collimated and prone to be reconstructed as a single photon.

Indeed the minimum angular separation of two photons from the $\pi^0 \to \gamma\gamma$ decay channel with $p_T^{\pi^0} > 15\,\text{GeV}$ is about the same as the ECAL barrel crystal size [88]. In order to improve the identification of the photon candidates from the primary vertex, against those originating in jets, a *cut based* approach is used where different selection requirements are applied sequentially to a set of individual identification variables. The cut based medium identification is implemented, giving a selection efficiency of approximately 80% and a background rejection higher than 80% for candidate photons reconstructed in either the ECAL barrel or endcaps. The selection criteria summarised in Table 5.3 are categorised depending on the photon location in the ECAL calorimeter and detailed in the following.

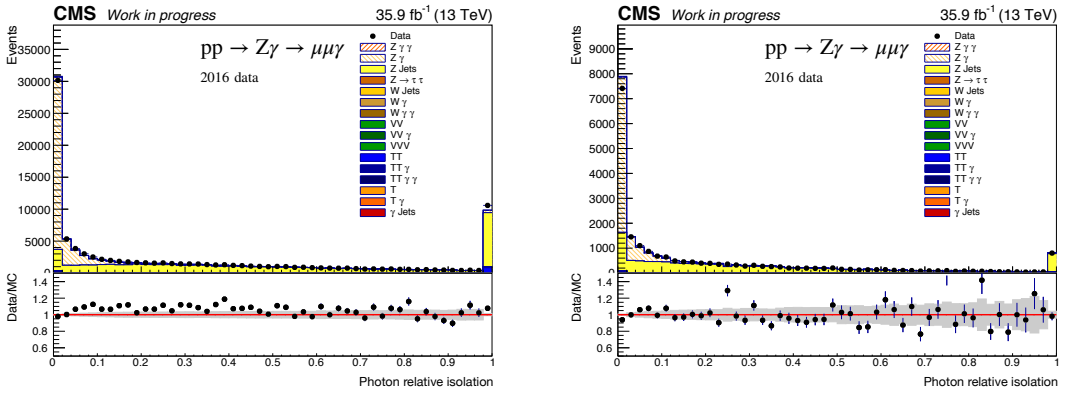
The first requirement imposed on the photon identification is meant to reject spurious electrons that are wrongly reconstructed as photons by the PF algorithm. This is achieved by rejecting any photon associated to a *pixel track seed*, defined as at least two hits in the pixel detectors that could suggest a charged particle trajectory approaching the photon ECAL cluster position. The candidate photon is required to have no pixel seed both when reconstructed in the ECAL barrel and endcap. Additional cuts are required on both the lateral and longitudinal shape of the photon energy deposit: the variables interested are similar to the ones described in the electron objects selection in Section 5.2.1. The lateral extension of the shower $\sigma_{i\eta i\eta}$ is required to be smaller than 0.01015(0.02720) while the hadronic leakage of the shower H/E is required to be smaller than 0.02197(0.03260) for identified photons in the ECAL barrel (endcap).

Three photon isolation variables are obtained by exploiting the information provided by the PF algorithm. The photon isolation variables are defined by summing the transverse momenta of charged hadrons $I_\gamma^{h^\pm}$, neutral hadrons $I_\gamma^{h^0}$ and photons I_γ^γ around the candidate photon direction in an isolation cone of radius $R = 0.3$. As for electrons, the contribution from pileup in the isolation cone is computed following the jet area approach. The effective areas are computed from simulation separately for charged hadrons, neutral hadrons and photons. When the extra contribution from pileup is subtracted from $I_\gamma^{h^\pm}$, $I_\gamma^{h^0}$ and I_γ^γ their dependence on the number of vertices in the event is removed. The ρ-corrected PF isolation variables are then defined as

$$I_{corr} = \max(I_{uncorr} - \rho \cdot A_{eff}, 0.).$$

The rejection of the jets misidentified as photon background relies heavily on photon isolation. Figure 5.4 shows the distribution of the three isolation variables described above combined together for signal (prompt) and background events selected in the $\text{Z}\gamma \to \text{ll}\gamma$ single photon phase space. Smaller values of the combined photon isolation variable correspond to prompt photons ($\text{Z}\gamma$ signal) while at larger values the events are dominated by jets misidentified as photons (Drell-Yan background).

To complement the medium cut based selection described above several additional kinematic cuts are implemented to reject photon candidates originating from any residual spurious effect. For all genuine photon candidates selected in the same event their angular separation is assessed. If a photon candidate is selected in a cone of radius $R = 0.4$ with respect to a previously selected photon that candidate is discarded. Furthermore, in order to reduce bremsstrahlung photons, each photon candidate selected in the event is matched to all the electrons and muons that pass the object selection criteria described



(a) (b)

Figure 5.4: Distributions of the combination of the three photon isolation variable $[I_{\gamma}^{h^{\pm}} + \max(0, I_{\gamma}^{h^{0}} + I_{\gamma}^{\gamma} - \rho \cdot A_{eff})]/p_T^{\gamma}$, obtained for signal (hatched) and background photons (solid yellow) selected the ECAL barrel (left) and endcaps (right). The distributions correspond to $Z\gamma \to \mu^{+}\mu^{-}\gamma$ events selected in this work for the 2016 dataset.

in sections 5.2.1 and 5.2.2. If the photon candidate is reconstructed within a cone of radius $R = 0.4$ centred around the lepton direction then the photon candidate is not considered.

Variable	ECAL barrel	ECAL endcap
Photon pixel seed	No	No
H/E	< 0.02197	< 0.0326
$\sigma_{i\eta i\eta}$	< 0.01015	< 0.0272
$I_{\gamma}^{h^{\pm}}$	< 1.141	< 1.051
$I_{\gamma}^{h^{0}}$	$< 1.189 + 0.01512E_{SC} + 2.259 \cdot 10^{-5}E_{SC}^{2}$	$< 2.718 + 0.0117E_{SC} + 2.3 \cdot 10^{-5}E_{SC}^{2}$
I_{γ}^{γ}	$< 2.08 + 0.004017E_{SC}$	$< 3.867 + 0.0037E_{SC}$

Table 5.3: The medium identification criteria used to select the photon objects used in this analysis.

5.2.4 Jets

As already described in Section 2.1.1, partons originating from the collision vertices hadronize by forming clusters of particles, collectively referred to as jets. A jet represents, to some extent, the characteristics and properties of the initial parton which originated it. In this analysis, jets are reconstructed by using the IRC anti-k_T algorithm. All the PF particles reconstructed in the events are clustered by imposing the radius parameter of Equation 2.1 to be equal to $R = 0.4$. Charged particles originating from pileup vertices are removed from the jets, thus reducing the influence of pileup interactions on the jet energy measurement. Besides, jets whose axes are within $\Delta R = 0.4$ of the direction of any selected electrons, muons, or photons, are not considered.

6 | Analysis strategy

In this chapter the measurement of the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section is presented. The cross section has been measured in a fiducial region defined selecting electron and muon pairs, coming from Z boson decays, accompanied by at least two isolated photons. The Z decay into taus is not considered as a signal process but is included as a background contribution. The other backgrounds sources are represented by vector boson production, such as $Z\gamma$+jets, $W\gamma$+jets, $WZ\gamma$+jets, $WW\gamma$+jets, $W\gamma\gamma$+jets events, and top plus photon production, such as $t\gamma$+jets, $tt\gamma$+jets and $tt\gamma\gamma$ events. The main background contribution is due to Z + jets (Drell-Yan) events where a Z boson is produced alongside at least two jets (or a photon and a jet) and the jets are subsequently misidentified as photon candidates. This background source cannot be reliably estimated from Monte Carlo (MC) simulation and a data-driven dedicated procedure has been developed to determine this contribution.

The chapter is organised as follows. In Section 6.1 the data and MC samples used are presented, together with the basic online trigger requirements. In Section 6.2 the event selection for the fiducial phase-space is defined while the MC object corrections and their calculations are presented in Section 6.3. The techniques exploited to estimate the background contributions are described in Section 6.4 while Section 6.5 describes the related systematic uncertainties. Finally, in Section 6.6 the distributions of events, obtained with the procedures described in the text, are presented.

6.1 Data and Monte Carlo samples

This analysis aims to measure the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section using the full LHC Run 2 statistics, which corresponds to three years of data taking. To take into account the evolution of the CMS detector characteristics during each year of data taking, the corresponding data and MC samples were considered separately and combined.

6.1.1 Data samples

Proton proton collisions at a centre of mass energy of 13 TeV are studied in this work. Data were collected by the CMS experiment during 2016, 2017 and 2018, for a corresponding integrated luminosity of $35.9\,\text{fb}^{-1}$ in 2016, $41.5\,\text{fb}^{-1}$ in 2017 and $59.4\,\text{fb}^{-1}$ in 2018. The overall luminosity obtained when the three years of data taking are combined is greater than $136\,\text{fb}^{-1}$. The data samples exploited for the event reconstruction were reconstructed, preselected and stored in different primary datasets with a single

lepton online trigger selection. The $e^+e^-\gamma\gamma$ events are reconstructed starting from the *SingleElectron* dataset for 2016 and 2017 while the *EGamma* dataset was used in 2018. The $\mu^-\mu^+\gamma\gamma$ events are reconstructed starting from the *SingleMuon* dataset for all the three years of data taking. All the primary datasets used are split into sub-periods, according to the corresponding technical conditions of the CMS detector and the LHC machine. A comprehensive summary of the data samples used in the electron and muon channels is given in Table 6.1

	Luminosity [fb^{-1}]	
Run	Electron dataset	Muon dataset
Run 2016 B	5.750	5.746
Run 2016 C	2.573	2.573
Run 2016 D	4.242	4.242
Run 2016 E	4.025	4.025
Run 2016 F	3.105	3.105
Run 2016 G	7.576	7.576
Run 2016 H	8.651	8.651
Run 2017 B	4.794	4.794
Run 2017 C	9.631	9.631
Run 2017 D	4.248	4.248
Run 2017 E	9.314	9.314
Run 2017 F	13.539	13.539
Run 2018 A	13.704	14.027
Run 2018 B	7.061	7.061
Run 2018 C	6.895	6.895
Run 2018 D	31.742	31.743
Full Run 2	136.85	137.17

Table 6.1: Integrated luminosity breakdown for the partial runs of the *SingleElectron*, *EGamma* and *SingleMuon* datasets over the LHC full Run 2 data taking period. Few runs are included only in the muon or in the electron datasets hence the slight luminosity difference between the two.

6.1.2 Monte Carlo samples

Several simulated samples are used in this work to model the expected signal contribution as well as the known background processes. All the samples were centrally computed while additional MC samples were specifically produced, for this analysis, to model signal contributions from possible Anomalous Quartic Gauge couplings between the Z boson and the photons.

Each MC sample is produced in a fiducial phase-space with different requirements on the generated particles kinematical variables, such as particle transverse momentum or pseudorapidity. The relevant kinematic requirements ("cuts") on the generated particles for each MC sample considered were specifically checked to be looser than the physics objects selection criteria used in the analysis. This procedure ensures that no efficiency loss was experienced when applying to the simulated samples the particle and

reconstruction level phase-space selections described in Section 6.2. Simulated events computed at NLO QCD were reweighed while this procedure was not necessary for the MC samples at LO (in this case each event weight is exactly 1). The MC samples considered were generated with a large number of events, thus providing a reduced statistical uncertainty on the final predictions. In order to be compared with the actual data, each simulated sample is normalised to the observed luminosity of the data samples by using the following equation

$$w_{MC} = \frac{\sigma_{MC} \cdot \mathcal{L}_{data}}{N_{MC,gen}}, \tag{6.1}$$

where σ_{MC} is the theoretical fiducial cross section of the specific MC simulated process, $N_{MC,gen}$ the generated number of events and $\mathcal{L}_{data}$ the recorded luminosity of the data sample to which the MC sample is compared with.

The relevant characteristics of each MC sample are detailed in the following, while a comprehensive summary of all the MC samples used in this analysis, as well as the most precise available fiducial cross cross section, are given in Table 6.2. For all the NLO QCD simulated samples obtained with MADGRAPH5_aMC@NLO, the software version v5 2.2 is used to generate events to be compared to the 2016 data samples while the more recent version v5 2.6 is used to generate events to be compared with the 2017 and 2018 data samples. Similarly, different versions of the same PDFs set are used when generating the events. The NNPDF 3.0 PDFs set is used to generate events for the signal sample to be compared with the 2016 data sample while the more recent NNPDF 3.1 PDFs set is used to generate events for the signal samples to be compared with the 2017 and 2018 data samples (see Section 2.2.1). The NLO QCD samples, explicitly including one or more jets at matrix element, are matched to the PS by using the FxFx matching scheme. The LO samples are matched to the PS following the MLM matching scheme. For all the simulated samples the PYTHIA8 v8.2 package is used to model the parton showering, as well as the hadronisation process and the underlying event simulation. The PYTHIA8 custom parameter set used for the 2016 samples corresponds to the CUETP8M1 tune, while for the 2017 and 2018 samples the more recent CP5 tune is used (see Section 2.3). For all the MC samples used in this analysis the CMS detector response is simulated using the GEANT4 package, with a detailed description of the detector conditions separately included for each year of data taking. The simulated samples are generated by including additional interactions in the same and neighbouring proton bunch crossing, commonly referred to as pileup events (see Section 6.3.1). The MC samples are classified according to the following notation: samples which include two photons in the hard process are referred to as *double photon exclusive samples* while samples which include one photon in the hard process are referred to as *single photon exclusive samples*. The "exclusive" label means that the corresponding sample is enriched in that kind of process. As a comparison, samples which do not include any explicit photon at matrix element are referred to as *inclusive samples*. Inclusive samples can simulate both processes with no photons in the final state and processes with one or more photons in the final state, thus overlapping with the predictions obtained by the exclusive ones. In order to avoid any possible event double counting when combining inclusive and exclusive MC samples, a dedicated overlap removal procedure has been applied to every MC sample considered (see Section 6.2.3)

Process	Generator	Cross section [pb]	Order
$Z\gamma\gamma$	MADGRAPH5_aMC@NLO	0.6832	NLO
$Z\gamma$ + jets	MADGRAPH5_aMC@NLO	49.77	NNLO
$W\gamma$ + jets	MADGRAPH5_aMC@NLO	199.5	NNLO
$t\gamma$ + jets	MADGRAPH5_aMC@NLO + MADSPIN	3.055	NLO
$t\bar{t}\gamma$ + jets	MADGRAPH5_aMC@NLO + MADSPIN	4.216	NLO
$t\bar{t}\gamma\gamma$	MADGRAPH5_aMC@NLO + MADSPIN	0.01687	NLO
$WW\gamma$	MADGRAPH5_aMC@NLO	0.2316	NLO
$ZW\gamma$	MADGRAPH5_aMC@NLO	0.04345	NLO
$Z\gamma$ + jets	SHERPA	93.96	NLO
Z + jets	MADGRAPH5_aMC@NLO	6077.22	NNLO
γ + jets	MADGRAPH5	29707.72	LO

Table 6.2: Summary of the signal and background simulated samples used in this analysis. The MC generators choice as well as the theoretical cross section used to normalise each corresponding sample are presented. The last column reports the precision level at which the theoretical cross section is obtained. The difference in the cross section of the $Z\gamma$ + jets MADGRAPH5_aMC@NLO and SHERPA sample is related to a distinct number of simulated partons at ME (the MADGRAPH5_aMC@NLO sample includes up to one jet at NLO while the SHERPA one includes up to two jets at NLO and a third jet at LO).

The signal contribution in the diphoton phase-space (see Section 6.2) is represented by two final state opposite sign electrons or muons with an invariant mass compatible with the $Z \to l^+l^-$ decay. The events need to contain at least two isolated photons, produced in the primary interaction vertex, either by ISR from the incoming quarks, by FSR from the outgoing leptons or by mixed ISR and FSR processes. The $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ events are simulated at NLO in perturbative QCD by the matrix element calculation of MADGRAPH5_aMC@NLO (see also Section 3.1.2). No jets are explicitly included in the hard process. The *Variable Flavour Scheme* [89], where the quark masses are considered equal to zero, is implemented in the QCD calculations for the computation of the matrix element. The $Z\gamma\gamma$ exclusive double photon signal sample includes all possible leptonic decays of the Z boson.

The main background contribution is represented by jets misidentified as photons. As it will be further explained in Section 6.4.2, a data-driven method has been developed exploiting both data and simulation, obtained in a control region orthogonal to the signal one, to extract this background in the diphoton phase-space. The remaining background contribution is represented by events with two reconstructed opposite signs leptons and two reconstructed photons. In order to model all the possible background contributions several additional exclusive double and single photon samples, which reproduce final states compatible with at least one photon and two reconstructed leptons, are needed. The $Z\gamma$ + jets and $W\gamma$ + jets processes are simulated at NLO in perturbative QCD with up to one explicit jet at matrix element. The $W\gamma$ + jets can contribute to the single photon phase-space if one jet is misreconstructed as a lepton. The predictions are obtained by the matrix element calculation of MADGRAPH5_aMC@NLO using the *Variable Flavour Scheme*. Both the $Z\gamma$ + jets and $W\gamma$ + jets theoretical cross sections are

rescaled to their corresponding next-to-next-leading order (NNLO) QCD predictions obtained with the MATRIX [90] package. The production of one or more photons, alongside one or two top quarks plus jets, can contribute both to the single and to the double photon phase-space if the jet or the top quark pairs are reconstructed as isolated leptons. These process are modelled by the $t\gamma$ + jets, $t\bar{t}\gamma$ + jets and $t\bar{t}\gamma\gamma$ samples, produced at NLO QCD by the matrix element calculation of MADGRAPH5_aMC@NLO and decayed using MADSPIN [91]. The production of a photon associated with WW or WZ pairs are expected to give a negligible contribution both to the single and to the double photon phase-space but are included for completeness. The $WW\gamma \to ll\gamma$ and $WZ\gamma \to lll\gamma$ processes are simulated at NLO QCD by the matrix element calculation of MADGRAPH5_aMC@NLO.

Additional samples are used in this work to validate the obtained results and to assess the systematic uncertainties related to the data-driven method. The $Z\gamma$ + jets process is additionally simulated at NLO QCD up to two jets, while the third jets is simulated at LO, by the matrix element calculation of SHERPA. This sample is used to assess the systematic uncertainty to the jet misidentified as photon background related to the photon FSR modelling. Drell-Yan events, where a Z boson is produced alongside jets, are simulated at NLO up to two jets at matrix element by the calculation of MADGRAPH5_aMC@NLO. The Z + jets theoretical cross section is rescaled to the corresponding NNLO QCD prediction obtained by the FEWZ [92] package. The Drell-Yan sample is mainly used in this analysis to validate the results in the single photon phase-space control region. It is also one of the main ingredients used to obtain the systematic uncertainty on the $Z\gamma$ + jets theoretical cross section, as further explained in Section 6.5.5.1. Finally, the production of a photon plus one or more jets can contribute to the single photon phase-space if the jets are reconstructed as opposite sign leptons. It also represents a relevant background in the dedicated single photon control region used to assess the systematic uncertainty on the jet misidentified as photon data-driven background. This contribution is modelled by the γ + jets sample which is produced at LO by the matrix element calculation of MADGRAPH5_aMC@NLO.

The anomalous Quartic Gauge Coupling samples, used to test the presence of new physics in the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ channel, are simulated at LO using the matrix element calculations of MADGRAPH5_aMC@NLO plus an appropriate set of Feynman diagrams representing the dimension-8 operators of Equation 3.1. The configuration of the input parameter to the matrix element generator is equivalent to the one used for the $Z\gamma\gamma \to l^+l^-\gamma\gamma$ signal sample.

6.1.3 Trigger selection

In order to reconstruct the leptonic decay of the Z boson candidate, electrons and muon objects, selected both in data and simulation samples, are required to pass online specific high-level trigger (HLT) criteria. For both the $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ event selection *single lepton* triggers are used, requiring at least one lepton with a transverse momentum greater than the specific trigger threshold value. Additionally, the HLTs require leptons to satisfy some minimal identification criteria designed to suppress the huge background from lepton misidentified objects. Separate HLTs are used when considering different years of data taking, thus reflecting the different operating conditions of the detector. The HLTs used to select events in this analysis are summarised

in Table 6.3.

Year	Electron trigger	Muon trigger
2016	HLT_Ele27_WPTight_Gsf	HLT_IsoMu24 OR HLT_IsoTkMu24
2017	HLT_Ele32_WPTight_Gsf_L1DoubleEG	HLT_IsoMu24 OR HLT_IsoMu27
2018	HLT_Ele32_WPTight_Gsf	HLT_IsoMu24

Table 6.3: The high-level triggers that electron and muon objects are required to satisfy for the $Z \to l^+l^-$ event selection.

The electron triggers require at least one electron object in the event to be reconstructed with a transverse momentum greater than 27 GeV for the 2016 and 32 GeV for the 2017 and 2018 data and simulated samples. Additionally, the electron object is required to satisfy the identification requirements of Section 5.2.1.

The muon triggers require at least one muon object in the event to be reconstructed with a transverse momentum greater than 24 GeV for the 2016, 2017 and 2018 data and simulated sample. Alternatively, only for 2017, a second HLT with a transverse momentum threshold of 27 GeV is used. Furthermore, the muon object is required to be isolated with respect to the activity of the nearby particles, according to the isolation variable described in Section 5.2.2.

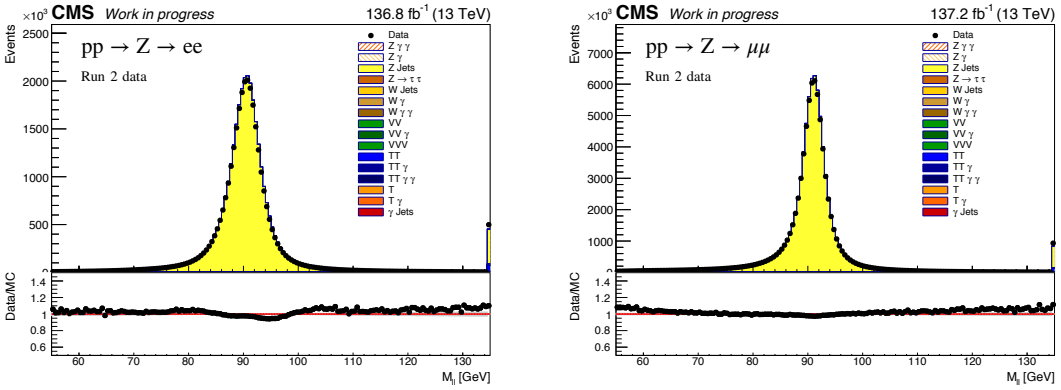
To be sure that the selected event contains the physics object that fired the online selection triggers, a *trigger matching* procedure has been applied. For each event of interest, the maximum distance in the (η, ϕ) plane between the most energetic selected lepton and the trigger object that fired the corresponding single lepton trigger has been assessed. The event of interest is further processed only if the selected lepton is reconstructed within a cone of radius $R = 0.3$ centred around the trigger object direction. Moreover, the trigger object is required to fulfil basic track quality criteria similar to that used in the lepton object selection. If these two requirements are not met the event it is not considered and is discarded.

6.2 Event selection

The selection criteria developed for this analysis aim to select events containing pairs of same flavour and opposite charged leptons, originating from the Z boson decay, produced alongside, at least, two photons. The electron and muon pairs could be either produced by direct decay of the Z boson or from τ intermediate states, such as $Z \to \tau^+\tau^- \to e^+e^-(\mu^+\mu^-)$. The $l^+l^-\gamma\gamma$ selected events represent the signal targeted by this analysis and thus belong to the so-called *double photon phase-space*. As previously stated, additional event selection phase-spaces are used for validation purposes or as control regions. Events selected with the same requirements applied to the double photon phase-space, but where only one photon is required to be present, belong to the *single photon phase-space*, while events where photons are not explicitly required belong to the *inclusive phase-space*.

