Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Corso di Laurea Magistrale in Fisica

Study of the performances of the $H \to ZZ^* \to 4l$ analysis with the CMS Phase-II Upgrade

Relatore:
Prof.ssa Sylvie Braibant

Presentata da:
Lisa Borgonovi

Correlatori:
Dott. Luigi Guiducci
Dott.ssa Francesca R. Cavallo

Sessione III
Anno Accademico 2013/2014

"Qualunque cosa tu possa fare, qualunque sogno tu possa sognare, comincia.
L'audacia reca in sé genialità, magia e forza.
Comincia ora."

J. W. Göethe

Abstract

Le prestazioni dell'analisi $H \rightarrow ZZ^* \rightarrow 4l$ sono studiate nel contesto di High Luminosity LHC, con il rivelatore CMS. L'alta luminosità (fino a $\mathcal{L} = 5 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$) dell'acceleratore crea ardue condizioni sperimentali. In particolare, il numero di eventi che si producono ad ogni incrocio dei fasci aumenterà fino a 140. Per far fronte a questa difficile condizione, il rivelatore CMS subirà una serie di migliorie, che verranno implementate in due fasi successive: Fase-1 e Fase-2. Sono stati utilizzati i campioni simulati con il software di simulazione di CMS (Full Simulation), e altri appositamente prodotti con una simulazione veloce e parametrizzata (Delphes). Inizialmente è stata realizzata una completa validazione di Delphes rispetto alla Full Simulation, usando campioni di riferimento della Fase-1. Delphes è stato quindi utilizzato per simulare la risposta del rivelatore di Fase-2. Tale configurazione di Fase-2 è stata infine confrontata con il rivelatore di Fase-1 e con uno stesso rivelatore di Fase-1, degradato da fenomeni di invecchiamento, entrambi modellati utilizzando la Full Simulation. Dal confronto di questi 3 scenari si possono trarre le seguenti conclusioni: il peggioramento delle prestazioni osservate nello scenario affetto da invecchiamento mostrano che un sostanziale miglioramento del rivelatore è necessario. La specifica configurazione di Fase-2 studiata, permette di mantenere le stesse prestazioni della Fase-1, e nel caso del canale in 4 muoni, perfino di migliorarle.

Abstract

The performances of the $H \to ZZ^* \to 4l$ analysis are studied in the context of the High Luminosity upgrade of the LHC collider, with the CMS detector. The high luminosity (up to $\mathcal{L} = 5 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$) of the accelerator poses very challenging experimental conditions. In particular, the number of overlapping events per bunch crossing will increase to 140. To cope with this difficult environment, the CMS detector will be upgraded in two stages: Phase-I and Phase-II. The tools used in the analysis are the CMS Full Simulation and the fast parametrized Delphes simulation. A validation of Delphes with respect to the Full Simulation is performed, using reference Phase-I detector samples. Delphes is then used to simulate the Phase-II detector response. The Phase-II configuration is compared with the Phase-I detector and the same Phase-I detector affected by aging processes, both modeled with the Full Simulation framework. Conclusions on these three scenarios are derived: the degradation in performances observed with the "aged" scenario shows that a major upgrade of the detector is mandatory. The specific upgrade configuration studied allows to keep the same performances as in Phase-I and, in the case of the four-muons channel, even to exceed them.

Contents

Introduction

In 2012 the Higgs boson discovery was announced by the two main experiments at the CERN Large Hadron Collider (LHC): CMS and ATLAS. This result was the conclusion of several decades of extensive searches aimed at verifying the last missing piece of the Standard Model of particles and interactions (SM). In fact, in 1964 F. Englert, R.Brout and P. Higgs proposed a mechanism that explained the masses of the vector bosons $W^{\pm}$ and Z^0 through the Spontaneous Symmetry Breaking (SSB) of the electroweak theory. The so-called Higgs mechanism predicted the existence of a new scalar boson with unknown mass, whose couplings to fermions and bosons are the origins of their mass.

Most of the future experimental program in particle physics will focus on measurements of the Higgs boson properties: for this purpose, the LHC is the only facility available in the world. In order to collect the high statistics needed for high precision measurements, an upgrade program has been proposed for the accelerator, increasing both energy and luminosity in the next twenty years. After the current "long shutdown" period (LS1), LHC will resume operations in spring 2015 with a center-of-mass energy increased from 8 TeV (2012) to 13 TeV. The LHC will be operated with a cycle of three years of data taking and then a long shutdown to perform ordinary maintenance of the accelerator complex. During the following shutdown (LS2), the CMS experiment will also undergo a substantial upgrade, referred to as Phase-I.

In the following long shutdown (LS3) scheduled in 2023, LHC will be substantially upgraded to reach luminosities as high as $\mathcal{L} = 5 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$. This will allow, in a ten year operating period to collect an integrated luminosity of 3000 fb^{-1}. The high instantaneous luminosity will come at the price of an extremely high number of overlapping events (pile-up), up to 140 for a bunch-crossing interval of 25 ns. To cope with these new conditions, the CMS detector will also need to be upgraded: this second stage is called Phase-II.

This thesis aims at verifying the performances of the analysis used for the Higgs boson search in the $H \to ZZ^* \to 4l$ channel, with the introduction of the CMS Phase-II detector upgrade. The effect of the Phase-II detector upgrade on the signal selection efficiency and on the level of accepted background were studied and compared to the corresponding results obtained in the Phase-I scenario, and in an aging scenario where

the same Phase-I detector was degraded by ten years of data taking and irradiation.

Monte Carlo simulations were used to model the three detector scenarios described above . In particular, while Full Simulation samples were available for the Phase-I and Phase-I "aged" detectors, the parametrized Delphes Simulation was used for the production of the Phase-II detector samples. This required tuning the Delphes input parameters, through a validation process that involved comparisons between results obtained with Delphes and Full Simulation for the Phase-I scenario.

In Chapter 1, an introduction to the Standard Model of particles and forces is presented, focusing on the electroweak symmetry breaking and the Higgs mechanism. The Higgs boson search and discovery at LHC are also presented.

In Chapter 2, an overview of the actual and the future upgraded configurations of the CMS detector is given. The reconstruction and identification procedures are also reported.

In Chapter 3, the two simulation tools, Full Simulation and Delphes, used to describe the future detector geometries are described in details.

Chapter 4 describes the $H \to ZZ^* \to 4l$ analysis performed in three different detector configurations: Phase-I, Phase-I "aged" and Phase-II. The results obtained are presented and conclusions regarding the performances of the three configurations are derived.

Chapter 1

Standard Model and Higgs mechanism

1.1 The Standard Model of elementary particles and interactions

The Standard Model (SM) of particles and forces [1–3] is a quantum field theory which successfully describes the electromagnetic, weak and strong interactions, explaining almost every experimental data from high energy physics experiments. The SM relies on two fundamental concepts: elementary particle and symmetry.

An elementary particle is the elementary point-like constituent of matter, with unknown structure up to the present experimental limits ($\lambda \simeq \mathcal{O}(10^{-19}$ cm)). The SM considers two types of elementary particles: matter and interaction exchange particles.
All matter particles are fermions, which obey the Fermi-Dirac statistics and have half-integer spin (s=1/2). They are divided in leptons, electron (e), muon (μ), tau (τ) and the associated neutrinos (ν_e, ν_μ, ν_τ respectively), and quarks, up (u), down (d), strange (s), charmed (s), bottom (b), top (t). Charged leptons are sensitive to the electromagnetic and weak interactions, neutrinos only to the weak ones. Quarks have an additional quantum number with respect to leptons, that is the color charge, which subjects them to the strong interaction. Both leptons and quarks are grouped in three families, as shown in Fig. 1.1. For every fermion, a corresponding anti-particle exists, with the same quantum numbers but opposite charge.
The interaction particles are bosons, which obey the Bose-Einstein statistics and have integer spin (s=1). The intermediate bosons are the carriers of the interactions. Electromagnetic interactions are mediated by photons (γ), weak interactions are mediated by vector bosons Z^0 and $W^\pm$, and strong interactions are mediated by eight gluons g. A summary of the SM bosons is shown in Fig. 1.2. The gravitational force is supposed to be mediated by the graviton (G), but it is not described in the SM because at the scale of high energy particle physics the gravitational force can be neglected.

Fermions	1^{st} family	2^{nd} family	3^{rd} family	Charge (e)	Spin	Interactions
Leptons	e	μ	τ	-1	1/2	EM + weak
	ν_e	ν_μ	ν_τ	0	1/2	weak
Quarks	u	c	t	$+2/3$	1/2	All
	d	s	b	$-1/3$	1/2	

Figure 1.1: SM Fermions.

Gauge Boson	Mass (GeV)	Coupling Constant	Range (cm)
γ	0	$\alpha_{EM}(Q \simeq O(m_e)) \simeq 1/137.036$	∞
Z^0	91.19	$G_F \simeq 1.16637 \cdot 10^{-5}$ GeV^{-2}	$\sim 10^{-16}$
$W^\pm$	80.39		
g_i $(i = 1, 2, ..., 8)$	0	$\alpha_S(Q \simeq O(m_Z)) \simeq 0.1184$	$\sim 10^{-13}$

Figure 1.2: SM Bosons.

The SM is a quantum field theory based on a gauge symmetry: it is a gauge theory in which the Lagrangian is invariant under a continuous group of local transformations. The SM is based on the symmetry of gauge unitary groups:

$$\mathrm{SU(3)_C} \otimes [\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}]\,. \tag{1.1}$$

This group includes:

- the symmetry group for strong interactions, $\mathrm{SU(3)_C}$, where C stands for color (red, blue, green) which is the generator of the group;
- the symmetry group for electroweak interactions, $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$, which unifies the electromagnetic and weak interactions, as theorized by Glashow, Weinberg and Salam [4–6]. The generator of the electroweak interaction is the third component of the weak isospin, I_3, which applies to left-handed (L) fermions. In this unified theory, the generator of the electromagnetic interaction is not simply the electric charge Q but the hypercharge Y defined by the Gell-Mann-Nishijima relation:

$$Q = I_3 + \frac{1}{2}Y. \tag{1.2}$$

1.2 The electroweak theory

The electroweak theory based on the symmetry group $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$, which unified the electromagnetic and weak interactions, was first proposed by S. L. Glashow in

1961 [4]. The four associated gauge bosons are the photon, $W^{\pm}$ and Z^0. Despite the unified description, an important asymmetry was present: the gauge bosons $W^{\pm}$ and Z^0 are massive, while the photon is massless. In the same year, the Goldstone's theorem proposed by Goldstone [7, 8] stated that The Spontaneous Symmetry Breaking (SSB) of any continuous global symmetry[1] implies the existence of a massless spinless boson called Nambu-Goldstone boson. Starting from this statement, in 1964 the SSB of a local gauge symmetry, the electroweak symmetry $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$, was independently proposed by F. Englert and R. Brout [11], by P. Higgs [12, 13], by G. Guralnik, C. R. Hagen and T. Kibble [14, 15]. This mechanism, known as the Higgs mechanism, allows the mediator of the weak interaction ($W^{\pm}$ and Z^0) to become massive while the mediator of the electromagnetic interaction (γ) remains massless. In 1967-1968 S. Weinberg and A. Salam formulated the electroweak theory including the Higgs mechanism [5, 6]. This theory was proved to be renormalizable in 1971 by G. 't Hooft.

In the following years this theory was tested in many experiments: in 1973 neutral current weak interactions were firstly observed in the bubble chamber Gargamelle at CERN [16], and a first estimation of the weak interaction gauge boson masses m_Z and m_W was provided; in 1983 the bosons $W^{\pm}$ and Z_0 were then observed at the Sp$\bar{\text{p}}$S collider at CERN [17]. At LEP (1989-2000) [18], more precise measurements confirmed the SM, but still missed the discovery of the fundamental ingredient of the electroweak theory, the Higgs boson. In 2012, the Higgs boson was finally discovered at the LHC (CERN) [19] proving the accuracy of the theory, and leading to the Nobel prize won by Englert and Higgs in 2013 [20].

1.2.1 Gauge Invariance in the Electroweak Theory

The Dirac Lagrangian density $\mathcal{L}$ for a free fermion is:

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi = \bar{\psi}(x)(i\gamma^{\mu}\partial_{\mu} - m)\psi(x) \tag{1.3}$$

where the first term represents the kinetic energy of the matter field ψ with mass m, and the second term, which is bilinear in ψ, is proportional to the mass m of the field ψ.
This Lagrangian is already invariant under U(1) rotations, that is:

$$\psi \to \psi' = U\psi \qquad \Rightarrow \qquad \mathcal{L}(\psi') = \mathcal{L}(\psi) \tag{1.4}$$

where U is a unitary matrix ($U^{\dagger}U = 1$).
To require local gauge invariance, instead, the derivative ∂_{μ} has to be replaced with the covariant derivative of the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry group:

$$\mathcal{D}_{\mu} = \partial_{\mu} + \frac{ig'}{2}B_{\mu}\mathbb{1} + \frac{ig}{2}W^a_{\mu}\sigma_a \tag{1.5}$$

where:

[1] For an introduction to gauge symmetries, see [9, 10]

- σ_a are the non-commutative Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \tag{1.6}$$

- B_μ is the massless mediator of the field of the $\mathrm{U(1)_Y}$ group with coupling constant g' and transforms as:

$$B_\mu \to B'_\mu = B_\mu - \frac{1}{g'}\partial^\mu \alpha(x); \tag{1.7}$$

- W^a_μ are the three massless quanta of $\mathrm{SU(2)}_L$ with coupling constant g and transform as:

$$W^a_\mu \to W'^a_\mu = W^a_\mu - \frac{1}{g}\partial_\mu \varepsilon^a(x) - \varepsilon_{abc}\varepsilon^b(x)W^c_\mu \tag{1.8}$$

  where ε^a are infinitesimal arbitrary parameters and ε_{abc} is the total antisymmetric tensor.

The Lagrangian 1.3 thus becomes:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\bar{\psi}\psi \tag{1.9}$$

that transforms as:

$$\mathcal{L}'(\psi') = i\bar{\psi}U^\dagger \gamma^\mu U \mathcal{D}_\mu \psi - m\bar{\psi}U^\dagger U\psi = i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\bar{\psi}\psi = \mathcal{L}(\psi) \tag{1.10}$$

This Lagrangian density describes the interaction between matter particles through the exchange of gauge bosons associated with the gauge fields, with the coupling constants g and g'. To include the propagation of gauge fields, the Lagrangian has to be completed with the Yang-Mills Lagrangian density:

$$\mathcal{L}_{YM} = -\frac{1}{4}W^a_{\mu\nu}(W^a)^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.11}$$

where $W^a_{\mu\nu}$ and $B_{\mu\nu}$ are defined as:

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\varepsilon_{abc}W^b_\mu W^c_\nu \tag{1.12}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.13}$$

In the Lagrangian density, there are no quadratic terms for the gauge fields ($m^2 B_\mu B^\mu$ or $m^2 W^a_\mu (W_a)^\mu$). For this reason, the gauge bosons associated with these gauge fields are massless. The real fields γ, $W^\pm$ and Z^0 are obtained from these massless fields after spontaneous symmetry breaking.

1.3 Spontaneous Symmetry Breaking and the Higgs mechanism

The mechanism that allows the mediators of the weak interaction to become massive and the mediator of the electromagnetic interaction to remain massless, is the so-called Higgs mechanism, that is the Spontaneous Symmetry Breaking which keeps the Lagrangian invariant under the group of gauge transformations $SU(2)_L \times U(1)_Y$. The Higgs mechanism is the local gauge-invariant extension of the Goldstone model described in Section 1.2: the would-be Goldstone bosons associated to the global symmetry breaking combine with the massless gauge bosons to produce massive bosons. The number of bosons acquiring mass is equal to the number of would-be Goldstone bosons. This mechanism requires the introduction of a new scalar boson, the Higgs boson, with the corresponding Higgs field. The self-interaction of the Higgs boson modifies the ground state so that it is no longer a hypercharge or a weak eigenstate. The interaction of this boson scalar field, which is supposed to be everywhere in the space-time, with the boson and fermion fields provides mass to the elementary particles.
In order to introduce the Higgs mechanism in the electroweak theorem an ad hoc system is adopted; the gauge invariant Lagrangian $\mathcal{L}$, corresponding to a self-interacting scalar isodoublet φ is:

$$\mathcal{L}_{\mathcal{H}} = \mathcal{L}_{\mathcal{D}} - \mathcal{L}_V + \mathcal{L}_{YM} \tag{1.14}$$

where the three terms are:

$$\mathcal{L}_{\mathcal{D}} = (\mathcal{D}_\mu \varphi)^\dagger (\mathcal{D}^\mu \varphi) \tag{1.15}$$

$$\mathcal{L}_V = V(\varphi^\dagger \varphi) \tag{1.16}$$

$$\mathcal{L}_{YM} = -\frac{1}{4} W^a_{\mu\nu} (W^a)^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \tag{1.17}$$

The potential is assumed to have the same form of the Ginzburg-Landau potential in the theory of superconductivity, used as a model[2]:

$$V(\varphi^\dagger \varphi) = \mu^2 \varphi^\dagger \varphi + \lambda (\varphi^\dagger \varphi)^2 \tag{1.18}$$

where μ^2 and λ are complex constants.
Two different cases are possible for μ^2:

- if $\mu^2 > 0$, the potential 1.18 has a parabolic shape, with a unique minimum;
- if $\mu^2 = -|\mu^2| < 0$ the potential 1.18 has the form of the so-called "Mexican Hat", as shown in Fig. 1.3. This form is due to the fact that the potential does not have a unique minimum and the ground state with $\varphi = 0$ corresponds to a local maximum of the potential, that is to an unstable equilibrium.

[2]For more information, see [21]

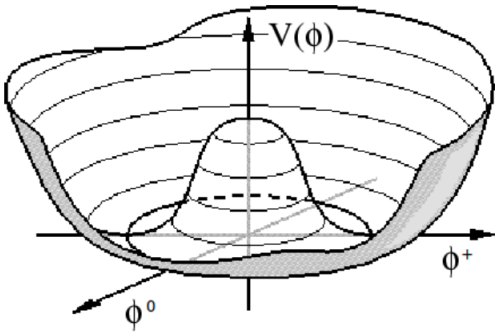

Figure 1.3: Potential $V(\phi)$ for $\mu^2 < 0$

The system is still invariant under global rotations but it is no longer invariant under local transformations.
The symmetry is broken simply choosing φ as a complex doublet with a given hypercharge ($Y = 1$), that is:

$$\varphi = \begin{pmatrix} \varphi^a \\ \varphi^b \end{pmatrix} \tag{1.19}$$

where:

$$\varphi^a = \frac{1}{\sqrt{2}}(\varphi_1 + \varphi_2) \tag{1.20}$$

$$\varphi^b = \frac{1}{\sqrt{2}}(\varphi_3 + \varphi_4). \tag{1.21}$$

Without loss of generality, we can choose $\varphi_1 = \varphi_2 = \varphi_4 = 0$ and $\varphi_3 = \sqrt{\frac{-\mu^2}{2\lambda}}$, in order to have:

$$\varphi_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.22}$$

where $v = \sqrt{\frac{-\mu^2}{\lambda}}$ is the vacuum expectation value of the Higgs field. The choice of a particular minimum spontaneously breaks the $SU(2)$ symmetry.
The Higgs scalar field can be expanded around the ground state:

$$\varphi = e^{i\boldsymbol{\xi}(x)\cdot\boldsymbol{\sigma}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.23}$$

where $\boldsymbol{\xi}(x)$ are real fields corresponding to excitations along the minimum of the potential: they are the so-called massless Goldstone bosons, that can be eliminated by a rotation:

$$\varphi' = e^{-i\boldsymbol{\xi}(x)\cdot\boldsymbol{\sigma}}\varphi(x) = \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.24}$$

This means that the Goldstone bosons do not have a physical meaning; this is the reason why they disappear after a gauge transformation. Instead, the real field $h(x)$ does not vanish, and can be interpreted as the Higgs boson, a real particle.
This expression of the scalar field φ can now be substituted in the Lagrangian 1.14, taking into account only the second order terms in the fields W^a_μ, B_μ and h, and neglecting higher order terms:

$$\mathcal{L}_\mathcal{D} = (\mathcal{D}_\mu\varphi)^\dagger(\mathcal{D}^\mu\varphi) = \frac{1}{2}(\partial_\mu h)(\partial^\mu) + \frac{1}{2}\left(\frac{g^2v^2}{4}\right)(W^1_\mu W^{1\mu} + W^2_\mu W^{2\mu}) + \frac{1}{8}v^2(gW^3_\mu - g'B_\mu)(gW^{3\mu} - g'B^\mu) \tag{1.25}$$

$$\mathcal{L}_V = V(\varphi^\dagger\varphi) = const. + \frac{1}{2}(-2\mu^2)h^2 \tag{1.26}$$

$$\mathcal{L}_{YM} = -\frac{1}{4}W^a_{\mu\nu}(W^a)^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.27}$$

where the term $W^a_{\mu\nu} = \partial_\mu W^a_\mu - \partial_\nu W^a_\mu - g\varepsilon_{abc}W^b_\mu W^c_\nu = \partial_\mu W^a_\mu - \partial_\nu W^a_\mu$ because only second order terms are considered.
In Eq. 1.25 the fields W^3_μ and B_μ appear in mixed products, so the corresponding bosons cannot appear with a physical mass. Therefore two orthogonal combinations of W^3_μ and B_μ can be defined as:

$$\mathcal{Z}_\mu = \cos\theta_W W^3_\mu - \sin\theta_\mathcal{W} B_\mu \tag{1.28}$$

$$\mathcal{A}_\mu = \sin\theta_W W^3_\mu + \cos\theta_\mathcal{W} B_\mu \tag{1.29}$$

where θ_W is the Weinberg angle, that is the mixing angle chosen in a way that the mixed products of $\mathcal{Z}_\mu$ and $\mathcal{A}_\mu$ disappear, that is:

$$\tan\theta_W = \frac{g'}{g} \tag{1.30}$$

The final Lagrangian is:

$$\begin{aligned}\mathcal{L}_H =& \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}(-2\mu^2)h^2 \\ &- \frac{1}{4}W^1_{\mu\nu}W^{1\mu\nu} + \frac{1}{2}\left(\frac{g^2v^2}{4}\right)W^1_\mu W^{1\mu} \\ &- \frac{1}{4}W^2_{\mu\nu}W^{2\mu\nu} + \frac{1}{2}\left(\frac{g^2v^2}{4}\right)W^2_\mu W^{2\mu} \\ &- \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} + \frac{1}{2}\left(\frac{g^2v^2}{4\cos^2\theta_W}\right)\mathcal{Z}_\mu\mathcal{Z}^\mu \\ &- \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + 0\mathcal{A}_\mu\mathcal{A}^\mu \\ &+ \mathcal{L}_{mixed}\end{aligned} \tag{1.31}$$

where:

$$\mathcal{L}_{mixed} = \frac{g^2 v}{4} h W^1_\mu W^{1\mu} + \frac{g^2 v}{4} h W^2_\mu W^{2\mu} + \frac{g^2 v}{4\cos^2\theta_W} h \mathcal{Z}_\mu \mathcal{Z}^\mu \tag{1.32}$$

is the part of the Lagrangian which contains mixed terms.
From the last term of the gauge invariant derivative in Eq. 1.5:

$$W^a_\mu \sigma^a = \begin{pmatrix} W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & -W^3_\mu \end{pmatrix} = \begin{pmatrix} W^3_\mu & \sqrt{2}W^+_\mu \\ \sqrt{2}W^-_\mu & -W^3_\mu \end{pmatrix}, \tag{1.33}$$

the complex fields W^+_μ and W^-_μ are defined as:

$$W^-_\mu = \frac{1}{\sqrt{2}}(W^1_\mu + iW^2_\mu); \tag{1.34}$$

$$W^+_\mu = \frac{1}{\sqrt{2}}(W^1_\mu - iW^2_\mu). \tag{1.35}$$

Replacing W^1_μ and W^2_μ with the complex fields W^+_μ and W^-_μ, Eq. 1.32 becomes:

$$\mathcal{L}_{mixed} = \frac{g^2 v}{4} h W^+_\mu W^{+\mu} + \frac{g^2 v}{4} h W^-_\mu W^{-\mu} + \frac{g^2 v}{4\cos^2\theta_W} h \mathcal{Z}_\mu \mathcal{Z}^\mu \tag{1.36}$$

It is important to notice that the Higgs boson does not couple with the massless gauge field $\mathcal{A}_\mu$ that is the photon. In fact in Eq. 1.31 no terms are associated with the kinetic term $B_{\mu\nu}B^{\mu\nu}$ corresponding to the propagation of the photon. From the same equation it can be seen that the mass terms have appeared implicitly without having introduced them in the Lagrangian. The mass terms are:

- for the $W^\pm$ bosons, which correspond to the gauge fields W^1_μ and W^2_μ:

$$m^2_W = \frac{g^2 v^2}{4} \tag{1.37}$$

- for the Z^0 boson, which corresponds to $\mathcal{Z}_\mu$:

$$m^2_Z = \frac{g^2 v^2}{4\cos^\theta_W} = \frac{m^2_W}{\cos^2\theta_W} \tag{1.38}$$

- for the photon, which corresponds to $\mathcal{A}_\mu$:

$$m^2_\gamma = 0 \tag{1.39}$$

- for the Higgs boson H, associated to h^2:

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda} v \tag{1.40}$$

The last term shows that the mass of the Higgs boson is not predicted by the theory, because μ (or λ equivalently) is an external parameter.
Finally, from Eq. 1.36, the Higgs couplings to the massive gauge bosons $W^{\pm}$ and Z^0 can be extracted:

$$g_{WWH} \equiv \frac{g^2 v}{2} = \frac{2m_W^2}{v}; \tag{1.41}$$

$$g_{ZZH} \equiv \frac{g^2 v}{4\cos^2\theta_W} = \frac{m_Z^2}{v}. \tag{1.42}$$

This means that the Higgs boson coupling to the vector bosons is proportional to the square of their mass.
From the diagonal of the gauge covariant derivative 1.5, and extracting W_μ^3 and B_μ from Eq. 1.28 and Eq. 1.29, the following identity comes out:

$$gW_\mu^3 \frac{\sigma_3}{2} + \frac{g'}{2} B_\mu \mathbb{1} = g\sin\theta_W \mathcal{Q}\mathcal{A}_\mu + \frac{g}{\cos\theta_W}\left(\frac{\sigma_3}{2} - \mathcal{Q}\sin^2\theta_W\right)\mathcal{Z}_\mu \tag{1.43}$$

where $\mathcal{Q}$ is defined as $\mathcal{Q} = \frac{\mathbb{1}+\sigma_3}{2}$.
Therefore, the coupling constant associated with the massless field $\mathcal{A}_\mu$, that represents the electric charge, is:

$$e = g\sin\theta_W \tag{1.44}$$

Other important relations, geometrically illustrated in Fig. 1.4, are:

$$\tan\theta_W = \frac{g'}{g} \tag{1.45}$$

$$\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \tag{1.46}$$

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}} \tag{1.47}$$

$$e = \frac{gg'}{\sqrt{g^2+g'^2}} \tag{1.48}$$

Finally, from Eq. 1.37 and 1.38, one finds:

$$m_W = m_Z\cos\theta_W \tag{1.49}$$

$$\sin^2\theta_W = 1 - \frac{m_W^2}{m_Z^2} \tag{1.50}$$

The importance of these equations lies in the fact that they can be used to find the numerical value of the Weinberg angle, through different experiments like the measurement of the gauge bosons $W^{\pm}$ and Z^0 masses. The most precise measurements of $\sin^2\theta_W$ and of the intermediate bosons $W^{\pm}$ and Z^0 masses are reported in Table 1.1, together with the value of the Fermi coupling constant G_F used to calculate the value of the masses of the bosons through the expression $v = (\sqrt{2}G_F)^{-\frac{1}{2}}$.

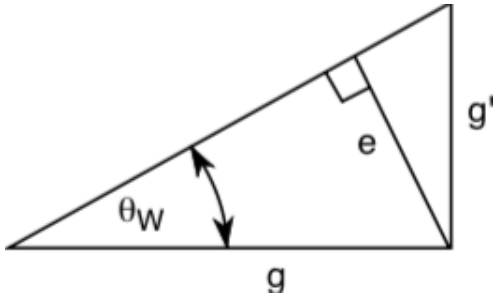

Figure 1.4: Geometrical illustration of the relations between the electroweak coupling constants e, g, g', θ_W.

$m_{W^\pm}$	0.23126 ± 0.00005	GeV/c^2
m_{Z^0}	80.385 ± 0.015	GeV/c^2
$\sin^2\theta_W$	91.1876 ± 0.0021	GeV/c^2
$G_F/(\hbar c)^3$	1.1663787 ± 0.0000006	GeV^{-2}

Table 1.1: Masses of the vector bosons $W^\pm$ and Z^0, and measurements of $\sin^2\theta_W$ and of the Fermi coupling constant G_F.

1.3.1 Fermion Masses

Since at this point fermions appear as massless ($m = 0$), the theory needs to be completed in order to give a non-zero mass to fermions: this is generated through the coupling of the Higgs field to the fermions. The following Yukawa Lagrangian density has thus to be added to the Standard Model Lagrangian:

$$\mathcal{L}_{\varphi-f_i} = -g_{si}\varphi\overline{\psi_i}\psi_i \tag{1.51}$$

where i stands for all types of fermions and the coupling constants g_{si} are arbitrarily chosen in order to reproduce the known physical masses of the fermions. Expanding the Higgs field around the ground state (as in Eq.1.23), the Lagrangian 1.51 becomes:

$$\mathcal{L}_{\varphi-f_i} = -g_{si}\frac{v}{\sqrt{2}}\overline{\psi_i}\psi_i - g_{si}\frac{h}{\sqrt{2}}\overline{\psi_i}\psi_i. \tag{1.52}$$

where:

- the first term is the mass term for each fermion:

$$m_{f_i} = g_{si}\frac{v}{\sqrt{2}}; \tag{1.53}$$

- the second term expresses the coupling between the fermion fields and the Higgs boson field:

$$\frac{g_{si}}{\sqrt{2}} = \frac{m_{f_i}}{v}. \tag{1.54}$$

The Higgs coupling constants to fermions are thus proportional to the corresponding fermion masses.

1.4 Search and discovery of the Higgs boson at LHC

One of the main motivations for the construction of the Large Hadron Collider (LHC) has been the Higgs boson discovery. In fact, this collider allowed to extend the search for the Higgs boson far beyond what was achieved by previous accelerators such as LEP [18] at CERN or Tevatron [22] at FermiLab. The experiments devoted to this purpose are ATLAS (A Toroidal LHC Apparatus) [23] and CMS (Compact Muon Solenoid) [24], large multi-purpose detectors. The Higgs boson mass range explored at the LHC was 100 GeV$< m_H <$1 TeV.

1.4.1 Bounds on the Higgs boson mass

Due to theoretical considerations, m_H should be smaller than 1 TeV, and precision electroweak measurements imply the upper bound $m_H < 152$ Gev at the 95% confidence level. Furthermore, direct measurements for the Higgs boson at LEP led to a lower bound: $m_H > 114.4$ GeV at the 95% CL [25,26]. The Tevatron excluded the mass region $162 \div 166$ GeV at the 95% CL [27,28].

1.4.2 Higgs production channels

The main processes contributing to the Higgs boson production in $p - p$ collisions are:

- gluon-gluon fusion: $gg \to H$ (see Fig. 1.5)
  This channel is dominant in the whole mass range; the t-quark contribution in the loop is significantly higher than the contributions from all the other fermions, because of the larger coupling constant, proportional to the fermion mass.

- vector-boson fusion: qq $\to qqH$ (Fig. 1.6)
  This channel represents the second contribution to the Higgs boson production. It is important because of its clear signature, due to the presence of two hadronic jets in the forward region.

- associated production (Fig. 1.7):

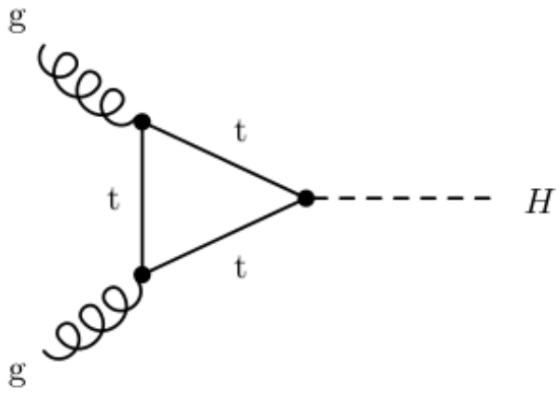

Figure 1.5: Higgs boson production: gluon-gluon fusion.

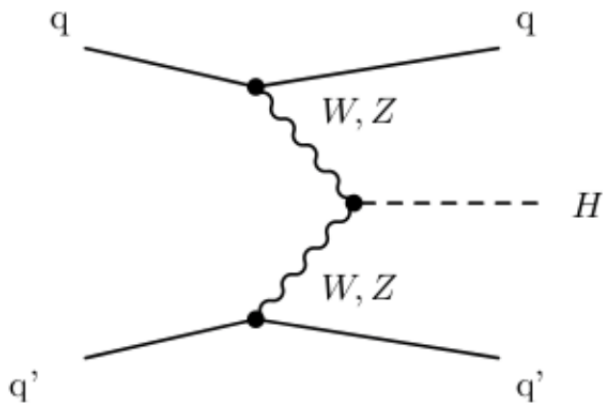

Figure 1.6: Higgs boson production: vector boson fusion.

1. gauge boson: $qq \to VH$
   This channel, where V stands for $W^{\pm}$ or Z^0, is often called *Higgstrahlung*.

2. $t\bar{t}$: $gg \to t\bar{t}H$
   This is the lowest contribution to the Higgs boson production, even if the presence of t-quarks in the final state provides a good experimental signature.

In Fig. 1.8 the Higgs boson production cross-sections are shown as a function of the Higgs mass.

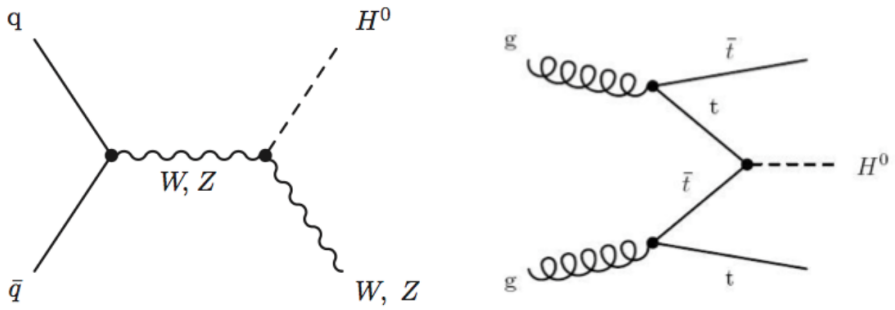

Figure 1.7: Higgs boson production: associated production.

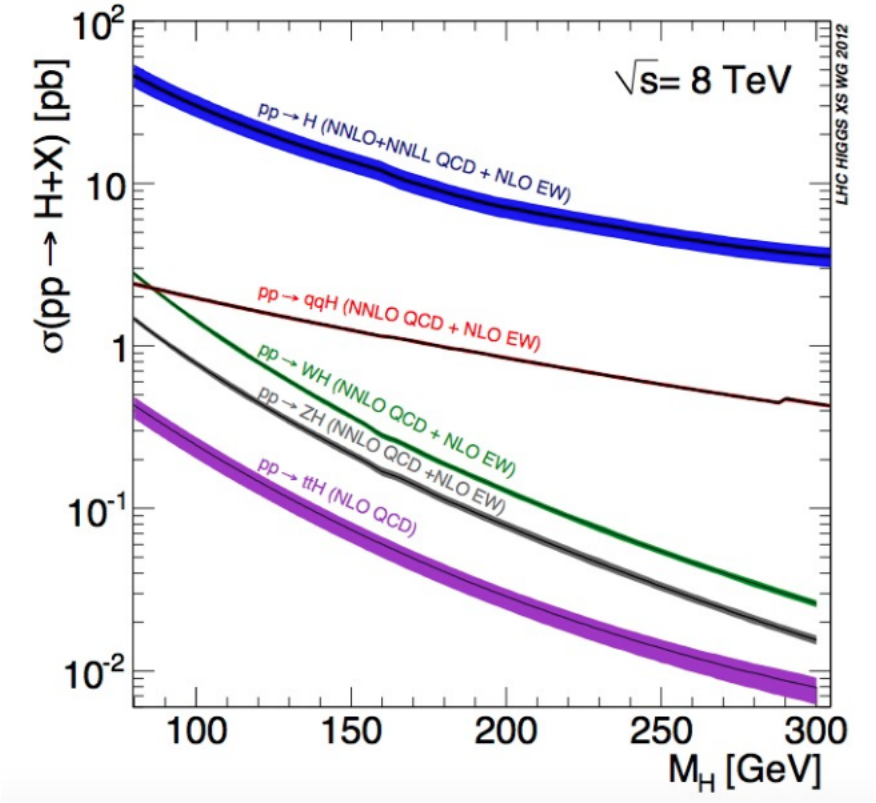

Figure 1.8: Higgs boson production cross-section as a function of the Higgs mass [**?**].