6.2.1 Reconstruction level event selection

An event is categorised as a Z boson decaying into electrons or muons if at least two opposite charged same flavour leptons are found, each of them satisfying the basic electron selection criteria of Section 5.2.1 or the muon selection criteria of Section 5.2.2. For each event, the lepton pairs found are ordered by decreasing values of the leptons transverse momenta and only the pair with the highest p_T is retained. One of the electron (muon) from the Z boson candidate is required to have $p_{T,\text{lead}} > 35(30)$GeV while the other $p_{T,\text{sublead}} > 15(15)$GeV. Only the leading lepton is required to match the trigger object that fired the event (see Section 6.1.3). The dilepton invariant mass, $M_{l^+l^-}$, is required to be greater than 55 GeV in order to match the generator requirements of the MC samples used and to exclude a region where additional SM dilepton resonances (such as the $J/\psi \to l^+l^-$ or the $\Upsilon \to l^+l^-$) are expected to contribute. The distributions of the Z invariant mass obtained from reconstructed events satisfying these selection criteria are presented in Figure 6.1. Selected distributions obtained in the inclusive phase-space are collected in Appendix A. The inclusive phase-space distributions are useful to



(a) (b)

Figure 6.1: Distribution of the dilepton invariant mass for events selected in the inclusive phase-space selection obtained processing data coming from the entire Run 2 dataset. The contribution from Z + jets events is in yellow. Other minor contributions are included as well but are negligible. In Figure 6.1(a) events are reconstructed in the Z boson electron decay channel, while in Figure 6.1(b) events are reconstructed in the Z boson muon decay channel. An overall good agreement between data and predictions is achieved in the inclusive phase-space, thus validating the basic lepton objects selection and the $Z \to l^+l^-$ events reconstruction.

validate both the basic lepton objects selection and the Z boson reconstruction. With no requirements on the number of selected photons in the event, the main contribution to the event selection is described by the Z + jets MC sample while the contribution of the inclusive $Z\gamma(\gamma)$ samples is negligible.

When an additional photon is required to be present, events are selected in the single photon phase-space. Events in this control region are enriched in the $Z\gamma$ + jets production and are used to determine the probability for a jet to be misidentified as a

genuine photon (see Section 6.4.2). Selected distributions obtained in the single photon phase-space are collected in Appendix B.

When at least two photons are required to be present events are selected in the double photon phase-space, which represents the signal region of this analysis. The $l^+l^-\gamma\gamma$ events are limited in statistic but enriched in the $Z\gamma\gamma$ signal. The main background contribution is represented by $Z\gamma$ + jets or Z + jets events, where one or more jets are misidentified as photons. It is expected to contribute by approximately 40% of the event yield in the double photon phase-space. The remaining background contribution is represented by events where a vector boson is produced alongside photons. This residual background is expected to contribute by approximately 5% of the event yield in the double photon phase-space and it is estimated from MC simulations. More details on the background calculation in the diphoton phase-space are given in Section 6.4. Selected distributions obtained in the double photon phase-space are collected in Section 6.6.

The event selections described in the text are summarised in Table 6.4.

OBJECT SELECTION	
Electrons	Muons
$p_T > 15\,\text{GeV}$	$p_T > 15\,\text{GeV}$
$\lvert\eta\rvert < 1.442$ OR $1.566 < \lvert\eta\rvert < 2.500$	$\lvert\eta\rvert < 2.4$
Cut based Tight ID	Cut based Tight ID
Isolation requirements	Isolation requirements
Photons	
$p_T > 20\,\text{GeV}$	
$\lvert\eta\rvert < 1.442$ OR $1.566 < \lvert\eta\rvert < 2.500$	
Cut based Medium ID	
Pixel Seed Veto	
$\Delta R(\gamma, \gamma/l) > 0.4$	

EVENT SELECTION
Double photon phase-space
At least two opposite charged same flavour leptons
$p_{T,\text{lead}}^{e(\mu)} > 35(30)\text{GeV}$
At least two photons
$M_{ll} > 55\,\text{GeV}$
Fiducial double photon phase-space
At least two opposite charged same flavour leptons
$p_{T,\text{lead}}^{e(\mu)} > 35(30)\text{GeV}$
At least two photons
$M_{ll} > 55\,\text{GeV}$

Table 6.4: Summary of the event selection used in this analysis. The inclusive and single photon phase-spaces are obtained when no explicit photon or at least one photon is required to be present in the event selection. The fiducial double photon phase-space is defined through selections at particle level.

6.2.2 Particle level event selection

In order to measure the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section, predictions of the $Z\gamma\gamma$ event yield are needed in a fiducial phase-space easily reproducible in theoretical calculations or in MC simulations. The $Z\gamma\gamma$ fiducial predictions are defined through selections at particle level. To minimise the dependency on the generator being used, these criteria are required to be as close as possible to the reconstruction level selection used to define the double photon phase-space (see Section 6.2.1). The particle level event selection criteria are also exploited to select the particle level objects used in the photon multiplicity overlap removal procedure, described in Section 6.2.3.

A *lepton dressing* procedure on the generated lepton, where the momenta of photons within a cone of radius $R = 0.1$ around the lepton direction are added to the lepton momentum, is applied to correct for final state photon radiation. The lepton dressing method is meant to reproduce, as close as possible, the spatial resolution of the detector. Dressed leptons are preselected by requiring a transverse momentum greater than $p_T^{\mathrm{l}} > 15\,\mathrm{GeV}$ in a pseudorapidity range $|\eta| < 2.5$. Only electrons and muons are considered in the definition of the fiducial phase-space. Particle level photons are preselected by requiring a transverse momentum greater than $p_T^{\gamma} > 15\,\mathrm{GeV}$ in a pseudorapidity range $|\eta| < 2.5$. The candidate photons are required to be isolated from other selected leptons and photons in a cone of radius $R = 0.4$ around the candidate photon direction. To reproduce the photon isolation criteria applied at the reconstructed level event selection, candidate photons are also required to be isolated with respect to any other stable generated particle in the event, different from leptons, photons and neutrinos. The isolation cone radius in this case is required to be $R = 0.1$. For each event, N_γ is defined as the number of generated photons which satisfy the criteria described above and have a transverse momentum greater than $p_T^{\gamma} > 20\,\mathrm{GeV}$.

Events are selected in the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ channel by requiring at least two opposite charged same flavour leptons, at least one of them with a transverse momentum $p_{T,\mathrm{lead}}^{\mathrm{l}} > 30\,\mathrm{GeV}$, and $N_\gamma \geq 2$. Additionally, the invariant mass of the lepton pair is required to be $M_{\mathrm{ll}} > 55\,\mathrm{GeV}$.

6.2.3 Photon multiplicity overlap removal

The $Z\gamma\gamma \to l^+l^-\gamma\gamma$ events of interest for this analysis are simulated by the $Z\gamma\gamma$ MC sample. However, as explained in Section 6.1.2, additional MC samples in the single photon phase-spaces are needed to estimate the background contribution or for validation purposes. A problem arises when combining the double and single photon MC samples since photons candidates, even if not explicitly included at matrix element generator, can be present in the process final state due to the PYTHIA8 hadronisation procedure. In the single photon phase-space, signal and background contributions include events from the $Z\gamma + \mathrm{jets}$ as well as from the Drell-Yan $Z + \mathrm{jets}$ samples. If an additional photon is generated during the parton showering process, there will clearly be an event *double counting* between the $Z\gamma + \mathrm{jets}$ and the $Z + \mathrm{jets}$ samples. The same issue applies when combining the signal and background MC samples in the double photon phase-space.

In order to avoid any double counting between different MC samples, an overlap removal procedure has been implemented. The basic idea of the method is to exploit the MC information to count the number of generated photons. For each MC sample,

events are classified according to the number of generated photons and discarded if the number of generated photons is not consistent with its expected value. Events from inclusive samples are then discarded if the number of generated photons is greater than zero. Similarly, events from the single photon exclusive samples are discarded if the number of generated photon is found to be different from one. Finally, events from the double photon exclusive samples are discarded if the number of generated photon is found to be smaller than two.

The overlap removal procedure is summarised in Table 6.5. For each MC sample N_γ is obtained by counting the number of photons satisfying the particle level phase-space criteria detailed in the previous section.

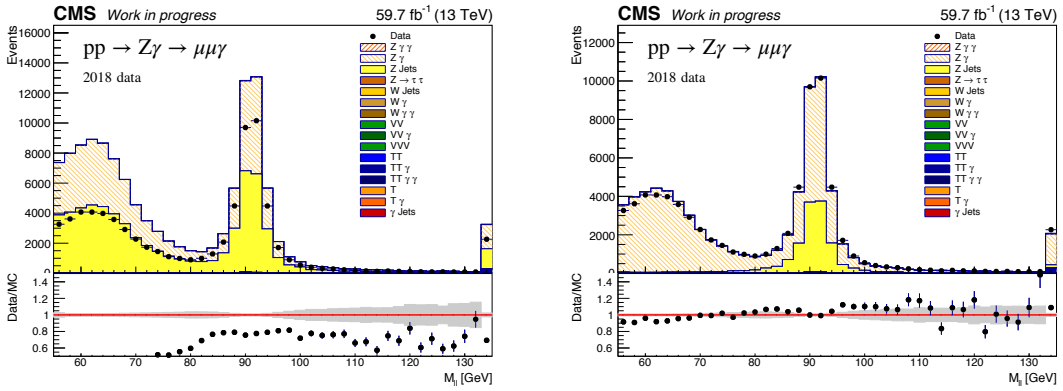
Sample	$N_\gamma = 0$	$N_\gamma = 1$	$N_\gamma \geq 2$
inclusive	Keep event	Discard event	Discard event
single photon exclusive	Discard event	Keep event	Discard event
double photon exclusive	Discard event	Discard event	Keep event

Table 6.5: The requirements on the number of generated photons applied to the simulated events when combining the inclusive and exclusive MC samples. N_γ is the number of generated photon in each sample, as explained in the text.

In order to validate the photon multiplicity overlap removal event distributions of different photon and lepton kinematical variables have been used. For example, Figure 6.2 shows the distribution of the reconstructed level invariant mass of the dimuon system for events reconstructed in the single photon phase-space. This distribution shows a good discriminating power between prompt photon events, simulated by the $\mathrm{Z}\gamma$ + jets signal sample, and the background contribution, which is mainly driven by the Z + jets background sample. In Figure 6.2(a) the predicted signal and background contributions are added together without any overlap removal between the samples: as can be seen, the ratio between observed data and predicted contributions is spoiled and the MC samples clearly overestimate the observed data. A good agreement between data and prediction is achieved when the photon multiplicity overlap removal procedure is applied, as can be seen in Figure 6.2(b).

Another distribution considered to validate the photon overlap removal procedure is the angular separation ΔR between the reconstructed photon and the closest reconstructed lepton. This kinematical variable is sensitive to the final state radiation modelling of the simulated samples and offers an insight into the stability of the overlap removal procedure when applied to different MC samples. Figure 6.3(a) shows the photon ΔR distribution when the inclusive Drell-Yan sample is split in events where no photons are selected at particle level, in yellow, and one photon is selected at particle level in hatched orange. The disagreement between data and predictions is due to the fact in the Z + jets inclusive sample the photon is produced at LO by the parton shower algorithm, while in the $\mathrm{Z}\gamma$ + jets exclusive sample the photon is produced more accurately at NLO, in which case the hard and large-angle contributions are better modelled. This is the main reason why additional exclusive samples are used in this analysis to simulate the prompt photon production in both the single and double photon phase-spaces. In Figure 6.3(b) the region corresponding to the $\mathrm{Z}\gamma$ + jets events is taken from the exclusive samples. A good agreement between data and prediction is achieved

when combining inclusive and exclusive MC samples after applying the photon overlap removal procedure.



(a) (b)

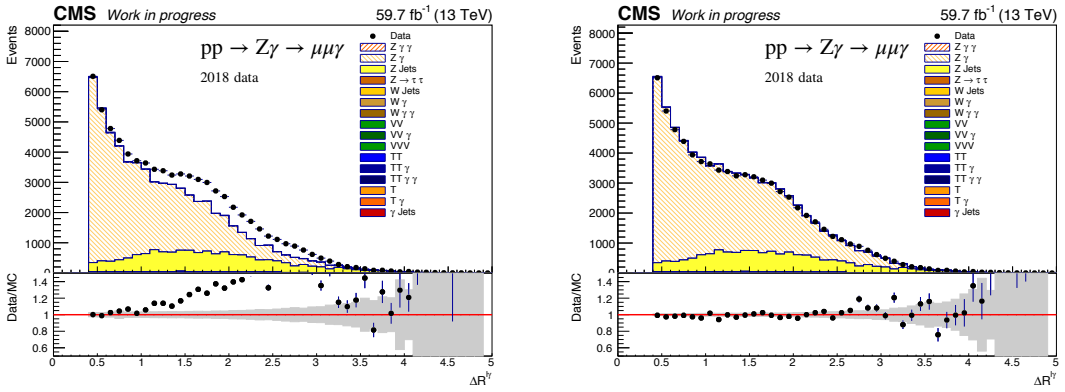
Figure 6.2: Distribution of the dimuon invariant mass for events selected in the single photon phase-space obtained processing data coming from the 2018 dataset. The signal contribution from $Z\gamma$ + jets events is in hatched orange while the main background contribution, which comes from Z+jets events, is in yellow. In Figure 6.2(a) the predicted signal and background contributions are added together without any overlap removal between the MC samples while in Figure 6.2(b) the photon multiplicity overlap removal procedure is applied to each MC sample.

6.3 Monte Carlo corrections

The simulated events used for this work are meant to describe, as close as possible, the observed data. In order to so, generated data are passed to a detailed detector simulation and then reconstructed with the same online and offline algorithms used for the observed data. Anyhow, residual discrepancies are observed when comparing the final event yields in data and simulated events. These residual differences cannot be included in the simulation from the beginning, because MC samples are produced before the actual data are collected. Being not possible to predict with perfect accuracy what might happen to the detector during the data taking period additional correction factors are needed. The MC corrections factors considered in this analysis address different data to simulation discrepancies. The correction to the pileup event simulation is presented in Section 6.3.1, while the correction to a trigger efficiency loss due to the ECAL conditions during the 2016 and 2017 years of data taking is presented in Section 6.3.2. Finally, the corrections to the object selection efficiencies, commonly known as *scale factors*, are presented in Section 6.3.3.

6.3.1 Pileup reweighting

The presence of multiple simultaneous proton proton collisions in each bunch crossing is a common problem when analysing events from high luminosity hadronic colliders.



(a) (b)

Figure 6.3: Distribution of the angle between the reconstructed photon and the reconstructed muons for events selected in the single photon phase-space obtained processing data coming from the 2018 dataset. Figure 6.3(a) shows the photon ΔR distribution when the inclusive $\mathrm{Z}+\mathrm{jets}$ sample is split for events where no photons are selected at particle level, in yellow, and one photon is selected at particle level in hatched orange while in Figure 6.3(b) the region corresponding to the $\mathrm{Z}\gamma+\mathrm{jets}$ events is taken from the exclusive samples while the yellow region corresponds to the $\mathrm{Z}+\mathrm{jets}$ events where no generator photons are selected.

These multiple collisions, referred as pileup (PU) events, can take place *in-time*, that is in the same bunch crossing, or can come *out-of-time* from events in different bunch crossings. Either in-time or out-of-time pileup events pollute the reconstruction of the physics objects of interest, by adding spurious particle tracks in the tracker as well as energy deposits in the calorimeters. This problem is of particular concern when studying the data from the LHC Run 2 era, where the average number of pileup vertices accompanying the primary vertex of interest was about 20.

Each simulated event used in this analysis is generated with a certain "true" pileup distribution superimposed. This distribution is assumed to be "Poissonian like" and it is obtained from a large dataset of minimum bias events, which includes elastic and low energy transfer proton proton collisions generated using the PYTHIA8 event generator. Events are sampled randomly from the simulated pileup dataset and mixed to the simulated hard scattering event, such that the average number of pileup events match the one expected in data. The MC samples are generated with a reasonable guess of the pileup distribution, but differences could arise when comparing to the actual data depending on beam conditions, instantaneous luminosity and so on. Since the $\mathrm{pp}\to\mathrm{Z}\gamma\gamma\to\mathrm{l}^{+}\mathrm{l}^{-}\gamma\gamma$ channel under study could potentially be affected by pileup, each MC event is reweighed so that the distribution of the number of pileup interactions, N_{PU}, matches the one observed in data. The event weights are determined in the following way. The average value of the pileup distribution $\overline{\mu}_{data}$ is estimated for each luminosity section by using

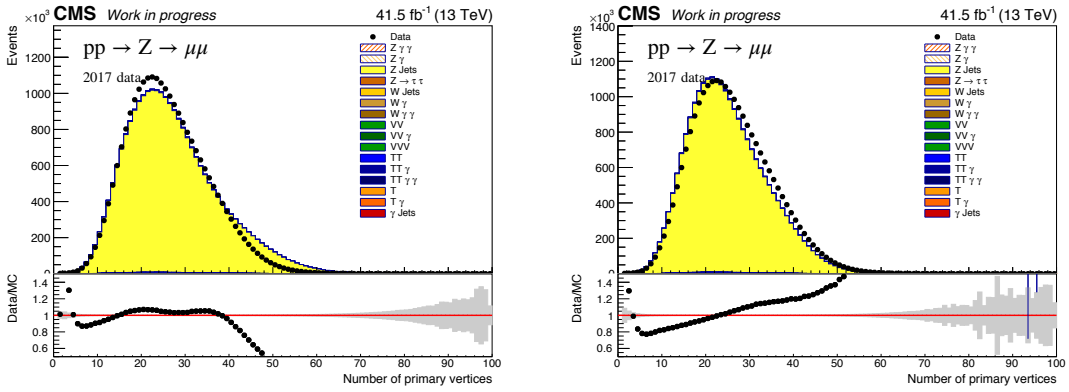
$$\overline{\mu}_{data} = \sigma_{min.\ bias} \cdot \mathcal{L}_{data},$$

where $\mathcal{L}_{data}$ is the recorded luminosity and $\sigma_{min.\ bias}$ the experimental measured proton

proton minimum bias cross section. For each year of data taking the corresponding luminosity is used, while a value of $\sigma_{min.\ bias} = 69.2\,\text{mb}$ is assumed for 13 TeV proton proton collisions, according to reference [93]. The μ_{data} distribution is then compared to the one used to generate the MC events, μ_{MC} and from the ratio between the two appropriate scale factors SF_{PU} are derived:

$$\text{SF}_{PU} = \frac{\mu_{data}}{\mu_{MC}}.$$

In order to validate the MC pileup reweighting procedure, the distributions of the number of the reconstructed primary vertices in data and simulation are compared before and after applying the pileup scale factors to the MC samples. In Figure 6.4 the number of primary vertices is reconstructed by looking in the $\text{Z} \to \mu^+\mu^-$ channel: Figure 6.4(a) is obtained before the reweighting procedure, while Figure 6.4(b) is obtained after applying the pileup scale factors to the simulated events. The agreement between data and predictions is assessed by the ratio plot at the bottom, which shows the improved agreement obtained after the reweighting procedure.



(a) (b)

Figure 6.4: Distribution of the reconstructed number of primary vertices before (left) and after (right) the pileup reweighting procedure to the MC samples for events selected in the inclusive phase-space obtained processing data coming from the 2017 dataset. The $\text{Z} \to \mu^+\mu^-$ events are selected in the inclusive photon phase-space. Data are compared to the Z + jets inclusive sample with the other minor background contributions included. The distributions correspond to Z + jets events selected in this work for the 2017 dataset.

6.3.2 ECAL trigger prefiring

During the 2016 and 2017 years of data taking, a gradual time shift of the ECAL readout signals was not properly propagated to the online L1 ECAL trigger primitives (TPs) (see Section 4.3.7). Indeed, the change in the transparency of the calorimeter crystals caused the light reaching the photodetectors to become increasingly delayed. The effect is worse at high pseudorapidity values, where the particle flux is particularly high and so the radiation damage more pronounced.

The mistiming of signals from the ECAL crystals is of particular concern since it leads to a reduction of the L1 trigger efficiency. This effect, known as *prefiring*, was not corrected for during the data taking period. It was thus possible for an L1 ECAL TP to be wrongly associated with the previous bunch crossing. Since the L1 trigger rules forbid the two following bunch crossing to fire, the trigger events could self veto if a significant amount of ECAL energy was found in the range $2 < |\eta| < 3$.

The effect of prefiring is not described by the MC samples and it is therefore corrected for by applying, to each simulated events, dedicated prefiring scale factors $\mathrm{SF}_{\mathrm{pref.}}$. The $\mathrm{SF}_{\mathrm{pref.}}$ scale factors are obtained from the following equation

$$\mathrm{SF}_{\mathrm{pref.}} = \prod_i [1 - \epsilon_i^{\mathrm{pref.}}(\eta, p_T^{EM})],$$

as the product of the non prefiring probability of all physics objects reconstructed in a dedicated datasets obtained from events separated by exactly two bunch crossings from the previous trigger fire, which makes them unprefireable. These events are $\approx 0.2\%$ of the total dataset.

6.3.3 Object selection efficiency

Minor residual discrepancies between data and simulated events could arise from an imperfect modelling of the objects used in the selection criteria of interest. Selection efficiencies, ϵ_i, are combined together when considering the lepton and photon objects selection in simulated events. The ϵ_i values are separately derived, both in data and MC, using the *Tag & Probe* approach [94]. The related MC scale factors, SF_i, are then computed as a function of the object transverse momentum and pseudorapidity by using the event weights

$$\mathrm{SF}_i = \frac{\epsilon_i^{data}(p_T, \eta)}{\epsilon_i^{MC}(p_T, \eta)}.$$

Events containing simulated electrons, reconstructed and selected in the $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ channel, are corrected for their reconstruction efficiency, ϵ_{reco}. This scale factor accounts for any difference, between data and MC, in the reconstruction of an electron track. Additionally, reconstructed electrons are corrected for their identification efficiency, ϵ_{id}, referring to the identification criteria described in Section 5.2.1, and for their trigger efficiency, ϵ_{HLT}, which is the efficiency of the online HLT selection defined in Section 6.1.3.

Events containing simulated muons, reconstructed and selected in the $\mathrm{Z} \to \mu^+\mu^-$ channel, are corrected for their isolation efficiency, ϵ_{iso}, their identification efficiency, ϵ_{id}, referring to the identification criteria described in Section 5.2.2 and for their trigger efficiency, ϵ_{HLT}.

Finally, events containing simulated photons are corrected for their identification efficiency, ϵ_{id}, referring to the identification criteria described in Section 5.2.3, and for the pixel seed veto efficiency, ϵ_{veto}, which is the efficiency of the photon pixel seed veto requirement described in Section 5.2.3.

All the aforementioned scale factors come from standard certified values, centrally available for all the analysis targeting the full Run 2 dataset. An important exception is represented by the electron trigger selection efficiency, ϵ_{HLT}, which is specific to this

analysis event selection phase-space. A dedicated Tag & Probe workflow, presented in the following section, has been developed to compute the related scale factor.

6.3.4 Electron trigger efficiencies using the Tag&Probe method

The main idea of the Tag & Probe approach is to measure the object selection efficiency directly from data while deriving appropriate correction factors to be applied to the efficiency measured in simulated events. In the Tag & Probe data-driven approach a know SM mass value, the Z boson resonance, can be used to select particles of the desired type (in this case electrons). The efficiency of a particular selection criterion is then probed on those selected particles.

In order to compute the efficiency of the electron HLTs selection criteria, the single electron datasets of Table 6.1 and the Z + jets Drell-Yan sample of Table 6.2 are used. The *tag* object is required to satisfy a set of very stringent selection criteria, designed to isolate the required particle type with a very small fake rate. Tag electron candidates are selected with a transverse momentum greater than 35 GeV and are required to be reconstructed within the detector acceptance by imposing $|\eta| < 2.1$, excluding the gap region between the ECAL barrel and endcaps $1.442 < |\eta| < 1.566$. Tag electrons are also required to pass the single electron HLT trigger of Table 6.3 and a tight cut based selection is applied to them. The *probe* object is selected with looser selection criteria and it is required to be paired with a tag object, such that the invariant mass of the combination of the two is compatible with the mass of the Z boson resonance. Probe electron candidates are selected with a transverse momentum greater than 35 GeV and are required to be reconstructed within the detector acceptance by imposing $|\eta| < 2.5$, excluding the gap region between the ECAL barrel and endcaps $1.442 < |\eta| < 1.566$. A tight cut based selection is then applied. The tag and probe pairs are required to be formed by opposite charge particles, with a reconstructed invariant mass in the range $50 < M_{\mathrm{ee}} < 130$ GeV. The trigger efficiency is then measured as

$$\epsilon_{HLT} = \frac{N_{pass}}{N_{pass} + N_{fail}},$$

where N_{pass} is the number of pairs where the probe electron has passed the HLT trigger requirement, and N_{fail} is the number of pairs where the probe electron has failed to pass these criteria. In order to measure N_{pass} and N_{fail}, the distributions of the reconstructed invariant mass of the pair are fitted separately, using a chi-square minimisation procedure, with an appropriate signal plus background model. An example of the fitted pass and fail distributions is presented in Figure 6.5. The fitted signal models are then integrated to obtain the desired N_i event yield. The whole procedure described above is repeated in different bins of the probe variables and the efficiency is measured in data as a function of the electron transverse momentum and pseudorapidity. The efficiency has been derived separately for each year of data taking to match the different running conditions. The final results are presented in Figure 6.6. The same plots show, at the bottom, the data to MC ratios which correspond to the MC scale factors used in this analysis.

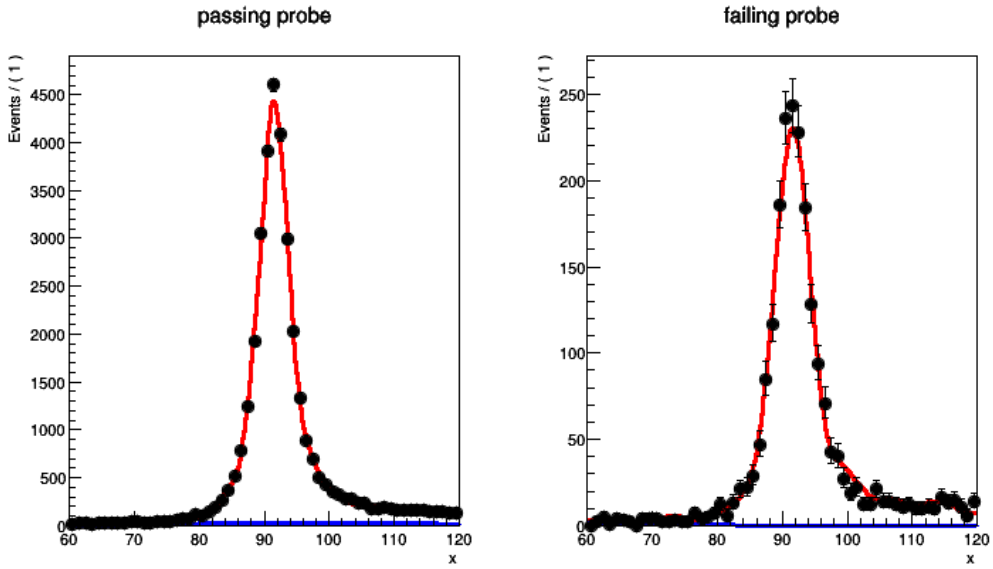


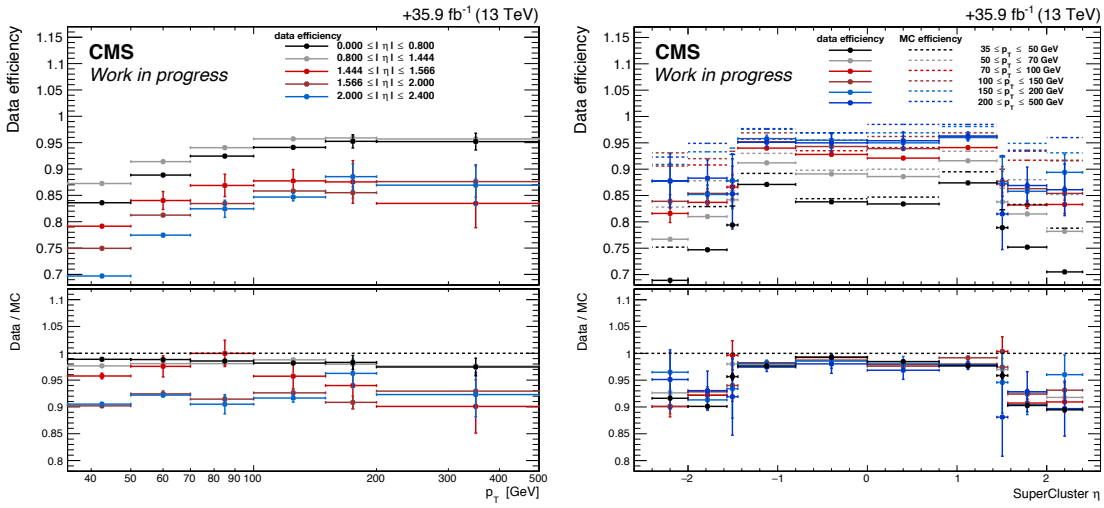
Figure 6.5: An example of the fitted "tag plus passing probe" and "tag plus failing probe" line shapes for electrons selected in data in a particular (p_T, η) bin.

6.4 Background contributions

Several SM processes, with a final state topology similar to the $Z\gamma\gamma$ final state, can contribute to the double photon phase-space. The background composition can be divided into two main categories: a prompt background contribution, where the $l^+l^-\gamma\gamma$ final state comes from prompt photon candidates, and a non-prompt background, represented by $l^+l^-\gamma\gamma$ events where either one or both photons come from hadrons decaying to photons inside jets. The prompt background contributes to approximatively 5% of the final event yield in the diphoton phase-space and it is estimated from simulated events, while the non-prompt background represents the major contribution (approximatively 40%) and it is estimated from control samples in data.

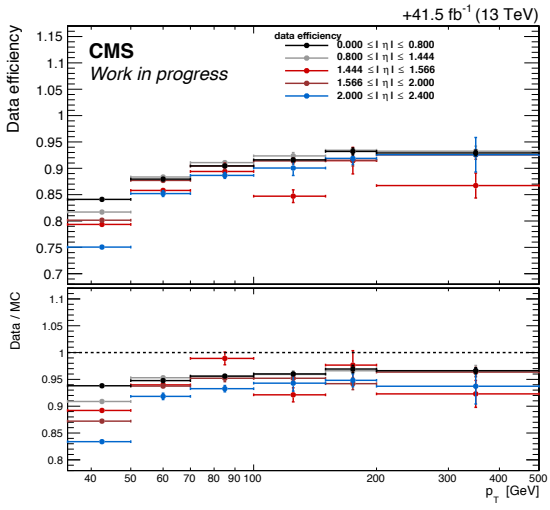
6.4.1 Prompt background contribution

A minor background contribution to the $l^+l^-\gamma\gamma$ double photon phase-space selection is represented by events where two opposite charged same flavour leptons plus a prompt photon are produced, and a second genuine photon coming from ISR/FSR processes is reconstructed. This background contribution accounts for approximately 12% of the total event yield in both the $Z\gamma\gamma \to e^+e^-\gamma\gamma$ and $Z\gamma\gamma \to \mu^+\mu^-\gamma\gamma$ decay channels, and it is estimated directly from MC simulations. The simulated samples used are the $Z\gamma$ + jets, the $W\gamma$ + jets, the $VV\gamma$ (where V could be either a Z or a W boson), the $t\gamma$ + jets, the $t\bar{t}\gamma$ + jets and the $t\bar{t}\gamma\gamma$ processes. All the aforementioned samples, where a Z is not explicitly required to be present in the hard process, could contribute if two opposite charged same flavour leptons with a $M_{ll} > 55\,\text{GeV}$ are found in the event (for example

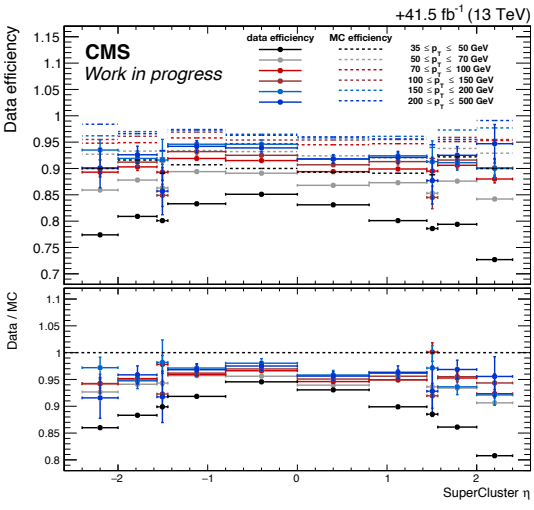


(a)

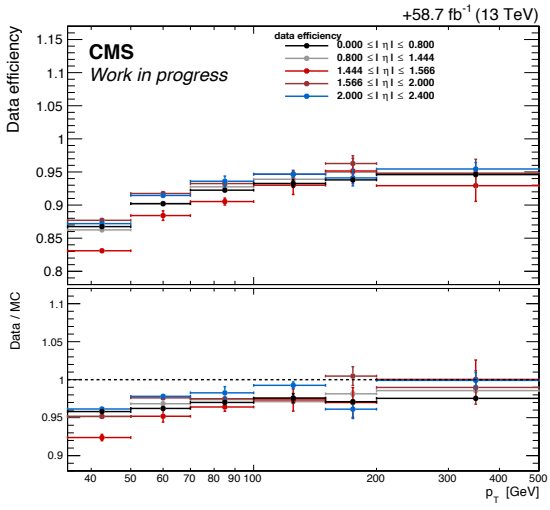
(b)



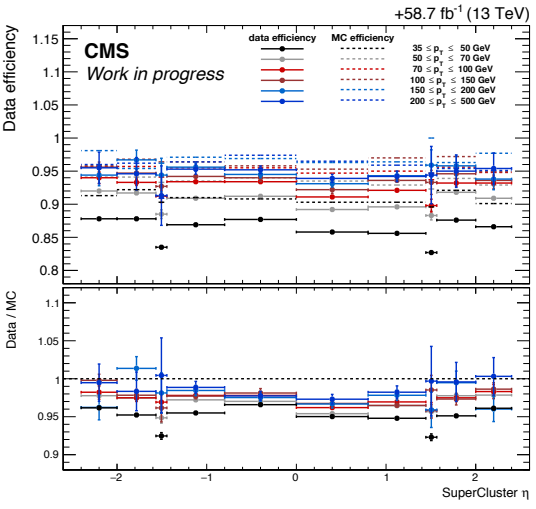
(c)



(d)



(e)



(f)

Figure 6.6: The electron HLT efficiencies, ϵ_{HLT}, and the corresponding SFs used in this work. Figure 6.6(a) and 6.6(b) show the results obtained for the 2016 year of data taking, Figure 6.6(c) and 6.6(d) those obtained in 2017 and Figure 6.6(e) and 6.6(f) those obtained in 2018.

coming from the top quark decays $t \to W + b \to l\nu b$). With the exception of the $tt\gamma\gamma$ sample, the single photon exclusive samples could contribute to the double photon phase-space if an additional photon is generated with a transverse momentum lower than 20 GeV, but it is then reconstructed with a transverse momentum greater than 20 GeV. In order to select only reconstructed events where two genuine photons are present each MC sample is matched at particle level, which means that each reconstructed level photon object must correspond to a genuinely generated photon. This matching procedure is possible since the history of each particle, from generation to reconstruction levels, is available in the simulated event.

6.4.2 Non-prompt background contribution

The major background contribution to the $l^+l^-\gamma\gamma$ double photon phase-space is made of events where two opposite charged same flavour leptons are produced alongside one or more non-prompt photons. This source is represented by jets which are reconstructed as photons, and it is so also labelled as *jet misidentified as photon* background. A high energetic jet can be reconstructed as a photon when it contains a neutral hadron, such as a π^0 meson, decaying to two overlapping photons via the electromagnetic decay process $\pi^0 \to \gamma\gamma$. If the photons carry a large fraction of the total jet energy the other hadronisation products in the jet cone will have low momentum and, eventually, the photons can be reconstructed as single energy deposit with no associated tracks. The photon object selection criteria used in this analysis, in particular the photon isolation variables, are specifically designed to reject this kind of background source. However, the $Z + \text{jets}$ and $Z\gamma + \text{jets}$ cross sections are orders of magnitude larger than the predicted $Z\gamma\gamma$ contribution (see Table 6.2) and, despite the dedicated object selection criteria, a considerable fraction of misreconstructed photons is expected to enter into the double photon phase-space signal region. The $Z + \text{jets}$ events are expected to contribute to this background when at least two separate jets in the event are reconstructed as photons, while the $Z\gamma + \text{jets}$ and all other single photon exclusive samples will contribute when at least one jet in the event is reconstructed as a photon. The probability for a jet to be misidentified as a photon is sensitive to how jet fragmentation is modelled. In this particular case, the MC predictions are not completely reliable and the contribution from background process with non-prompt photons is estimated with a dedicated data-driven method, similar to the one developed in other analyses by the CMS and ATLAS Collaborations [57, 58, 95].

The data-driven method developed in this analysis aims to extract the probability, for a prompt photon and a jet, to be reconstructed as isolated photons. These probabilities are derived from events selected in the single photon phase-space and combined to obtain a set of *templates*. The templates are then used, jointly with the signal region and three additional control regions, to derive the non-prompt contribution in the double photon phase-space. Different templates are obtained as a function of the single photon transverse momentum and pseudorapidity values in order to estimate, simultaneously, both the normalisation and the shape of the background contribution, differentially with respect to the diphoton distribution of interest.

Events satisfying the full analysis selections, with the exception of the photon isolation criteria (see Section 5.2.3), are selected in the double photon phase-space separately

for each year of data taking. The photon isolation requirements act as discriminating variables, able to separate, for each event, candidates likely to be associated with prompt photons from non-prompt ones. For each event, the photon candidates which satisfy simultaneously the three photons isolations criteria are labelled as *tight* (T) otherwise as *loose* (L). Three control regions are then defined in the double photon phase-space by the identification criteria of the most energetic (leading) and the second most energetic (sub-leading) photons in each photon pair:

- events where the leading photon is tagged as tight and the sub-leading one as loose constitutes the N_{TL} control region;
- events where the leading photon is tagged as loose and the sub-leading one as tight constitutes the N_{LT} control region;
- finally, events where the leading photon is tagged as loose and the sub-leading one as loose constitutes the N_{LL} control region.

The N_{TL} and N_{LT} regions are separately considered in order to account for any difference in the misidentification probabilities related to the candidate photon transverse momentum and pseudorapidity. These control regions do not overlap with the signal region (N_{TT}, where both photons are tagged as tight) and are enriched in events where at least one jet is likely to have been reconstructed as a prompt photon. Events selected in the N_{TT} signal region and in the N_{TL}, N_{LT} and N_{LL} control regions receive contributions by four possible sources of diphoton candidates:

- events where the leading and sub-leading photon candidates are prompt photons constitute the $\alpha_{\gamma\gamma}$ event source;
- events where the leading photon candidate is a prompt photon the sub-leading one a jet constitute the $\alpha_{\gamma j}$ event source;
- events where the leading photon candidate is a jet and the sub-leading one a photon constitute the $\alpha_{j\gamma}$ event source;
- finally, events where the leading and sub-leading photon candidates are jets constitute the α_{jj} event source.

The link between the observed diphoton events, selected in the signal and control regions, and the expected contribution from each source of diphoton candidates, is represented by the probabilities for prompt photons or jets misidentified as photons to be reconstructed as a tight (loose) photon candidates. These probabilities are derived by selecting events both in the single photon exclusive MC samples and in data where a dimuon pair, compatible with the $\mathrm{Z} \to \mu^{+}\mu^{-}$ boson decay, is produced alongside a photon. The only difference with respect to the single photon phase-space described in Section 6.2.1 is that the photon candidate is not required to satisfy the photon isolation criteria.

The probability ϵ for a single prompt photon to be reconstructed as a tight photon candidate is obtained from simulated $\mathrm{Z}\gamma$ + jets events. For the MC selection, the reconstructed photon is required to match a particle level photon with the same procedure described in Section 6.4.1. This additional requirement aims to select a sample of simulated events as pure as possible in prompt photons. Simulated events are

then split into two categories, being $n_T^{\gamma,MC}$ the number of simulated events where the reconstructed photon was labelled as tight and $n_L^{\gamma,MC}$ those where the reconstructed photon was labelled as loose. The $n_X^{\gamma,MC}$ account for the weights defined in Section 6.1.2 and Section 6.3 and are derived as a function of the reconstructed photon transverse momentum and pseudorapidity. For each bin the corresponding $\epsilon(p_T, \eta)$ values are obtained according to

$$\epsilon(p_T, \eta) = \frac{n_T^{\gamma,MC}}{n_T^{\gamma,MC} + n_L^{\gamma,MC}}.$$

The $\epsilon(p_T, \eta)$ probabilities are derived in five ranges of the photon transverse momentum, namely $20 < p_T < 40\,\text{GeV}$, $40 < p_T < 60\,\text{GeV}$, $60 < p_T < 80\,\text{GeV}$, $80 < p_T < 180\,\text{GeV}$ and $p_T > 180\,\text{GeV}$, separately for photons reconstructed in the ECAL barrel (EB) or in the ECAL endcaps (EE).

The probability f for a jet to be reconstructed as a tight photon candidate is obtained from data in a sample enriched as much as possible in jets misidentified as photons. To obtain such a sample, the prompt photon contribution must be subtracted using the single photon exclusive MC samples. The outcomes of this procedure are $n_T^{\gamma,Data}$, the number of events selected in data where the reconstructed photon was labelled as tight and $n_L^{\gamma,Data}$, where the reconstructed photon was labelled as loose. The $n_X^{\gamma,Data}$ are derived in the same ranges of the photon transverse momentum and pseudorapidity as the corresponding $n_X^{\gamma,MC}$ ones. The $f(p_T, \eta)$ values are then obtained as

$$f(p_T, \eta) = \frac{n_T^{\gamma,Data}}{n_T^{\gamma,Data} + n_L^{\gamma,Data}}.$$

The prompt photon yields, obtained from simulated events, are compared to the corresponding data yields, obtained after selecting tight and loose photon candidates in Table 6.6. Due to the high non-prompt background rejection power of the photon object selection criteria used it is possible that, for some (p_T, η) bins, the prompt photon yield from MC is larger than the one observed data. This is particularly evident for high values of the candidate photon p_T. In such cases, the $n_X^{\gamma,Data}$ would be negative and thus the $f(p_T, \eta)$ probabilities not physical. To avoid this situation, when computing the $f(p_T, \eta)$ probabilities, if $f(p_T, \eta) \leq 0$ for a particular (p_T, η) bin the $f(p_T, \eta)$ probability derived for the lower p_T interval range is considered.

The probabilities for prompt photons and for jets to be reconstructed as a tight isolated photon candidates are summarised in Table 6.7, separately for each year of data taking. The corresponding probabilities for a prompt photon and a jet to be reconstructed as a loose isolated photon candidate are $1 - \epsilon$ and $1 - f$.

To estimate the non-prompt photon background contribution in the double photon phase-space, the ϵ and f coefficients obtained above are combined to form a 4×4 probability matrix M. Each entry of the matrix M_{AB}^{XY} represents the probability for each source α_{AB} of diphoton candidates to fall within the signal or control regions N_{XY}. Being ϵ and f probabilities, the entries in each column of the matrix M sum to unity by construction. The normalisations of the four sources of diphoton candidates are then

2016	EB		EE	
p_T range	Tight (%)	Loose (%)	Tight (%)	Loose (%)
$20 < p_T < 40$	87 ± 1	12 ± 1	75 ± 1	10 ± 1
$40 < p_T < 60$	88 ± 2	10 ± 1	78 ± 3	11 ± 1
$60 < p_T < 80$	90 ± 4	12 ± 1	81 ± 6	14 ± 2
$80 < p_T < 180$	89 ± 4	16 ± 1	93 ± 7	14 ± 2
$180 < p_T < 300$	92 ± 10	25 ± 4	99 ± 23	10 ± 6
Total	88 ± 1	12 ± 1	77 ± 1	11 ± 1

2017	EB		EE	
p_T range	Tight (%)	Loose (%)	Tight (%)	Loose (%)
$20 < p_T < 40$	79 ± 1	11 ± 1	70 ± 1	13 ± 1
$40 < p_T < 60$	74 ± 2	9 ± 1	70 ± 3	10 ± 1
$60 < p_T < 80$	76 ± 3	9 ± 1	73 ± 6	11 ± 1
$80 < p_T < 180$	82 ± 3	11 ± 1	76 ± 6	12 ± 2
$180 < p_T < 300$	79 ± 8	14 ± 2	86 ± 20	13 ± 5
Total	79 ± 1	11 ± 1	71 ± 1	13 ± 1

2018	EB		EE	
p_T range	Tight (%)	Loose (%)	Tight (%)	Loose (%)
$20 < p_T < 40$	83 ± 1	12 ± 1	72 ± 1	15 ± 1
$40 < p_T < 60$	83 ± 2	9 ± 1	74 ± 3	13 ± 1
$60 < p_T < 80$	86 ± 3	10 ± 1	85 ± 5	13 ± 1
$80 < p_T < 180$	89 ± 3	12 ± 1	91 ± 6	15 ± 1
$180 < p_T < 300$	105 ± 9	18 ± 2	108 ± 20	21 ± 6
Total	83 ± 1	11 ± 1	73 ± 1	15 ± 1

Table 6.6: Estimated prompt photon fraction in the Tight and Loose regions. The values are obtained selecting events, in data and MC, in the single photon phase-space for the $Z \to \mu^+\mu^-$ decay channel. The results are presented separated by years, binned with respect to the transverse momentum value and pseudorapidity of the photon candidate.

extracted by solving the matrix equation

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} \epsilon_1\epsilon_2 & \epsilon_1 f_2 & f_1\epsilon_2 & f_1 f_2 \\ \epsilon_1(1-\epsilon_2) & \epsilon_1(1-f_2) & f_1(1-\epsilon_2) & f_1(1-f_2) \\ (1-\epsilon_1)\epsilon_2 & (1-\epsilon_1)f_2 & (1-f_1)\epsilon_2 & (1-f_1)f_2 \\ (1-\epsilon_1)(1-\epsilon_2) & (1-\epsilon_1)(1-f_2) & (1-f_1)(1-\epsilon_2) & (1-f_1)(1-f_2) \end{pmatrix} \cdot \begin{pmatrix} \alpha_{\gamma\gamma} \\ \alpha_{\gamma j} \\ \alpha_{j\gamma} \\ \alpha_{jj} \end{pmatrix}, \quad (6.2)$$

where the indexes "1" and "2" refer to the leading and sub-leading photons candidate in the event. Once the α_{AB} are known, the data-driven method can be used to predict the number of events from each diphoton candidate source in every signal or control regions. As an example, the number of events from source $\alpha_{\gamma j}$ reconstructed in the control region N_{TL} is given by the product $\alpha_{\gamma j} \cdot M_{\gamma j}^{TL}$. The final number of background events from non-prompt photons in the double photon phase-space signal region $N_{TT}^{j\to\gamma}$ is the sum of

2016	EB		EE	
p_T range	ϵ	f	ϵ	f
$20 < p_T < 40$	0.827 ± 0.001	0.083 ± 0.004	0.885 ± 0.002	0.225 ± 0.009
$40 < p_T < 60$	0.778 ± 0.003	0.054 ± 0.007	0.844 ± 0.005	0.152 ± 0.019
$60 < p_T < 80$	0.776 ± 0.005	0.050 ± 0.014	0.829 ± 0.009	0.153 ± 0.037
$80 < p_T < 180$	0.776 ± 0.006	0.073 ± 0.019	0.840 ± 0.010	0.057 ± 0.047
$180 < p_T < 300$	0.786 ± 0.015	0.100 ± 0.088	0.913 ± 0.030	0.009 ± 0.219

2017	EB		EE	
p_T range	ϵ	f	ϵ	f
$20 < p_T < 40$	0.810 ± 0.001	0.124 ± 0.003	0.852 ± 0.002	0.274 ± 0.008
$40 < p_T < 60$	0.760 ± 0.003	0.098 ± 0.005	0.822 ± 0.005	0.174 ± 0.015
$60 < p_T < 80$	0.758 ± 0.005	0.096 ± 0.009	0.805 ± 0.009	0.155 ± 0.026
$80 < p_T < 180$	0.764 ± 0.005	0.075 ± 0.011	0.796 ± 0.009	0.144 ± 0.029
$180 < p_T < 300$	0.772 ± 0.013	0.126 ± 0.040	0.842 ± 0.026	0.113 ± 0.119

2018	EB		EE	
p_T range	ϵ	f	ϵ	f
$20 < p_T < 40$	0.813 ± 0.001	0.105 ± 0.003	0.840 ± 0.003	0.272 ± 0.008
$40 < p_T < 60$	0.773 ± 0.004	0.068 ± 0.006	0.796 ± 0.008	0.171 ± 0.016
$60 < p_T < 80$	0.760 ± 0.007	0.058 ± 0.010	0.795 ± 0.012	0.098 ± 0.029
$80 < p_T < 180$	0.764 ± 0.007	0.050 ± 0.012	0.796 ± 0.013	0.061 ± 0.033
$180 < p_T < 300$	0.768 ± 0.018	0.050 ± 0.012	0.788 ± 0.038	0.061 ± 0.033

Table 6.7: The probabilities for prompt photons and for a jets to be reconstructed as a tight isolated photon candidates. The ϵ and f values are obtained selecting events, in data and MC, in the single photon phase-space for the $Z \to \mu^+\mu^-$ decay channel. The results are presented separated by years, binned with respect to the transverse momentum value and pseudorapidity of the photon candidate. The prescription to avoid negative values when computing the $f(p_T, \eta)$ probabilities is applied only to the last p_T bin in the 2018 dataset.

the contribution of the different diphoton candidate sources involving at least one jet and it is estimated to be

$$\begin{aligned} N_{TT}^{j\to\gamma} &= M_{\gamma j}^{TT} \cdot \alpha_{\gamma j} + M_{j\gamma}^{TT} \cdot \alpha_{j\gamma} + M_{jj}^{TT} \cdot \alpha_{jj} \\ &= \epsilon_1 f_2 \cdot \alpha_{\gamma j} + f_1 \epsilon_2 \cdot \alpha_{j\gamma} + f_1 f_2 \cdot \alpha_{jj}. \end{aligned} \tag{6.3}$$

The distribution of the sum of the transverse momenta of the photon candidate pair $p_T^{\gamma\gamma}$ is exploited to extract the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross sections in both the electron and muon channels and to put limits on the anomalous Quartic Gauge Couplings. For this reason, both the normalisation and the shape of the non-prompt photon background need to be determined. This task is accomplished by the data-driven method by summing, bin-by-bin, each $N_{TT}^{j\to\gamma}$ contribution originating from candidate photon pairs sharing the same $p_T^{\gamma\gamma}$ value. This is possible since the N_{XY} and α_{AB} vectors are obtained with the same intervals scheme used for deriving the ϵ and f probabilities. The total number of M matrices needed to obtain the non-prompt photon background

shape is 100, corresponding to $5 \times 5 = 25$ possible combinations of the photon momenta and $2 \times 2 = 4$ possible locations of the reconstructed candidate photons pair in the ECAL detector (EB-EB, EB-EE, EE-EB and EE-EE). The background contribution from non-prompt photons is estimated separately for each year of data taking, thus accounting for any possible time dependence of the ϵ and f probabilities. This approach is justified by the fact that the f probabilities are expected to depend on the detector operating conditions. The background contributions obtained for each separate years of data taking are then summed to obtain the entire contribution expected for the full Run 2 dataset.