1.4.3 Higgs decay modes

After being produced, the Higgs boson may decay in many final states with different branching ratios, as shown in Fig. 1.9, where the branching ratio (BR) is defined as the fraction of the Higgs bosons decaying in a given channel with respect to the total number

of Higgs bosons produced.

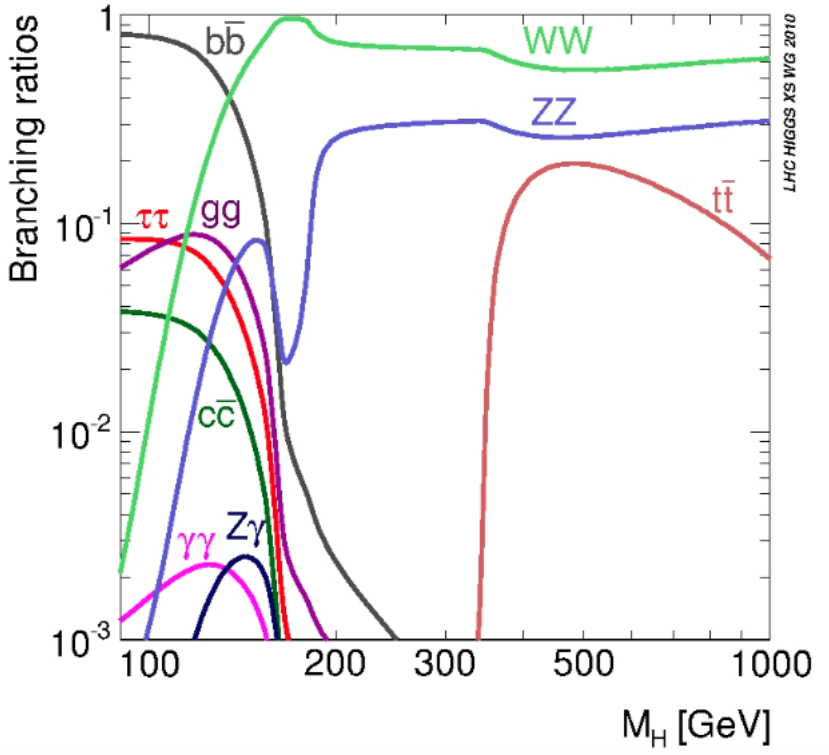

Figure 1.9: Higgs boson decay branching ratios as a function of the Higgs mass [29].

At low Higgs mass ($m_H < 130$ GeV) fermion decay channels are dominant, especially $H \to b\bar{b}$, since the quark b is the most massive fermion kinematically accessible. At higher masses ($m_H > 130$ GeV), the Higgs boson decays to gauge bosons are preferred. At very high masses ($m_H > 350$ GeV), the $H \to t\bar{t}$ decay becomes possible.
Summing all the decay modes, the total width of the SM Higgs resonance peak, Γ_H, which is strongly dependent on m_H can be estimated: $\Gamma_H = 3 \times 10^{-3}$, 0.1 and 200 GeV at $m_H = 120$, 160 and 400 GeV, respectively.
In order to estimate the sensitivity of a channel, the production cross-section must be multiplied by the corresponding branching ratio, as reported in Fig. 1.10.
Since the lepton and photon energy are better measured than the energy of hadron jets, a clearer signature is expected for these channels. For this reason, among the many different signatures searched at LHC, the lepton and photon final states are favored with respect to hadronic final states (they suffer large QCD background at hadron colliders, especially in the forward region).

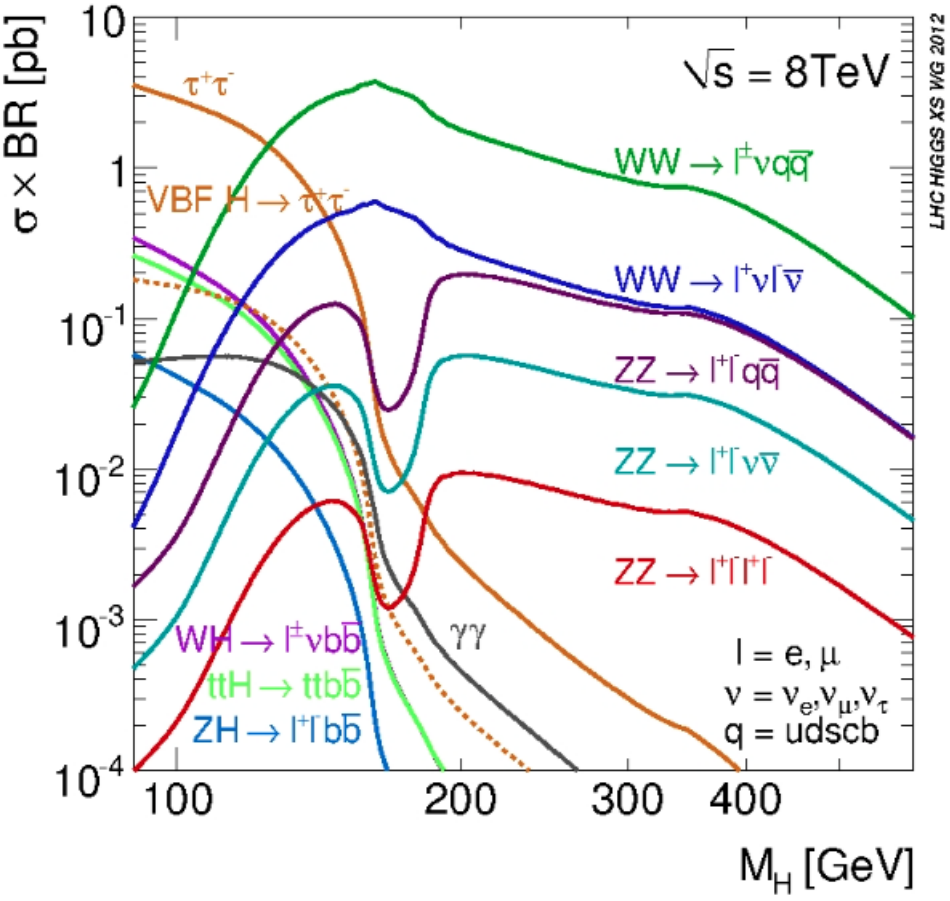

Figure 1.10: Higgs boson production cross-sections times branching ratios as a function of the Higgs mass [29].

1.4.4 The $gg \to H \to ZZ^* \to 4l$ process

Since in this thesis the gluon-gluon production mechanism and the Higgs decay in the four-lepton final state via ZZ^* production has been studied, these processes are reviewed in more details.
Gluon-gluon fusion is the main Higgs production mechanism, and the corresponding Feynman diagram with the four-lepton decay is shown inf Fig 1.11.

The main background source is composed of the non-resonant $ZZ \to 4l$ process, where ZZ or $Z\gamma^*$ produce the four-lepton final state via $q\bar{q}$ annihilation and gluon fusion, as shown in the Feynman diagrams of Fig:1.12. This background is called irreducible because it produces the same final state as the signal.

The reducible background, which produces a different final state which is mis-identified

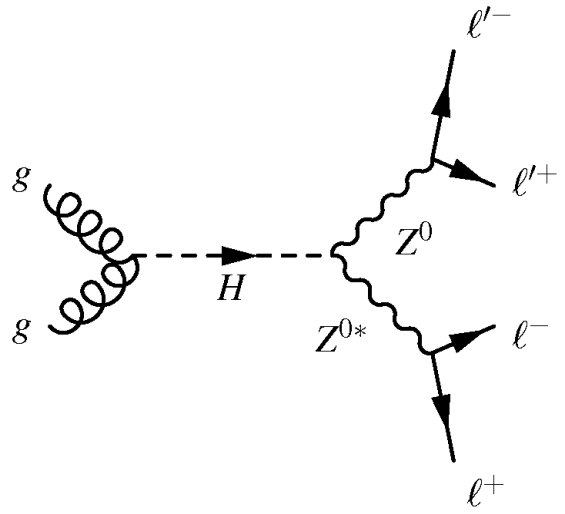

Figure 1.11: Feynman diagram for the process $gg \to H \to ZZ^* \to 4l$ process.

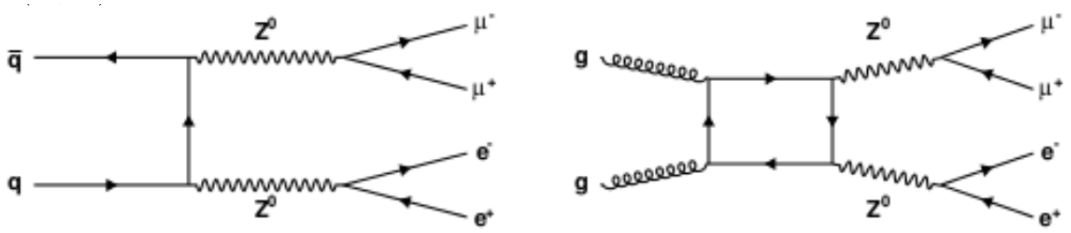

Figure 1.12: Irreducible background $ZZ \to 4l$ Feynman diagrams.

as a $H \to 4l$ final state, is composed of $Z + jets$, $t\bar{t}$ and $WZ + jets$ processes. This background contains non-isolated leptons coming from heavy-flavor quark decays (b quarks in the $t\bar{t}$ decay), mis-reconstructed jets (in $Z + jets$ and $WZ + jets$ processes) and electrons from photon conversions. In Fi.g 1.13 the Feynman diagrams related to the $Z + jets$ background processes are shown.

At a center-of-mass energy $\sqrt{s} = 14$ TeV, and with a simulated Higgs mass of $m_H = 125$ GeV, the production cross-sections times branching ratio for the signal processes are [29]:

- 1.5 fb for 4μ and $4e$ final states;
- 3.0 fb for the $2e2\mu$ final state.

For the non resonant background processes:

- 20 fb for 4μ and $4e$ final states;
- 40 fb for the $2e2\mu$ final state.

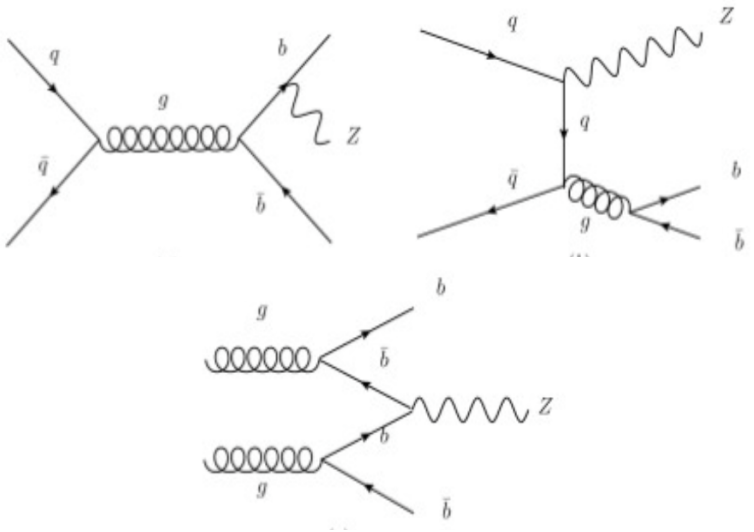

Figure 1.13: Reducible background $Z + jets$ Feynman diagrams.

1.4.5 Higgs boson discovery

The channels studied by the CMS and ATLAS experiments, in the region $110 \div 160$ GeV, are:

- $H \to \gamma\gamma$
- $H \to ZZ^*$
- $H \to W^+W^-$
- $H \to \tau^+\tau-$
- $H \to b\bar{b}$

During the two years of data taking at the LHC, the CMS experiment excluded the mass range $127 \div 600$ GeV at 95% CL, while the ATLAS experiment excluded the ranges $111.4 \div 116.6$, $119.4 \div 122.1$ and $129.2 \div 541$ GeV, all at the 95% CL.
In addition to these excluded mass regions, the CMS and ATLAS experiments found an excess of events at 125 GeV in the data corresponding to an integrated luminosity of 5 fb$^-$1 and a center-of-mass energy $\sqrt{s} = 7$ TeV. In 2012, when the center-of-mass energy was increased to 8 TeV, more data became available. Summing the data recorded at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV, both experiments, ATLAS and CMS, observed an excess of events above the expected background, with a local significance[3] of 5.1 σ and 5.0 σ respectively, at a mass near 125 GeV [31, 32].

[3]For more details on Statistical Methods, see [30]

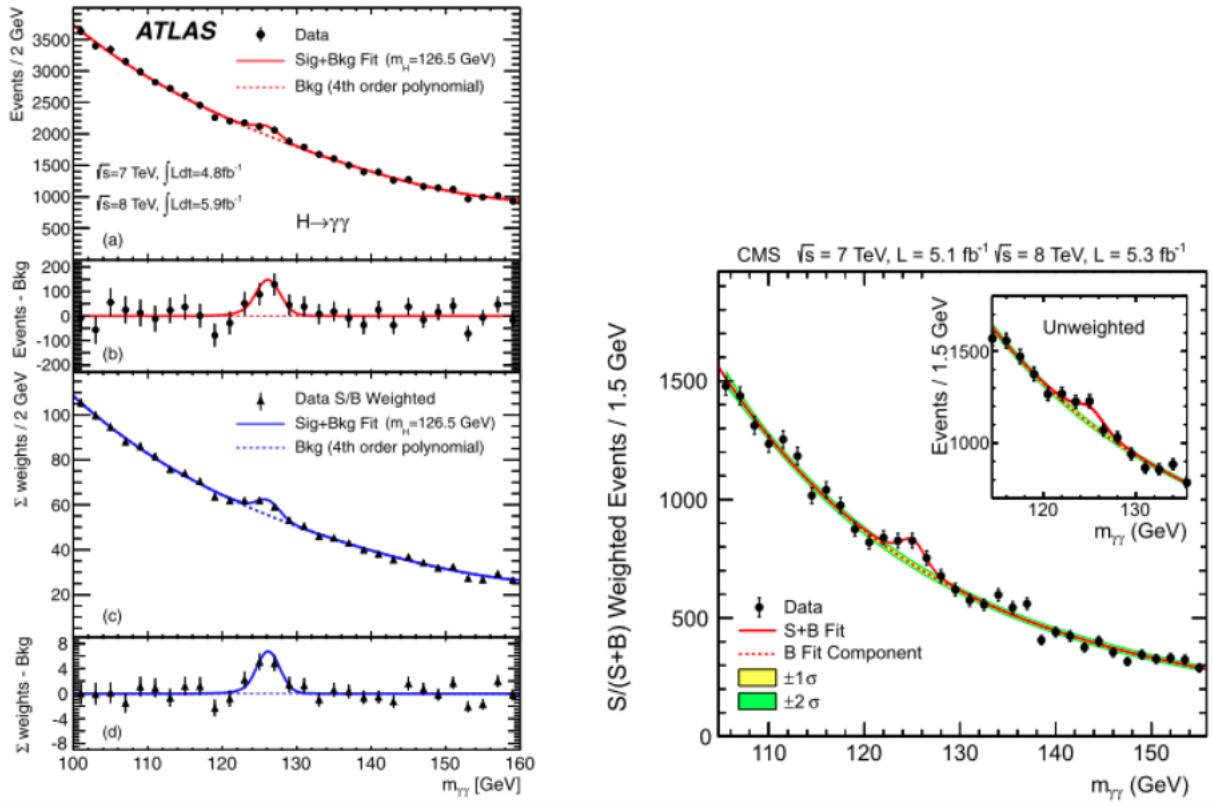

Figure 1.14: ATLAS (on the left) and CMS (on the right) di-photon invariant mass distributions [31, 32].

The results obtained by the two experiments, are:

- $H \to \gamma\gamma$
  The search for the SM Higgs boson through this decay has been performed in the mass range $110 \div 150$ GeV, where the dominant background is di-photon production and the production of jets mis-identified as photons. The invariant mass of the $\gamma\gamma$ system is reported in Fig. 1.14. A clear peak, over a steeply falling background, corresponding to $m_H \approx 125$ GeV, is present.

- $H \to ZZ^*$
  In this decay mode a search has been made for a peak in the four-lepton invariant mass distribution in the presence of a small continuum background: the invariant mass of the four-lepton final states obtained by the CMS experiment is shown in Fig. 1.15.

These decays into vector bosons indicated that the new particle had to be a boson; in addition, the di-photon decay implied that its spin was different from 1 (the spin 2 hy-

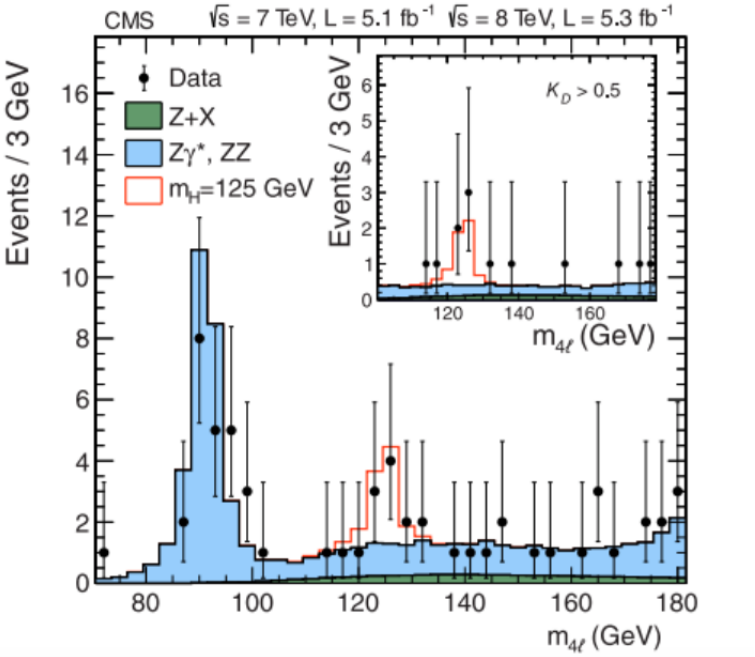

Figure 1.15: CMS distribution of the four-lepton invariant mass for the $ZZ^* \rightarrow 4l$ analysis [32].

pothesis is now excluded at more than 99% C.L. by the analysis of the kinematic of the decay into vector bosons [33]). Combining the individual results, the Higgs boson was excluded at the 95% CL in the range $110 \div 121.5$ GeV and above, and a significant excess was seen in the range $121.5 \div 128$ GeV.
The consistency of the observed excess with the background-only hypothesis is visible in Fig. 1.16, which shows the value of the local p-value for the 7 and 8 TeV data sets and their combination.
In the CMS experiment, the 7 and 8 TeV data sets exhibit an excess of 3.2 σ and 3.8 σ significance, respectively, for a Higgs boson mass of approximately 125 GeV. A significance of 5.0 σ for $m_H = 125.5$ GeV was obtained using the overall combination of the data. In the ATLAS experiment instead, the combined data exhibit an excess of 5.1 σ.

The mass of the Higgs boson is estimated using the $\gamma\gamma$ and ZZ^* decay modes: in Fig. 1.17 the two-dimensional 68% CL regions for the signal strength σ/σ_{SM} versus m_H for these channels and the values of the signal strength for the individual decay modes and for their combination are shown for the CMS data. The mass obtained for the Higgs boson was, on the 4th July 2012, the date of the announcement of the Higgs discovery:

$$m_H^{ATLAS} = 126.0 \pm 0.4(stat.) \pm 0.4(sys.),$$

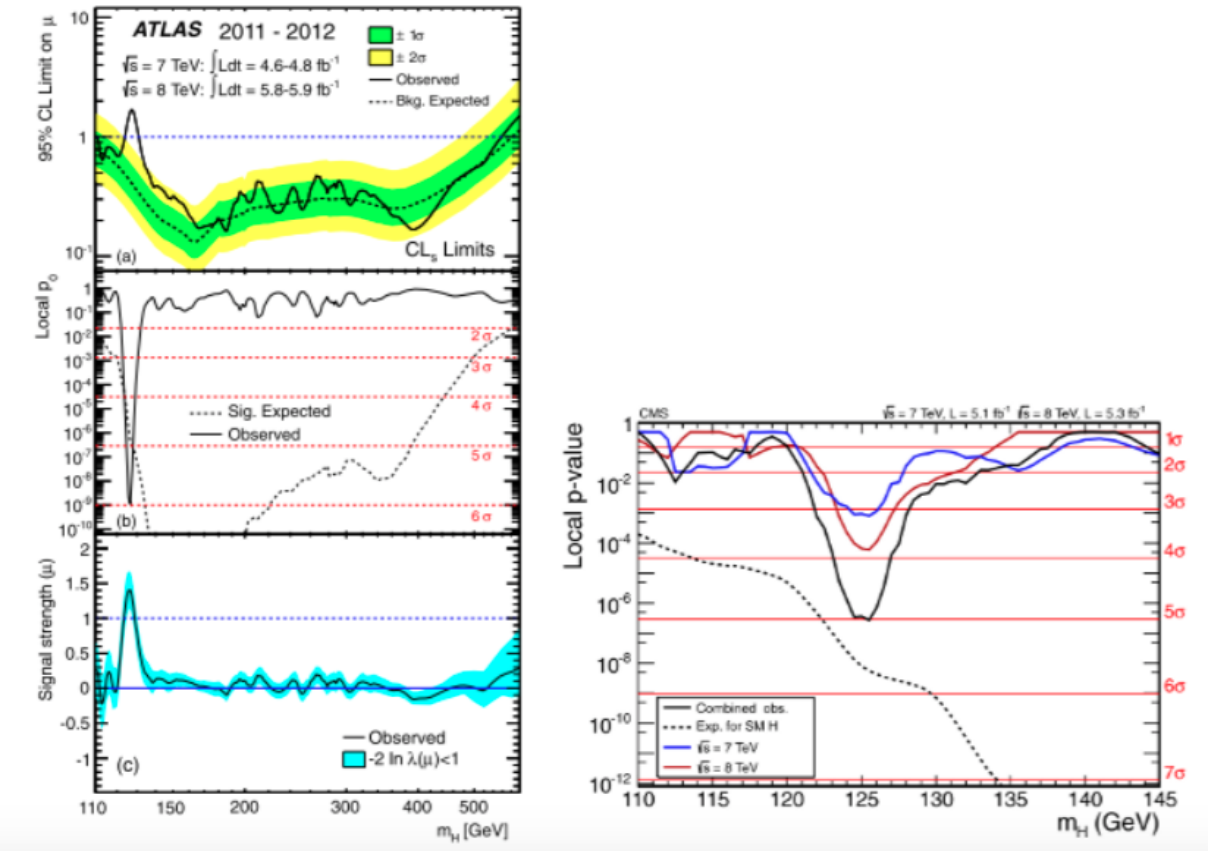

Figure 1.16: ATLAS (on the left) and CMS (on the right) observed (solid) local p-value as a function of the Higgs boson mass and the expectation (dashed) for a Higgs boson signal hypothesis [31, 32].

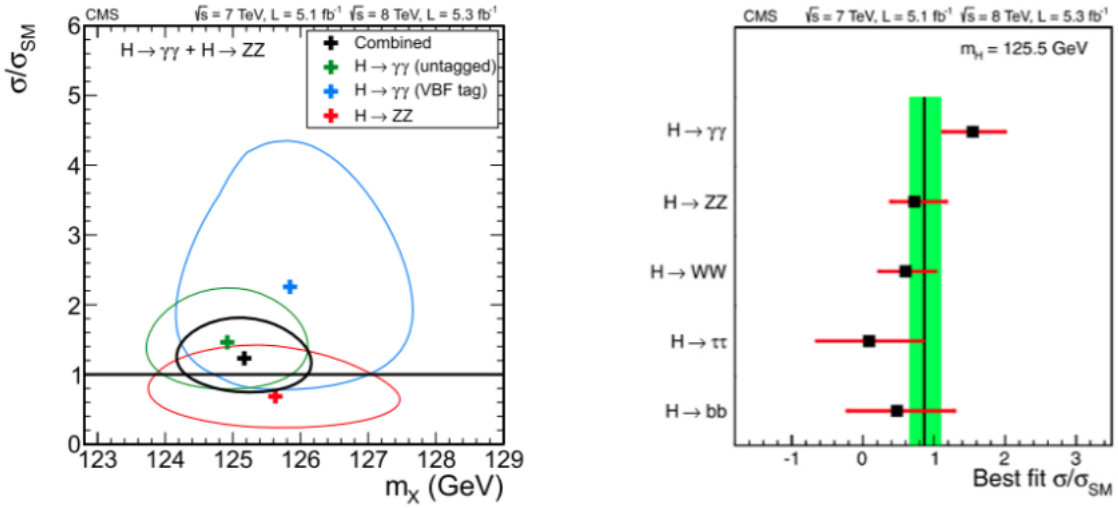

Figure 1.17: On the left, the 68% CL contours for the signal strength σ/σ_{SM} versus the Higgs boson mass measured by CMS; on the right, values of σ/σ_{SM} for the combination (solid vertical line) and for individual decay modes (points) for the CMS experiment [32].

$$m_H^{CMS} = 125.3 \pm 0.4(stat.) \pm 0.5(sys.).$$

Chapter 2

The CMS Detector at LHC

2.1 The Large Hadron Collider at CERN

The *Large Hadron Collider* (LHC) [19] superconducting proton-proton (and heavy ions for short periods) collider is installed at CERN (*European Organization for Nuclear Research*) [34]. LHC is located in the 26.7 km long tunnel which was previously hosting the LEP (*Large Electon Positron*) [18] collider, about 100 m under the French-Swiss border near Geneva. The goal of this unprecedented collider in terms of size, luminosity and center-of-mass energy is the search for rare physics events, the validation of the Standard Model (SM) and the search for new physics Beyond the Standard Model (BSM), e.g. SuperSymmetric (SUSY) particles.
The 23^{rd} November 2009 the first $p-p$ collisions were produced, at a center-of-mass energy of $\sqrt{s} = 900$ GeV; then, after some pilot runs at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 2.36$ TeV, the 30^{th} March 2010 the energy of each beam was set to 3.5 TeV for a total energy of 7 TeV in the center-of-mass. Finally on the 5^{th} April 2012 the energy was raised to $\sqrt{s} = 8$ TeV, until the end of 2012, when the LHC was shut down for maintenance and upgrade. In spring 2015, LHC will restart with a center-of-mass energy of $\sqrt{s} = 13$ TeV.

Proton injection and acceleration

The proton production, injection in the LHC ring and acceleration up to 4 TeV for each beam proceeds in several steps:

1. Hydrogen atoms are ionized to produce protons;
2. protons are firstly accelerated up to 50 MeV through the LINAC (*Linear Accelerator*);
3. protons are then accelerated up to 1.4 GeV using the PSB (*Proton Synchrotron Booster*);

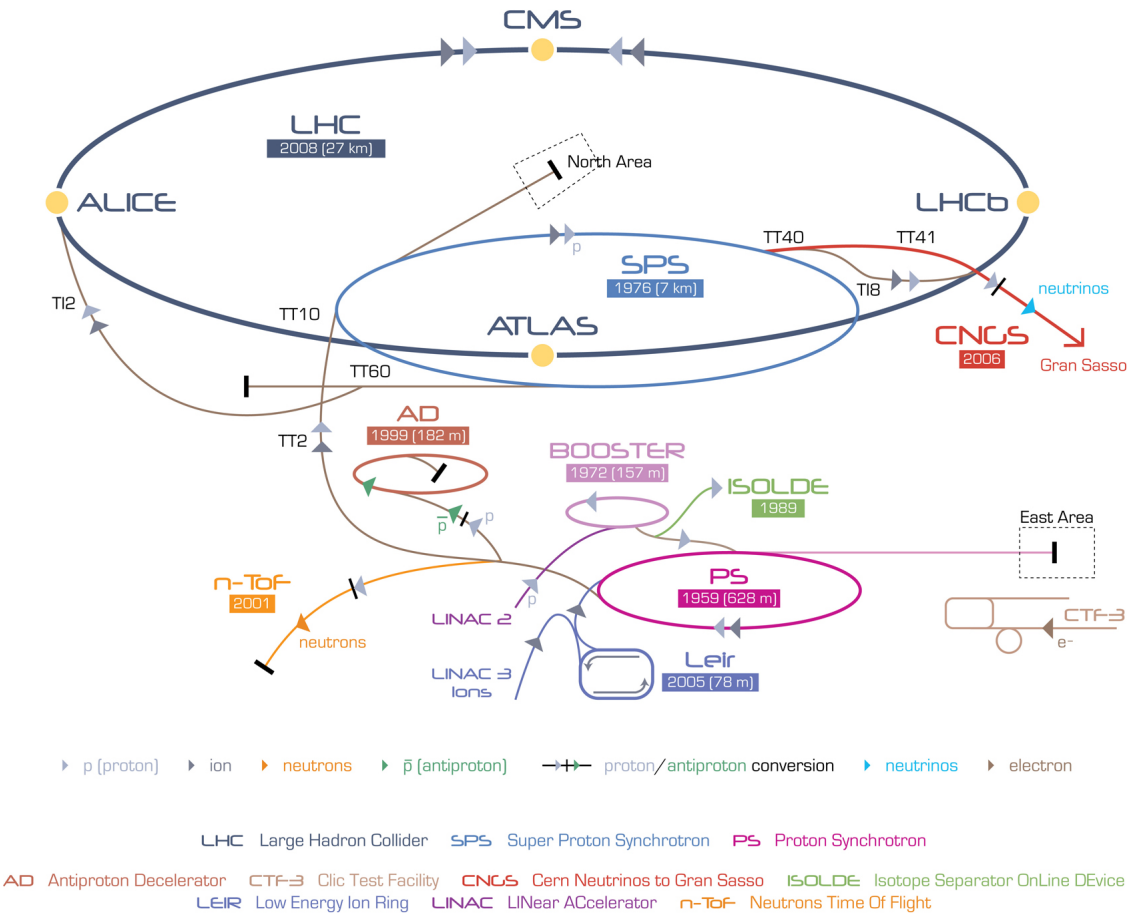

Figure 2.1: CERN LHC ring, detectors and pre-accelerators layout

4. the PS (*Proton Synchrotron*) accelerates the protons up to 26 GeV;

5. protons reach the SPS (*Super Proton Synchrotron*) where they are accelerated up to 450 GeV;

6. protons are finally injected in bunches into the LHC, where they are accelerated up to the final energy of 3.5 TeV.

In Fig. 2.1 a layout of the LHC with all experiments and pre-accelerators is shown.

In the LHC protons run in two adjacent parallel beam pipes, separated by 194 mm. Beam 1 circulates clockwise while Beam 2 circulates counterclockwise. Each beam is composed of about 1400 bunches. The two beam pipes intersect at four interaction points, in correspondence of the four experiments which detect the particles resulting from the $p - p$ collisions:

- ATLAS (*A Toroidal LHC Apparatus*) [23] and CMS (*Compact Muon Solenoid*) [24]: general purpose experiments;
- LHCb (*LHC beauty experiment*) [35]: b-quark physics devoted experiment;
- ALICE (*A Large Ion Collider Experiment*) [36]: heavy ion experiment analyzing $Pb-Pb$ collisions to study the quark-gluon-plasma state of matter.

To maintain a circular trajectory, a strong magnetic field is needed. Since the particles colliding have the same charge, each beam must be provided with opposite magnetic fields. Twin bore coil dipole magnets, hosting the two beam lines, are installed inside the same mechanical structure and cryostat, reducing the cold mass (to be kept at 1.9 K) and the space required by the equipment. The intensity of the magnetic field necessary to bend protons in the LHC accelerator is given by:

$$p[\text{TeV}] = 0.3\ B[\text{T}]\ r[\text{km}], \tag{2.1}$$

where p is the proton momentum ($p \simeq 7$ TeV) and r is the LHC radius ($r \simeq 4.2$ km). In these conditions, the magnetic field has to be $B = 5.4$ T. About 2/3 of the beam line is equipped with 1232 magnet coils of 14.3 m length each, made with copper-clad niobium-titanium cables. Each magnet can provide a magnetic field $B = 8.33$ T at a temperature of T=1.9 K. Over 96 tons of liquid helium are used to keep the temperature down to this critical value. To focus the beam, 392 quadrupole magnets are employed.

The event rate produced (that is the number of events per second) is given by:

$$R = \mathcal{L}\ \sigma_{\text{proc}} \tag{2.2}$$

where σ_{proc} is the production cross-section of the physics process under study, and $\mathcal{L}$ is the luminosity, a machine characteristic expressed by:

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \tag{2.3}$$

where f is the bunch crossing frequency, n_1 and n_2 are the particles contained in Bunch 1 and Bunch 2 respectively, and σ_x and σ_y are the transverse dimensions of the beam. The design bunch crossing interval (BX), 25 ns, will be adopted in Run2, starting in spring 2015, while in Run1 (up to 2012) it was 50 ns. In Fig. 2.2 the total integrated luminosity is reported: in 2010 and 2011, with $\sqrt{s} = 7$ TeV, LHC reached a peak instantaneous luminosity of $5{\cdot}10^{33}$ cm^{-2}s^{-1}, delivering 44.2 pb^{-1} and 6.1 fb^{-1} of integrated luminosity in 2010 and 2011 respectively; in 2012, with $\sqrt{s} = 8$ TeV, the peak instantaneous luminosity increased to $7 \cdot 10^{33}$ cm^{-2}s^{-1} for an integrated luminosity of 23.3 fb^{-1} [37]. The ATLAS and CMS experiments are designed to operate at high luminosity (10^{34} cm^{-2}s^{-1}), while the instantaneous luminosity delivered to LHCb and ALICE are 10^{32} cm^{-2}s^{-1} and 10^{27} cm^{-2}s^{-1}, respectively.

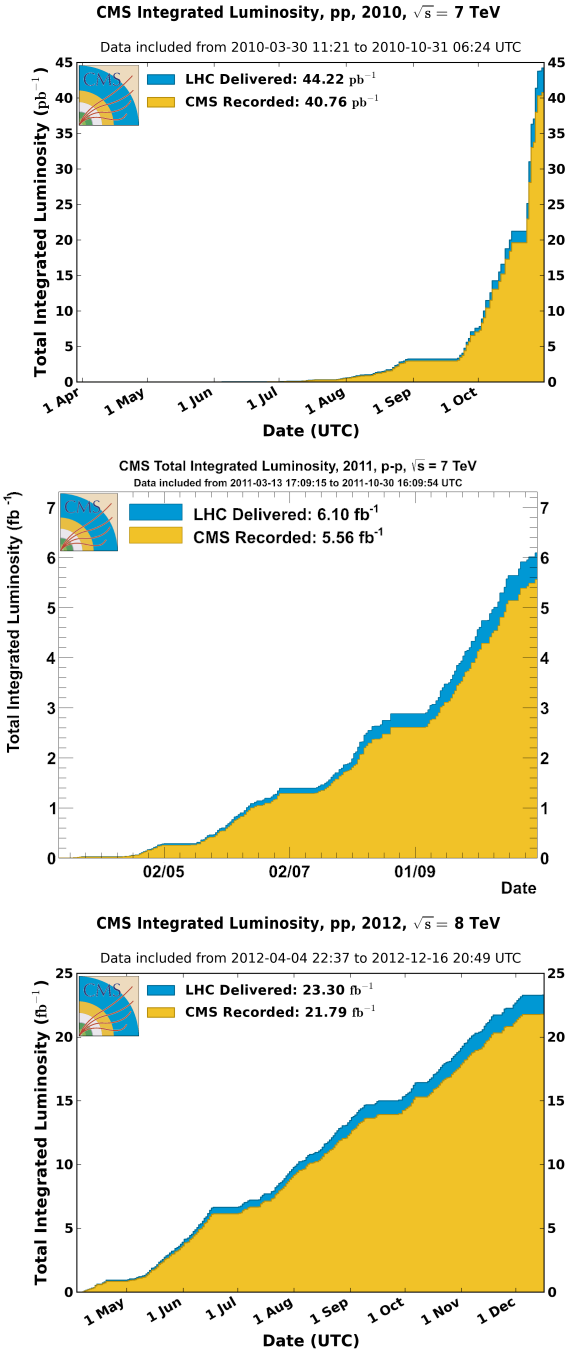

Figure 2.2: Total integrated luminosity delivered by LHC (in blue) and recorded by the CMS detector (in yellow) for 2010 (top), 2011 (middle) and 2012 (bottom) [37].

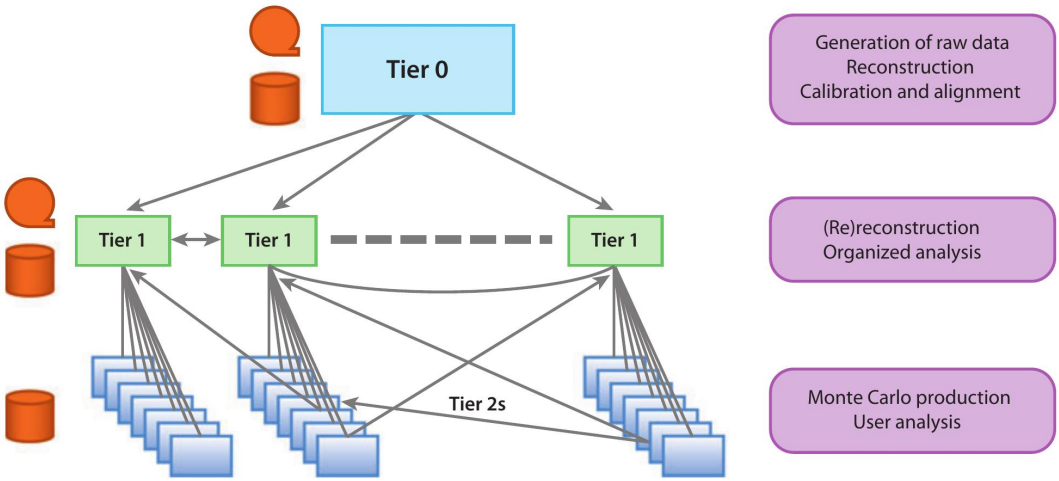

Figure 2.3: Scheme of the Tiered structure of the WLCG [38].