The statistical uncertainty $\sigma_{N_{TT}^{j\to\gamma}}$ on this background source is the sum in quadrature of two independent sources. The first and most important one is represented by the statistical uncertainty resulting from the limited number of events selected in the N_{XY} control regions. This contribution is taken into account when the matrix M is inverted to estimate the four possible source of diphoton candidates α_{AB}. The N_{XY} yields are assumed to be Poissonian distributed and uncorrelated between each other such that their associated statistical uncertainty is $\sigma_{N_{XY}} = \sqrt{N_{XY}}$. The statistical uncertainties on the N_{XY} yields are then propagated to each α_{AB} value using

$$\sigma_{\alpha_{AB}} = \sqrt{\sum_{ij}((M_{AB}^{ij})^{-1} \cdot \sigma_{N_{ij}})^2},$$

where the pair of indices ij runs over the four possible combinations TT, TL, LT and LL and the $((M_{AB}^{ij})^{-1}$ are the coefficients of the inverted matrix M, obtained by solving the linear system of Equation 6.2. The $\sigma_{\alpha_{AB}}$ uncertainties are then propagated to $\sigma_{N_{TT}^{j\to\gamma}}$ by using Equation 6.3.

Another contribution to $\sigma_{N_{TT}^{j\to\gamma}}$ is represented by the propagation of the statistical uncertainties on the ϵ and f probabilities which enter the matrix M. The statistical uncertainties $\sigma_\epsilon = \sqrt{n_T^{\gamma,MC}}$ and $\sigma_f = \sqrt{n_T^{\gamma,Data}}$ are defined under the assumption that the $n_T^{\gamma,MC}$ and the $n_T^{\gamma,Data}$ yields are Poissonian distributed. The statistical uncertainties on the ϵ probabilities are expected to be small, being the latter derived from high statistics MC simulations, while the corresponding statistical uncertainties on f are larger, especially for high photon p_T ranges, as can be seen in Table 6.7. This is motivated by the fact that the prompt photon contribution in data can be especially large, as can be seen in Table 6.6, and the remaining data yield $n_T^{\gamma,Data}$, after the MC prompt photon subtraction, can be small resulting in higher $\sigma_{n_T^{\gamma,Data}}$ values. The statistical uncertainties on the ϵ and f probabilities are propagated when solving the system of Equations 6.2 by constructing additional matrices M_i where, for each value of the i index, the ϵ or f probabilities referring to the leading or sub-leading photon are shifted up or down by their corresponding 1σ uncertainty. The output of the procedure just described consists on four pairs of additional estimates of $N_{TT}^{j\to\gamma}$, each pair corresponding to the $\pm1\sigma$ shifts. For each pair, the average between the up and down variation with respect to the nominal value of $N_{TT}^{j\to\gamma}$ is summed in quadrature to the $\sigma_{N_{TT}^{j\to\gamma}}$ value.

For events that have two jets misidentified as photons the matrix-element coefficients M_{jj}^{XY} are used, where the f probabilities of the leading and sub-leading photon candidate to be reconstructed in the N_{XY} region are multiplied together. Possible correlations

between the leading and sub-leading photons candidates have been investigated using dedicated differential distributions and no appreciable correlations effects were found.

The data-driven matrix method developed to estimate the non-prompt photon contribution in the diphoton phase-space has been validated in the single photon phase-space and consistent results between MC and data-driven predictions were found.

6.5 Systematic uncertainties

Determining the possible sources of systematic uncertainties related to the double photon phase-space event selection is of fundamental importance, both for the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section measurements, but also to put appropriate limits on the anomalous Quartic Gauge Couplings parameters. Several sources of systematic uncertainties have been considered to contribute to the signal region of this analysis. The different sources, described in the following, have been treated as correlated or uncorrelated between the different years of data taking and the Z boson electron and muons decay channels. The systematic uncertainties impact directly the predicted event yields, both for signal and background events. The general approach to treat the systematic uncertainties has been to obtain the predicted event yields, separately for the signal and for each background contributions, when each source is varied by its corresponding $\pm 1\sigma$ uncertainty. The nominal predicted event yield and all its related systematic variations represent the input parameters of the maximum likelihood fit procedure to the observed data, presented in Section 7.1.

6.5.1 Luminosity uncertainty

The systematic uncertainties on the luminosity recorded by the CMS detector during the different years of data taking considered in this work are described in Section 4.4. When combining all the contribution for the Run 2 dataset, a luminosity uncertainty of 1.8% is obtained. The systematic uncertainty on luminosity affects directly the signal and prompt photon background yield predictions by entering the MC normalisation factors w_{MC} of Equation 6.1. It also indirectly affects the non-prompt background, since the probability for a jet to be reconstructed as a tight photon candidate f is obtained subtracting from data the prompt photon contribution yield from simulation.

6.5.2 Pileup uncertainty

The systematic uncertainty on the pileup reweighting procedure (described in Section 6.3.1) is obtained by shifting the minimum bias cross section $\sigma_{min.\ bias}$ by its $\pm 1\sigma$ uncertainty. The pileup weights are then recomputed and propagated to the analysis. According to the most recent measurement of the inelastic proton proton cross section with the CMS detector at a centre of mass energy of $\sqrt{s} = 13\,\text{TeV}$, the total uncertainty on $\sigma_{min.\ bias}$ amounts to 4.6% [93]. The systematic uncertainty on the pileup reweighting procedure affects the signal and all the background contributions and it is considered as correlated between the different years of data taking.

6.5.3 Object selection uncertainties

The systematic uncertainties on the object selection efficiencies (described in Section 6.3.3) come from the Tag & Probe method. Each object selection efficiency uncertainty is taken into account by shifting the corresponding scale factor up and down by its corresponding 1σ uncertainty and then recomputing the predicted event yield for each signal and background source. The correlations of the systematic uncertainties on the efficiencies between the different years of data taking are derived under the following assumptions: since the systematic uncertainty of a measured value is dominated by biases associated with the measurement method, as long as the method remains the same, it is assumed that the method bias will be the same. The corresponding uncertainties are then considered as correlated when combining the different years of data taking. Conversely, when the systematic uncertainty is extrapolated directly through MC simulations its value can change between years if new criteria, new variables or new detectors are introduced in the computation. These uncertainties are then considered as uncorrelated between years. The systematic uncertainties considered for the objects selection criteria used in this work are described in the following.

6.5.3.1 Electron and photon objects selection uncertainties

Events containing electron candidates are corrected for their reconstruction, identification and trigger selection efficiencies, while events containing photons are corrected for their identification and their pixel seed veto efficiency. The uncertainties on the electron reconstruction efficiencies are considered as correlated between the different years of data taking, but uncorrelated with respect to the electron identification efficiencies. The electron and photon identification efficiencies are considered as fully correlated between years. The trigger selection efficiencies, specific to this analysis, are considered as correlated between years. Finally, the uncertainties on the photon pixels seed veto selection efficiencies are considered as uncorrelated between different years due to different pixel detector and running conditions.

6.5.3.2 Muon objects selection uncertainties

Events containing muon candidates are corrected for their identification, isolation and trigger selection efficiencies. Being all the respective scale factors derived using the same Tag & Probe workflow across the different years, the systematic uncertainties are considered as correlated.

6.5.4 ECAL trigger prefiring uncertainty

The systematic uncertainties related to the ECAL trigger prefiring corrections are obtained by shifting the prefiring probabilities $\epsilon_i^{\text{pref.}}$ (see Section 6.3.2) plus or minus their uncertainties and then recomputing the expected event yield. The uncertainty is taken as the convolution between an estimated systematic 20% uncertainty on the $\epsilon_i^{\text{pref.}}$ and the statistical uncertainty associated to the dedicated data sample used to derive the $\epsilon_i^{\text{pref.}}$ themselves.

6.5.5 Background contributions uncertainties

Several sources of systematic uncertainties affect the background contributions in the double photon phase-space. The prompt photon background yield is estimated entirely from simulated single photon and double photon exclusive samples, normalised to their theoretical cross section. The uncertainty on to the value of their cross sections must then be taken into account and propagated to the final predicted event yield. Moreover, the uncertainty on the cross section of the single photon exclusive samples plays a role also when estimating the non-prompt photon background. This a direct consequence of the subtraction of the prompt photon contribution to data reconstructed in the single photon phase-space to estimate the f probabilities (see also Section 6.4.2). Additional systematic uncertainties are considered to affect solely the non-prompt photon background, and are intended to cover all the particular features of the specifically developed data-driven method.

6.5.5.1 Prompt background contribution uncertainty

The prompt photon background in the $l^+l^-\gamma\gamma$ event selection is derived from MC simulations. Each simulated sample yield is normalised, using Equation 6.1, to its nominal cross section (the cross section value of each MC sample, together with the order of computation, is summarised in Table 6.2). To take into account the uncertainties related to the cross sections the estimated prompt photon background yield is reobtained by shifting the value of the cross section of each process by its $\pm 1\sigma$ uncertainty.

The process which accounts for almost all of the prompt photon background yield in both the $e^+e^-\gamma\gamma$ and $\mu^+\mu^-\gamma\gamma$ signal regions is the $Z\gamma$+jets sample, while the contribution from the other single photon exclusive sample considered is almost negligible.

There are different methods to estimate the systematic uncertainty to the $Z\gamma + \text{jets}$ cross section calculation. One way to do it is to compare directly the predictions of the prompt photon yield to the observed data for events reconstructed in the single photon phase-space. This procedure has the advantage of being completely independent of the theoretical models used to compute the theoretical cross section. Events with at least one isolated prompt photon have been reconstructed, separately in the Z electron and muon decay channels, both in data and simulation. The predicted signal and background contributions have been fitted to the observed data: the signal contribution is obtained from the $Z\gamma + \text{jets}$ simulated process, where each reconstructed level photon object must correspond to a genuinely generated photon, while the major background contribution is represented by Z+jets Drell-Yan process. Any other background contribution is assumed to be negligible. This assumption is well motivated as can be seen from the distribution of the invariant mass of the lepton pairs M_{ll} obtained with the selection criteria just described in the $Z \to e^+e^-$ (Figure 6.7(a)) and the $Z \to \mu^+\mu^-$ (Figure 6.7(b)) decay channels.

The choice of M_{ll} as the fitted distribution is motivated by the fact that this kinematical variable is able to separate events containing prompt photons from those where the photon candidates are likely to come from either prompt or non-prompt photons. Indeed, when looking at the invariant mass of the lepton pairs, events where a prompt photon is emitted by one of the leptons originating from the $Z \to l^+l^-$ boson decay (FSR photons) are well separated from events where the photon candidate could be

either a prompt photon, emitted by one of the two initial colliding quarks (ISR photons), or a non-prompt photon resulting from a jet in a Drell-Yan event reconstructed as a photon. This separation between FSR and ISR plus non-prompt photon is particularly visible in Figure 6.7: events where $M_{\ell\ell}$ is reconstructed in the $55 < M_{\ell\ell} < 75\,\text{GeV}$ range are almost entirely dominated by the $\text{Z}\gamma$ + jets processes while the region corresponding to the Z boson mass peak is contaminated by non-prompt photons.

The MC $\text{Z}\gamma$ + jets signal and Z + jets background contributions were fitted simultaneously to data using a binned χ^2 minimisation. Outside the Z boson mass peak region the only contribution to data is determined by prompt photon events and the χ^2 minimisation procedure is driven by the level of agreement between data and the $\text{Z}\gamma$ + jets MC sample prediction. The parameter of interest of the procedure is the fitted yield of the $\text{Z}\gamma$ + jets MC prediction. The results of the fitting procedure in the $\text{Z} \to \text{e}^+\text{e}^-$ and $\text{Z} \to \mu^+\mu^-$ decay channels are presented in Figure 6.7(c) and 6.7(d), while the fitted signal and background yields are summarised in Table 6.8. An improved agreement between data and predictions is obtained in the post-fit distribution in both the electron and muon decay channels.

Process	$\text{Z} \to \text{e}^+\text{e}^-$ decay channel	$\text{Z} \to \mu^+\mu^-$ decay channel
$\text{Z}\gamma$ + jets	0.946 ± 0.007	0.948 ± 0.004
Z + jets	1.22 ± 0.03	1.24 ± 0.02

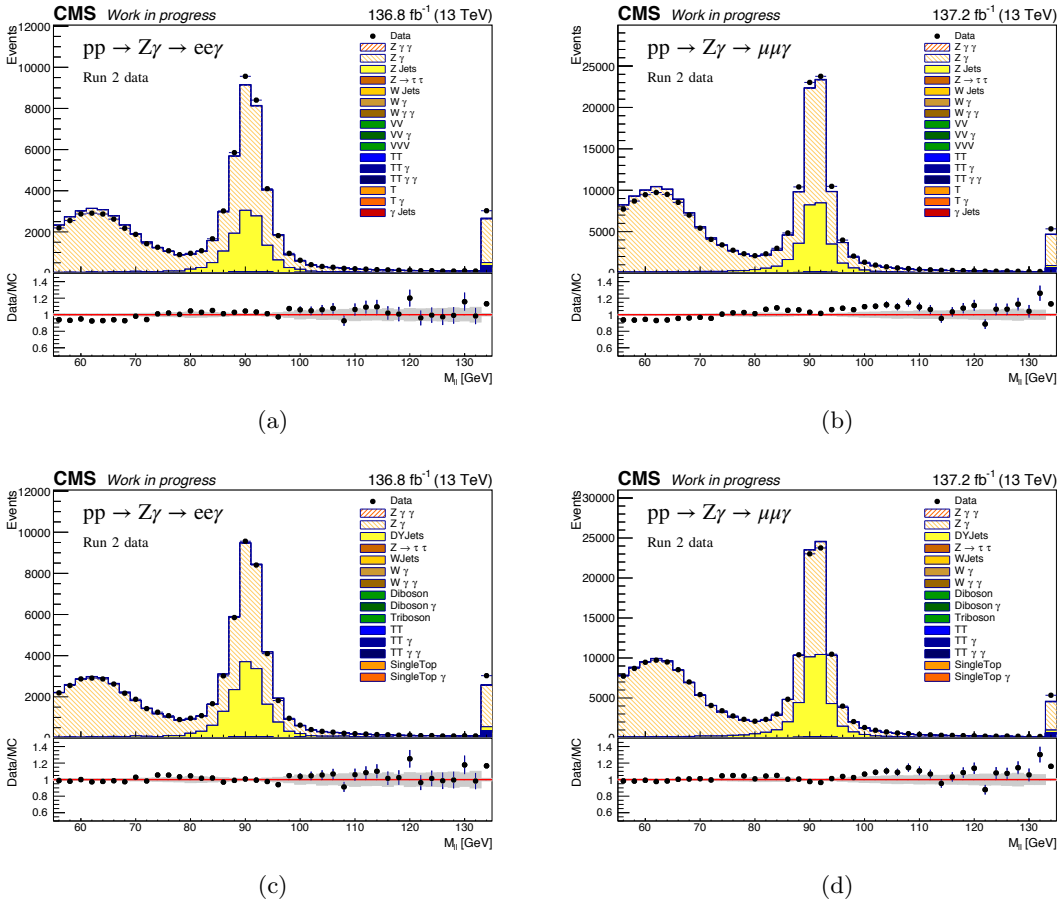
Table 6.8: The fitted percentage yield for the $\text{Z}\gamma$ + jets signal and Z + jets background predictions in the $\text{Z} \to \text{e}^+\text{e}^-$ and $\text{Z} \to \mu^+\mu^-$ decay channels. The numbers represent the percentage of each process yield with respect to the nominal one after the fitting procedure described in the text. The statistical uncertainty on the fitted parameters is shown as well. The post-fit data-based normalisations are used only to estimate the $\text{Z}\gamma$ + jets theoretical cross-section uncertainty.

The results of the fitting procedure are that, in order to improve the agreement with data, the $\text{Z}\gamma$ + jets signal yield should be reduced by an overall 5% with respect to the nominal pre-fit yield. The obtained results are validated by the consistent results obtained in the electron and muon decay channels, separately considered. Based on the fit results, an uncertainty on the $\text{Z}\gamma$ + jets theoretical cross section of 2.5% is assumed to cover any discrepancy between data and predictions. The theoretical uncertainty on the other single photon exclusive samples is conservatively assumed to be 7.5%, that is three times larger the one obtained with the $\text{Z}\gamma$ dedicated fitting procedure described above. Nevertheless, the contribution of the other single photon exclusive samples in the prompt background estimate is negligible.

6.5.5.2 Non-prompt background contribution uncertainties

The systematic uncertainties considered to affect the non-prompt background contribution are related to the methods used to derive the photon ϵ and f probabilities which enter in the matrix M of the double photon phase-space. Three sources of systematic uncertainties have been considered, and the methods developed to estimate their contributions are detailed in the following.

The first source of uncertainty affects only the f probabilities calculation and,



(a) (b)

(c) (d)

Figure 6.7: Distributions of the invariant mass of the lepton pairs M_{ll} obtained with the selection criteria described in the text for the $Z \to e^+e^-$ (left) and $Z \to \mu^+\mu^-$ (right) decay channels. The pre-fit (post-fit) distributions are shown on the top (bottom). The distributions are obtained exploiting the full Run 2 statistics. An improved agreement is observed between data and predictions in the post-fit distributions.

as explained at the beginning of this section, it is related to the subtraction of the prompt photon contribution from data in the single photon phase-space. Indeed the systematic uncertainty on the cross sections of the single photon exclusive samples must be propagated through the matrix method. This is done by shifting the theoretical cross sections of the MC samples up and down by their 1σ uncertainty and then by subtracting the new estimate for the prompt photon yield from data in the single photon phase-space to compute the f probabilities. The new M matrices are then used to estimate the non-prompt photon background contribution in the double photon phase-space. The impact of this systematic variation on the final result corresponds to the difference between the nominal background yield and the one obtained with the varied method.

The second source of uncertainty affects both the ϵ and f probabilities, and it is related to the characteristics of the single photon phase-space event selection. An alternative estimation of the ϵ and f probabilities is performed by varying the single

photon phase-space selection. Another single photon phase-space is defined in a W-*like* region, obtained with the same selection criteria of the nominal single photon phase-space described in Section 6.2.1, but requiring only one lepton in the event instead of two. To enrich the contribution from multi-jet events where a jet is misrecontructed as a lepton, the selected lepton is required to be not isolated. The single photon phase-space obtained with such selection is labelled as W-*like QCD enriched* region. The use of a W-*like* selection to obtain new estimates of the ϵ and f probabilities is necessary because a similar Z-*like* selection, requiring two non-isolated lepton candidates, would result in almost no selected events. The non-prompt background is then derived using the M matrices built from new estimates of the ϵ and f probabilities derived in the W-*like QCD enriched* region. The same procedures of Section 6.4.2 are followed. The obtained probabilities cannot be used directly in the double photon phase-space of this analysis: instead the percentage difference between the f probabilities obtained in the W-*like QCD enriched* region just described (f_{W}^{QCD}) and the ones obtained in a W-*like* region ($f_{\mathrm{W}}^{Nominal}$), similar to the W-*like QCD enriched* one but requiring the lepton to be isolated, has been computed. The nominal f probabilities of Table 6.7 are then modified according to

$$f_{\mathrm{Z}}^{QCD} = f_{\mathrm{Z}}^{Nominal} \cdot \left[1 + \frac{f_{\mathrm{W}}^{QCD} - f_{\mathrm{W}}^{Nominal}}{f_{\mathrm{W}}^{Nominal}} \right].$$

The new set of matrices obtained from the f_{Z}^{QCD} coefficients is then exploited to compute the non-prompt background contribution in the double photon phase-space. The systematic uncertainty is assessed by comparing the non-prompt background yield obtained through the procedure just described with the results obtained by following the standard procedure of Section 6.4.2. Half the difference between the two estimates is considered to be the systematic uncertainty related to the single photon phase-space event selection.

The third source of uncertainty is connected to the modelling of the prompt photon production features in MC simulation, such as the initial and final state radiation modelling or the energy scales of the particles. These effects could impact the shape of the simulated prompt photon distribution in the single photon phase-space that is subtracted to data when the f probabilities are calculated. In order to test possible *shape* effects, a different generator and hadronisation algorithm is used to obtain alternative MC samples. In this work, the Sherpa matrix element generator has been used to simulate the $\mathrm{Z}\gamma$+jets process instead of the nominal Madgraph5_aMC@NLO samples (see also Table 6.2). The aim is to compare the ϵ and f probabilities, obtained when using the nominal prompt photon prediction, with the probabilities obtained using the alternate prompt photon prediction. Since only the shape of the prompt photon alternative distribution is of interest, the prompt photon event yield in the single photon phase-space obtained by the $\mathrm{Z}\gamma$ + jets Sherpa samples is normalised to the yield obtained by the nominal Madgraph5_aMC@NLO samples. The alternative Sherpa MC samples are directly compared with the nominal Madgraph5_aMC@NLO ones for each year of data taking. The comparison is performed directly for the 2016 and 2017 years of data taking while the 2018 Madgraph5_aMC@NLO sample is compared with the 2017 alternative Sherpa sample since at the time of this analysis no 2018

corresponding alternative samples were available. The systematic uncertainty related to the MC shape effect is assessed as half the difference between the non-prompt background yield obtained by the nominal M matrix and the one obtained with a varied matrix obtained by using the alternative SHERPA samples.

6.6 Characterisation of the double photon phase-space signal region

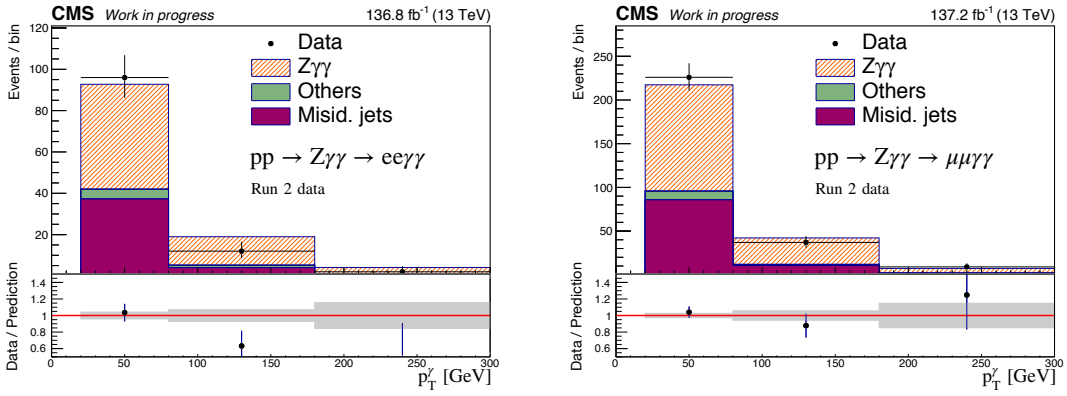
In this section, several events distributions obtained in the double photon phase-space with the procedures previously described, are presented. The $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ event yield, obtained differentially with respect to the transverse momentum of the leading (subleading) photon is shown in Figure 6.8 (6.9) while the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ event yield, obtained differentially with respect to the sum of the leading and subleading photon transverse momenta, is shown in Figure 6.10. The binning used in the different distributions is not uniform but optimised for data display purpose. The last bin of each distribution includes the overflow. The agreement between data and predictions is shown in the ratio at the bottom of each plot. In the ratio, the grey hashed area represents the statistical uncertainty on the sum of signal and background contributions while the uncertainty on the black dots is due to the statistical uncertainty of the data. The systematic uncertainties are not included in the plot.

The diphoton transverse momentum distributions are of central importance in this analysis, being the ones used to extract the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section (see Chapter 7) and to place limits on the $Z\gamma\gamma$ anomalous Quartic Gauge Couplings (see Chapter 8).

All the distributions presented in this section are obtained separately in the Z boson electron and muon decay channels by exploiting the full Run 2 statistics. The final event yield for data and predicted signal and background contributions are presented in Table 6.9. An overall good agreement, given the limited statistics available, is observed between data and predictions in both the $e^+e^-\gamma\gamma$ and $\mu^+\mu^-\gamma\gamma$ decay channels.

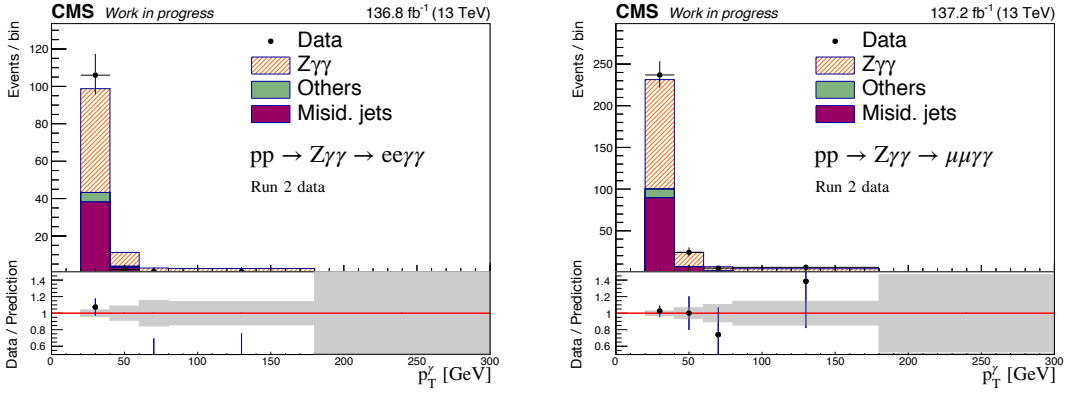
Process	$e^+e^-\gamma\gamma$ channel	$\mu^+\mu^-\gamma\gamma$ channel
Signal	$67.5 \pm 1.8(\text{stat.}) \pm 5.1(\text{syst.})$	$157.4 \pm 3.0(\text{stat.}) \pm 10.6(\text{syst.})$
Prompt bkg.	$6.1 \pm 1.2(\text{stat.}) \pm 0.5(\text{syst.})$	$11.1 \pm 2.1(\text{stat.}) \pm 0.7(\text{syst.})$
Non-prompt bkg.	$42.3 \pm 3.5(\text{stat.}) \pm 9.2(\text{syst.})$	$97.9 \pm 5.3(\text{stat.}) \pm 26.9(\text{syst.})$
Predicted bkg.	$48.3 \pm 3.7(\text{stat.}) \pm 9.0(\text{syst.})$	$109.0 \pm 5.7(\text{stat.}) \pm 26.5(\text{syst.})$
Total prediction	$115.8 \pm 4.1(\text{stat.}) \pm 8.0(\text{syst.})$	$266.3 \pm 6.4(\text{stat.}) \pm 23.3(\text{syst.})$
Observed data	$110.0 \pm 10.5(\text{stat.})$	$272.0 \pm 16.5(\text{stat.})$

Table 6.9: Summary of the observed and predicted number of events, obtained with the full Run 2 statistics for the double photon phase-space event selection in the electron and muon channels. The systematic uncertainty on the predicted background and total background categories is obtained by summing the different contributions (taking into account the statistical uncertainty) and then computing the variation from the reference result. The larger non-prompt background systematic uncertainty in the muon channel is related to the alternative estimation of the ϵ and f probabilities, performed by varying the single photon phase-space selection. In the electron channel the non-prompt background distribution is affected in such a way that the final event yield is similar to the reference one while this is not the same in the muon channel.



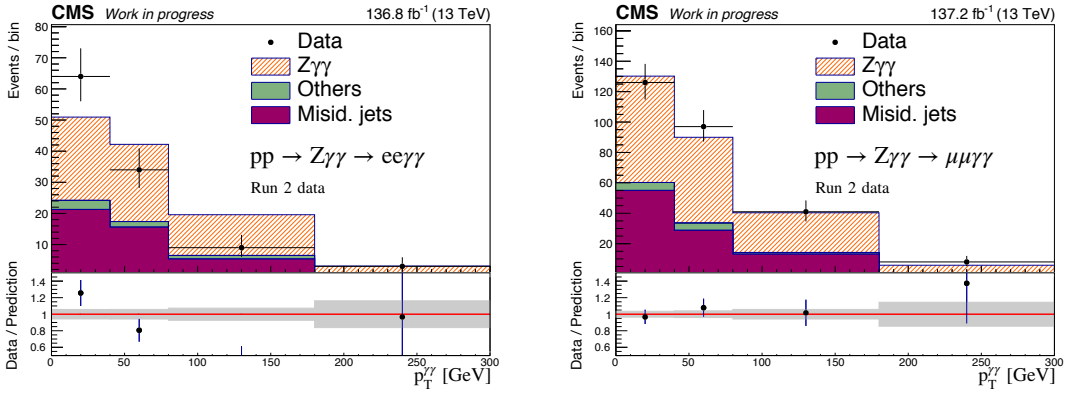
(a) (b)

Figure 6.8: Distribution of the leading photon transverse momentum in the double photon phase-space selection, obtained processing data coming from the entire Run 2 dataset. In Figure 6.8(a) events are reconstructed in the Z boson electron decay channel while in Figure 6.8(b) events are reconstructed in the Z boson muon decay channel. The $Z\gamma\gamma$ signal is in hatched orange, the jet misidentified as photon data-driven background in purple and the residual background in green. An overall good agreement between data and predictions, given the limited statistics available, can be observed.



(a) (b)

Figure 6.9: Distribution of the subleading photon transverse momentum in the double photon phase-space selection, obtained processing data coming from the entire Run 2 dataset. In Figure 6.9(a) events are reconstructed in the Z boson electron decay channel while in Figure 6.9(b) events are reconstructed in the Z boson muon decay channel. The Zγγ signal is in hatched orange, the jet misidentified as photon data-driven background in purple and the residual background in green. An overall good agreement between data and predictions, given the limited statistics available, can be observed.



(a) (b)

Figure 6.10: Distribution of the diphoton transverse momentum in the double photon phase-space selection, obtained processing data coming from the entire Run 2 dataset. In Figure 6.10(a) events are reconstructed in the Z boson electron decay channel while in Figure 6.10(b) events are reconstructed in the Z boson muon decay channel. The Zγγ signal is in hatched orange, the jet misidentified as photon data-driven background in purple and the residual background in green. An overall good agreement between data and predictions, given the limited statistics available, can be observed.

7 | Zγγ production cross section measurement at 13 TeV

In this chapter, events selected in the double photon phase space signal region are used to measure the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section. The cross section is obtained with a maximum likelihood fit to the observed data yield, separately in the $Z\gamma\gamma \to e^+e^-\gamma\gamma$ electron channel, in the $Z\gamma\gamma \to \mu^+\mu^-\gamma\gamma$ muon channel and in the combined $Z\gamma\gamma \to l^+l^-\gamma\gamma$ leptonic decay channel. The signal strength μ, which represents the ratio of the measured cross section to the theoretical one, is treated as a free parameter in the fitting procedure. The sources of systematic uncertainties, described in Section 6.5, are included in the fitting procedure as nuisance parameters.

The text is organised as follows. In Section 7.1 the general approach used to measure the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section is given. In Section 7.1.1 the statistical tools are presented followed by the post-fit results of Section 7.1.2. Finally a discussion of the results obtained is given in Section 7.2.

7.1 Fiducial cross-section measurement

One of the most common task in physics is to infer some model parameter from data. To accomplish this task an *estimator* is defined and then used to obtain the true value of the parameters of interest $\vec{\alpha}$. This analysis exploits the most widely used estimator in physics, the *maximum likelihood estimator* or *best-fit estimator*, which is defined as the value of $\vec{\alpha}$ which maximises the *likelihood function* $L(\vec{\alpha})$. The generic likelihood function is defined as

$$L(\vec{\alpha}) \sim p(data|\vec{\alpha}),$$

where $\vec{\alpha}$ are parameters of the likelihood and $p(data|\vec{\alpha})$ is the probability of observing the actual data, for a particular value of $\vec{\alpha}$. It has to be noticed that the likelihood function must not be interpreted as a probability density for $\vec{\alpha}$. Indeed its value at maximum is of no use, while the only interesting outcome of the maximisation procedure is the best-fit estimator $\hat{\vec{\alpha}}$.

The vector of parameters $\vec{\alpha}$ contains both the physics parameter of interest (POI) and the nuisance parameters $\vec{\theta}$ related to the systematic uncertainties contribution on the background determination. In this analysis, the specific POI is the Standard Model (SM) prediction of the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ production cross section $\sigma^{meas.}_{Z\gamma\gamma}$ and the data are the observed event yield in the double photon phase space. The final goal is to compare the measured value of $\sigma^{meas.}_{Z\gamma\gamma}$ against the SM prediction.

To separate the POI from the other uninteresting parameters, the $\vec{\alpha}$ vector is then rewritten as

$$\vec{\alpha} = (\sigma_{\mathrm{Z}\gamma\gamma}, \vec{\theta}).$$

Typically the nuisance parameters are constrained by some external measurements. These constraints are parametrised by additional *constraint terms*

$$\pi(\vec{\theta}_0|\vec{\theta}) \sim p(\vec{\theta}|\vec{\theta}_0),$$

where $\vec{\theta}_0$ are the measured values of the nuisance parameters to be simultaneously constrained with the POI in the fitting procedure. The global likelihood function, which includes the POI and the nuisance parameters constrain, is defined as

$$L(\sigma_{\mathrm{Z}\gamma\gamma}, \vec{\theta}) \sim p(data|\sigma_{\mathrm{Z}\gamma\gamma}, \vec{\theta}) \cdot \pi(\vec{\theta}_0|\vec{\theta}). \tag{7.1}$$

In the following section the specific model used in the analysis to obtain the best-fit estimator $\sigma^{meas.}_{\mathrm{Z}\gamma\gamma}$ is presented.

7.1.1 Building the statistical model

In general, the outcome of a scattering experiment is a number n of observed events, which can be related to the predicted expected event count ν by a probability distribution function $f(n, \nu)$. The probability to observe exactly n events, given an expected count ν, is assumed to be Poisson distributed so that $f(n, \nu)$ takes the following explicit form

$$f(n, \nu) = e^{-\nu} \frac{\nu^n}{n!}. \tag{7.2}$$

The approximation is appropriate given the relatively modest final observed event yield in the signal region of this analysis (see Table 6.9).

Events selected in the double photon phase space are assumed to be represented as the sum of signal and background contributions, thus the number of expected events can be written as

$$\nu = \nu_s + \nu_b,$$

where ν_s and ν_b are, respectively, the expected event yields for the signal and background processes. The SM is assumed to be the reference theory predicting a value $\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}$ for the $\mathrm{pp} \to \mathrm{Z}\gamma\gamma \to \mathrm{l}^+\mathrm{l}^-\gamma\gamma$ production cross section. The number of predicted signal events can be obtained by combining the production cross section, the measured luminosity, and the reconstruction efficiency $\epsilon = n_{reco.}/n_{theo.}$, which is given by the ratio of the number of reconstructed events in the double photon phase space (see Section 6.2.1) to the expected event yield in the fiducial phase space (see Section **??**), as obtained from a MC simulation. The number of predicted signal events is then explicitly given by

$$\nu_s = \mu \cdot \epsilon \cdot \sigma^{theo.}_{\mathrm{Z}\gamma\gamma} \cdot \mathcal{L}$$

where $\mathcal{L}$ is the luminosity, $\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}$ is the theoretical predicted cross section in the fiducial phase space used in this work and μ is the signal strength parameter. This rate modifier is intended to reproduce any difference between the observed data and the expected yield. A value of μ equal to one means that the observed data are in perfect agreement

with the expected signal and background process yields (the *signal hypothesis*). On the other side, a value of μ equal or close to zero corresponds to the *background-only hypothesis*, such as the predicted background contribution is sufficient to describe the observed data and no hint of signal events is present.

Having defined the signal strength μ and the theoretical fiducial cross section $\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}$, Equation 7.2 can be rewritten as

$$f(n|\mu,\nu_b) = e^{-(\mu\cdot\epsilon\cdot\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}\cdot\mathcal{L}+\nu_b)} \cdot \frac{(\mu\cdot\epsilon\cdot\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}\cdot\mathcal{L}+\nu_b)^n}{n!}. \tag{7.3}$$

The likelihood function for the observed data n is a function of the signal strength μ and it is defined as

$$L(\mu) = f(n|\mu,\nu_b). \tag{7.4}$$

In the absence of any systematic uncertainty, the best-fit estimate $\hat{\mu}$ and its associate confidence interval are simply obtained by maximising the likelihood function in Equation 7.4. The measured fiducial cross section is then given by

$$\sigma^{meas.}_{\mathrm{Z}\gamma\gamma} \equiv \hat{\mu}\cdot\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}. \tag{7.5}$$

Equation 7.5 does not imply that the measured cross section $\sigma^{meas.}_{\mathrm{Z}\gamma\gamma}$ depends directly on the theoretical predicted one $\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}$, as the signal strength μ will adapt to describe the observed data for different $\sigma^{theo.}$ input values (provided that the reconstruction efficiency ϵ does not change).

The likelihood function used in this work is more sophisticated than the one defined in Equation 7.4, but it shares with it the same basic principles. The signal strength is estimated by maximising a binned likelihood function built from the distribution of events reconstructed in the double photon phase space as a function of the diphoton transverse momentum. The distributions being fitted are the ones in Figure 6.10. Given a set of n_{bin} bins, the probabilities f_i to observe a number of event n_i in each bin i have the same functional form of Equation 7.3, and are given by

$$f_i(n_i|\mu,\nu_{b,i}) = e^{-(\mu\cdot\epsilon_i\cdot\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}\cdot\mathcal{L}+\nu_{b,i})} \cdot \frac{(\mu\cdot\epsilon_i\cdot\sigma^{theo.}_{\mathrm{Z}\gamma\gamma}\cdot\mathcal{L}+\nu_{b,i})^{n_i}}{n_i!},$$

where ϵ_i is the reconstruction efficiency for the i-th bin. The observed number of events is assumed to be distributed independently in each bin, such as the global likelihood function can be written as the product of the single probabilities f_i

$$L(\mu) = \prod_{i=1}^{n_{bin}} f_i(n_i|\mu,\nu_{b,i}). \tag{7.6}$$

Up to now the uncertainties in the determination of the underlying models, that are the systematic uncertainties on the predicted number of background events $\nu_{b,i}$, were not included in the discussion. Any such systematic uncertainty will reduce the accuracy with which the predicted number of background events is known. The systematic uncertainties considered to affect the background predictions are the ones presented in Section 6.5. The Barlow and Beeston approach [96] is used to incorporate, in the maximum likelihood

fit, the fact that the MC samples statistics used are finite and thus subject to statistical fluctuations [97]. The systematic uncertainties are modelled by introducing in the likelihood function additional nuisance parameters $\vec{\theta}$. Each nuisance parameter is assumed to be constrained by a Gaussian distribution function $\pi(\theta_0|\theta) = G(\theta|\theta_0, \sigma_\theta)$, where θ_0 is the measured pre-fit value of the nuisance parameter and σ_θ its related pre-fit uncertainty. The likelihood function in Equation 7.6 is then expanded to include these additional constrains

$$L(\mu, \vec{\theta}) = \prod_{i=1}^{n_{bin}} f_i(n_i|\mu, \nu_{b,,i}(\vec{\theta}_i)) \cdot G(\vec{\theta}_i|\vec{\theta}_{0,i}, \sigma_{\vec{\theta}_i}), \tag{7.7}$$

where it has to be noticed that $\vec{\theta}_i$ and not $\vec{\theta}_{0,i}$ are used to maximise the likelihood. This is of particular importance since, as the best-fit parameters $\vec{\theta}$ diverges from their input values $\vec{\theta}_{0,i}$, the value assumed by likelihood function is reduced. The systematic uncertainties result in an overall distortion in the shape of the observed spectrum, which is parametrised by the *vertical morphing* technique [97].

The observed events in the double photon phase space are interpreted in terms of the SM predictions via the maximisation of the likelihood function defined in Equation 7.7. To avoid dealing with small or large values of the likelihood function its negative logarithm is actually considered and the maximisation procedure is translated into a minimisation problem. The negative logarithm of the likelihood function is minimised using the MINUIT software package in the RooFit framework [98]. The outcomes of the minimisation procedure are referred to as *best-fit* parameters, and distributions obtained with the best-fit parameters are labelled as *post-fit distributions*. Since the only parameter of interest of this analysis is the signal strength μ, the other nuisance parameters are removed in the fitting procedure using the *profiling method*. The idea is to find the values for $\vec{\theta}$ which maximise the likelihood at each possible value of μ. The value of $L(\mu, \vec{\theta})$, along the multidimensional manifold defined by $\hat{\vec{\theta}}(\mu)$ for which $L(\mu, \vec{\theta})$ is maximum, is referred as the *profiled likelihood*

$$L(\mu, \theta) \rightarrow L(\mu, \hat{\vec{\theta}}(\mu)) \equiv \max_{\vec{\theta}} L(\mu, \vec{\theta}).$$

After being profiled, the implicit dependence of the likelihood function on $\vec{\theta}$ can be dropped obtaining $L(\mu, \theta) \rightarrow L(\mu)$. The interval of confidence of the POI best fit value $\hat{\mu}$ is obtained by an explicit scan of the profiled likelihood on a chosen grid of parameter values. The scanning procedure allows for a straightforward check that no unexpected features in the shape of the likelihood function are present. Moreover, it will also account for possible deviations of the likelihood function from the simple parabola approximation. The obtained likelihood values are interpolated to find the likelihood curve around the minimum $\hat{\mu}$ [99]. In the following, for display purposes, the likelihood curve is redefined by subtracting the value at its minimum, that is

$$-\ln L(\mu) \rightarrow -\ln L(\mu) - (-\ln L(\hat{\mu})) \equiv -\Delta \ln L(\mu).$$

The confidence interval around the best fit value $\hat{\mu}$ is obtained by applying the Wilks theorem, such as that the 68% confidence interval corresponds to the region for which $-\Delta \ln L(\mu) < 1$ [100].

The likelihood fit can also be exploited to quantify the statistical significance of an observed excess of data over the background-only hypothesis. This is a standard hypothesis test where one wants to exclude the possibility that the observed data are produced only by background processes. The test-statistic q_0 is then defined according to

$$q_0 = -2 \ln \frac{L(\mu = 0, \hat{\hat{\vec{\theta}}}_0)}{L(\hat{\mu}, \hat{\hat{\vec{\theta}}})},$$

that is the ratio of two profiled likelihoods, one in which the signal strength is set to zero and the other in which it is free to float. The *significance*, or p-value, for rejecting the background hypothesis $\mu = 0$ can therefore be calculated as

$$p_0 = \int_{q_0^{meas.}}^{+\infty} f(q_0|\mu = 0) \mathrm{d}q_0,$$

where the distribution of the test-statistics $f(q_0|\mu = 0)$ can be determined using Wilks theorem, provided the number of events is large enough [100]. Conventionally, a discovery can be claimed if the measured p-value is found to be equal or less than 0.287×10^{-6}, which corresponds to a significance of 5σ of the signal plus background hypothesis.

7.1.2 Post-fit results

In Figure 7.1 the scans of the profiled likelihood in the electron, muon and combined lepton channels are shown. The total uncertainty σ_{tot} on the best-fit value $\hat{\mu}$ is obtained from the abscissa corresponding to the intersection of the $-\Delta \ln L(\mu)$ curve with a straight line at $-\Delta \ln L(\mu) = 1$. The total uncertainty is also split into two components, corresponding to the contribution from the statistical and systematic uncertainties, in order to illustrate which aspects of them dominate. The statistical component $\sigma_{stat.}$, which is represented in Figure 7.1 by the red parabola, is defined as the post-fit uncertainty one would have measured if all the systematic uncertainties were fix to zero. It is obtained by freezing the nuisance parameters to their best fit value when performing the scan of the POI. The systematic component is defined as $\sigma_{syst} = \sqrt{\sigma_{tot}^2 - \sigma_{stat.}^2}$. The best-fit value of the measured signal strength in the different Z boson decay channels is summarised, together with the observed signal significance, in Table 7.1.

Channel	Signal strength	Significance (σ)
$pp \to Z\gamma\gamma \to e^+e^-\gamma\gamma$	$0.73^{+0.18}_{-0.17}(\text{stat.})^{+0.12}_{-0.13}(\text{syst.})$	3.4
$pp \to Z\gamma\gamma \to \mu^+\mu^-\gamma\gamma$	$1.06^{+0.11}_{-0.11}(\text{stat.})^{+0.10}_{-0.10}(\text{syst.})$	5.4
$pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$	$0.91^{+0.10}_{-0.10}(\text{stat.})^{+0.11}_{-0.12}(\text{syst.})$	4.8

Table 7.1: The best-fit value for the signal strength μ obtained by fitting, separately, the $pp \to Z\gamma\gamma \to e^+e^-\gamma\gamma$ and the $pp \to Z\gamma\gamma \to \mu^+\mu^-\gamma\gamma$ decay channels. The electron and muon channel signal strengths are compatible within two standard deviations when considering only the statistical uncertainties. In the last row the electron and muon decay channels are combined together for the fitting procedure. The signal significance for each channel is reported as well.

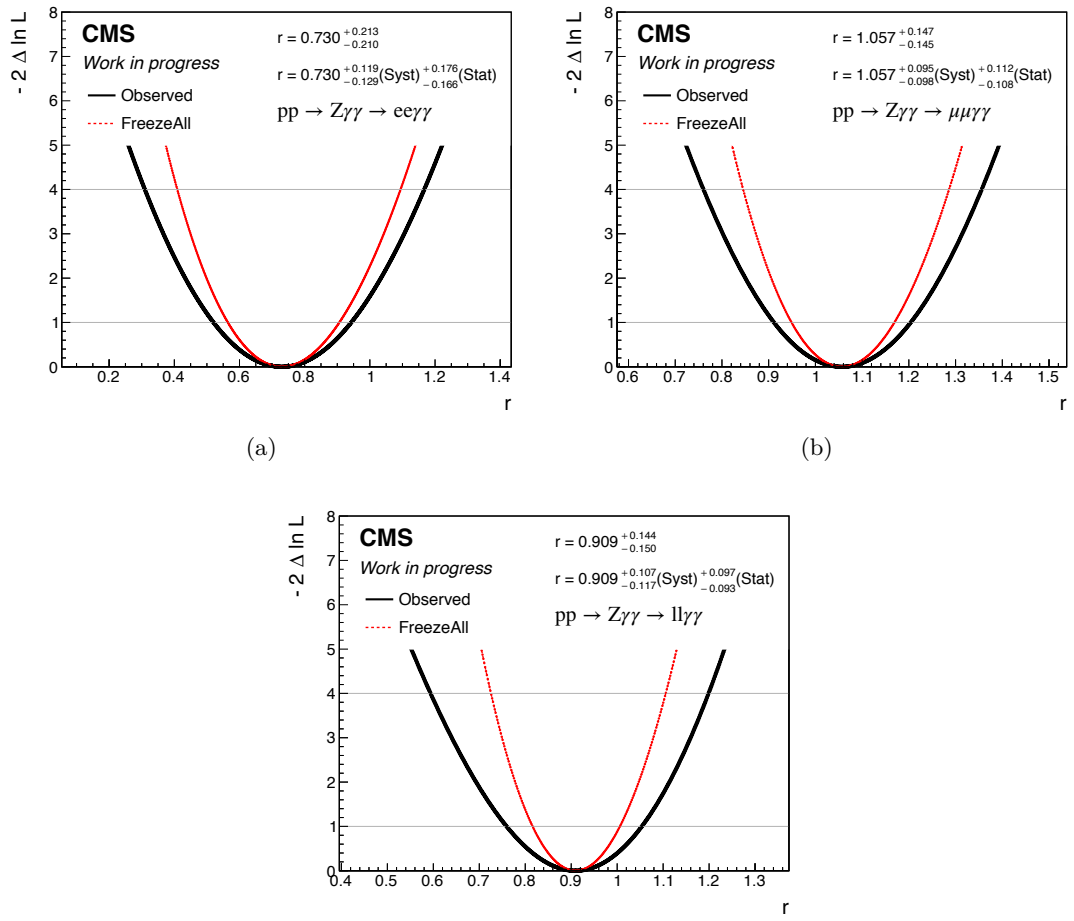


Figure 7.1: Scan of the profiled likelihood $-\Delta \ln L(\mu)$ for the double photon phase space event selection in the Z boson electron (7.1(a)), muon (7.1(b)) and combined decay channels (7.1(c)). The "freeze all" parabola corresponds to the likelihood scan when all the systematic uncertainties are fixed to their best fit outcomes. The signal strength parameter μ is labelled as r in the plots.