2.1.1 The LHC data processing

The LHC produces ∼15 PB of data per year, which have to be accessible to thousands of individual users for their analyses. CERN collaborates with institutions in 34 different countries to provide a distributed computing and data storage infrastructure, the Worldwide LHC Computing Grid (WLCG) which is the largest world's computing grid [39, 40]. In the WLCG infrastructure and operation costs are shared among participating institutions, and single point of failures are avoided (multiple copies of data are stored in many different places). In fact it is structured in several layers, called Tiers. The Tier-0 is located at CERN: it stores on tape the raw data (RAW) from the Data Acquisition system (DAQ), performs event reconstruction, and distributes both RAW and reconstructed (RECO) data to Tier-1 sites. They are large computer centers located in Germany, Canada, France, Italy, the Netherlands, the Nordic countries, Spain, Taipei, the UK, and the USA. They have enough storage capability to store a copy of the RAW and RECO data, and serve as no-stop support for the computing grid. They distribute the data to 200 Tier-2 so that scientists and experiments can access LHC data from their countries through Tier-3 computing resources (local clusters in Universities or individual PCs).

2.2 The CMS Experiment

CMS is a general purpose experiment [24]. Its main goals are:

- to study the electroweak symmetry breaking due to the Higgs mechanism, observe

the Higgs boson and measure its properties;

- to search for physics Beyond the Standard Model (BSM), which could involve the presence of SuperSymmetric particles (SUSY) or Z' and W' new heavy gauge bosons;
- to explore in more detail QCD processes at extreme conditions of temperature, density and energy;
- to confirm the Standard Model (SM) with very precise measurements of known electroweak and flavor physics phenomena.

To achieve this wide variety of goals, a versatile experimental structure is needed, with an optimal particle identification (ID).

2.3 The CMS Detector

The CMS detector [41, 42] has a cylindrical structure, 21.6 m long with a diameter of 14.6 m and a total weight of approximately 14500 tons. In order to describe physics quantities and the detector geometry, a proper reference frame is defined.

Coordinate Frame

A right-handed cartesian reference frame is used, with the origin centered in the CMS interaction point, defined as follows:

- the x-axis is horizontal, pointing towards the center of the LHC ring;
- the y-axis is vertical, pointing upwards;
- the z-axis is tangent to the beam line.

The $x - y$ plane, which is orthogonal to the beam pipe, is called the transverse plane, while the z-axis direction is called longitudinal. Since the CMS detector has a cylindrical symmetry, cylindrical coordinates can be used in reconstruction algorithms. They are defined using:

- r: distance from the interaction point in the transverse plane $x - y$, $r = \sqrt{x^2 + y^2}$;
- ϕ: azimuthal angle, measured from the x-axis in the transverse plane;
- θ: polar angle, measured from the z-axis in the longitudinal plane $z - y$.

Using these coordinates, many useful variables can be defined:

- the particle momentum can be split in the longitudinal and transverse components:

$$p = \sqrt{p_z^2 + p_T^2} \qquad \text{where } p_T = \sqrt{p_x^2 + p_y^2}; \tag{2.4}$$

- the transverse energy and mass can be defined respectively as:

$$E_T = E \ \sin\theta \qquad m_T = \sqrt{p_T^2 + m^2}; \tag{2.5}$$

- the missing transverse energy, according to momentum conservation, can be computed as:

$$E_T^{miss} = -\sum_i \vec{p}_T^{\,i} \qquad \text{with } \left|E_T^{miss}\right| = \sqrt{\left(\sum_i p_x^i\right)^2 + \left(\sum_i p_y^i\right)^2} \tag{2.6}$$

  where i represents every final state particle;

- the rapidity of a particle, which is a Lorentz-invariant variable, is:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}; \tag{2.7}$$

- the pseudo-rapidity of a particle, which approximates the rapidity in the case of high energy particles, is defined as:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{2.8}$$

Operational and physics challenges

The LHC regime is very challenging for the CMS detector.

1. With the design luminosity ($\mathcal{L} \simeq 10^{34}$ cm^{-2}s^{-1}) and the expected $p-p$ cross-section at $\sqrt{s} = 14$ TeV ($\sigma_{pp} \simeq 100$ mb), the expected interaction rate is $R = \mathcal{L}\ \sigma_{pp} \simeq 10^9$ Hz.

2. The number of protons per bunch is $N = 1.1\ 10^{11}$: this leads to pile-up effects, that is the overlapping of many events in the same data acquisition time interval. For the designed luminosity, and at $\sqrt{s} = 8$ TeV, the pile-up is ~ 20 overlapped events.

3. The radiation level is very high, especially in the regions close to the beam pipe, causing damage and aging of the sensitive detector materials.

The main features of the CMS detector and its Data Acquisition System (DAQ) are:

1. efficient on line trigger system to select physics signals and reduce the rate from 10^9 Hz to $0.1 - 1$ kHz;

2. high granularity of the sub-detectors to reduce the occupancy and very good time resolution to resolve multiple interaction vertices;

3. radiation hard detectors and electronics devices.

Proper particle identification and good event reconstruction are required to achieve the physics goals of the experiment. CMS provides high quality physics objects (muons, electrons, jets, etc) with excellent characteristics:

- good muon identification, muon charge determination and momentum resolution in a wide range of momenta and angles, yielding to a high precision in the di-muon invariant mass calculation;

- good charged particle momentum resolution and reconstruction efficiency, with a good tagging of b quarks and τ leptons to identify τ's and b-jets;

- good electromagnetic energy resolution, precise di-photon and di-electron mass resolution in a wide η range;

- good missing transverse energy and di-jet mass resolution, achievable with a hermetic and fine segmentation calorimeter structure.

Detector Structure

In Fig. 2.4 a schematic representation of the CMS detector is reported. It can be divided in three main sections:

- *barrel*: central region, composed of five wheels, coaxial to the beam axis;

- *endcaps*: two regions orthogonal to the beam axis that hermetically close the *barrel* at both ends; composed of three disks each;

- *very forward regions*: sub-detectors very close to the beam axis to allow the detection of particles in a very high pseudo-rapidity range.

From the inside-out, the detector presents the following sub-systems:

- **Tracker**, composed of high granularity silicon pixel detectors in the inner region, and of silicon microstrip detectors in the outer region. This detector allows an efficient charged particle track reconstruction, with primary and secondary vertices identification;

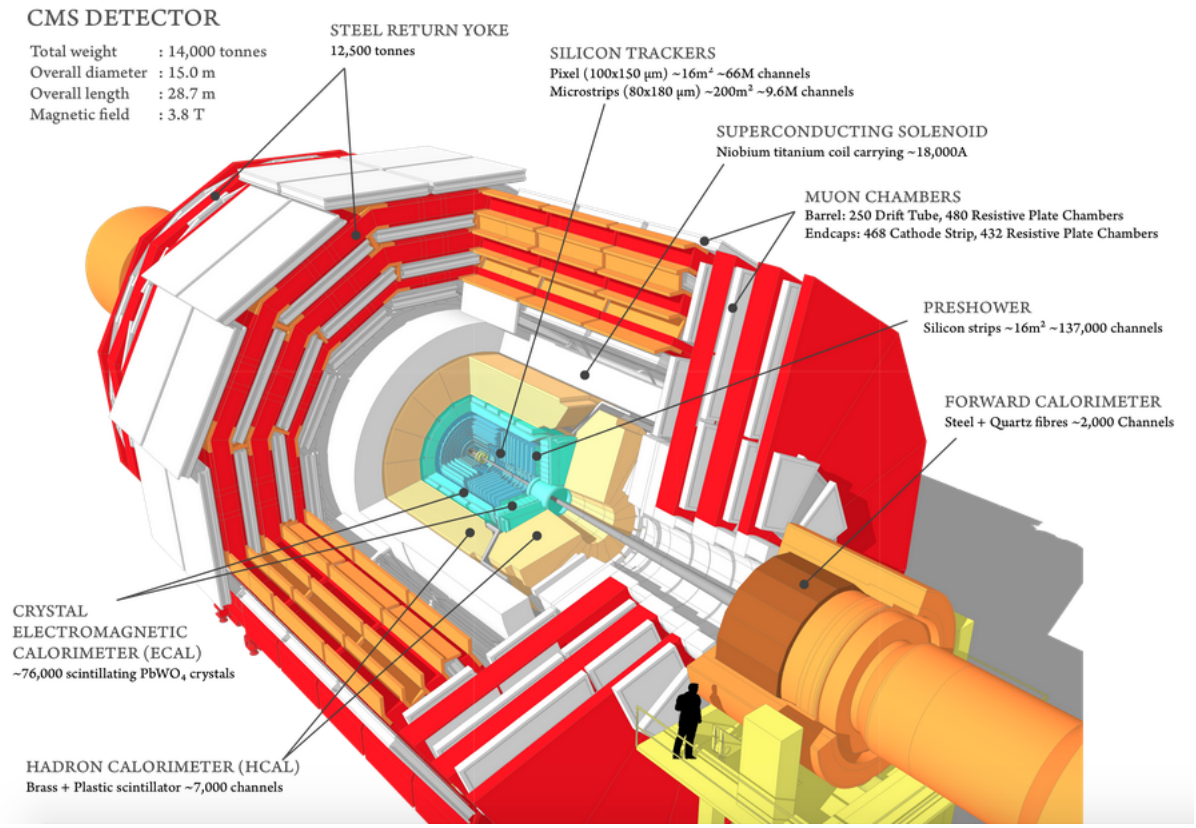

Figure 2.4: A perspective view of the CMS detector.

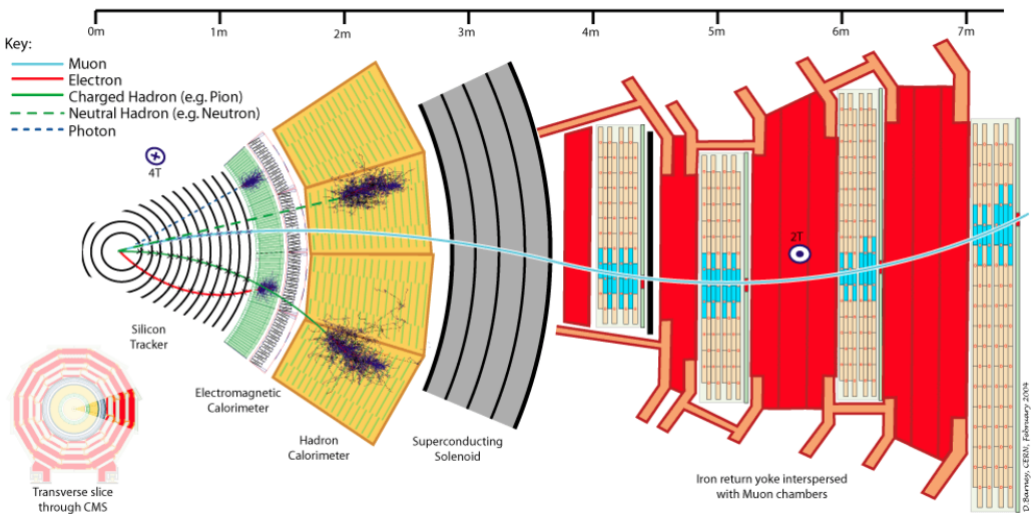

Figure 2.5: Transverse view of the CMS detector, with the signature of muons (cyan line), electrons (red line), charged hadrons (green line), neutral hadrons (green dotted line), photons (blue dotted line).

- **Electromagnetic Calorimeter (ECAL)**, made of lead tungstate scintillating crystals, preceded in the endcaps by a pre-shower detector. This part of the detector allows the identification of photons and electrons ($e^{\pm}$) thanks to the energy they deposit in the material;

- **Hadron Calorimeter (HCAL)**, which is a brass-scintillator sampling calorimeter, used to identify hadrons and reconstruct jets;

- **Magnet**, made of a superconducting solenoidal coil providing a 3.8 T magnetic field. This is used to bend charged particle tracks in the tracker and identify the particle charge and momentum;

- **Muon System**, consisting of Drift Tubes in the central region and Cathode Strip Chambers in the endcaps, both paired with Resistive Plate Chambers to ensure redundancy. The muon detector is hosted in the iron of the return yoke of the superconducting magnet; it allows muon identification, together with their charge determination, and a standalone p_T measurement important for the trigger system.

In Fig. 2.5 a transverse view of the CMS detector is shown, with the signature that different particles leave in the detector. In the next Sections the sub-detectors and the electronics devices that form the CMS experimental apparatus are described in more detail.

2.3.1 Magnet

The CMS magnet [43] is constituted of a superconducting solenoid made of niobium-titanium (NbTi) cables wrapped with copper. It is kept at $T = 4$ K to maintain the superconducting mode in a vacuum cylinder which isolates it from the outside. The potential 4 T magnetic field which is able to produce is lowered to 3.8 T in order to maximize the longevity of the material. The structure, 12.5 m long, with an inner diameter of 6 m and a total weight of 220 t, is completed by an external iron yoke which is responsible for the return of the magnetic flux. The yoke is made of 5 layers in the barrel and 3 disks for each endcap, with a total weight of 10000 tons; it extends up to 14 m in length and absorbs all particles except for muons and neutrinos. The magnet provides a significant bending power, which allows precise measurements of the transverse momentum of charged particles, either in the tracker or in the iron yoke.

2.3.2 Tracker

The CMS Tracker [44], which is the largest tracker system ever built for a collider experiment, is the sub-detector closest to the interaction point. Its purpose is the reconstruction of charged particle tracks and vertices in a high particle density condition. Moreover, the strong magnetic field provided by the CMS solenoid allows precise momentum measurements. The main requirement for this detector are:

- high efficiency in all the η range down to very low p_T ($p_T \gtrsim 0.5$ GeV);
- good particle momentum resolution;
- efficient primary and secondary vertex reconstruction;
- good pattern recognition, which consists in the recognition of all the hits produced by a single particle in the sensitive material.

These goals are achievable thanks to:

- low occupancy, thanks to the high detector granularity;
- fast detector response;
- large redundancy, using many layers to collect more than 10 hits along the particle trajectory.

For these reasons, the tracker is entirely made of silicon detectors: they cover the region $|\eta| < 2.5$ with a radius $r < 1.2$ m and for $|z| < 2.7$ m, for 5.8 m in length and a total surface of 210 m^2. The thickness of the silicon sensors changes as a function of the pseudo-rapidity, being 0.35 radiation lengths (X_0) at small η, 1.8 X_0 in the transition

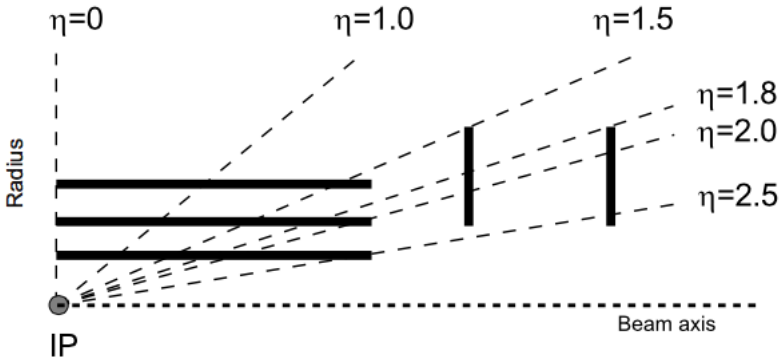

Figure 2.6: Geometrical layout of the pixel detector.

region between barrel and endcap, and 1.1 X_0 at $|\eta| \simeq 2.5$. The silicon detectors can be divided into two categories: Pixels and Microstrips. The Pixels provide very low occupancy, high resolution and precise vertex reconstruction: for this reason they are chosen to cover the region close to the beam pipe. Instead, Microstrips cover the more extended region outside the Pixel detector; the Microstip silicon detectors allow to reduce the number of read-out channels, maintaining a good resolution.

Pixel Detector

The Pixel Detector is the closest detector to the interaction point, where the particle flux is highest ($\sim 10^7$ particles per second). It is composed of ∼66 million pixel cells, each of 100×150 μm^2, grouped in 1400 sensors for a total surface of 1.06 m^2. They are disposed in two regions, as shown in Fig. 2.6:

- three layers in the barrel region (BPix), each 53 cm long and at a radius $r = 4.4$ cm, $r = 7.3$ cm, $r = 10.2$ cm respectively;
- two disks for each endcap (FPix), made of 24 blades in a turbine-like shape each, at radius $r = 7.3$ cm and $r = 15$ cm respectively.

A spatial resolution of 10 μm in the transverse plane $r - \phi$ and of 15 μm in the $z-$coordinate are achieved in the barrel region, while lower resolutions (15 μm and 20 μm respectively) are achieved in the endcaps.

Silicon Strip Tracker

The Microstrip detector region extends from $r = 20$ cm to $r = 120$ cm. It can be divided in two different regions, as shown in Fig. 2.7:

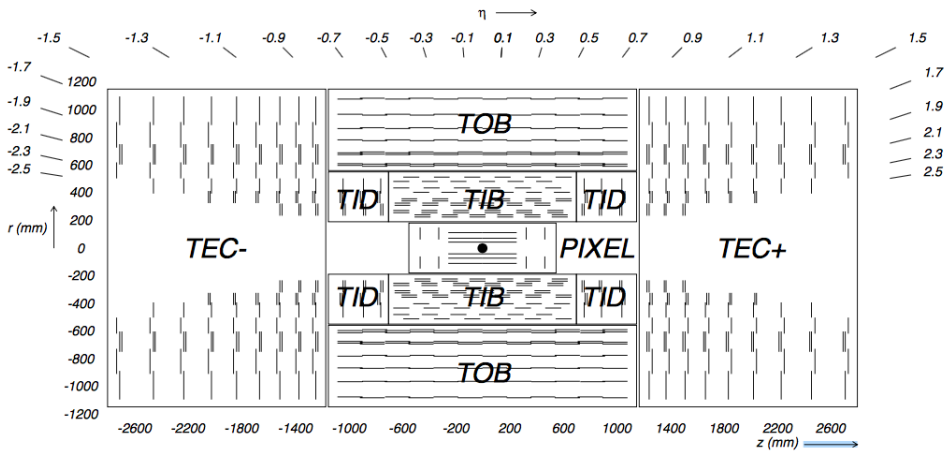

Figure 2.7: Transverse view of the CMS tracker: each line represents a detector module.

- the inner region (20 cm $< r <$ 55 cm), which is composed of four layers in the barrel (TIB, Tracker Inner Barrel) and three disks in each endcap (TID, Tracker Inner Disk) with a minimum cell size of 10 cm$\times$80 μm;

- the outer region, ($r >$ 55 cm), which is composed of six barrel layers (TOB, Tracker Outer Barrel) and nine disks in each endcap (TEC, Tracker EndCap) with a size of 25 cm$\times$180 μm^2.

This detector provides a spatial resolution ranging from 40 to 60 μm in the $r-\phi$ transverse plane, and of 500 μm in the $z-$coordinate.

2.3.3 Electromagnetic calorimeter

The CMS electromagnetic calorimeter (ECAL) [45] is a homogeneous and hermetic calorimeter made of more than 75000 lead tungstate ($PbWO_4$) crystals. This scintillating material detects the electromagnetic shower produced through Bremsstrahlung and pair production. The shower shape allows to measure the energy of the emitting particle and identify electrons and photons. The material was chosen for its high density ($\rho = 8.28$ g/cm^3), short radiation length ($X_0 = 0.89$ cm), small Molière radius[1] (2.2 cm) and very short scintillation time (in 25 ns almost 80% of the light is collected by silicon avalanche photo-diodes in the barrel and vacuum photo-triodes in the endcaps). These characteristics allow the ECAL to be compact, fast, and with a fine granularity. The ECAL is divided in two regions: barrel ECAL (EB) and endcap ECAL (EE), as shown in Fig. 2.8.

[1] $R_M = 21.2\frac{X_0}{\varepsilon_c}$ MeV [46].

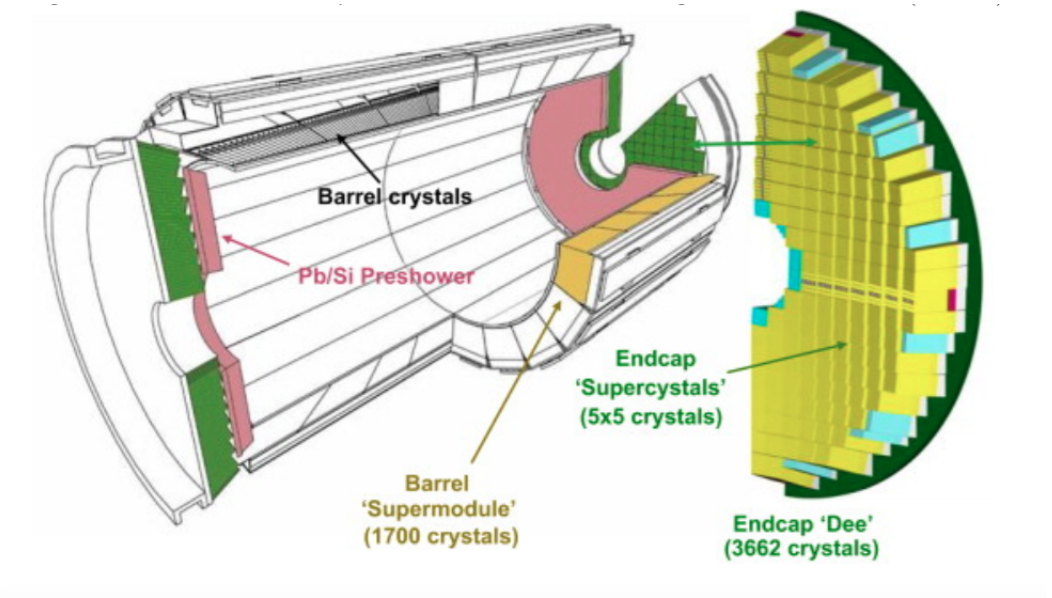

Figure 2.8: Layout of the CMS ECAL, showing the arrangement of crystal modules, supermodules and endcaps, with the pre-shower in front.

Barrel ECAL

The EB covers the pseudo-rapidity range $|\eta| < 1.479$, with 61200 crystals contained in a thin-walled alveolar structure set at a radius $r = 1.29$ m. Each crystal has a surface of 26×26 mm^2 and a length of 230 mm which correspond to 25.8 X_0. The crystals are mounted in a truncated pyramid geometry, tilted of 3° with respect to the axis from the interaction vertex, in both the ϕ and η directions, to avoid cracks aligned with particle trajectories.

Endcap ECAL

The EE covers the pseudo-rapidity range $1.479 < |\eta| < 3$. It consists of 7324 crystals 220 mm long (24.7 X_0) and with a surface of 30×30 mm^2, grouped in supercrystals. A pre-shower detector (ES) is placed in front of the EE to identify π^0 in the pseudo-rapidity region $1.653 < |\eta| < 2.6$. The ES is a sampling calorimeter composed of two layers: lead radiators to initiate the shower and silicon strip sensors to measure the deposited energy and the transverse shower profiles, for a total thickness of 20 cm.

The energy resolution of the ECAL can be parametrized with three terms:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{2.9}$$

where E is the particle energy and:

- S is the stochastic term, which accounts for fluctuations in the number of photo-electrons produced and fluctuations in the shower-containment;
- N is the noise term, due to the electronics and pile-up noise;
- C is a constant term, which is related to the calorimeter calibration, and to the energy leakage of the crystals.

Using test beams, a typical energy resolution was found to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2 \tag{2.10}$$

where E is expressed in GeV.

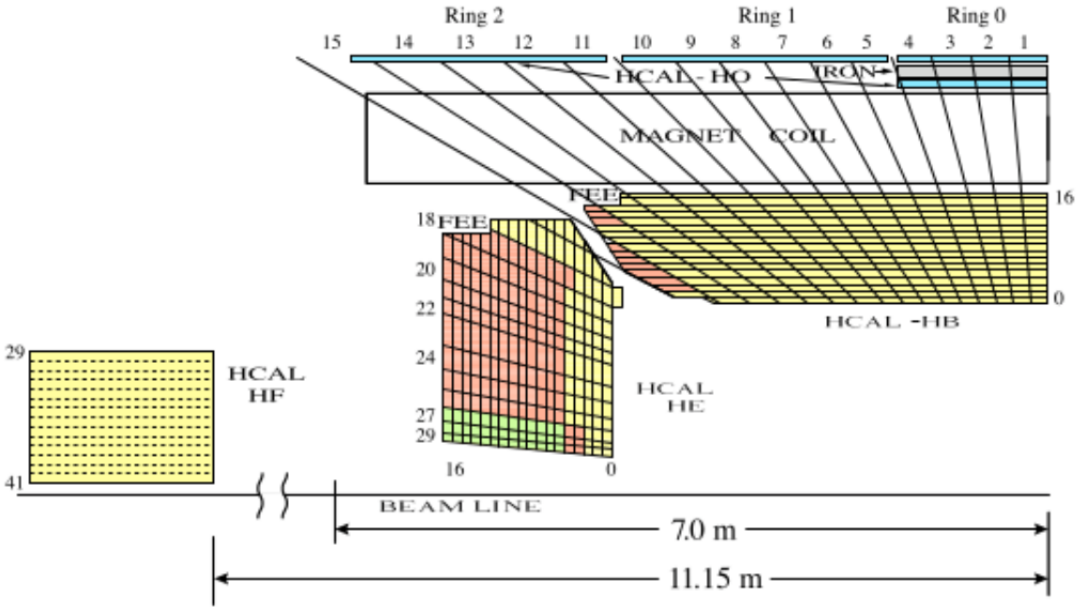

Figure 2.9: Longitudinal view of the CMS hadron calorimeter (HCAL)

2.3.4 Hadron calorimeter

The Hadron Calorimeter (HCAL) [47] is used to identify hadrons and measure the hadron energy deposits in order to reconstruct jets and measure the missing transverse energy. For this reason the calorimeter has to be hermetic up to the maximum η region possible, here up to $|\eta| = 5$. The calorimeter can be divided in two regions, as shown in Fig. 2.9:

- the Barrel Hadron Calorimeter (HB), up to $|\eta| < 1.26$, and two Endcap Hadron Calorimeters (HE), up to $|\eta| < 3$, constrained between the ECAL and the magnet, from $r = 1.77$ m to $r = 2.95$ m. They are sampling calorimeters, made of brass layers (absorber) alternated to plastic scintillators (active material), coupled to hybrid photo-diodes using wavelength-shifting fibers. Being non-magnetic and with a short interaction length (λ_0), the brass is particularly useful to obtain small shower dimensions.

- Two Forward Calorimeters (HF), around the beam-pipe at $|z| = 11.2$ m which cover up to $|\eta| < 5$, to increase the hermeticity. This part is built with radiation hard materials, being close to the beam line. In fact, steel plates are used as absorbers, while quartz fibers are used as active material (producing Cherenkov light at the passage of relativistic particles).

The Outer Calorimeter (HO), outside the magnet coil, is added to improve the energy resolution of the barrel calorimeters, catching the tails of the hadron showers.

The depth of the calorimeter is a function of the pseudo-rapidity, being 5.25 λ_0 at $|\eta| = 0$, 9.1 λ_0 at $|\eta| = 1.3$ and 10.5 λ_0 at $|\eta| \simeq 5$. The energy resolutions, for the different regions, are:

- $\frac{\sigma}{E} \simeq \frac{65\%}{\sqrt{E}} \oplus 5\%$ for HB
- $\frac{\sigma}{E} \simeq \frac{85\%}{\sqrt{E}} \oplus 5\%$ for HE
- $\frac{\sigma}{E} \simeq \frac{100\%}{\sqrt{E}} \oplus 5\%$ for HF

where E is the energy of the particle measured in GeV and $\oplus$ stands for the sum in quadrature.

2.3.5 Muon system

The CMS muon system [48] is the outer part of the detector and is designed to:

- identify muons, since they are the only charged particles which can penetrate all the inner detector layers, reaching the muon spectrometer;

- measure the muon transverse momentum.

Since the system is contained into the iron yoke, the measurement of the transverse momentum is possible thanks to the magnetic bending power ($B \simeq 1.8$ T) created by the return flux. The detector covers the pseudo-rapidity region $|\eta| < 2.4$ and is entirely made of gaseous detectors, that use the ionization electrons created by the passage of charged particles in the gas to produce the signal.

Three different types of gaseous detectors are employed:

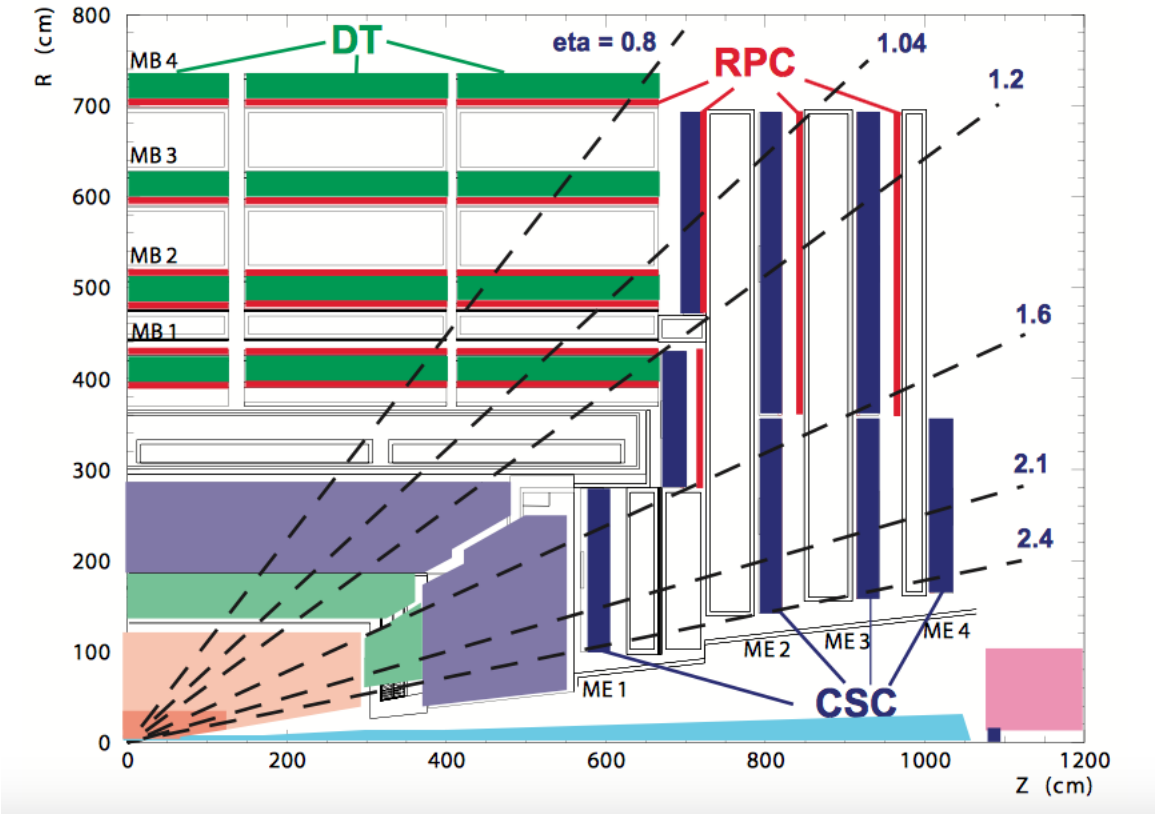

Figure 2.10: Longitudinal view of the CMS muon system: DTs are colored in green, CSCs in blue, RPCs in red.

- Drift Tubes (DTs) used in the barrel region, covering up to $|\eta| < 1.2$;
- Cathod Strip Chambers (CSCs) used in the endcaps covering the region $0.9 < |\eta| < 2.4$;
- Resistive Plate Chambers (RPCs) in the region $|\eta| < 1.6$, to improve the DTs and CSCs performances and ensure redundancy.

As shown in Fig. 2.10 and in Fig. 2.11, two different regions are formed: Muon Barrel (MB), where four stations of DTs and RPCs are located, divided into five wheels in the z-direction, and Muon Endcap (ME) composed of four disks orthogonal to the beam axis where the CSCs and RPCs are located.

Drift Tubes

Drift chambers with rectangular drift tubes are used in the barrel region. A schematic representation of a DT cell is shown in Fig. 2.12. They are organized in four stations:

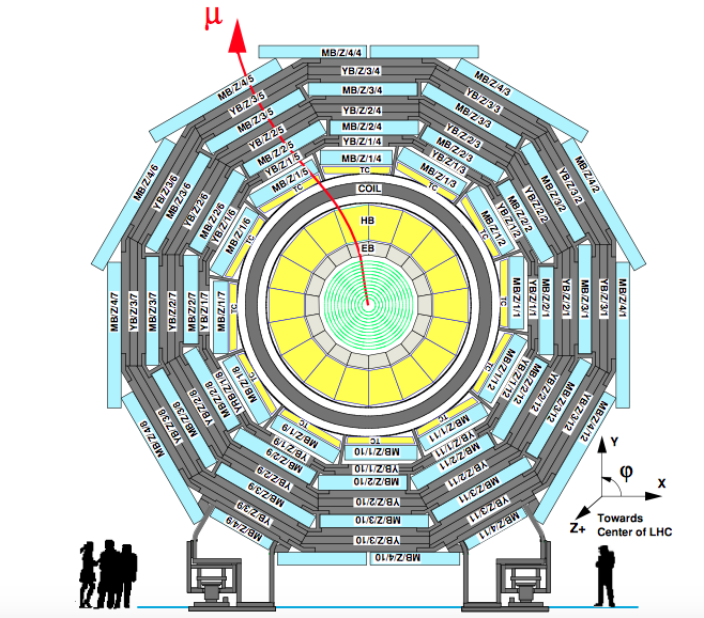

Figure 2.11: Transverse view of the CMS barrel muon system. Each wheel consists of twelve sectors formed by DTs (light blue) embedded in the yoke (gray).

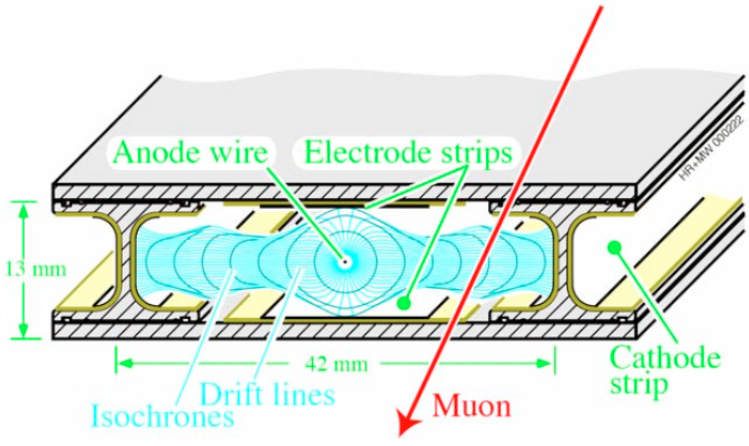

Figure 2.12: Section of a DT cell used in the barrel region of the CMS muon system.

MB1, MB2 and MB3 contain eight layers of drift cells used to measure the muon position in the $r - \phi$ plane and four layers to measure the coordinate in the z-direction, while MB4 has only the 8 layers to measure the coordinate in the $r - \phi$. Each cell, which covers an area of 4.2×1.3 cm^2, is filled with a mixture of Ar (85%) and CO_2 (15%). Consecutive layers are offset by half a cell width to improve the coverage and efficiency and provide an accurate BX identification using the Mean Timer algorithm [49]: with this configuration, an excellent time resolution is possible.

Cathode Stripe Chambers

These detectors, located in the endcap region, are multi-wire proportional chambers with the cathode plane segmented in strips orthogonal to the anode wire, in order to have a 2-D information about the muon position. The structure of a CSC is shown in Fig. 2.13. Thanks to their fan-shape, they can easily be arranged in the endcap regions. In fact, each CSC is formed by trapezoidal panels mounted on eight disks, four in each endcap, partially overlapping in the $\phi-$plane to improve the coverage and efficiency. CSCs are filled with a mixture of Ar (30%), CO_2 (50%) and CF_4 (20%).

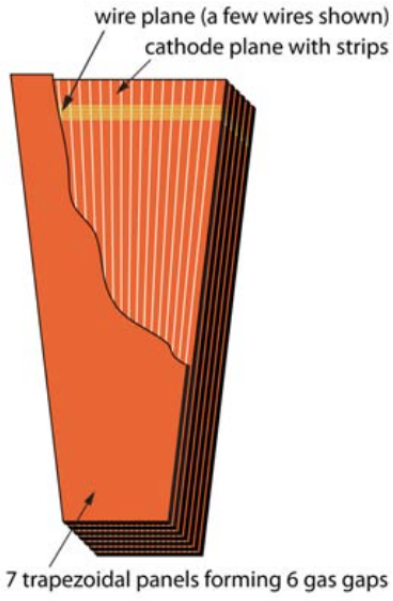

Figure 2.13: Schematic view of a CSC used in the endcap region of the CMS muon system.