It is also useful to examine in detail the effects of each systematic uncertainty on the signal strength measurement. This is often referred to as calculating the *impact* of each nuisance parameters $\vec{\theta}$ on the best-fit value of the parameter of interest μ. The impact is defined as the shift $\Delta\mu$ that is induced as each best-fit value θ is shifted to its $\pm 1\sigma$ post-fit uncertainty, with all other parameters profiled as normal. This is effectively a measure of the correlation between the nuisance parameters and the POI and it is useful for determining which nuisance parameters have the largest effect on the POI uncertainty. The post-fit impacts of each nuisance parameter on the best-fit value of the signal strength μ measured in the electron, muon and combined decay channel of the Z boson are summarised in Table 7.2 and shown in Appendix C.

Systematic uncertainty	$e^+e^-\gamma\gamma$ (%)	$\mu^+\mu^-\gamma\gamma$ (%)	$l^+l^-\gamma\gamma$ (%)
Luminosity	3	1	3
Pileup	2	< 1	1
Electron scale factors	3	< 1	1
Muon scale factors	2	< 1	1
Photon scale factors	6	5	5
L1 ECAL prefiring	< 1	< 1	< 1
Non-prompt normalisation	6	5	6
Non-prompt shape	2	1	1
Prompt normalisation	7	5	4
MC samples stat.	7	3	4

Table 7.2: The post-fit impacts (in percent, rounded to the closest integer) of each nuisance parameter on the best-fit signal strength (see also Appendix C).

7.2 Discussion of the results

The measured cross sections are obtained by fitting the data while leaving the signal strength and the nuisance parameters free to float. The results obtained in the electron and muon channels are

$$\sigma(\mathrm{Z}(\mathrm{e}^+\mathrm{e}^-)\gamma\gamma)^{\mathrm{meas.}} = 4.38^{+1.06}_{-0.99}(\mathrm{stat.})^{+0.72}_{-0.77}(\mathrm{syst.})\,\mathrm{fb}$$

and

$$\sigma(\mathrm{Z}(\mu^+\mu^-)\gamma\gamma)^{\mathrm{meas.}} = 6.32^{+0.67}_{-0.64}(\mathrm{stat.})^{+0.57}_{-0.58}(\mathrm{syst.})\,\mathrm{fb}.$$

The dominant uncertainties are those associated to the non-prompt normalisation and to the photon scale factors. Other relevant sources are represented by the prompt photon background normalisation and the limited statistics of the MC samples. The total systematic uncertainties for the cross section measurements in the $pp \to Z\gamma\gamma \to e^+e^-\gamma\gamma$ and $pp \to Z\gamma\gamma \to \mu^+\mu^-\gamma\gamma$ channels are found to be 18% and 10%, while the statistical uncertainties amount to 25% and 10% respectively. The fact that systematic and statistical uncertainties give a similar contribution is in agreement with the larger event yield in the muon channel with respect to the electron one, and it is also reflected by the larger observed significance. When the electron and muon channels are combined,

the total systematic uncertainty is found to be 13%, while the statistical uncertainty amounts to 11%, and the measured cross section is

$$\sigma(\mathrm{Z}(\mathrm{l}^+\mathrm{l}^-)\gamma\gamma)^{\mathrm{meas.}} = 5.44^{+0.58}_{-0.56}(\mathrm{stat.})^{+0.64}_{-0.70}(\mathrm{syst.})\,\mathrm{fb},$$

The measured cross section is in good agreement with the theoretical value

$$\sigma(\mathrm{Z}(\mathrm{l}^+\mathrm{l}^-)\gamma\gamma)^{\mathrm{theo.}} = 6.00 \pm 0.07(\mathrm{pdf} + \mathrm{scale})\,\mathrm{fb},$$

where the theoretical uncertainty on the signal cross section predicted by MadGraph5_aMC@NLO comes from the parton distribution functions set, the factorisation and renormalisation scales and the strong coupling constant determination. These different contributions have been estimated following the prescriptions given in reference [101] with a dedicated Rivet [102, 103] routine. Theoretical uncertainties on the selection efficiency, related to the parton distribution function set, are estimated using the set of 100 replicas of the NNPDF3.1 PDFs set and are found to be 0.8%, while uncertainties related to the renormalisation scale μ_R and factorisation scale μ_F choice are estimated to be 0.6%. Finally, uncertainties related to the value of the strong coupling constant are estimated to be 0.4%.

8 | Limits on Anomalous Quartic Gauge Couplings on Zγγ processes at 13 TeV

In this chapter, the pp $\rightarrow$ Z$\gamma\gamma$ $\rightarrow$ $l^+l^-\gamma\gamma$ events are used to investigate the presence of new physics parametrised as anomalous couplings between the Z boson and the photons. For such studies, the $l^+l^-\gamma\gamma$ event yield as a function of a kinematic variable particularly sensitive to the anomalous couplings is used. Events are selected in the same phase space used for the fiducial cross section measurement. No evidence of new physics is obtained, and limits on the values of the anomalous couplings are placed.

This chapter is organised as follows. In Section 8.1 the production of the dedicated anomalous coupling signal sample is presented, while in Section 8.2 the statistical analysis of the results is given together with the obtained limits. Finally, a discussion of the obtained results and a comparison with other previous measurements are presented in Section 8.3.

8.1 Anomalous quartic gauge couplings sample production

For the study of the anomalous Quartic Gauge Couplings (aQGC), a dedicated Monte Carlo (MC) sample has been produced using the MADGRAPH5_aMC@NLO matrix element generator. The same initialisation parameters of the reference Z$\gamma\gamma$ signal sample were used, with the only difference being the underlying physics model. For the aQGC studies, events are simulated at leading order using the *AQGC_ALL_UFO_new* model [104]. This model adds, to the Standard Model (SM) Lagrangian, eight independent dimension 8 operators $\mathcal{O}_{T,i}$ which are expected to contribute to the pp $\rightarrow$ Z$\gamma\gamma$ channel (see Section 3.2).

To simulate the effect of the aQGC into the Z$\gamma\gamma$ sample a *reweighting procedure* has been used. It consists in generating Z$\gamma\gamma$ events under the SM assumption and associating to each of them additional weights which correspond to a scan of the aQGC parameters linked to the eight $\mathcal{O}_{T,i}$ operators. The range of scan of the aQGC parameters is presented in Table 8.1. The additional weights w_{aQGC} are based only on the matrix

element computation, and are obtained by using the following relation

$$w_{aQGC} = \frac{|M_{aQGC}|^2}{|M_{SM}|^2} \cdot w_{SM},$$

where the M_i terms represent the matrix element of the process. This approach has the advantage of not requiring the generation of multiple MC samples for each parameter under study. The SM situation, where all the parameters are set to zero, has been used as a reference.

Parameter	Range of scan [TeV^{-4}]
$f_{T,0}$	-20, 20
$f_{T,1}$	-20, 20
$f_{T,2}$	-40, 40
$f_{T,5}$	-40, 40
$f_{T,6}$	-40, 40
$f_{T,7}$	-40, 40
$f_{T,8}$	-20, 20
$f_{T,9}$	-20, 20

Table 8.1: The range of variation for the scan of the aQGC coupling parameters used in this work.

The presence of nonzero aQGC parameters is expected to enhance the event yield in the right tail of the diphoton transverse momentum $p_T^{\gamma\gamma}$ distribution. Figure 8.1 shows the transverse momentum of the diphoton system, in blue for the SM prediction, and in red after the introduction of an arbitrary value for the aQGC parameters $f_{T,0}$ associated to the $\mathcal{O}_{T,0}$ operator. The different behaviour of the two sets can be clearly seen and hints to the fact that the most sensitive region for the study of deviations from the SM predictions is the one at high momenta.

8.2 Anomalous quartic gauge coupling limits extraction

A statistical framework, similar to the one developed in Section 7.1.1, has been used to study different aQGC parameters. The systematic uncertainties are included in the analysis and electron and muon channels are combined to increase the sensitivity to the aQGC parameters. The distribution of the transverse momentum of the diphoton system $p_T^{\gamma\gamma}$ is the one fitted in the limits extraction procedure.

8.2.1 Building the statistical model

Fits to the likelihood function are used in this section to establish confidence intervals for the aQGC parameters. The test statistics, used to probe exclusion limits for the alternative models, is defined as

$$t_\alpha = -2 \ln \lambda(\alpha)$$

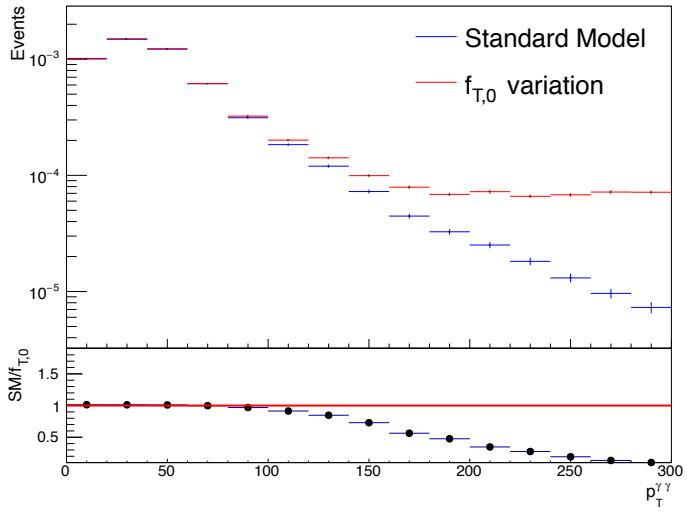


Figure 8.1: Transverse momentum of the diphoton system. The SM prediction is in blue while the predicted event yield after the introduction of an arbitrary value for the aQGC parameter $f_{T,0}$ is in red.

where $\lambda(\alpha)$ is the profiled likelihood ratio

$$\lambda(\alpha) = \frac{L(\alpha, \hat{\vec{\theta}})}{L(\hat{\alpha}, \hat{\vec{\theta}})}. \tag{8.1}$$

The likelihood function in Equation 8.1 is very similar to the one in Equation 7.7, with the important difference that the signal strength μ is now a function of α. The rate modifier $\mu(\alpha)$ is defined, for each point α of the aQGC parameter scan, as the ratio of the BSM predicted yield to the SM one. The ratio is calculated for each point of the parameter scan and then fitted with a second-order polynomial to extrapolate its dependence over the full range of the values. An example of the polynomial fit is presented Figure 8.2. This procedure is implemented for each coupling parameter considered, separately for the electron and muon channels. Limits on the aQGC values are then extracted from the best-fit value $t_\alpha^{meas.}$, by assuming that the distribution of the test statistics can be approximated by a χ^2 distribution with one degree of freedom.

8.2.2 Post-fit results

The measured limits on the aQGC values, obtained separately in the electron, muon and combined channels, are shown in Figure 8.3. The expected and observed limits obtained after combining the electron and muon channels are summarised in Table 8.2. The expected limits are obtained replacing the observed data with the MC predictions.

8.3 Discussion of the results

The measured limits on the coupling parameters $f_{T,i}$, corresponding to the aQGC operators $\mathcal{O}_{T,i}$, can be compared with the same quantities obtained in other analyses

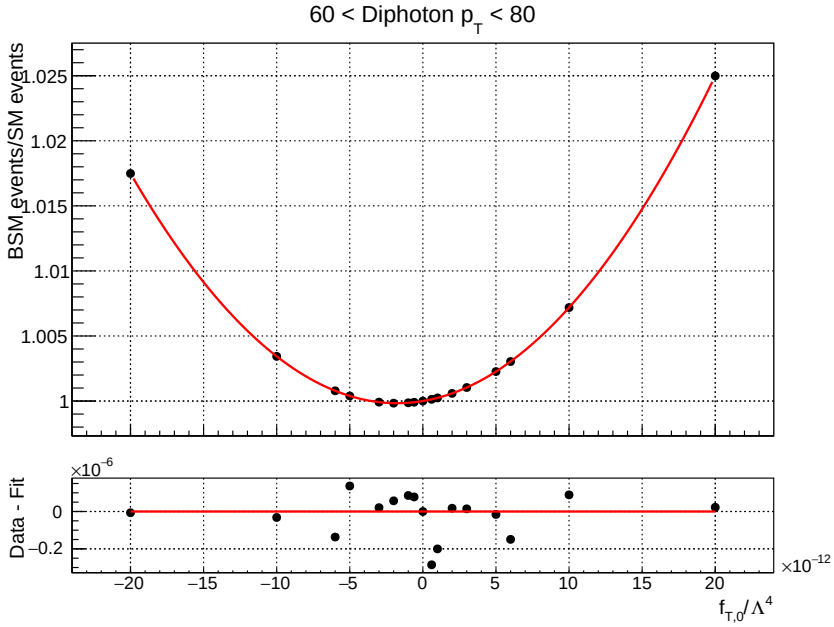


Figure 8.2: Example of the ratio of the aQGC to SM yield, fitted with a quadratic function in the electron channel, for the $f_{T,0}$ parameter. The red curves correspond to the functional form of $\mu(\alpha)$ in a particular diphoton transverse momentum bin.

	$e^+e^-\gamma\gamma$ (TeV^{-4})		$\mu^+\mu^-\gamma$ (TeV^{-4})		$l^+l^-\gamma\gamma$ (TeV^{-4})	
Parameter	Expected	Observed	Expected	Observed	Expected	Observed
$f_{T,0}/\Lambda^4$	$[-6.46, 6.26]$	$[-6.38, 6.14]$	$[-5.50, 5.30]$	$[-6.54, 6.38]$	$[-4.86, 4.66]$	$[-5.70, 5.46]$
$f_{T,1}/\Lambda^4$	$[-6.46, 6.26]$	$[-6.38, 6.14]$	$[-5.50, 5.30]$	$[-6.54, 6.38]$	$[-4.86, 4.66]$	$[-5.70, 5.46]$
$f_{T,2}/\Lambda^4$	$[-12.9, 12.5]$	$[-12.8, 12.3]$	$[-11.1, 10.7]$	$[-13.2, 12.8]$	$[-9.72, 9.32]$	$[-11.4, 10.9]$
$f_{T,5}/\Lambda^4$	$[-3.32, 3.40]$	$[-3.24, 3.32]$	$[-2.76, 2.84]$	$[-3.32, 3.40]$	$[-2.44, 2.52]$	$[-2.92, 2.92]$
$f_{T,6}/\Lambda^4$	$[-4.36, 4.36]$	$[-4.28, 4.36]$	$[-3.64, 3.72]$	$[-4.36, 4.44]$	$[-3.24, 3.24]$	$[-3.80, 3.88]$
$f_{T,7}/\Lambda^4$	$[-8.92, 8.84]$	$[-8.92, 8.76]$	$[-7.56, 7.48]$	$[-9.00, 8.92]$	$[-6.68, 6.60]$	$[-7.88, 7.72]$
$f_{T,8}/\Lambda^4$	$[-1.22, 1.26]$	$[-1.22, 1.22]$	$[-1.02, 1.06]$	$[-1.22, 1.26]$	$[-0.90, 0.94]$	$[-1.06, 1.10]$
$f_{T,9}/\Lambda^4$	$[-2.06, 2.06]$	$[-2.06, 2.06]$	$[-1.74, 1.74]$	$[-2.10, 2.10]$	$[-1.54, 1.54]$	$[-1.82, 1.82]$

Table 8.2: Expected and observed limits for the different anomalous couplings parameters in the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ channel.

by the CMS and ATLAS Collaborations at 8 TeV and 13 TeV [58, 105–112]. Limits extracted in the diboson channels VV (where $V = W$, Z or γ) currently represent the most stringent values on these parameters. This analysis, on the other side, by targeting the rare triboson production process $Z\gamma\gamma$, is not specifically designed to be competitive with the results obtained in the diboson channel. The latter is not statistically limited and thus provide a higher sensitivity to the aQGCs. However, the extraction of the aQGC limits in the $Z\gamma\gamma$ channel represents an independent method of testing aQGCs, beyond the one using the diboson production channel. The presented results represent the first extraction of limits on the aQGCs in the $pp \to Z\gamma\gamma \to l^+l^-\gamma\gamma$ channel at

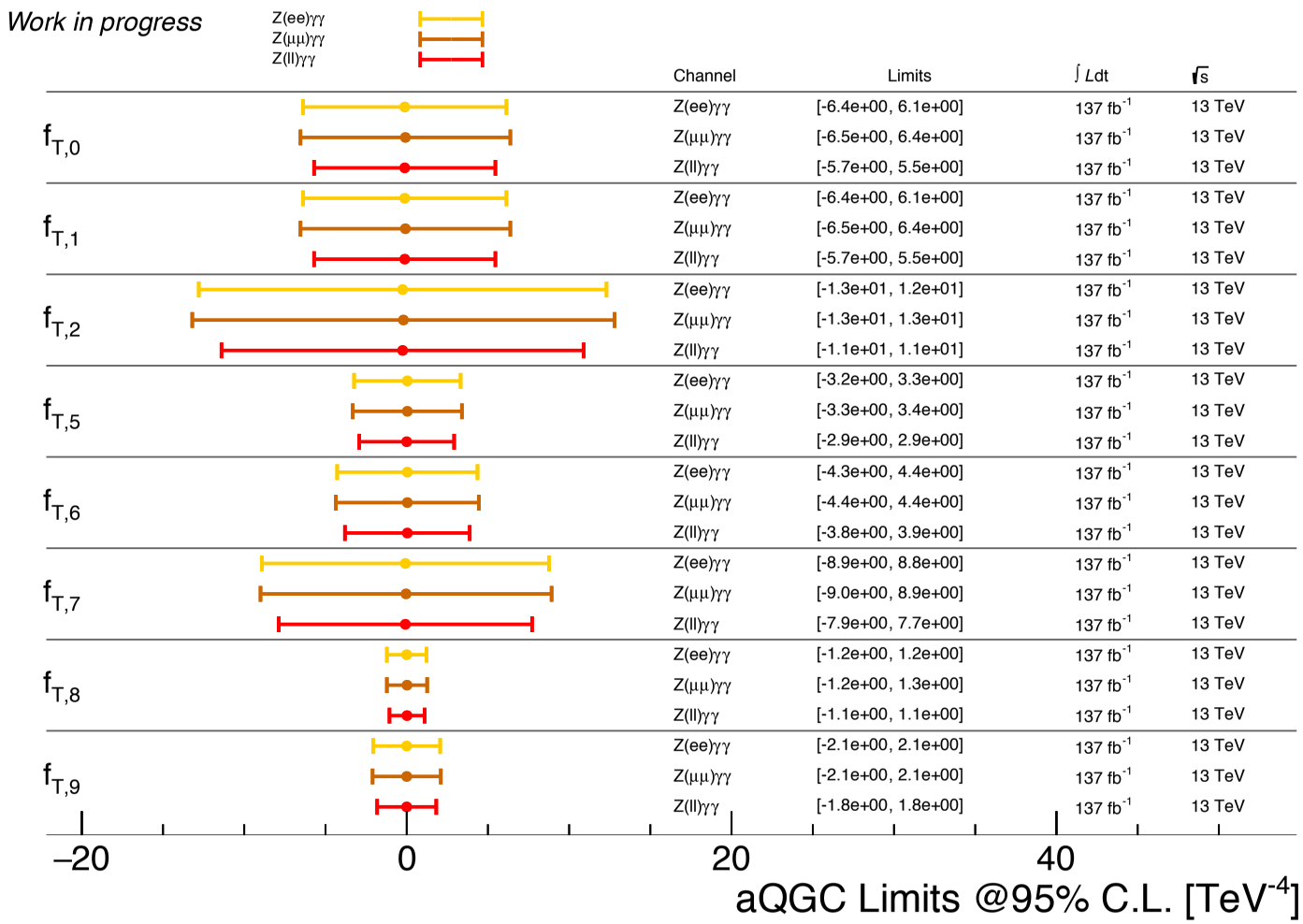


Figure 8.3: Observed limits on the aQGC coupling parameters at 95% confidence interval obtained in this work. In yellow for the electron channel, in orange for the muon channel and in red for the combination of the two.

13 TeV.

The limits obtained for the $f_{T,0}$, $f_{T,1}$ and $f_{T,2}$ coupling parameters are compared to other analyses in Figure 8.4 while the same is done for the $f_{T,5}$, $f_{T,7}$, $f_{T,8}$ and $f_{T,9}$ coupling parameters in Figure 8.5. Results obtained in this work are shown in red, while those obtained by the CMS (ATLAS) Collaboration in green (cyan). The limit placed by this analysis on the $f_{T,5}$ parameter can be directly compared with the one obtained in the same channel by the ATLAS Collaboration exploiting the LHC Run 1 data. The result presented in this work represents a clear improvement, being three times more precise, and clearly benefits from the increased centre of mass energy and statistic collected during the LHC Run 2. Limits obtained for the $f_{T,8}$ and $f_{T,9}$ coupling parameters are of particular interest because they can be extracted only by studying the production of EW neutral bosons. As can be seen in Figure 8.5, the obtained results are competitive with the ones obtained by the CMS and ATLAS Collaborations.

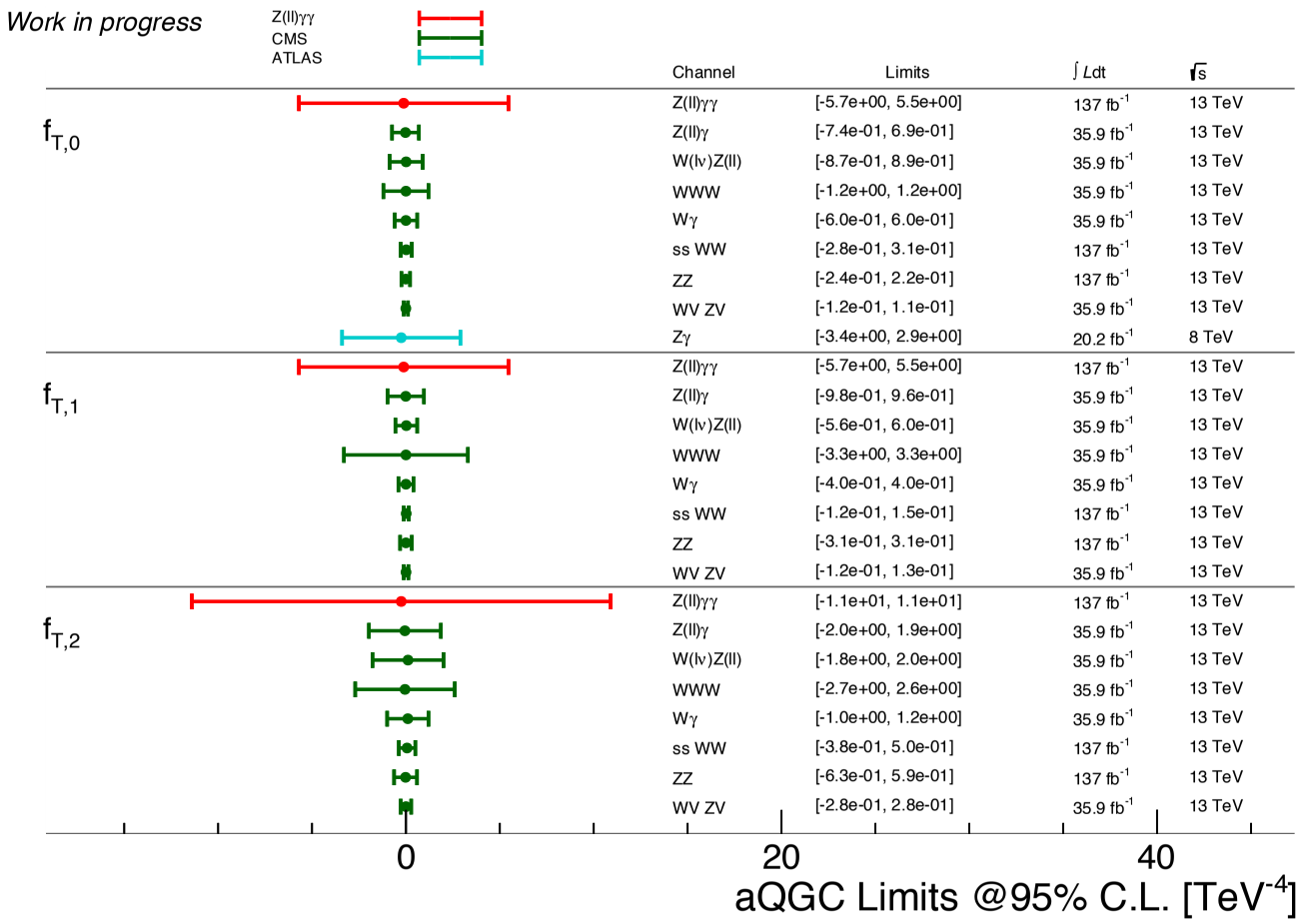


Figure 8.4: Observed limits at 95% confidence interval for the aQGC parameters: in red for the Z combined lepton channel of this analysis, in green (cyan) the results obtained by the CMS (ATLAS) Collaborations in other multi-boson analyses at 8 TeV and 13 TeV.

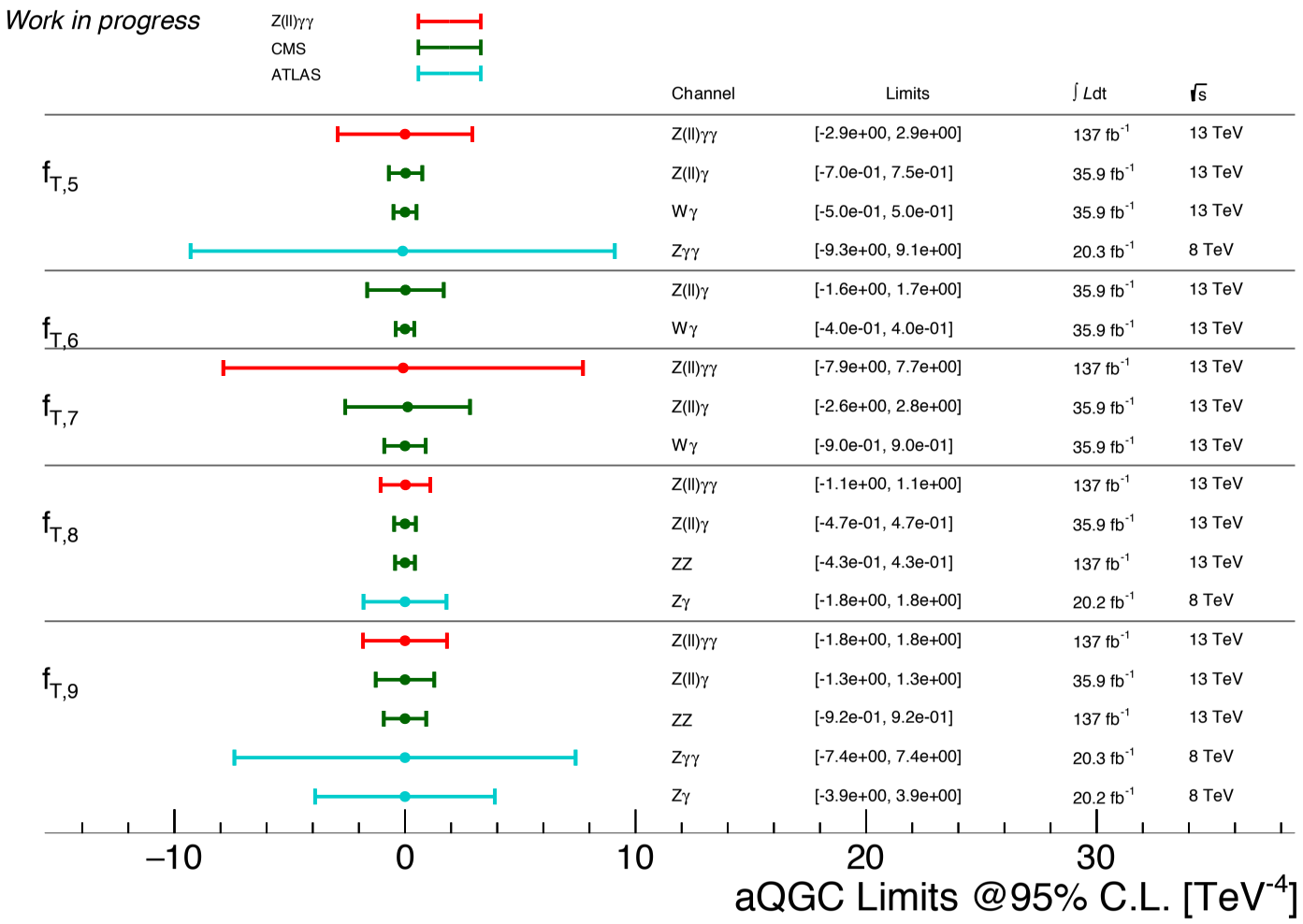


Figure 8.5: Observed limits at 95% confidence interval for the aQGC parameters: in red for the Z combined lepton channel of this analysis, in green (cyan) the results obtained by the CMS (ATLAS) Collaborations in other multi-boson analyses at 8 TeV and 13 TeV.

Conclusions

A measurement of the production of a Z boson in association with two or more photons has been presented. The analysis is based on proton proton collisions data collected by the CMS experiment, at a centre of mass energy of 13 TeV, and corresponding to an integrated luminosity of about $137\,\text{fb}^{-1}$. The $\text{l}^+\text{l}^-\gamma\gamma$ events are reconstructed in a fiducial phase space by selecting pairs of same flavour opposite sign leptons ($\text{l} = \text{e}, \mu$), originating from the Z boson decay, and at least two isolated photons.

Several Standard Model background processes share the same final state topology with the signal process targeted by this analysis. The background sources can be divided in two classes, "prompt" and "non-prompt", the former including events with genuine photons while the latter representing events where one or more reconstructed photons originate from hadrons decay inside jets. The non-prompt background source is the dominant one and is estimated with a data-driven technique, while the prompt background source is estimated by using simulated events and cross-checked using control samples.

The $\text{pp} \to \text{Z}\gamma\gamma \to \text{l}^+\text{l}^-\gamma\gamma$ cross section is measured with a maximum likelihood fit to the observed data yield. The signal strength of the process $\mu(\text{Z}(\text{l}^+\text{l}^-)\gamma\gamma)$ is treated as a free parameter in the fitting procedure. The measured fiducial cross section amounts to $\sigma(\text{Z}(\text{l}^+\text{l}^-)\gamma\gamma)^{meas.} = 5.44^{+0.58}_{-0.56}(\text{stat.})^{+0.64}_{-0.70}(\text{syst.})\,\text{fb}$, corresponding to a fitted signal strength of $\mu(\text{Z}(\text{l}^+\text{l}^-)\gamma\gamma) = 0.91^{+0.10}_{-0.10}(\text{stat.})^{+0.11}_{-0.12}(\text{syst.})$, in agreement with the Standard Model prediction. The measurement is slightly dominated by the systematic uncertainty, which mainly comes from the data-driven method and the prompt background source normalisation. The observed (expected) significance corresponds to 4.8 (5.8) standard deviation. This is the first measurement of the $\text{pp} \to \text{Z}\gamma\gamma \to \text{l}^+\text{l}^-\gamma\gamma$ cross section performed at a centre of mass energy of 13 TeV.

Events selected in the same phase space used for the fiducial cross section measurement are also exploited to set limits on several anomalous Quartic Gauge Couplings, in terms of dimension 8 effective field theory operators. The obtained constraints are compatible and complementary to other studies which use alternative production and decay channels. In particular, limits obtained for the $f_{T,8}/\Lambda^4$ and $f_{T,9}/\Lambda^4$ coupling parameters are competitive with the results obtained by the CMS Collaboration in the diboson production channel. Furthermore, limits obtained for the $f_{T,5}/\Lambda^4$ and $f_{T,9}/\Lambda^4$ coupling parameters can be directly compared to the ones obtained by the ATLAS Collaboration, in the same production channel, by exploiting collisions at a centre of mass energy of 8 TeV collected during the LHC Run 1. The results presented in this analysis show an improvement in precision of a factor of three with respect to the ATLAS measurements and benefit from the improved statistics delivered by the LHC Run 2. This work represents the first extraction of limits on the anomalous Quartic Gauge Couplings parameters, in the $\text{pp} \to \text{Z}\gamma\gamma \to \text{l}^+\text{l}^-\gamma\gamma$ channel, at a centre of mass energy of 13 TeV.

A | Inclusive phase space distributions

This appendix contains plots showing selected distributions obtained in the inclusive phase space of this analysis. In Figure A.1 the invariant mass for the events in which a Z boson is reconstructed in the electron or muon channel, without any requirements on the photons, is shown. The same distributions but for the entire Run 2 dataset are shown in Figure 6.1. The pseudorapidity distributions for the $Z \to l^+l^-$ channel are shown, for the separate years, in Figure A.2 and for the full Run2 in Figure A.3.

The last bin of each distribution includes the overflow. The agreement between data and predictions is shown in the ratio at the bottom of each plot. In the ratio, the grey hashed area represents the statistical uncertainty on the sum of signal and background while the uncertainty on the black dots is due to the statistical uncertainty of the data. The systematic uncertainties are not included in the plot.

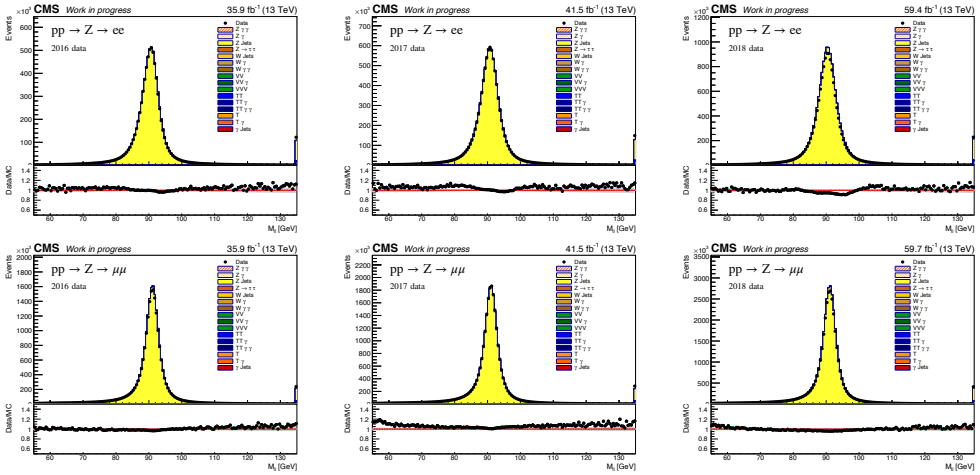


Figure A.1: Reconstructed invariant mass for the $Z \to l^+l^-$ events in the electron (top) and muon (bottom) channels for data collected in 2016 (left), 2017 (centre) and 2018 (right).

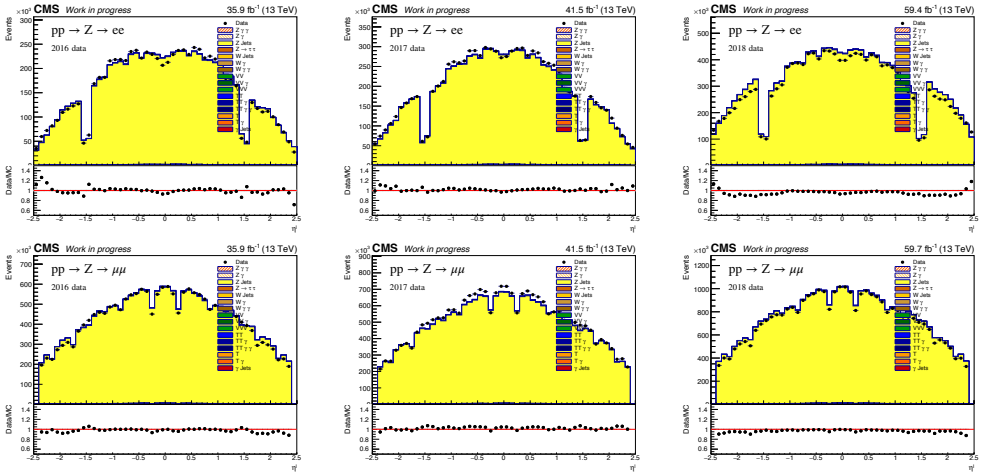


Figure A.2: Lepton pseudorapidity for the $Z \to l^+l^-$ events in the electron (top) and muon (bottom) channels for data collected in 2016 (left), 2017 (centre) and 2018 (right).

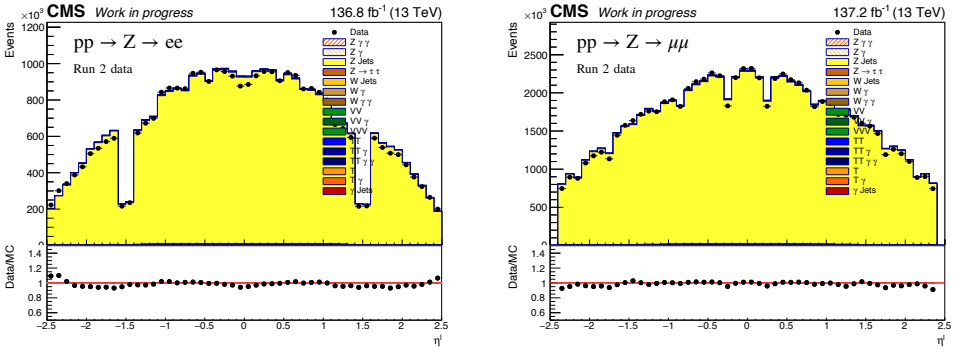


Figure A.3: Lepton pseudorapidity for the $Z \to l^+l^-$ in the electron (left) and muon (right) channels for data corresponding to all Run 2 dataset.

B | Single photon phase space distributions

This appendix contains plots showing selected distributions obtained in the single photon phase space of this analysis. In Figure B.1 the invariant mass for the events in which a Z boson is reconstructed in the electron or muon channel with at least one photon in the final state is shown. The same distribution, but for the entire Run 2 dataset, is shown in Figure B.3.

In Figure B.2 the transverse momentum of the leading photon for the events in which a Z boson is reconstructed in the electron or muon channel, with at least one photon in the final state, is shown. The same distribution, but for the entire Run 2 dataset, is presented in Figure B.4.

The last bin of each distribution includes the overflow. The agreement between data and predictions is shown in the ratio at the bottom of each plot. In the ratio, the grey hashed area represents the statistical uncertainty on the sum of signal and background while the uncertainty on the black dots is due to the statistical uncertainty of the data. The systematic uncertainties are not included in the plot.

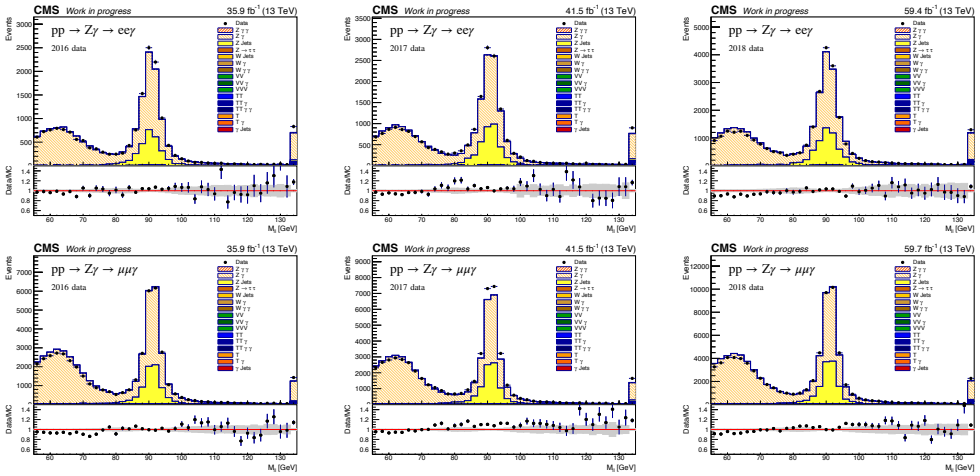


Figure B.1: Reconstructed invariant mass for the $Z \to l^{+}l^{-}\gamma$ events in the electron (top) and muon (bottom) channels for data collected in 2016 (left), 2017 (centre) and 2018 (right).

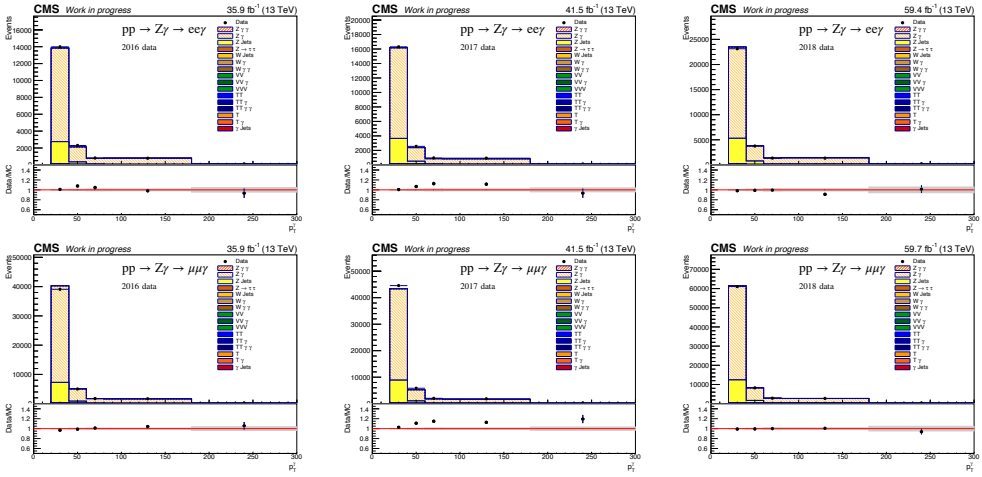


Figure B.2: Transverse momentum of the leading photon for the $Z \to l^+l^-\gamma$ events in the electron (top) and muon (bottom) channels for data collected in 2016 (left), 2017 (centre) and 2018 (right).

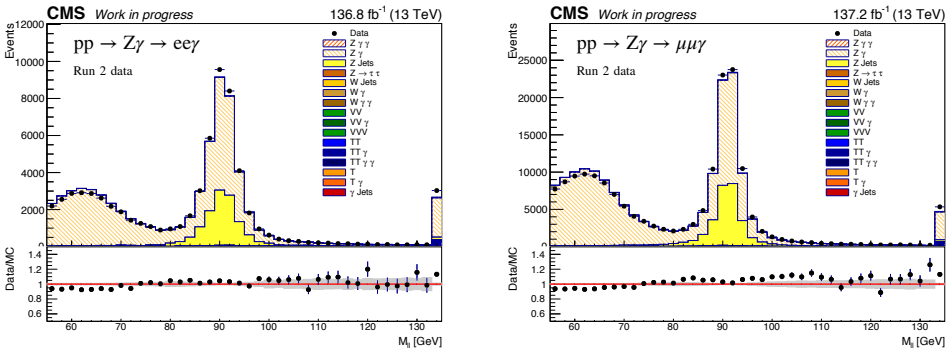


Figure B.3: Reconstructed invariant mass for the $Z \to l^+l^-\gamma$ events in the electron (left) and muon (right) channels for data corresponding to all Run 2 dataset.

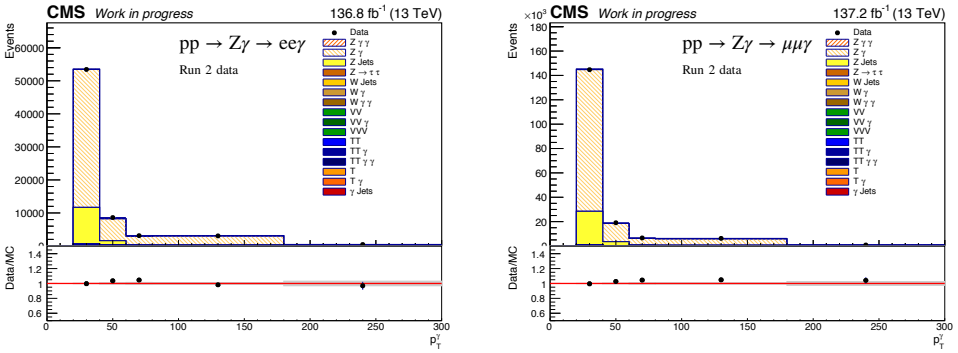


Figure B.4: Transverse momentum of the leading photon for the $Z \to l^+l^-\gamma$ events in the electron (left) and muon (right) channels for data corresponding to all Run 2 dataset.

C | Impact plots for the best-fit value of the signal strength

This appendix contains plots showing the impact of each systematic uncertainty on the best-fit value of the signal strength μ. The impact is defined as the shift $\Delta\mu$ that is induced as each best-fit value of the nuisance parameter θ is shifted to its $\pm 1\sigma$ post-fit uncertainty, with all other parameters profiled as normal (see Section 7.1.2).

The left panel in the impact plots shows the value of $(\theta - \theta_0)/\Delta_\theta$ where Δ_θ is the pre-fit uncertainty associated to the specific nuisance parameter and θ_0 the corresponding pre-fit value. The asymmetric error bars show the pre-fit uncertainty divided by the post-fit uncertainty. This means that parameters with error bars smaller than ± 1 are constrained in the fit. The direction of the $\pm 1\sigma$ impacts on the parameter of interest, shown in the right panel, indicates whether the nuisance parameter is correlated or anti-correlated with μ.

Table C.1 connects the naming convention used to label the systematic uncertainties in Table 7.2 to the one used the impact plots in Figure C.1, Figure C.2 and Figure C.3.

Text	Impact plot
Prompt normalisation	xsec_syst_*
Non-prompt normalisation	jet_misid_qcd
Photon scale factors	sf_pho_*
MC sample stat.	prop_binch_*
Luminosity	lumi
Electron scale factors	sf_ele_*
Pileup	pileup
Non-prompt shape	jet_misid_sherpa
L1 ECAL prefiring	l1prefiring
Muon scale factors	sf_muo_*

Table C.1: Table showing the connection between the naming convention for the systematic uncertainties as used in Section 7.1.2 to the one used in the impact plots in Figure C.1, Figure C.2 and Figure C.3.

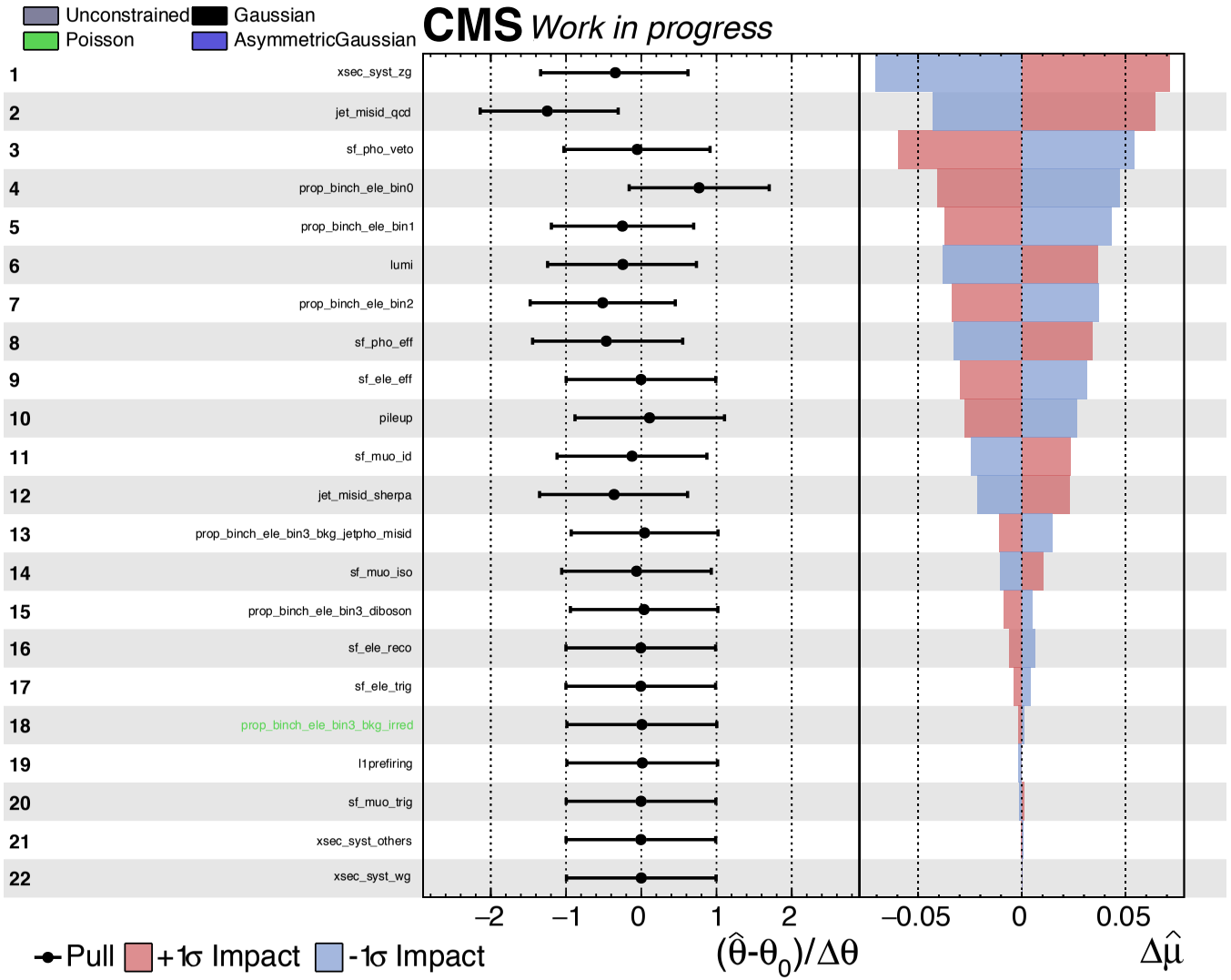


Figure C.1: Normalised pulls and constraints of the nuisance parameters related to the modelling uncertainties for the extraction of the $\sigma(\mathrm{Z}(\mathrm{e}^{+}\mathrm{e}^{-})\gamma\gamma)^{\mathrm{meas.}}$ cross section.