Resistive Plate Chambers

The RPCs are placed both in the barrel and in the end-cap, to provide redundancy. They are made of four Bakelite planes which form two gaps filled with a mixture of $C_2H_2F_4$ (96,5%) and C_4H_{10} (3.5%). Since their response time is about 3 ns, they ensure an excellent time resolution and are thus used for triggering. Moreover, the presence of a double gap provides a high efficiency with lower electric fields, with respect to single gap chambers.

The overall space resolution of the muon system is of the order of 250 μm in the $r-\phi$ plane and of 500 μm in the z-direction, and the reconstruction efficiency is close to 100 %.

2.3.6 Trigger and Data Acquisition system

The CMS Trigger and Data Acquisition System (DAQ) system [50, 51] is designed to collect and analyze the detector information every 25 ns, that is at every bunch crossing. A rate of 10^9 interactions per second is expected and with $\sim 10^8$ data channels per event, a total amount of $\sim$ 1 MB should be stored for each event. This quantity is too high to be easily handled by the storing process. For this reason, the trigger is a fundamental part of the experiment, making a real-time selection of the events to store. A multi-level trigger is adopted: the first, called Level-1 Trigger (L1), is based on custom hardware electronics, and reduces the rate from $\sim$40 MHz to $\sim$100 kHz with a latency of 4 μs; the second, called High Level Trigger (HLT), performs a software event building, selected reconstruction and event selection on commercial processors, reducing the rate to $\sim$100 Hz.

Level-1 Trigger

The L1 Trigger has to take a decision on accepting an event every 25 ns. For this reason it is based on the rough identification of particles in the sub-detectors. This trigger operates using a pipelined structure, which allows the temporary storage of the full detector information in pipeline memories, for up to 4 μs from the collision. The trigger is organized in three sub-sectors:

- L1 Calorimeter Trigger, which identifies electrons ($e^{\pm}$), photons, jets and missing transverse energy;
- L1 Muon Trigger, which identifies muons from the muon system;
- L1 Global Trigger, which takes the final decision according to pre-defined algorithms.

Each one of them is organized in sub-structures, as shown in Fig. 2.14.

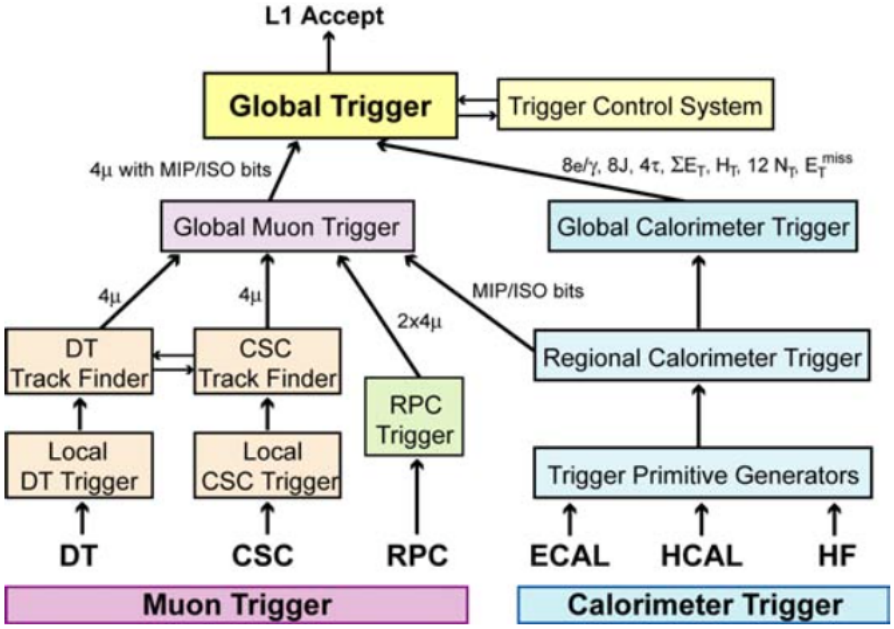

Figure 2.14: Overview of the CMS Level 1 Trigger.

High Level Trigger

The HLT receives data filtered from the L1 Trigger and performs a better reconstruction and selection of the events. This process is based on signatures to assign events to specific data-sets, suitable for different studies, such as the search for top-quarks, Higgs boson, SUSY particles, etc. The selection proceeds in two steps:

- Level-2 Trigger (L2): a first selection algorithm which relies only on the calorimeter and muon system information; this part is fast and runs on all events.

- Level-3 Trigger (L3): it includes the reconstruction of the complete tracks in the tracker, a process that requires a large amount of CPU time and is thus conditioned to the L2 decision.

The result of the final selection is called RAW data, containing the detector information, the L1 Trigger and HLT results. It is stored in the Tier-0 at CERN for the reconstruction step and sent to the GRID for access by all the CMS collaboration.

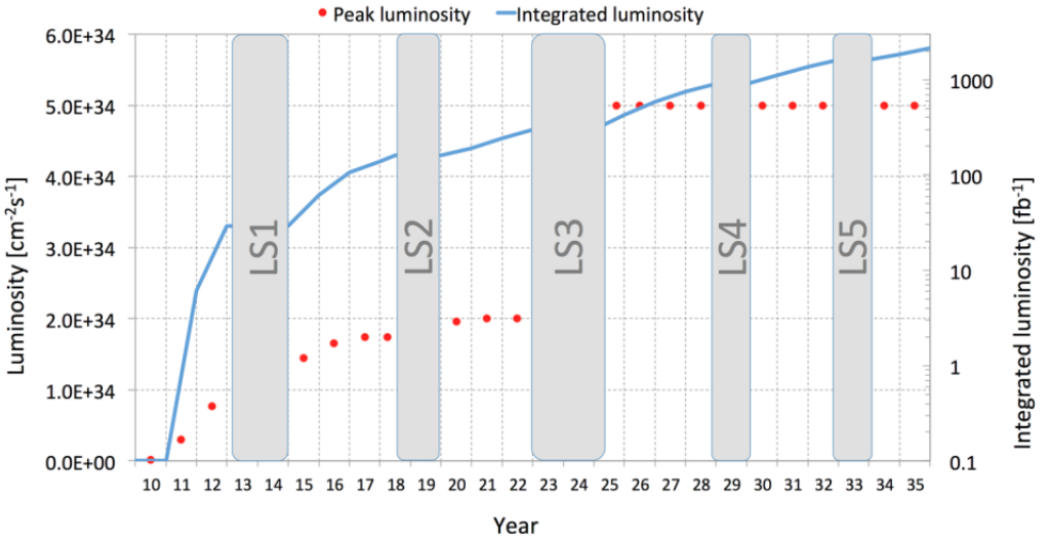

Figure 2.15: Scheme of the planning of shutdown and LHC performances for the next 20 years.

2.4 The CMS upgrade

2.4.1 LHC future performances

As reported in Fig. 2.15 the LHC will be operated with cycles of three years of data taking interleaved with shutdown periods used to maintan and upgrade the machine. After the current Long Shutdown 1 (LS1), which is going to finish in the spring of 2015, Run2 will start: the beam energy will be 6.5 TeV (for a center-of-mass energy $\sqrt{s} = 13$ TeV) with a bunch crossing interval of 25 ns. After 2016, the beam energy will possibily be increased to 7 TeV ($\sqrt{s} = 14$ TeV). The instantaneous peak luminosity will exceed $\mathcal{L} = 1 \cdot 10^{34}$ cm^{-2}s^{-1}, providing an integrated luminosity per year of ~ 45 fb^{-1}. In 2019, the Long Shutdown 2 (LS2) will start: this shutdown period will be used to apply the so-called Phase-I upgrade. At the end of the LS2 shutdown period, the center-of-mass energy should be raised up to $\sqrt{s} = 14$ TeV, and Run3 will begin with an instantaneous luminosity that will reach $2 \cdot 10^{34}$ cm^{-2}s^{-1}, that is twice as much as the 2015 conditions. At the end of Run3, the integrated luminosity will be of the order $\sim 300 - 400$ fb^{-1}. The Long Shutdown 3 (LS3) is planned to start in 2023 to significantly increase the luminosity of the LHC, with a major upgrade of the accelerator. This new operating regime is called High Luminosity LHC (HL-LHC). The peak instantaneous luminosity will be levelled at $\mathcal{L} = 5 \cdot 10^{34}$ cm^{-2}s^{-1} with $\sqrt{s} = 14$ TeV, and at the end of a 10 year data taking period, the integrated luminosity will reach ~ 3000 fb^{-1}. During LS3 a major upgrade of the detector, the so-called Phase-II upgrade, will also take place [52].

2.4.2 Technical and Physics motivations for the upgrade

In order to cope with these unprecedented LHC performances, the CMS experiment will need to improve the detector ability to select and measure the final states produced in the $p-p$ collisions. The aging of the detector material must also be addressed.

The expected number of overlapping events will be ~ 50 after the Phase-I upgrade and will increase to over a hundred of events at the HL-LHC. High values of pile-up (PU) increase the probability of fake rate in tracking, and reduce the energy resolution in both ECAL and HCAL. More granularity in these detectors is thus needed, to be able to distinguish the main event particles from the pile-up ones.

Another important factor to consider is the effect of radiation on the sensitive detector material, that causes a worsening of the detector performances. In some cases (e.g. the tracker) this effect is so critical that it does not allow any event reconstruction at all.

The CMS physics program of the next 20 years is challenging, and aims at answering fundamental questions still incomplete, with precise measurements of known SM processes, exploration of high-energy physics processes, and the study of very rare processes. The study of the Higgs boson discovered in 2012 will continue to be crucial: it will include precise measurements of the Higgs boson couplings, mass, J^{PC} and width, and the search for rare SM and BSM decays. Moreover, a search for new physics at higher mass scales will be possible with more statistics, probing or not the existence of SUSY particles, or other BSM physics.

2.4.3 Phase-I detector geometry upgrade

The Phase-I Upgrade, scheduled to be installed during the LS2 in 2019, consists of three sub-detector upgrades:

- replacement of the Pixel detector with a new detector with four-layers;
- improvement of the L1 Trigger system with higher granularity and additional processing capabilities;
- upgrade of the photo-detectors and electronics of the HCAL, to improve the measurement of jets and missing transverse energy.

The main reason for the Phase-I upgrade, in addition to the improvement of the detector performances is to face the high radiation damage to the detector material, especially in the regions very close to the beam pipe.

Pixel Tracker Upgrade

The current Pixel tracker will be replaced with a new high efficiency and low mass silicon pixel detector, with four barrel layers and three endcap disks, to provide 4 pixel hits in the full $|\eta| < 2.5$ acceptance region [55].

L1 Trigger Upgrade

The L1 Trigger was designed to operate a 100 kHz rate. With the increase in luminosity, energy and pile-up, a substantial increase in the trigger threshold will be required to remain within the 100 kHz limit. In order not to loose interesting physics events, the L1 Trigger will undergo an upgrade of the electronics of the calorimeter trigger, the muon trigger and the global trigger [53].

HCAL Upgrade

The HCAL upgrade is divided into two parts: one concerns the HF, the other the HB and HE calorimeters.
The HF currently uses Photomultiplier Tubes (PTM) to collect light from the absorber material and produce electronic signals that will be replaced with multi-anode tubes.
The HB and the HE instead use Hybrid Photo-diodes transducers (HPD) that will be replaced by Silicon Photo-multipliers (SiPM), achieving better HB and HE performances [54].

2.4.4 Proposals for the Phase-II detector geometry

The Phase-II upgrade, planned for the LS3 in 2023, is still under study. The primary goal of this upgrade is to face the very high luminosity conditions of HL-LHC, maintaining or improving the excellent detector performances achieved until 2012. The focus is to identify changes which are mandatory for the Phase-II conditions, and provide performance projections based on real data taken during Run1 and radiation doses studies for HL-LHC. Full Simulations[2] of the detector were thus made, to evaluate the performance of the detector with and without upgrades; the configurations simulated are:

- Phase-I detector (BX = 25 ns, PU 50, $\mathcal{L} = 1 \cdot 10^{34}$ cm^{-2}s^{-1}) without radiation aging, to establish a benchmark for the Phase-II detector performances. In the following this will be referred to as Scenario 1.

- Phase-I detector (BX = 25 ns, PU 140, $\mathcal{L} = 5 \cdot 10^{34}$ cm^{-2}s^{-1}) with the modeling of the effects of radiation damage to the detector after an integrated luminosity of 1000 fb^{-1}, in order to understand which parts of the detector will suffer the most from radiation. This configuration will be called Scenario 2.

- Phase-II detector (BX = 25 ns, PU 140, $\mathcal{L} = 5 \cdot 10^{34}$ cm^{-2}s^{-1}), to evaluate the Phase-II upgraded detector performances. The exact definition of this configuration (Scenario 3) is still ongoing.

[2]For details on the Full Simulation procedure, see Section 3.2.

Tracker

The Tracker will suffer from significant radiation aging and high pile-up levels and will have to be completely replaced [56]. To maintain the same tracking efficiency as in Phase-I, the granularity of both the pixel and outer trackers will be increased by a factor of four. The new tracker design will be capable of providing track information to the L1 Trigger [57], contributing to the background rejection. Moreover, the extension of the tracker acceptance up to $|\eta| \approx 4$ is under study to increase the η coverage.

Endcap Calorimeter

The electromagnetic and hadron endcap calorimeters will dramatically suffer from radiation damage [58], and thus will have to be entirely replaced with new detectors able to sustain higher radiation levels and provide higher granularity. Two projects are under study:

- an EE calorimeter with a new design (Shashlik [59]) followed by a re-built HE with radiation hard components;
- a High Granularity Calorimeter (HGC) [60] including both the ECAL and the HCAL endcap calorimeters, followed by a HE with reduced depth.

Muon endcaps

Most of the muon system will sustain the radiation damage and will be operational until the end of HL-LHC. To improve redundancy, the muon system will be equipped with four new detectors: two of them will be made of Gas Electron Multiplier (GEM) [61] chambers, for good position resolution, and the other two will be RPCs with lower granularity but very good time resolution to reject background [62]. Moreover, related to the tracker η extension, the implementation of another GEM station in the forward region is planned, to increase the muon detector coverage beyond $|\eta| = 2.4$, as shown in Fig. 2.16

Trigger and Read-Out electronics

To cope with the Phase-II pile-up conditions, the L1 Trigger acceptance rate will be increased to 500 kHz: this should allow to maintain almost the same trigger thresholds of the Phase-I trigger. This higher rate will be necessarily correlated to the upgrades of the read-out electronics, especially for the front-end electronics of EB, CSCs and DTs.

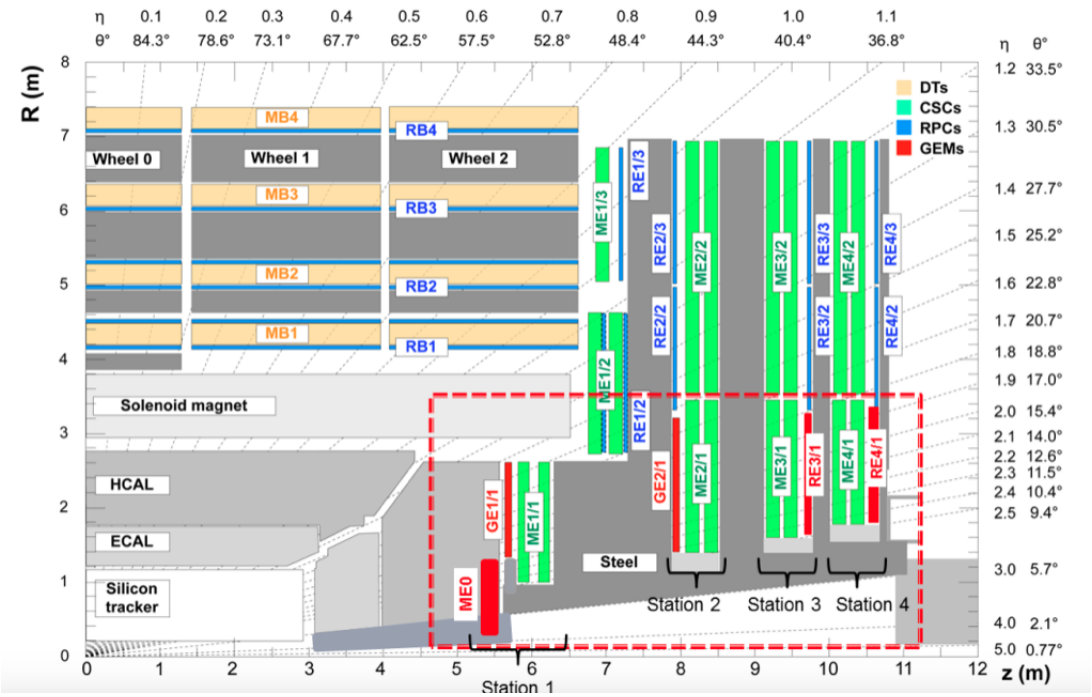

Figure 2.16: Longitudinal view of the CMS muon system with the new planned detectors framed by the red dash line.

2.5 The CMS Reconstruction and Identification

Starting from RAW data, the CMS software (CMSSW[3]) is used to process the events and to provide reconstructed physics objects (RECO data) [64]. This collection, or its Analysis Object data (AOD) subsets, is used in the majority of the CMS physics analyses.

2.5.1 Tracks and Primary Vertex

The algorithms implemented in CMS to reconstruct tracks use the hits of the charged particles in the silicon tracker to determine their helicoidal trajectories and measure their direction and momentum. The sequence of algorithms used in CMS is called Combinatorial Track Finder (CTF) [65] and is composed of three steps: track seeding, finding and fitting.

1. The Seeding step consists in searching for pairs of hits which can be candidates of charged tracks. The pixel information is used for better precision, except for the

[3]CMSSW is a collection of over thousand sub-packages (geometry, simulation, reconstruction, identification, etc) created to provide an extensive toolkit which is used in the analysis of the CMS data [63].

region $2 < |\eta| < 2.5$, where the information from the inner strips is added not to loose efficiency.

2. The Track Finding stage is based on the Kalman Filter pattern recognition approach [66]. Starting from the seeds, the track trajectory is extrapolated to the neighboring layers, and the compatible hits are added to the track. The Kalman Filter updates the track information with the added hits.

3. The Track Fitting process is obtained using again the Kalman Filter. It applies the least-squares fit method in two ways: firstly from the interaction region to the outer hits and secondly from the outside to the inside smoothing the trajectory and giving the best estimation of the track parameters.

An improvement to this procedure is obtained applying the track reconstruction in an iterative procedure. The CTF algorithm is thus applied three or four times, leading to a final collection of traks reconstructed with high precision.

Starting from a collection of well reconstructed tracks, the Primary Vertex is reconstructed using the transverse impact parameter and the longitudinal coordinate z. Then the vertices originating from the subsequent prompt decays, are referred to as Secondary Vertices.

2.5.2 The Particle-Flow Reconstruction

Thanks to the CMS granularity, especially in the tracker, and to the high bending power of the magnetic field, the Particle-Flow reconstruction (PF) can be implemented [67]: this is a reconstruction method which reconstructs the stable particles in the event combining all the sub-detectors information together, for optimal determination of their direction, energy and identification. PF particles are used to build jets (and tag τ's and b-jets), to compute the missing transverse energy, to determine the lepton isolation, etc. The PF method shows good performances, and is generally used in most of the physics analyses.

2.5.3 Muons

The muon reconstruction and identification [68] is based on the tracker and the muon system information, and additionally on the calorimeter energy deposit information.

Muon Reconstruction

In the CMS muon reconstruction procedure, tracks are built from the tracker and from the muon system independently. Different collections are produced:

- Stand-Alone Muons: Stand-Alone muons are reconstructed using the muon system information only;

- Global Muons (outside-in): starting from a Stand-Alone muon, a matching track is associated and then the total track is fitted.

- Tracker Muons (inside-out): all tracks are extrapolated to the muon system and if at least one matching muon segment is found, the corresponding track is considered a muon track. This method is more efficient than the Global Muon algorithm at low transverse momenta.

Thanks to the high granularity of the tracker and to the high efficiency of the muon system, almost 99% of the muons produced in a collision are reconstructed as Tracker or Global Muons. However, the reconstructed muons contain a significant part of misreconstructed charged hadrons. To get rid of them, the following identification requirement is applied.

Muon Identification

The typical identification selections produce different muon collections:

- Soft Muons: Tracker Muons with the additional requirement that a segment in the muon system is matched in both the x and y coordinates with the extrapolated muon track.

- Tight Muons: Global Muons with five additional requirements: normalized χ^2 of the Global Muon track fit < 10, at least one muon system hit is included in the final track fit, the track includes hits from at least two muon stations, there are more than 10 tracker hits, the transverse impact parameter is less than 2 mm.

- Particle-Flow Muons: they are identified applying three different selection criteria. The first one is the isolation: the reconstructed muon is considered isolated if in a cone of size $R = \sqrt{\phi^2 + \eta^2}$ centered on the muon, the sum of the transverse momenta of all the other particles except the muon is less than a certain threshold. Then two selections are applied, called pf-tight and pf-loose respectively, to the remaining muons. They both require a minimum number of hits in the tracker and the compatibility of the muon segment and energy calorimeter deposits with some template models based on simulations. However, the pf-loose requirements are more relaxed than the pf-tight ones and the template-based compatibility criteria is replaced by a matching requirement of the tracker tracks with muon tracks.

2.5.4 Electrons

The electron reconstruction and identification [69] is based on the Tracker and Electromagnetic Calorimeter information, searching for a match between the cluster deposit in the ECAL and a track in the Tracker.

Electron Reconstruction

The energy deposits in the ECAL are grouped in superclusters[4] which collect the energy of Bremsstrahlung photons emitted by electrons in the active detector material, and use them to find matching hits in the Tracker. This ECAL-driven process is particularly efficient for high transverse momentum electrons ($p_T > 10$ GeV), while for low p_T electrons it has to be complemented with a tracker-driven approach. Tracks are then reconstructed using appropriate algorithms and fitted with the Gaussian Sum Filter (GSF) [70]. Finally GSF tracks and associated ECAL superclusters allow the single electron reconstruction. The electron transverse momentum is found using the GSF information and is added to the Tracker one. The electron reconstruction efficiency is about 90% over all the ECAL acceptance. However, the sample of reconstructed electrons contain other particle misreconstructed as electrons. For this reason, an additional identification requirement has to be applied.

Electron Identification

The electron identification is based on the Boosted Decision Tree (BDT) multivariate technique [71]. The BDT uses three categories of variables: observables related to the purely calorimeter information, to the purely tracker information, and to the matching between the calorimeter and tracker information. The information is collected in a unique statistic test (variable). The training of the BDT is performed in three different η regions, and in two p_T intervals, to optimize the cut values in all the acceptance region.

2.5.5 Photons, Jets, Missing Transverse Energy

Photon candidates are reconstructed using the energy deposit information in the ECAL plus the information of the Tracker to identify the electrons from photon conversions. Energy corrections are applied to account for the interaction of particles with the detector material and for the shower containment. Photons are identified and reconstructed using the particle-flow algorithm which is efficient down to $E_\gamma \simeq 230$ MeV in the barrel and 600 MeV in the endcaps. Finally, in order to discriminate against pile-up photons, the isolation requirement can be applied.

[4] A supercluster is a group of cells of energy deposits in the ECAL, associated with specific algorithms.

The jets can be reconstructed in two different ways: using the standard jet reconstruction algorithm, which employs the information of the HCAL only, or using the particle-flow process. In the first method, only the calorimeter information is used. Cells in the HCAL are organized in towers: the energy of each tower is computed as the sum of the relative cells. The tower information is then used as input in the jets reconstruction algorithms. In CMS the *Iterative cone* [72], *Midpoint cone* [73] and the *Inclusive* k_T [74] jet algorithms have been developed and optimized. In the second method, jets are reconstructed using the particle-flow information, from all the sub-detectors. Since the ECAL and the tracker have higher p_T resolution, this technique leads to a more precise collection of jets.

The Missing Transverse Energy (E_T^{miss}) can be efficiently calculated using the PF-algorithm: the high granularity ECAL and HCAL, the redundancy of the muon system and the optimal tracker resolution allow to obtain precise E_T^{miss} measurements.

Chapter 3

Event Generation and Simulation

As explained in Chapter 2, final state events produced in the $p-p$ collisions are reconstructed through specific algorithms which transform RAW data into physics quantities (such as tracks, particle transverse momenta, etc). Physics analyses are then performed on these reconstructed objects.

The simulation of physics processes through Monte Carlo generators is crucial for high energy physics experiments. In general simulation is needed to:

- study the phenomenology of physics processes;
- measure and optimize detector performances;
- study new physics scenarios;
- optimize the analysis strategy;
- estimate corrections to data to take into account the detector response.

The simulation of a high energy physics event is performed in two steps: the event generation and the simulation of the detector response to the passage of particles through the detector materials. At this point simulated objects have the same format and structure of the real data one, and can undergo the same reconstruction process and analysis described before.

3.1 Event Generator

An event generator is a tool which performs the generation of high energy physics events. The generation is based on a series of physics models which allow the evolution from a few-body hard process to a complex multi-hadron final state, considering many different processes such as QCD and electroweak processes.

The event generation proceeds in three steps:

1. a parton from each one of the showers produced from two incoming particles (described with partonic distributions) starts the hard process, and the nature of the event is thus decided;

2. the generation of multi-particle interactions, final state radiations and beam remnants is performed, using both perturbative and non-perturbative models: a partonic structure of the jets and of the underlying event is produced;

3. the hadronisation of the partonic structure is performed, through fragmentation and unstable particle decays. This is achieved using a complex parametrization of non-perturbative models.

The result is a collection of generated final state particles which represents a physics process. The most commonly used event generators by the LHC experiments are PYTHIA [76, 77], MadGraph [78] and POWHEG [79].

3.2 Full Simulation

After the generation of the event, the interaction of the particles with the detector material has to be simulated. This task is usually referred to as *detector simulation* and is performed by GEANT4 [80, 81], which is a C++ object-oriented toolkit developed by a worldwide collaboration started in 1993. This tool, which accepts as input the output of event generators, is based on a rich set of physics models that allow a complete description of the energy loss and interaction of particles in matter, for a wide range of energies (from 250 eV to PeV). Some of the main features provided by GEANT4 are:

- description of the detector geometry and material, with the possibility to distinguish active from passive material;

- collection of physics processes that describe the interactions of particles in matter such as Bremmstrahlung, pair production, nuclear interactions, multiple scattering, photon conversion, etc;

- effect of the magnetic field on the tracking process;

- simulation of the electronic response and digitization of the simulated hits;

- management of pile-up effects;

- detector and particle trajectory visualization.

Thanks to the object-oriented structure, the user is able to add and modify modules, and to choose different implementations and approaches during the simulation.

At the end of this Full Simulation process, the simulated events are in the same format as the real data events. Therefore they can undergo the same Reconstruction process that transforms RAW events into RECO events, as described in Section 2.5.

3.3 Delphes Simulation

For the production of large background samples or for phenomenological studies, such as the comparison between several different configurations of a detector, a complete simulation is very demanding in terms of computing time and resources. For these purposes, Delphes [82, 83], which is a C++ modular framework for a parametrized and fast multipurpose detector simulation, has been developed by S. Ovyn, X. Rouby and V. Lemaître of the Center for Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, Belgium in 2010. The simulation takes into account a tracking system embedded in a magnetic field, electromagnetic and hadron calorimeters with their granularity, and a muon identification system. It allows the reconstruction of physics objects useful for the analysis, such as charged leptons, photons, jets and missing energy.

The Delphes tool takes as input the event generators output, in different file formats like ProMC [84], HepMC [85], LHEF [86], and produces an output ROOT [87, 88] file which contains the simulated and generated objects.

As it is a parametrized and fast simulation, it has some limitations: the detector geometry is assumed to be ideal, with no cracks nor dead materials and symmetric around the beam axis. Moreover, secondary interactions, energy loss due to Bremmstrahlung or multiple scattering and photon conversion are not simulated.

Delphes can be easily tuned to reproduce the performances of any existing multipurpose detector (such as CMS and ATLAS) or to design detectors which could be used in future colliders. The structure and the features of this framework are described in the next sections.

3.3.1 Structure of the Simulated Detector

As shown in Fig. 3.1, Delphes simulates the structure of a general purpose detector, composed of an inner tracker embedded in a magnetic field, electromagnetic and hadron calorimeters, and a muon system, designed concentrically and symmetrically around the beam axis.

Tracker

The first step of the simulation is the propagation of particles through the inner tracker volume. The tracker is embedded in a uniform magnetic field, parallel to the beam axis.

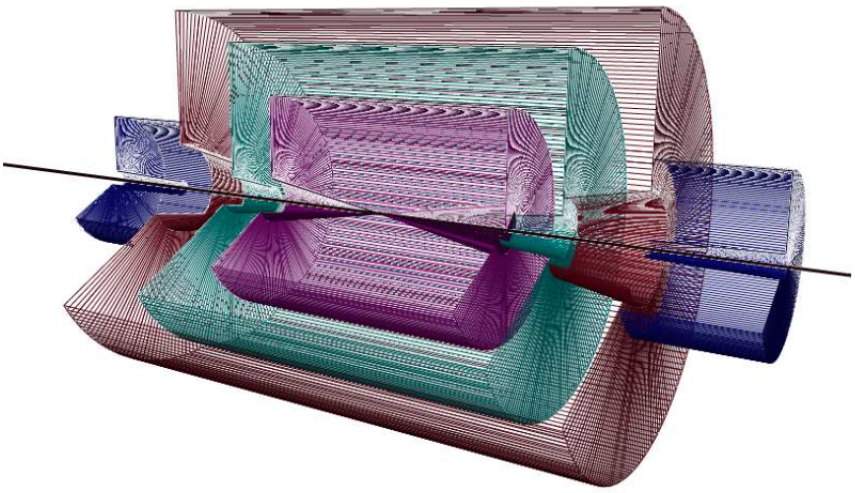

Figure 3.1: Layout of a generic detector geometry simulated by Delphes: from the inside-out the tracker (in purple), the calorimeters (in green) and the muon system (in brown) can be appreciated; finally the two forward calorimeters are shown in blue [82].

For this reason, charged particles follow a helicoidal trajectory, while neutral particles have a straight trajectory up to the calorimeters. As in real detectors, only charged particles have a user-defined probability to be reconstructed as tracks in the tracker volume; no smearing is applied to the track parameters, except for the module of the transverse momentum, which is smeared at this point of the propagation. The user can specify the energy and momentum resolutions, as well as the tracking reconstruction efficiency, as a function of the particle type, of the transverse momentum and of the pseudo-rapidity.

Calorimeters

For computational reasons, Delphes simulates the electromagnetic calorimeter (ECAL) and the hadron calorimeter (HCAL) with the same segmentation. The design is such that ECAL and HCAL are perfectly overlaid so that one particle reaches exactly one cell in the ECAL and one cell in the HCAL. The segmentation in the (η, ϕ) plane is shown in Fig. 3.2. No longitudinal segmentation is available. The size of the elementary unit of the calorimeters (*cell*) can be defined, but the segmentation in ϕ is set uniform, and the detector is assumed to be symmetric in ϕ and with respect to the $\eta = 0$ plane. The coordinate of the energy deposit (in the following labeled *tower*) is computed at the center of the cell.

The particle energy loss can be shared by the ECAL and HCAL according to pre-defined fractions (f_{ECAL} and f_{HCAL}) depending on the nature of the particle. By default,

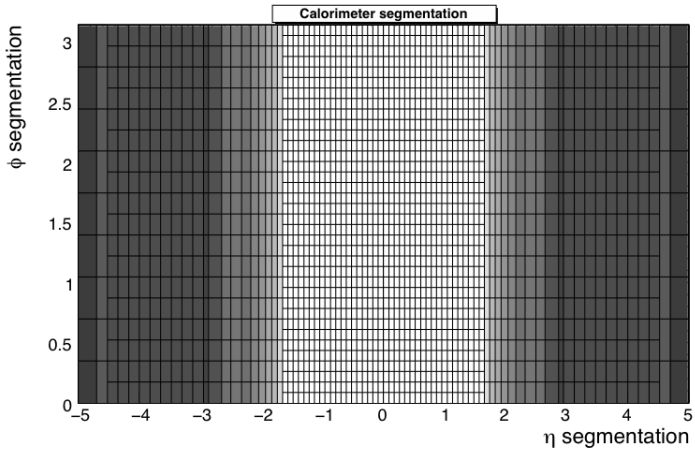

Figure 3.2: Default calorimeters segmentation in the (η, ϕ) plane.

electrons and photons leave all their energy in the ECAL ($f_{ECAL}^{e\pm,\gamma} = 1$, $f_{HCAL}^{e\pm,\gamma} = 0$), hadrons deposit all their energy in the HCAL ($f_{ECAL}^{h} = 0$, $f_{HCAL}^{h} = 1$), kaons and Λs share their energy between the two calorimeters ($f_{ECAL}^{K,\Lambda} = 0.3$, $f_{HCAL}^{K,\Lambda} = 0.7$), while muons, neutrinos and neutralinos do not loose energy in the calorimeters; however, these values can be modified to specify the energy loss fraction for every particle.

The resolutions of the electromagnetic and hadron calorimeters are independently parametrized as a function of the particle energy and pseudo-rapidity:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S(\eta)}{\sqrt{E}}\right)^2 + \left(\frac{N(\eta)}{E}\right)^2 + C(\eta)^2, \tag{3.1}$$

where S, N and C are respectively the stochastic, noise and constant terms. The electromagnetic and hadronic energy deposits are independently smeared by a log-normal distribution, and the final tower energy is computed as:

$$E_{Tower} = \sum_{particles} \ln \mathcal{N}(f_{ECAL} \cdot E, \sigma_{ECAL}(E,\eta)) + \ln \mathcal{N}(f_{HCAL} \cdot E, \sigma_{HCAL}(E,\eta)). \tag{3.2}$$

where σ_{ECAL} and σ_{HCAL} are the resolutions for ECAL and HCAL respectively, and are computed using Eq. 3.1. The energy of each particle is concentrated in one single tower and all the particles that reach that tower are used to calculate the energy of this particular tower.

Particle-Flow

Delphes implements a simplified particle-flow approach (see Section 2.5.2) based on the tracker and the calorimeters information. The particle momenta are always computed

using only the tracker information (even if in real detectors the tracker resolution is better than the calorimeter one only if the particle energy is lower than a certain threshold).

In Delphes the algorithm produces particle-flow tracks and particle-flow towers, starting from:

- E_{ECAL} and E_{HCAL}: the total energy deposit in ECAL and HCAL respectively;
- $E_{ECAL,trk}$ and $E_{HCAL,trk}$: the total energy detected in ECAL and HCAL that is associated to charged particles reconstructed in the tracker;

defining:

$$\Delta_{ECAL} = E_{ECAL} - E_{ECAL,trk} \qquad \Delta_{HCAL} = E_{HCAL} - E_{HCAL,trk} \tag{3.3}$$

and computing:

$$E_{Tower}^{eflow} = \max(0, \Delta_{ECAL}) + \max(0, \Delta_{HCAL}). \tag{3.4}$$

At this point,

- PF tracks are built from particles reconstructed in the tracker, estimated with a good resolution;
- a PF tower is created with energy E_{Tower}^{eflow} if $E_{Tower}^{eflow} > 0$: it contains the information from neutral particles, charged particles with no corresponding reconstructed track and additional energy deposits, with degraded resolution.

3.3.2 Object Reconstruction

The object reconstruction and identification produces as output a collection of physics quantities with transverse momentum, energy, and other related quantities.

Muons, electrons and photons

Muons ($\mu^{\pm}$) and electrons ($e^{\pm}$) generated in the interaction are reconstructed with a user-defined probability (as a function of the transverse momentum and of the pseudo-rapidity), only if they are inside the tracker acceptance and have a transverse momentum higher than a certain threshold. The reconstructed momentum is the result of a Gaussian smearing of the generated momentum: the resolution can be defined by the user as a function of the transverse momentum and the pseudo-rapidity.
For electrons, the information from the tracker is combined with the one from the calorimeter: at low energies the tracker resolution is more precise while at high energies the calorimeter energy resolution dominates.
Photons are reconstructed using only the information from the electromagnetic calorimeter. Two important approximations are made: photon conversion is neglected and electrons without an associated track in the tracker volume are reconstructed as photons.

The energy resolution is calculated using Eq. 3.1.

Isolation

Many important physics processes yield final states with isolated particles, that is not surrounded by other particles. In Delphes, the isolation is computed considering a cone of radius R around the particle and computing the quantity:

$$I(P) = \frac{\sum_{i \neq P} p_T(i)}{p_T(P)}, \tag{3.5}$$

where the denominator is the transverse momentum of the particle P, and the numerator is the sum of the transverse momentum above a p_T^{min} threshold of all particles inside the cone built around the particle P. If this quantity is higher than a certain threshold I_{min}, the particle is considered not to be isolated. On the contrary, if $I(P) < I_{min}$ the lepton is isolated. All the parameters that enter in this calculation, such as R, p_T^{min}, and I_{min} are user-defined.