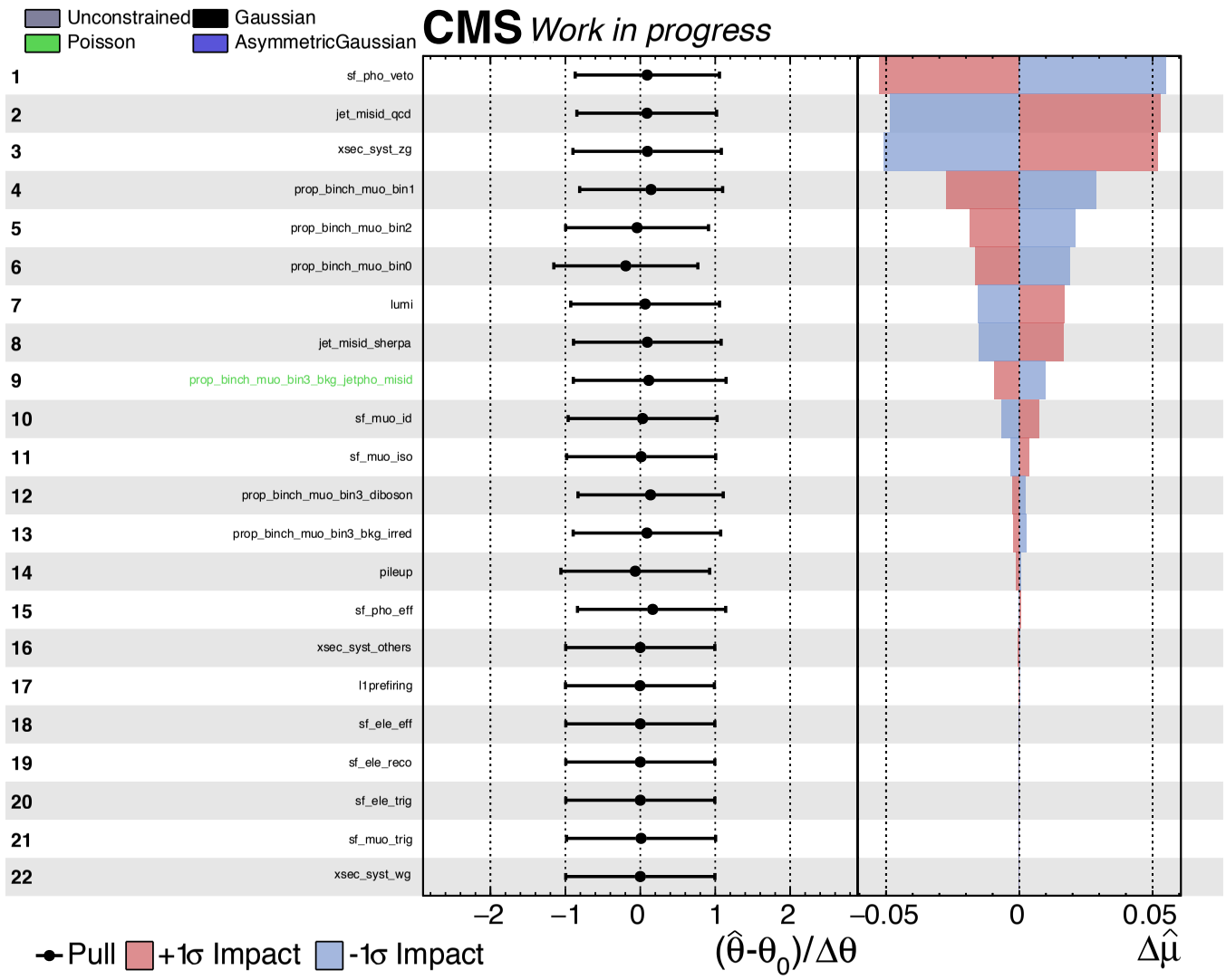


Figure C.2: Normalised pulls and constraints of the nuisance parameters related to the modelling uncertainties for the extraction of the $\sigma(Z(\mu^+\mu^-)\gamma\gamma)^{\text{meas.}}$ cross section.

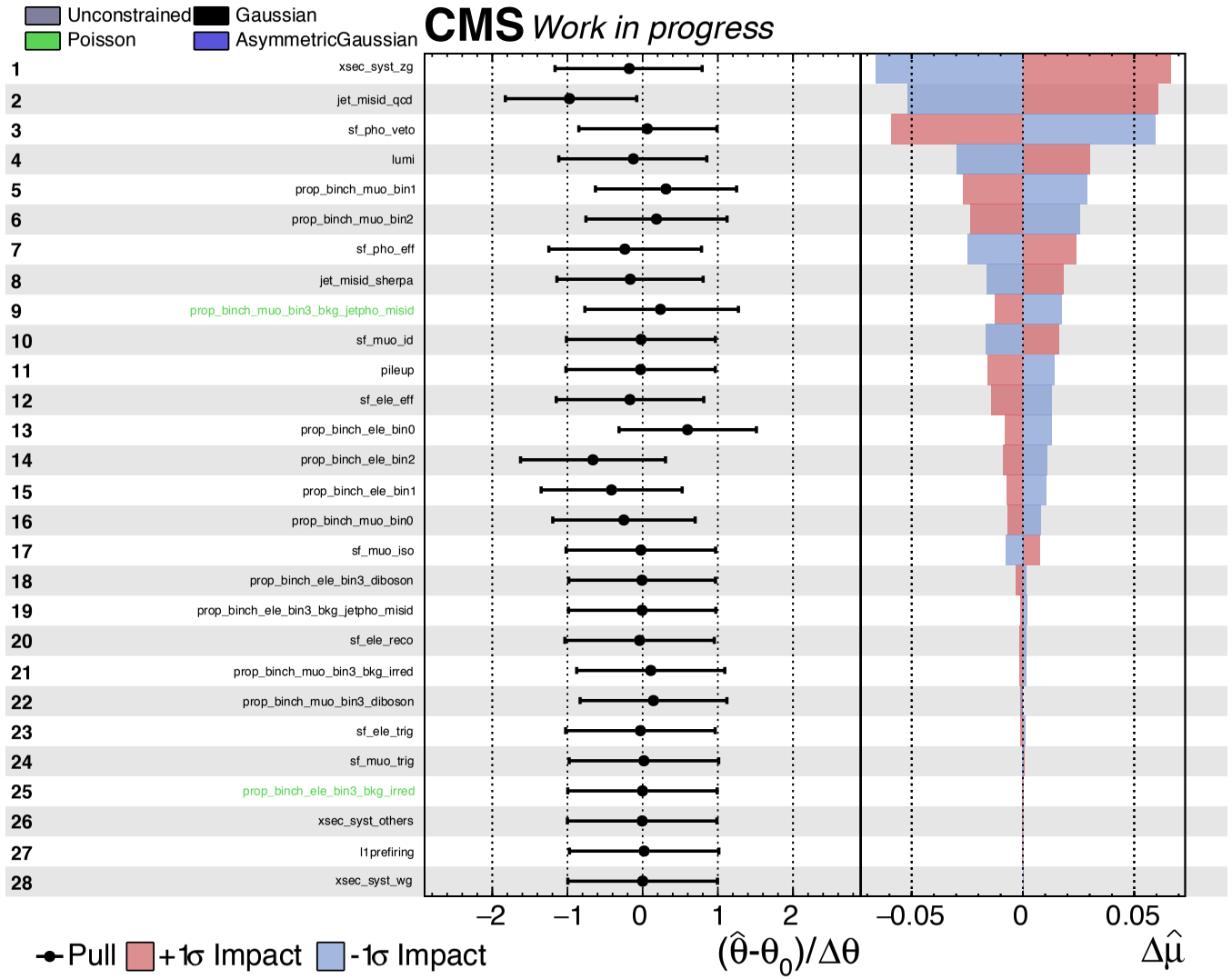


Figure C.3: Normalised pulls and constraints of the nuisance parameters related to the modelling uncertainties for the extraction of the $\sigma(\mathrm{Z}(\mathrm{l}^{+}\mathrm{l}^{-})\gamma\gamma)^{\mathrm{meas.}}$ cross section.

Bibliography

[1] G. Altarelli, "Collider Physics within the Standard Model: a Primer", Technical Report arXiv:1303.2842. RM3-TH-13-1. CERN-PH-TH-2013-020, CERN, 2013.

[2] V. Chiochia, G. Dissertori, and T. Gehrmann, "Phenomenology of particle physics". CreateSpace Independent Publ., [S.l.], 2014.

[3] M. E. Peskin, "Concepts of elementary particle physics". Oxford master series in particle physics, astrophysics, and cosmology. Oxford University Press, Oxford, 2019. doi:0198812191.

[4] S. Weinberg, "The Making of the Standard Model", *Eur. Phys. J. C* **34** (2004) 5–13, doi:10.1140/epjc/s2004-01761-1.

[5] G. 't Hooft, "The making of the Standard Model", *Nature* **448** (2007) 271–273, doi:10.1038/nature06074.

[6] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys. Rev. Lett.* **10** (1963) 531–533, doi:10.1103/PhysRevLett.10.531.

[7] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction", *Prog. Theor. Phys.* **49** (1973) 652–657, doi:10.1143/PTP.49.652.

[8] MissMJ and Nasfarley88, "Standard Model of Elementary Particles". http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[9] C.-N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance", *Phys. Rev.* **96** (1954) 191–195, doi:10.1103/PhysRev.96.191.

[10] S. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* **22** (1961) 579–588, doi:10.1016/0029-5582(61)90469-2.

[11] A. Salam and J. C. Ward, "Electromagnetic and weak interactions", *Phys. Lett.* **13** (1964) 168–171, doi:10.1016/0031-9163(64)90711-5.

[12] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.

[13] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508.

[14] CMS Collaboration, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021.

[15] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020.

[16] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the Standard Model", *Phys. Rept.* **457** (2008) 1–216, doi:10.1016/j.physrep.2007.10.004.

[17] Particle Data Group, "W boson mass". https://pdg.lbl.gov/2020/listings/rpp2020-list-w-boson.pdf.

[18] Particle Data Group, "Z boson mass". https://pdg.lbl.gov/2020/listings/rpp2020-list-z-boson.pdf.

[19] H. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys. Rev. Lett.* **30** (1973) 1346–1349, doi:10.1103/PhysRevLett.30.1346.

[20] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories", *Phys. Rev. Lett.* **30** (1973) 1343–1346, doi:10.1103/PhysRevLett.30.1343.

[21] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group Collaboration, "Precision electroweak measurements on the Z resonance", *Phys. Rept.* **427** (2006) 257–454, doi:10.1016/j.physrep.2005.12.006.

[22] ALEPH, DELPHI, L3, OPAL, LEP Electroweak Collaboration, "Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP", *Phys. Rept.* **532** (2013) 119–244, doi:10.1016/j.physrep.2013.07.004.

[23] Particle Data Group Collaboration, "Review of Particle Physics", *PTEP* **2020** (2020), no. 8, 083C01, doi:10.1093/ptep/ptaa104.

[24] C. Degrande et al., "Effective Field Theory: A Modern Approach to Anomalous Couplings", *Annals Phys.* **335** (2013) 21–32, doi:10.1016/j.aop.2013.04.016.

[25] A. Schälicke et al., "Event generator for particle production in high-energy collisions", *Prog. Part. Nucl. Phys.* **53** (2004) 329–338, doi:10.1016/j.ppnp.2004.02.031.

[26] G. P. Salam, "Towards Jetography", *Eur. Phys. J. C* **67** (2010) 637–686, doi:10.1140/epjc/s10052-010-1314-6.

[27] S. Catani et al., "New clustering algorithm for multi-jet cross-sections in $e^+ e^-$ annihilation", *Phys. Lett. B* **269** (1991) 432–438, doi:10.1016/0370-2693(91)90196-W.

[28] R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and Collider Physics". Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996. doi:10.1017/CBO9780511628788.

[29] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+ e^- Annihilation by Perturbation Theory in Quantum Chromodynamics", *Sov. Phys. JETP* **46** (1977) 641–653.

[30] V. Gribov and L. Lipatov, "Deep inelastic e p scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[31] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys. B* **126** (1977) 298–318, doi:10.1016/0550-3213(77)90384-4.

[32] A. Buckley et al., "LHAPDF6: parton density access in the LHC precision era", *Eur. Phys. J. C* **75** (2015) 132, doi:10.1140/epjc/s10052-015-3318-8.

[33] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040.

[34] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5.

[35] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079.

[36] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024.

[37] S. Hoeche et al., "Matching parton showers and matrix elements", in *HERA and the LHC: A Workshop on the Implications of HERA for LHC Physics: CERN - DESY Workshop 2004/2005 (Midterm Meeting, CERN, 11-13 October 2004; Final Meeting, DESY, 17-21 January 2005)*, pp. 288–289. 2005. doi:10.5170/CERN-2005-014.288.

[38] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **06** (2002) 029, doi:10.1088/1126-6708/2002/06/029.

[39] B. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference", *Nucl. Phys. B* **238** (1984) 492–528, doi:10.1016/0550-3213(84)90333-X.

[40] T. Sjöstrand, "Jet fragmentation of multiparton configurations in a string framework", *Nucl. Phys. B* **248** (1984) 469–502, doi:10.1016/0550-3213(84)90607-2.

[41] T. Gleisberg et al., "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007.

[42] D. Amati and G. Veneziano, "Preconfinement as a Property of Perturbative QCD", *Phys. Lett. B* **83** (1979) 87–92, doi:10.1016/0370-2693(79)90896-7.

[43] A. Buckley et al., "General-purpose event generators for LHC physics", *Phys. Rept.* **504** (2011) 145–233, doi:10.1016/j.physrep.2011.03.005.

[44] B. Andersson, G. Gustafson, and B. Söderberg, "A General Model for Jet Fragmentation", *Z. Phys. C* **20** (1983) 317, doi:10.1007/BF01407824.

[45] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x.

[46] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4.

[47] GEANT4 Collaboration, "GEANT4–a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[48] CMS Collaboration, T. J. Orimoto, "First CMS Results with LHC Beam", in *24th Lake Louise Winter Institute: Fundamental Interactions*. 5, 2009. arXiv:0905.4814.

[49] G. Belanger and F. Boudjema, "Probing quartic couplings of weak bosons through three vector production at a 500 GeV NLC", *Phys. Lett. B* **288** (1992) 201–209, doi:10.1016/0370-2693(92)91978-I.

[50] D. R. Green, P. Meade, and M.-A. Pleier, "Multiboson interactions at the LHC", *Rev. Mod. Phys.* **89** (2017), no. 3, 035008, doi:10.1103/RevModPhys.89.035008.

[51] F. Campanario et al., "NLO corrections to processes with electroweak bosons at hadron colliders", *Nucl. Part. Phys. Proc.* **261-262** (2015) 268–307, doi:10.1016/j.nuclphysbps.2015.03.019.

[52] G. Bozzi, F. Campanario, M. Rauch, and D. Zeppenfeld, "Zγγ production with leptonic decays and triple photon production at next-to-leading order QCD", *Phys. Rev. D* **84** (2011) 074028, doi:10.1103/PhysRevD.84.074028.

[53] M. Baak et al., "Working Group Report: Precision Study of Electroweak Interactions", in *Community Summer Study 2013: Snowmass on the Mississippi*. 2013. arXiv:1310.6708.

[54] O. Eboli, M. Gonzalez-Garcia, and J. Mizukoshi, "$pp \to jje^{\pm}\mu^{\pm}\nu\nu$ and $jje^{\pm}\mu^{\mp}\nu\nu$ at $\mathcal{O}(\alpha_{em}^6)$ and $\mathcal{O}(\alpha_{em}^4\alpha_s^2)$ for the study of the quartic electroweak gauge boson vertex at CERN LHC", *Phys. Rev. D* **74** (2006) 073005, doi:10.1103/PhysRevD.74.073005.

[55] L3 Collaboration, "The $e^+e^- \to Z\gamma\gamma \to q\bar{q}\gamma\gamma$ reaction at LEP and constraints on anomalous quartic gauge boson couplings", *Phys. Lett. B* **540** (2002) 43–51, doi:10.1016/S0370-2693(02)02127-5.

[56] OPAL Collaboration, "Constraints on anomalous quartic gauge boson couplings from $\nu\bar{\nu}\gamma\gamma$ and $q\bar{q}\gamma\gamma$ events at CERN LEP2", *Phys. Rev. D* **70** (2004) 032005, doi:10.1103/PhysRevD.70.032005.

[57] CMS Collaboration, "Measurements of the $pp \to W\gamma\gamma$ and $pp \to Z\gamma\gamma$ cross sections and limits on anomalous quartic gauge couplings at $\sqrt{s} = 8\,\text{TeV}$", *JHEP* **10** (2017) 072, doi:10.1007/JHEP10(2017)072.

[58] ATLAS Collaboration, "Measurements of $Z\gamma$ and $Z\gamma\gamma$ production in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector", *Phys. Rev. D* **93** (2016), no. 11, 112002, doi:10.1103/PhysRevD.93.112002.

[59] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.

[60] L. I. S. Group, "LEP design report". CERN, Geneva, 1983. By the LEP Injector Study Group.

[61] J. Blewett, "200-GeV Intersecting Storage Accelerators", *eConf* **C710920** (1971) 501.

[62] CMS Collaboration, "The CMS Experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[63] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[64] LHCb Collaboration, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[65] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[66] A. Blas et al., "The PS complex as proton pre-injector for the LHC: Design and implementation report". Benedikt, M., 2000. doi:10.5170/CERN-2000-003.

[67] E. Mobs, "The CERN accelerator complex - 2019", *CERN-GRAPHICS-2019-002* (2019). General Photo.

[68] R. R. Wilson, "The Tevatron", *Phys. Today* **30N10** (1977) 23–30, doi:10.1063/1.3037746.

[69] ALEPH Collaboration, "ALEPH: A detector for electron-positron annnihilations at LEP", *Nucl. Instrum. Meth. A* **294** (1990) 121–178, doi:10.1016/0168-9002(90)91831-U. [Erratum: Nucl.Instrum.Meth.A 303, 393 (1991)].

[70] CMS Collaboration, "Performance of the CMS Drift-Tube Local Trigger with Cosmic Rays", *JINST* **5** (2010) T03003, doi:10.1088/1748-0221/5/03/T03003.

[71] T. Sakuma and T. McCauley, "Detector and Event Visualization with SketchUp at the CMS Experiment", *J. Phys. Conf. Ser.* **513** (2014) 022032, doi:10.1088/1742-6596/513/2/022032.

[72] J. G. Layter, "The CMS muon project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

[73] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[74] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13\,\text{TeV}$", Technical Report CMS-PAS-LUM-17-004, CERN, Geneva, 2018.

[75] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13\,\text{TeV}$", Technical Report CMS-PAS-LUM-18-002, CERN, Geneva, 2019.

[76] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017), no. 10, P10003, doi:10.1088/1748-0221/12/10/P10003.

[77] W. Adam, B. Mangano, T. Speer, and T. Todorov, "Track Reconstruction in the CMS tracker", Technical Report CMS-NOTE-2006-041, CERN, Geneva, 2006.

[78] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC", Technical Report CMS-NOTE-2005-001, CERN, Geneva, 2005.

[79] H. Bethe and W. Heitler, "On the Stopping of fast particles and on the creation of positive electrons", *Proc. Roy. Soc. Lond. A* **146** (1934) 83–112, doi:10.1098/rspa.1934.0140.

[80] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009.

[81] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems", *IEEE Proc.* **86** (1998), no. 11, 2210–2239, doi:10.1109/5.726788.

[82] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", arXiv:2012.06888.

[83] CMS Collaboration, "Electron and Photon performance in CMS with the full 2017 data sample and additional 2016 highlights for the CALOR 2018 Conference", Technical Report CMS-DP-2018-017, CERN, 2018.

[84] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119–126, doi:10.1016/j.physletb.2007.09.077.

[85] CMS Collaboration, "Muon reconstruction performance during Run II", Technical Report CMS-DP-2019-022, CERN, 2019.

[86] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **13** (2018), no. 06, P06015, doi:10.1088/1748-0221/13/06/P06015.

[87] A. Bodek et al., "Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments", *Eur. Phys. J. C* **72** (2012) 2194, doi:10.1140/epjc/s10052-012-2194-8.

[88] CMS Collaboration, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015), no. 08, P08010, doi:10.1088/1748-0221/10/08/P08010.

[89] J. M. Campbell, R. Ellis, F. Maltoni, and S. Willenbrock, "Associated production of a Z Boson and a single heavy quark jet", *Phys. Rev. D* **69** (2004) 074021, doi:10.1103/PhysRevD.69.074021.

[90] M. Grazzini, S. Kallweit, and M. Wiesemann, "Fully differential NNLO computations with MATRIX", *Eur. Phys. J. C* **78** (2018), no. 7, 537, doi:10.1140/epjc/s10052-018-5771-7.

[91] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015.

[92] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Comput. Phys. Commun.* **182** (2011) 2388–2403, doi:10.1016/j.cpc.2011.06.008.

[93] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161.

[94] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7\,\text{TeV}$", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080.

[95] ATLAS Collaboration, "Measurement of the isolated di-photon cross-section in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS detector", *Phys. Rev. D* **85** (2012) 012003, doi:10.1103/PhysRevD.85.012003.

[96] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993) 219–228, doi:10.1016/0010-4655(93)90005-W.

[97] J. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", in *PHYSTAT 2011*, pp. 115–120. 2011. doi:10.5170/CERN-2011-006.115.

[98] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling", *eConf* **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[99] ATLAS, CMS Collaboration, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045.

[100] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0. [Erratum: Eur.Phys.J.C 73, 2501 (2013)].

[101] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001.

[102] A. Buckley et al., "Rivet user manual", *Comput. Phys. Commun.* **184** (2013) 2803–2819, doi:10.1016/j.cpc.2013.05.021.

[103] G. Sorrentino, "Characterization of Zγγ and Wγγ events in pp collisions at the LHC and limits on anomalous couplings", 2019. https://cds.cern.ch/record/2696065.

[104] "Anomalous gauge coupling models". http://feynrules.irmp.ucl.ac.be/wiki/AnomalousGaugeCoupling.

[105] CMS Collaboration, "Measurement of the cross section for electroweak production of a Z boson, a photon and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on anomalous quartic couplings", *JHEP* **06** (2020) 076, doi:10.1007/JHEP06(2020)076.

[106] CMS Collaboration, "Measurement of electroweak WZ boson production and search for new physics in WZ + two jets events in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **795** (2019) 281–307, doi:10.1016/j.physletb.2019.05.042.

[107] CMS Collaboration, "Search for the production of $W^{\pm}W^{\pm}W^{\mp}$ events at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **100** (2019) 012004, doi:10.1103/PhysRevD.100.012004.

[108] CMS Collaboration, "Observation of electroweak production of Wγ with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:2008.10521.

[109] CMS Collaboration, "Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **809** (2020) 135710, doi:10.1016/j.physletb.2020.135710.

[110] CMS Collaboration, "Evidence for electroweak production of four charged leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:2008.07013.

[111] CMS Collaboration, "Search for anomalous electroweak production of vector boson pairs in association with two jets in proton-proton collisions at 13 TeV", *Phys. Lett. B* **798** (2019) 134985, doi:10.1016/j.physletb.2019.134985.

[112] ATLAS Collaboration, "Studies of Zγ production in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *JHEP* **07** (2017) 107, doi:10.1007/JHEP07(2017)107.

Non devi mai guardare la cima ma il metro di terreno che ti sta davanti al naso. Superato quello ne arriva un altro, e un altro ancora. Avanti così, senza alzare gli occhi, finchè ti ritrovi a calpestare aria. Lì è la cima.

Celio