As previously mentioned, every experiment has the opportunity to customize the Delphes simulation (definitions of modules, parameters, etc.) in order to obtain a coherent description of its experiment detector. In the CMS Delphes version, the definition of the isolation variable in Eq. 3.5 has been modified, as will be explained in more detail in Section 3.3.3

Jets

In Delphes jets are reconstructed using jet clustering algorithms and parameters included in the FastJet package [75], the same used in CMSSW [63], which is integrated in Delphes. The most common cone and recombination algorithms such as *CDF Jet Clusters* [72], *CDF MidPoint* [73], *longitudinal invariant* k_t [74], *Cambridge/Aachen* [89] and *Anti* k_t [90] algorithms are available. These methods can be used to produce jets starting from different collections of objects, depending on the user needs, and leading to:

- *Generated jets*, which are clustered starting from the generator level particles;
- *Calorimeter jets*, built using calorimeter towers;
- *Particle-flow jets*, obtained from the clustering of particle-flow tracks and towers.

The reconstruction of jets that come from a b quark or from a τ lepton proceeds in two steps:

- a jet is labeled as a b-jet or τ-jet if a b or τ is found within the distance from the jet axis $\Delta R = \sqrt{(\eta^{jet} - \eta^{b,\tau})^2 + (\phi^{jet} - \phi^{b,\tau})^2} < \Delta R^{min}$;
- the b or τ jet is reconstructed depending on a user-defined efficiency.

For b and τ jets, unlike for charged leptons and photons, a mis-tagging efficiency (the probability that a particle which is not a b nor a τ is reconstructed as a b or a τ) is implemented and can be modified by the user.

Missing transverse energy and scalar transverse energy

In the Delphes simulation the missing transverse energy and the scalar transverse energy are computed respectively as follows:

$$\vec{E}_T^{miss} = -\sum_i \vec{p}_T(i), \qquad H_T = \sum_i |\vec{p}_T(i)|, \tag{3.6}$$

where the index i runs over the selected input collections, such as *Calorimeter Towers*, *Generator-level particles* or *PF particles*. These variables are useful to take into account neutrinos or other neutral particles whose energy loss can only be indirectly calculated.

3.3.3 High-level Corrections

The collections of reconstructed objects obtained with the methods described above need to undergo some high-level corrections because of the pile-up contamination and the non-uniformity in the energy response.

Energy Scale correction

The Energy Scale Correction is a procedure that compensates the mismatching of the reconstructed objects particle momenta compared to their generator-level momenta. In Delphes this correction is applied only to jets, as a function of the reconstructed jet pseudo-rapidity and transverse momentum.

Pile-up subtraction

For simplicity, in Delphes the pile-up subtraction is only applied to jets and to the isolation. The algorithm is composed of two steps:

- thanks to the combined action of vertexing and tracking reconstruction, tracks that originate at a distance from the primary vertex $|z| > \delta Z_{vtx}$ (where δZ_{vtx} is the spatial vertex resolution of the detector) can be identified as belonging to pile-up

events and then removed from the event. If the particle-flow procedure is used, the particle-flow tracks recognized as pile-up tracks are removed from the particles used in the jet clustering and isolation calculation;

- the residual contributions, due to particles that are too close to the vertex, charged particles without an associated reconstructed track and neutral particles, are subtracted using the Jet Area method. Thanks to the FastJet [75] package, this procedure allows the calculation of the average pile-up contamination density ρ, used to correct the observables sensitive to the residual contamination:

$$p_{jet} \rightarrow p_{jet} - \rho \cdot A_{jet} \tag{3.7}$$
$$I(P) \rightarrow I(P) - I_C \tag{3.8}$$

where A_{jet} is the jet area estimated using the FastJet package, R is the radius of the cone built around the particle to compute the isolation variable, and the correction I_C is defined as:

$$I_C = \frac{\rho \cdot \pi R^2}{p_T(P)}. \tag{3.9}$$

Isolation and pile-up correction in the CMS Delphes version

In the Delphes version customized for the CMS experiment, the definition of the isolation variable was modified as follows:

$$I(P) = \frac{\sum\limits_{i \neq P} p_T^{charged}(i) + \max\left[\sum\limits_{i \neq P} p_T^{neutral}(i) - I_C,\ 0\right]}{p_T(P)} \tag{3.10}$$

where $p_T^{\min}$ is the p_T threshold of particles accepted in the cone of radius R. This means that the pile-up correction defined in Eq. 3.9 is applied only to neutral particles, since the charged part of the sum contains particles that come from the primary vertex only and is therefore already corrected for the pile-up contamination.

3.3.4 Validation

The Delphes simulation framework has been validated comparing the results of the simulation to the results of real experiments, such as the two multipurpose detectors operating at the LHC, ATLAS and CMS.

Only a brief summary of the validation results is reported here. All the Monte Carlo samples used in this procedure were produced using the MadGraph5 event generator, and then hadronized with PYTHIA6; then the Delphes simulation was applied, using the specific configuration of the two detectors ATLAS and CMS. In Fig. 3.3, the muon

transverse momentum resolution is reported as a function of η for Delphes and CMS Full Simulation, and of p_T for Delphes and ATLAS Full Simulation. In Fig. 3.4, the electron and photon energy resolution is shown as a function of the energy for Delphes and CMS Full Simulation. For all the distributions there is a good agreement between the parametrized response of the Delphes simulated detector and the real experiment ones. For this reason Delphes can be considered a reliable tool to produce a fast simulation of a multipurpose detector, useful in phenomenological studies and quick productions of large background samples.

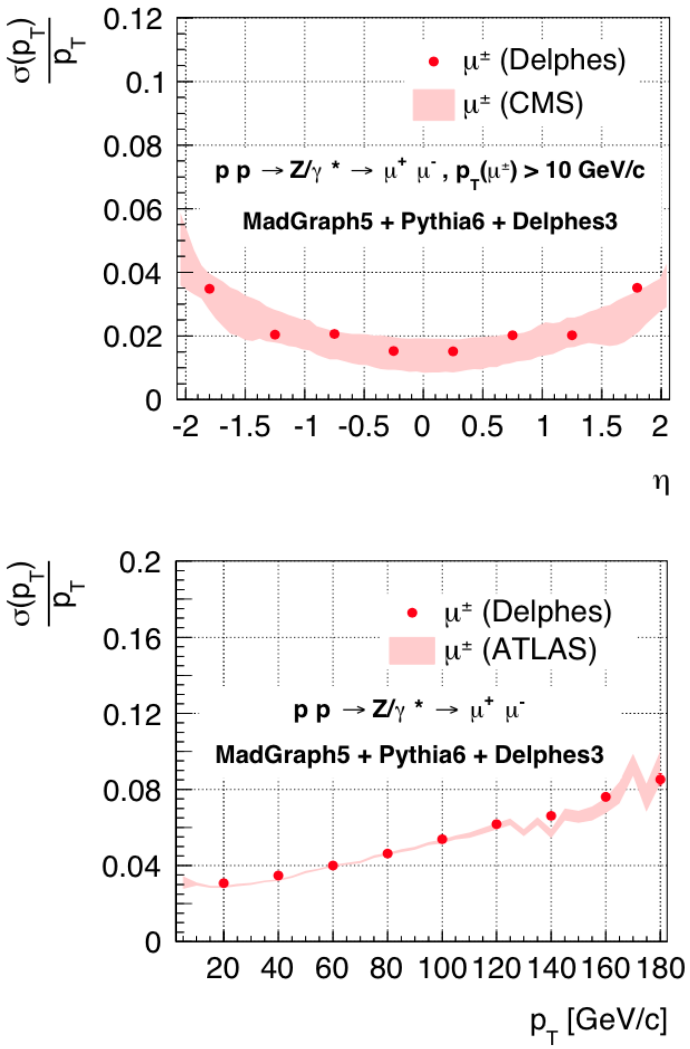

Figure 3.3: Left: muon p_T resolution as function of η for Delphes and CMS; Right: muon p_T resolution as function of p_T for Delphes and ATLAS [83].

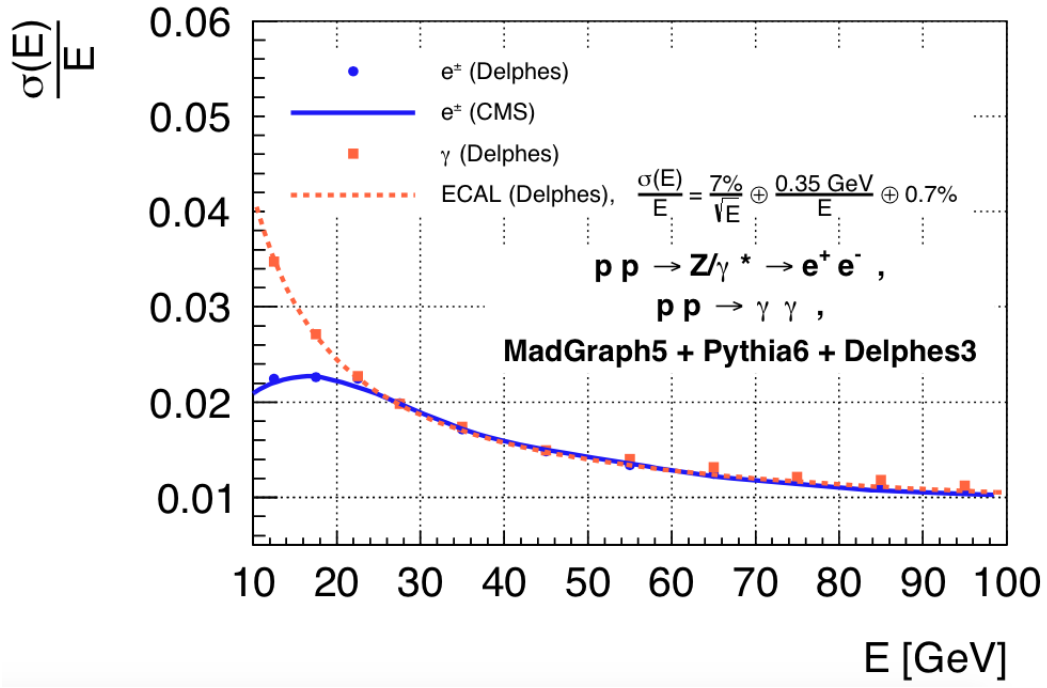

Figure 3.4: Electron and photon energy resolution as function of the energy for Delphes and CMS.

3.3.5 Software Implementation

Data-Flow

In Fig. 3.5 a simplified work flow chart is reported. As previously mentioned, Delphes takes as input the most common event generators output files in various formats, which are decoded by a reader module. Then, the number of pile-up events is randomly extracted from a user defined Poisson distribution; the pile-up events are randomly extracted from a pre-generated file containing only low-Q^2 QCD interactions, and randomly placed along the beam axis according to a user-defined longitudinal spread and overlaid to the hard scattering event; stable particles are then propagated from the tracker to the calorimeters, within the magnetic field. At this point, the objects are processed by a series of modules: muons, electrons, photons and particle-flow particles are reconstructed with user-defined efficiencies and resolutions, and pile-up is subtracted to these collections of objects. Then the isolation variable is computed and the reconstruction of jets is performed. Finally, duplicates of reconstructed objects are removed, and the resulting collections of global event quantities and physics objects are stored in a ROOT Tree file format, together with the initial Monte Carlo generated object collections.

Technical Performance

As Delphes was designed to be a fast simulation, the resources needed for the simulation have to be reduced as much as possible. In Fig. 3.6 the memory usage as a function of time is shown: after an initial increase due to memory allocation, it remains constant and always under a few hundreds of MegaBytes during the full process. In Fig. 3.7 the relative CPU time used by the different Delphes modules to process a $t\bar{t} + jets$ sample is reported: the FastJet package, which performs the jet reconstruction, uses the majority of the CPU.

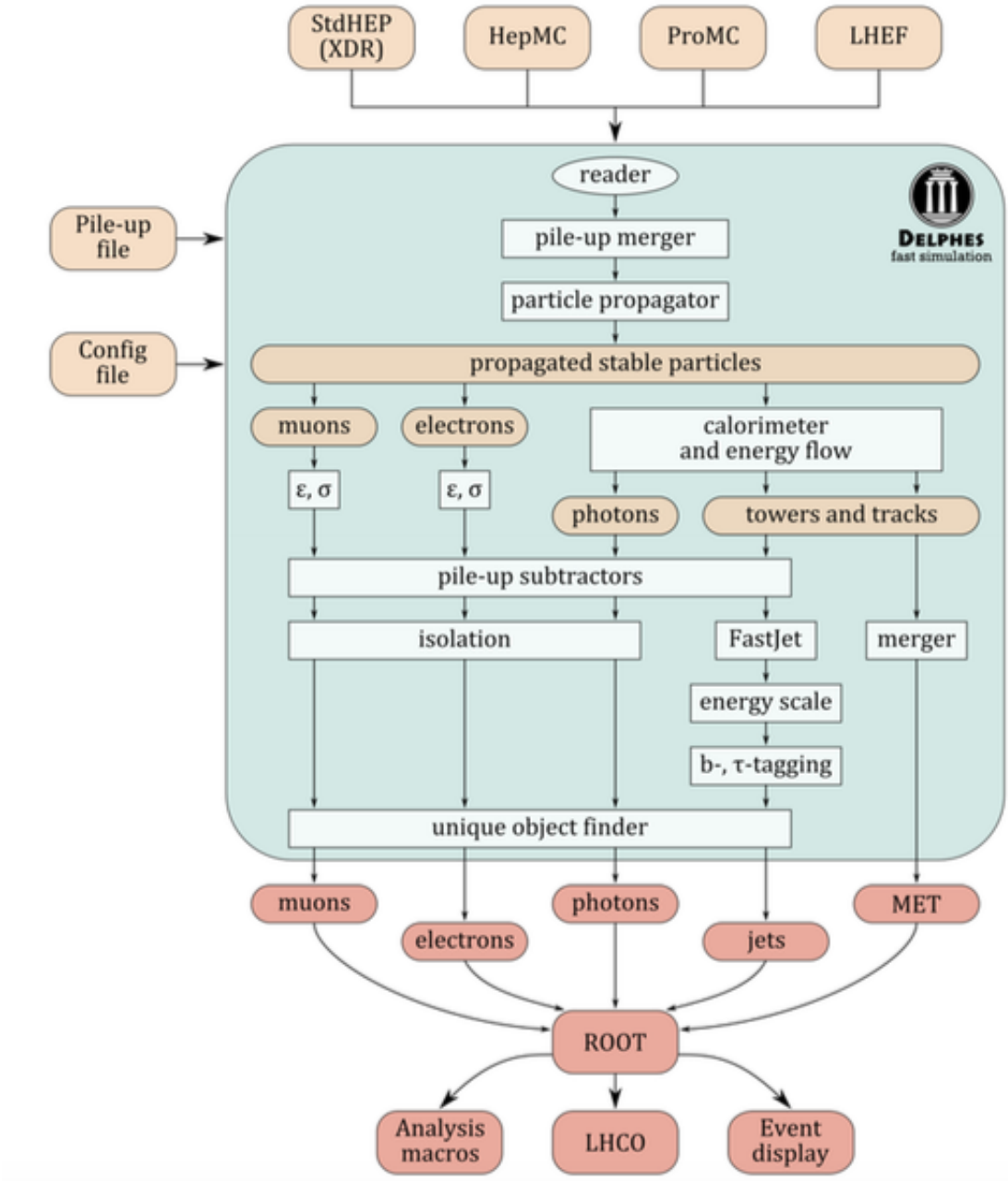

Figure 3.5: Simplified Data-Flow chart of the Delphes simulation.

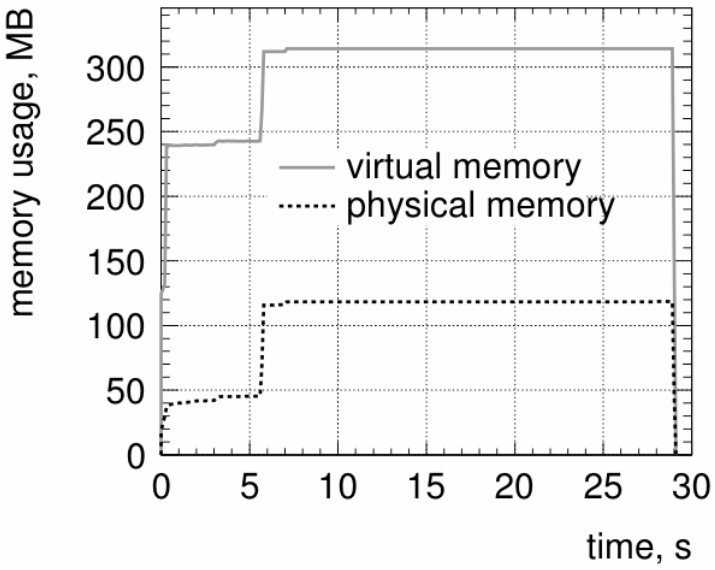

Figure 3.6: Memory usage as function of processing time.

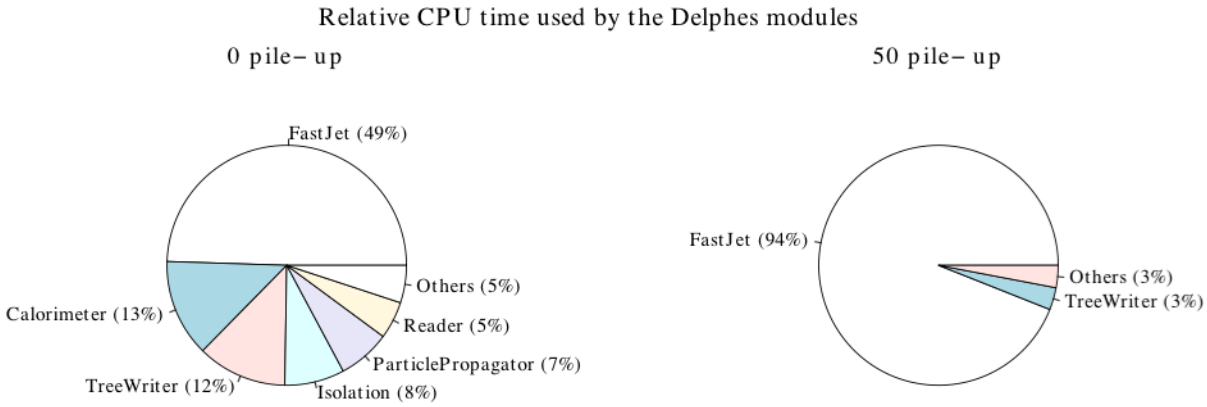

Figure 3.7: Relative CPU time usage for different Delphes modules to process a $t\bar{t}+$jets sample, for two different pile-up conditions.

Chapter 4

The $H \to ZZ^* \to 4l$ analysis study for the upgrade

As already mentioned in Section 2.4, the Phase-II detector upgrade is being designed to cope with the HL-LHC beam conditions.

In this thesis, the baseline analysis used for the Higgs boson discovery and the measurement of its properties [32, 91, 92] in the $H \to ZZ^* \to 4l$ channel was implemented within the simulation of the Phase-II detector upgrade [93].

4.1 Signal and Background Processes

In this analysis, the signal process $H \to ZZ^* \to 4l$ was studied in the three following decay modes:

- $H \to ZZ^* \to 4\mu$,
- $H \to ZZ^* \to 4e$,
- $H \to ZZ^* \to 2e2\mu$.

The dominant background for the $H \to ZZ^* \to 4l$ is composed of the $ZZ \to 4l$ process, where the ZZ intermediate state is produced via $q\bar{q}$ annihilation and gluon fusion. This background is called irreducible because four real leptons coming from the ZZ decay produce the same final state as the signal. The other main contribution to the background arises from $Z+jets$, $WZ+jets$ and $t\bar{t}$ processes. This is called reducible background and the events are composed of real leptons from W and Z decays plus $jets$ mis-identified as leptons.

The signal and irreducible background events were produced using the POWHEG NLO

Simulation tools	Scenario 1 Phase-I PU 50	Scenario 2 Phase-I aged PU140	Scenario 3 Phase-II upgrade PU140
FullSim	✓	✓	✗
Delphes	✓	✗	✓

Table 4.1: Signal and irreducible background samples available for the three scenarios studied in the analysis, for the two simulation tools used.

generator [79] and PYTHIA6 [76], respectively for the event generation and fragmentation. The reducible background samples were produced using Madgraph for event generation and PYTHIA for photon showering.
Only the gluon-gluon fusion production process was considered in the generation of the signal.

The Full Simulation (FullSim) and Delphes Simulation (Delphes) were then applied to the event generator output for the detector simulation and event reconstruction.

4.2 Simulation Samples

In Table 4.1 a summary of the available samples is reported for each detector geometry configuration. Since the CMS Full Simulation samples with a description of the Phase-II upgraded detector (Scenario 3) were not ready yet, the parametrized Delphes simulation was used to simulate a possible Phase-II detector upgrade configuration. In addition, the reduced computing power needed by Dephes makes it particularly useful for the production of large background samples. To demonstrate the reliability of this tool, a set of Delphes samples with a Phase-I detector (Scenario 1) parametrization was produced. The validation of Delphes Phase-I configuration was performed comparing the results with the FullSim Phase-I . Then a reasonable parametrization of the Phase-II detector was implemented in Delphes. In addition to Scenario 3, a Phase-I detector geometry without optimizations was studied (Scenario 2). Full Simulation samples produced considering the detector aging process after an integrated luminosity of 1000 fb^{-1} were used. Finally a comparison between the three Scenarios was made, to study the benefits and disadvantages of these detector configurations.

4.3 Analysis Strategy

The same analysis used in the 2012 search for the Higss boson was applied [91].

4.3.1 Signal and Irreducible Background

For the event selection, muons were accepted with $p_T > 5$ GeV and pseudo-rapidity $|\eta| < 2.4$ for Scenario 1 and 2, and $|\eta| < 3.0$ for Scenario 3; electrons were accepted with a $p_T > 7$ GeV and pseudo-rapidity $|\eta| < 2.5$ for the first two Scenarios and $|\eta| < 3.0$ for Scenario 3. The sequence of consecutive steps of the event selection $H \to ZZ^* \to 4l$ is reported below.

1. At least four leptons (4μ, $4e$ or $2e2\mu$) were required to be reconstructed inside the acceptance of the detector.

2. At least four reconstructed leptons were required to be identified as muons or electrons. In Delphes, reconstruction and identification are embedded in the same parametrization formula, and therefore cannot be separated in the analysis. For this reason the first Delphes requirement is to have at least four leptons (4μ, $4e$ or $2e2\mu$) inside the detector acceptance region.

3. Leptons selected as described above were requested to be isolated: the Isolation requirement is essential for the reducible background rejection, in which leptons are mostly produced in jets and are therefore not isolated. However this selection criteria does not affect the irreducible background, that presents the same final state as the signal: four isolated leptons coming from the two Z bosons.

4. To reconstruct the first Z candidate, a pair of leptons with the same flavor but with opposite charge was requested. Among all the pairs, the one with the invariant mass closest to the nominal Z mass was selected and labeled Z_1.

5. The mass of the reconstructed Z_1 was requested to be in the mass interval 40 GeV $< M(Z_1) <$ 120 GeV.

6. To reconstruct the second Z candidate, a second pair of same flavor but opposite charge leptons was requested among the remaining possible pairs. The pair that presented the higher transverse momentum was then selected and labeled Z_2.

7. The reconstructed Z_2 was required to have a mass lower than 120 GeV.

8. Among the selected four leptons forming the two Z candidates, at least one was required to have $p_T > 20$ GeV and another to have $p_T > 10$ GeV. These thresholds were chosen in order to have a trigger efficiency close to 100% (plateau).

9. Any opposite charged lepton pair, irrespectively of flavor, was required to satisfy $m_{l^+l^-} > 4$ GeV. This selection criteria reduced the background due to leptons from hadron decays in jet fragmentation and from the decay of low-mass hadron resonances.

10. A lower bound on the Z_2 mass was set, in order to optimize the background rejection: $M(Z_2) > 12$ GeV.

11. Finally, the four-leptons invariant mass was requested to be larger than 100 GeV.

The selection criteria described above were implemented within both the Full Simulation and the Delphes Simulation, following the same sequence. The two analyses were then compared at each step of the selection process.

4.3.2 Reducible Background

As explained in Section 4.1, the reducible background for this analysis is composed by Z and WZ processes with associated $jets$, and $t\bar{t}$ events. The $Z+jets$ gives the dominant contribution due to its large production rate. These events contaminate the signal region (i.e. defined by the baseline analysis) since heavy-flavor quark jets containing non-prompt leptons and light-flavor quark jets mis-identified as leptons produce the same final state as the signal. A small fraction also comes from real leptons originating from photon conversions. These fakeable lepton objects are collectively referred to as "fake leptons". In Delphes, which was the simulation framework used to study this background, only the real leptons part was simulated: this leads to a slight underestimation of the reducible background. More work is ongoing to evolve the existing Delphes framework to perform a complete and precise estimation of these processes.

The $Z + jets$ samples contain $Z + b\bar{b}$ and $Z + light\ jets$ events to estimate the contribution of fake leptons from heavy flavor quark decay.

The same method used in the baseline analysis [91] was used to estimate the background in these studies for the upgrade: the object fake rate was firstly measured, and then used to estimate the yield of $Z + jets$ events into the signal region.

The lepton fake rate f_l is defined as the probability that one loosely selected lepton (i.e. a jet in $Z + jets$) passed the lepton selection used in the analysis. The term "loosely" refers to the process of relaxing some selection criteria in order to apply a less strict selection. Since in Delphes the identification (ID) requirement cannot be separated from the reconstruction one, only the Isolation cut was relaxed. In addition, the Impact Parameter (IP)[1] variable was used to achieve the same level of fake rate as that obtained in the 2012 analysis and the optimized cut value was found to be 2.2. The fake rate was then derived from the sample in bins of transverse momentum and pseudo-rapidity:

$$f_l = f(p_T, \eta) \tag{4.1}$$

In Fig. 4.1 the measured fake rate is shown for a Phase-I detector, with an average pile-up of 50 and 140 events.

[1]The impact parameter is defined as the distance of the particle trajectory from the interaction vertex in the point where the distance is smallest.

A control region (C) is then defined, selecting events containing two real leptons from the Z and two additional objects failing the Isolation or IP requirement. The number of background events contributing to the signal region was then estimated as:

$$N_B = N_C F_1 F_2, \tag{4.2}$$

where N_C is the number of events in the control region, and $F_i = \frac{f_i}{1-f_i}$, where i runs over the additional two objects. Table 4.2 shows the expected number of events in the signal regions from the reducible background processes, both for the PU 50 and PU 140 scenario at 14 TeV. The first error is the statistical uncertainty. The second error is a systematic uncertainty due to the statistical uncertainty on the fake rates.

Baseline	4e	4μ	$2e2\mu$	$2\mu2e$
PU 50	30.6 ± 3.4 ± 15.3	56.5 ± 6.6 ± 28.3	60.2 ± 7.4 ± 30.1	55.7 ± 6.1 ± 27.8
PU 140	143.9 ± 18.8 ± 72.0	150.1 ± 17.5 ± 75.0	194.7 ± 21.7 ± 97.5	170.8 ± 20.2 ± 85.4

Table 4.2: Contribution of reducible background processes in the signal region ($m_{4l} >$ 100 GeV). The predictions correspond to the scenario with number of pileup events equal to 50 and 140 on average, at 14 TeV with an integrated luminosity of 3000 fb^{-1}. The first error denotes the uncertainty due to the limited statistics in the control regions. The second error is a systematic uncertainty due to the statistical uncertainty on the fake rates; it was assumed to be about 50 %.

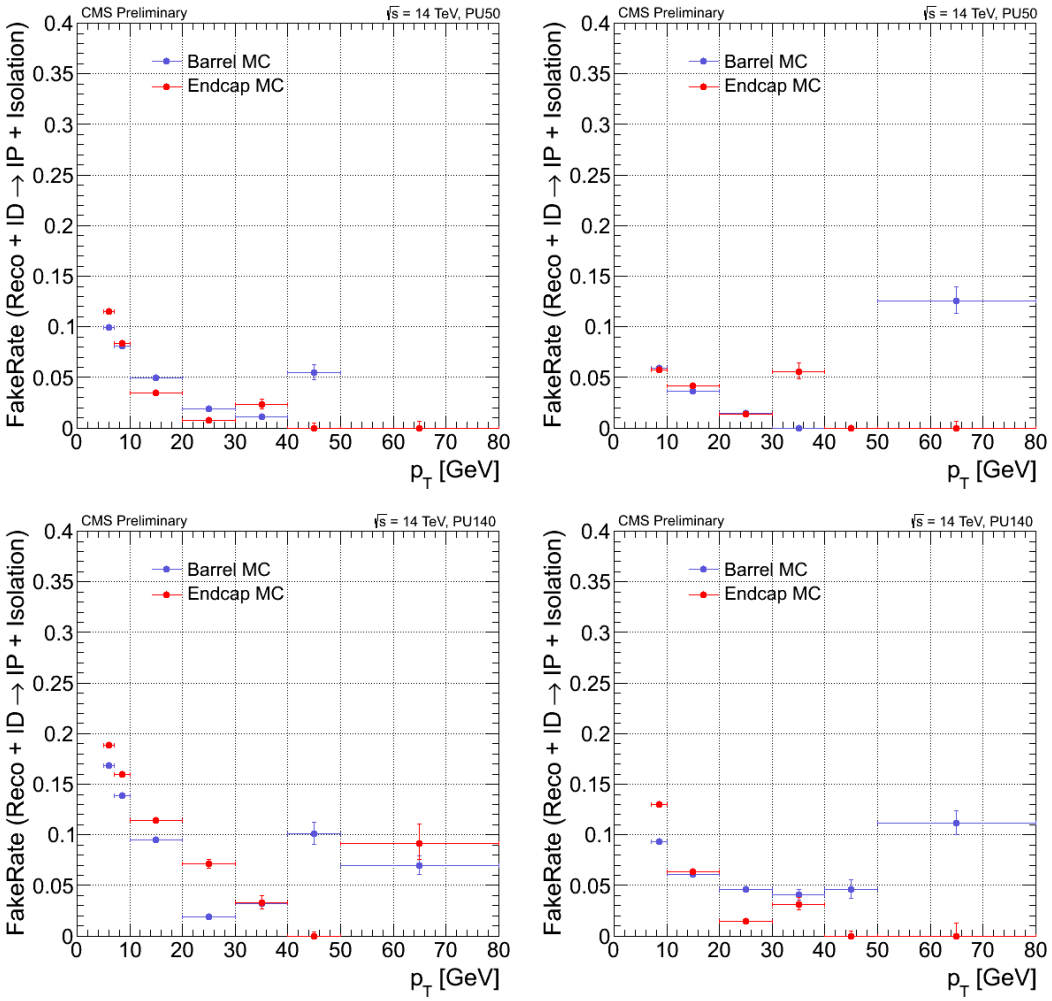

Figure 4.1: Measured fake rates for muons (left) and electrons (right) using $Z(ll) + \mu$ (left) and $Z(ll) + e$ (right) events respectively, at $\sqrt{s} = 14$ TeV. Plots are obtained with the simulation of the events in the Phase-I and Phase-II upgraded scenarios, respectively expecting 50 (top) and 140 (bottom) pile-up events on average. The blue (red) symbols correspond to muons within $\mid \eta \mid < 1.2$ ($\mid \eta \mid > 1.2$) and electrons within $\mid \eta \mid < 1.479$ ($\mid \eta \mid > 1.479$.

4.4 Validation of the Delphes Simulation for the Phase-I Detector

In order to use the Delphes Simulation for the Phase-II detector performances predictions, a complete validation with respect to the Full Simulation was needed. Extensive

Samples	Scenario 1 FullSim N evts. $\times 10^3$	Scenario 1 Delphes N evts. $\times 10^3$
$H \to ZZ^* \to 4\mu$	94	100
$H \to ZZ^* \to 4e$	94	10
$H \to ZZ^* \to 2e2\mu$	95	100
$ZZ \to 4\mu$	58	61
$ZZ \to 4e$	99	100
$ZZ \to 2e2\mu$	98	100

Table 4.3: Samples used for Delphes validation.

comparison and tunings between the two simulations were therefore performed using the Phase-I detector geometry samples, available for both simulation frameworks (for more information see Section 2.4.3 and Table 4.1). The distributions of relevant physics quantities such as reconstruction efficiencies, resolution, isolation, etc, were compared, tuned and parametrized in Delphes. Each step of the selection process was then tested. In Table 4.3 the samples used for the Delphes validation are listed.

In the figures shown in the following, FullSim samples are reported with a blue solid line while the Delphes Simulation is superimposed using red dots.

4.4.1 Efficiency and Resolution Distributions

The lepton reconstruction efficiency was computed as the ratio between the number of reconstructed leptons and the number of generated leptons in bins of p_T or η:

$$\varepsilon = \frac{\text{\# of reconstructed leptons}}{\text{\# of generated leptons}} \tag{4.3}$$

The denominator was calculated as the number of all generated $\mu^{\pm}$, $e^{\pm}$ in the final state, inside the η acceptance region. The numerator was calculated as the number of reconstructed leptons inside the η acceptance region. The matching algorithm that finds the reconstructed leptons associated to the generated leptons is different for the two simulations. In Delphes, a specific characteristic of the framework was used: reconstructed particles contain a "reference" to their associated generated particle. For every lepton of the denominator, a loop on all the reconstructed leptons was made searching for the one with the reference to its generated particle. Instead, in the FullSim, a matching in the $(\eta,\ \phi)$ plane was performed. The distance between reconstructed and generated leptons was defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{4.4}$$

If this distance satisfies $\Delta R < \Delta R_{\text{threshold}} = 0.4$, the two objects are considered as matched, and the reconstructed lepton is accepted. To avoid mismatching, the lepton is rejected if too close to another reconstructed one.
The Delphes efficiency was parametrized in order to match the FullSim efficiency in the best possible way. In Fig. 4.3 and Fig. 4.4 the muon and electron reconstruction efficiencies in $H \to 4l$ events are shown separately as functions of the lepton transverse momentum and of the lepton pseudo-rapidity.

In Fig. 4.5 the single lepton transverse momentum resolution in $H \to 4l$ events is shown. The p_T resolution σ_{p_T} is defined as the width of the distribution of the variable:

$$\frac{p_T^{\text{reco}} - p_T^{\text{gen}}}{p_T^{\text{gen}}} \tag{4.5}$$

The numerator contains the difference between the generated and reconstructed lepton transverse momentum, while the denominator contains the generated lepton transverse momentum. For the muons, an excellent agreement between FullSim and Delphes results is observed, while for electrons, the Delphes parametrization gives a slightly worse resolution than the FullSim one.

4.4.2 Kinematic Discriminant and Isolation Distributions

The lepton Isolation variable and the Kinematic Discriminant (K_D) are two key quantities used to discriminate signal from background.

The Isolation variable is useful to reject the reducible background. It takes into account the energy of charged and neutral particles within a cone around the lepton in the (η, ϕ) plane. Since the irreducible background is characterized by four isolated leptons in the final state, as the signal, the Isolation distribution does not significantly differ from the signal one. This can be seen in Fig. 4.6, where the Isolation variable distributions are shown, both for muons and for electrons from $H \to 4l$ events. Instead, a significant difference in the Isolation distribution can be seen when comparing the irreducible and the reducible backgrounds[2], as shown in Fig. 4.7. In the analysis, the lepton Isolation variable was requested to be smaller than 0.4, as in the 2012 analysis.

The Kinematic Discriminant is a variable which takes into account the kinematics of the Higgs boson decay to discriminate the signal from the irreducible background. This variable can be used in combination with the four-lepton invariant mass to estimate the significance of the signal hypothesis. Fig. 4.2 shows the production and decay of a Higgs boson in the process $gg \to H \to ZZ^* \to 4l$. Since the Higgs boson is spin-less, its

[2]The Isolation distribution of the irreducible background is similar to the signal one but provides more statistics.

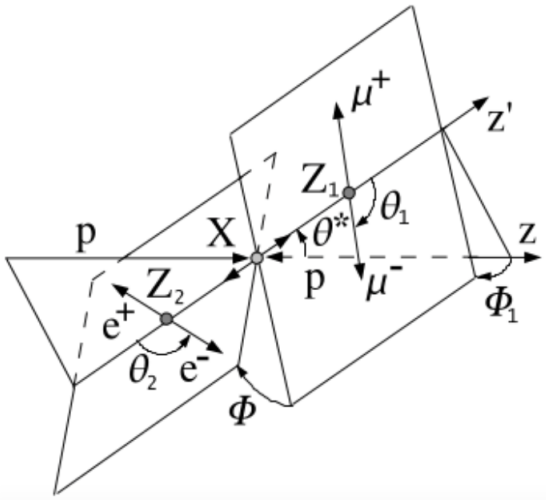

Figure 4.2: Illustration of a particle (X) production and decay in a $gg \to X \to ZZ^* \to 4l$ process.

angular distribution does not depend on the production mechanism. The kinematics of the $H \to 4l$ process, at a given four-lepton invariant mass, in their center-of-mass frame, is fully described by 7 parameters:

- θ^* and Φ_1: production angles in the Higgs rest frame;
- θ_1, θ_2, Φ: decay angles in the Z_1, Z_2 and Higgs rest frame respectively;
- m_{Z_1}, m_{Z_2}: the Z_1 and Z_2 boson masses.

A Matrix Element Likelihood Approach (MELA) [91] is used to construct a kinematic discriminant, function of these 7 parameters:

$$K_D = f(\theta^*,\ \Phi_1,\ \theta_1,\ \theta_2,\ \Phi,\ m_{Z_1},\ m_{Z_2})$$

created through the evaluation of the probability for the event to come from the signal or from the background:

$$K_D = \frac{P_{sig}}{P_{sig} + P_{bkg}} = \left[1 + \frac{P_{bkg}(\theta^*,\ \Phi_1,\ \theta_1,\ \theta_2,\ \Phi,\ m_{Z_1},\ m_{Z_2}|m_{4l})}{P_{sig}(\theta^*,\ \Phi_1,\ \theta_1,\ \theta_2,\ \Phi,\ m_{Z_1},\ m_{Z_2}|m_{4l})}\right]^{-1} \tag{4.6}$$

for every value of m_{4l}. By construction, the K_D value is constrained to be between zero and one. In Fig. 4.8 the distribution of the K_D variable is shown both for the signal and for the irreducible background, for 4μ, $4e$ and $2e2\mu$ final states.

For both the Isolation and K_D distributions, a good agreement was obtained between the FullSim and the Delphes simulations.

4.4.3 Invariant Mass distributions and Event Selection

The invariant mass distributions of the two Z bosons as reconstructed in the analysis are shown in Fig. 4.9 and in Fig. 4.10 for 4μ, $4e$ and $2e2\mu$ final state events.

In Fig. 4.11 the fractions of events passing each step of the selection with respect to the initial number of events (cut flow table) are shown, in order to compare the Delphes and Full Simulation performances at each step of theselection analysis. As a reference,the cut flow table of the analysis performed during Run1 at 8 TeV is reported with a black solid line. The same selection was applied to the irreducible background, and the final efficiency was computed in the same way as for the signal. In Table 4.4 the event selection efficiencies are summarized for signal and background events, in the three channels. It has to be noted that the efficiency of the 8 TeV distribution is larger than the Scenario 1 distribution, in all three channels: the reason is that no optimization was implemented in the analysis, and the selection criteria used were optimized for the 8 TeV selection.

In Fig. 4.12 the four-lepton invariant mass distributions are shown, computed at the end of the analysis for the three signal final states. In Fig. 4.13 and in Fig. 4.14 the invariant mass distributions, referred to an integrated luminosity of 3000 fb^{-1}, for the signal $H \to 4l$ and backgrounds are shown, for the three decay channels separately and together.

4.4.4 Summary

A good agreement between Delphes and the reference FullSim has been achieved for Scenario 1: there are only very small discrepancies between the distributions obtained with the two types of simulation. Thanks to this validation, Delphes is considered to be a reliable tool able to simulate the detector geometry with sufficient precision. Therefore, it can be used to simulate the Phase-II upgraded detector geometry (Scenario 3), as explained in the next Section.

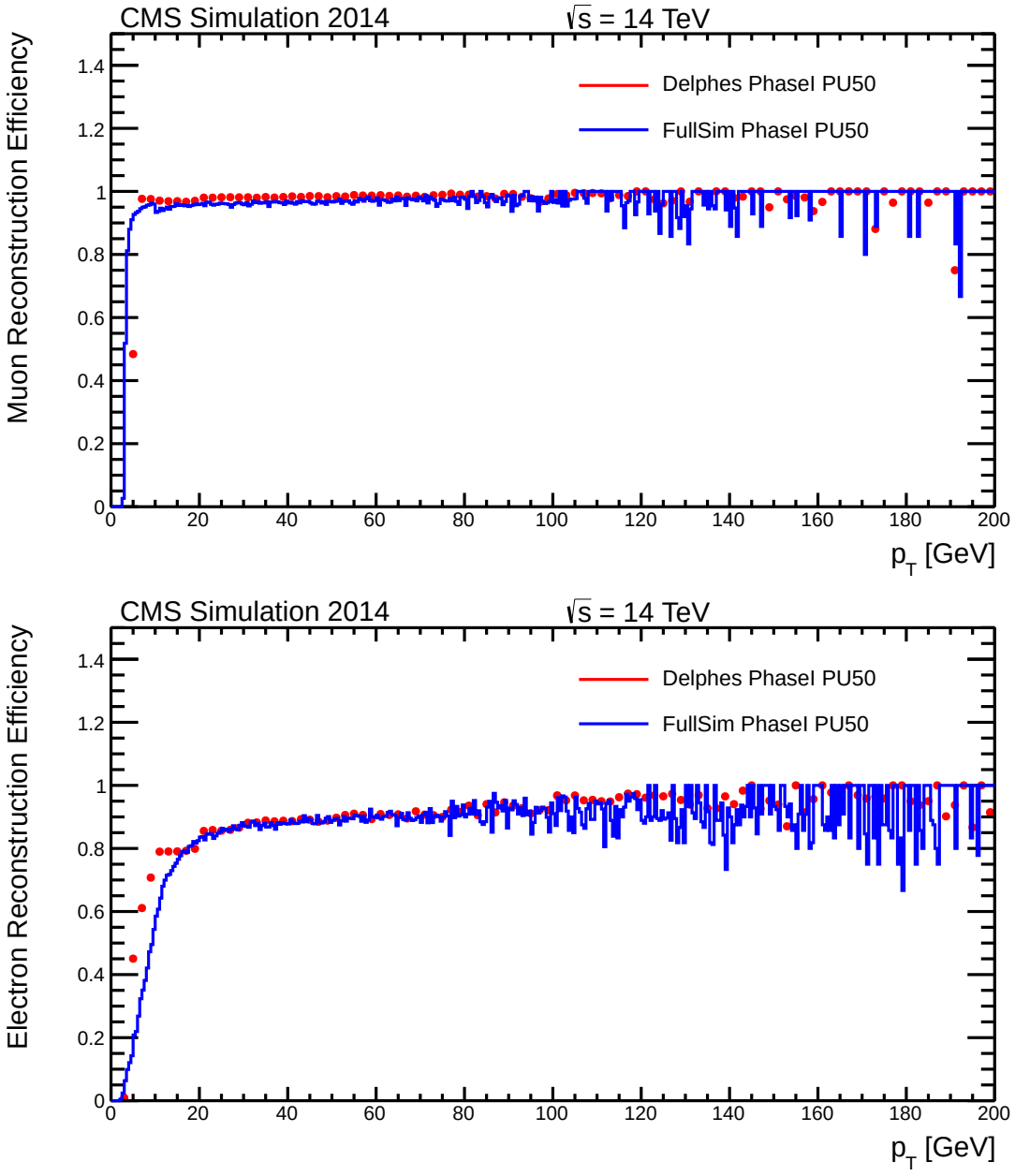

Figure 4.3: Scenario 1: lepton reconstruction efficiency as a function of p_T for FullSim (blue solid line) and Delphes (red dots). The upper plot is for muons, the lower plot is for electrons.

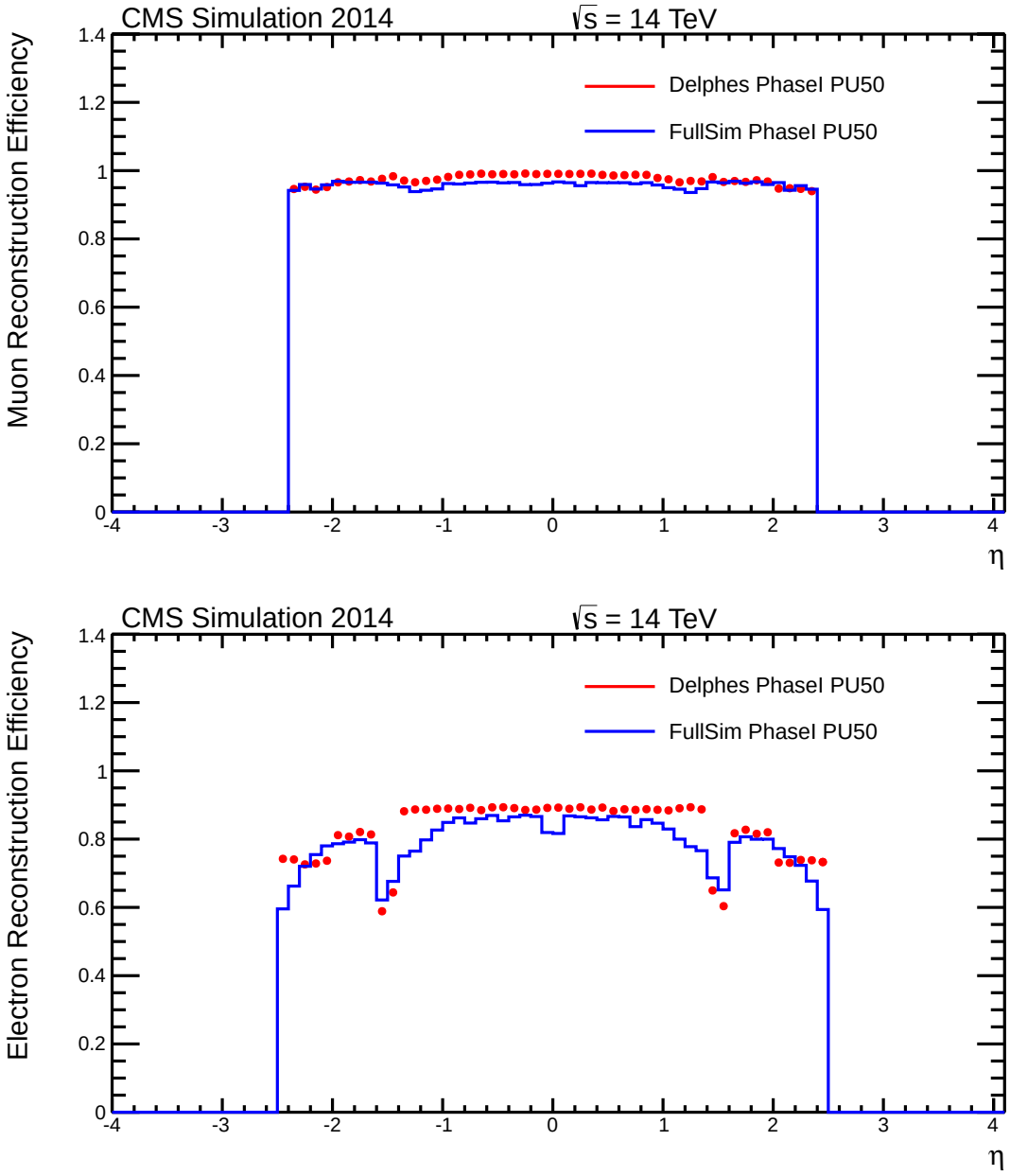

Figure 4.4: Scenario 1: lepton reconstruction efficiency as a function of η for FullSim (blue solid line) and Delphes (red dots). The upper plot is for muons, the lower plot is for electrons.

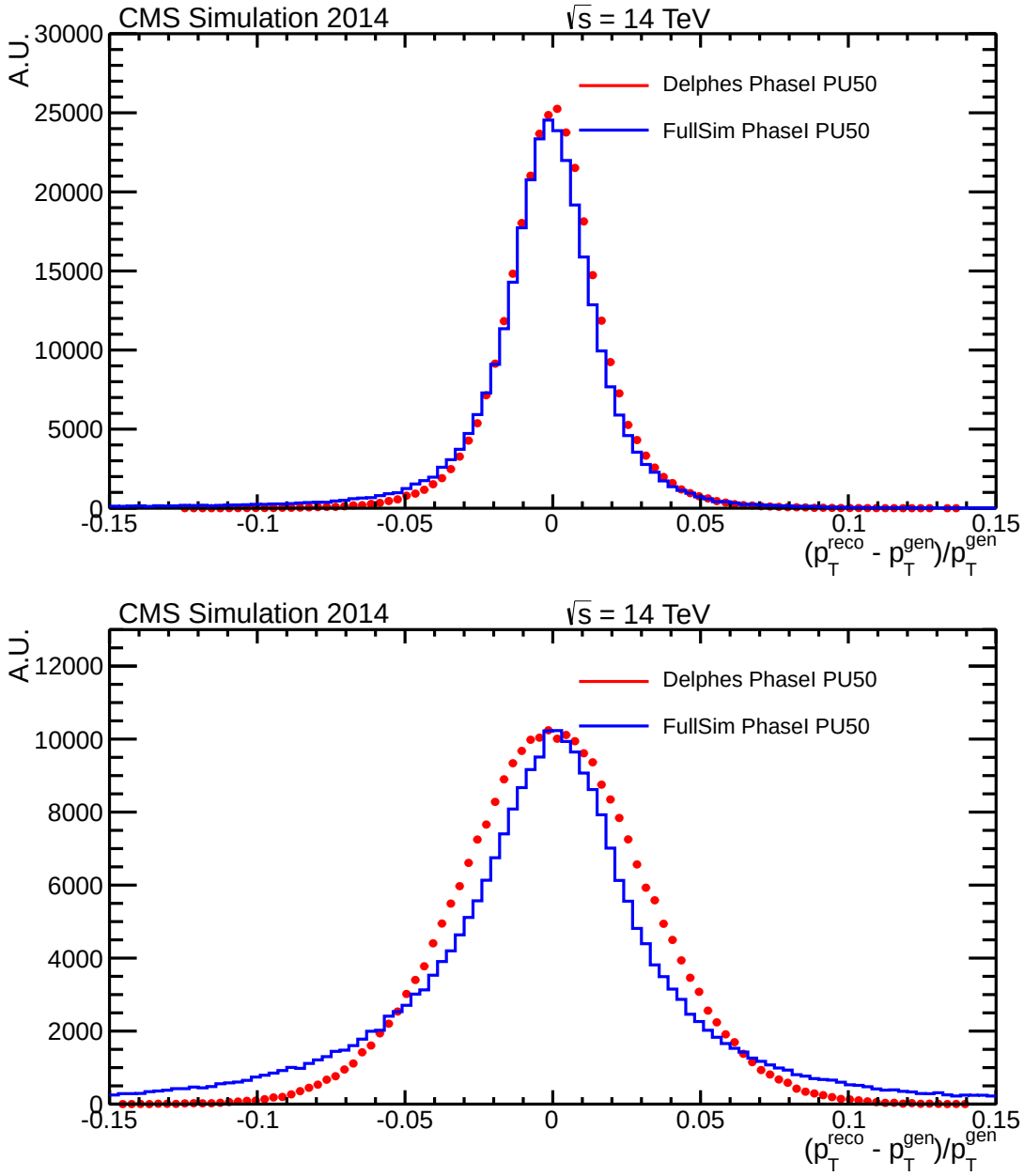

Figure 4.5: Scenario 1: single lepton $\frac{p_T^{\mathrm{reco}} - p_T^{\mathrm{gen}}}{p_T^{\mathrm{gen}}}$ distributions for FullSim (blue solid line) and Delphes (red dots). The transverse momentum resolution is defined as the width of the distribution. The upper plot is for muons, the lower plot is for electrons.

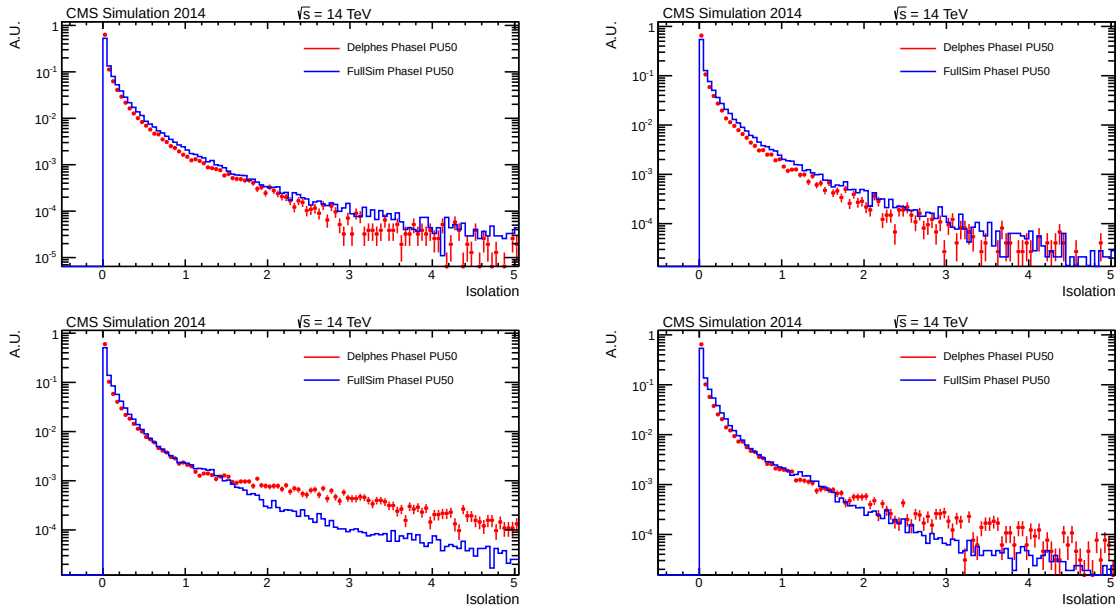

Figure 4.6: Scenario 1: Isolation variable distributions for leptons in signal $H \rightarrow ZZ^* \rightarrow 4l$ events (left plots) and in background $ZZ^* \rightarrow 4l$ events (right plots), for FullSim (blue solid line) and Delphes (red dots). The upper plots are for muons, the lower plots are for electrons.

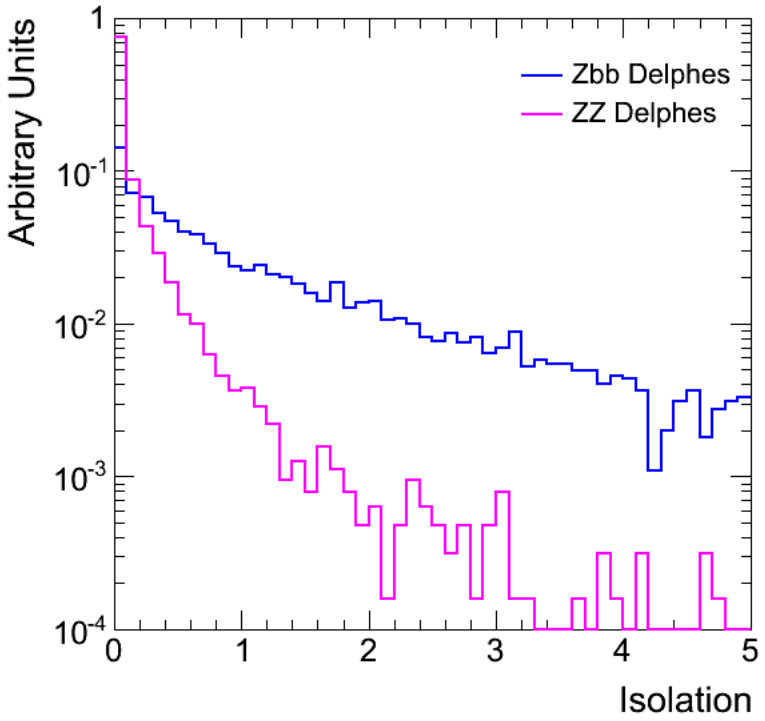

Figure 4.7: Scenario 1: Isolation variable distributions for leptons in irreducible background $ZZ \to 4l$ events (pink solid line) and in reducible background $Zb\bar{b}$ events (blue solid line).

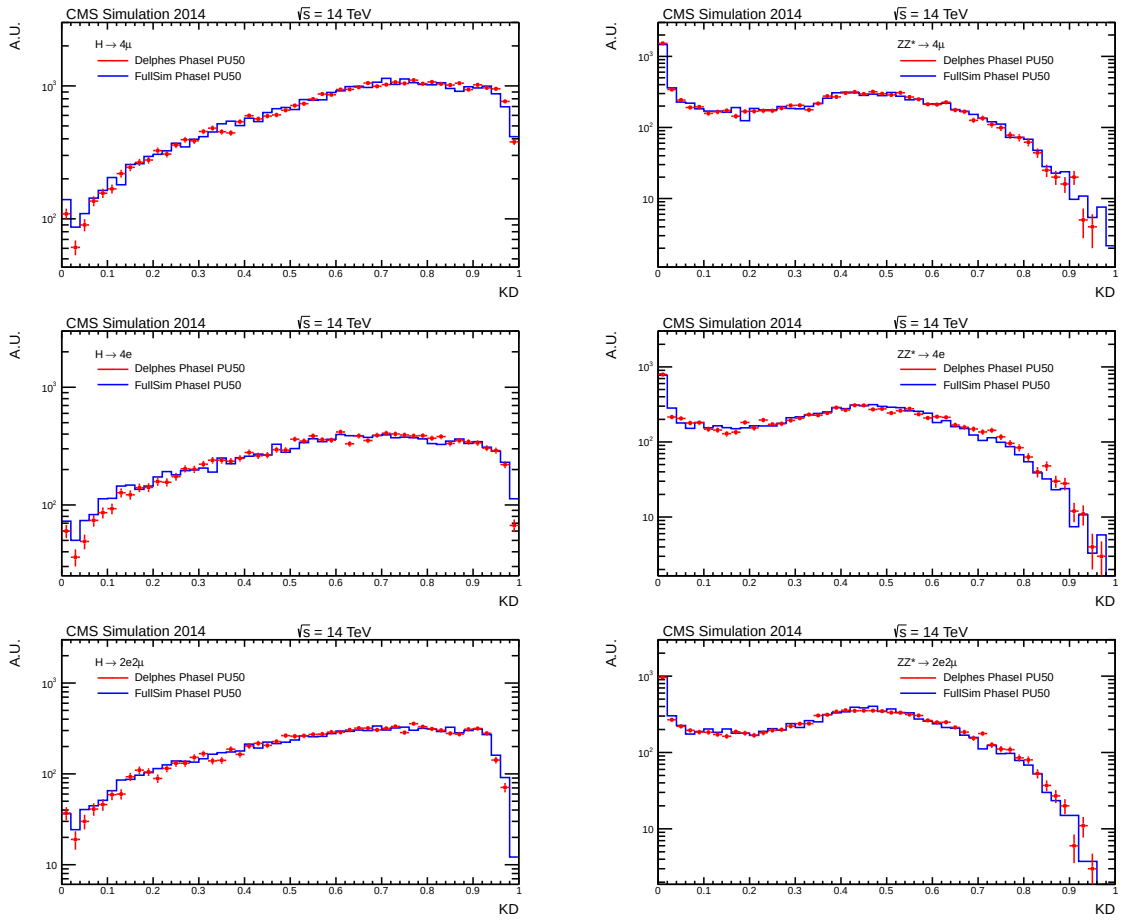

Figure 4.8: Scenario 1: Kinematic discriminant variable distribution for signal $H \rightarrow ZZ^* \rightarrow 4l$ events (left plots) and for irreducible background $ZZ^* \rightarrow 4l$ (right plots) events, for FullSim (blue solid line) and Delphes (red dots). The upper plots are for 4μ, the middle plots are for $4e$ and the bottom plots are for $2e2\mu$ final states.

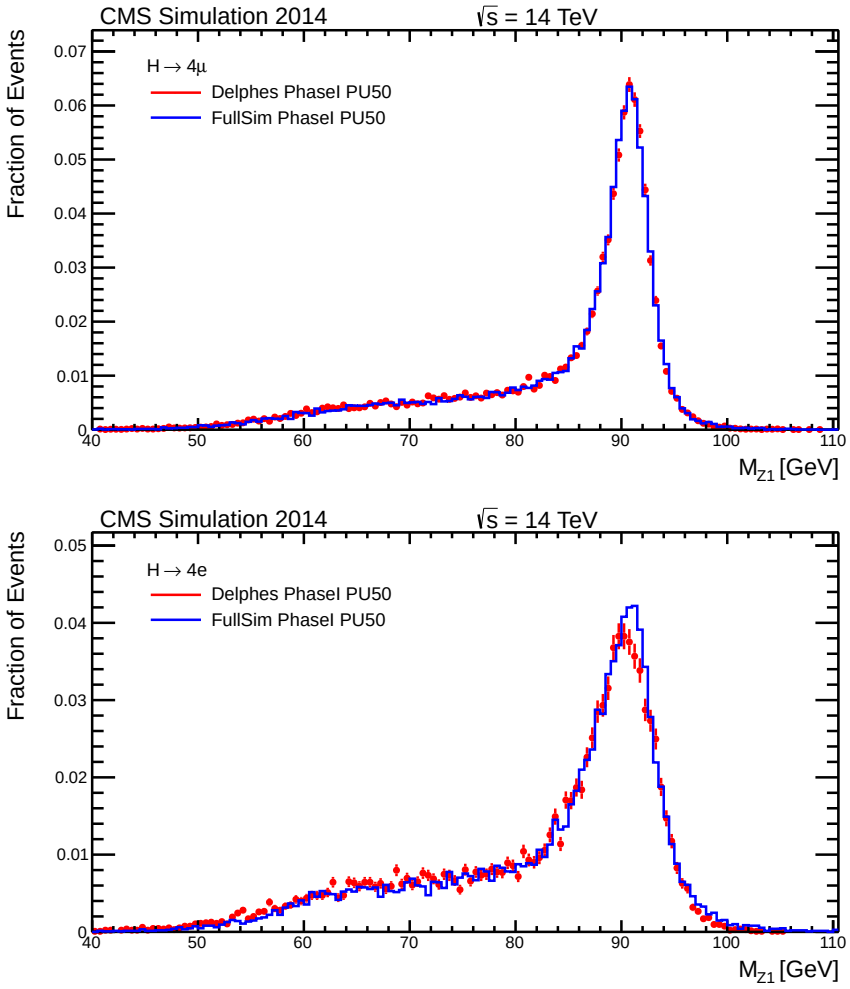

Figure 4.9: Scenario 1: Z_1 invariant mass distributions for FullSim (blue solid line) and Delphes (red dots), for the signal $H \to ZZ^* \to 4l$ sample. The upper plot is for muon pairs, while the lower plot is for electron pairs.

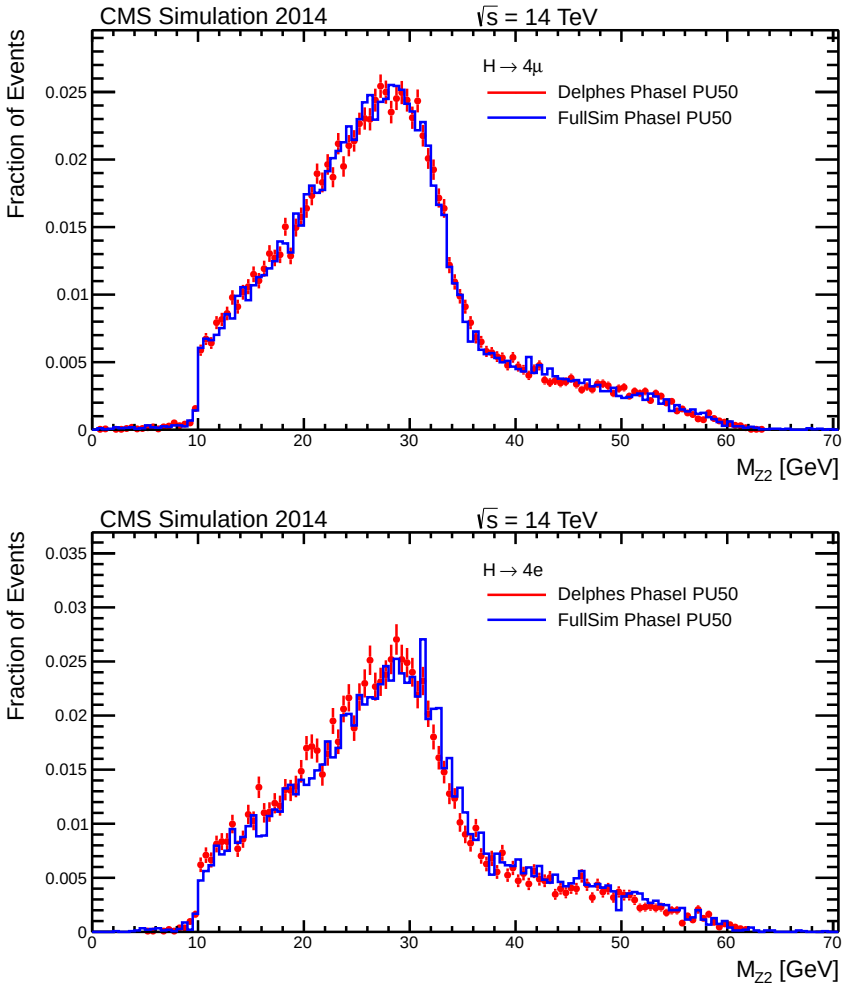

Figure 4.10: Scenario 1: Z_2 invariant mass distributions for FullSim (blue solid line) and Delphes (red dots), for the signal $H \to ZZ^* \to 4l$ sample. The upper plot is for muon pairs, while the lower plot is for electron pairs.

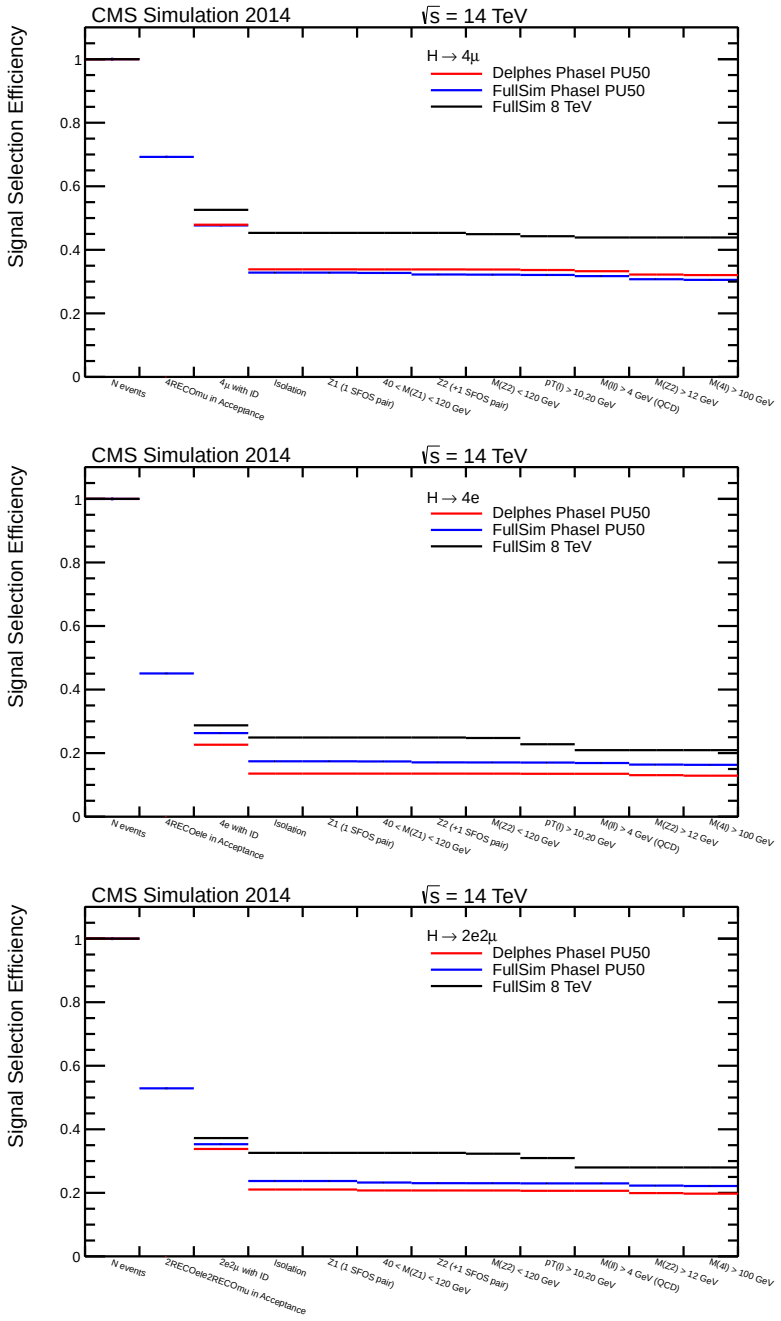

Figure 4.11: Scenario 1: cut flow table showing selection efficiencies at each step of the selection analysis for $H \rightarrow ZZ^* \rightarrow 4l$ events for FullSim (blue solid line) and Delphes (red dots) and 8 TeV real data (2012). The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is for $2e2\mu$ final states.

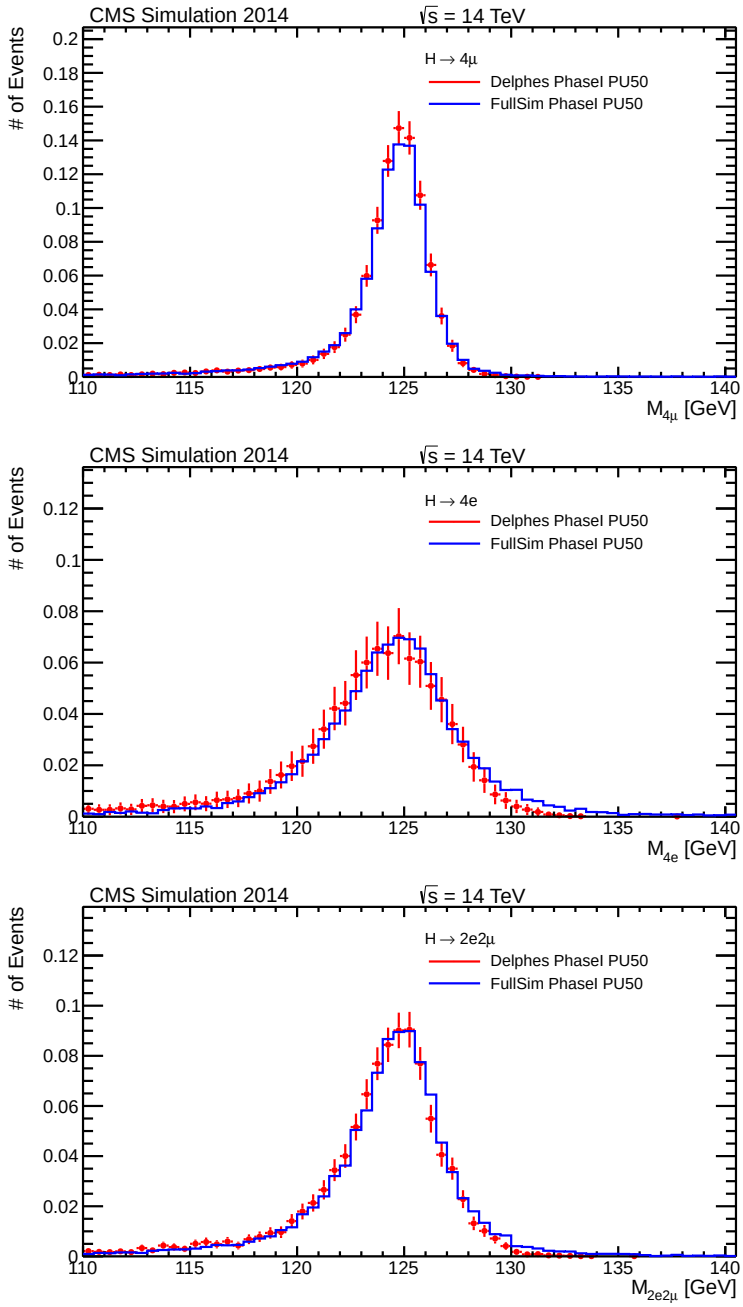

Figure 4.12: Scenario 1: four-lepton invariant mass distributions in $H \rightarrow ZZ^* \rightarrow 4l$ events, at the end of the selection analysis for FullSim (blue solid line) and Delphes (red dots). The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is for $2e2\mu$ final states.

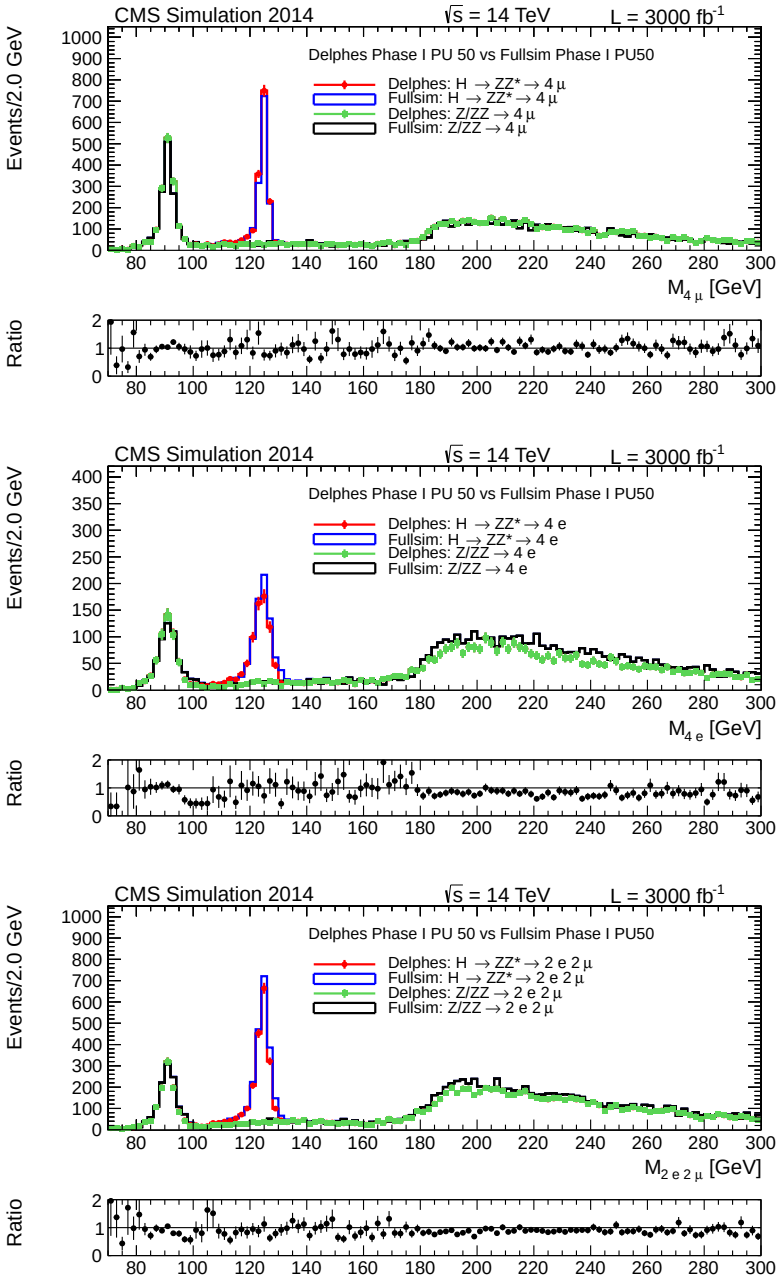

Figure 4.13: Scenario 1: four-lepton invariant mass distributions obtained with 3000 fb^{-1} considering the signal $H \rightarrow ZZ \rightarrow 4l$ (blue solid line and red dots), and the irreducible background $ZZ \rightarrow 4l$ (black solid line and green dots). The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is for $2e2\mu$ final states. Under each plot, the Ratio graph shows the ratio between the Delphes and FullSim distributions for the $ZZ \rightarrow 4l$ background, bin per bin.

Samples	Scenario 1 FullSim	Scenario 1 Delphes
$H \to ZZ \to 4\mu$	$(30.7 \pm 0.1)\%$	$(32.0 \pm 0.1)\%$
$H \to ZZ \to 4e$	$(16.4 \pm 0.1)\%$	$(12.9 \pm 0.01)\%$
$H \to ZZ \to 2e2\mu$	$(22.3 \pm 0.1)\%$	$(19.7 \pm 0.2)\%$
$ZZ \to 4\mu$	$(15.6 \pm 0.1)\%$	$(13.3 \pm 0.1)\%$
$ZZ \to 4e$	$(10.83 \pm 0.09)\%$	$(7.98 \pm 0.08)\%$
$ZZ \to 2e2\mu$	$(11.3 \pm 0.1)\%$	$(9.40 \pm 0.09)\%$

Table 4.4: Scenario 1: event selection efficiencies at the end of the analysis, for the signal $H \to ZZ \to 4l$ and for the irreducible background $ZZ \to 4l$ under the three detector Scenarios studied.

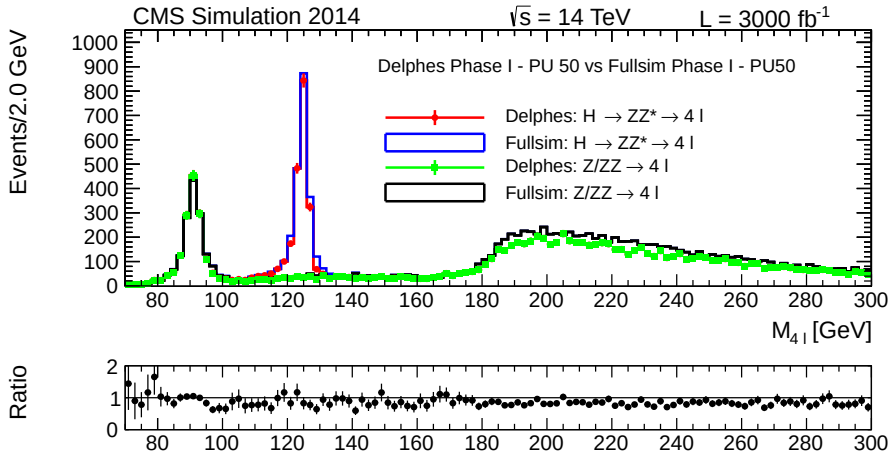

Figure 4.14: Scenario 1: four-lepton invariant mass distributions obtained with 3000 fb^{-1} for the signal $H \to ZZ \to 4l$ (blue solid line and red dots), and for the irreducible background $ZZ \to 4l$ background (black solid line and green dots). All four lepton final states are summed in this plot. The Ratio graph shows the ratio between the Delphes and FullSim distributions for the $ZZ \to 4l$ background, bin per bin.

4.5 Studies of the Phase-II Detector Performances

After the validation process, a study of the Phase-II detector upgrade (Scenario 3) was made, comparing the Delphes samples to:

- Full Simulation Phase-I detector geometry samples, used also in the validation process (Scenario 1);

Samples	Scenario 1 FullSim N evts. $\times 10^3$	Scenario 2 FullSim N evts. $\times 10^3$	Scenario 3 Delphes N evts. $\times 10^3$
$H \to ZZ \to 4\mu$	94	85	42
$H \to ZZ \to 4e$	94	93	42
$H \to ZZ \to 2e2\mu$	95	91	40
$ZZ \to 4\mu$	58	56	61
$ZZ \to 4e$	99	99	100
$ZZ \to 2e2\mu$	98	97	100

Table 4.5: Samples used for detector upgrade studies.

- Full Simulation Phase-I aged detector geometry samples (Scenario 2, as explained in Section 2.4.4);

In Table 4.5 the samples used for Scenarios 2 and 3 are listed. For the simulation of Scenario 3, the samples were produced taking into account the studies of the detector performances explained in Section 2.4.4

In the following, the results from Scenarios 1 and 3, respectively obtained with a Phase-I upgraded detector and a Phase-II upgraded detctor, are compared to the results from Scenario 2, corresponding to an aged detector. The distributions for these three scenarios are respectively shown as a blue solid line for Scenario 1, as a green solid line for Scenario 2, and as a red solid line for Scenario 3.

4.5.1 Efficiency and Resolution Distributions

In Fig 4.15 and in Fig. 4.16 the reconstruction efficiency as a function of the lepton transverse momentum and pseudo-rapidity respectively are shown. For the Phase-II samples, an improvement in the efficiency is clearly visible, due to the increased detector acceptance. However, for the four-electron case, due to the high pile-up, some efficiency losses are observed. For the aged detector samples, an evident loss in reconstruction efficiency is visible, as expected.

Since the forward electromagnetic calorimeter has a lower efficiency, in Fig. 4.17 the p_T resolution shows a worsening of the Scenario 3 distribution, with respect to Scenario 1, for the electron case. As expected, Scenario 2 remains the worst one. For the muon case, a slight improvement in the p_T resolution is visible, with respect to Scenario 1 and 2, due to the improvement of the tracker performances.

4.5.2 Kinematic Discriminant and Isolation Distributions

The lepton Isolation distributions, for muons and electrons coming from signal and irreducible background are shown in Fig. 4.18. As expected, for the upgrade scenario the shape of the distribution remains similar for both the signal and irreducible background, and becomes even more discriminant. On the contrary, the aged scenario shows an opposite behavior, reducing its discriminant capabilities.

The K_D distributions are shown in Fig. 4.19 for both signal and background. A reduced discrimination power is observed for the aged detector samples, while it is restored in the Phase-II Scenario.

4.5.3 Invariant Mass Distributions and Event Selection

In Fig. 4.20 and Fig. 4.21 the Z_1 and Z_2 invariant mass distributions are shown. In the upgraded detector scenarios 1 and 3, the mass resolution remains high, while for the aged detector it degrades.

In Fig. 4.22 the efficiency of the event selection is reported step by step: the final selection efficiency is larger for the 4μ and $2e2\mu$ final states for the Phase-II detector upgrade, while it remains similar to the Phase-I results for the $4e$ final state. As expected, the aged detector scenario is the worst one, for all three channels. It has to be noted that no tuning of the reconstruction and identification algorithms for the aged detector configuration were attempted. Proper tuning of such algorithms, which may result in an increase of the final selection efficiency, was out of the scope of this analysis. However some preliminary studies to improve the electron identification were performed: they are presented in Section 4.6.2. The event selection efficiencies at the end of the analysis are reported in Table 4.6, for the signal and background.

The invariant mass distribution of the four selected leptons m_{4l} is shown in Fig. 4.23. The improvement for the upgraded detector simulation is clearly visible for the 4μ and $2e2\mu$ final state, while for the $4e$ channel the improvement is not substantial. On the contrary, the aged detector scenario manifests a very low efficiency for all three channels.

In Fig. 4.24 and Fig. 4.25 the invariant mass distributions for the signal and the irreducible background, scaled to an integrated luminosity of 3000 fb^{-1} are shown for the three channels separately, and then added together. In this case the Phase-II detector simulation is compared to the Phase-I aged detector only.

A direct comparison between the Phase-I and the Phase-II scenarios is shown in Fig. 4.27: the improvements provided by an upgraded detector are clearly visible. Finally the 4-leptons invariant mass obtained with the Phase-II upgraded detector only is shown in Fig. 4.26, for the signal and the irreducible background, scaled to an integrated luminosity of 3000 fb^{-1}.

4.5.4 Summary

A degraded event selection efficiency is observed in the aged detector scenario, due to the decrease of the single lepton reconstruction efficiency, and the reduced Isolation and K_D discrimination power. On the contrary, the improvement on the single lepton reconstruction efficiency and resolution obtained with the Phase-II upgraded detector results into a better four-lepton invariant mass reconstruction and to a larger yield of signal events.

Samples	Scenario 1 FullSim	Scenario 2 FullSim	Scenario 3 Delphes
$H \to ZZ \to 4\mu$	(30.7 ± 0.1)%	(10.6 ± 0.1)%	(41.3 ± 0.1)%
$H \to ZZ \to 4e$	(16.4 ± 0.1)%	(1.08 ± 0.04)%	(15.6 ± 0.1)%
$H \to ZZ \to 2e2\mu$	(22.3 ± 0.1)%	(4.73 ± 0.07)%	(25.0 ± 0.2)%
$ZZ \to 4\mu$	(15.6 ± 0.1)%	(5.48 ± 0.09)%	(18.9 ± 0.1)%
$ZZ \to 4e$	(10.83 ± 0.09)%	(2.42 ± 0.05)%	(9.92 ± 0.09)%
$ZZ \to 2e2\mu$	(11.3 ± 0.1)%	(3.13 ± 0.05)%	(12.5 ± 0.1)%

Table 4.6: Event selection efficiencies at the end of the selection analysis. These are reported for the signal $H \to ZZ \to 4l$ and for the irreducible background $ZZ \to 4l$ under the three detector scenarios studied.

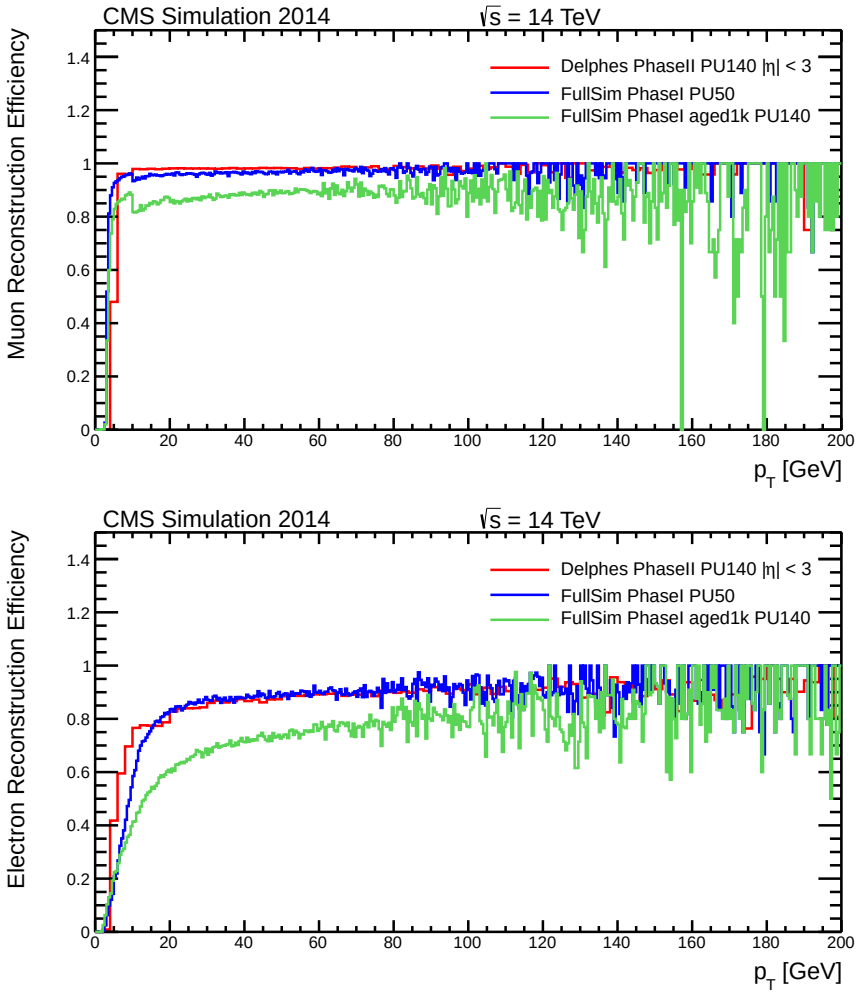

Figure 4.15: Lepton reconstruction and identification efficiencies as a function of the lepton p_T, for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The upper plot is for muons, the lower plot is for electrons.

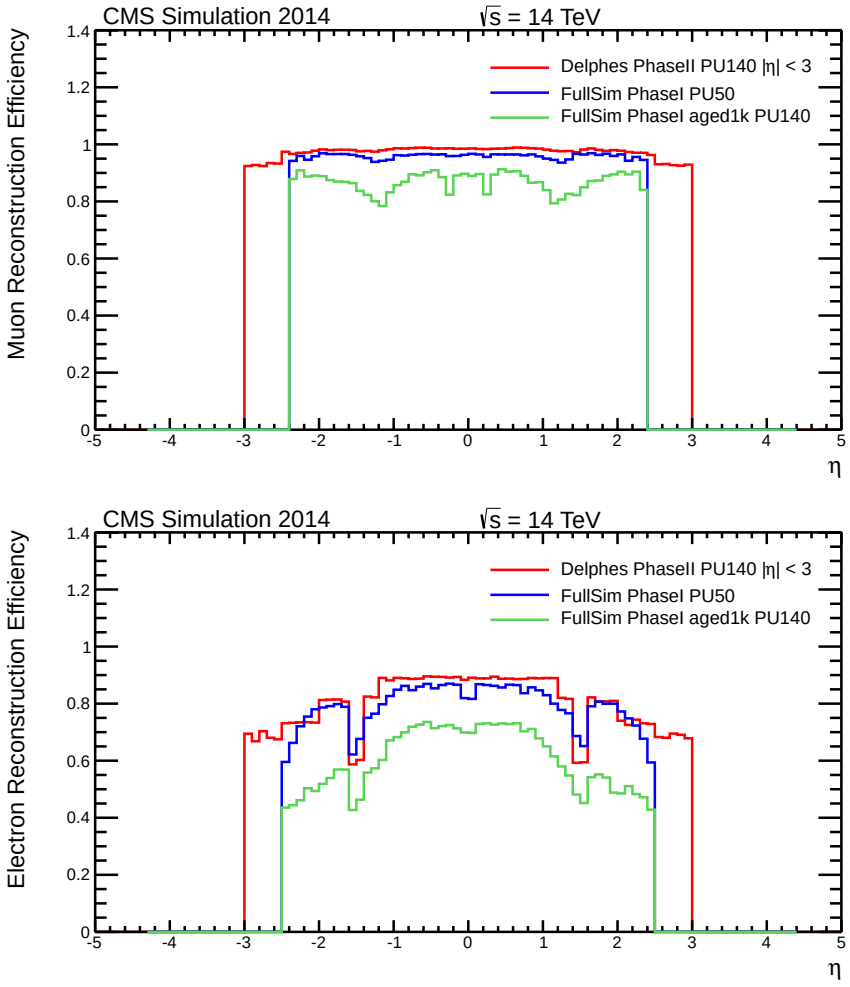

Figure 4.16: Lepton reconstruction and identification efficiencies as a function of the lepton η, for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The upper plot is for muons, the lower plot is for electrons.

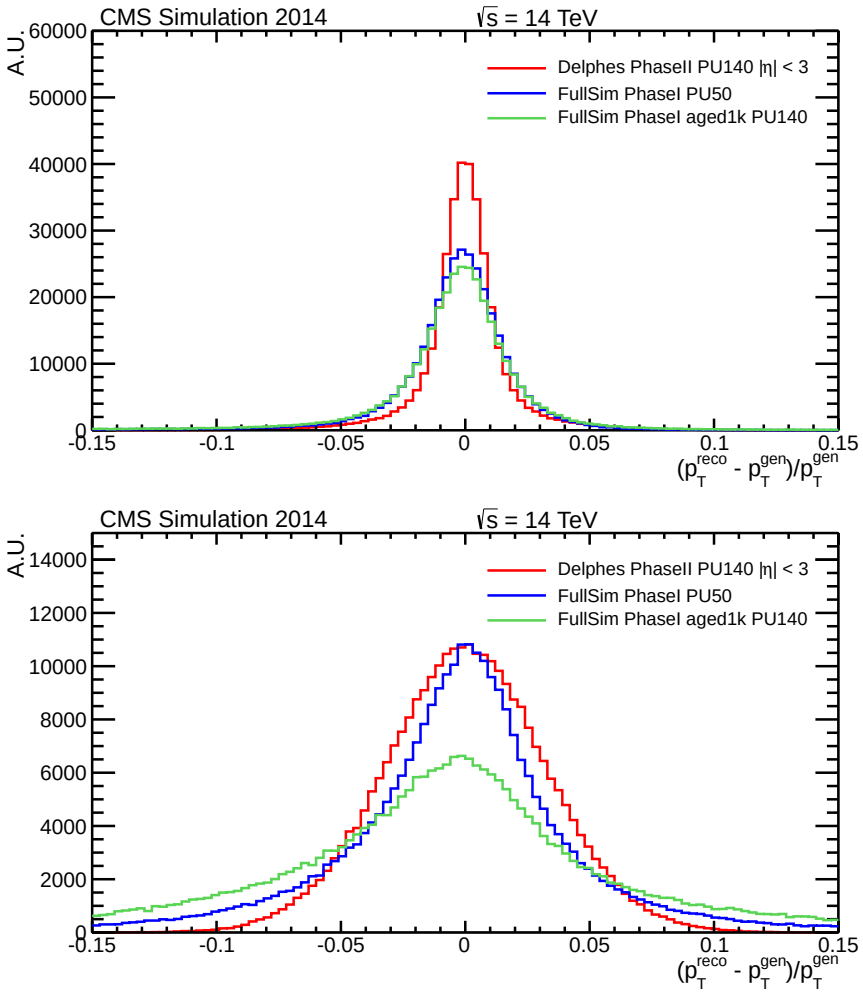

Figure 4.17: Single lepton $\frac{p_T^{\mathrm{reco}} - p_T^{\mathrm{gen}}}{p_T^{\mathrm{gen}}}$ distributions for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The transverse momentum resolution is defined as the width of the distribution. The upper plot is for muons, the lower plot is for electrons.

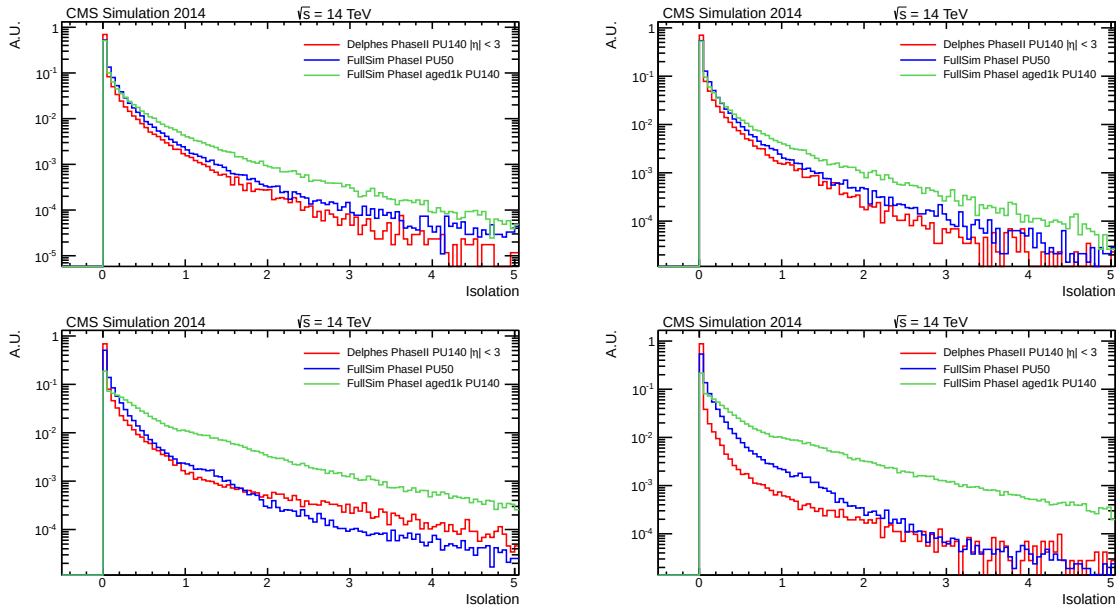

Figure 4.18: Isolation variable distributions for leptons in signal $H \to ZZ^* \to 4l$ events (left plots) and in background $ZZ^* \to 4l$ events (right plots), for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The upper plots are for muons, the lower plots are for electrons.

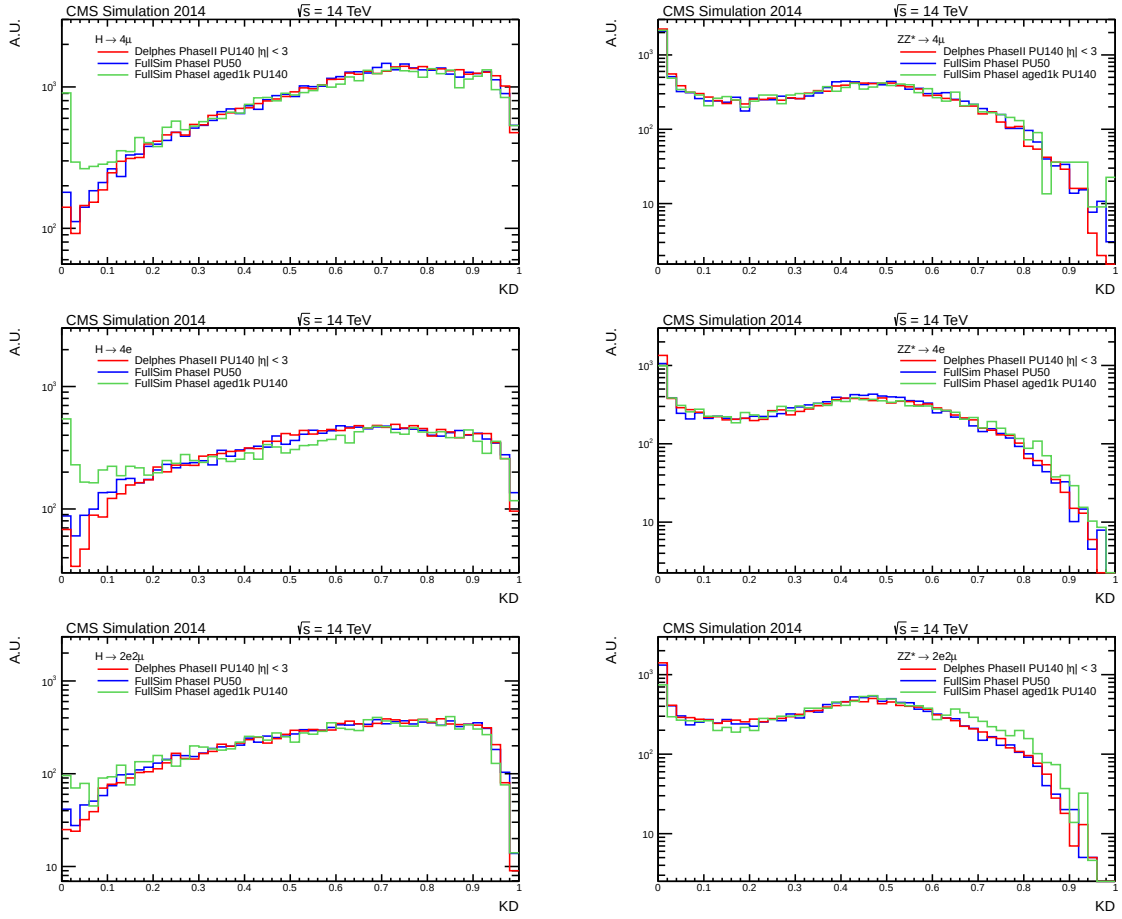

Figure 4.19: Kinematic discriminant variable distributions for signal $H \rightarrow ZZ^* \rightarrow 4l$ events (left plots) and for background $ZZ^* \rightarrow 4l$ events (right plots), for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The upper plots are for 4μ, the middle plots are for $4e$, the lower plots are for $2e2\mu$ final states.

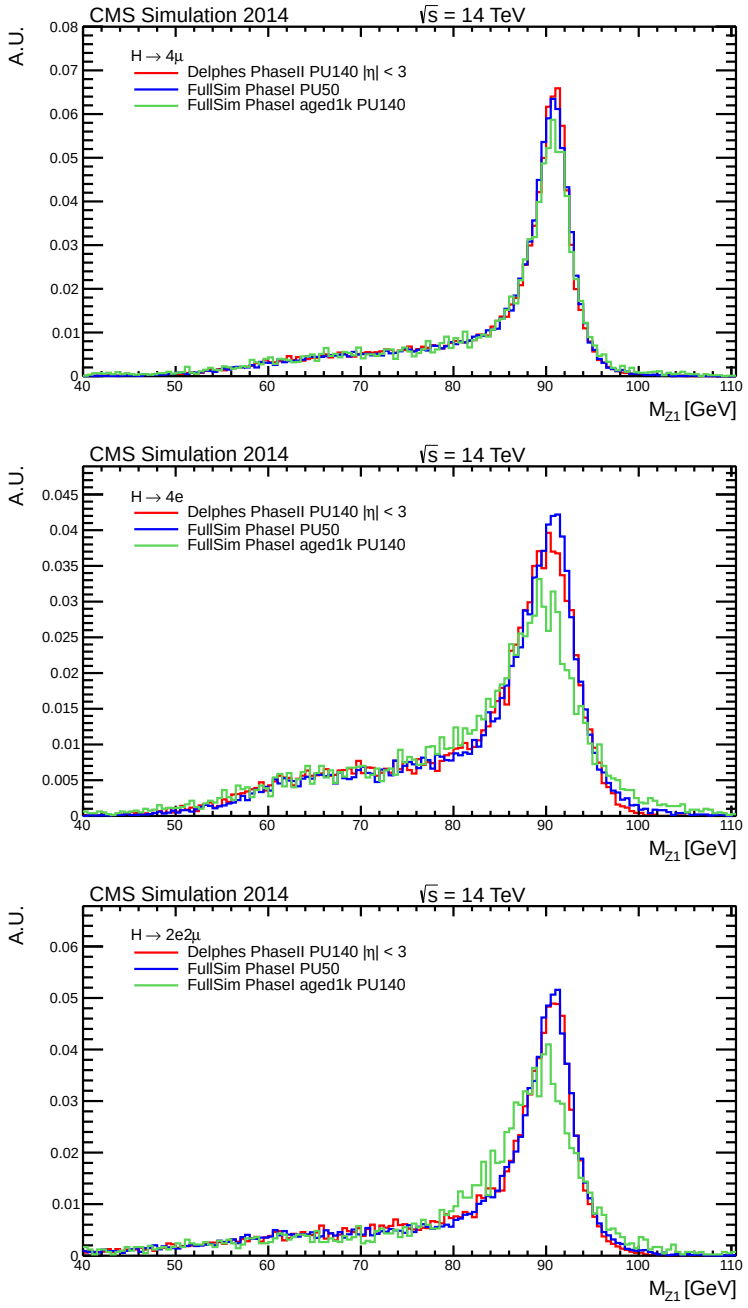

Figure 4.20: Z_1 invariant mass distributions for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is for $2e2\mu$ final states.

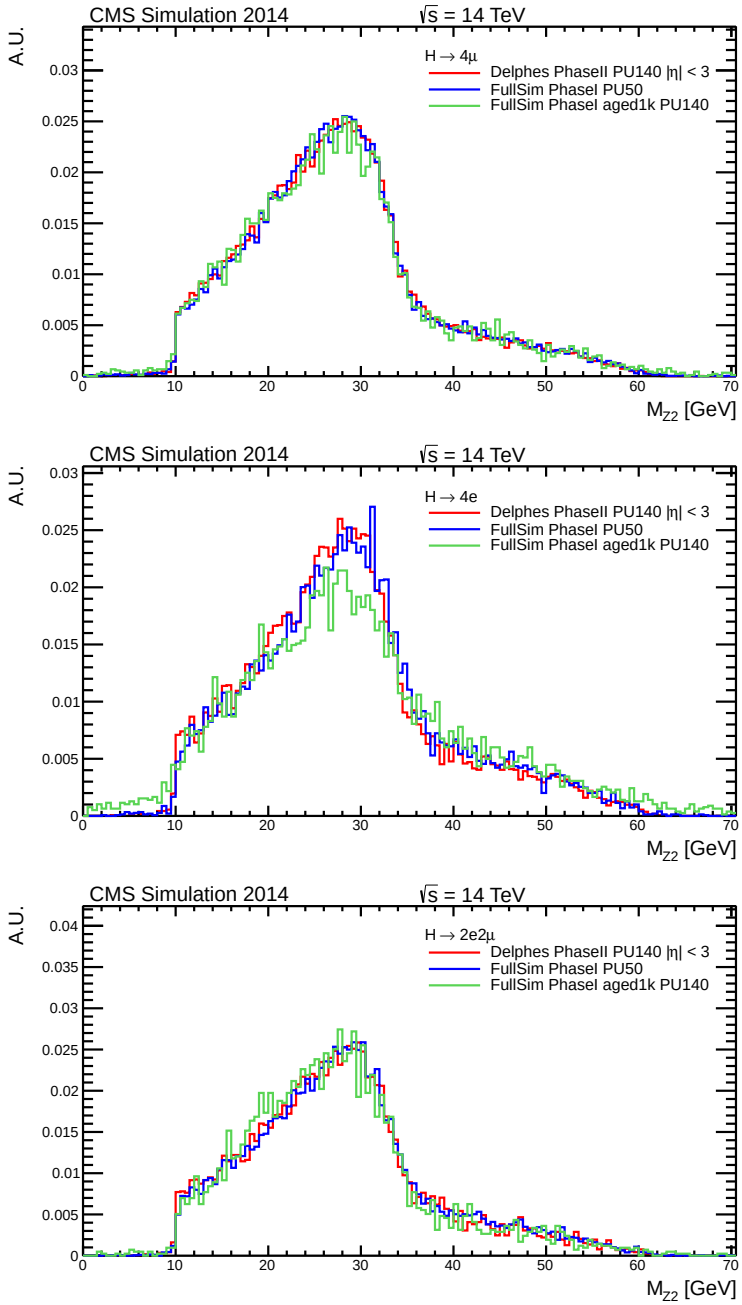

Figure 4.21: Z_2 invariant mass distributions for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is for $2e2\mu$ final states.

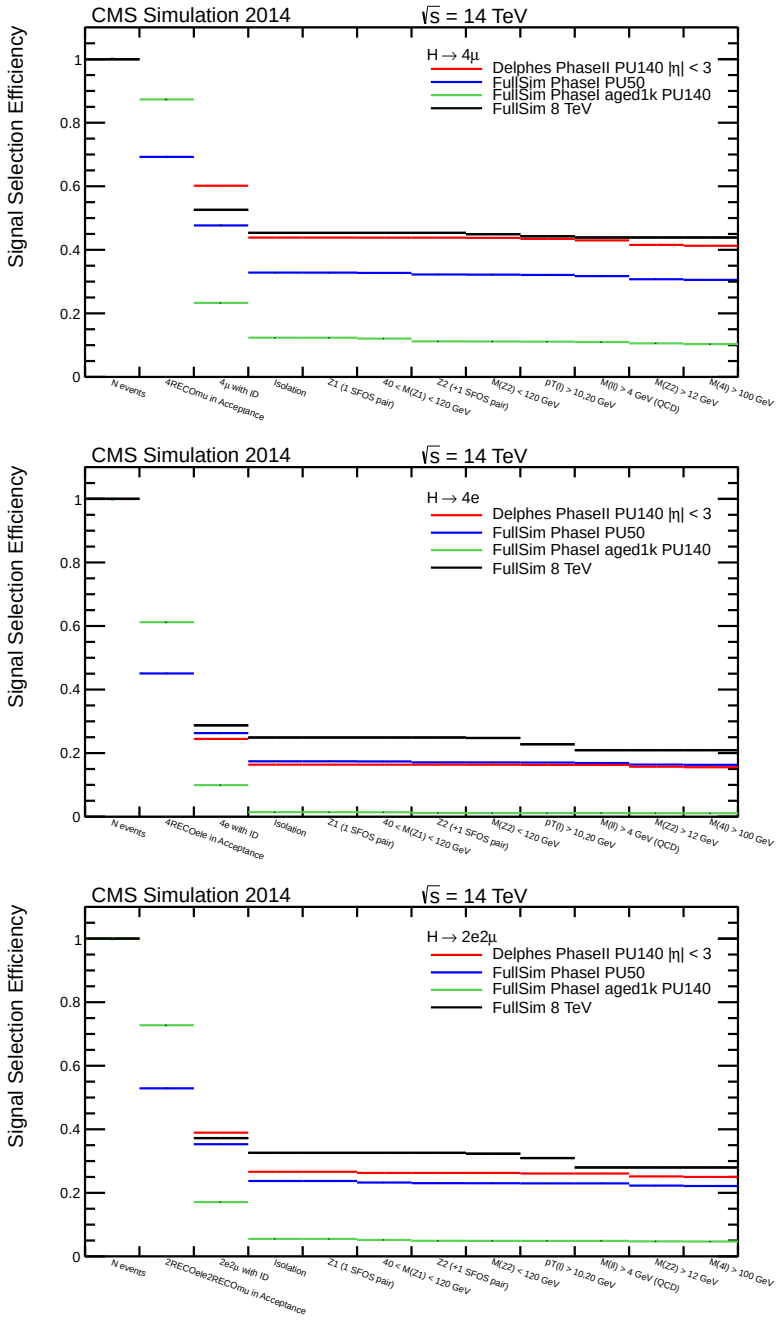

Figure 4.22: Cut flow table showing selection efficiencies at each step of the selection analysis for signal $H \rightarrow ZZ^* \rightarrow 4l$ events, for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line) and 8 TeV real data (2012). The upper plot is for 4μ, the middle plot is for $4e$, the lower plots is for $2e2\mu$ final states.

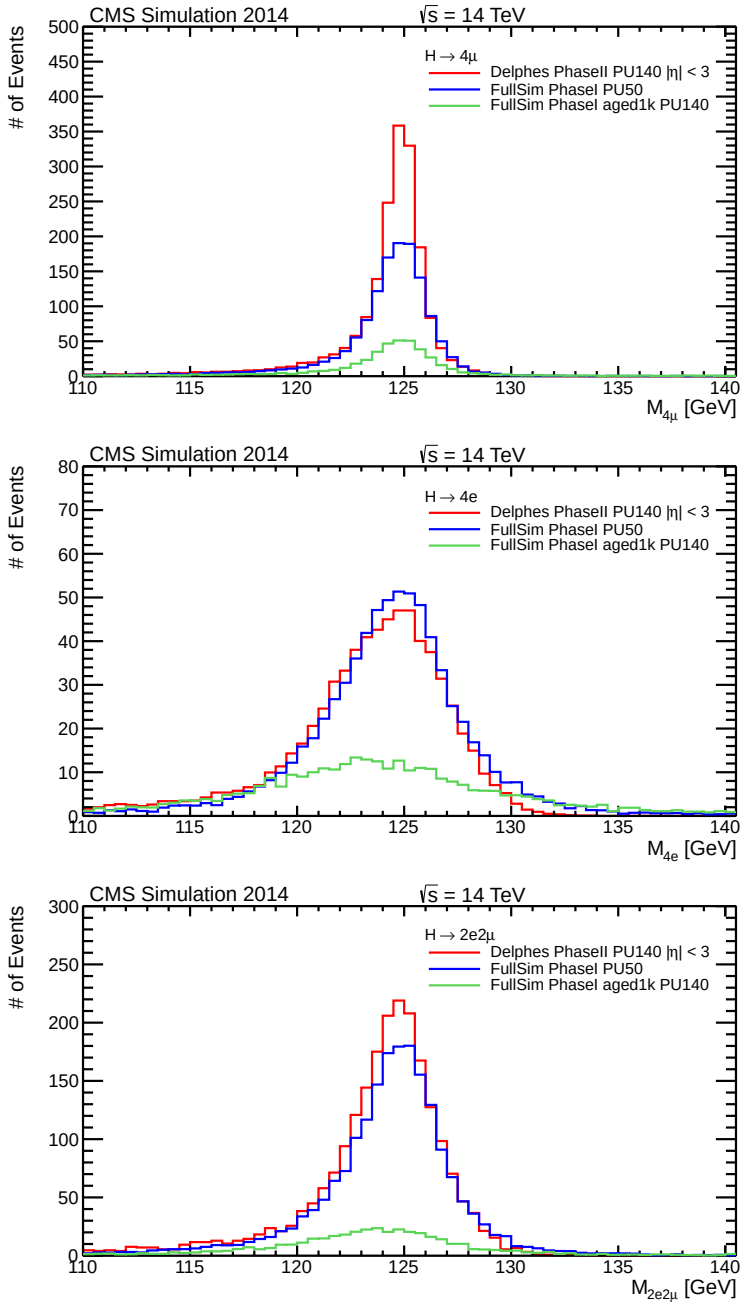

Figure 4.23: Four-lepton invariant mass distributions, for Scenario 1 (blue solid line), Scenario 2 (green solid line) and Scenario 3 (red solid line). The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is for $2e2\mu$ final states.

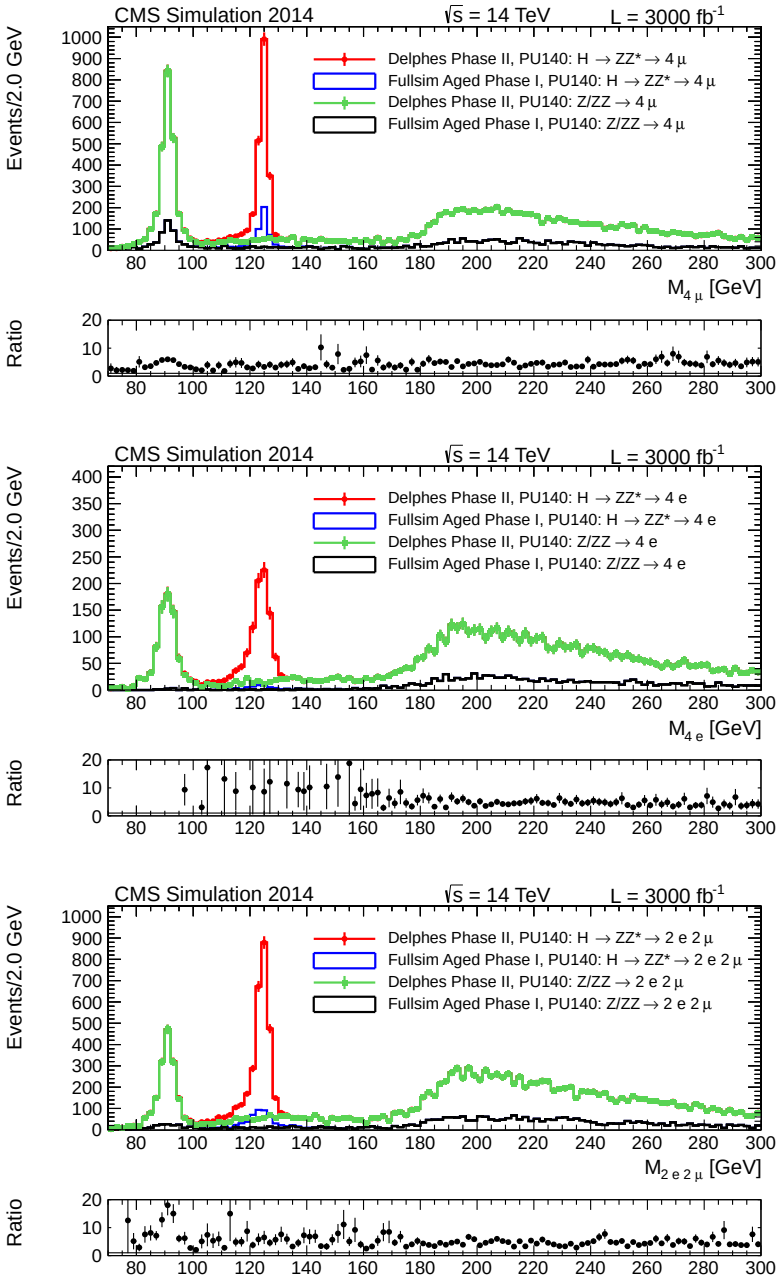

Figure 4.24: Four-lepton invariant mass distributions obtained with 3000 fb^{-1} for the signal $H \to ZZ \to 4l$ (blue solid line and red dots), and for the irreducible background $ZZ \to 4l$ (black solid line and green dots), for Scenario 2 and Scenario 3. The upper plots is for 4μ, the middle plots is for $4e$, the lower plots is for $2e2\mu$ final states. Under each plot, the Ratio graph shows the ratio between the Delphes and FullSim distributions for the $ZZ \to 4l$ background, bin per bin.

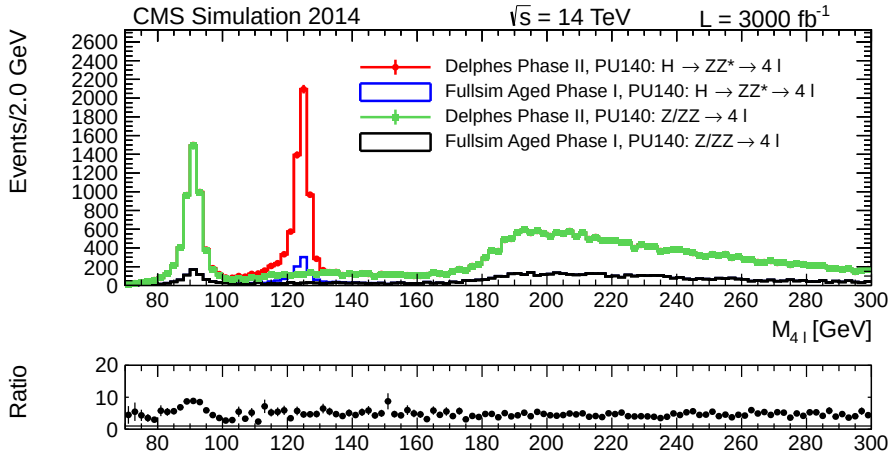

Figure 4.25: Four-lepton invariant mass distributions obtained with 3000 fb^{-1} for the signal $H \to ZZ \to 4l$ (blue solid line and red dots), and for the irreducible background $ZZ \to 4l$ (black solid line and green dots), for Scenario 2 and 3. All four-lepton final states are summed in this plot. The Ratio graph shows the ratio between the Delphes and FullSim distributions for the $ZZ \to 4l$ background, bin per bin.

4.6 Other Detector Performance Studies

Further studies were made in order to check the performances of the Phase-II detector geometry simulation in different pseudo-rapidity ranges. An attempt to optimize the electron identification for Scenario 3 was also performed.

4.6.1 Phase-II Detector: Different η Extensions

In order to understand the effect of an extended geometry on the event selection, the Phase-II detector selection efficiency was also computed with a restriction in the η acceptance region to $\eta < 2.4$. Lepton reconstruction efficiency distributions in the two configurations are shown in Fig. 4.28 as a function of η. These different efficiencies lead to different final event selection efficiencies for the three channels, as visible in Fig. 4.29. The final efficiencies increase with the increasing of the η acceptance region. However this effect is very small for the four-electron final states, because of the low reconstruction efficiency of very high pseudo-rapidity electrons.

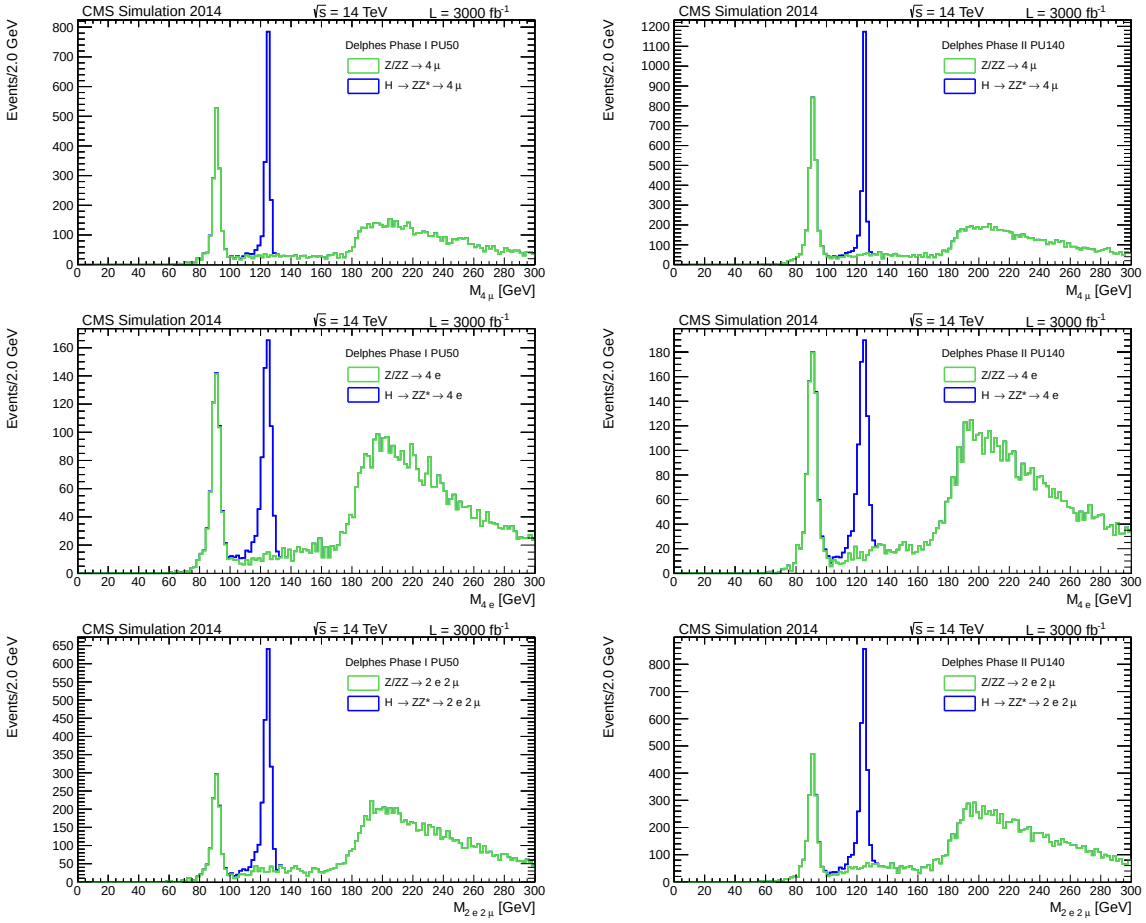

Figure 4.26: Four-lepton invariant mass distributions obtained with 3000 fb^{-1} for the signal $H \to ZZ \to 4l$ (blue), and for the irreducible background $ZZ \to 4l$ (green). Plots on the left are for Scenario 1, plots on the right are for Scenario 3. The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is $2e2\mu$ final states.

4.6.2 Phase-I Aged Detector: Electron ID

The optimization of the electron identification for an aged detector in high pile-up conditions was beyond the scope of this analysis. However, different approaches were used and applied to the event selection in order to compare the results with the ones obtained with the multivariate analysis (BDT, see Section 2.5.4).

1. Cut-based Identification: the electron identification is based on subsequent cuts on different physics variables of the particle. In this particular case, the selection criteria were optimized for the Phase-I analysis at pile-up 50 and applied to five variables, with different requirements for two different pseudo-rapidity regions.

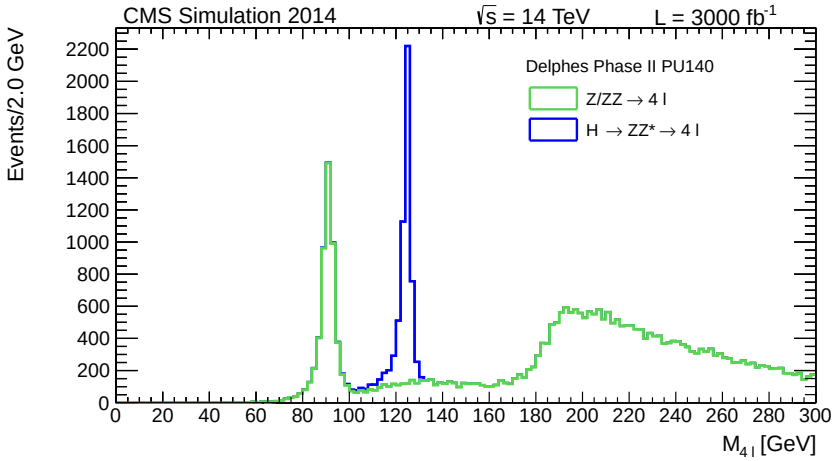

Figure 4.27: Four-lepton invariant mass distributions obtained with 3000 fb^{-1} for the signal $H \to ZZ \to 4l$ events (blue), and for the irreducible background $ZZ \to 4l$ (green). The three decay channels are added together.

2. No Identification: the identification requirement is completely removed, and all the electrons reconstructed within the detector acceptance are selected.

3. Isolation+Impact Parameter Identification: the BDT Identification was removed and only the Isolation cut was maintained but relaxed in order to achieve an efficiency of approximately 20% on the four-lepton signal. Then, a proper cut on the Impact Parameter was set, in order to keep the same fake-rate than in the rest of the analysis.

In Fig.4.30 the three identification methods are compared. The 8 TeV data selection efficiency is also reported, as a reference. All three methods show very low efficiencies.

1. The cut-based ID leads to an even smaller efficiency than the BDT ID. A thorough retuning of the selection criteria needs to be done to cope with the different pile-up conditions of the Phase-II detector configuration.

2. Removing completely the ID requirement does not provide significant improvements: the final efficiency remains below 5%. This is due to the high discriminant power of the Isolation requirement, that rejects the majority of the electrons. Moreover this method is not reliable because of the presence of potentially non-identified electrons in the final state.

3. The Isolation+IP ID leads to a higher efficiency, even if it remains very small, under 10%.

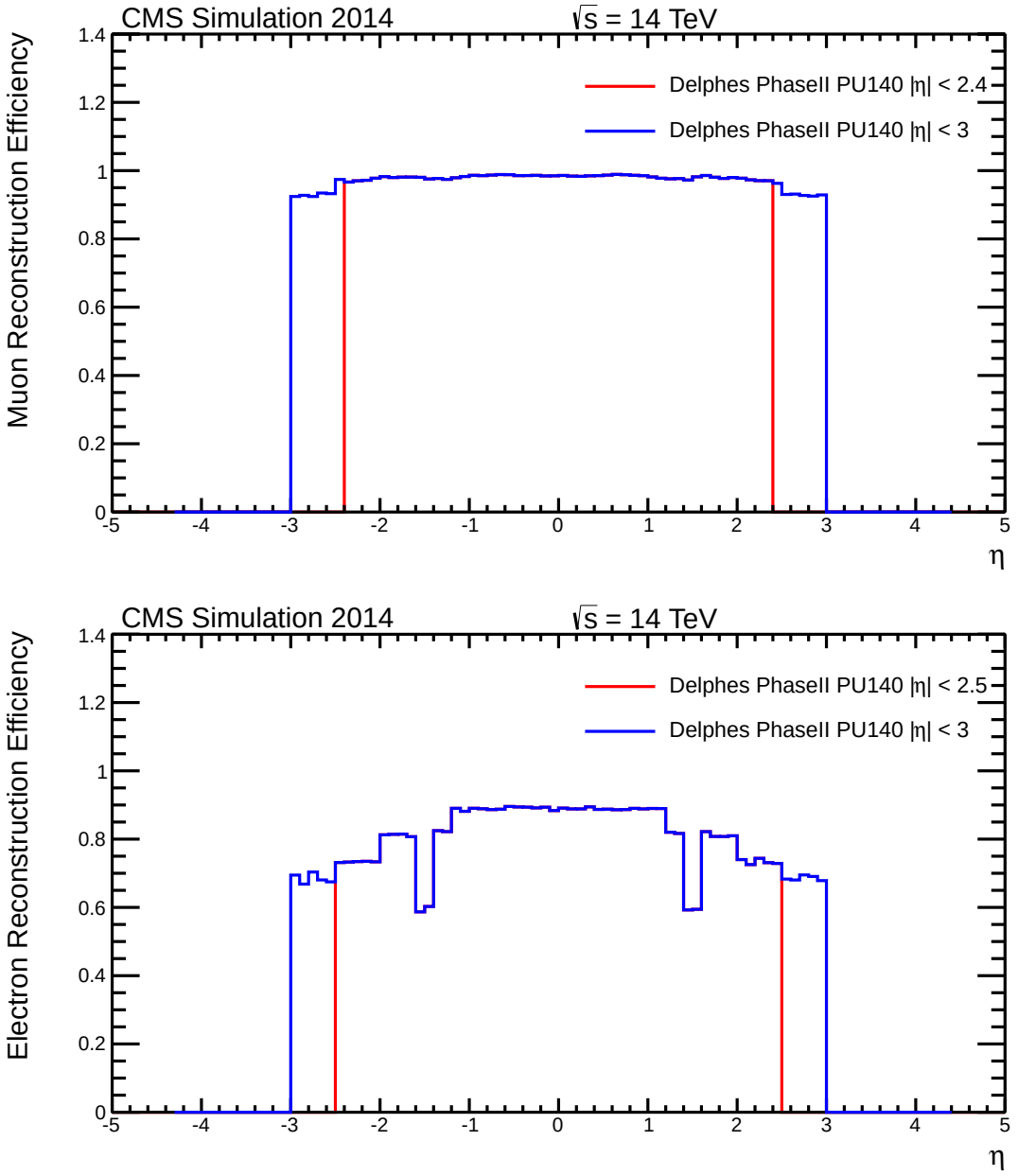

Figure 4.28: Lepton reconstruction and identification efficiencies as a function of the lepton η, for Scenario 3 with $\eta < 2.4$ (red solid line) and $\eta < 3.0$ (blue solid line). The upper plot is for muons, the lower plot is for electrons.

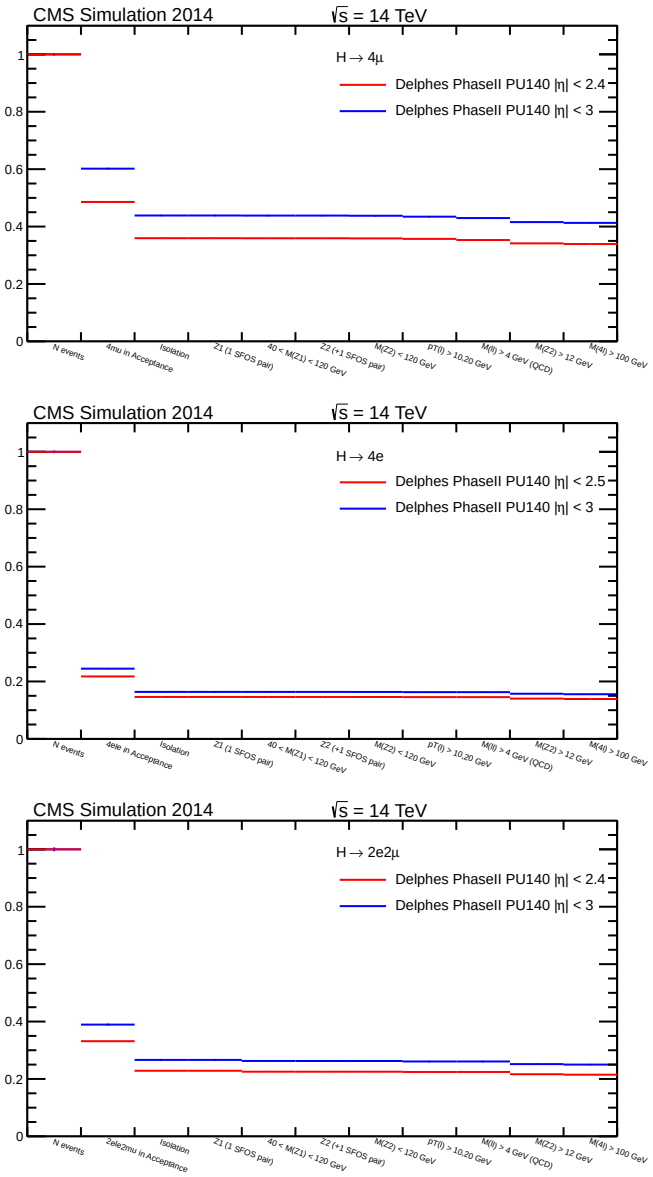

Figure 4.29: Cut flow table showing selection efficiencies at each step of the selection analysis for signal $H \to ZZ^* \to 4l$ events, for Scenario 3 with $\eta < 2.4$ (red solid line) and $\eta < 3.0$ (blue solid line). The upper plot is for 4μ, the middle plot is for $4e$, the lower plot is for $2e2\mu$ final states.

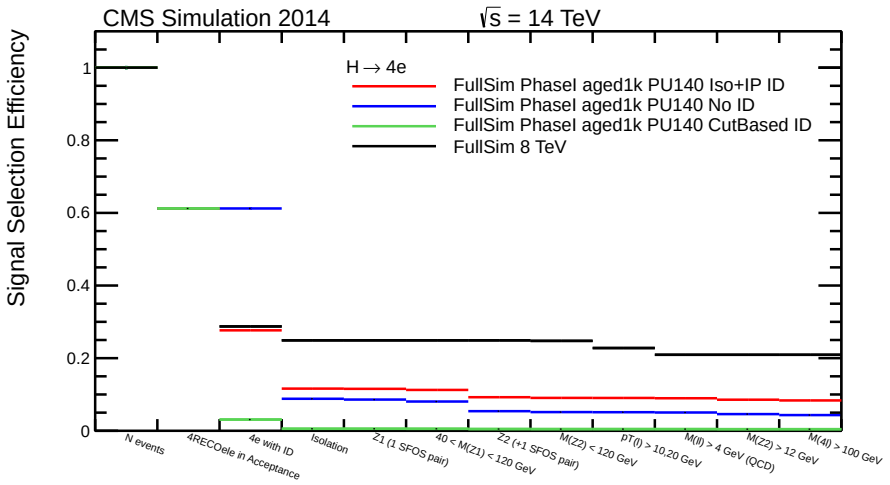

Figure 4.30: Cut flow table showing selection efficiencies at each step of the selection analysis for signal $H \rightarrow ZZ^* \rightarrow 4e$ events, for Scenario 2 with different electron ID: Isolation+IP ID (red line), no ID (blue line), Cutbased ID (green line).

Conclusions

The Large Hadron Collider (LHC) will soon restart operations after a "long shutdown" period (LS1) during which the accelerator was upgraded in order to reach a center-of-mass energy of 13 TeV. The LHC will be operated with a cycle of three years of data taking and then a long shutdown to perform ordinary maintenance of the accelerator complex. During the following long shutdown (LS2), the CMS experiment will also undergo a substantial upgrade, referred to as Phase-I. In the third long shutdown (LS3) scheduled in 2023, LHC will be substantially upgraded to reach much higher luminosities. To cope with these improvements the CMS detector will also need to be upgraded: this second stage is called Phase-II.

In this thesis, the impact of the Phase-II CMS detector upgrade on the $H \to ZZ^* \to 4l$ analysis was studied, comparing the results of the Higgs analysis in three different future scenarios. The Phase-II detector performance was compared to that of the Phase-I detector (used as reference) and to another future scenario corresponding to the Phase-I detector affected by the aging due to ten years of data-taking, i.e. as it would be after LS3 in the absence of any intervention. Options of the possible upgraded detector considered include an extended coverage in the forward region of tracker, muon system and electromagnetic calorimeter.

Since the three scenarios correspond to future configurations of the detector, Monte Carlo simulations were used to study them. The Phase-I reference detector configuration and the Phase-I aged configuration were simulated with the CMS Full Simulation, which provides a complete and detailed description of the detector geometry and allows a full particle reconstruction and identification. However, since no Full Simulation samples were available for the Phase-II detector, the Delphes framework was used instead. Delphes, which is a parametrized fast simulation, had to be validated to confirm its reliability. For this reason, a comparison between its results and the Full Simulation ones was performed. Since the results were satisfying, Delphes was then used to parametrize the Phase-II detector response.

From the obtained results, two main conclusions could be derived. The performance of the aged detector is significantly worse than the reference Phase-I detector performances. The event selection efficiency degrades from $\sim 30\%$ to $\sim 10\%$ for the 4μ channel, from $\sim 15\%$ to $\sim 0\%$ for the $4e$ channel, and from $\sim 22\%$ to $\sim 5\%$ for the $2e2\mu$ channel.

These results show that in order to maintain a high efficiency, an upgrade of the detector is mandatory.

This goal can be achieved with the Phase-II detector upgrade, capable of coping with the HL-LHC challenging conditions. The event selection efficiency would increase to $\sim 40\%$ in the 4μ channel, would remain $\sim 15\%$ for the $4e$ channel, and would increase to $\sim 25\%$ for the $2e2\mu$ channel. These results should be interpreted keeping in mind the very challenging experimental conditions of HL-LHC. In particular, the high pile-up affects the reconstruction and identification of electrons much more than muons. Moreover, the same baseline analysis as that used in 2012 was applied, without any optimization.

In the future, the reconstruction and identification, together with the analysis selection strategy will be revised and optimized to fully exploit the performances of the upgraded CMS detector.

Bibliography

[1] F. Halzen, A. D. Martin, *Quarks and Leptons: An introductory course in Modern Particle Physics*, Jhon Wiley and Sons (1984)

[2] S. Braibant, G. Giacomelli, M. Spurio, *Particles and Fundamental Interactions: An Introduction to Particle Physics*, Springer (2012)

[3] D. H. Perkins, *Introduction to High Energy Physics*, Cambridge Univ. Press (2000)

[4] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl.Phys. **22** (1961) 579

[5] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264

[6] A. Salam, *Elementary Particle Physics: Relativistic Groups and Analyticity (Nobel Symposium No. 8)*, N. Svartholm, Almquist and Wiksells, Stockholm (1968) 367

[7] J. Goldstone, *Field Theories with Superconductor Solutions*, Nuovo Cimento **19** (1961) 154.

[8] J. Goldstone, A. Salam and S. Weinberg, *Broken Symmetries*, Phys. Rev. **127** (1962) 965

[9] T. P. Cheng, L. F. Li, *Gauge theory of elementary particle physics*, Oxford University Press (1984)

[10] C. Quiqq, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Princeton Univ. Press (2013)

[11] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321

[12] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132

[13] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys.Rev. **145** (1966) 1156

[14] G. S. Guralnik, C. R. Hagen, T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585

[15] T.W.B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, Phys. Rev. **155** (1967) 1554

[16] *Gargamelle* http://home.web.cern.ch/about/experiments/gargamelle

[17] *The Super Proton Synchrotron* http://home.web.cern.ch/about/accelerators/super-proton-synchrotron

[18] *The Large Electron-Positron Collider* http://home.web.cern.ch/about/accelerators/large-electron-positron-collider

[19] *LHC Design Report*, CERN 2004-003 (2004)

[20] *The Nobel Prize in Physics 2013*, http://www.nobelprize.org/nobel_prizes/physics/laureates/2013

[21] R. Graham, T. Tél, *Potential for the Complex Ginzburg-Landau Equation*, Europhys. Lett. **13** (1990) 715

[22] *The Tevatron* http://www.fnal.gov/pub/tevatron/

[23] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S0B003

[24] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004

[25] M. Goebel, *Status of the global fit to electroweak precisions data*, arXiv:1012.1331v1 [hep-ph]

[26] *The LEP Electroweak Working Group*, http://lepewwg.web.cern.ch/LEPEWWG

[27] The TEVNPH Working Group for the CDF and D0 Collaborations, *Combined CDF and D0 Upper Limits on Standard Model Higgs-Boson Production with up to 6.7 fb^{-1} of Data*, arXiv:1007.4587 [hep-ex]

[28] The TEVNPH Working Group, *Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production woth up to 10.0 fb^{-1} of Data*, FERMILAB-CONF-12-318-E, CDF Note 10884, D0 Note 6348 (2012), arXiv:12007.0449 [hep-ex]

[29] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections

[30] G. Cowan, *Statistical Data Analysis*, Oxford Science Pub. (1998)

[31] The 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

[32] The 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

[33] The CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, arXiv:1411.3441 [hep-ex]

[34] *CERN* http://home.web.cern.ch

[35] The LHCb Collaboration, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030 (2003)

[36] S. Beole et al., *ALICE technical design report: Detector for high momentum PID*, CERN-LHCC-98-19

[37] https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults

[38] I. Bid, *Computing for the Large Hadron Collider*, Ann. Rev. of Nuclear and Particle Science **61** (2011) 99

[39] The LCG TDR Editorial Board, *The LHC Computing Grid: Technical Design Report*, CERN/LHCC 2005/024, LCG Technical Design Report 1 (2005)

[40] *WLCG* http://home.web.cern.ch/about/computing/worldwide-lhc-computing-grid

[41] The CMS Collaboration, *The CMS Physics Technical Design Report, Volume I: Detector Performance and Software*, CERN/LHCC 2006/001, CMS Technical Design Report 8.1, CERN (2006)

[42] The CMS Collaboration, *The CMS Physics Technical Design Report, Volume II: Physics Performance*, CERN/LHCC 2006/021, CMS Technical Design Report 8.2, CERN (2006)

[43] The CMS Collaboratio, *The Magnet Projet Technical Design Report*, CERN/LHCC 97/10, CMS Technical Design Report 1, CERN (1997)

[44] The CMS Collaboration, *The CMS tracker system project: Technical Design Report*, CERN/LHCC 1998/006, CMS Technical Design Report 5, CERN (1998); CMS Technical Design Report 5, Addendum CERN/LHCC 2000/016 (2000)

[45] The CMS Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*, CERN/LHCC 1997/033, CMS Technical Design Report 4 (1997)

[46] W. R. Leo, *Techniques For Nuclear And Particle Physics Experiments*, Springer-Verlag (1994)

[47] The CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*, CERN/LHCC 1997/031, CMS Technical Design Report 2 (1997)

[48] The CMS Collaboration, *The CMS muon project: Technical Design Report*, CERN/LHCC 1997/032, CMS Technical Design Report 3 (1997)

[49] M. Andlinger et al., *Bunch crossing identification at lhc using a mean timer technique*, Nucl. Instr. Meth. A **336** (1993) 91

[50] The CMS Collaboration, *The TriDAS Project Technical Design Report, Volume 1: The Trigger Systems*, CERN/LHCC 2000/38, CMS Technical Report 6.1 (2000)

[51] The CMS Collaboration, *The TriDAS Project Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger*, CERN/LHCC 2002/26, CMS Technical Report 6.2 (2002)

[52] The CMS Collaboration, *Technical Proposal for the upgrade of the CMS Detector though 2010*, CMS UG–TP–1 (2011)

[53] The CMS Collaboration, *CMS Technical Design Report for the Level-1 Trigger Upgrade*, CMS-TDR-12 (2012)

[54] The CMS Collaboration, *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*, CMS-TDR-010 (2012)

[55] The CMS Collaboration, *CMS Technical Design Report for the Pixel Detector Upgrade*, CMS-TDR-11 (2012)

[56] CMS Tracker Collaboration, *Phase II tracker upgrade: beam test with prototype 2S modules*, in CMS Tracker week (2014)

[57] C. C. Foudas, A. Rose, J. Jones, G. Hall, *A Study for a Tracking Trigger at First Level for CMS at SLHC*, arXiv:0510227

[58] G. Dissertori et al., *A study of high-energy proton induced damage in cerium fluoride in comparison with measurements in lead tungstate calorimeter crystals*, Nucl.Instrum.Meth. **622** (2010) 41, doi:http://dx.doi.org/10.1016/j.nima.2010.07.052

[59] A.C. Benvenuti et al., *An electromagnetic shashlik calorimeter with longitudinal segmentation*, Nuclear Instruments and Methods in Physics Research A **432** (1999) 232

[60] C. Adloff et al., *Response of the CALICE Si-W electromagnetic calorimeter physics prototype to electrons*, Nucl.Instrum.Meth. A **608** (2009) 372, doi:10.1016/j.nima.2009.07.026

[61] D. Abbaneo, M. Abbrescia, M. Alfonsi, C. Armaingaud, P. Aspell, M. G. Bagliesi, Y. Ban, S. Bally et al., *An overview of the design, construction and performance of large area triple- GEM prototypes for future upgrades of the CMS forward muon system*, JINST **7** (2012) C05008

[62] Y. Haddad, G. Grenier, I. Laktineh, N. Lumb and S. Cauwenbergh, *High Rate Resistive Plate Chamber for LHC de- tector upgrades*, Nucl. Instrum. Meth. A **718** (2013) arXiv:1211.5698 [physics.ins-det]

[63] *CMSSW* https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBook

[64] A. K. Nayak, *Reconstruction of physics objects in the CMS detector*, arXiv:1310.7408 [hep-ex]

[65] P. Azzurri, *Track Reconstruction Performance in CMS*, arXiv:0812.5036 [physics.ins-det] (2008)

[66] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444, doi:10.1016/0168-9002(87)90887-4

[67] CMS Collaboration, *The CMS Particle Flow Algorithm*, arXiv:1401.8155

[68] The CMS Collaboration, *Muon reconstruction and identification in CMS*, AIP Conf.Proc. **1200** (2010) 701

[69] The CMS Collaboration, *Electron reconstruction and identification at* $\sqrt{s} = 7$ *TeV*, CMS Physics Analysis Summary CMS PAS EGM-10-004 (2010)

[70] T. Speer, R. Frẗwirth, *A Gaussian-Sum Filter for Vertex Reconstruction*, CERN-CMS-NOTE-2005-005 (2005)

[71] B.P. Roe et. al., *Boosted decision trees as an alternative to artificial neural networks for particle identification*, NIM A543 (2005) 577

[72] F. Abe et al., *Properties of events with large total transverse energy produced in proton-antiproton collisions at* $\sqrt{s} = 1.8$ *TeV*, Phys. Rev. D **45** (1992) 1448

[73] G. C. Blazey et al., *Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop*, arXiv:0005012 [hep-ex]

[74] S. Catani, Yu. L. Dokshitzer, M. H. Seymour, B. R. Webber, *Longitudinally-invariant k_T-clustering algorithms for hadron-hadron collisions* Nuclear Physics B **406** (1993) 187

[75] M. Cacciari, G. P. Salam, G. Soyez, *FastJet User Manual*, Eur.Phys.J. C **72** (2012) 1896

[76] T. Sjostrand, S. Mrenna, P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv:hep-ph/0603175

[77] T. Sjostrand, S. Mrenna, P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **178** (2008) 852-867, arXiv:0710.3820 [hep-ph]

[78] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. -S. Shao, T. Stelzer, P. Torrielli, M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **1407** (2014) 079, arXiv:1405.0301 [hep-ph]

[79] The POWHEG box *http://powhegbox.mib.infn.it*

[80] GEANT4 Collaboration (S. Agostinelli et al.), *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250-303, DOI: 10.1016/S0168-9002(03)01368-8

[81] S. Banerjee, *CMS Simulation Software*, J. Phys. Conf. Ser **396** (2012) 022003

[82] S. Ovyn, X. Rouby, V. Lemaître, *Delpes, a framework for fast simulation of a generic collider experiment*, CP3-09-0, arXiv:0903.2225 [hep-ph]

[83] DELPHES 3 Collaboration (J. de Favereau et al.), *Delpes 3: A modular framework for fast simulation of a generic collider experiment*, JHEP **1402** (2014) 057, arXiv:1307.6346 [hep-ex]

[84] S.V. Chekanova, E. Maya, K. Strandb, P. Van Gemmerena, *ProMC: Input-output data format for HEP applications using varint encoding*, arXiv:1311.1229 [physics.comp-ph]

[85] M. Dobbs, J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics* Comput. Phys. Commun. **134**, 41 (2001)

[86] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos, A. Buckley, J.M. Butterworth, L. Dudko, S. Frixione, L. Garren et al. *A Standard format for Les Houches event files*, Comput. Phys. Commun. **176** (2007) [hep-ph/0609017]

[87] I. Antcheva, M. Ballintijn, B. Bellenot at al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Comm. **180** (2009) 2499

[88] *ROOT: Data Analysis Framework*, https://root.cern.ch/drupal/

[89] M. Wobisch, T. Wengler, *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering*, arXiv:9907280 [hep-ph]

[90] M. Cacciari, G. P. Salam, G. Soyex, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063

[91] The CMS Collaboration, *Measurement of the properties of a Higgs boson in the four lepton final state*, Phys. Rev. D 89 (2014) 092007 arXiv:1312.5353 [hep-ex]

[92] The CMS Collaboration, *Studies of the mass and spin-parity of the Higgs boson candidate via its decays to Z boson pairs*, Phys.Rev.Lett. **110** (2013) 8, 081803 arXiv:1212.6639 [hep-ex]

[93] The CMS Collaboration, *Studies of the performances of the $H \to ZZ^* \to 4\mu$ analysis with the CMS Phase II detector upgrade*, CMS-PAS-FTR-13-003

[94] *Cut Based Electron ID* https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaCutBasedIdentification