UNIVERSITÀ DEGLI STUDI DI BARI

DOCTORAL THESIS

DOTTORATO IN FISICA XXX CICLO

Search for Dark Matter Produced in Association with a Higgs Boson in the four lepton final state at 13 TeV

Author:
Giorgia MINIELLO

Supervisor:
Prof. Giorgio Pietro MAGGI
Dr. Nicola DE FILIPPIS

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Dipartimento Interateneo di Fisica "Michelangelo Merlin"

Anno 2017

"My mother made me a scientist without ever intending to. Every other Jewish mother in Brooklyn would ask her child after school: So? Did you learn anything today? But not my mother. "Izzy," she would say,"did you ask a good question today?" That difference, asking good questions, made me become a scientist."

Isidor Isaac Rabi

UNIVERSITÀ DEGLI STUDI DI BARI

Abstract

Dipartimento Interateneo di Fisica "Michelangelo Merlin"

Doctor of Philosophy

Search for Dark Matter Produced in Association with a Higgs Boson in the four lepton final state at 13 TeV

by Giorgia MINIELLO

After the discovery of the Higgs boson in 2012 made by the ATLAS and CMS experiments together at the Large Hadron Collider (LHC), a fundamental milestone of the Standard Model has been established. Nevertheless, there are other fundamental physics phenomena not yet explained by this model such as *gravity*, *dark matter* and *dark energy*, *neutrino oscillations*, and *matter and anti-matter asymmetry*. In particular, although most of the matter in the Universe is dark matter, its nature remains still unknown. Even though cosmological evidence for dark matter has been confirmed, so far no dark matter evidence and also no evidence of non-gravitational interactions between dark matter and Standard Model particles has been observed in particle physics experiments. If these kinds of interactions exist in nature it can be possible in principle to observe them at colliders. Since dark matter particles themselves cannot produce any signal in the LHC detectors, one possible way to detect them is to study their associated production with a *visible* SM particle as the Higgs boson. Such kind of reactions in which a SM-like Higgs recoils against an *invisible* state are called "Mono-Higgs" reactions, in which the final signature to be look for is an associated production of the Higgs boson with a large value of missing transverse energy (MET) that should provide the imprint for possible particle-antiparticle dark matter candidates. The aim of this thesis is indeed to examine the idea that, If the dark matter is associated with the scale of the electroweak symmetry breaking, the Higgs can be used as a probe to search for dark matter candidates at colliders. Latest results from a search for dark matter candidates using a SM-like Higgs decaying into four leptons ($H \rightarrow ZZ \rightarrow 4l$, with $l = e, \mu$) as a probe will be presented. The experimental final signature for this analysis provides for the four leptons produced by the Higgs decay along with a large value of MET. The study has been performed using data collected during 2016 at an integrated luminosity of 35.9 fb^{-1} at a center of mass energy of 13 TeV with the CMS experiment at LHC. Two simplified models have been used for interpreting data, the $Z'2HDM$ and the $Z'Baryonic$ model. 95% CL upper limits on the cross section for both the models have been derived. Even though no evidence of dark matter candidates has been found, a good improvement in the sensitivity of the channel with respect to previous 2015 studies has been observed. This improvement opens a quite exciting scenario, especially in the prospective of a significant luminosity increase to be expected during next years at LHC, since it ensures that, despite the low branching fraction of the $H \rightarrow ZZ \rightarrow 4l$ channel, this Mono-Higgs search, as a result of the efforts spent to optimize the cut-based analysis and the definition of the control and the signal regions presented in this work, will be able to exclude some points of the parameter space just adding the statistics acquired by CMS with the 2017 data, encouraging further studies using also different benchmark models.

Acknowledgements

…

Contents

To my beloved daughter Zoe.
Sometimes in a life, my love, hearts, like symmetries,
can be broken, but this could be the only way for
good things to start getting mass...

Introduction

After the discovery of the Higgs boson in 2012 during the first operational run (2009 - 2013) of the Large Hadron Collider (LHC), also known as Run I, a fundamental milestone of the Standard Model has been established.
Nevertheless, there are other fundamental physics phenomena not adequately explained by this model such as *gravity*, *dark matter* and *dark energy*, *neutrino oscillations*, and *matter and anti-matter asymmetry*.
In particular, although most of the matter in the Universe is dark matter, its nature remains still unknown.
Other models in connection with New Physics, *Beyond Standard Model* [8], are necessary to expand the physics scenario and propose possible dark matter candidates at the electroweak symmetry breaking scale [9, 10]. Hence, the discovery of the Higgs boson opened new opportunities to dark matter search; in terms of importance, dark matter discovery would represent the challenge of RUN II as the Higgs boson discovery was for Run I.
The existence of a Dark Matter component of the universe is now firmly established, due to the observation of gravitational effects both on astrophysical and cosmological scales [34]. So far, the most accredited paradigm provides for a *cold* Dark Matter made up of some weakly interacting massive particles (WIMPs).
The WIMP theory remains a very successful one for a series of reasons, such as the great number of theoretical models which support it and the relic abundance of DM which can be obtained through the thermal freeze-out mechanism for reasonable ranges of WIMP's mass and annihilation cross section [58].
Nevertheless, so far no dark matter evidence and also no evidence of non-gravitational interactions between dark matter and Standard Model particles has been observed in particle physics experiments. If these kinds of interactions exist in nature it can be possible in principle to observe them at colliders. Since Dark Matter particles themselves cannot produce any signal in the LHC detectors, one possible way to detect them is to study their associated production with a *visible* SM particle X (where X stands for a jet, a photon, a quark, a vector boson Z or W, or a Higgs). Such reactions in which particles or jet recoil against an invisible state are called "Mono-X" reactions in which the final signature consists of X plus a large value of missing transverse momentum (a.k.a. MET) that can be observed in the detector.
The aim of this thesis is indeed to examine the idea that, if the dark matter is associated with the scale of the electroweak symmetry breaking, the Higgs can be used as a probe to search for dark matter candidates at colliders [11]. Then, the X to be looked for is a Higgs boson associated with large value

of MET, which should provide the signature for possible dark matter candidates.
The principle of a so-called "Mono-Higgs" search is to use the SM Higgs as a *probe* for tagging dark matter events taking the advantage of a direct coupling DM-SM due to the small coupling of the Higgs to quarks (see Chapter 1, Section 1.2). The Mono-Higgs search presented here aims at analysing events registered by the CMS experiment in which a Higgs boson decays into a four leptons via the channel $H \rightarrow ZZ \rightarrow 4l$, where the 4l combination considered can be 4μ, $4e$ or $2e2\mu$. The first chapter of this work will show a theoretical overview of the Standard Model along with the two simplified models, $Z'2HDM$ and $Z'Baryonic$, used for the analysis presented.
The second Chapter will provide a description of the LHC experiment and the CMS (Compact Muon Solenoid) detectors, while in the third chapter the data and Monte Carlo simulation samples for signals and background will be described.
The fourth chapter will provide an overview of all the main physics object used along with their distributions, then in the fifth chapter the real analysis will be described in detail. Finally, the sixth chapter will present all the results of this study.

Chapter 1

Theoretical Models for Dark Matter Search: Beyond the Standard Model

In this first chapter the theoretical models considered for this analysis will be presented. First section 1.1 will provide a quick overview of the Standard Model, then section 1.2 will introduce the main features of a Mono-Higgs search. In section 1.3, the necessity of focusing just on particular kinds of models for DM search at colliders will be discussed along with the latest Dark Matter search results at colliders. Finallly, in section 1.4 the two models used to simulate the signal for the analysis studied in this thesis will be presented. Even though Dark Matter is not expected in the Standard Model, we look for a possible extension of the SM with a final signature consisting of a SM-like Higgs boson plus a couple of particle-antiparticle DM candidates in order to look for possible extension of the Standard Model itself.

1.1 Introduction to the Standard Model

The theory about the Standard Model [2] was stated for the first time by Weinberg and Salam and later completed by 't Hooft. This theory soon turned out to be a powerful one, being the only model able to explain a wide variety of phenomena in nature. Through many decades and a lot of experiments, this model has become the one which has the greatest number of experimental confirmations. It was developed starting from the need to find a single unified symmetry group describing both electromagnetic and weak interactions. The main constituents of the SM are summarised in Figure 1.1.
The elementary particles are divided in subgroups: tre generations of quarks, three lepton families, four gauge bosons and the outstanding "young" Higgs boson. They are all characterized by different mass, charge, spin and quantum number (see Ref. [1]).
The electroweak symmetry group for the Standard Model is:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

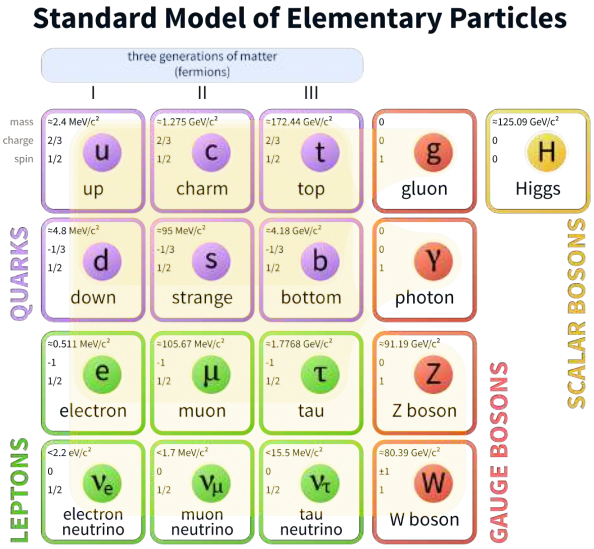


FIGURE 1.1: Elementary Particles from the Standard Model

where $SU(3)_Y$ is the symmetry group fro the strong interaction generated by the operator C, $SU(2)_L$ is the left-handed component of the symmetry group for the weak current and $U(1)_Y$ is the symmetry group generated by the hyper-charge operator Y which includes the electromagnetic interactions. The eight colored gluons are the mediator of the strong force, they have no mass and no charge. The charged bosons, $W^{\pm}$, and the neutral one, Z, are massive and are the mediators of the weak force, while the photon is the mediator of the electromagnetic interaction. It is neutral and massless. All the gauge boson mediator have spin 1.
All quarks and leptons are fermions of spin 1/2. Furthermore, they are massive. In order to give mass to gauge bosons and fermions, an electroweak spontaneous symmetry breaking (EWSB) is predicted in the SM gauge symmetry. After the EWSB the SM gauge symmetry group 1.1 follows the pattern:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \Rightarrow SU(3)_C \otimes U(1)_{em}. \tag{1.2}$$

The spontaneous electroweak symmetry breaking ensures to give mass to fermions and bosons and to keep photons massless while generating a new massive (m_H=125 GeV/c) spinless scalar, the Higgs boson, whose existence was theorised by Peter Higgs and Francois Englert and Robert Brout independently. This mechanism has also the great merit To delete from the SM lagrangian the unwanted term associated to a massless scalar particle, actually not observed in nature, known as the *Goldstone boson*.
The complete version of the SM lagrangian can be written as follows:

$$\mathcal{L} = -\tfrac{1}{4}\mathbf{W}_{\mu\nu}\cdot\mathbf{W}^{\mu\nu} - \tfrac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad \left\{ \begin{array}{l} \mathrm{W}^{\pm},\mathrm{Z},\gamma \text{ kinetic} \\ \text{energies and} \\ \text{self-interactions} \end{array} \right.$$

$$\begin{array}{l} +\bar{L}\gamma^{\mu}\left(i\partial_{\mu} - g\tfrac{1}{2}\boldsymbol{\sigma}\cdot\mathbf{W}_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)L \\ +\bar{R}\gamma^{\mu}\left(i\partial_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)R \end{array} \quad \left\{ \begin{array}{l} \text{lepton and quark} \\ \text{kinetic energies} \\ \text{and their} \\ \text{interactions with} \\ \mathrm{W}^{\pm},\mathrm{Z},\gamma \end{array} \right.$$

$$+\left|\left(i\partial_{\mu} - g\tfrac{1}{2}\boldsymbol{\sigma}\cdot\mathbf{W}_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)\phi\right|^{2} - V(\phi) \quad \left\{ \begin{array}{l} \mathrm{W}^{\pm},\mathrm{Z},\gamma, \text{ and Higgs} \\ \text{masses and} \\ \text{couplings} \end{array} \right.$$

$$-(G_1\bar{L}\phi R + G_2\bar{L}\phi_c R + \text{hermitian conjugate}). \quad \left\{ \begin{array}{l} \text{lepton and quark} \\ \text{masses and} \\ \text{coupling to Higgs} \end{array} \right.$$

where L denotes a left-handed fermion (lepton or quark) doublet, and R denotes a right-handed fermion singlet.
For more details on each term of the lagrangian, see Ref. [2].
The Higgs boson was actually discovered in 2012 at CERN by the ATLAS [7] and the CMS [6] collaborations and in 2013 both Englert and Higgs were awarded with the Nobel Prize in Physics. More details on the Standard Model electroweak spontaneous symmetry breaking can be found in Refs. [3], [4], [5].

1.2 The Mono Higgs Search

As already introduced in the preliminary section of this thesis, a "Mono-X" search looks for final states whose signature is characterized by a SM particle X plus large $\not{E}_T$ values. In these signatures, X can be identified as a jet [12,13], a t/b quark [14–16], a photon [17–19], a lepton [20,21], a W/Z boson [22–24], or a Higgs boson.
Recently, the production of Dark Matter candidates in association with a top quark pair [28,29], a boosted top quark in the hadronic final state [30], a vector boson [31], a photon [32], have been investigated and latest results have been published. As already pointed, no evidence of Dark Matter at colliders has been found so far, but the Higgs boson discovery opened definitely a new interesting window for searching DM at the LHC via the Mono-Higgs signature.
While in other Mono-X signatures, the "X" can be emitted as initial state radiation (ISR) or from a new physics vertex coupling DM to Standard Model particles, in Mono-Higgs final states the ISR is highly suppressed due to the small coupling of the Higgs to quarks. Therefore, one Higgs boson is preferentially emitted from the new physics vertex. This would directly probe the

effective DM-SM coupling.
Figure 1.2 shows the Feynman diagram of Mono-Higgs production in proton-proton collisions having an electroweak boson (H, Z, γ), a new vector mediator particle (referred as $Z\prime$ in the figure) or a scalar singlet (referred as S in the fiigure). The interaction between DM and the Higgs is represented by the grey circle.

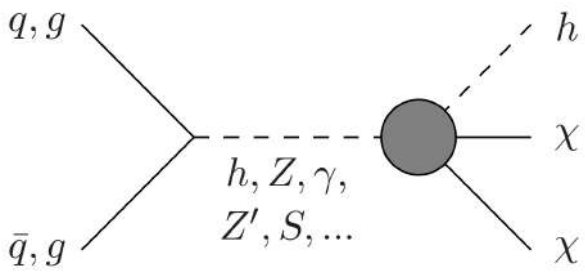


FIGURE 1.2: Feynman diagram of Mono-Higgs production in proton−proton collisions mediated by electroweak bosons, new mediator particles or scalar singlet.

1.3 Dark Matter Benchmark Models and the ATLAS/CMS Dark Matter Forum

The study of the Dark Matter search at colliders can be considered a complementary approach with respect to other strategies like *direct searches* of DM looking for scattering events of DM with heavy nuclei and *indirect searches* in which the final products of DM annihilations are detected in the galaxy or in the Sun, such as gamma-rays or neutrinos. The Feynman diagram which summarises the three strategies is shown in Figure 1.3, where M stands for "mediator", while g_q and g_χ are the coupling constants of the mediator with quarks and DM particles respectively.

The synergy that comes from the complementarity of these different approaches opens the real possibility to improve the discovery potential using all of them and, consequently, the definition of a common language into which the results can be interpreted and compared becomes mandatory.
For the Mono-X analyses two main theoretical model approaches have been considered: the Effective Field Theory (EFT) and the Simplified Models [33]. The exploration of new model-indipendent approaches to DM searches at colliders exploited the use of the EFT which describes the unknown DM interaction with the SM particles.
Unfortunately, the use of EFT in the context of DM search at colliders has been questioned and limitations have been put on it, mainly due to the hierarchy problem and suppression of the non-renormalizable operators by the

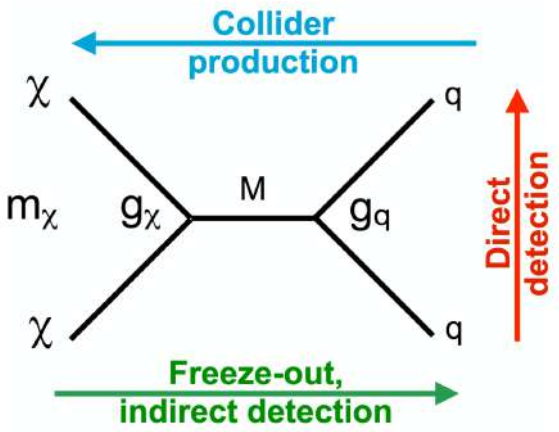


FIGURE 1.3: Feynman diagram summarising the three search strategies. Here M stands for "mediator", while g_q and g_χ are the coupling constants of the mediator with quarks and DM

powers of $1/\Lambda$, where Λ stands for the effective energy scale of the ultraviolet (UV) particles which are integrated [11, 35, 36].
The use of the so-called *simplified models* stands exactly in the middle between the effective-operator approximation and complete ultraviolet models [33, 38, 39]. In simplified models, the UV particles are kept as degrees of freedom [11, 33] and vector (axial) or scalar (pseudoscalar) *mediators* plays a fundamental role in the interaction between DM candidates and particles from the SM model sector.
Although more model-dependent than the EFT approach, simplified models ensure to be more reliable when the typical parton energies in the events are comparable to Λ [41] and to be not *blind* to possible constraints on ultraviolet physics involved in the generation of the corresponding model operators [11].
The pressing need to establish new official guidelines about the choice of new models for DM search studies, both for the ATLAS and the CMS experiments, brought to the birth of the ATLAS/CMS Dark Matter Forum, which summarized all the Dark Matter benchmark models for Early LHC Run-2 searches in an official public report [42].
The very latest complete summary of the DM searches in CMS has been presented during the LHCP 207 Conference []. Figure 1.4 shows the 95% Confidence Level upper limits (see Chapter 6) for scalar (left) and pseudoscalar (right). Following the recommendation of the LHC DM working group , the exclusions are computed the coupling constant to quark is $g_q = 1$ and the coupling constant to DM is $g_{DM} = 1$. For the pseudoscalar mediator the dark matter search in jets+MET final state is the most sensitive, since the corresponding expected upper limits are lower respect with the theoretical ones for a wide range of the mediator mass. Figure 1.5 shows the 90% CL upper limits. The couplings to quark and Dark Matter are $g_q = 0.25$ and $g_{DM} = 1$ respectively.
The CMS exclusion contour is compared with the LUX 2016, PandaX-II 2016,

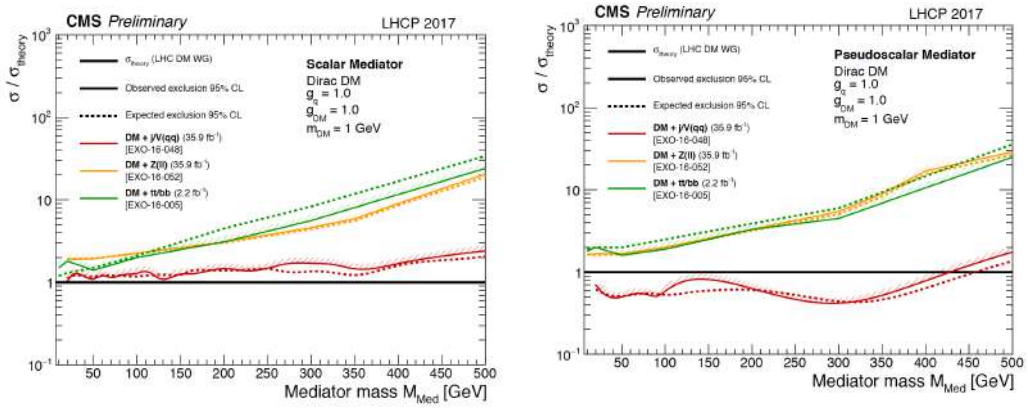


FIGURE 1.4: 95% CL observed (full-line) and expected (dashed-line) exclusion limits for the Scalar (left) and Pseudoscalar (right) model as a function of the mediator mass for different missing transverse energy based DM searches from CMS.

CDMSLite 2015 and CRESST-II 2015 limits for the vector mediators and with the PICASSO, PICO experiments, the IceCube limit for the $t\bar{t}$, $b\bar{b}$ annihilation channels and the Super-Kamiokande limit for the $b\bar{b}$ annihilation channel for the axial-vector mediators.
These constitute the strongest documented constraints obtained from *direct* DM searches. As it can be observed from the plots the DM search at colliders is much more sensitive with respect to cosmological dark matter experiments.

1.4 Signal Benchmarks

The analysis reported in this thesis is a Mono-Higgs analysis, with the Higgs boson decaying as follows:

$$\mathrm{H} \rightarrow ZZ \rightarrow 4l$$

where $l = e, \mu$.
For this study, two simplified signal models have been chosen as benchmarks:

- the Z' - Two Higgs Doublet Model ($Z'2HDM$) simplified model,
- the Z'Baryonic model.

In the following subsections, some theoretical details about these two models will be presented.

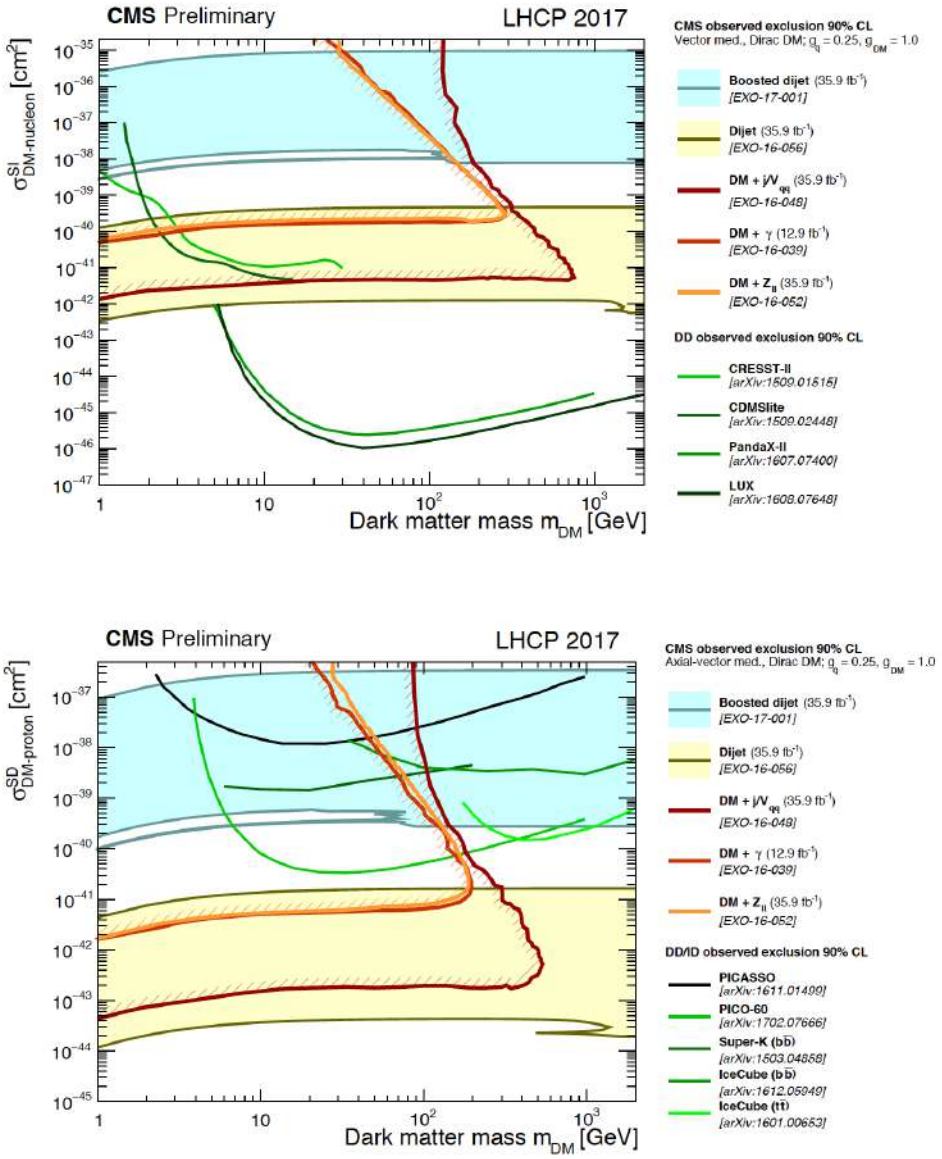


FIGURE 1.5: A comparison of CMS results to the $m_{DM} - \sigma_{SI}$ plane for vector mediators (top) and to $m_{DM} - \sigma_{SD}$ plane for axial-vector mediators.

1.4.1 Z′ - Two Higgs Doublet Model (Z'2HDM) simplified model

The first signal model used in this analysis is the Z′ - Two Higgs Doublet Model ($Z'2HDM$) simplified model. In this model, the SM gauge group is extended including a new symmetry, $U(1)_{Z'}$, with the Z′ as the corresponding gauge boson [43].
This symmetry is spontaneously broken by a scalar singlet ϕ, generating a Z′ mass above the EW symmetry breaking scale. The right-handed quarks are charged under $U(1)_{Z'}$ and all other SM particles are neutral. The constraints

on the Z' coupling to quarks g_z are defined by EW global fits [44] and di-jet resonance searches [48,49] so that

$$g_z < 0.03 \frac{g}{\cos\theta_w \sin^2\beta} \frac{\sqrt{M_{Z'}^2 - M_Z^2}}{M_Z}. \tag{1.3}$$

where $M_{Z'}$ represents the mass of Z', while, additionally, a second Higgs doublet is added, introducing states Φ_u and Φ_d, which couple to up and down quarks respectively as

$$\mathcal{L} \supset -y_u Q \bar{\Phi}_u \bar{u} - y_d Q \Phi_d \bar{d} + y_e L \Phi_d \bar{e} + h.c.. \tag{1.4}$$

Φ_u is chosen to be charged under $U(1)_{Z'}$, while Φ_d is neutral. The two Higgs doublets can be parametrized as follow

$$\Phi_d = \frac{1}{\sqrt{2}} \begin{pmatrix} -\sin(\beta) H^+ \\ v_d - \sin(\alpha) h + \cos(\alpha) H - i \sin(\beta) A^0 \end{pmatrix} \tag{1.5}$$

$$\Phi_u = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(\beta) H^+ \\ v_u + \cos(\alpha) h + \sin(\alpha) H + i \cos(\beta) A^0 \end{pmatrix} \tag{1.6}$$

where h and H are neutral CP-even scalars and A^0 is CP-odd, and v_u and v_d are the two vacuum expectation values after EW symmetry breaking. The angle α is defined as the angle that diagonalizes the $h - H$ mass mixing matrix, and the angle β is defined as $\tan(\beta) = v_u / v_d$. h is assumed to be the SM Higgs boson with mass $m_H = 125$ GeV, while the other scalars have masses 300 GeV. Due to perturbativity and previous constraints, α and β are chosen such that $\tan(\beta) > 0.3$ and $\alpha = \beta - \pi/2$.
Mono-Higgs signals arise when the pseudoscalar A^0 has a large branching ratio to DM. The Feynman diagram leading to Higgs+E_{miss}^T production via A^0 is sketched in Figure 1.6. As shown, the pseudoscalar A^0 comes from the decay of the *on-shell* resonance Z' [42]. The new particles and parameters of the Z'2HDM model are summarized in Table 1.1. For this model, the reference values for m'_Z and m_{A_0} are reported in Table 1.1, Chapter 3.

1.4.2 Z'Baryonic Model

In the Z'Baryonic simplified model, the gauge group of the SM is extended to include a new symmetry, $U(1)_B$ for the baryon number B, with the Z' being the gauge boson of $U(1)_B$ [50–52].
Z' couples to quarks and fermionic DM as

$$\mathcal{L} \supset g_q \bar{q} \gamma^\mu q Z'_\mu + g_\chi \bar{\chi} \gamma^\mu \chi Z'_\mu \tag{1.7}$$

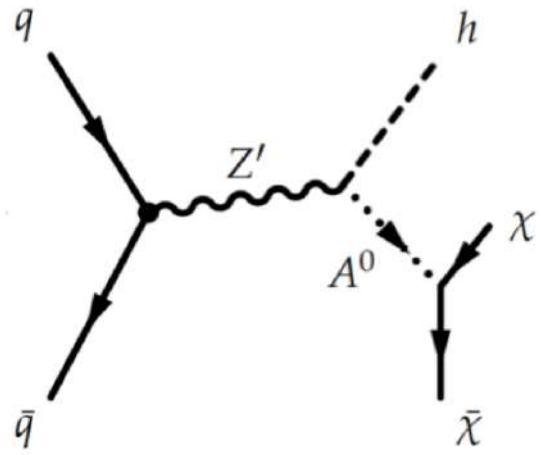


FIGURE 1.6: Feynman diagram for Higgs + E_miss^T production via a new invisibly decaying pseudoscalar A^0

Particle	Description	
χ	Fermionic DM particle	
Z′	$U(1)_{Z'}$ gauge boson	
ϕ	Z′ sector scalar	
Φ_u, Φ_d	Two Higgs doublets	
h, H	Neutral CP-even scalars	
$H^{\pm}$	Charged heavy Higgs	
A^0	Neutral CP-odd pseudoscalar	
Param.	Description	Value
m_χ	DM mass	100 GeV
$m_{Z'}$	Z′ mass	see Table 3.3
m_{A_0}	A^0 mass	see Table 3.3
g_z	Z′-quark coupling	0.8 or Eq. 1.3
$\tan\beta$	$\Phi_{u/d}$ VEV angle	1

TABLE 1.1: Z'2HDM simplified model: parameter description and values used.

The $U(1)_B$ simmetry is spontaneously broken by a "baryonic Higgs" scalar h_B, which mixes with the SM Higgs via a mixing angle θ. This mixing induces an h-Z′ interaction $-g_{hZ'Z'}hZ'_\mu Z'^\mu$ with coupling

$$g_{hZ'Z'} = \frac{m^2_{Z'}\sin(\theta)}{v_B} \tag{1.8}$$

where $m_{Z^{prime}}$ is the mass of the Z' and v_B is the vacuum expectation value (vev) of h_B. At energies less than $m_{Z'}$, these operators combine to yield an effective Lagrangian

$$\mathcal{L}_{eff} = -\frac{g_q g_\chi}{m_{Z\{prime}}}^2 \bar{q}\gamma^\mu q\bar{\chi}\gamma_\mu\chi(1+\frac{g_{hZ'Z'}}{m^2_{Z'}}h) \tag{1.9}$$

While the first term gives rise to mono-jet and mono-EW boson interaction terms, the second one corresponds to the Mono-Higgs production diagram shown in Figure 1.7.

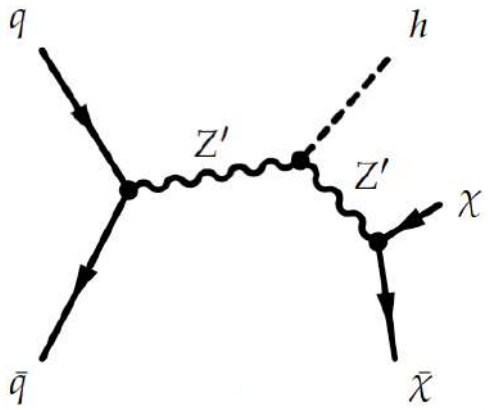


FIGURE 1.7: Collider production diagram for Z′ Baryonic model.

The new particles and parameters of the Z′Baryonic model are summarized in Table 1.2.
The Z′-quark coupling is required to be less than 4π [11] due to perturbativity arguments. The values of the parameters chosen are given in Table 1.2.
For this model, the reference values for m'_Z and m_{A_0} are reported in Table 3.5, Chapter 3.

Particle	Description	
χ	Fermionic DM particle	
Z'	$U(1)_B$ gauge boson	
h_B	Baryonic Higgs	
Param.	**Description**	**Value**
m_χ	DM mass	Table 3.5
$m_{Z'}$	Z′ mass	Table 3.5
g_q	Z′-quark coupling	1/3
g_χ	Z′-DM coupling	1
$g_{hZ'Z'}$	Z′-h coupling	$m_{Z'}$

TABLE 1.2: Z′Baryonic simplified model: parameter description and values used.

Chapter 2

The Large Hadron Collider and the CMS Detector

In this chapter an overview of the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) experiment, whose 2016 collected data have been used in this study, is provided.

2.1 The Large Hadron Collider at CERN

The CERN (European Organization for Nuclear Research) was established in 1954 as the world's largest particle physics laboratory. The acronym CERN originally stood, in french, for *Conseil Europèen pour la recherche Nuclèaire* (European Council for Nuclear Research), which was a provisional council for setting up the laboratory, established by twelve European countries in 1952. It is located in the northwest part of Geneva, on the Franco-Swiss border, and currently the organization includes twentytwo European member states. One of the CERN main function has been to provide the particle accelerators and other infrastructures needed for high-energy physics research.
Since its birth, many experiments have been built at CERN by international collaborations but, at present, most of the activities at CERN are directed towards operating the Large Hadron Collider (LHC), and the experiments for it. The LHC [59, 60] represents a large-scale worldwide scientific cooperation project. It is placed in a tunnel located approximately 100 meters underground, This tunnel was previously occupied by the LEP (The Large Electron-Positron Collider) experiment, which also made important headway in hunting for the Higgs boson. It was closed down in November 2000. The LHC inherited from LEP some advanced devices as the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) accelerator systems (see Figure 2.1).

The acceleration of protons starts from a linear accelerator (LINAC) that injects the protons to the Proton Synchrotron (PS), which accelerates them to 25 GeV. In the following stage, the Super Proton Synchrotron (SPS) accelerates the beams to 450 GeV and then injects them into the LHC ring.

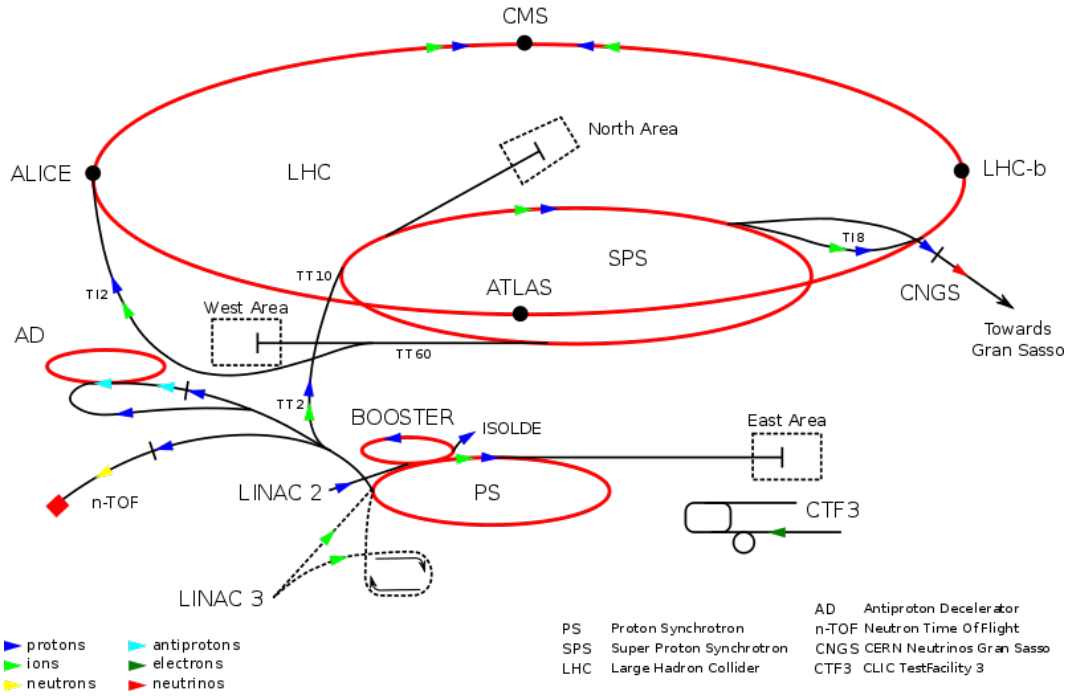


FIGURE 2.1: The LHC accelerator complex.

The main experiments ATLAS, CMS, ALICE and LHCb, are located at the four interaction regions. Two of them, CMS and ATLAS, are particularly focused on the Higgs boson search within the SM context and on physics beyond it. LHC has been designed for two kinds of collisions: collisions of protons, and collisions of heavy ions.

2.1.1 Performance Goals

The LHC was designed to investigate the scalar sector of the Standard Model and also the physics beyond the SM.
The number of events per second of a given physics process is related to the cross section[1] of the corresponding process, via the luminosity $\mathcal{L}$ of the machine, by the following relation

$$N = \mathcal{L}\sigma. \tag{2.1}$$

The relevant events for physics searches, such as Higgs physics and physics beyond the Standard Model, are predicted to have a quite low production cross sections in proton-proton collisions. Figure 2.2 shows that the cross section for the production of a Higgs boson is several orders of magnitude smaller than the total inelastic cross section. Besides, it increases significantly with the center-of-mass energy of the collisions and that is the reason why,

[1] In particle physics, the cross section is used to express the normalized rate or probability of a given particle interaction. It has the dimension of a surface and is usually expressed in barns (b): 1b=10^{-28}m^2.

in order to reach a high event rate, both the collision luminosity[2] and the center-of-mass energy must be as high as possible.

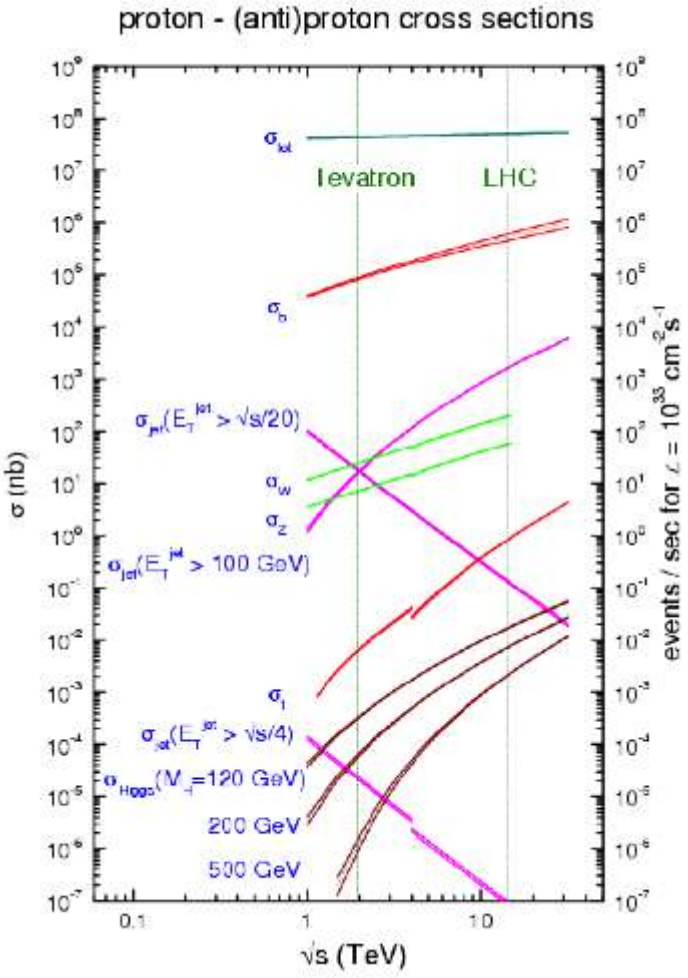


FIGURE 2.2: Expected cross section vs energy

The nominal center-of-mass energy for LHC collisions is $\sqrt{s} = 14$ TeV (7 TeV per beam), and the nominal peak luminosity is $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1} for the CMS and ATLAS experiments.
Since LHC is a proton accelerator with a constrained circumference, the maximal energy per beam is related to the strength of the dipole field that maintains the beams in orbit.
A high technology global magnet system allows to reach the nominal LHC beam energy of 7 TeV. The system uses a total of about 9600 magnets. The 1232 dipole magnets use niobium-titanium (NbTi) cables. By pumping superfluid helium into the magnets, they are brought to a temperature of 1.9 K . For this purpose, a total of 120 t of superfluid helium is used.
At that temperature, the dipoles are in a superconducting state and a field of 8.33 T can be provided. Such a magnetic field is necessary to bend the 7 TeV

[2]The luminosity is defined as the number of the particles per unit area per unit time times the opacity of the target. It is usually expressed in $cm^{-2}s^{-1}$

beams around the 27-km ring of the LHC.
Among the other magnets, quadrupoles play a major role at collision points: they are used to focus the beam, and maximize the probability of collision. The very high LHC design luminosity implies many constraints on the proton beam parameters.
The nominal luminosity can be reached with a number of bunches per beam $n_b = 2.808$ and a number of protons per bunch $N_b = 1.15 \cdot 10^{11}$.
Such a high beam intensity could not be reached with antiproton beams, hence a simple particle-antiparticle accelerator collider configuration cannot be used at LHC.
The LHC has therefore been designed with two separate rings. The common sections are located at the insertion regions, which are equipped with the experimental detectors. The configuration is presented in Figure 2.3.

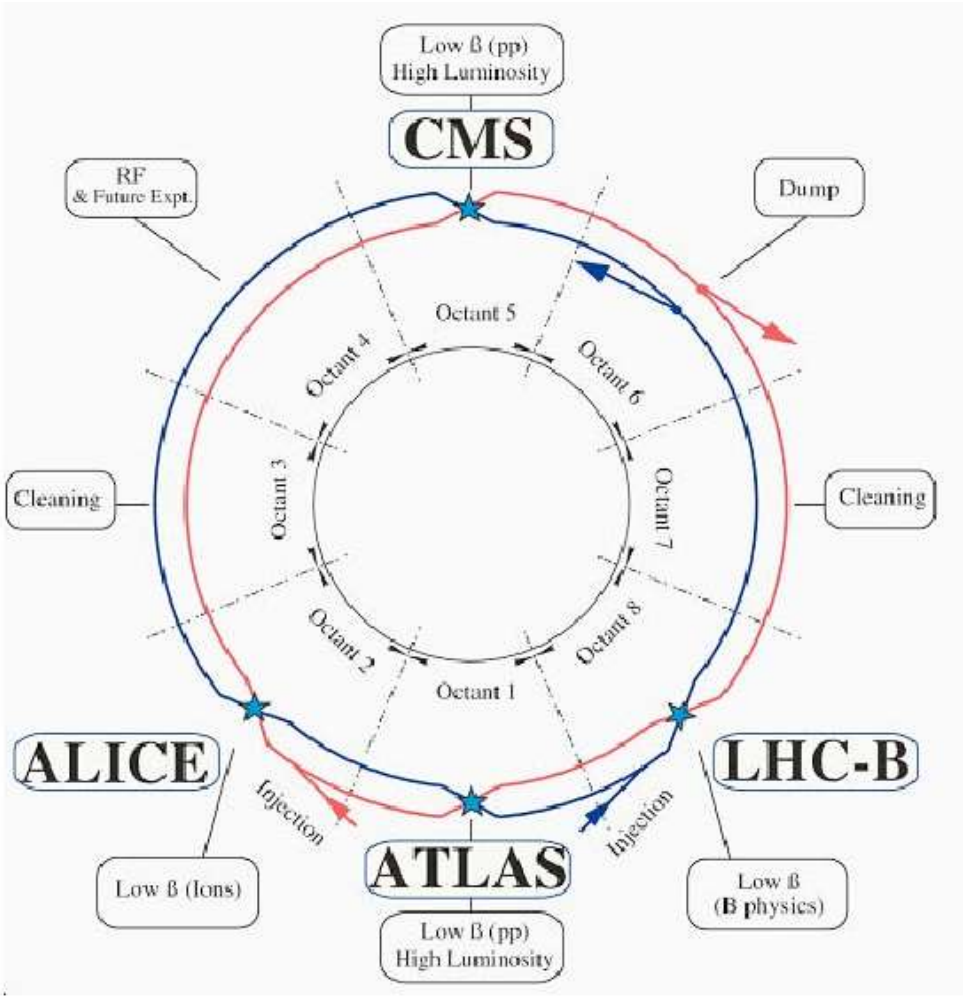


FIGURE 2.3: Schematic layout of LHC (Beam1-clockwise, Beam2-anticlockwise).

A summary of the machine parameters [61] is given in Table 2.1. The numbers indicated correspond to the nominal values. In addition to the previously mentioned parameters, the luminosity lifetime is an important parameter at LHC and colliders in general. The luminosity tends to be reduced during a physics run, because of the degradation of intensities and emittances of the circulating and colliding beams.

Cironference	26.659 km
Center-of-mass energy ($\sqrt{s}$)	14 TeV
Nominal Luminosity ($\mathcal{L}$)	10^{34} cm^{-2}s^{-1}
Luminosity lifetime	15 hrs.
Time between two bunch crossings	24.95 ns
Distance between two bunches	7.48 m
Longitudinal max. size of a bunch	7.55 cm
Number of bunches (n_b)	2808
Number of protons per bunch (N_b)	1.15×10^{11}
beta function at impact point ($\beta^\star$)	0.55 m
Transverse RMS beam size at impact point ($\sigma^\star$)	16.7 $^-$m
Dipole field at 7 TeV (B)	8.33 T
Dipole temperature (T)	1.9 K

TABLE 2.1: List of the nominal LHC parameters, for proton-proton collisions, relevant for the detectors.

2.1.2 LHC Collision Detectors

Reaching the high luminosity values previously discussed imposes tight constraints on the design of the detectors. Under nominal conditions, the LHC can produce 10^9 inelastic collision[3] events per second.
The corresponding bunch crossing rate is 40 MHz (i.e. a bunch crossing spacing of 25 ns), with $\sim$ 20 collisions events expected per bunch crossing. A significant number of inelastic collisions are so expected to occur at each crossing (corresponding to $\sim$ 1000 particles per bunch), due to the large number of protons per bunch.
To distinguish one event from another, a high detector granularity is mandatory. Besides, for a good pile-up control, the detectors must provide a fast response (i.e. a response concentrated in a single bunch spacing) along with a good time resolution (few ns) in order to distinguish the events coming from two consecutive bunch crossings. The limit where two consecutive signals start to overlap is called *out-of-time pile-up* and it affects the shape of the signal.
Considering the elevated number of events per second, which can be hardly handled even by the very high performance facilities of the LHC, it must be noticed that events can be recorded only at a rate of $\sim$ 300 Hz. Hence, an

[3] An inelastic collision is the collision of two partons, one from a proton of the first beam and one from a proton of the second beam. The energy of each parton is an unknown fraction of the proton energy and so it can be observed that the collision energy is not a fixed parameter.

online selection system, which could determine if an event is worth to be considered, is mandatory. In this way, an event reduction of seven orders of magnitude can be performed.

2.2 The Compact Muon Solenoid (CMS) Detector

The CMS detector was designed to be a general-purpose detector at LHC, having a quite broad physics program ranging from studying the Standard Model (aiming at the Higgs boson discovery in particular) to searching from extra dimensions and particles that could make up Dark Matter, which is studied in this thesis. The next subsections will go through details not only of all the CMS components but also of the standard reconstruction and identification algorithms used by the experiment.

2.2.1 Coordinate System

The coordinate system of the CMS detector is illiustrated in Figure 2.4. The

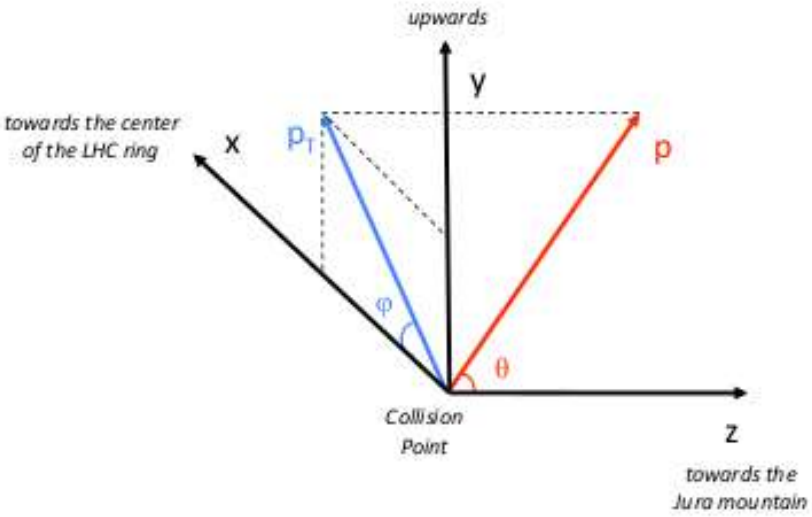


FIGURE 2.4: The CMS coordinate system.

detector has a cylindrical shape around the beam axis (z axis). As origin, the nominal collision point inside the experiment is considered. The x axis points horizontally towards the center of the LHC, and the y axis points vertically upwards, while the z (longitudinal) axis, horizontal and collinear to the beam trajectory, points to the direction of the vector perpendicular to xy plane, considering the anti-clockwise rotation of the x axis towards the y axis. In the transverse (x-y) plane, the azimuthal angle ϕ is measured from the x axis and the radial coordinate is denoted r. The polar angle θ is measured from

the z axis. The pseudorapidity variable η can be defined as

$$\eta = -ln\ \tan(\theta/2). \tag{2.2}$$

A particle trajectory direction at production point is then described by the coordinates (η, ϕ). Two parts of the subdetectors can be considered relying on the cylindrical shape of the detector:

1. the 'barrel', which corresponds to the central cylindrical region
2. the 'endcaps', which are two disks at the extremities closing the detector along the beam axis.

The parton momentum, before the collision, is expected to be longitudinal (along the beam axis). Being the transverse momentum of each parton negligible and the total transverse momentum conserved during an interaction, the transverse momentum of the collision is expected to be negligible too.

2.2.2 The CMS detector structure and the Magnet

The CMS detector [62] is a multi-purpose detector. Its length is 21.6 m and it has a diameter of 14.6 m. It contains two calorimeters: an electromagnetic one and an hadronic one. In the first, electromagnetic particles are stopped and measured, while hadronic particles are stopped in the second but measured in both of them.
The trajectories of all the charged particles are measured by an inner tracking device. Charged particles crossing both the calorimeters (i.e. muons) are measured instead by an outer tracking device.
The tracking devices are exposed to a magnetic field which curves the trajectories of charged particles. As reminded by its denomination, CMS detector has been designed [63] paying a particular attention to muons: their energy is measured through their track curvature information combined in both the inner and the outer tracking devices.
The charge and the p_T measurements are performed by using the curvature radius of the trajectory that is inverse proportional to the p_T.
For the Higgs boson search in the decay channel $H \rightarrow ZZ \rightarrow 4\mu$, an extremely precise measurement of the four-lepton mass is mandatory, so a precise measurement of the muon momentum is necessary, at least for p_T values up to 100 GeV/c and, along with it, a precise measurement of the muon track curvatures. Hence, a large bending power in the tracker region is also strongly recommended. For this purpose, a 3.8 T superconducting solenoid is used.
The tracker, and both calorimeters are located inside the solenoid, and exposed to its longitudinal magnetic field. The magnetic flux is returned through a 10,000 t iron yoke comprising 5 wheels and 2 endcaps, composed of three disks each. Four muon stations are included along the detector's length. The geometry of the CMS detector [64] is illustrated in Figure 2.5. The subdetectors and the online selection ('trigger') system are presented in the next sections.

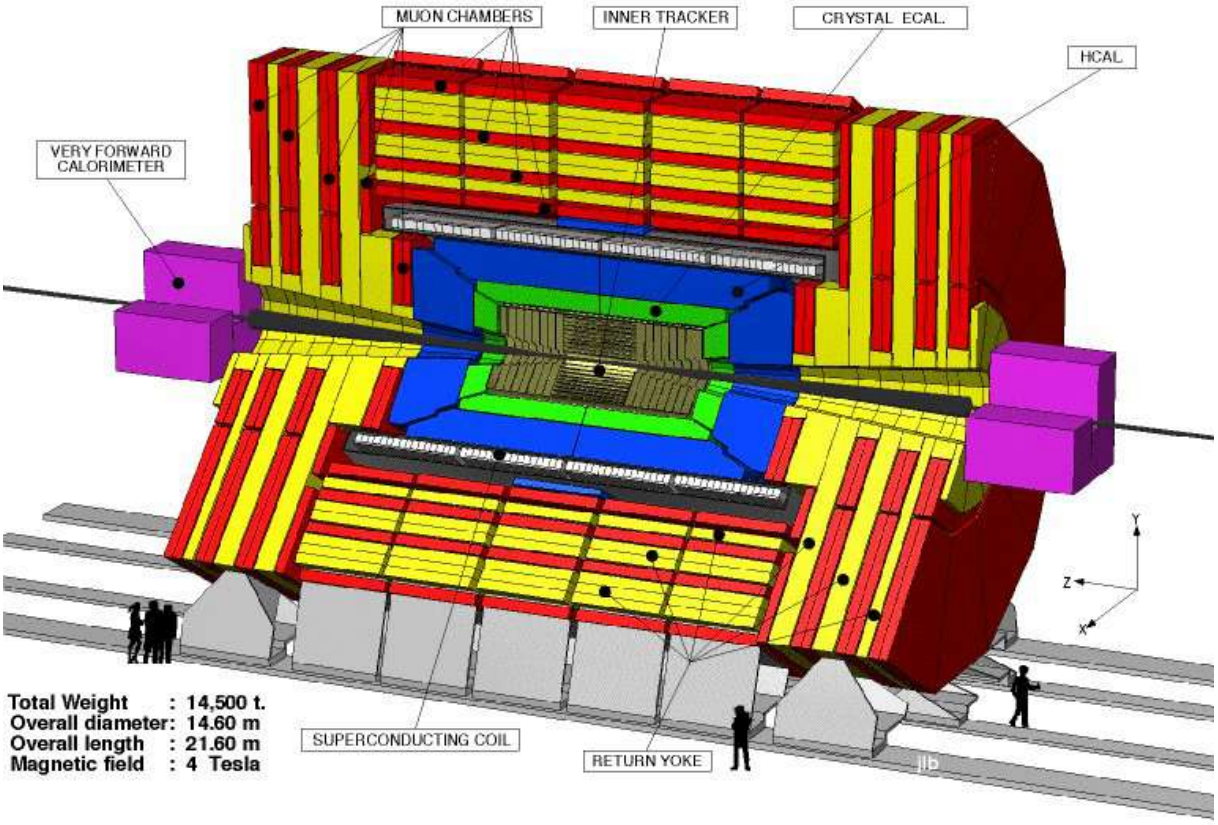


FIGURE 2.5: A perspective view of CMS detector.

2.2.3 Inner Tracking System

The CMS tracker is a fundamental tool for the charge and momentum measurements of charged particles. It has a length of 5.8 m and a diameter of 2.5 m around the interaction point. It covers a pseudorapidity range of $|\eta| < 2.5$. Since it is located directly around the collision point, the tracker material must be very resistant to radiation.
The very fine granularity in the innermost part is essential feature for the identification of the different vertices in a bunch crossing. While the primary vertex correspond to the interaction point of the collision, secondary vertices can indicate other interactions that can occur during the same bunch crossing (pile-up), or the presence of long-live particles[4].
A tracker design entirely based on silicon detector technology has therefore been chosen. However this very powerful system has some disadvantages:

- it implies a high density of detector electronics, which requires an efficient cooling system;
- the particles coming from collisions may interact with this dense material while crossing the tracker detector. It implies a complicated reconstruction procedure (see section 4.1) and a loss in the detector efficiency.

The high hit occupancy, which imposes constraints to the detector granularity, is the result of the high number of particles crossing the tracker.
The CMS tracker is made of two kinds of silicon sensors:

1. silicon pixels, which constitute the pixel detector in the most inner part;
2. silicon strips, which constitute the rest of the tracker.

The outer tracker region is made of thicker silicon sensors since the spatial density of tracks decreases far from the interaction point. In Figure 2.6 a schematic cross section of the CMS tracker is presented. Each line represents a detector module. The double lines indicate back-to-back modules which deliver stereo hits.
The pixel detector contains barrel and endcap modules; the silicon strip detector contains two collections of barrel modules, the Tracker Inner Barrel (TIB) and the Tracker Outer Barrel (TOB), and two collections of endcap modules, the Tracker Inner Discs (TID) and the Tracker EndCaps (TEC). The tracker structure contains several parts of central barrel layers and are completed by endcap disks on both sides. The pixels have to:

- provide the three first hits of the track of a primary particle;
- allow a precise measurement of a particle impact parameter (see chapter 4);
- allow secondary vertices identification.

[4]Leptons coming from late decays indicate a background event in the $H \to ZZ^{(*)\cdot} \to 4\ell$, where $\ell = e, \mu$

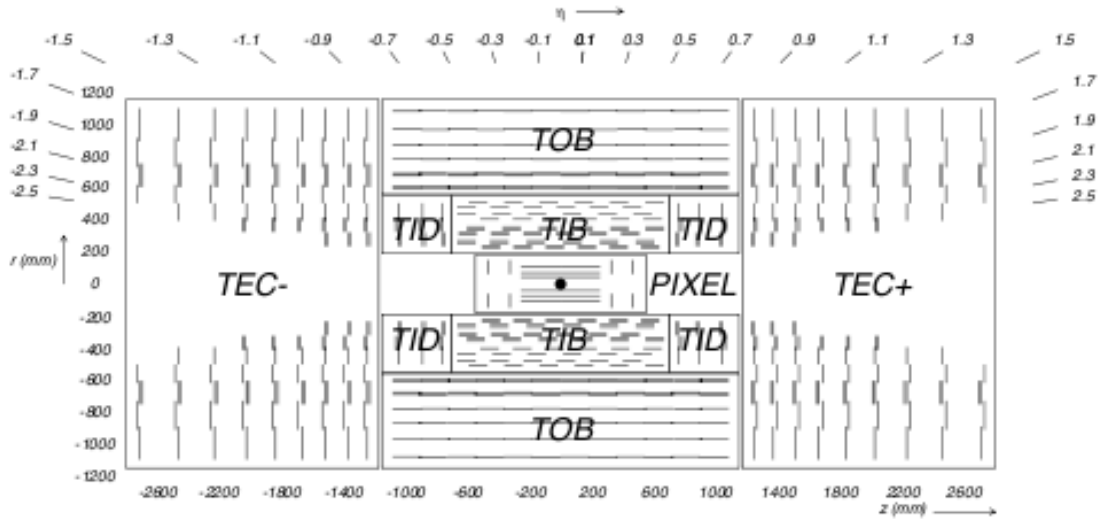


FIGURE 2.6: Schematic cross section through the CMS tracker.

The silicon tracker is coupled to a cooling system made of liquid Perfluorohexane (C_6F_14) and operates at temperature close to 0°C to prevent thermal risks.
The transverse momentum resolution varies according to the tracker modules crossed. A ∼1% resolution in the most central region, raising to ∼3% for high pseudorapidity values, is measured in the p_T range of W and Z boson decays ($p_T \sim 40$ GeV/c).

2.2.4 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [65] has been calibrated according to the requirements of the $H \rightarrow \gamma\gamma$ search. It is the only subdetector to provide information about photons. For an accurate di-photon mass reconstruction (~ 0.1 GeV/c^2), a very precise position and energy measurement is provided by the ECAL. The ECAL is also of primary importance for the electron reconstruction in a Higgs boson analysis in a multi-lepton final state. The combination of its information with the one from the tracker can ensure a very precise measurement of electron position and momentum and a significant background removal. A good segmentation is essential to distinguish the energy deposit shape of an electromagnetic particle from the one belonging to a hadronic particle. The CMS ECAL is a hermetic and homogeneous calorimeter, that covers the rapidity range $|\eta| < 3$. It is made of lead tungstate (PbWO4) crystals, mounted in a barrel ($|\eta| < 1.479$) and two endcaps ($1.479 < |\eta| < 3.0$).
A longitudinal view of the detector is shown in Figure 2.7. The crystals are followed by photodetectors that read and amplify their scintillation. Avalanche photodiodes (APDs) are used in the barrel. In the endcaps a very high resistivity to radiation and to the magnetic field is recommended.

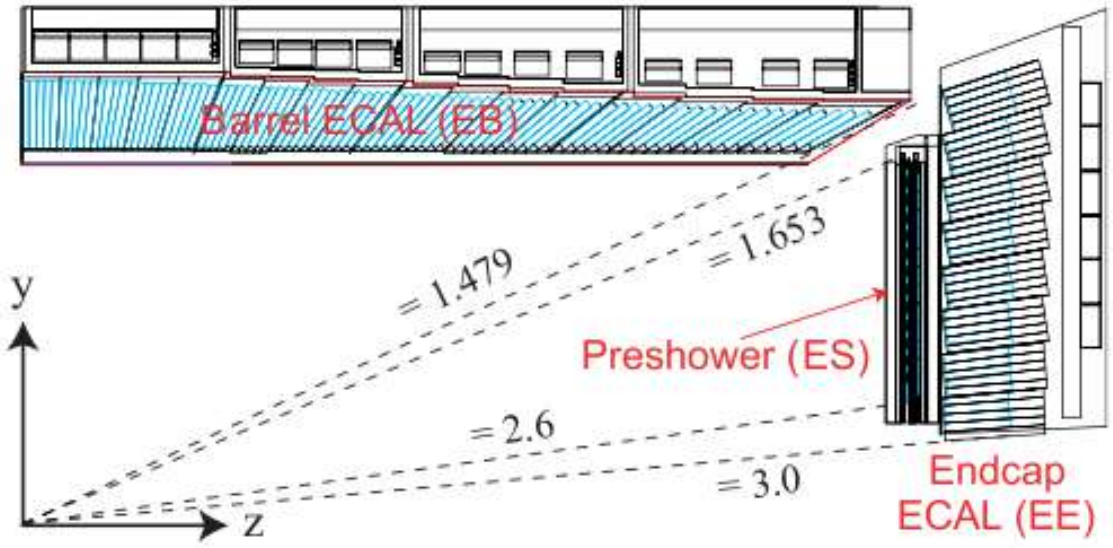


FIGURE 2.7: Longitudinal view of part of CMS electromagnetic calorimeter showing the ECAL barrel and an ECAL endcap with the preshower in front.

In the forward region the pion population becomes particularly important, and the π_0 decaying into two photons is hard to distinguish from a single photon. Hence, a better photon identification is ensured by a preshower detector installed in front of the ECAL endcaps (see Figure 2.7).
The preshower is a 20-cm thick sampling device, made of two parts located at both ends of the tracker, in front of the ECAL endcaps, covering the pseudorapidity range $1.653 < |\eta| < 2.6$ (see Fig. 2.8). Electromagnetic showers from incoming electrons and photons are initiated by its absorber, made of lead radiators. Two layers of silicon strip sensors are positioned, orthogonal oriented, behind each radiator. These sensors measure the deposited energy and the transverse shower profiles for better identification of electromagnetic particles. An electron or a photon emitted in the direction of the preshower, deposits 5% of its energy in the preshower, and the rest in the ECAL endcap.

The choice of the lead tungstate crystals relies on some constraints assigned to the detector:

- the compactness of the ECAL, needed to include both calorimeters inside the magnet;
- the good separability of electromagnetic showers due to the smallness of Molière radius[5] (2.2 cm) of lead tungstate;

[5]The Molière radius R_μ is a characteristic constant of a material, giving the scale of the transverse dimension of the fully contained electromagnetic showers initiated by an incident high energy electron or photon. It is defined as the mean deflexion of an electron of critical energy after crossing a width $1X_0$, where X_0 is defined as the radiation length, i.e. the average distance covered by an electron in a material through which it loose a fraction of its energy equal to $1/e$. A cylinder of radius R_μ contains on average 90% of the shower's energy deposition.

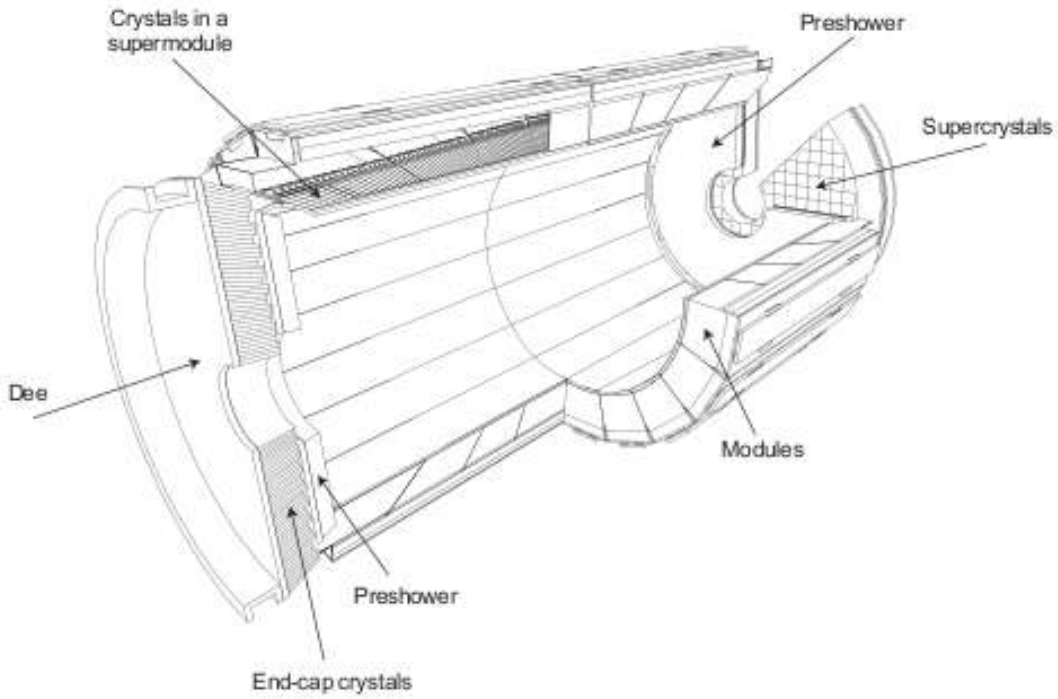


FIGURE 2.8: Layout of the CMS ECAL showing the arrengement of crystal modules, supermodules and endcaps, with preshower in front.

- the scintillation decay time of the crystals, which is fast enough for the LHC necessities.

The ECAL barrel is made of 36 identical *supermodules*, each covering half the barrel length ($-1.479 < \eta < 0$ or $0 < \eta < 1.479$), with a width of 20° in ϕ . Each supermodule is composed by four *Modules* in the η direction (see Figure 2.8). The presence of acceptance gaps, called *cracks*, between the modules, makes the energy reconstruction more complicated. At $\eta = 0$ a larger crack is present between supermodules, and an even larger one marks the barrel-endcap transition.
Each ECAL endcap is made of two semi-circular plates called *dees* (see Figure 2.8). Small cracks are also present between the endcap dees, but they can be considered negligible.
The energy loss can be measured by comparing the energy measured in the ECAL with the momentum measured in the tracker on electrons with little bremsstrahlung, considering that the difference is due to energy loss in cracks. In order to cancel these losses a recovery method has been conceived, except for the border corresponding to $\eta = 0$ and the barrel-endcap transition, where energy losses are 5% and 10% respectively.
Finally, the energy resolution was measured on one barrel supermodule, using incident electrons [66]. It consists of a stochastic, a noise and a constant contribution as follows:

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right) + \left(\frac{0.12\%}{E}\right)^2 + (0.30\%)^2. \tag{2.3}$$

and the result is shown in Figure 2.9. A resolution higher than 1% is achieved for electrons of energy higher than 15 GeV; for 40 GeV electrons it is of 0.6%.

2.2.5 Hadron Calorimeter

The hadron calorimeter (HCAL) plays a major role in the detection of hadron jets. It is located behind the Tracker and the Electromagnetic Calorimeter (from the interaction point of view). Its purpose is then to provide a sufficient containment of the hadron showers. Moreover, a wide extension in pseudorapidity is mandatory to have a precise description of the total collision event, allowing a more reliable measurement of the missing transverse energy.
The importance of the HCAL from the point of view of a Higgs boson analysis in a multi-lepton final state is that it allows to distinguish electrons from hadron jets, which can be misidentified as leptons (see Chapter 5). It is a sampling calorimeter. Such as the ECAL, it is composed of a barrel part (HB) and an endcap part (HE).
The HCAL Barrel covers the pseudorapidity range $|\eta| < 1.3$. It is limited in radial dimension, between the outer extent of the ECAL and the inner extent of the magnet coil (1.77 m $<$ R $<$ 2.95 m). Moreover, the HCAL is extended outside the solenoid with a tail catcher called the outer calorimeter, HO, just

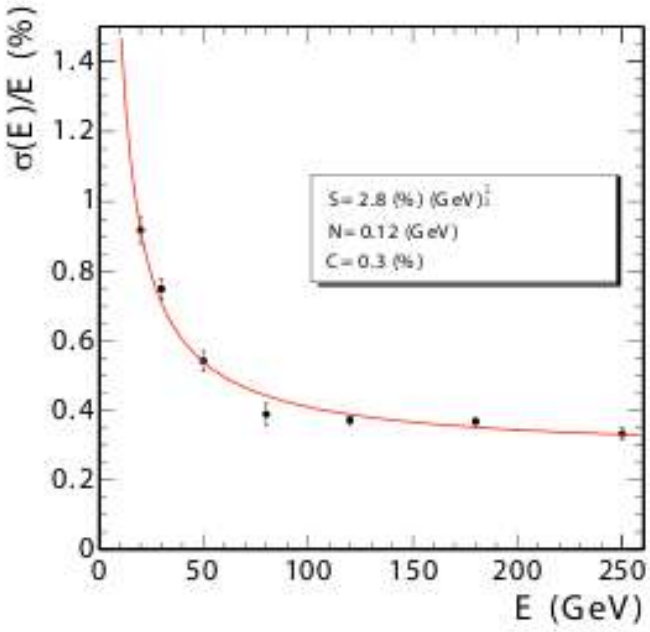


FIGURE 2.9: ECAL barrel energy resolution, $\sigma(E)/E$, as a function of electron energy as measured from a beam test (see text). The points correspond to events taken restricting the incident beam to a narrow (4×4 mm^2) region. The stochastic (S), noise (N), and constant (C) terms are given.

to ensure adequate sampling depth for $|\eta| < 1.3$.
The HCAL Endcaps cover a wide rapidity range: $1.3 < |\eta| < 3$. The forward hadron calorimeters (HF) are placed at 11.2 m from the interaction point extend. They extend the pseudorapidity coverage down to $|\eta| < 5.2$. The structure of the Hadron Calorimeter is illustrated in Figure 2.10.

The HB effective thickness increases with polar angle (θ) as $1/\sin\theta$. It results in 10.6 λ_I at $\eta = 1.3$, where λ_I is the radiation lenght[6].
The HO uses the solenoid coil as an additional absorber equal to $1.4/\sin\theta$ interaction lengths and is used to identify late-starting showers and to measure the shower energy deposited after HB. The material in the HCAL Endcaps must resist to the radiation, and handle high counting rates.
The absorber consists of a 40-mm-thick front steel plate, followed by eight 50.5-mm-thick brass plates, six 56.5-mm-thick brass plates, and a 75-mm-thick steel back plate.
Finally, the HE has to fully contain hadronic showers. The calorimeter barrel energy resolution $(EB + HB + HO)$ has been measured on pions which energy varies in a range of 3-500 GeV by test beams. It has been found to be:

$$\left(\frac{\sigma(E)}{E}\right) = \left(\frac{84.7\%}{\sqrt{E}}\right) \oplus 7.4\%. \tag{2.4}$$

[6]Nuclear interaction length is defined as the mean path length in which the energy of relativistic charged hadrons is reduced by the factor of $1/e$ as they pass through matter.

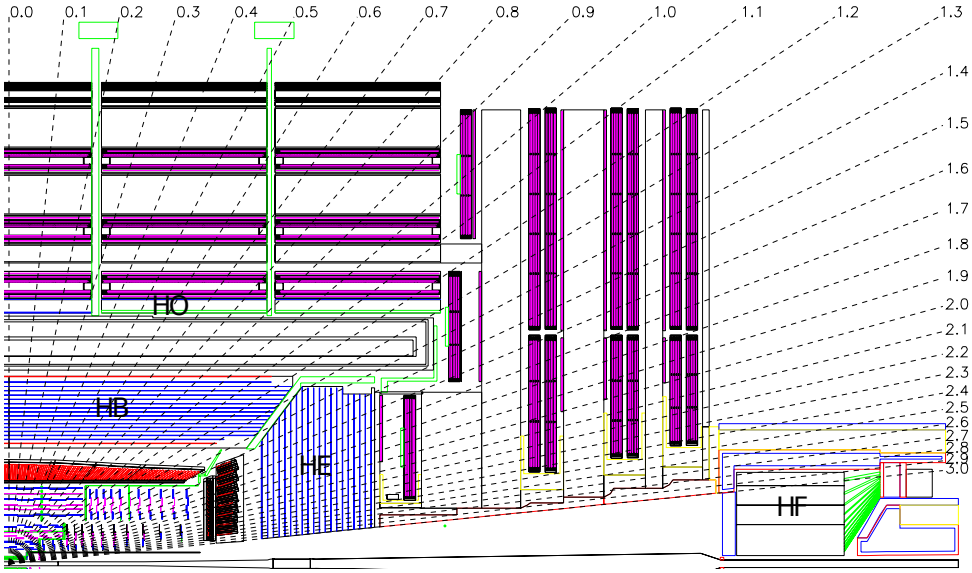


FIGURE 2.10: Longitudinal view of the CMS detector. The locations of the hadron barrel (HB), the endcap (HE), the outer (HO) and the forward (HF) calorimeters.

It can be observed that the energy resolution is dominated by the HCAL contribution.

2.2.6 The Muon System

The topology of the final state of $H \to ZZ \to 4\mu$ analysis give reasons for the construction of a muon system with a wide angular coverage and no acceptance gap.
Muons are particularly easy to identify and distinguish from backgrounds with CMS detector, thanks to the absorbing function played by the calorimeters.
The muon systems are divided into a cylindrical barrel section and two planar endcap regions. Less background, a low muon rate and a uniform 4-T magnetic field, mostly contained in the steel yoke, is measured in the barrel. A longitudinal view of the muon detectors can be found in Figure 2.11.

Muon System Subdetectors

Drift tube (DT) chambers have been used. They cover the pseudorapidity region $|\eta| < 1.2$. Chambers measuring the muon coordinate in the $r - \phi$ bending plane are alternated with chambers providing a measurement in the z direction. Each of the four stations contains four chambers of each kind. The most relevant problem of this design is the presence of 'cracks', i.e. dead efficiency spots in efficiency between the chambers. It has been solved by the presence of an offset of the drift cells between neighbor chambers.

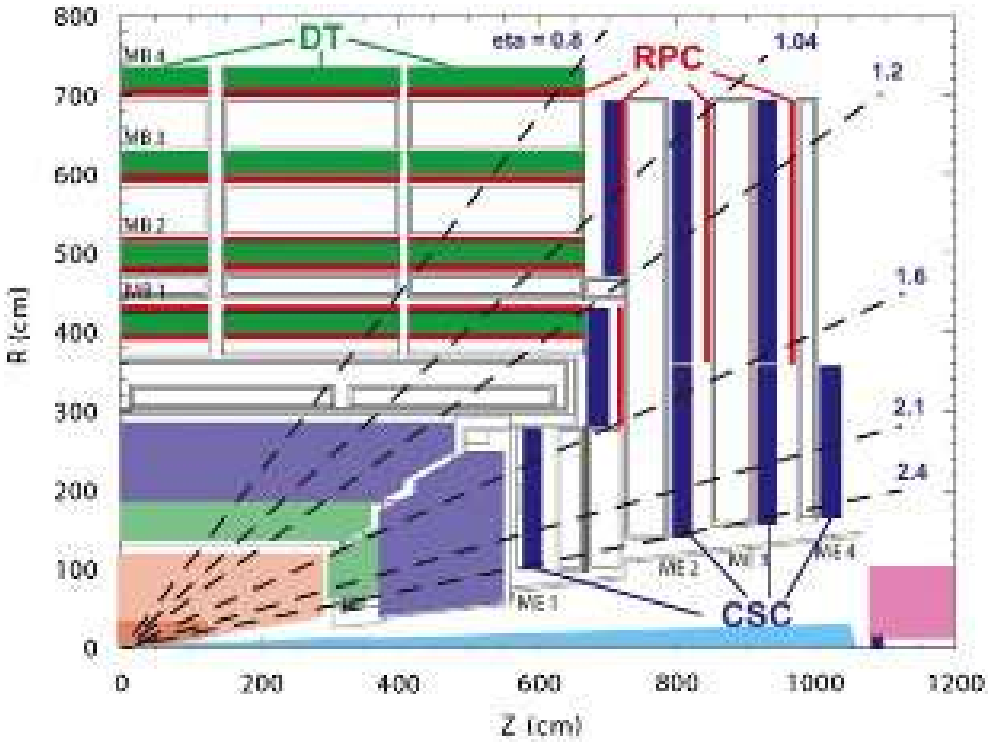


FIGURE 2.11: Longitudinal view of the muon detectors: DT, RPC and CSC.

The endcaps cover a region of higher rates. In this region the magnetic field appears large and non-uniform.
Cathode strip chambers (CSC) are instead used to cover the pseudorapidity region $0.9 < |\eta| < 2.4$. Each of the four stations contains six layers of chambers and anode wires. The chambers are positioned perpendicular to the beam line and provide a precision measurement in the $r - \phi$ bending plane, while the anode wires provide measurements of and the beam-crossing time of muons. Other tools are included to reject non-muon backgrounds and to match hits to those in the other stations and in the inner tracker.
A system of resistive plate chambers (RPC) has been added in barrel and endcap regions, over a large portion of the pseudorapidity range ($|\eta| < 1.6$). They consists of double-gap chambers, operated in avalanche mode to ensure good operation at high rates. Six layers are present in the barrel and three in each endcap. They produce a fast response, with good time resolution but coarser position resolution than the DTs or CSCs. They provide an independent trigger system with an optimal time resolution. Moreover, they help to reduce ambiguities in track reconstruction.
The muon momentum resolution is optimized by a high technology alignment system, which measures the relative positions of the muon detectors along with their positions respect with the inner tracker system.

2.2.7 Trigger

The Trigger system can be considered as the first event selection step.
The main feature of this step, which makes it different from the other selection steps, is that it is not reversible. It indeed performs a fast selection of the events which seem to be of interest for physics analysis among the huge amount of those produced by LHC collisions. This selection can drastically reduce the extremely high event rate (the LHC nominal bunch crossing rate is $\sim$ 40 MHz) to a reasonable rate, more suitable for data recording ($\sim$ 300 Hz). Obviously, all collision data must be kept until the trigger decision has been taken, so requiring a fast response.
To fulfill these requirements, a two-level trigger system has been designed. The Level-1 (L1) Trigger is a hardware system made of largely programmable electronics. It provides a first rate reduction to 100 kHz, scanning events fastly in 3.2 μs. This timing constraint are satisfied considering coarse granularity objects from the calorimeters and from the muon system.
A positive L1 decision is converted in a transfer of the complete event information to the next level: the High Level Trigger (HLT). Unlike the previous one, the HLT is a software system which is based on algorithms of increasing complexity that use the fine granularity of the event. So, the HLT decision time is not a fixed value as the L1 trigger one. It may vary according to the event, with a mean value of 50 ms.
In the case of the Higgs boson analysis in multi-lepton final state here presented, the trigger relies on events containing electron and muon signals. For the Level-1 Trigger, an electron signature can be identified with a narrow and highly energetic deposit in the ECAL, while a muon signature is based on a track segment or a hit pattern in muon chambers.
The High-Level Trigger considers higher granularity objects, because it reconstructs the total energy deposits in the calorimeters and muon tracks, and combines them with the tracker and preshower information.

Level-1 Trigger

The Level-1 Trigger architecture is described in Figure 2.12. It is divided into two parallel trigger systems, one corresponding to the calorimeters and the other to the muon chambers. Each system is composed of a local, a regional, and a global part, then merged into a Global Trigger for the final L1 decision. The candidate categories of the Level-1 Trigger are:

- Muons, built in the Muon Trigger;
- Electrons/Photons (isolated and non-isolated: $e - \gamma$);
- Central and forward Jets;
- Taus, built in the Regional Calorimeter Trigger;

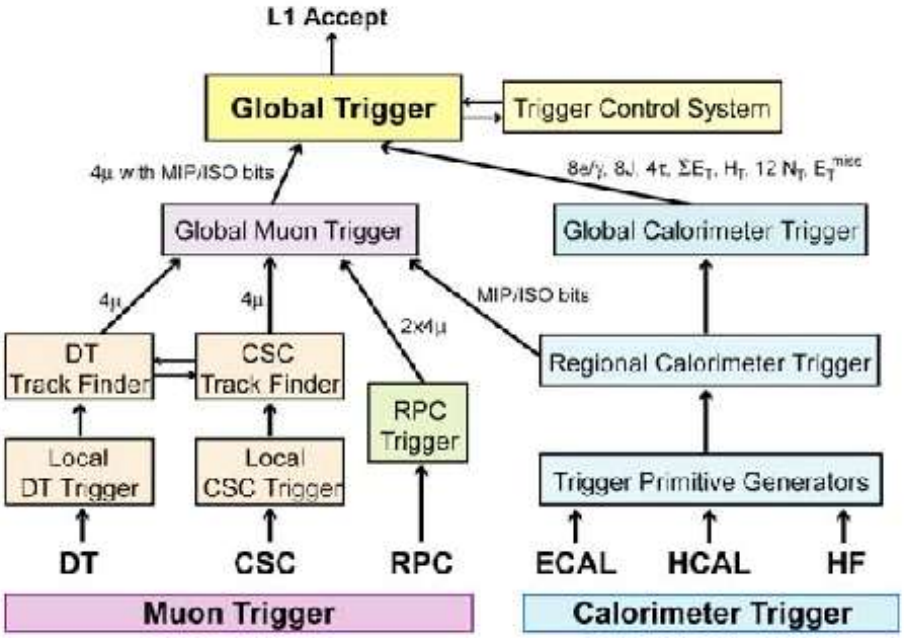


FIGURE 2.12: Level-1Trigger Architecture.

- Total Transverse Energy ($\sum E_T$), Missing Transverse Energy E_T^{miss}, Scalar Transverse Energy Sum of all Jets (above a given threshold: H_T, built in the Global Calorimeter Trigger;

Local Triggers

The local trigger creates coarse-granularity information . In the calorimeters, this information is a collection of Trigger Primitives.

Regional Triggers

The Regional Calorimeter Trigger collects the local information to build Level-1 Trigger candidates, combining all the information of both calorimeters. A DT track finder and a CSC track finder collect the local DT and CSC information to build Level-1 Trigger Candidates as tracks for the muon trigger. The RPC trigger is directly regional.
The four most relevant Candidates of each category are sent to the Global Calorimeter Trigger or to the Global Muon Trigger respectively. The regional summed transverse energy is also sent to the Global Calorimeter Trigger by the Regional one.

Global Calorimeter Trigger and Global Muon Trigger

The Global Calorimeter Trigger has finally the task of sorting the Level-1 Trigger Candidates to send the four most relevant ones of each category to the Global Trigger. It also calculates the summed E_T and the E_T^{miss} information of the event, as well as the scalar transverse energy sum of all jets above a given

threshold (HT). The Global Trigger receives this information as well.
The Global Muon Trigger collects and compares the candidates from the DT, CSC and RPC Triggers and combines them into four Muon Candidates. It also uses some information from the Regional Calorimeter Trigger for isolation considerations. The Global Trigger finally collects the information about the four Muon Candidates.

Global Trigger

The Global Trigger collects the candidates produced by the Global Calorimeter Trigger and the Global Muon Trigger, and compares them to the Level-1 (L1) Trigger Menu. This menu is a list of Level-1 enabled triggers. If at least one of the listed triggers is satisfied by a candidate collection, the Level-1 Trigger response is positive and the fine granularity event information can be sent to High-Level Trigger. The Level-1 Trigger also follows some rules to prevent memory overload (e.g. L1 Trigger cannot accept two events separated by only one single bunch crossing).

The trigger algorithms consist in a threshold applied to the highest energetic candidate of each category. For background reduction, a combination of triggers is often required.

High-Level Trigger

The higher and last level trigger step is the High-Level Trigger which builds candidates corresponding to all kinds of reconstructed objects considered in the offline analyses. The algorithms used are very similar to the previous ones. Its inner sub-structure is made of several increasing complexity levels, starting from Level-2.
The High Level Triggers can access to all the information used in Level-1 since this is stored locally in the L1 trigger crates.
The Level 2 starts generally with the Level-1 Trigger information, and builds fine granularity objects around the Level-1 candidates, using only the information from the calorimeters and the muon system. The tracker information is also used, only when necessary, at the next 2.5 Level. Eventually, the High Level Triggers use the full event data to decide to keep an event.

2.3 CMS upgrades for Run II

As previously mentioned, the second operational Run, also called Run II, started in 2013. Many upgrades on the various software and hardware CMS components have been developed to cope with the higher energy reached (among others, see for example Refs. [54], [55], [56]).

Figure 2.13 shows the integrated luminosity reached during 2017 in proton-proton collisions at center of mass energy $\sqrt{s} = 13$TeV, while in Figure 2.14 the comparison among the growing trends of the integrated luminosity registered by the CMS experiment from 2010 to 2017 can be observed. With respect to 2012, a milestone year for the discovery of the Higgs boson, the luminosity trend grows much faster, doubling its value in 2017 at the end of the year (see November bin in the Figure). Other upgrades for Phase-1 (Run III in 2019-2022) and Phase-2 (2024, HL-LCH runs) are scheduled to allow CMS to deal a luminosity of $L = 2 \times 10^{34} cm^{-2} s^{-1}$ with 25 ns of bunch crossing (average PU of 50 ns) and a luminosity of $L = 5 \times 10^{34} cm^{-2} s^{-1}$ with 25 ns of bunch crossing (average PU of 140 ns) respectively.

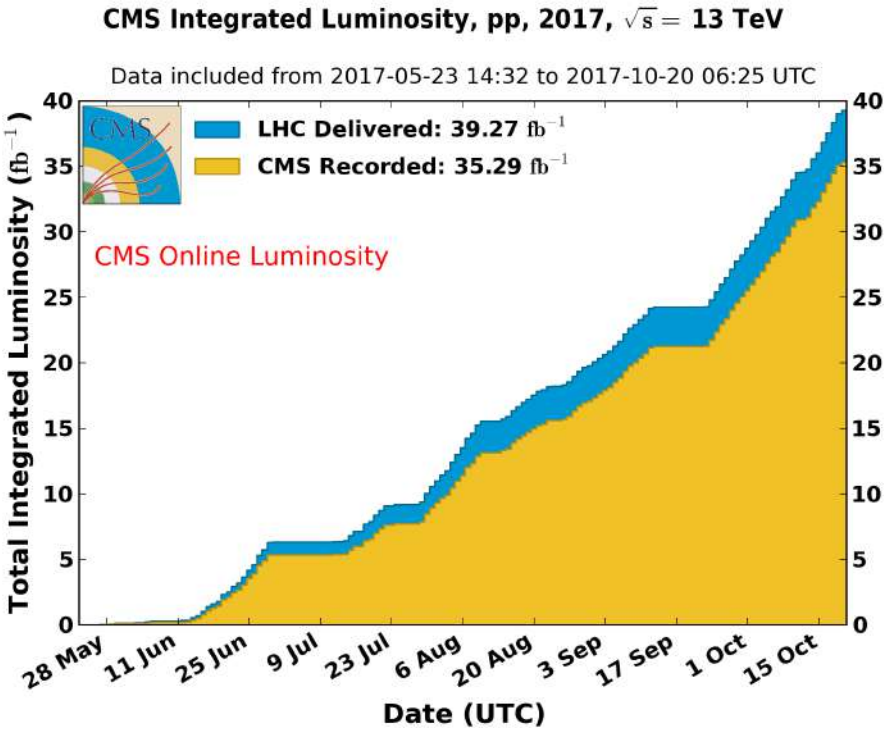


FIGURE 2.13: CMS Integrated Luminosity reached during 2017 in proton-proton collisions at center of mass energy $\sqrt{s} = 13$TeV

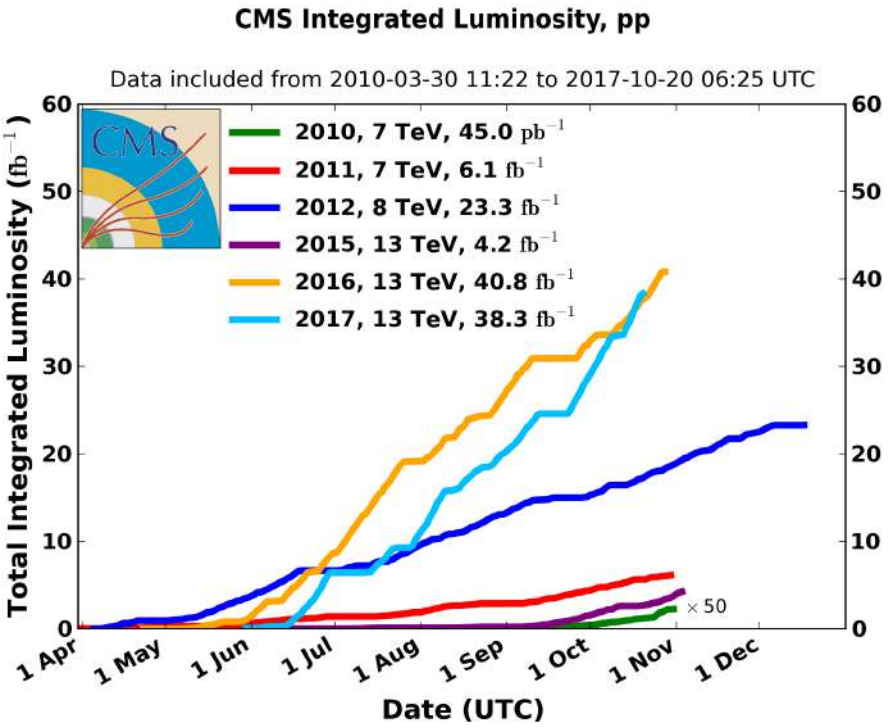


FIGURE 2.14: CMS Integrated Luminosity in proton-proton collisions. The colored lines refer to different years of data-taking ranging from Run I (2010-2012) to Run II (2015-2017)

Chapter 3

Data and Monte Carlo Simulation

In this chapter the data and Monte Carlo background and signal samples used for the Mono-Higgs analysis described will be presented in details. While for the 2016 analysis the official MC samples for the Mono-Higgs signals in the $Z'2HDM$ model including a Higgs boson decaying into four leptons were available (see the following sections), for the 2015 analysis they were privately produced on the local Tier-2 using ReCaS Data Center [57] machines using the same generators described in subsection 3.2.1.

3.1 Data

3.1.1 Triggers and Datasets

This analysis uses a sample of data which includes only the runs certified by a CMS Official Group established on purpose. This data have been recorded during 2016 and correspond to an integrated luminosity of 35.867 fb^{-1}. Table 3.1 shows the list of all the datasets along with the related integrated luminosity.

Five different primary datasets (PDs)[1] are used for this analysis:

- *DoubleEG*,
- *DoubleMuon*,
- *MuEG*,
- *SingleElectron*,
- *SingleMuon*.

[1] CMS primary datasets are AOD (Analysis Object Data) files which contain the information that is needed for the analysis.: all the high-level physics object (such as muons, electrons, etc..), the tracks with associated hits, calorimetric clusters with associated hits, vertices and also information about event selection (High Level Trigger, HLT), data needed for further selection and identification criteria for the physics oblects.

Run range	Dataset	Integrated luminosity
5*273150-275376	/DoubleMuon/Run2016B-23Sep2016-v3/MINIAOD /DoubleEG/Run2016B-23Sep2016-v3/MINIAOD /MuonEG/Run2016B-23Sep2016-v3/MINIAOD /SingleElectron/Run2016B-23Sep2016-v3/MINIAOD /SingleMuon/Run2016B-23Sep2016-v3/MINIAOD	5*5.892 fb^{-1}
5*275656-276283	/DoubleMuon/Run2016C-23Sep2016-v1/MINIAOD /DoubleEG/Run2016C-23Sep2016-v1/MINIAOD /MuonEG/Run2016C-23Sep2016-v1/MINIAOD /SingleElectron/Run2016C-23Sep2016-v1/MINIAOD /SingleMuon/Run2016C-23Sep2016-v1/MINIAOD	5*2.646 fb^{-1}
5*276315-276811	/DoubleMuon/Run2016D-23Sep2016-v1/MINIAOD /DoubleEG/Run2016D-23Sep2016-v1/MINIAOD /MuonEG/Run2016D-23Sep2016-v1/MINIAOD /SingleElectron/Run2016D-23Sep2016-v1/MINIAOD /SingleMuon/Run2016D-23Sep2016-v1/MINIAOD	5*4.353 fb^{-1}
5*276831-277420	/DoubleMuon/Run2016E-23Sep2016-v1/MINIAOD /DoubleEG/Run2016E-23Sep2016-v1/MINIAOD /MuonEG/Run2016E-23Sep2016-v1/MINIAOD /SingleElectron/Run2016E-23Sep2016-v1/MINIAOD /SingleMuon/Run2016E-23Sep2016-v1/MINIAOD	5*4.117 fb^{-1}
5*277932-278808	/DoubleMuon/Run2016F-23Sep2016-v1/MINIAOD /DoubleEG/Run2016F-23Sep2016-v1/MINIAOD /MuonEG/Run2016F-23Sep2016-v1/MINIAOD /SingleElectron/Run2016F-23Sep2016-v1/MINIAOD /SingleMuon/Run2016F-23Sep2016-v1/MINIAOD	5*3.186 fb^{-1}
5*278820-280385	/DoubleMuon/Run2016G-23Sep2016-v1/MINIAOD /DoubleEG/Run2016G-23Sep2016-v1/MINIAOD /MuonEG/Run2016G-23Sep2016-v1/MINIAOD /SingleElectron/Run2016G-23Sep2016-v1/MINIAOD /SingleMuon/Run2016G-23Sep2016-v1/MINIAOD	5*7.721 fb^{-1}
15*281207-284068	/DoubleMuon/Run2016H-PromptReco-v1/MINIAOD /DoubleEG/Run2016H-PromptReco-v1/MINIAOD /MuonEG/Run2016H-PromptReco-v1/MINIAOD /SingleElectron/Run2016H-PromptReco-v1/MINIAOD /SingleMuon/Run2016H-PromptReco-v1/MINIAOD /DoubleMuon/Run2016H-PromptReco-v2/MINIAOD /DoubleEG/Run2016H-PromptReco-v2/MINIAOD /MuonEG/Run2016H-PromptReco-v2/MINIAOD /SingleElectron/Run2016H-PromptReco-v2/MINIAOD /SingleMuon/Run2016H-PromptReco-v2/MINIAOD /DoubleMuon/Run2016H-PromptReco-v3/MINIAOD /DoubleEG/Run2016H-PromptReco-v3/MINIAOD /MuonEG/Run2016H-PromptReco-v3/MINIAOD /SingleElectron/Run2016H-PromptReco-v3/MINIAOD /SingleMuon/Run2016H-PromptReco-v3/MINIAOD	15*8.857 fb^{-1}

TABLE 3.1: Campioni di dati utilizzati nell'analisi.

As the name of these PDs suggest, they collect data passing single-lepton, dilepton, or trilepton triggers. Each of them is required to combine a collection of HLT paths.
The High Level Trigger (HLT) paths used are listed in Table 3.2.

To avoid duplicate events from different primary datasets, events are taken:

- from DoubleEG if they pass the diElectron or triElectron triggers;
- from DoubleMuon if they pass the diMuon or triMuon triggers and fail the diEle and triEle triggers;
- from MuEG if they pass the MuEle or MuDiEle or DiMuEle triggers and fail the diEle, triEle, diMuon and triMuon triggers;

HLT path	L1 seed	prescale	primary dataset
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	L1_DoubleEG_15_10	1	DoubleEG
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	L1_DoubleEG_22_10	1	DoubleEG
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL	(Multiple)	1	DoubleEG
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL	L1_TripleEG_14_10_8	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL	L1_DoubleMu_11_4	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL	L1_DoubleMu_11_4	1	DoubleMuon
HLT_TripleMu_12_10_5	L1_TripleMu_5_5_3	1	DoubleMuon
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL	L1_Mu5_EG15	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL	L1_Mu5_EG20	1	MuonEG
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	L1_Mu12_EG10	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	L1_Mu20_EG10	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL	L1_SingleMu*	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL	L1_Mu6_DoubleEG10	1	MuonEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL	L1_DoubleMu7_EG7	1	MuonEG
HLT_Ele25_eta2p1_WPTight	L1_SingleEG*	1	SingleElectron
HLT_Ele27_WPTight	L1_SingleEG*	1	SingleElectron
HLT_Ele27_eta2p1_WPLoose_Gsf	L1_SingleEG*	1	SingleElectron
HLT_IsoMu20 OR HLT_IsoTkMu20	L1_SingleMu*	1	SingleMuon
HLT_IsoMu22 OR HLT_IsoTkMu22	L1_SingleMu*	1	SingleMuon

TABLE 3.2: Trigger paths used in 2016 collision data.

- from SingleElectron if they pass the singleElectron trigger and fail all the above triggers;
- from SingleMuon if they pass the singleMuon trigger and fail all the above triggers.

The trigger requirement does not depend on the selected final state: it is always the OR of all 10 HLT paths. The trigger efficiency measured using $4l$ events is found to be $> 99\%$ for each final state.

3.2 Simulation

3.2.1 Signal samples

Official Monte Carlo signal samples for both Z' - Two Higgs Doublet Model and the Z'Baryonic model were simulated for this study.
MADGRAPH5_AMCATNLO [67] generator was used for Z'2HDM and the Z'Baryonic model simulation. Showering of parton-level events is done using PYTHIA8.209 [68], and in the matching is performed by allowing QCD emissions at all energies in the shower. All the values of the mass parameters used for the simulation of signal sample for the Z'2HDM model and the related cross sections are reported in Table 3.3 and Table 3.4.

The values for the mass parameters used for the simulation of signal sample for the Z'Baryonic model and the related cross sections are reported in Table 3.5 and Table 3.6.

m_{A^0} [GeV]	$m_{Z'}$ [GeV]							
300	600	800	1000	1200	1400	1700	2000	2500
400	600	800	1000	1200	1400	1700	2000	2500
500		800	1000	1200	1400	1700	2000	2500
600		800	1000	1200	1400	1700	2000	2500
700			1000	1200	1400	1700	2000	2500
800			1000	1200	1400	1700	2000	2500

TABLE 3.3: Mass values used for Z'2HDM.

m_{A^0} [GeV]	$m_{Z'}$ [GeV]							
300	0.45217	0.27765	0.14383	0.075451	0.041208	0.017786	0.0083417	0.0025458
400	0.063916	0.091124	0.058824	0.034039	0.019609	0.0088466	0.0041934	0.0013246
500	0.0046949	0.0367	0.03431	0.022935	0.014225	0.0068221	0.003341	0.0010859
600	0.0011927	0.010101	0.01897	0.015763	0.010822	0.0056149	0.0028691	0.00096781
700	0.00049587	0.0017325	0.0085776	0.010028	0.0079659	0.0045823	0.0024651	0.00086993
800	0.00024874	0.00059868	0.0027483	0.005655	0.0055221	0.0036537	0.0020981	0.0007806

TABLE 3.4: Z'2HDM model production cross sections [pb] corresponding to mass points in Table 3.3

m_χ [GeV]	$m_{Z'}$ [GeV]									
1	10	20	50	100	200	300	500	1000	2000	10000
10	10	15	50	100						10000
50	10		50	95	200	300				10000
150	10				200	295	500	1000		10000
500	10						500	995	2000	10000
1000	10							1000	1995	10000

TABLE 3.5: Mass values used for the Z'Baryonic model.

m_χ [GeV]	$m_{Z'}$ [GeV]									
1	2.5948	2.7224	3.2561	3.1802	2.5517	2.2672	1.0937	0.20177	0.013935	1.1916e-08
10	0.011929	0.041810	3.2386	3.1735						1.1931e-08
50	0.00020487		0.0071829	0.18559	2.0924	1.9471				1.1839e-08
150	3.1486e-06				0.0031894	0.13389	0.67973	0.19601		1.1585e-08
500	4.5072e-09						2.0570e-05	0.011122	0.012672	9.5642e-09
1000	2.7729e-11							5.5155e-07	0.00072969	6.5979e-09

TABLE 3.6: Z'Baryonic model production cross sections [pb] corresponding to mass points in Table 3.5

3.2.2 Background Samples

Two kinds of background have been considered for this analysis: SM Higgs samples, in which the four leptons are the products of the decay of a Higgs boson and all the other samples, in which the leptons do not come from the Higgs. The missing transverse energy in the background events comes from neutrinos.
In the following sections all the background samples considered for this analysis will be discussed.

Process	Dataset Name	$\sigma \times BR(\times\epsilon_{\text{filter}})$
$gg \to H \to ZZ \to 4\ell$	/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8/[1]	12.18 fb
$qq \to Hqq \to ZZqq \to 4\ell qq$	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8/[1]	1.044 fb
$q\bar{q} \to W^+H \to W^+ZZ \to 4\ell + X$	/WplusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUgenV6_pythia8/[1]	0.232 fb
$q\bar{q} \to W^-H \to W^-ZZ \to 4\ell + X$	/WminusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUgenV6_pythia8/[1]	0.147 fb
$q\bar{q} \to ZH \to ZZZ \to 4\ell + X$	/ZH_HToZZ_4LFilter_M125_13TeV_powheg2-minlo-HZJ_JHUgenV6_pythia8/[1]	0.668 fb
$gg \to ttH \to ttZZ \to 4\ell + X$	/ttH_HToZZ_4LFilter_M125_13TeV_powheg_JHUgen_pythia8/[1]	0.393 fb

[1] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6-v1

TABLE 3.7: Signal Monte Carlo samples and corresponding cross section values.

SM Higgs samples

The SM Higgs simulated samples are produced using the POWHEG V2 [69–71] generator for the main production modes:

- gluon-gluon fusion (ggH) including quark mass effects [73],
- vector boson fusion (VBF) [74],
- associated production (WH, ZH, and $t\bar{t}H$ [75]).

An extension of POWHEG, MiNLO HVJ, has been used [76] in the case of WH and ZH.
The decay of the Higgs boson to four leptons is obtained using the JHUGEN generator [77]. In the case of WH, ZH, and ttH, the Higgs boson is allowed to decay to H$\to$ ZZ $\to$ 2l2X such that 4-lepton events where two leptons originate from the decay of associated Z, W bosons, or top quarks are also taken into account in the simulation.
The parton shower and the hadronization processes are managed by PYTHIA8.209. All samples are generated with the NNPDF 3.0 NLO (Next to the Leading Order) parton distribution functions (PDFs) [72]. Table 3.7 shows the list of signal samples and their cross sections.

Other background samples

The production of ZZ via quark-antiquark annihilation is generated at next-to-leading order using POWHEG V2 [45] and PYTHIA8 (used for the parton shower and hadronization)
Since this simulation covers a large range of ZZ invariant masses, dynamical QCD factorization and renormalization scales have been chosen, equal to m_{ZZ}. The $gg \to ZZ$ process is simulated at leading order (LO) with MCFM [47, 94].
In order to match the $gg \to$ H $\to ZZ$ transverse momentum spectra predicted by POWHEG at NLO, the showering for MCFM samples is performed with different PYTHIA8 settings, allowing only emissions up to the parton-level scale.
Additional MC samples of WZ, Drell-Yan (DY) +jets, $t\bar{t}$, $t\bar{t}$ + vector bosons, and tribosons are generated using MADGRAPH5_AMCATNLO either inclusively or merging several jet multiplicities.
Table 3.8 summarizes the MC simulation datasets used for this analysis.

Process	Dataset Name	$\sigma \cdot BR$
$qq \to ZZ \to 4\ell$	/ZZTo4L_13TeV_powheg_pythia8/[1]	1.256pb
$qq \to ZZ \to 4\ell$	/ZZTo4L_13TeV-amcatnloFXFX-pythia8/[1]	1.212pb
$gg \to ZZ \to 4e$	/GluGluToContinToZZTo4e_13TeV_MCFM701/[1]	0.00159pb
$gg \to ZZ \to 4\mu$	/GluGluToContinToZZTo4mu_13TeV_MCFM701/[1]	0.00159pb
$gg \to ZZ \to 4\tau$	/GluGluToContinToZZTo4tau_13TeV_MCFM701/[1]	0.00159pb
$gg \to ZZ \to 2e2\mu$	/GluGluToContinToZZTo2e2mu_13TeV_MCFM701/[1]	0.00319pb
$gg \to ZZ \to 2e2\tau$	/GluGluToContinToZZTo2e2tau_13TeV_MCFM701/[1]	0.00319pb
$gg \to ZZ \to 2\mu2\tau$	/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701/[1]	0.00319pb
$Z \to \ell\ell$ + jets	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	6104pb
$Z \to \ell\ell$ + jets	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	18610pb
$WZ \to 3\ell\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8/[1]	4.430pb
$t\bar{t}$	/TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	815.96pb
$t\bar{t} \to 2\ell2\nu2b$	/TTTo2L2Nu_13TeV-powheg/[1]	87.31pb
ZZZ	ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.01398pb
WWZ	WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.16510pb
WZZ	WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.05565pb
$t\bar{t}$ W	TTWJetsToLNu_TuneCUETP8M1$_1$3$TeV - amcatnloFXFX - madspin - pythia$8	0.2043pb
$t\bar{t}Z$	TTZToLLNuNu_M-10_TuneCUETP8M1$_1$3$TeV - amcatnlo - pythia$8	0.2529pb

[1] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6-v1

TABLE 3.8: Background Monte Carlo samples and cross sections.

On top of the generated samples, GEANT4 [78] has been used to simulate the CMS detector and the interaction of the particles with that; then the events are reconstructed by using official tools provided by the collaboration.
All MC samples have been reweighted to take into account the pileup contribution in 2016 data.

Chapter 4

Physics Objects

The final signature for the Mono-Higgs analysis reported includes four leptons plus a large value of missing transverse energy due to the DM candidates expected. Since the four lepton combinations studied are 4μ, 4e, $2e2\mu$, in the following sections the muon and electron reconstruction, identification and isolation criteria will be discussed along with the reconstruction of the missing transverse energy and jets.
Before going through the cut-based selection applied to the objects used this analysis, they will be introduced in detail in sections 4.2, 4.3, 4.4, 4.5 and 4.6, while section 4.1 will previously present a general overview on how the leptons used in this analysis, electrons and muons, are reconstructed with the CMS detectors.

4.1 Lepton Reconstruction in CMS

4.1.1 Electron Reconstruction

The electron reconstruction [79] combines the information from the electromagnetic calorimeter (ECAL) and the silicon tracker. It starts by the reconstruction of clusters seeded by hot cells in the ECAL. Electron seeds are then used to form superclusters (clusters of clusters) to collect the electron energy radiated by bremsstrahlung in the tracker and spread in ϕ by the solenoidal magnetic field and to initiate a track building and a fitting procedure.
The superclusters are first preselected using a hadronic veto (defined by the ratio H/E of the hadronic energy estimated by summing HCAL towers energy within a cone of $\Delta\mathcal{R} = 0.15$ behind the supercluster position over the supercluster energy) and applying a 4 GeV cut on the supercluster transverse energy. The superclusters are also used to search for hits in the innermost tracker layers which are used to accomplish the seeding of the tracks.
The ECAL driven seeding algorithm has been used. It has been optimized for isolated electrons in the p_T^e range relevant for Z or W decay, down to 5 GeV/c. For lower electron p_T^e values the ϕ window used for the superclusters becomes too small and the electrons which radiate lead to electron and photon clusters separated by a distance greater than 0.3 rad (the maximum

limit) in the magnetic field.
Moreover, for electrons in jets, the energy collected in the superclusters could include neutral contribution from jets so biasing the energy measurement used to seed the tracks. For these reasons, the driven seeding strategy has been complemented by a tracker driven seeding algorithm. It can be illustrated with two extreme cases:

- electrons which do not radiate energy by bremsstrahlung while crossing the tracker;
- electrons which undergo a significant energy loss by bremsstrahlung.

In the first case, the electron creates a single cluster in the ECAL and its track may be reconstructed well enough by the standard Kalman Filter, which is able to collect hits up to ECAL. The track reconstructed is then matched with a particle flow[1] [81] cluster and the ratio E/p of the cluster energy over the track momentum can be evaluated. If the value of this ratio is close to unity, the seed of the track is considered as an electron seed.
Instead, in the second case, the Kalman Filter cannot follow the change of curvature and a small number of hits belongs to the track. In this case, the electron tracks are selected using the silicon tracker as a preshower and evaluating the different characteristics of a pion track and an electron track reconstructed by Kalman Filter.
A merging procedure of the seeds of the two algorithms is then carried out so keeping the track of seed provenance. It can be also noticed that the tracker driven algorithm for non-isolated electrons brings, if applied, an efficiency enhancement on isolated electrons too, in particular in the ECAL cracks regions ($\eta \simeq 0$ and $|\eta| \simeq 1.5$) and as expected, at low p_T^e values.
The trajectories in the silicon tracker volume are reconstructed using a dedicated modeling of the electron energy loss and fitted with a Gaussian Sum Filter, which relies on a modeling of electron radiative energy loss. The seeding algorithm combines the information from pixel and TEC layers so to get an efficiency gain in the forward region where the coverage by the pixel layers is limited. To perform the selection, a matching between superclusters and trajectory seeds built from hit pairs or triplets is required.
The electron momentum is estimated by combining the tracker and ECAL measurements. The electron candidates preselection is performed applying loose cuts on track-cluster matching observables, so preserving a high efficiency value while removing part of QCD background.
To resolve ambiguous cases (due to conversion legs of radiated photons) in which several tracks are reconstructed, a cleaning procedure is carried out. The mis-identification arising from the early conversions of radiated photons is coped with electron charge determination which is performed comparing different charge measurement observables.

[1] The aim of the CMS particle flow event-reconstruction algorithm is to identify and reconstruct individually each particle arising from the LHC proton-proton collision, by combining the information from all subdetectors.
The resulting global event description leads to an improved performance for the reconstruction of jets and for the identification of electrons, muons, and taus.

Electrons are classified using observables sensitive to the bremsstrahlung emission and showering in the tracker material.
For the analysis presented here, the electron candidates have been required to have transverse momentum p_T^e larger than 7 GeV/c and a reconstructed $|\eta^e|$ <2.5. The reconstruction efficiency for isolated electron is expected to be above $\approx$ 90% over the full ECAL acceptance, apart from some narrow "crack" regions. The one for basic electron objects integrated over the acceptance rises to reach $\approx$ 90% at $p_T = 10$ GeV/c, and then more slowly to reach a plateau of $\approx$ 95% for $p_T^e = 30$ GeV/c. The application of identification requirements on top of the reconstructed electron objects collection allows the enhancement of the purity of the sample of electron candidates.
The electron objects are separated into classes according to the amount of energy lost by bremsstrahlung processes. A series of different cuts are then applied to each category.
The variables used, which are sensitive to bremsstrahlung processes, are the fraction of radiated energy as measured from the innermost and outermost state of the electron track and the ratio E/p between the supercluster energy and the measured track momentum at the vertex. This procedure allows to handle the non gaussian fluctuations induced on the ECAL and on the tracker measurements by the presence of material in the tracker. Three different categories are defined with quite different measurement characteristics and purity:

- “brem";
- “lowbrem";
- “badtrack".

In addition, two others categories are defined to separate electron objects in transition regions:

- "pure tracker-driven objects";
- "crack objects".

Subcategories are also defined for "brem", "lowbrem", "badtrack" and "pure tracker-driven" objects according to pseudorapidity regions (barrel and endcap), leading to a total number of nine categories. The cuts on each category are applied just to optimize the signal to background ratio (s/b).
As previously mentioned, the shape of most of the discriminating variables strongly depends on the transverse energy (E_T) of the electron, and so the selection cuts are made E_T-dependent.
The cuts are defined for the following variables:

- $|\Delta\eta_{\rm in}| = |\eta_{\rm sc} - \eta_{\rm in}^{\rm extrap.}|$, where $\eta_{\rm sc}$ is the energy weighted position in η of the supercluster and $\eta_{\rm in}^{\rm extrap.}$ is the η coordinate of the position of closest approach to the supercluster position, extrapolating from the innermost track position and direction;

- $|\Delta\phi_{\text{in}}| = |\phi_{\text{sc}} - \phi_{\text{in}}^{\text{extrap.}}|$, where $\Delta\phi_{\text{in}}$ is a quantity similar to the previous one but in azimuthal coordinates;
- $E_{\text{seed}}/p_{\text{in}}$, where E_{seed} is the seed cluster energy and p_{in} the track momentum at the innermost track position;
- H/E: ratio of energy deposited in the Hadronic Calorimeter directly behind the ECAL cluster (H) and the energy of the electron supercluster (E);
- $\sigma_{i\eta i\eta}$: supercluster η width taken from cluster shape covariance.

Electron charge mis-identification has been measured on 2010 data using Z events and a charge mis-ID of 0.004 ± 0.001 (0.028 ± 0.003) has been measured in the ECAL barrel (ECAL endcaps) in very good agreement with the simulation and no significant p_T dependency has been observed in the range of on-shell Z boson decays.
The electron classification also allows the identification of electrons accompanied by low bremsstrahlung with smaller measurement error.

4.1.2 Muon Reconstruction

Muon detection and reconstruction [80] requires an excellent detection of muons over the full acceptance of the CMS detector. The muon recostruction is performed combining the information of the tracking and the calorimeter devices.
Three are the high level physics objects (particles travelling through the detector) involved, depending on the muon track reconstruction step:

- *stand-alone muons*;
- *global muons*;
- *tracker muons*.

The reconstruction starts in the muon spectrometer with the reconstruction of the hits positions in DT, CSC and RPC chambers. A hit corresponds to a signal from a particle recorded by detector components. So the signal is reconstructed as individual points in space called *recHits*. The hits in DT and CSC are then associated together to form "segments" (track stubs).
The seeding procedure is then accomplished collecting and matching the segments. The seeds are so used to perform the actual track fit using also the RPC hits. The reconstructed track that results in the muon spectrometer is called *stand-alone muon*. The *stand-alone muon* tracks are then matched with those from the silicon tracker, generating *global-muon* tracks.
The third high level object, the *tracker muon*, is reconstructed with an algorithm which starts from the silicon tracker tracks and then requests the matching with segments in the muon chambers. From all these three kinds of objects a unique collection of muons is obtained.
The muon system in CMS has three distinct functions:

- muon identification
- momentum measurement
- triggering over the whole kinematic range

The hits are analyzed using a recognition algorithm to associate measurements with trajectories. The procedure used to extract tracks from hits consists of the following steps:

- trajectory seeding: the determination of the track reconstruction initial point is accomplished using an estimated trajectory state or collection of hits compatible with the assumed physics process;
- trajectory building: it starts from the position specified by the trajectory seed, proceeding in the direction specified by the seed and searching for compatible hits on the subsequent detector layers. The track finding and fitting procedure is performed using a Kalman Filter. This last method uses an iterative approach to update the trajectory estimate, using track parameters and covariance and propagating them to the next detector layer;
- trajectory cleaning: the trajectory building procedure produces a large amount of possible trajectories, sharing a lot of hits. This step is then finalized to remove the ambiguities, keeping a maximum number of tracks candidates;
- trajectory smoothing: in this step a backward fitting is performed. The Kalman filter is also used in this case because of its feature of being linear in the measurements and its backward complement capability to use the whole information package;

Once the hits are fitted and the fake trajectories are removed, the remaining tracks are extrapolated to the point of closest approach to the beam line, where the information from the track is measured in the transverse plane.
In order to improve the p_T resolution, a beam spot constraint has been applied. It can be observed that the muon tracks are not re-fitted to the common vertex.
The reconstructed muons are required to have transverse momentum p_T^μ larger than 5 GeV/c and $|\eta^\mu| < 2.4$: the first requirement ensures an efficiency for the reconstruction of muons above the 80%, while the second relies on the geometric acceptance of the tracker detector where the muon reconstruction is fully efficient.
For ensuring an accurate measurement of the track momentum, more than ten silicon tracker hits ($N_{hits} > 10$) have to be included in the track fit.

Detailed study on muons

My personal contribution on the study of the objects used for this analysis is a detailed on monitoring of muon identification, reconstruction and isolation

perfomances at CMS.
The results of this motoring have been reported in the official database of the CMS Collaboration"PdmV Validation DB", where the distributions of all the variables corresponding to *standalone*, *tracker* or *global* muons are compared using two different and, most of the time, consecutive releases of the official CMS software (CMSSW). The main aim of this work is a continuous monitoring of these performances so that possible issues can be identified and solved passing from one version to another. Figure 4.1 shows as an example the comparison of the distributions of some variables related to the reconstructed muons in the reference release (blu markers) and in the new one (red markers). As it can be observed, all the plots show relevant discrepancies passing from one sofware version to another. In particular, for the validation campaign these distribution belong to (CMSSW_9_4_0_pre1_design versus CMSSW_9_3_0_pre5_design), the differences have been marked as "changes expected" due to changes in the tracker seeding (see Subsection 4.1.2) which can cause a different number of tracks depending on the values of η. This thing implies that all the distributions of the variables - such as muon fake rate vs η (top left), the number of the hits in the trackers vs η (top right) the resolution of the impact parameter along the transverse plane (center left) and longitudinal one (center right), the χ^2 of the standalone muons (bottom left) and the pt resolution of the global muons (bottom right) - related to the tracks should change.

4.2 Electrons

The electrons used for the analysis are classified as *loose* and *tight* electrons according to the cuts applied. Electrons, which are preselected using loose cuts on track-cluster matching observables, are called *loose* electrons.
This first selection step preserves the highest possible efficiency while rejecting part of the QCD background. Loose electrons must satisfy the following properties:

- transverse momentum $p_T > 7$ GeV,
- pseudorapidity $|\eta| < 2.5$,
- transverse impact parameter $|d_{xy}| < 0.5$ and longitudinal impact parameter $|d_z| < 1$.

4.2.1 Electron Reconstruction

The early runs of Run 2 during 2016 data-taking exhibit a tracking inefficiency originating from a reduced hit reconstruction efficiency in the strip detector, called the *Highly Ionizing Effect* (HIP) effect.
As a result, a discrepancy in data-MC agreement was observed. This mismatch was corrected using data to MC scale factors as it is done for the

electron selection with efficiencies measured in data using the same tag-and-probe technique described in section 4.2.5. These studies are carried out by the EGM Particle Object Group and electron reconstruction scale factors as a function of the super cluster η are derived and used for this analysis (it was shown that the p_T dependence of the scale factor is negligible).
More details on electron reconstruction can be found in Ref. [87].

4.2.2 Electron Identification

All the reconstructed electrons are identified by means of a multivariate classifier algorithm, also known as Gradient Boosted Decision Tree (GBDT).
This algorithm exploits observables from the electromagnetic cluster, the matching between the cluster and the electron track as well as observables based exclusively on tracking measurements.
The classifier is trained on a Drell-Yan+jets MC sample for both signal and background. The performance of the BDT has been improved studying the new 2016 running conditions (e.g. a new splitting of the BDT training bins has been considered where electrons falling into the gap regions of the ECAL - EB-EE transition region - were trained separately from the electrons away from the gap regions).
Finally, the setup currently used has been found already able to properly take into account the significantly different input distributions in those regions.
Additional variables were also studied, including more cluster-shape observables. None of those variables helped improving the efficiency in the relevant $> 95\%$ signal efficiency region, although an improved background rejection of up to 20% was found for the 80% working point.

Table 4.1 summarizes the full list of observables used as input to the classifier and Table 4.2 lists the cut values applied to the BDT score for the chosen working point. For the analysis, **tight electrons** as the loose electrons that pass this MVA identification working point.

observable type	observable name
6*cluster shape	RMS of the energy-crystal number spectrum along η and φ; $\sigma_{i\eta i\eta}$, $\sigma_{i\varphi i\varphi}$
	super cluster width along η and ϕ
	ratio of the hadronic energy divided by the supercluster energy, H/E
	circularity $(E_{5\times5} - E_{5\times1})/E_{5\times5}$
	sum of the seed and the adjacent 9 crystals divided by the supercluster energy, R_9
	for Endcap training bins: energy fraction in pre-shower, E_{PS}/E_{raw}
2*track-cluster matching	energy-momentum agreement E_{tot}/p_{in}, E_{ele}/p_{out}, $1/E_{tot} - 1/p_{in}$
	position matching $\Delta\eta_{in}$, $\Delta\varphi_{in}$, $\Delta\eta_{seed}$
5*tracking	fractional momentum loss $f_{brem} = 1 - p_{out}/p_{in}$
	number of hits of the KF and GSF track N_{KF}, N_{GSF} $(\cdot)$
	reduced χ^2 of the KF and GSF track χ^2_{KF}, χ^2_{GSF}
	number of expected but missing inner hits $(\cdot)$
	probability transform of conversion vertex fit χ^2 $(\cdot)$

TABLE 4.1: Overview of input variables to the electron identification classifier. Variables not used in the Run I MVA are marked with $(\cdot)$.

minimum BDT score	$\|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$\|\eta\| > 1.479$
$5 < p_T < 10$ GeV	-0.211	-0.396	-0.215
$p_T > 10$ GeV	-0.870	-0.838	-0.763

TABLE 4.2: Minimum BDT score required for passing the electron identification.

4.2.3 Electron Isolation

The relative isolation for electrons is defined as:

$$\text{RelPFiso} = (\sum_{\text{charged}} p_T + \sum_{\text{neutral}}^{\text{corr}} p_T)/{p_T}^{\text{lepton}}. \tag{4.1}$$

where the corrected neutral component of the isolation[2] is computed using the formula:

$$\sum_{\text{neutral}}^{\text{corr}} p_T = \max(\sum_{\text{neutral}}^{\text{uncorr}} p_T - \rho \times A_{\text{eff}}, 0). \tag{4.2}$$

and the mean pile-up contribution to the isolation cone is obtained as:

$$PU = \rho \times A_{\text{eff}} \tag{4.3}$$

where ρ is the mean energy density in the event and the effective area A_{eff} is defined as the ratio between the slope of the average isolation and that of ρ as a function of the number of vertices.

The electron isolation working point was optimized in Ref. [83] and the electron isolation working is chosen to be $\text{RelPFiso}(\Delta R = 0.3) < 0.35$.

4.2.4 Electron Energy Calibrations

Electrons in data are corrected for features in ECAL energy scale in bins of p_T and $|\eta|$. Corrections are calculated on a $Z \to ee$ sample to align the di-electron mass spectrum in the data to the one in the simulation, and to minimize the width of the distribution.
The $Z \to ee$ mass resolution in simulation is made to match data by applying a pseudo-random Gaussian smearing to electron energies, with Gaussian parameters varying in bins of p_T and $|\eta|$. This results in a convolution of the electron energy spectrum with a Gaussian.

[2]The neutral component of the isolation corresponds to the neutral hadrons identified by the Particle Flow algorithm

4.2.5 Electron Efficiency Measurements

Once the final object selection is defined, it is important to study how efficient that selection is. Many bias sources (such as trigger inefficiency, misidentified hadrons(jet), photon conversions, non-isolated electrons from heavy-flavors decay) could result in a wrong evaluation of electron efficiency itself. The most common way to proceed is to use the so called *tag and probe* method, in which the dilepton decays of some measured resonances (e.g. Z, JΨ are exploited. If the invariant mass of lepton pairs is plotted around the Z mass, the result is a peak at the mass of the Z with a full width half maximum (FWHM) approximately given by the width of the Z.
The idea of tag and probe is that such a pair of electrons are formed by one electron which is the tag and the other one which is the *probe*. The tag is a fully identified electron, constrained by a very tight selection on it so that it is almost certain that a real electron corresponds to it, while the *probe* is only given a loose selection. The constraint for the tag and probe pair is to have an invariant mass close to the Z nominal mass. This condition results in a good probability also for the probe to be a real electron and, when this basic conditions are applied, the method has to check if the *probe* can be identified as an "full" electron. To accomplish this, the probe can be subjected to various cuts which are used in the analysis to identify electrons and the number of probes passing and failing this constraints provides the efficiency of the cut to select real electrons.
For the analysis here presented, electrons on which tag-and-probe (T&P) method was applied have been used. It was performed on the single electron primary datasets listed in Table 3.1 using a JSON corresponding to 36.8 fb^{-1}. More details on the tag-and-probe method applied can be found in [83]. Tagged electrons used in this study need to satisfy the following quality requirements:

- trigger matched to HLT_Ele27_eta2p1_WPTight_Gsf_v*;
- $p_T > 30$ GeV, super cluster (SC) $|\eta| < 2.1$ but on in EB-EE gap ($1.4442 < |\eta| < 1.566$);
- "tight" working point of the Spring16 cut-based electron ID.

Probe electrons only need to be reconstructed as Gsf Electron. The Final State Radiation (FSR) recovery[3] algorithm used in the main analysis is used consistently throughout the efficiency measurement: the isolation is calculated after removing any FSR photon matched to the electron and the electron p_T itself, and the di-electron invariant mass is estimated by including the FSR photons.

[3]The identification of photons from final state radiation is very important in this analysis for two main reasons: the inclusion of FSR photons in the computation of the invariant mass of the system results in a sharper peak for the Higgs boson signal compared to what achievable using only the momenta of the four leptons and also the removal of FSR photons from the isolation cone of the leptons improves the selection efficiency [88].

The nominal MC efficiencies are evaluated from the LO MadGraph DY sample, while the NLO systematics use the 0,1 jet MadGraph_AMCatNLO sample listed in Table 3.8. The m_{ee} signal shape of the passing and failing probes is taken from MC and convoluted with a Gaussian. The data are then fitted with the convoluted MC template and a CMSShape (an error-function with a one-sided exponential tail). All the prescription followed to apply the Tag and Probe method are due to the usage of the T&P tool developed by the EGM POG.

The electron selection efficiency is measured as a function of the probe electron and η, and separately for electrons falling in the ECAL gaps.

Figure 4.2 shows the p_T turn-on curves measured in data; final scale factors for the efficiencies are derived and used within the analysis.

The EGM recommendations on the evaluation of tag-and-probe uncertainties for efficiency measurements are followed.

4.3 Muons

4.3.1 Muon Reconstruction and Identification

In this analysis, *loose* muons are defined as the muons that satisfy $p_T > 5$ GeV, $|\eta| < 2.4$, $d_{xy} < 0.5$cm, $d_z < 1$cm, where d_{xy} and d_z are the transverse and the longitudinal impact parameters with respect to the primary vertex, and using the 'muonBestTrack' prescription. This last constraint implies that muons have to be reconstructed by either the Global Muon or Tracker Muon algorithm. If a muon is classified as a Tracker muon, it must be arbitrated (i.e. the number of matches between the tracker and the cluster object must be greater than 2). Standalone Muon tracks, which are only reconstructed in the muon system, must be rejected. Standalone Muons are discarded even if they are marked as global or tracker muons. Loose muons with p_T below 200 GeV are considered **tight muons** if they also pass the Particle Flow (PF)[4] [81] muon ID (note that the naming convention used for these IDs differs from the muon POG naming scheme, in which the "tight ID" used here is called the "loose ID").
Loose muons with p_T above 200 GeV are considered tight muons if they pass the PF ID or the Tracker High-p_T ID, the definition of which is shown in Table 4.3.
When a very heavy resonance decays into two Z bosons, both bosons are very boosted and the leptons coming from the decay of a highly boosted Z are nearly collinear in the laboratory frame, so that PF ID loses efficiency for muons separated by approximately $\Delta R < 0.4$, which roughly corresponds to muons originating from Z bosons with $p_T > 500\ GeV$.

[4]Particle Flow is a CMS algorithm The Particle Flow (PF) algorithm whose aim is to identify and reconstruct all the particles produced by the collisions by combining optimally the information of the different subdetectors.

TABLE 4.3: The requirements for a muon to pass the tracker high-$p_T t$ ID. Note that these are equivalent to the Muon POG high-p_T ID with the global track requirements removed.

Plain-text description	Technical description
Muon station matching	Muon is matched to segments in at least two muon stations **NB: this implies the muon is an arbitrated tracker muon.**
Good measurement	$\frac{p_T}{\sigma_{p_T}} < 0.3$
Vertex compatibility $(x-y)$	$d_{xy} < 2$ mm
Vertex compatibility (z)	$d_z < 5$ mm
Pixel hits	At least one pixel hit
Tracker hits	Hits in at least six tracker layers

An additional "ghost-cleaning" step is performed to deal with situations when a single muon is incorrectly reconstructed as two or more muons. More details on muon reconstruction can be found in Ref. [83].

4.3.2 Muon Isolation

The PF-based isolation, described for electrons in Section 4.2.3, is also used for the muons. The only difference with respect to the electrons is the way the pileup contribution is subtracted; for the muons, the $\Delta\beta$ correction is applied, where $\Delta\beta = \frac{1}{2}(\sum_{\text{PU}}^{\text{charged had.}} p_T)$ gives an estimate of the energy deposit of neutral particles (hadrons and photons) from pileup vertices. The relative isolation for muons is then defined as:

$$RelPFiso = \frac{\sum^{\text{charged had.}} p_T + \max(\sum^{\text{neutral had.}} E_T + \sum^{\text{photon}} E_T - \Delta\beta, 0)}{{p_T}^{\text{lepton}}} \tag{4.4}$$

The muon isolation was optimized in Ref. [83] and the working point is chosen to be the same as for the electrons, namely RelPFiso$(\Delta R = 0.3) < 0.35$.

4.3.3 Muon Efficiency Measurements

Muon efficiencies are measured with the tag-and-probe (T&P) method, already discussed for the electrons, performed on $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ events. More details on the methodology can be found in Ref. [83]. The Z sample is used to measure the muon reconstruction and identification efficiency at high p_T and the efficiency of the isolation and impact parameter requirements at all p_T values, while the J/Ψ sample is used to measure the reconstruction efficiency at low p_T, as it benefits from a better purity in that kinematic regime.

Reconstruction and identification

Results for the muon reconstruction and identification efficiency for $p_T > 20$ GeV have been derived by the Muon Particle Object Group.
The probe in this measurement are tracks reconstructed in the inner tracker, and the passing probes are those that are also reconstructed as a global or tracker muon and passing the Muon POG Loose muon identification. Results for low p_T muons were derived using events, with the same definitions of "probe" and "passing probe". The systematic uncertainties are estimated by varying the analytical signal and background shape models used to fit the dimuon invariant mass.
The efficiency and scale factors used for low p_T muons are the ones derived using single muon prompt-Reco dataset.
The efficiencies in data and in simulation are shown in Figure 4.3.

Impact parameter requirements

The measurement is performed using Z events selected with the `|HLT_IsoMu20_v*|` or `|HLT_IsoMu22_v*|` triggers. For this measurement, the probe is a muon passing the POG "loose" identification criteria, and it is considered a "passing" probe if it satisfies the SIP3D, d_{xy}, d_z cuts of this analysis.
Figure 4.4 shows the efficiency of the muon impact parameter requirements.

Isolation requirements

The isolation efficiency is measured using events from the Z decay corresponding to any p_T. The isolation of the muons are calculated after recovery of the FSR photons (see 4.4) and subtracting their contribution to the isolation cone of the muons.
Figure 4.5 shows the efficiency of the muon isolation requirement in the barrel and in the endcaps.

Tracking

The efficiency to reconstruct a muon track in the inner detector is measured using as probes the tracks reconstructed in the muon system alone.
The efficiency and data-to-simulation scale factors are measured from Z events as a function of η, for $p_T > 10$ GeV and $p_T < 10$ GeV. The values of these scale factors used in this analysis are those from the ReReco version of the full dataset collected in 2016.
In Figure 4.6, the tracking efficiency in data and in simulation as a function of η is shown.

Overall results

Finally, the overall product of all data-to-simulation scale factors for muon tracking, reconstruction, identification, impact parameter, and isolation requirements is used for this analysis.

4.4 Photons and FSR Recovery

The problem of the inclusion of the Final State Radiation photons has been already mentioned in 4.2.5.
The selection of FSR photons is done per-lepton and does not depend on any Z mass criterion.
In the association of photons with leptons, the rectangular cuts on $\Delta R(\gamma, \ell)$ and $E_{T,\gamma}$ separately have been replaced by a single cut on the object $\Delta R(\gamma, \ell)/E^2_{T,\gamma}$.
The selection of the photons, starting from the collection of Particle Flow photons (provided by the PF algorithm), and also the selection of their association to a lepton proceeds follows these steps:

1. The preselection of PF photons is done by requiring $p_{T,\gamma} > 2$, $|\eta^{\gamma}| < 2.4$, and a relative PF isolation smaller than 1.8. The latter variable is computed using a cone of radius $R = 0.3$, a threshold of 0.2 on charged hadrons with a veto cone of 0.0001, and 0.5 on neutral hadrons and photons with a veto cone of 0.01, also including the contribution from pileup vertices (with the same radius and threshold as per charged isolation);

2. Supercluster veto: we remove all PF photons that match with any electron passing both the loose ID and SIP cuts. The matching is performed by directly associating the two PF candidates;

3. Photons are associated to the closest lepton in the event among all those that pass both the loose ID and SIP cuts;

4. We discard photons that do not satisfy the cuts $\Delta R(\gamma, \ell)/E^2_{T,\gamma} < 0.012$ and $\Delta R(\gamma, \ell) < 0.5$;

5. If more than one photon is associated to the same lepton, the one with the lowest $\Delta R(\gamma, \ell)/E^2_{T,\gamma}$ is selected;

6. For each FSR photon that was selected, we exclude that photon from the isolation sum of all the leptons in the event that pass both the loose ID and SIP cuts. This concerns the photons that are in the isolation cone and outside the isolation veto of said leptons ($\Delta R < 0.4$ AND $\Delta R > 0.01$ for muons, and $\Delta R < 0.4$ AND ($\eta^{\mathrm{SC}} < 1.479$ OR $\Delta R > 0.08$) for electrons).

All the details on the optimization of the FSR photon selection can be found in [82, 83].

4.5 Jets

After the rejection of the charged hadrons that are associated to a pileup primary vertex, the jets are reconstructed by using the anti-k_T clustering algorithm [89] out of PF candidates.
In order to reduce instrumental background, the loose working point for the jet identification suggested by the JetMET Particle Object Group is applied.
Jets are required to have p_T >30 GeV and $|\eta| < 4.7$ and are cleaned from any of the tight leptons (passing the SIP and isolation cut computed after FSR correction) and FSR photons by a separation criterion defined as
$\Delta R(\text{jet,lepton/photon}) > 0.4$.
However, it must be noticed that the calorimeter response to particles is not linear and Jet Energy Corrections (JEC) are needed to "translate" the measured jet energy to the true particle or parton energy.
Standard JEC are applied to all the reconstructed jets.

B-tagging

In order to reject some +vector boson background events in this analysis a b-tagging discriminator is used to tag b-jets.
The *Combined Secondary Vertex* (CSV) algorithm is used as our b-tagging algorithm. It combines information about impact parameter significance, the secondary vertex and jet kinematics. The variables are combined using a likelihood ratio technique to compute the b-tag discriminator. In this analysis, a jet is considered to be b-tagged if it passes a well-defined working point corresponding to a value of the discriminator larger than 0.8484 [90].
Data-to-simulation scale factors for b-tagging efficiency are provided for this working point for the full dataset as a function of jet , η and flavour.

4.6 Missing Transverse Energy

The missing transverse energy, denoted as E_T^{miss} or MET, has a very special place in the physics object description of this Mono-Higgs analysis, being tightly related to dark matter evidence. It is defined as the magnitude of the imbalance in the transverse momentum of all the *visible* particles, which interact with the electromagnetic or strong forces, in the final state of collisions. Since, as well known, momentum is conserved in each direction, MET is the transverse momentum that must have been carried by *invisible* particles (e.g. neutrinos). Therefore, MET is an estimate of transverse momentum of neutrinos. We use MET in measurements of W bosons, top quarks, and tau leptons as these particles can decay into neutrinos. Further, many models of physics beyond the Standard Model predict the existence of particles or something that is invisible and can carry momentum like Dark Matter candidates or supersymmetric particles, unparticle models. MET is consequently used to test

the related models.
A thorough reconstruction of MET turns to be quite demanding since it relies on an accurate reconstruction of all invisible particles in a certain event. A very *hermetic* detector which can detect all the particles interacting via electromagnetic or strong forces are then mandatory. The highly granular electromagnetic calorimeter, the hermetic hadronic calorimeter, a redundant muon systems and a silicon tracker surrounded by a strong magnetic field (4 T) of the CMS detector ensure that this requirement is satisfied.

Fake MET and MET Filters

An evidence of large values of missing transverse energy observed in some events could not be only due to the production of invisible particles, but more often to uninteresting causes such as detector noise, cosmic rays, and beam-halo particles. MET with uninteresting causes is called false MET, anomalous MET, or *fake* MET. Then, for an accurate reconstruction of MET, it is not sufficient to reconstruct all visible particles produced in collisions.
In order to identify the *fake* MET in the events a series of MET Filters based on appropriate algorithms have been implemented and recommended by the CMS Jet MET Particle Object Group for Run 2 analyses which use missing transverse energy [91]. These algorithms use timing, pulse shape, and topology of signal. After the identified false MET is removed by the application of the filters, the agreement of the MET spectrum with MC typically improves significantly. The MET filters applied for this Mono-Higgs analysis are the following:

- HBHENoiseFilter;
- HBHENoiseIsoFilter;
- EcalDeadCellTriggerPrimitiveFilter;
- goodVertices;
- eeBadScFilter;
- globalTightHalo2016Filter;
- BadPFMuonFilter;
- BadChargedCandidateFilter.

The first two filters (*HBHENoiseFilter* and *HBHENoiseIsoFilter*) remove noisy events from the HCAL, where the HBHE (HB stands for HCal Barrel and HE stands for HCal Endcap) scintillator produce anomalous signals with pulse shapes and pixel multiplicities different from those coming from a clean signal.

The *EcalDeadCellTriggerPrimitiveFilter* module removes events with non-functioning ECAL data links, comparing the sum of energy deposited in each supercluster cell to the trigger primitive saturation energy.

The *goodVertices* filter events with noisy vertex reconstruction from pileup effects by requiring the reconstruction of at least one good vertex fullfilling the following criteria: high number of degree of freedom ($N_{PV} > 4$), collisions restricted along the $z-$axis ($z_{PV} < 24$ cm) and small radius of the PV ($r_{PV} < 2$ cm).

The *eeBadScFilter* filter removes events with noisy ECAL Endcap superclusters, while the *globalTightHalo2016Filter* filter removes events with enhanced MET from beam-halo particles which are in time with the beam.

The last two filters, the *BadPFMuonFilter* and *BadChargedCandidateFilter* remove events with mis-reconstructed muon and charged hadron PF candidates.

Raw MET

The MET object denoted as Particle Flow MET (PFMET) in this analysis is indeed the *raw* MET. It is defined as the magnitude of the negative vectorial sum of the transverse momentum of all reconstructed particle flow candidates:

$$\overrightarrow{\not{E}_T}^{raw} = -\sum_{i \in \text{all}} \overrightarrow{p}_{T,i} \tag{4.5}$$

The vector quantity that is the negative sum of reconstructed particle momenta is the most appropriate definition of missing transverse momentum, although this term is used to denote its magnitude as well. The raw MET is systematically different from true MET, i.e. the transverse momentum carried by invisible particles as previously defined, for many reasons such as the non-compensating nature of the calorimeters and detector misalignment. In order to make MET a better estimate of true MET, MET *corrections* can be applied. In the analysis here described only one correction, called Type-I, has been applied and it is described in the following subsection.

Type-I correction

The Type-I correction is a propagation of the jet energy corrections (JEC) to MET. The Type-I correction replaces the vector sum of transverse momenta of particles which can be clustered as jets with the vector sum of the transverse momenta of the jets to which JEC is applied.

The particles can be classified into two disjoint sets: clustered (jets) or unclustered. The raw MET can be rewritten as follows:

$$\overrightarrow{E_T}^{raw} = -\sum_{i\in\text{jets}} \overrightarrow{p}_{T,i} - \sum_{i\in\text{uncl}} \overrightarrow{p}_{T,i}. \tag{4.6}$$

The first vector sum can be written also as:

$$\sum_{i\in\text{jets}} \overrightarrow{p}_{T,i} = \sum_{jet} \overrightarrow{p}^{raw}_{T,jet}. \tag{4.7}$$

Then it can be said that the vector sum of the p_T of the clustered particles is the same as the vector sum of the p_T of all the jets. In equation 4.7 the superscript "raw" means that no jet energy corrections are applied to the jets.

The Type-I correction replaces the raw jet $\overrightarrow{p}_T$ with the corrected jet $\overrightarrow{p}_T$ so it can be written as the difference between two vector sums:

$$\vec{C}_{\text{T}}^{\text{Type-I}} = \sum_{\text{jet}} \vec{p}_{\text{T jet}}^{\text{raw}} - \sum_{\text{jet}} \vec{p}_{\text{T jet}}^{\text{JEC}} \tag{4.8}$$

Using the equivalence 4.7, the equation 4.8 can be also written as follows:

$$\vec{C}_{\text{T}}^{\text{Type-I}} = \sum_{i\in\text{jets}} \vec{p}_{\text{T}i} - \sum_{\text{jet}} \vec{p}_{\text{T jet}}^{\text{JEC}}. \tag{4.9}$$

The equation 4.9 is exactly the vector term that has to be added to the raw MET to produce a corrected MET:

$$\vec{E}_{\text{T}}^{\text{Type-I}} = \vec{E}_{\text{T}}^{\text{raw}} + \vec{C}_{\text{T}}^{\text{Type-I}}. \tag{4.10}$$

Substituing 4.6 and 4.9 in equation 4.10 the Type-I corrected MET can be written:

$$\vec{E}_{\text{T}}^{\text{Type-I}} = -\sum_{\text{jet}} \vec{p}_{\text{T jet}}^{\text{JEC}} - \sum_{i\in\text{uncl.}} \vec{p}_{\text{T}i}. \tag{4.11}$$

Systematic uncertainties related to the modelling of the real MET are obtained by varying the JEC and jet energy resolution (JER) and measuring the propagation of these variations to the MET uncertainty. These measurements are described in greater detail in the Section 5.5.

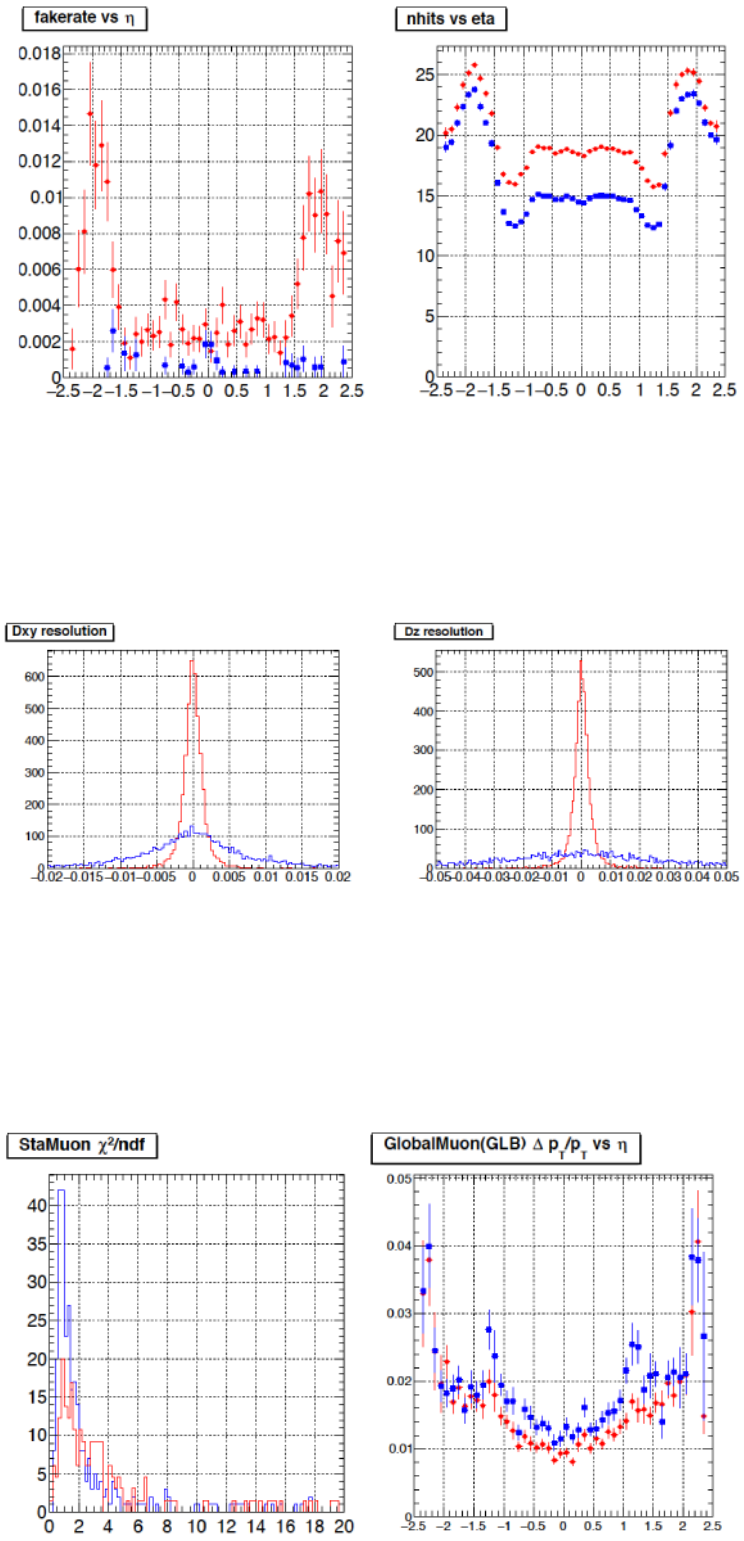
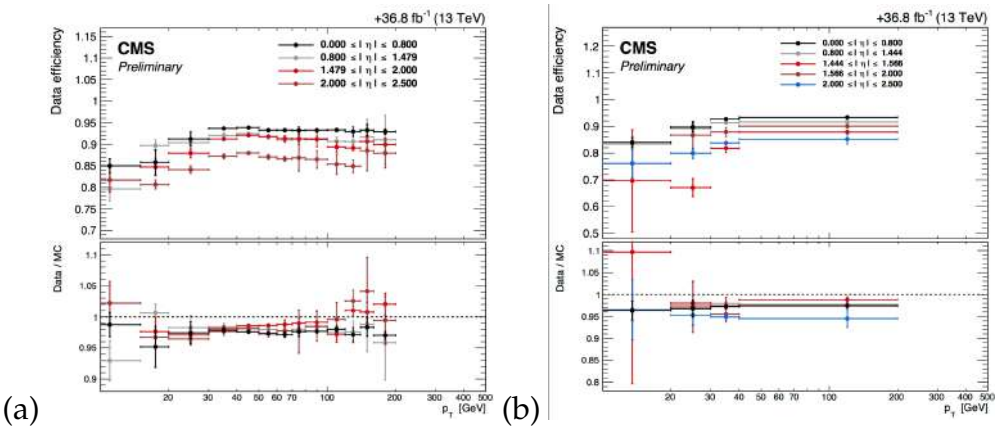
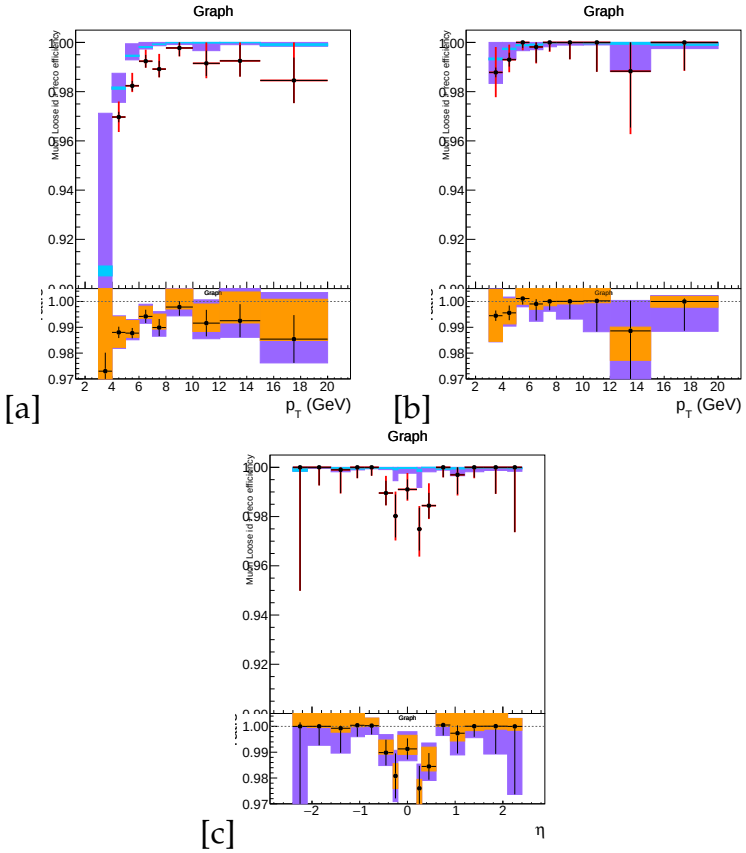


FIGURE 4.1: Comparison of the distributions of some variables related to the tracks in the official CMS muon validation campaign CMSSW_9_4_0_pre1_design versus CMSSW_9_3_0_pre5_design. From top to bottom: muon fake rate vs η (top left), the number of the hits in the trackers vs η (top right) the resolution of the impact parameter along the transverse plane (center left) and longitudinal one (center right), the χ^2 of the standalone muons (bottom left) and the pt resolution of the global muons (bottom right).



(a) (b)

FIGURE 4.2: Electron selection efficiencies measured using the tag-and-probe technique described in the text, non-gap electrons (a) and gap electrons (b).



[a] [b] [c]

FIGURE 4.3: Muon reconstruction and identification efficiency at low p_T, measured with the tag-and-probe method on J/Ψ events, as a function of p_T in the Barrel (a) and Endcaps (b, and as a function of η for p_T> 7 GeV (c). In the upper panel, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. In the lower panel showing the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematical uncertainty, and the violet rectangles include both uncertainties.

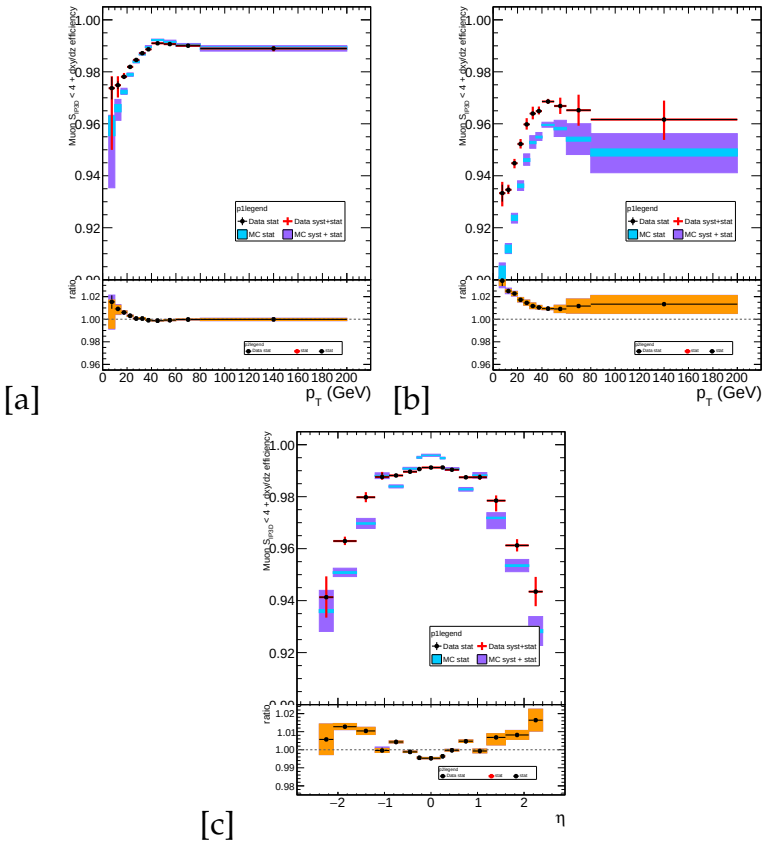
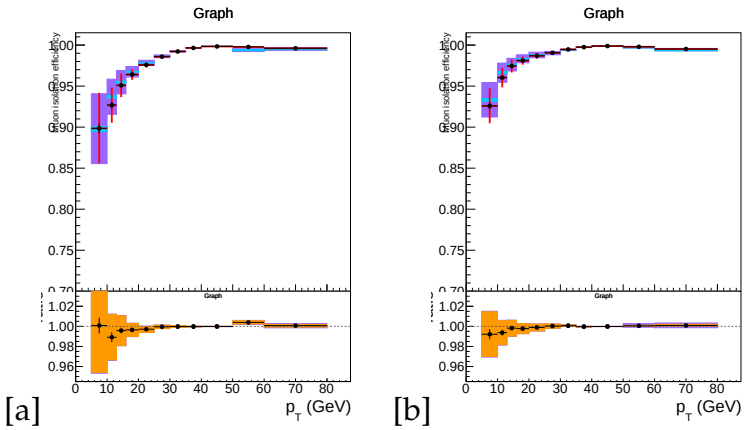
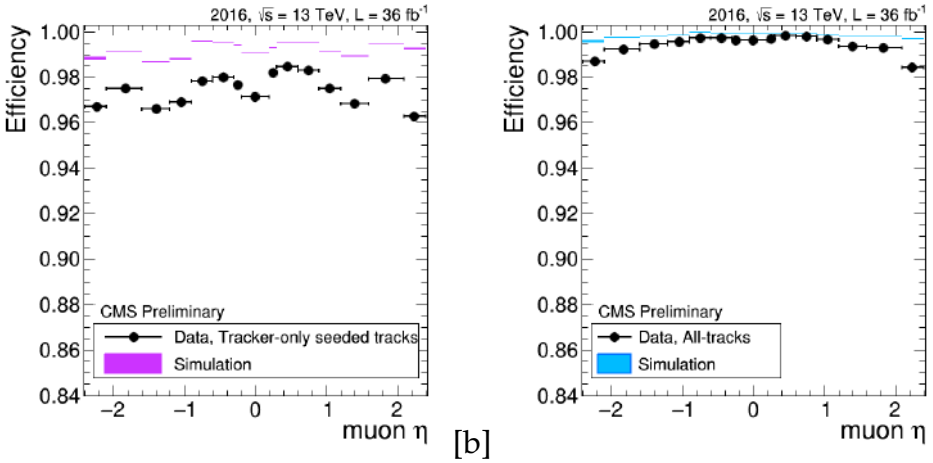


FIGURE 4.4: Efficiency of the muon impact parameter requirements, measured with the tag-and-probe method on events, as a function of p_T in the Barrel (a) and Endcaps (b), and as a function of η for $p_T > 20$ GeV (c). In the upper panel, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. In the lower panel showing the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematical uncertainty, and the violet rectangles include both uncertainties.



[a] [b]

FIGURE 4.5: Efficiency of the muon isolation requirement, measured with the tag-and-probe method on Z events, as a function of p_T in the Barrel (left) and Endcaps (right). In the upper panel, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. In the lower panel showing the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematical uncertainty, and the violet rectangles include both uncertainties.



[a] [b]

FIGURE 4.6: Tracking efficiency in data and in simulation as a function of η for muon $p_T < 10$ GeV (a) and $p_T > 10$ GeV (b) with ReReco data.

Chapter 5

The Mono-Higgs Analysis

This chapter focuses on the analysis selection for the Mono-Higgs search.
The criteria to select the four lepton objects in the events are described along with the optimized definition of the appropriate signal region (SR) chosen for this analysis.
The choice of the SR turns out to be a fundamental part of it, since the definition of an appropriate strategy to reject the background improves the sensitivity of this study which suffers from a low Higgs to 4l branching fraction ($\sim 10^{-4}$).
Then, in Section 5.4, a background estimation for both reducible and irreducible background samples is described while section 5.5 presents the main sources of systematic uncertainties.

5.1 Event Selection

This section describes the selection of the four lepton candidates that is used to build the Higgs boson which, along with a large value of missing transverse energy, constitutes the final signature of this analysis.
Great care must be taken when choosing these criteria in order to optimize the efficiency of the selection.

5.1.1 Vertex Selection

At least one *good* primary vertex (PV) must be found in each event in order to consider it for the analysis. A PV can be identified as a *good* only if it fullfills the following criteria:

- high number of degree of freedom of the vertex fit ($N_{PV} > 4$);
- collisions restricted along the longitudinal distance w.r.t the nominal interaction point smaller than 24 cm ($z_{PV} < 24$ cm);
- small distance of the PV w.r.t. to the nominal interaction point in the transverse plane less than 2 cm ($r_{PV} < 2$ cm).

This criteria are intended to discard events with fake vertices related to failures of the vertex fit and pileup events.

5.1.2 ZZ Candidate Selection

The four-lepton candidates are built from what we call the **selected leptons**, which are the tight leptons already defined in sections 4.2.2 and 4.3.1). In addition they have to pass a cut on the significance of the impact parametersignificance of the impact parameter is defined as the ratio between the three-dimensional impact paramter and its error (a.k.a. SIP, where |SIP<4|) and they have to be isolated following the critieria explained in Sections 4.2.3 and 4.3.2.
The FSR photons are subtracted as already mentioned in Section 4.4.
A lepton cross cleaning is applied by discarding electrons which are within $\Delta R < 0.05$ of the selected muons. The building of four-lepton candidates proceeds through the following steps:

1. First, the **Z candidates** are defined as a pair of selected leptons of opposite charge (OC) and same flavor (SF), such as e^+e^- or $\mu^+\mu^-$, that satisfy $12 < m_{\ell\ell(\gamma)} < 120$ GeV/c^2. Here the index γ means that the Z candidate mass includes the contribution of the selected FSR photons if there is any.

2. Then the **ZZ candidates** are defined as pairs of non-overlapping Z candidates. The Z candidate with reconstructed mass $m_{\ell\ell}$ closest to the nominal Z boson mass is denoted as Z_1, and the second one is denoted as Z_2. ZZ candidates are required to satisfy the following list of requirements:

   - **Ghost removal** : $\Delta R(\eta, \phi) > 0.02$ between each of the four leptons.
   - **lepton** p_T: at least two of the four selected leptons should pass $p_{T,i} > 20$ and $p_{T,j} > 10$ GeV/c.
   - **QCD suppression**: all four OS pairs that can be built with the four leptons (regardless of lepton flavor) must satisfy $m_{\ell\ell} > 4$ GeV/c^2. Here, selected FSR photons are not used in computing $m_{\ell\ell}$, since a QCD-induced low mass dilepton (eg. J/Ψ) may have photons nearby (e.g. coming from π_0).
   - Z_1 **mass**: $m_{Z1} > 40$ GeV/c^2
   - **'smart cut'**: defining Z_a and Z_b as the mass-sorted alternative pairing Z candidates (where Z_a is the one closest to the nominal Z boson mass), requiring not passing the following condition: $|m_{Za} - m_Z| < |m_{Z1} - m_Z|$ and $m_{Zb} < 12$).
     Here, selected FSR photons are included in m_Z's computations.
     This cut ensures to discard 4μ and $4e$ candidates where the alternative pairing looks like an on-shell Z + low-mass $\ell^+\ell^-$. It also

allows to choose the best ZZ candidate after all kinematic cuts and so after an already optimized selection.

- **four-lepton invariant mass**: $m_{4\ell} > 70$ GeV/c^2
- **choice of the best ZZ candidate**: if more than one ZZ candidate survives the above selection, the one with the highest value of the scalar sum of the four lepton p_T is chosen.

3. **"SM Higgs selection"**: events with at least one selected ZZ candidate passing all the previous selection criteria are then used to tag the Higgs decaying in the four-lepton final state;
4. **Signal Region:** Additional cuts are then added to define the signal region for this analysis with a signature of four-leptons and missing transverse energy, as discussed later in the text (see Sec. 5.3);
5. **Control Region:** Two control regions are built in order to study both the impact of "fake" leptons contributing to the $m_{4\ell}$ spectrum and the modeling of the missing transverse energy distribution (see Sec. 5.4).

5.2 Physics Observables

Several observables have been studied at different selection steps of this analysis. The most important ones are the four-lepton invariant mass, $m_{4\ell}$, and the missing transverse energy, E_T^{miss}.
After the opposite sign dilepton pairs are selected as described in Section 5.1 ("Z candidate", step#1), the agreement between the data and MC distributions has been checked.
From now on, in all the figures the real data will be described by the black markers superimposed to the simulated samples from the SM background (alwasys stacked each other) and for the signals, represented by the two benchmark models chosen the $Z'2HDM$ and $Z'Baryonic$ models.
The distributions are built for events where at least one Z candidate is reconstructed, either decaying to two muons or to two electrons, separately.
Figure 5.1 shows how data are well described by the simulation for the MET. As it can be observed, at this step the signal is completely overwhelmed by the background; for this reason it is superimposed to the E_T^{miss} distribution of the backgrounds and data.
The ratio between the E_T^{miss} distribution for the data and the sum of all the SM backgrounds is found to be almost flat and close to 1 as a function of E_T^{miss}. The agreement is within 10%.
Only few events are found with $E_T^{miss} > 200$ GeV and so there is a large uncertainty on the ratio.

The ZZ candidates are built as described in Section 5.1 and the reconstruction of the Higgs boson in four leptons is checked (SM Higgs selection, step#3). At this step of the selection, the background is modeled well by simulation

as it can be appreciated in the distribution of the four-lepton invariant mass as obtained with real data and with simulation for background and signal in Figure 5.2. The distributions for $4e$, 4μ, $2e2\mu$ are summed together.
A boson with a mass around 125 GeV and compatible with the prediction of the SM Higgs boson is clearly observed in data and confirms the discovery made in 2012. The four-lepton invariant mass of the $Z'2HDM$ and $Z'Baryonic$ signal expected is concentrated around a mass of 125 GeV.
However, due to the much lower cross section, the signals are completely overwhelmed by the background also at this step of the selection.

Figure 5.3 shows the E_T^{miss} distribution as obtained with data and simulation for background and signal processes.
Only a few events are observed in data at large values of MET. This obviously implies tha twe need to rely on the simulation in order to describe the tail of the distribution.
As a consequence, the reasonably good agreement between data and the sum of all SM MC backgrounds can be appreciated only at low E_T^{miss} range.
Furthermore, it can be noticed how the events contributing to the tail come mostly from $t\bar{t}$+ vector boson gauge events, triple gauge boson production VVV (ZZZ, WZZ, WWZ), associated production of the Higgs with vector boson fusion (ZH and WH), and ZZ+jet events.
The very small signal appears at large values of MET, which can be considered a hint for DM candidate production.

Other observables have also been studied at this step of the selection.
The transverse mass for the 4l and MET system, for example, has been considered an interesting observable because it actually combines the information of the four-lepton invariant mass and the E_T^{miss} as follows:

$$m_T(4l + E_T^{miss}) = \sqrt{m_{4l}^2 + 2 \times p_{T,4l} \times E_T^{miss} - \vec{p}_{T,4l} \cdot \vec{E}_T^{miss}} \tag{5.1}$$

where $m_{4\ell}$ and $p_{T,4\ell}$ are is the invariant mass and the transverse momentum of the four lepton candidate system, respectively, and $\vec{p}_{T,4\ell}$ is the vector of the four lepton momentum projected on the transverse plane.
The distribution of this variable is shown in Figure 5.4 (top); the signal distributions are superimposed to the background distribution, and compared to data.

Another useful observable that has been studied is the angle between $\vec{E}_T^{miss}$ and the four-lepton momentum in the transverse plane $\Delta\phi(4\ell - E_T^{miss})$. For some models these vectors are likely to be back-to-back depending on the kinematics of the production.
Figure 5.4 (bottom) shows the distribution of $\Delta\phi(4\ell - E_T^{miss})$ for a specific signal model and the background events.
As it can be observed, the distribution for the background is quite flat and reflects well the real data distribution.
Looking at the previous plots, it does not seem useful to apply any cut on

those variables without suppressing a large fraction of the signal.

On the other hand, further studies on the distribution of the number of good leptons and on the number of b-tagged jets turned to be quite useful for background suppression.
Indeed, by requiring exactly four leptons passing the full selection reduces the contribution of $t\bar{t}$+V, VVV and ZH and WH background events without a relevant impact on the signal, as it can be observed from Figure 5.5 (top).

In addition, events from $t\bar{t}$+V and from Higgs boson production in vector boson fusion and $t\bar{t}$H topologies can be reduced by requiring that the number of b-tagged jets is strictly less that two, as shown in Figure 5.5 (bottom).

The cuts on the last two observables contributed to determine the signal region for this analysis, as previously defined in section 5.3. The third requirement of the SR on the difference between the four lepton invariant mass and the Higgs nominal mass to be less than 10 GeV/c^2 is reasonably due to the fact that we have to reconstruct the product of the decay of the SM Higgs, whose mass value is now well measured and fixed at 125 GeV/c^2.

5.3 Cut-based Selection

In the signal models considered for this analysis, the Higgs boson decaying to four leptons is produced in association with DM particles that do not interact with the detector, thus generating a source of *real* E_T^{miss}.
Once the SM Higgs selection is accomplished, the remaining main background comes from processes like $t\bar{t}$+V, VVV and VH events, as shown in the distributions in Section 5.2.

The following selection cuts have been originally studied and optimized to define the SR (already introduced in the subsection 5.1.2) :

- $|m_{4\ell} - 125| < 10$ GeV;
- Number of good leptons: $N_\ell = 4$;
- Number of b-jets: $N_b \leq 1$.

The logic of this choice is clear once the distributions of the observables are built.
The cut-based selection is optimized by performing a scan over a range of cuts for the remaining variables and selecting the set of cuts that maximizes the sensitivity, measured directly by the 95% CL upper limits on the cross section.
The two variables with the greatest discriminating power between signal and background are the invariant mass of the four lepton object $m_{4\ell}$ and the missing transverse energy E_T^{miss}.

The strategy of this analysis does not include any cut on MET because the shape of the MET the distribution is used as input for the statistical analysis described in Chapter 6.
Furthermore, the E_T^{miss} distribution strongly depends on the particular model and on the parameters used.

5.4 Background Estimation

The background coming from the five main production modes of the SM Higgs, gluon fusion, vector boson fusion, associated production (WH, ZH, and $t\bar{t}H$) is estimated from MC by normalizing the yields according to the cross sections times the branching ratio computed by the LHC Higgs Cross Section Working Group [93].

The ZH production with four charged leptons from the same vertex and real E_T^{miss}, coming either from $Z \to \nu\nu$ and $H \to ZZ \to 4\ell$ or from $Z \to \ell\ell$ and $H \to ZZ \to \ell\ell\nu\nu$, is considered an irreducible background.
Events coming from VVV ($V = W, Z$) processes could also yield four charged lepton from the same vertex and real E_T^{miss} from neutrinos: this is another source of irreducible background and it is estimated from simulation.

5.4.1 Reducible Background

$qq \to ZZ$ Modelling

The $qq \to ZZ$ background is generated at NLO, while the fully differential cross section has been computed at NNLO [95], but are not yet available in a partonic level event generator. Therefore NNLO/NLO k-factors for the qq background process are applied to the POWHEG sample. The inclusive cross sections obtained using the same PDF and renormalization and factorization scales as the POWHEG sample at LO, NLO, and NNLO are used [95]. The NNLO/NLO k-factors are applied in the analysis differentially as a function of m_{ZZ}.

Additional NLO electroweak corrections which depend on the initial state quark flavor and kinematics are also applied to the $qq \to ZZ$ background process in the region $m_{ZZ} > 2m_Z$ where the corrections have been computed as described in Ref. [84].

$gg \to ZZ$ Modelling

Event simulation for the background is done at LO with the generator MCFM 7.0 [94], [47].

Although no exact calculation exists beyond the LO for the $gg \to ZZ$ background, it has been recently shown [96] that the soft collinear approximation is able to describe the background cross section and the interference term at NNLO.
Further calculations also show that the k-factors are very similar at NLO for the signal and background [97] and at NNLO for the signal and interference terms [98]. Therefore, the same k-factor is used for the signal and background [100].

Z+X background

The $Z + X$ reducible background originates from processes which contain one or more than one non-prompt leptons in a four-lepton final state, together with additional missing transverse energy. This contribution is estimated from data and is described in this section.
Non-prompt leptons may originate from decay of heavy-flavour mesons, mis-reconstructed jets, or photon-to-electron conversion. These leptons are also called *fake* leptons. The lepton "fake rate" is defined as the probability of a lepton passing the "loose" selection criteria in a Z+lepton control region to also pass the "tight" selection criteria. In this CR, the reconstructed Z is formed from two tight leptons and the single lepton component is expected to be dominated by fakes. The loose and tight lepton selection criteria have been already defined in Chapter 4.
The electron and muon fake rates, f_e and f_μ, has been estimated from $Z(\to \ell\ell)$+e and $Z(\to \ell\ell)+\mu$ events, where the leptons forming the Z are opposite sign, same flavor, with $p_T > 20(10)$ GeV.
The third lepton, either electron or muon, has to pass the loose selection criteria and is used as a probe to measure the *fake rate*.
In order suppress QCD contamination, the invariant mass of this loose lepton with the opposite-sign lepton from the Z has to satisfy the minimal condition $m_{\ell\ell} > 4$ GeV.
The dilepton invariant mass of the two tight leptons has to satisfy the condition $|M_{inv}(\ell_1, \ell_2) - M_Z| < 7$ GeV, to suppress the presence of photon conversions to electrons.
Contamination from WZ and $t\bar{t}$ is suppressed by imposing $E_T^{miss} < 25$ GeV.
Muon and electron fake rates, binned by p_T in the barrel and endcap regions are shown in Figure 5.6.

The strategy for applying the lepton fake rates discussed above in order to estimate the Z+X background from data in the signal region relies on the definition of two control regions.
Starting from the events which pass the step where a Z candidate is built, those with an additional pair of loose leptons of same flavour and opposite charge and passing the $|\mathrm{SIP}_{3D}| < 4$ cut are selected; those events should also satisfy all kinematic cuts applied for the SM Higgs selection (see Subsec. 5.1.2).

The **first Control Region** is defined by requiring that the two loose leptons, which do not form the Z_1 candidate, do not pass the final identification and isolation criteria, while the other two leptons should pass the final selection criteria because of the definition of the Z_1.
This control region defines the "2 Prompt + 2 Fail" ("2P+2F") sample.
This sample is expected to be populated with events that have only two prompt real leptons (mostly DY, with a small fraction of $t\bar{t}$ and $Z\gamma$ events).

The **second Control Region** is defined by requiring one of the four leptons not to pass the final identification and isolation criteria; instead the other three leptons should pass the final selection criteria. This control region defines the "3 Prompt + 1 Fail" ("3P+1F") sample.
This sample is expected to be populated with the type of events that populate also the 2P+2F region, even if with different relative proportions, as well as with WZ events that can provide three prompt real leptons.

The control samples built in this way are enriched with fake leptons and can be considered quite orthogonal, just by construction, to the SM Higgs signal region, as they should be. The samples obtained are used to estimate the reducible background in the SM Higgs signal region by using the *fake rate*.

Figures 5.7 and 5.8 show the 4-lepton invariant mass distributions of events selected in the 2P+2F and 3P+1F control samples respectively.
Discrepancies between data and simulation are related to the fact that the *fake rate* do not properly take into account the background composition of the 2P+2F.

The final expression for the prediction of the Z+X background in the SM Higgs signal region is:

$$N_{\mathrm{SR}}^{bkg} = \sum \frac{f_i}{(1-f_i)}(N_{\mathrm{3P1F}} - N_{\mathrm{3P1F}}^{\mathrm{bkg}} - N_{\mathrm{3P1F}}^{\mathrm{ZZ}}) + \sum \frac{f_i}{(1-f_i)}\frac{f_j}{(1-f_j)}N_{\mathrm{2P2F}} \quad (5.2)$$

where N_{3P1F} is the number of observed events in the 3P+1F control region, $N_{\mathrm{3P1F}}^{\mathrm{ZZ}} + N_{\mathrm{3P1F}}^{\mathrm{bkg}}$ is the expected contribution from the 3P+1F region and ZZ processes in the SM Higgs signal region.

Uncertainty on the Z+X background

The limited size of the samples in the control regions is the source of the statistical uncertainty of the method, which is typically in the range of 1-10%.

A source of systematic uncertainty of the methods presented above potentially arises from the different composition of background processes (DY, $t\bar{t}$, WZ, $Z\gamma$) in the region where we measure and apply the fake ratios; the emphfake rate' method corrects for the resulting bias via the "3P1F component"

of its prediction. Compositions of reducible background processes (DY, $t\bar{t}$, WZ, $Z\gamma^{(*)}$) in the region where the fake ratio method are measured and applied are typically not the same. This reopresents the main source of the systematic uncertainty of the fake ratio method. This uncertainty can be estimated by measuring the fake ratios for individual background processes in the $Z + \ell$ region in simulation. The weighted average of these individual fake ratios is the fake ratio that we measure in simulation. The exact composition of the background processes in the 2P+2F region where we plan to apply the fake ratios is determined from simulation, and the individual fake ratios can be reweighed according to the 2P+2F composition. The difference between the reweighed fake ratio and the average one can be used as an estimate of the uncertainty on the measurement of the fake ratios, as detailed in Ref. [84]. The effect of this systematic uncertainty is propagated to the final estimates, and it amounts to about 43% for the $4e$, 40% for the $2e2\mu$, and 36% for the 4μ final states, accordingly.

In order to estimate the uncertainty on the $m_{4\ell}$ shape we looked at the differences between the shapes of predicted background distributions for all three channels, and between both methods. The envelope of the differences between these distributions is used as an estimate of the shape uncertainty. The uncertainty is estimated to be in the range of 5%-15%.

The Z+X event yields in the signal region is obtained after applying the additional cuts for the signal region defined in Section 5.3 and is reported in Table 5.1.

Z+X	4e	4μ	$2e2\mu$	$2\mu2e$	Total 4ℓ
SM Higgs signal region	27.2 ± 11.7	31.0 ± 11.2	26.4 ± 10.6	35.6 ± 14.2	120.2 ± 29.6
Signal region (mono-Higgs)	3.2 ± 1.4	3.9 ± 1.4	5.9 ± 2.4	4.8 ± 1.9	17.8 ± 4.7

TABLE 5.1: Event yield contribution of the reducible Z+X background processes in the SM Higgs signal region (cut-based), and the mono-Higgs signal region predicted from measurements in data using the "fake rate" method. The total uncertainty on the measurements is quoted.

5.5 Systematic uncertainties

This section covers the treatment of systematic uncertainties associated with the Mono-Higgs analysis here presented. The variables in common with the $H \to ZZ \to 4\ell$ search are treated in the same way, including uncertainties on the background estimated from data, instead the uncertainties associated with the *real* and *fake* E_T^{miss} modelling have been computed and reported in the following subsections.

5.5.1 Experimental uncertainties

Experimental uncertainties affecting both signal and background are the uncertainty on the integrated luminosity (2.6%) and the uncertainty on the lepton identification and reconstruction efficiency (ranging from 2.5–9% on the overall event yield for the 4μ and $4e$ channels, respectively).
The uncertainty on the lepton energy scale is determined by considering the $Z \to \ell\ell$ mass distributions in data and in the simulation.
The uncertainty is determined to be 0.04% (0.3%) for the 4μ ($4e$) channel.
The uncertainty on the 4ℓ mass resolution coming from the uncertainty on the per-lepton energy resolution is 20% (see Ref. [84]).

Systematics on *Real* MET

Experimental uncertainties for the *fake rate* based background estimation, described in the previous section, vary between 36% (4μ) and 43% ($4e$), as described in the previous Section 5.4.1.

Furthermore, there are two types of systematic uncertainties related to the modelling of E_T^{miss}, those from the measurement of real E_T^{miss}, as from the samples with neutrinos, and those from fake E_T^{miss} due to the mis-measurement of jets and the other physics objects.

The uncertainties from the modelling of *real* E_T^{miss} are measured by varying the parameters associated to the corrections used to calculate E_T^{miss}, as suggested by the MET Particle Object Group and propagating those variations to the E_T^{miss} distribution after the full mono-Higgs analysis.
Each correction is varied up and down by one standard deviation of the input distribution. The corrections used in this calculation are:

- jet energy,
- jet resolution,
- muon energy,
- electron energy,
- photon energy,
- unclustered jet energy.

The E_T^{miss} distributions are derived for the ZZ background sample and a benchmark signal sample after passing the full mono-Higgs analysis once each correction is varied up and down; those distributions are reported in Figure5.9 for the ZZ background sample.

In order to derive the final systematic uncertainty, a fit of the ratio of the difference between the corrected E_T^{miss} and the E_T^{miss} has been made in the range

of E_T^{miss} of [0,110] GeV. Results are shown for the ZZ sample in Table 5.2. Finally, a systematic uncertainty of about 14% on the *real* E_T^{miss} has been measured in the range of [0,110] GeV summing in quadrature the deviations corresponding to each correction.

Correction	Variation Up	Variation Down
Jet energy	4.9%	2.3%
Jet resolution	3.5%	3.5%
Muon energy	3.6%	3.6%
Electron energy	3.3%	3.3%
Photon energy	3.7%	3.7%
Unclustered jet energy	6.5%	3.7%

TABLE 5.2: Systematic uncertainty on the E_T^{miss} obtained by the variation of the correction up and down, for ZZ sample after pass the full mono-Higgs selection.

Systematics on *Fake* MET

Another source of systematic uncertainty is the modeling of *fake* E_T^{miss}, and it is primarily due to the mis-measurement of jets; this uncertainty is measured from the CR, as defined in Section 5.4.1, as the percentage difference between the efficiency for the data and total background sample to pass the full mono-Higgs selection; a difference of around 40% is measured and it is taken as the systematic uncertainty on the *fake* E_T^{miss}. Additional uncertainties may come from the inaccurate knowledge of the jet energy scale and it has bee estimated by applying the official jet uncertainties on the jet energy scale (which depend on η and of the jet). The impact on this amounts up to 10% for the VBF sample. The uncertainties on the selection of b-jets is taken into account by looking at the b-jet efficiency and mis-identification rate (up to 6% for the $t\bar{t}H$ sample). Following the b-tagging POG recommendation to compute the data-to-MC scale factor for those quantities (and their uncertainties), the uncertainty on the full selection efficiency is about 1%.

A summary of the experimental uncertainties is given in Table 5.3.

5.5.2 Theoretical uncertainties

Theoretical uncertainties affect both signal and background estimation. They include uncertainties from the renormalization and factorization scale and they depend on the choice of PDF set. The uncertainty from the renormalization and factorization scale is determined by varying these scales between 0.5 and 2 times their nominal value while keeping their ratio between 0.5 and 2. The uncertainty from the PDF set is determined by taking the root mean square of the variation when using different replicas of the default NNPDF set. An additional uncertainty of the 10% on the k-factor used for the $ggZZ$

Summary of experimental systematic uncertainties	
Common experimental uncertainties	
Luminosity	2.6 %
Lepton identification/reconstruction efficiencies	2.5 – 9 %
Background related uncertainties	
Z+X	36 – 43 %
Signal related uncertainties	
Lepton energy scale	0.04 – 0.3 %
Lepton energy resolution	20 %
Real E_T^{miss} uncertainties	
E_T^{miss}	up to 14 %
Fake E_T^{miss} uncertainties	
E_T^{miss}	40 %
Jet Energy Scale uncertainties	
Selection efficiency	up to 10 %
b-tagging uncertainties	
Selection efficiency	1 %

TABLE 5.3: Summary of the experimental systematical uncertainties for this analysis.

prediction is applied as described in Section 5.4.1. A systematic uncertainty of 2% on the branching ratio of $H \to ZZ \to 4l$ only affects the signal yield.

A summary of the theoretical uncertainties is given in Table 5.3.

Summary of theoretical systematic uncertainties	
QCD scale (gg)	$\pm$ 3.9 %
PDF set (gg)	$\pm$ 3.2 %
Bkg K factor (gg)	$\pm$ 10 %
QCD scale (VBF)	+0.4/-0.3 %
PDF set (VBF)	$\pm$ 2.1 %
QCD scale (WH)	+0.5/-0.7 %
PDF set (WH)	$\pm$ 1.9 %
QCD scale (ZH)	+3.8/-3.1 %
PDF set (ZH)	$\pm$ 1.6 %
QCD scale ($t\bar{t}H$)	+5.8/-9.2 %
PDF set ($t\bar{t}H$)	$\pm$ 3.6 %
BR(H$\rightarrow 4\ell$)	2 %
QCD scale (qq)	+3.2/-4.2 % %
PDF set (qq)	+3.1/-3.4 %
Electroweak corrections (qq)	$\pm$ 0.1 %

TABLE 5.4: Summary of the theoretical systematical uncertainties for this analysis.

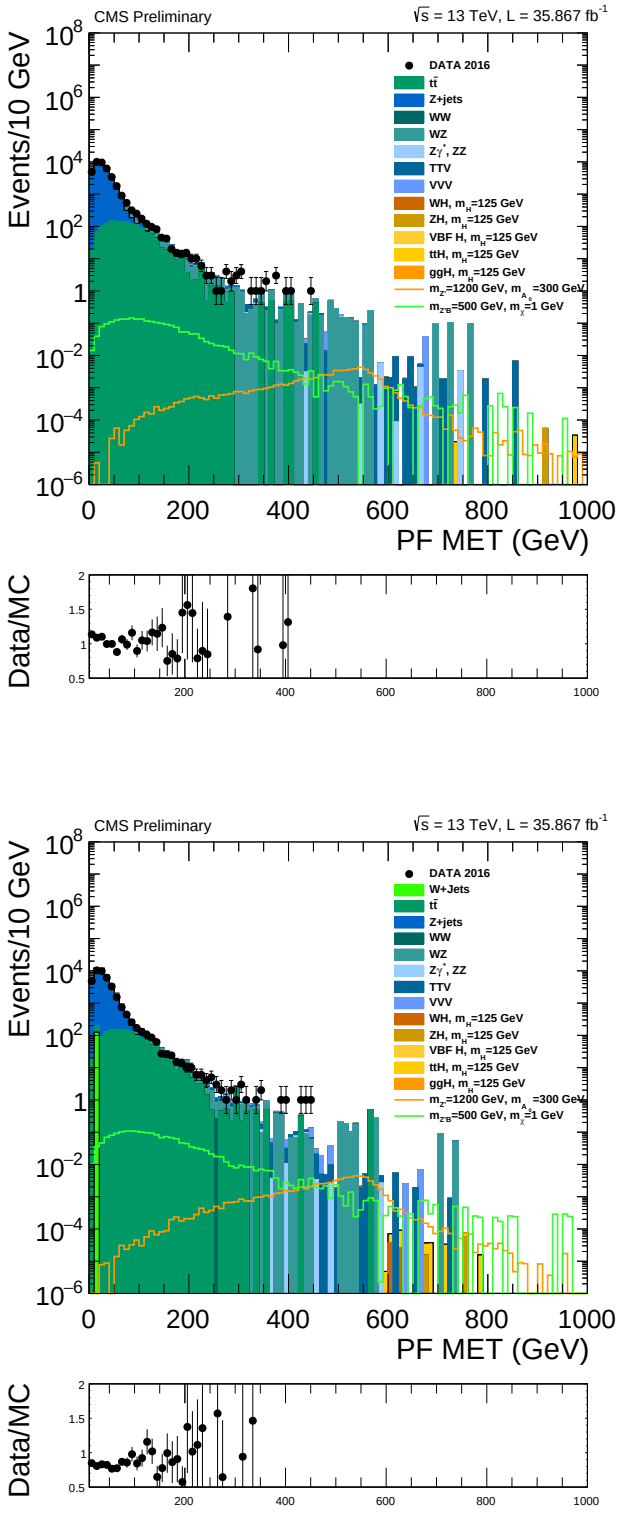


FIGURE 5.1: Event selection at step #1: Missing transverse energy distributions for a sample of inclusive $Z \to \mu\mu$ (top) and $Z \to ee$ (bottom) candidates. Markers correspond to the data collected in 2016, while the distributions from all SM background simulated processes are stacked. The distribution for the two signal samples of the $Z'2HDM$ and $Z'Baryonic$ models are superimposed. The ratio of the data and the sum of all the SM backgrounds is also shown at the bottom.

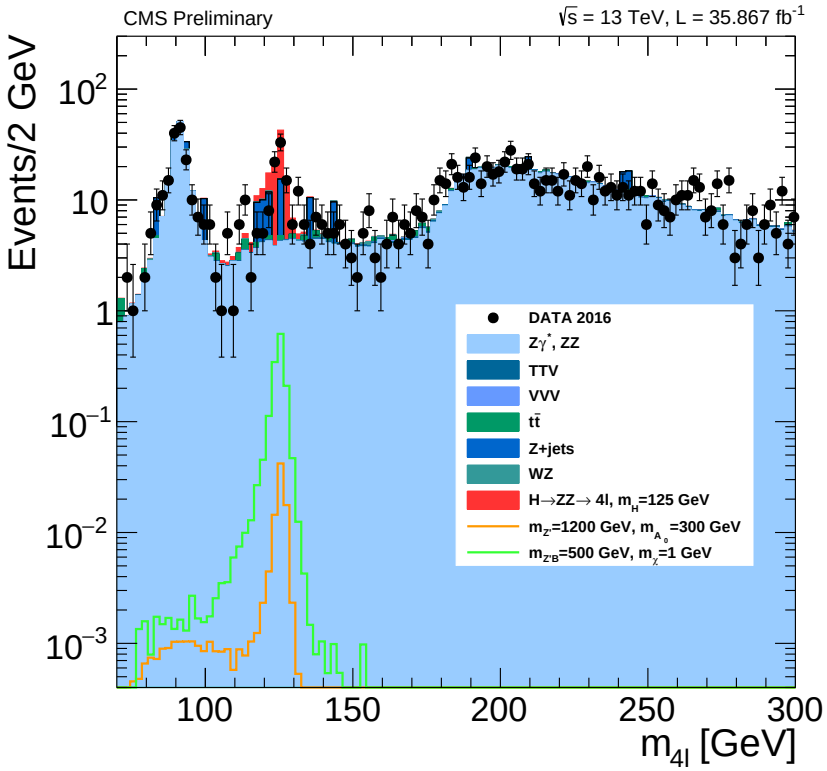


FIGURE 5.2: Event selection at step #3: Four-lepton invariant mass distributions from data and simulation, for both background and signal. Markers correspond to the data, while the distributions from all simulated SM background processes are stacked. The distribution for the two signal samples for the $Z'2HDM$ and $Z'Baryonic$ models are superimposed.

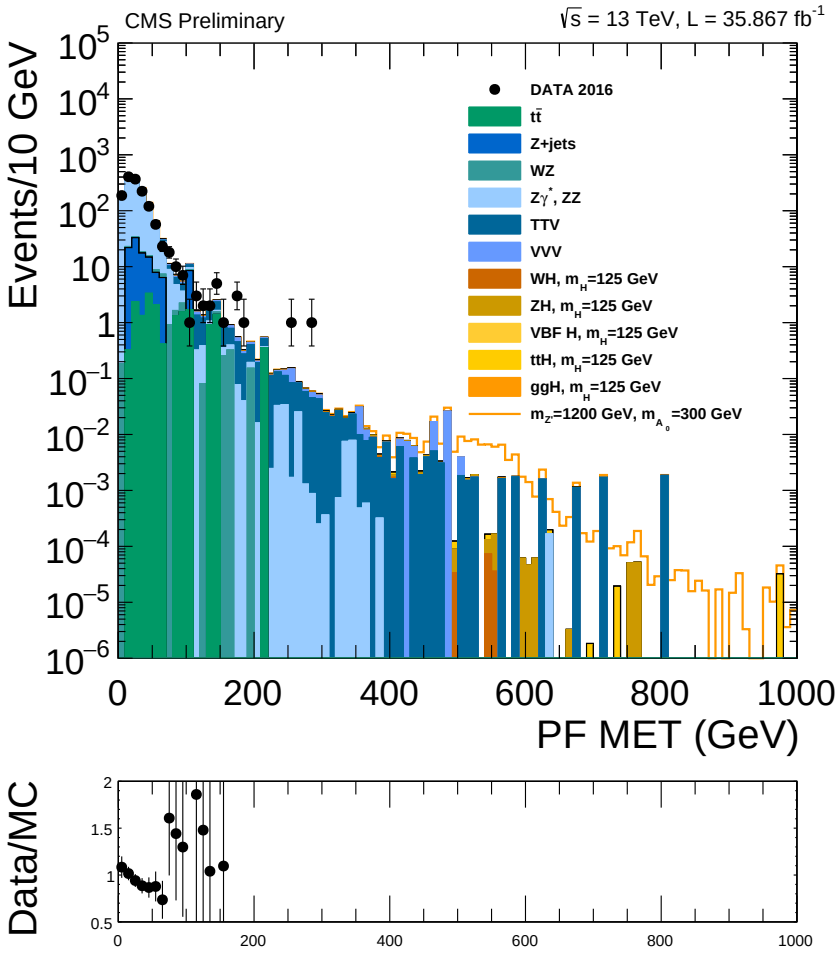


FIGURE 5.3: Event selection at step #3: Missing transverse energy distribution for data and all SM backgrounds, and for the DM signal. SM backgrounds are from simulation and are stacked, while the signal is superimposed (not stacked). The ratio of the data and the sum of all the SM backgrounds is also shown at the bottom.

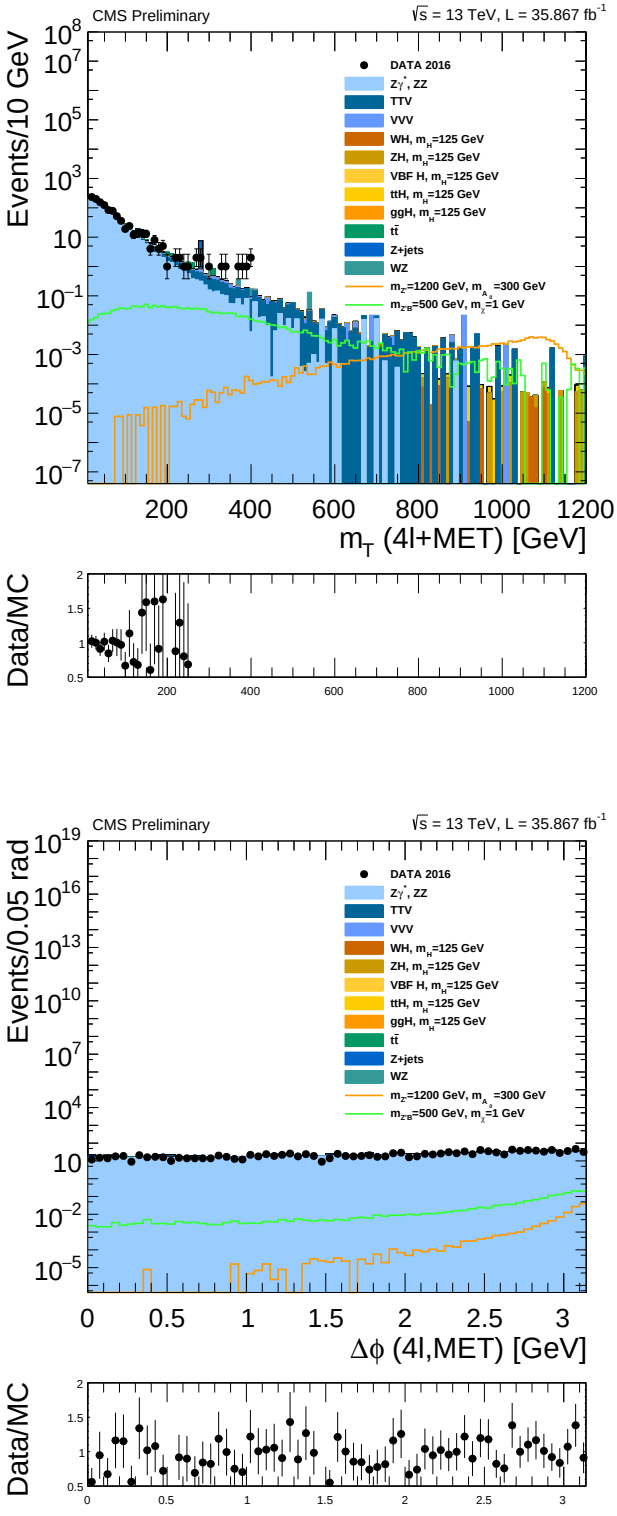


FIGURE 5.4: Transverse mass of the four-lepton system and E_T^{miss}, $m_T(4\ell + E_T^{miss})$ (top), and azimuthal angle difference between the four-lepton system and the E_T^{miss} ($\Delta\phi(4\ell - E_T^{miss})$) (bottom). Distributions are shown for data, and simulated background and signal events. Markers correspond to the data collected in 2016, while the distributions from SM background simulated processes are stacked. The distribution for the signal corresponding to two signal samples for the $Z'2HDM$ and $Z'Baryonic$ models are also superimposed.

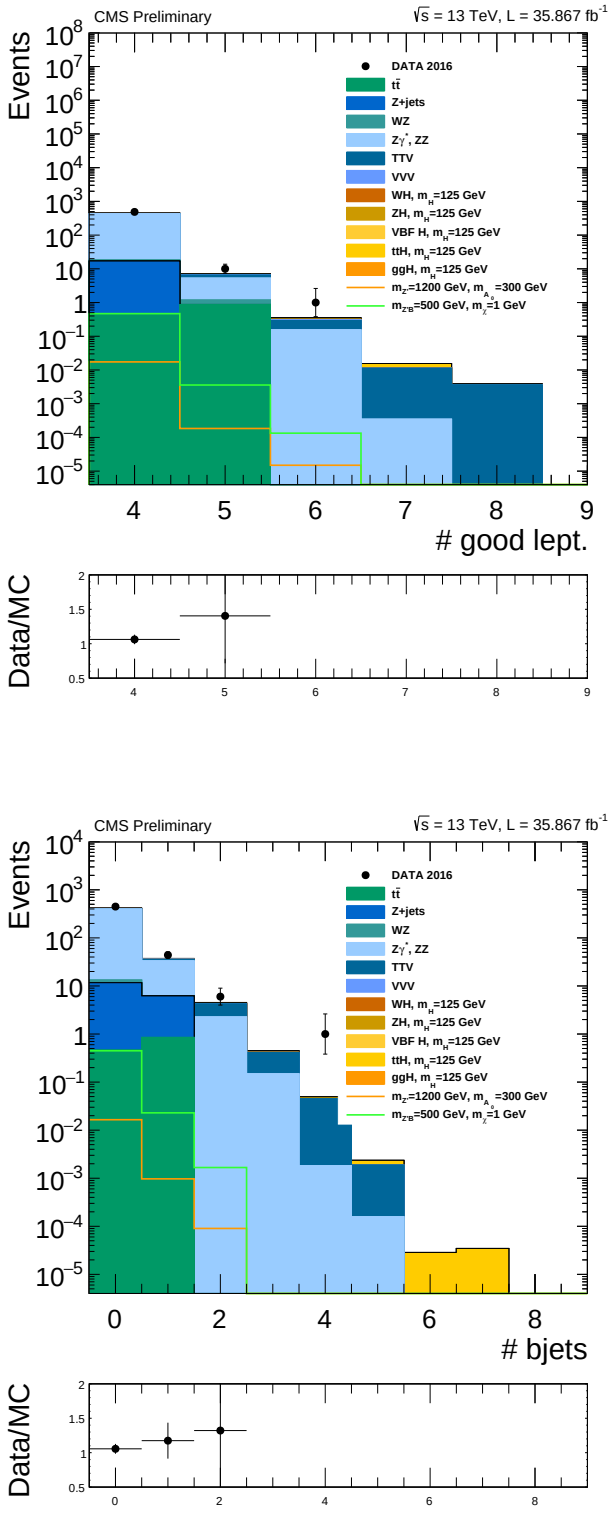


FIGURE 5.5: Number of selected good leptons (top) and b-tagged jets (right) for data, and simulated background and signal events. Markers correspond to the data collected in 2016, while the distributions from SM background simulated processes are stacked. The distribution for the signal corresponding to two signal samples for the $Z'2HDM$ and $Z'Baryonic$ models are superimposed.

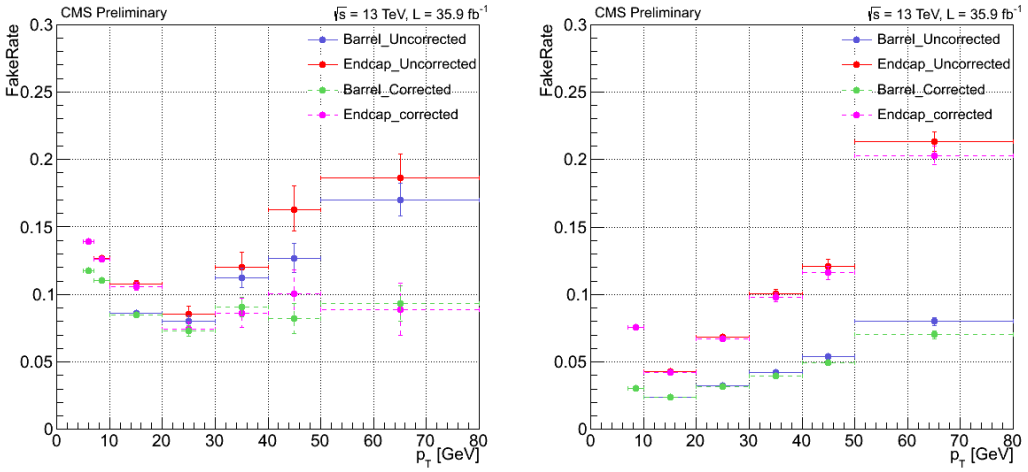


FIGURE 5.6: Muon (left) and electron (right) fake rates as a function of their transverse momenta.

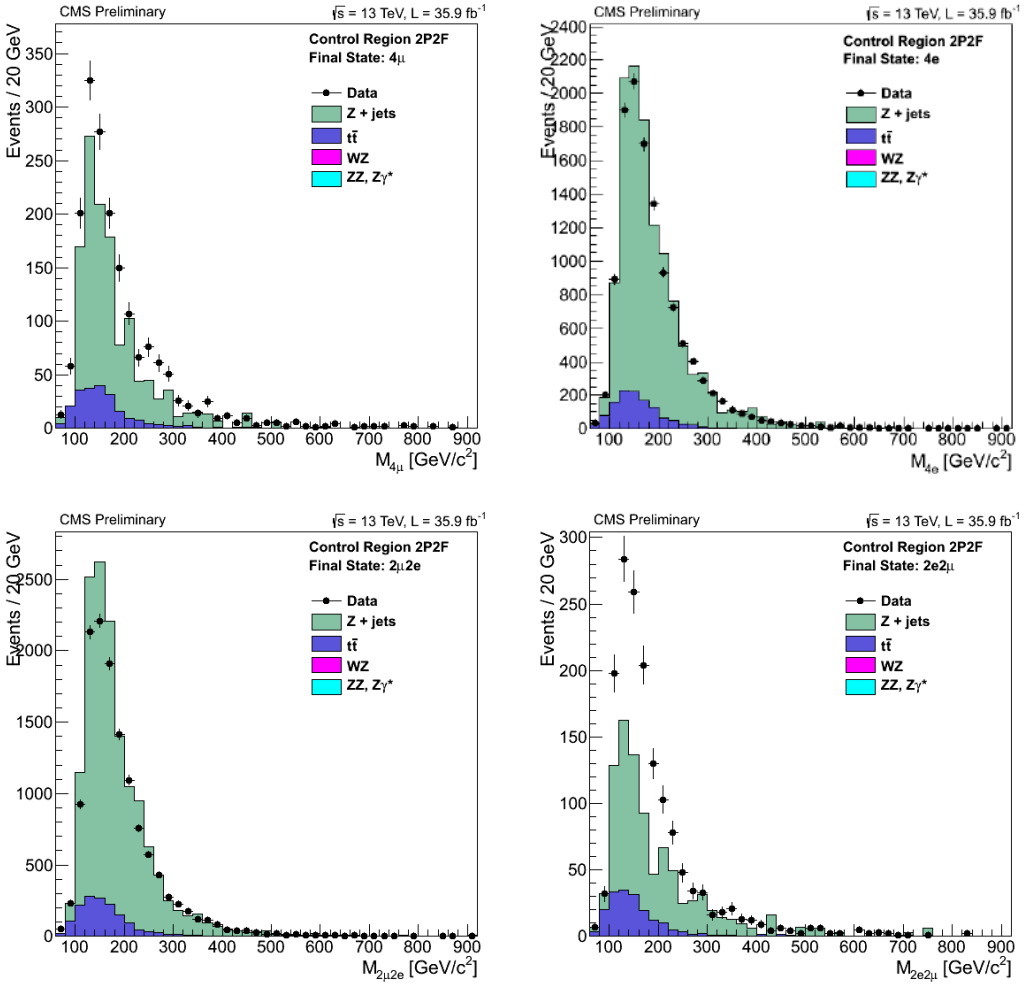


FIGURE 5.7: Invariant mass distribution of the events selected in the 2P+2F control sample in the 4μ (top left), 4e (top right), 2μ2e (bottom left), and 2e2μ (bottom right) final states.

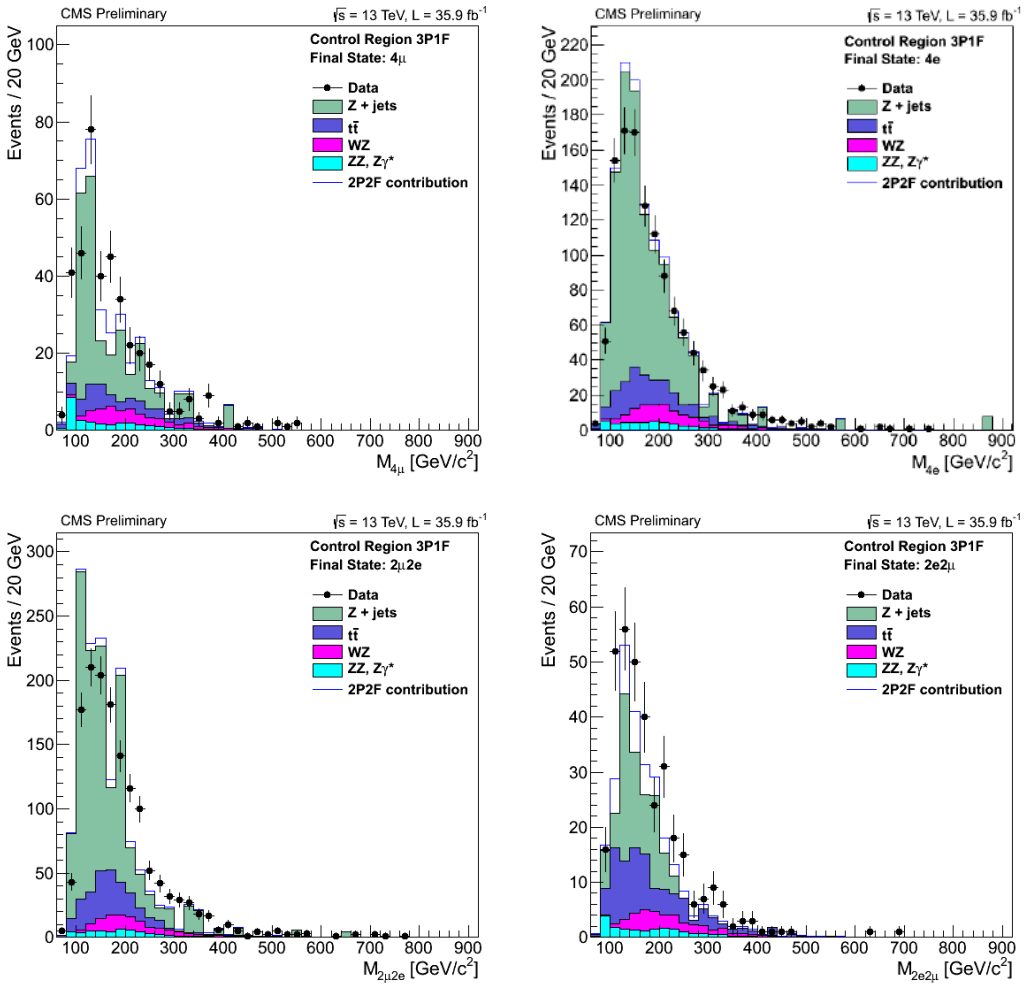


FIGURE 5.8: Invariant mass distribution of the events selected in the 3P+1F control sample in the 4μ (top left), 4e (top right), 2μ2e (bottom left), and 2e2μ (bottom right) final states.

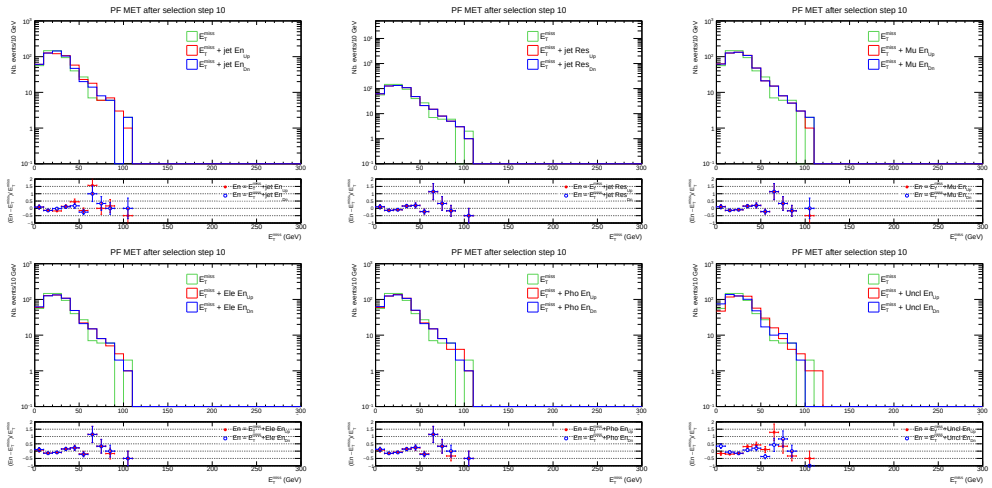


FIGURE 5.9: E_T^{miss} distributions for ZZ sample after passing the full Mono-Higgs selection and after varying the following corrections up and down (from left to right, from top to bottom): jet energy, jet resolution, muon energy, electron energy, photon energy, and unclustered jet energy.

Chapter 6

Results

In this last chapter the latest results of this Mono-Higgs analysis are reported. Section 6.1 reports the yields after the main selection steps for each final state of the Higgs decay and the final distributions of the four-lepton invariant mass and the missing transverse energy.
Finally Section 5.5 describes the statistical analysis and the results obtained in terms of 95% upper limits on production cross section for both of the two models used to simulate the Dark Matter signal.

6.1 Final yields and distributions

Table 6.1 shows the event yields after each selection cut for the three final states of the Higgs decay, 4μ 4e and 2e2μ for SM backgrounds, for a benchmark Z'*Baryonic* signal sample, and for data. The tables includes the Z'*Baryonic* signal events since it turned to be the most sensitive model for this analysis (see Section 6.2). The signal sample correspond to the Z'Baryonic model with $m_{Z'} = 500$ GeV and $m_{\chi} = 1$ GeV. As it can be observed, the number of events observed in real data are quite compatible with the total amount of background events within the errors (only the statistical uncertainty is quoted in the Table).

6.1.1 SM Higgs selection results

The four lepton invariant mass spectrum obtained after including the Z+X background estimated from data, at the step of the SM Higgs selection is shown in Figure. 6.1, for the three Higgs decay channel, while the full 4ℓ final state invariant mass distribution is shown in Figure 6.2. The Z+X distribution is here smoothed by using the Landau function. The data/MC comparison show a quite good agreement between data and simulation. The pink peak corresponds to the Higgs signal (m_H = 125 GeV/c^2)which is expected to be reconstructed in the four lepton invariant mass distribution.

Channel: 4μ	$Z\gamma^*, ZZ$	$Z+js$	$WZ, WW, W+js$	$t\bar{t}$	$t\bar{t}V$	VVV	QCD	$SM\,H$	$Tot.\,Bkg$	$Z'Baryonic$	Obs.
HLT	3.15e+04 ± 1.31e+01	9.21e+07 ± 2.33e+04	5.58e+08 ± 2.53e+05	2.18e+06 ± 9.44e+02	6.18e+03 ± 2.97e+00	2.12e+03 ± 6.06e+00	2.81e+08 ± 9.65e+06	4.05e+02 ± 4.10e-01	9.33e+08 ± 9.66e+06	8.29e+00 ± 3.07e-02	5.27e+08
l^+l^-, $12 < m_{l^+l^-} < 120$	1.71e+03 ± 3.47e+00	3.00e+04 ± 4.21e+02	2.20e+03 ± 1.33e+01	1.45e+03 ± 2.46e+01	5.80e+01 ± 2.37e-01	2.16e+01 ± 5.80e-01	0.00e+00 ± 0.00e+00	5.79e+01 ± 1.56e-01	3.55e+04 ± 4.22e+02	1.15e+00 ± 1.14e-02	4.14e+04
Z_1, Z_2	587.47 ± 1.99	475.00 ± 54.49	2.78 ± 0.51	8.87 ± 1.85	6.12 ± 0.10	2.50 ± 0.16	0.00 ± 0.00	28.94 ± 0.11	1111.70 ± 54.56	0.58 ± 0.01	1205
$p_T > 20/10$, $m_{l^+l^-} > 4$, $m_{Z_1} > 40$	447.27 ± 1.71	18.70 ± 10.80	1.81 ± 0.38	1.43 ± 0.71	5.84 ± 0.10	2.42 ± 0.15	0.00 ± 0.00	26.00 ± 0.10	503.47 ± 10.96	0.51 ± 0.01	505
$m_{4l} > 70$	416.05 ± 1.59	16.57 ± 9.57	1.67 ± 0.35	1.32 ± 0.66	5.49 ± 0.09	2.28 ± 0.15	0.00 ± 0.00	23.96 ± 0.10	467.33 ± 9.73	0.47 ± 0.01	503
$\lvert m_{4l} - 125\rvert \leq 10$, $N_l = 4$, $n_b \leq 1$	14.87 ± 0.31	0.00 ± 0.00	0.25 ± 0.14	0.00 ± 0.00	0.06 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	21.75 ± 0.09	36.93 ± 0.35	0.45 ± 0.01	44

Channel: $4e$	$Z\gamma^*, ZZ$	$Z+js$	$WZ, WW, W+js$			VVV	QCD	$SM\,H$	$Tot.\,Bkg$	$Z'Baryonic$	Obs.
HLT	3.15e+04 ± 1.31e+01	9.21e+07 ± 2.33e+04	5.58e+08 ± 2.53e+05	2.18e+06 ± 9.44e+02	6.18e+03 ± 2.97e+00	2.12e+03 ± 6.06e+00	2.81e+08 ± 9.65e+06	4.05e+02 ± 4.10e-01	9.33e+08 ± 9.66e+06	8.29e+00 ± 3.07e-02	6.61e+08
$\ell^+\ell^-$, $12 < m_{\ell^+\ell^-} < 120$	1.12e+03 ± 2.80e+00	4.16e+04 ± 5.02e+02	1.70e+03 ± 1.25e+02	1.48e+03 ± 2.47e+01	4.63e+01 ± 2.16e-01	1.70e+01 ± 5.15e-01	0.00e+00 ± 0.00e+00	4.44e+01 ± 1.36e-01	4.60e+04 ± 5.18e+02	8.98e-01 ± 1.01e-02	4.31e+04
Z_1, Z_2	322.86 ± 1.46	166.17 ± 30.34	3.28 ± 0.49	9.23 ± 1.92	5.10 ± 0.09	2.16 ± 0.16	0.00 ± 0.00	16.63 ± 0.08	525.42 ± 30.44	0.35 ± 0.01	438
$> 20/10$, $m_{\ell^+\ell^-} > 4$, $m_{Z_1} > 40$	278.03 ± 1.34	48.95 ± 16.32	2.85 ± 0.45	5.07 ± 1.46	4.93 ± 0.09	2.16 ± 0.16	0.00 ± 0.00	15.15 ± 0.08	357.15 ± 16.45	0.31 ± 0.01	288
$m_{4\ell} > 70$	261.09 ± 1.26	46.77 ± 15.59	2.73 ± 0.43	4.73 ± 1.36	4.64 ± 0.09	2.02 ± 0.15	0.00 ± 0.00	14.42 ± 0.08	336.39 ± 15.71	0.30 ± 0.01	288
$\lvert m_{4\ell} - 125\rvert \leq 10$, $N_\ell = 4$, $n_b \leq 1$	6.74 ± 0.21	12.47 ± 8.82	0.06 ± 0.06	0.00 ± 0.00	0.08 ± 0.01	0.04 ± 0.03	0.00 ± 0.00	12.87 ± 0.07	32.27 ± 8.82	0.28 ± 0.01	23

Channel: $2e2\mu$	$Z\gamma^*, ZZ$	$Z+js$	$WZ, WW, W+js$	$t\bar{t}$	$t\bar{t}V$	VVV	QCD	$SM\,H$	$Tot.\,Bkg$	$Z'Baryonic$	Obs.
HLT	3.15e+04 ± 1.31e+01	9.21e+07 ± 2.33e+04	5.58e+08 ± 2.53e+05	2.18e+06 ± 9.44e+02	6.18e+03 ± 2.97e+00	2.12e+03 ± 6.06e+00	2.81e+08 ± 9.65e+06	4.05e+02 ± 4.10e-01	9.33e+08 ± 9.66e+06	8.29e+00 ± 3.07e-02	1.28e+09
l^+l^-, $12 < m_{l^+l^-} < 120$	3.49e+03 ± 4.91e+00	7.16e+04 ± 6.55e+02	3.90e+03 ± 1.26e+02	2.96e+03 ± 3.50e+01	1.09e+02 ± 3.27e-01	4.21e+01 ± 7.97e-01	0.00e+00 ± 0.00e+00	1.38e+02 ± 2.41e-01	8.22e+04 ± 6.68e+02	2.85e+00 ± 1.80e-02	8.54e+04
Z_1, Z_2	657.98 ± 2.06	43.60 ± 16.48	6.00 ± 0.71	26.29 ± 3.19	9.74 ± 0.13	4.31 ± 0.22	0.00 ± 0.00	36.31 ± 0.12	784.23 ± 16.93	0.80 ± 0.01	723
$p_T > 20/10$, $m_{l^+l^-} > 4$, $m_{Z_1} > 40$	630.67 ± 2.01	37.00 ± 15.10	5.43 ± 0.67	11.84 ± 2.09	9.44 ± 0.13	4.31 ± 0.22	0.00 ± 0.00	35.25 ± 0.12	733.95 ± 15.40	0.79 ± 0.01	672
$m_{4l} > 70$	590.65 ± 1.88	32.69 ± 13.35	5.12 ± 0.63	11.09 ± 1.96	8.86 ± 0.12	4.06 ± 0.21	0.00 ± 0.00	32.96 ± 0.11	685.42 ± 13.64	0.73 ± 0.01	671
$\lvert m_{4l} - 125\rvert \leq 10$, $N_l = 4$, $n_b \leq 1$	18.17 ± 0.35	15.24 ± 8.80	0.48 ± 0.18	0.92 ± 0.65	0.10 ± 0.01	0.03 ± 0.03	0.00 ± 0.00	29.12 ± 0.11	64.06 ± 8.83	0.68 ± 0.01	45

TABLE 6.1: Cut flow table: number of events passing the full selection for the 4μ (top), $4e$ (middle), and $2e2\mu$ (bottom) final states, as obtained from simulation for background and signal and from real data with a luminosity of 35.9 fb^{-1}. The signal sample correspond to the Z'Baryonic model with $m_{Z'} = 500$ GeV and $m_\chi = 1$ GeV. Only statistical errors are quoted for the simulated signal and background samples.

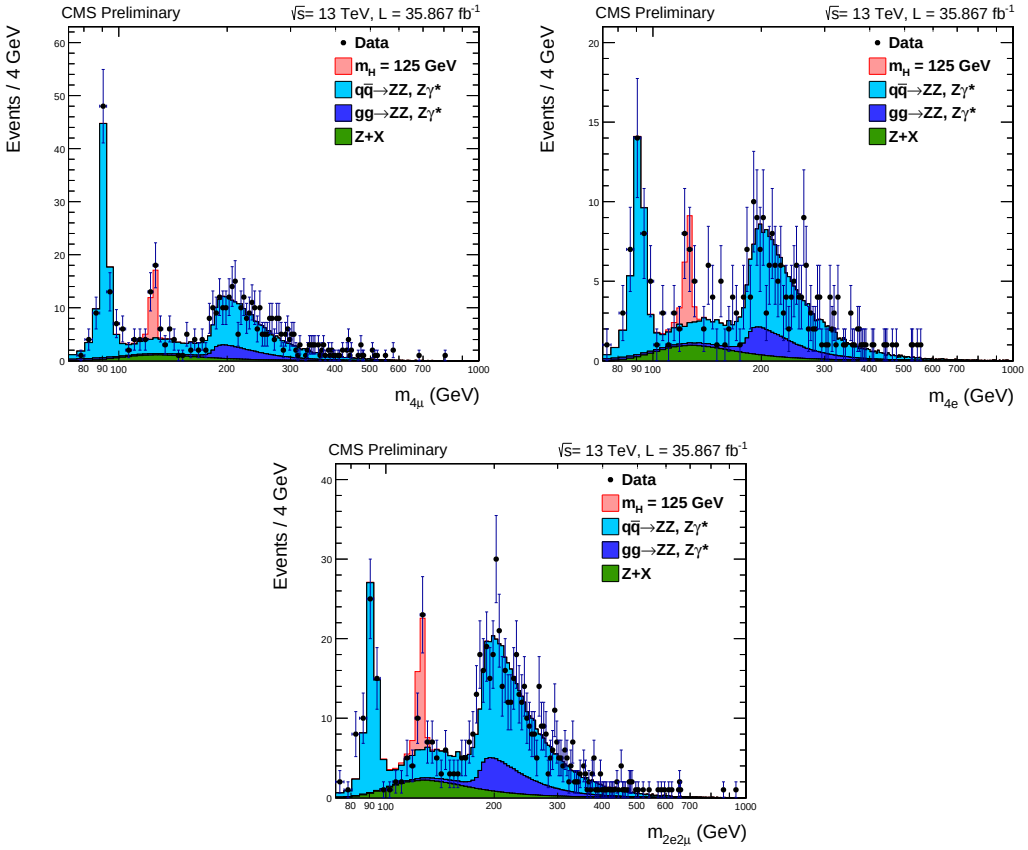


FIGURE 6.1: The observed four-lepton invariant mass spectrum after the SM Higgs selection, overlaid on the background prediction from simulation.

The counting of the total events remaining after the Standard Model Higgs signal selection for the 4e, 4μ, and 2e2μ final states is reported in Table 6.2. Only statistical errors are quoted for the simulated signal and background samples, while the "Z+X" contribution includes the systematic uncertainties estimated and presented in Chapter 5, Subsection 5.4.1.

6.1.2 Mono-Higgs selection results

The final distribution of the E_T^{miss} after applying the full Mono-Higgs event selection (sse Chapter 5, Section 5.3) is shown in Fig. 6.3.
The markers correspond to real data, while the background distributions from simulated processes are stacked; simulations of DM signal with different mass points are also stacked.
The ratio of the data and the sum of all the backgrounds is shown at the bottom.

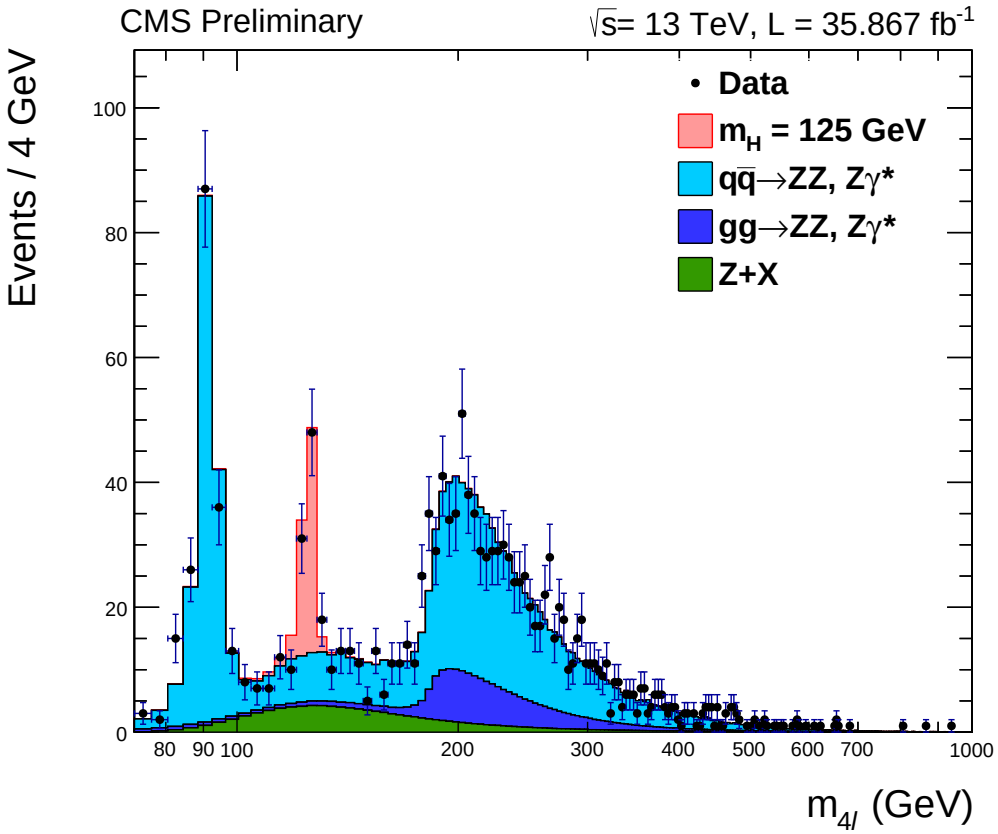


FIGURE 6.2: The observed four-lepton invariant mass spectrum after the SM Higgs selection, overlaid on the background prediction from simulation.

Table 6.3 shows the number of events remaining after the full Mono-Higgs event selection, as expected from signal and from background, estimated both from simulation and from data (for Z+X processes), and those observed in data. Only statistical errors are quoted in the Table except for "Z+X" which includes the systematic uncertainty.
No evidence of any excess on the Standard Model background can be highlighted from the numbers in 4μ and $4e$, and in the $2e2\mu$ Higgs decay channel. Therefore 95% upper limits are derived and presented in Section 6.2. Moreover, a deficit of events in the $2e2\mu$ channel can be observed in data with respect to those estimated from SM background simulation.

Figure 6.4 shows the event display corresponding to the event observed in real data after the full Mono-Higgs selection with the larger E_T^{miss} value (a 4μ event with E_T^{miss} 120.5 GeV).
In addition to the four muons and E_T^{miss} there are also three jets passing the loose jet identification criteria.

Channel	$4e$	4μ	$2e2\mu$	$4l$
$q\bar{q} \to ZZ$	217.76 ± 1.26	352.32 ± 1.58	490.84 ± 1.88	1060.92 ± 2.76
$gg \to ZZ$	43.33 ± 0.07	63.73 ± 0.09	99.81 ± 0.13	206.88 ± 0.18
$Z + X$	27.20 ± 11.70	31.00 ± 11.16	62.00 ± 24.80	120.20 ± 29.60
Sum of bkgs	288.29 ± 11.76	447.05 ± 11.27	652.65 ± 24.87	1387.99 ± 29.73
SM Higgs ($m_H = 125$ GeV)	14.42 ± 0.08	23.96 ± 0.10	32.96 ± 0.11	71.34 ± 0.17
Total expected	302.72 ± 11.76	471.01 ± 11.27	685.61 ± 24.87	1459.33 ± 29.73
Observed	288	503	671	1462

TABLE 6.2: Number of estimated background and signal events and number of observed candidates, after the SM Higgs selection, in the mass range $m_{4\ell} > 70$ GeV. SM Higgs signal and ZZ background are estimated from simulation, while Z+X is estimated from data; only statistical errors are quoted for the simulated signal and background samples, while the "Z+X" contribution includes the systematic uncertainties estimated.

Channel	$4e$	4μ	$2e2\mu$	4ℓ
$Z\gamma*, ZZ$	6.74 ± 0.21	14.87 ± 0.31	18.17 ± 0.35	39.78 ± 0.51
$t\bar{t}V$	0.08 ± 0.01	0.06 ± 0.01	0.10 ± 0.01	0.25 ± 0.02
VVV	0.04 ± 0.03	0.00 ± 0.00	0.03 ± 0.03	0.07 ± 0.04
$Z + X$	3.20 ± 1.38	3.90 ± 1.40	10.70 ± 4.28	17.80 ± 4.71
SM Higgs ($m_H = 125$ GeV))	12.87 ± 0.07	21.75 ± 0.09	29.12 ± 0.11	63.74 ± 0.16
Sum of bkgs	22.94 ± 1.39	40.58 ± 1.44	58.12 ± 4.30	121.64 ± 4.74
Signal: Z'Baryonic ($m_{Z'} = 500$ GeV, $m_\chi = 1$ GeV)	0.28 ± 0.01	0.45 ± 0.01	0.68 ± 0.01	1.41 ± 0.01
Observed	23	44	45	112

TABLE 6.3: Number of expected background and signal events and number of observed candidates, after applying the full Mono-Higgs selection; only statistical errors are quoted for the simulated signal and background samples except for "Z+X" which includes the systematic uncertainty.

6.2 The Statistical Analysis

The essential ingredients to perform a reliable statistical analysis are the event yields for the both the signal and background, the number of events observed in real data, and, for this particular analysis, an accurate study of the the distribution of the E_T^{miss} after the full event selection, together with a good description of the systematic uncertainties (see Chapter 5, Section 5.5).
All this information has been used to set the 95% upper limits on the cross section for the two signal models considered in this analysis; these limits can eitherallow to exclude a m'_Z window corresponding to the benchmark model used or, at least, to estimate the sensitivity of the analysis.
Before presenting the limit results for this Mono-Higgs analysis in Subsection 6.2.2, a quick view of the theory behind them will be provided in the next subsection.

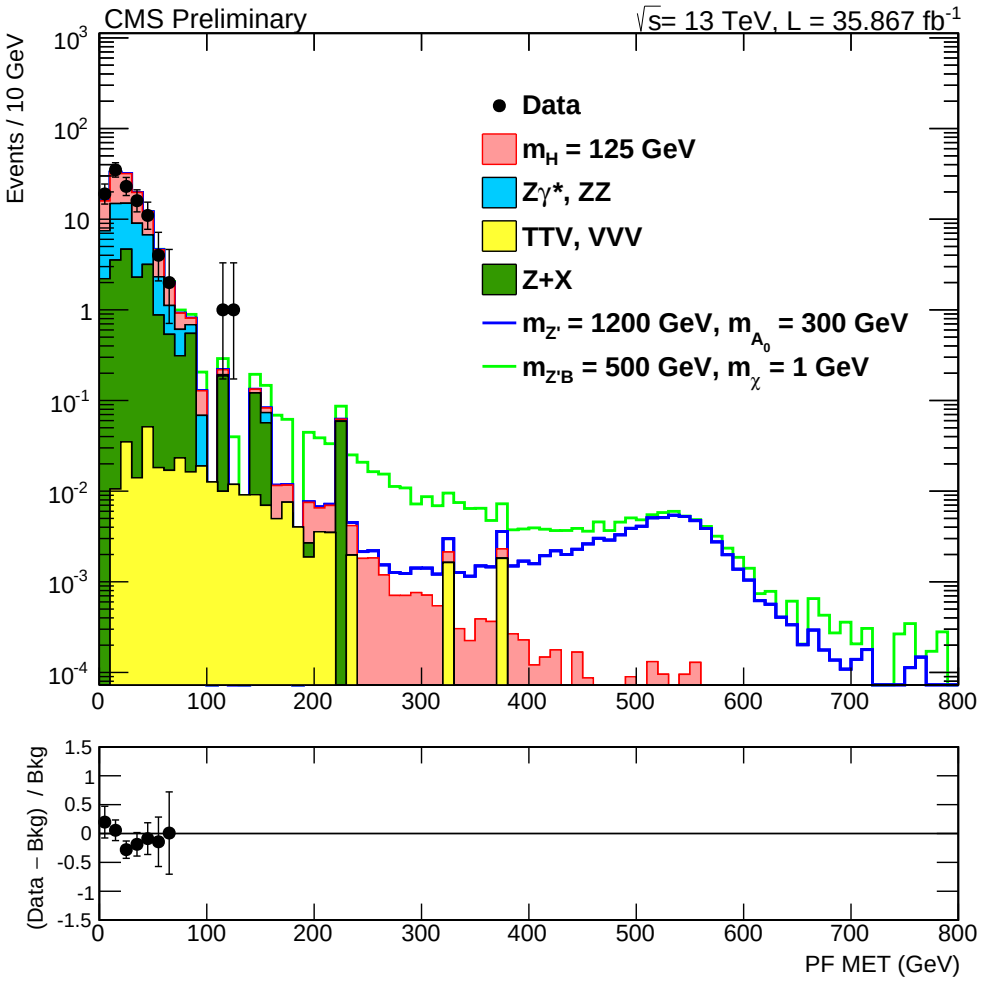


FIGURE 6.3: Distributions of the E_T^{miss} after applying the full event selection. The markers correspond to data, while the background distributions from simulated processes are stacked; simulated DM signals with different mass points are also stacked. The ratios of the data and the sum of all the SM backgrounds are shown at the bottom.

6.2.1 The Statistical Method for the interpretation of the results: the theory behind the limits

The statistical approach for the interpretation of results based on the upper limits set on the production cross section (or, equivalently, on the signal strenght modifier[1]) aims to characterize a non-observation of a particular signally or, otherwise, to establish a significant excess of events over the simulated background.
Both the Bayesian and the classical frequentist approach are two statistical

[1] the signal strength modifier is defined as the ratio between the cross section - computed at a certain Confidence Level - as expected from Monte Carlo and the one predicted by the model considered

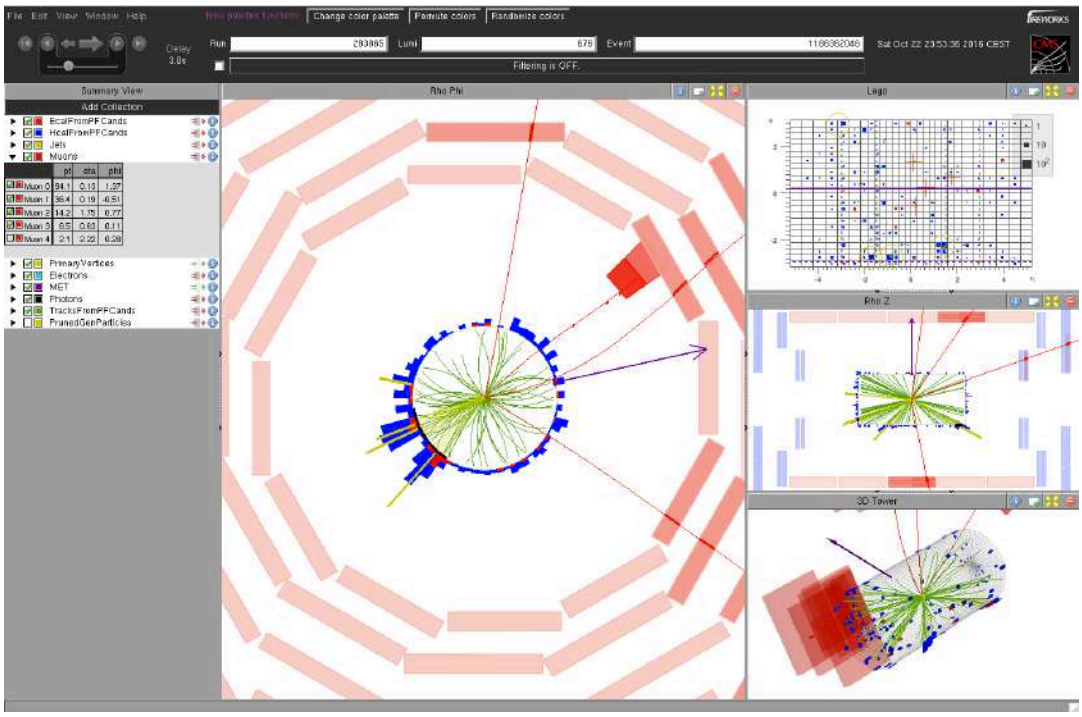


FIGURE 6.4: Event display of the event with the largest E_T^{miss} passing the full Mono-Higgs selection; four muons are visible.

method commonly used in high energy physics for characterizing the absence of a signal. Both these two approaches allow to quantify the level of incompatibility of data with a signal hypothesis, which is expressed as a Confidence Level (CL). Generally, for a pure convention, it is common to require a 95% C.L. for excluding a signal.
The classic frequentist approach here reported is formulated for the case of no systematic uncertainties and starts with the definition of a test statistic q_{mu} which is built to discriminate signal-like from background-like events. It can be said that q_{mu} compresses all signal-vs-background discriminating information into just one number.
By the Neyman-Pearson lemma, it can be shown that the ratio of likelihood functions Q is the most powerful discriminator. For many different reasons, the actual quantity used is a logarithm of the ratio, or more accurately, $-2lnQ$:

$$q_\mu = -2ln\frac{\mathcal{L}(data|\mu s + b)}{\mathcal{L}(data|b)} \tag{6.1}$$

where s stands for *signal*, b stands for *background*,μ is the signal strenght and $\mathcal{L}(data|rate)$ is a product of Poisson probabilities for number of either observed or simulated events, given the expected signal and background rates. After defining the test statistic, the corresponding probability distribution

functions (*pdfs*) must be built under *signal+background* hypothesis.This is accomplished by tossing toy pseudo-observations according to the same Poisson probabilities. Using these *pdfs*, the probability $P(q_\mu \geq q^data_\mu | \mu s + b)$ for the observed value q^data_μ to be as or less compatible with the *background+signal hypothesis*. Such probability is denoted as CL_{s+b}.
In the classical frequentist approach, it can be said that the signal is excluded at 95% C.L., if CL_{s+b} = 0.05.
This classical frequentist approach was then modified by Feldman and Cousins [102], who introduced a method to build one/two-sided confidence intervals based on the likelihood-ratio test statistic:

$$q_\mu = -2ln\frac{\mathcal{L}(data|\mu s + b)}{\mathcal{L}(data|\hat{\mu} s + b)},$$

where $\hat{\mu}$ is chosen so that it maximizes the likelihood function $\mathcal{L}(data|\mu s + b)$.
This construction is intended to protect the limits on signal stregth from background fluctuations and, above all, to avoid the under-coverage due to having to make decisions between reporting one-sided upper limits (no excess) and two-sided intervals when a significant excess of events is observed.
In particular, the one-sided test statistics which does not allow the signal to become negative is the profile likelihood test statistic that can be defined as follows:

$$\hat{q}_\mu = -2ln\frac{\mathcal{L}(data|\mu, \hat{\theta}_\mu)}{\mathcal{L}(data|\hat{\mu}, \hat{\theta})},$$

where $\hat{\theta}_\mu$ and $\hat{\theta}$ are the maximum likelihood estimators for the signal+background hypothesis (with the signal strength factor μ) and for the background-only hypothesis (μ = 0) respectively.
Removing the $\hat{\mu} > 0$ from the above equation the test statistic obtained is expected, following the Wilks theorem on the asymptotic regime, to have half χ^2 distribution for one degree of freedom (under signal+background experiments). The asymptotic regime is the on used to obtained the upper limits for this analysis.
For more details on the modified frequentist approach, also referred to as CL_s Method, and the asyptotic approximation see Ref. [101].
Systematic uncertainties on signal and background rates, s and b, not included in the above method can be introduced by modifying the test statistic itself and/or the pseudo-data generation and are represented by individual nuisance parameters with log-normal distributions.
In the Mono-Higgs analysis here presented, the exclusion limits are set on the cross section times the BR($H \to ZZ \to 4\ell$) for the various benchmark signal models, by using a procedure based on the modified frequentist approach previously discussed.
Statistical uncertainties from MC simulated events are taken into account for the limit setting.

The uncertainties affect the overall normalization of the signal and backgrounds as well as the shape of the distribution of the observables used as input for the statistical tool.
The 95% confidence level (CL) expected and observed limits on the signal strength parameter μ (which is, in this case, the the ratio of the 95% CL expected cross section to the theoretical one from the model), as well as one and two sigma deviations from the expected limits, have been computed for both the Z'2HDM and the Z'Baryonic models.
The points where the upper limit on the cross section is lower than the cross section from the theory are interpreted as excluded regions. Limits are set individually for the three decay channels, $4e$, 4μ, and $2e2\mu$, and then they are also combined together in the final 4ℓ limit.
In the context of the the Z'2HDM model the value of the m_{A0} parameter is fixed to 300 GeV since that corresponds to the highest cross section and the largest branching fraction to DM production.
In the case of the Z'Baryonic model the value of the parameter m_χ is fixed to 1 GeV since that corresponds to the highest cross section.

6.2.2 Results from the cut-based selection

During my first sensitivity studies using 2015 data corresponding to a luminosity of 2.8 fb^{-1} at a center of mass energy of 13 TeV, preliminary 95% upper limits have been derived using $Z'2HDM$ signal models and background official samples for the Monte Carlo campaign known as *Fall15*. No official Monte Carlo signal sample for this model in the four lepton Higgs' decay channel was available, so, in order to produce the related statistical analysis, I produced privately the Monte Carlo simulation for the signal considered using MADGRAPH5_AMCATNLO and, on top of it, PYTHIA8.209 in order to add the parton-showering (see chapter 3, section 3.2).
As in 2016, also in 2015 analysis the 95% upper limits on the cross section have been obtained performing a mono dimensional shape analysis on the distribution of the missing transverse energy after the full Mono-Higgs analysishas been applied.
The mass of the pseudoscalar mediator chosen for the generation was $A_0 = 300/c^2$ corresponding to the highest value of the production cross section for the $Z'2HDM$.
Figure 6.5 shows the preliminary 95% upper limits obtained using 2015 data collected by CMS.

As it can be observed, the expected and the theoretical curves are quite far so that they do not cross at all for all the mass range of the Z' parameter considered for the model. Therefore, no point of the parameter space can be excluded.
The latest results (using 2016 data and all MC Moriond17 official samples for

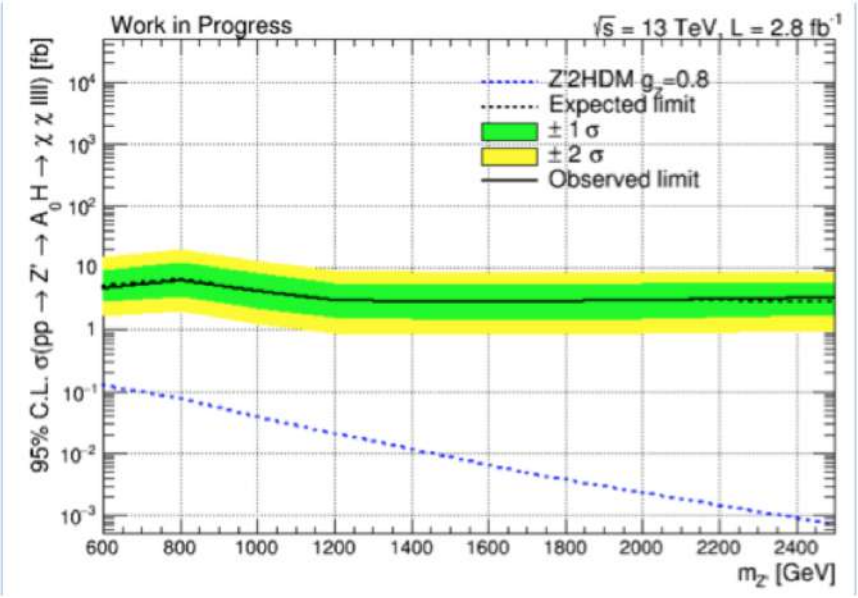


FIGURE 6.5: 95% upper limits on the cross section for the combination of the three channels, 4μ, $4e$ and $2e2\mu$ obtained performing a shape analysis on E_T^{miss} distribution using the $Z'2HDM$ model as a benchmark.

signals and backgrounds) in terms of number of events after the full Mono-Higgs analysis (as shown in Table 6.3) and the E_T^{miss} distribution have been used to perform a one dimensional (1D) shape analysis on the missing transverse energy itself.
The 1D 95% C.L. limits on $\sigma \times BR$ for the Z'2HDM model are shown in Figure 6.6 (top) For each of the three channels, 4μ, $4e$ and $2e2\mu$, and in Figure 6.6 (bottom) for the combination of all of them.
As it can be observed, the sensitivity of the Mono-Higgs, $H \rightarrow ZZ \rightarrow 4\ell$, search to the predictions of the Z'2HDM model is limited and no point of the parameter space of the model can be excluded at 95% CL.

The 1D 95% CL limits on $\sigma \times BR$ for the $Z'Baryonic$ model are reported in Figure 6.7 (top) for each of the three channel, 4μ, $4e$, and $2e2\mu$, and in Figure 6.7 (bottom) for the combination of all of them.
The Mono-Higgs, $H \rightarrow ZZ \rightarrow 4\ell$ analysis is almost sensible to exclude at 95% CL points of the parameter space of the Z'Baryonic model once the three channels are combined.

It can be observed that for both the two models the observed limit (dotted black line) overlaps the expected limit (solid black line) ensuring that the Monte Carlo studies prior to the data analysis was well run.
In particular, comparing Figures 6.5 and 6.6 (bottom), a quite relevant improvement in the sensitivity of the Mono-Higgs analysis here reported can be observed, confirming also the improvement of the cut-based analysis selection and of the signal/control regions definition.

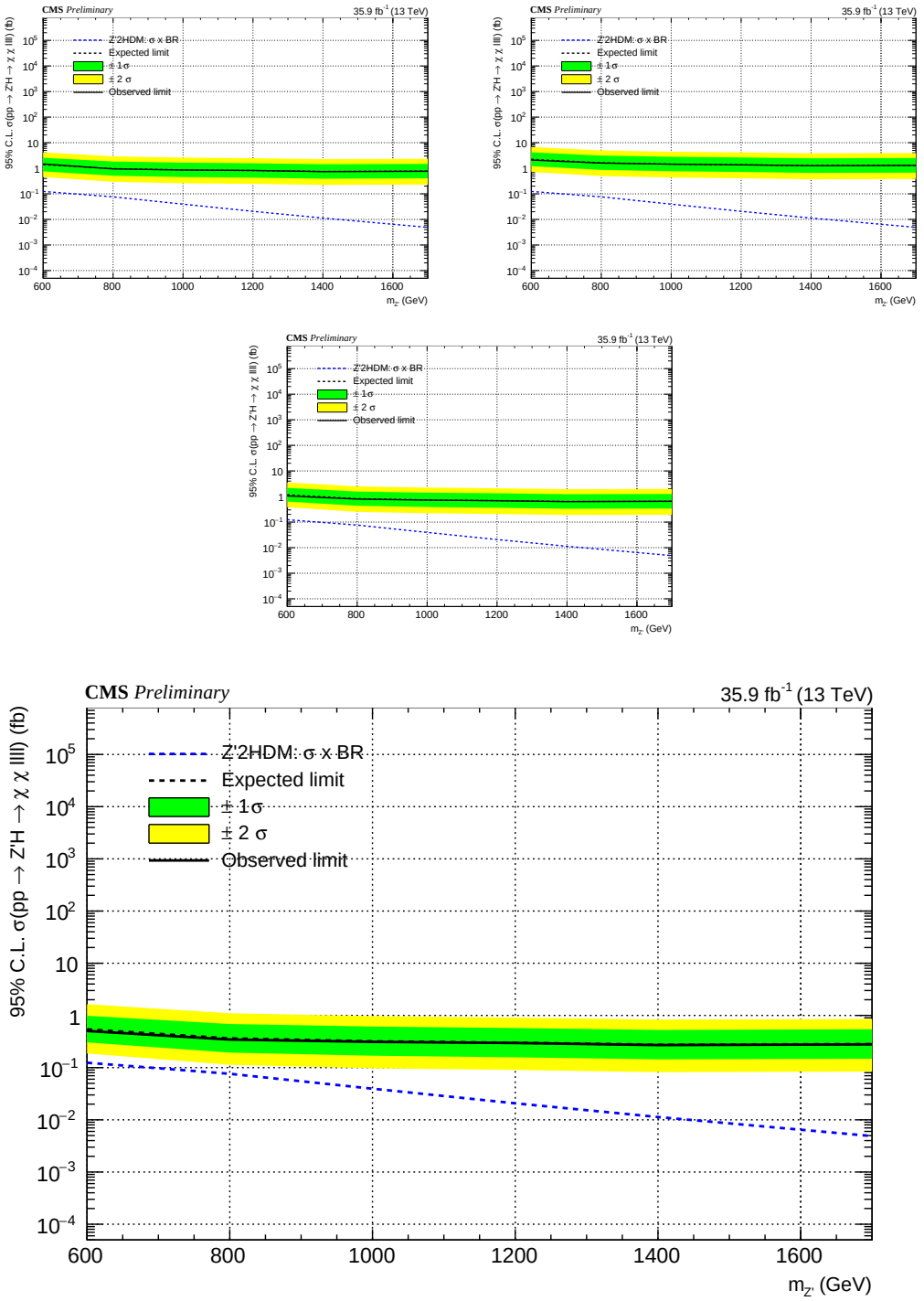
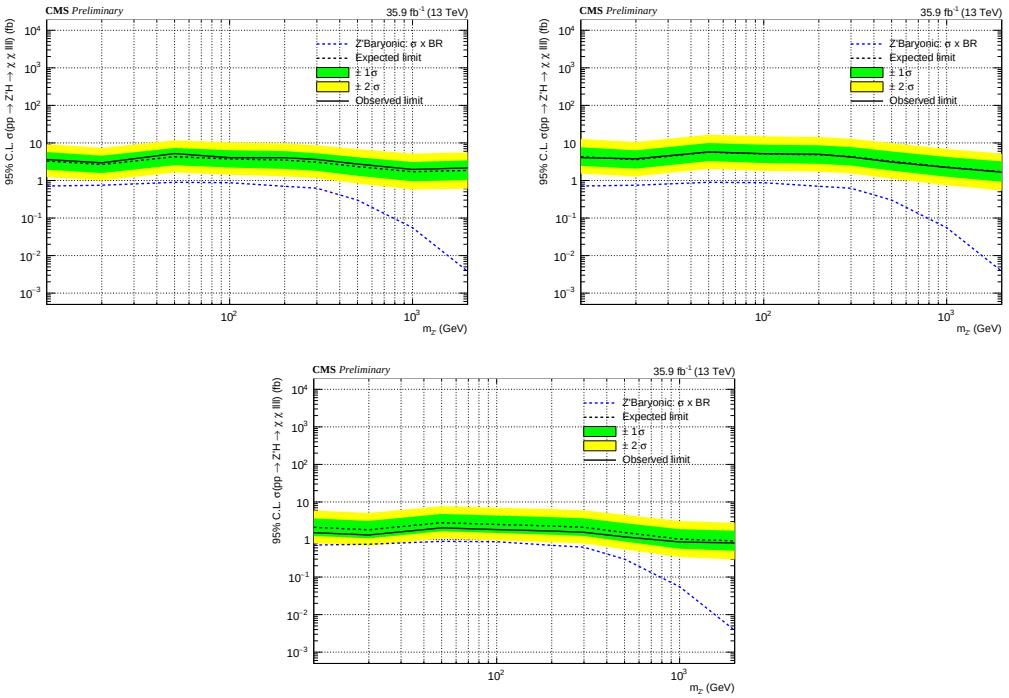


FIGURE 6.6: 95% C.L. upper limits on $\sigma \times$ BR for each of the channels, 4μ, $4e$, and $2e2\mu$ (top), and for the combination of all of them (bottom), as obtained performing a shape analysis on E_T^{miss} and using Z'2HDM with $m_{A0} = 300$ GeV as the benchmark model using 2016 data collected by CMS.



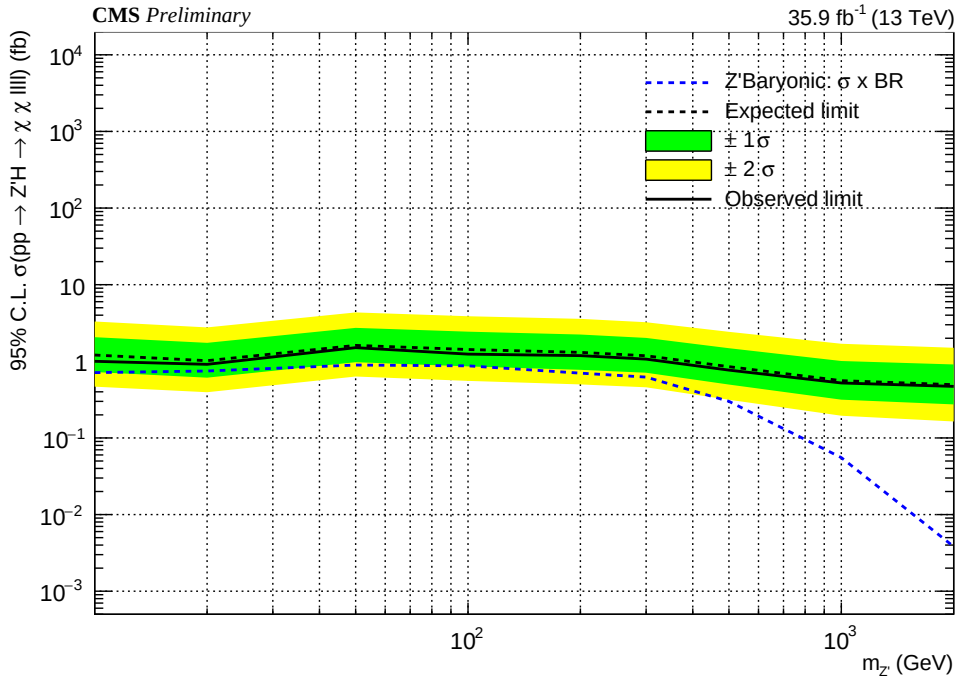


FIGURE 6.7: 95% upper limits on the cross section for each of the channels, 4μ, $4e$ and $2e2\mu$ (top), and for the combination of all of them (bottom) obtained performing a shape analysis on E_T^{miss} using Z'Baryonic model as a benchmark.

Conclusions

A search for dark matter candidates associated with a SM-like Higgs decaying into $ZZ \to 4l$ has been performed with the CMS experiment at LHC. For this study data collected during 2016 corresponding to an integrated luminosity of 35.9 fb^{-1} at a center of mass energy $\sqrt{s} = 13$ TeV have been used.
No relevant excess is observed in real data and 95% CL upper limits on the cross section are derived in the context of two simplified models, the $Z'2HDM$ and $Z'Baryonic$ models for some benchmark scenarios corresponding the largest cross section.
Even if limited by the low branching fraction of the Higgs decay into four leptons this Mono-Higgs is almost able to exclude points of the parameter space of the $Z'Baryonic$ model at 95% confidence level once the three channels are combined.
Furthermore,a remarkable improvement in sensitivity with respect to the upper limits computed with the 2015 data corresponding to an integrated luminosity of 2.8 $fb^{-}1$.
This opens a quite exciting scenario, especially in the prospective of a significant luminosity increase to be expected during next years at LHC since the data collected by CMS during 2017 will certainly allow to exclude some of those points.

Bibliography

[1] J. Beringer *et al.*, *Review of Particle Physics (RPP)* Phys.Rev., D86:010001, 2012

[2] F. Halzen Alan D. Martin, *QUARKS LEPTONS; An Introductory Course in Modern Particle Physics* Published by John Wiley and Sons, 1984

[3] F. Englert and R. Brout., *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. 13:321?323, Aug 1964

[4] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys.Rev. Lett., 13:508?509, Oct 1964

[5] P. W. Higgs. *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev., 145:1156?1163, May 1966

[6] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, 10.1016/j.physletb.2012.08.021, 2012

[7] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, 10.1016/j.physletb.2012.08.020, 2012

[8] Nima Arkani-Hamed, *Beyond the Standard Model theory* Phys. Scr.2013 014023, 2013

[9] G. Jungman, M. Kamionkowski, K. Griest, *Supersymmetric Dark Matter*, arXiv:hep-ph/9506380, 1995.

[10] G. Bertone, D. Hooper, and J. Silk, *Particle Dark Matter: Evidence, Candidates and Constraints* Phys.Rept. 405, 279 (2005), hep-ph/0404175, 2005.

[11] L. Carpenter, A. DiFranzo, M. Mulhearn, C. Shimmim, S. Tulin, D. Whiteson, *Mono Higgs: a new collider probe for dark matter*, 10.1103/PhysRevD.89.075017, 2013.

[12] Aad, Georges *et al.*, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at* $\sqrt{s}$ *=8 TeV with the ATLAS detector*, 10.1103/PhysRevD.94.032005, 2016.

[13] Khachatryan, Vardan *et al.*, *Search for dark matter, extra dimensions, and unparticles in monojet events in proton−proton collisions at* $\sqrt{s} = 8$ *TeV*, 10.1140/epjc/s10052-015-3451-4, 2015

[14] Aad, Georges *et al.*, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at* $\sqrt{s}$ *=8 TeV with the ATLAS detector*, 10.1103/PhysRevD.94.032005, 2016.

[15] Khachatryan, Vardan *et al.*, *Search for Monotop Signatures in Proton-Proton Collisions at* $\sqrt{s} = 8$ *TeV*, 10.1103/PhysRevLett.114.101801, 2015

[16] Khachatryan, Vardan *et al.*, *Search for the production of dark matter in association with top-quark pairs in the single-lepton final state in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, 10.1007/JHEP06(2015)121, 2015

[17] G. Aad *et al.*, *Search for new phenomena in events with a photon and missing transverse momentum in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector* 10.1103/PhysRevD.92.059903, 10.1103/PhysRevD.91.012008, 2015

[18] S. Chatrchyan *et al. Search for Dark Matter and Large Extra Dimensions in pp Collisions Yielding a Photon and Missing Transverse Energy*, 10.1103/PhysRevLett.108.261803 , 2012

[19] V. Khachatryan *et al.*, *Search for new phenomena in monophoton final states in proton-proton collisions at* $\sqrt{s} = 8$ *TeV* 10.1016/j.physletb.2016.01.057", 2014

[20] V. Khachatryan *et al.*, *Search for physics beyond the standard model in final states with a lepton and missing transverse energy in proton-proton collisions at sqrt(s) = 8 TeV*", 10.1103/PhysRevD.91.092005, 2014

[21] G. Aad *et al.*, *Search for new particles in events with one lepton and missing transverse momentum in pp collisions at* $\sqrt{s}$ *= 8 TeV with the ATLAS detector*, 10.1007/JHEP09(2014)037, 2014

[22] G. Aad *et al.*, *Search for dark matter in events with a Z boson and missing transverse momentum in pp collisions at* $\sqrt{s}$*=8 TeV with the ATLAS detector*, 10.1103/PhysRevD.90.012004, 2014

[23] G. Aad *et al.*, *Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, 10.1103/PhysRevLett.112.041802, 2013

[24] V. Khachatryan *et al. Search for dark matter and unparticles produced in association with a Z boson in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, 10.1103/PhysRevD.93.052011, 2016

[25] The ATLAS Collaboration, *Search for Dark Matter in association with a Higgs boson decaying to b-quarks in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, 2016

[26] G. Aad *et al.*, *Search for dark matter produced in association with a Higgs boson decaying to two bottom quarks in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector* 10.1103/PhysRevD.93.072007, 2015

[27] G. Aad *et al.*, *Search for Dark Matter in Events with Missing Transverse Momentum and a Higgs Boson Decaying to Two Photons in pp Collisions at*

$\sqrt{s} = 8$ *TeV with the ATLAS Detector*, 10.1103/PhysRevLett.115.131801, 2015

[28] The CMS Collaboration, *Search for dark matter in association with a top quark pair at* $\sqrt{s} = 13$ *TeV*, CMS PAS EXO-16-005, 2016

[29] The CMS Collaboration, *Search for dark matter in association with a top quark pair at* $\sqrt{s} = 13$ *TeV in the dilepton channel*, CMS PAS EXO-16-028, 2016

[30] The CMS Collaboration, *Search for dark matter in association with a boosted top quark in the hadronic final state at at* $\sqrt{s} = 13$ *TeV*, CMS PAS EXO-16-017, 2016

[31] The CMS Collaboration, *Search for dark matter production in association with jets,or hadronically decaying W or Z boson at* $\sqrt{s} = 13$ *TeV*, CMS PAS EXO-16-013, 2016

[32] The CMS Collaboration, *Search for Dark Matter and Large Extra Dimensions in the* γ + $\not{E}_T$ *final state in pp Collisions at* $\sqrt{s} = 13$ *TeV*, CMS PAS EXO-16-013, 2016

[33] A. De Simone, T. Jacques, *Simplified Models vs. Effective Field Theory Approaches in Dark Matter Searches*, 10.1140/epjc/s10052-016-4208-4, 2016

[34] Planck Collaboration, P. A. R. Ade emphet al., *Planck 2013 results. XVI. Cosmological parameters*, 10.1051/0004-6361/201321591, 2013

[35] Jessica Goodman, William Shepherd *LHC Bounds on UV-Complete Models of Dark Matte*, arXiv:1111.2359 [hep-ph], 2011

[36] O. Buchmueller, Matthew J. Dolan, Christopher McCabe *Beyond Effective Field Theory for Dark Matter Searches at the LHC*, 10.1007/JHEP01(2014)025, 2013

[37] J. Abdallah *et al.*, *Simplified Models for Dark Matter and Missing Energy Searches at the LHC*, arXiv:1409.2893 [hep-ph], 2014

[38] A. Boveia *et al.*, *Recommendations on presenting LHC searches for missing transverse energy signals using simplified s-channel models of dark matter*, arXiv:1603.04156, 2016

[39] LHC New PhysicsWorking Group Collaboration, D. Alves, *Simplified Models for LHC New Physics Searches*, 10.1088/0954-3899/39/10/105005 2011

[40] LHC New PhysicsWorking Group Collaboration, D. Alves, *Simplified Models for LHC New Physics Searches*, 10.1088/0954-3899/39/10/105005 2011

[41] M. Cirelli, E. Del Nobile, P. Panci, *Tools for model-independent bounds in direct dark matter searches*, JCAP10(2013)019 2013

[42] D. Abercrombie *et al.*, *Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum*, arXiv:1507.00966 [hep-ex], 2015

[43] Berlin, Asher and Lin, Tongyan and Wang, Lian-Tao, *Mono-Higgs Detection of Dark Matter at the LHC*, 10.1007/JHEP06(2014)078, 2014

[44] Particle Data Group and Beringer, J. emphet al., *Review of Particle Physics*, 10.1103/PhysRevD.86.010001, 2012

[45] P. Nason, G. Zanderighi, W^+W^- , *WZ and ZZ production in the POWHEG-BOX-V2*, 10.1140/epjc/s10052-013-2702-5, 2014

[46] J. M. Campbell, and R. K. Ellis, *MCFM for the Tevatron and the LHC*, 10.1016/j.nuclphysbps.2010.08.011, Nucl. Phys. Proc. Suppl., 2010

[47] Campbell, John M. and Ellis, R. Keith and Williams, *Bounding the Higgs width at the LHC using full analytic results for* $gg- > e^-e^+\mu^-\mu^+$, 10.1007/JHEP04(2014)060, JHEP, 2014

[48] Aaltonen, T. *et al.*, *Search for new particles decaying into dijets in proton-antiproton collisions at s**(1/2) = 1.96-TeV*, 10.1103/PhysRevD.79.112002, 2009

[49] Chatrchyan, Serguei *et al.*,, *Search for narrow resonances using the dijet mass spectrum in pp collisions at* $\sqrt{s}$*=8??TeV*, 10.1103/PhysRevD.87.114015, 2013

[50] C. D. Carone and H. Murayama, *Possible light U(1) gauge boson coupled to baryon number*, 10.1103/PhysRevLett.74.3122, 1995

[51] Agashe, Kaustubh and Servant, Geraldine, *Baryon number in warped GUTs: Model building and (dark matter related) phenomenology*, 10.1088/1475-7516/2005/02/002, 2005

[52] P. Fileviez Perez and M. B. Wise,

[53] , 10.1103/PhysRevD.82.079901, 10.1103/PhysRevD.82.011901, 2010

[54] C. Battilana *on behalf of the CMS Collaboration*, *The CMS muon system: status and upgrades for LHC Run-2 and performance of muon reconstruction with 13 TeV data*, Journal of Instrumentation, Volume 12, January 2017

[55] L. Cadamuro, *The CMS Level-1 trigger system for LHC Run II*, Journal of Instrumentation, Volume 12, March 2017

[56] T. Bawej *et al.*, *The New CMS DAQ System for Run-2 of the LHC*, IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 62, NO. 3, JUNE 2015

[57] `https://www.recas-bari.it/`

[58] L. Roszkowski, E. M. Sessolo, S. Trojanowski, *WIMP dark matter candidates and searches - current issues and future prospects*, arXiv:1707.06277 [hep-ph], 2017

[59] THE LHC STUDY GROUP Collaboration, *LHC - the Large Hadron Collider accelerator project*, CERN-AC-93-03, [http://cdsweb.cern.ch/record/87244]

[60] THE LHC STUDY GROUP Collaboration, *LHC - the Large Hadron Collider conceptual design*, CERN-AC-95-05, [http://cdsweb.cern.ch/record/291782]

[61] L. Evans, P. Bryant [LHC Collaboration], *LHC Machine*, JINST 3, S08001, 2008

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

[63] CMS Collaboration, *The Compact Muon Solenoid technical proposal*, CERN-LHCC-94-38, [http://cdsweb.cern.ch/record/290969]

[64] R. Adolphi *et al.* [The CMS Collaboration], *The CMS experiment at the CERN LHC*, JINST 3, S08004, 2008

[65] The CMS Collaboration *ECAL 2010 performance results*, CMS Detector Performance Summary DP-2011/008, 2011

[66] P.Adzic *et al.* [The CMS Collaboration], *Energy resolution of the barrel of the CMS electromagnetic calorimeter*, JINST 2, P04004, 2007

[67] Alwall, J. and Frederix, R. and Frixione, S. and Hirschi, V. and Maltoni, F. and Mattelaer, O. and Shao, H. -S. and Stelzer, T. and Torrielli, P. and Zaro, M., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, 10.1007/JHEP07(2014)079, 2014

[68] `home.thep.lu.se/ torbjorn/pythia82html/`

[69] Alioli, S. and Nason, P. and Oleari, C. and Re, E, *NLO vector-boson production matched with shower in POWHEG*, 10.1088/1126-6708/2008/07/060, 2008

[70] Nason, Paolo, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, 10.1088/1126-6708/2004/11/040, 2004

[71] Frixione, Stefano and Nason, Paolo and Oleari, Carlo, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, 10.1088/1126-6708/2007/11/070, 2007

[72] Ball, Richard D. and others", *Parton distributions for the LHC Run II*, 10.1007/JHEP04(2015)040, 2015

[73] Bagnaschi, E. and Degrassi, G. and Slavich, P. and Vicini, A, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM* 10.1007/JHEP02(2012)088, 2012

[74] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP, Vol. 02, 2010, 10.1007/JHEP02(2010)037, 2010

[75] Hartanto, Heribertus B. and Jager, Barbara and Reina, Laura and Wackeroth, Doreen, *Higgs boson production in association with top quarks in the POWHEG BOX* 10.1103/PhysRevD.91.094003, 2015

[76] Luisoni, Gionata and Nason, Paolo and Oleari, Carlo and Tramontano, Francesco, *$HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, 10.1007/JHEP10(2013)083", 2013

[77] Gao, Yanyan and Gritsan, Andrei V. and Guo, Zijin and Melnikov, Kirill and Schulze, Markus and Tran, Nhan V., *Spin determination of single-produced resonances at hadron colliders*, 10.1103/PhysRevD.81.075022, 2010

[78] Allison, John and others, *Geant4 developments and applications*, EEE Trans. Nucl. Sci., 10.1109/TNS.2006.869826, 2006

[79] S. Baffioni *et al.*, *Electron reconstruction in CMS*, Eur.Phys.J. C49, 10.1140/epjc/s10052-006-0175-5, 2007

[80] G. Abbiendi, N. Adam *et al.*, *Muon Recostruction in the CMS Detector*, CMS AN 2008/097, July 2009

[81] F.. Beaudette on behalf of the CMS Collaboration, *The CMS Particle Flow Algorithm*, arXiv:1401.8155 [hep-ex], Proceedings of the CHEF2013 Conference, 2013

[82] M. Ahmad *et al. New results on the study of Higgs boson production in the four-lepton final state at $\sqrt{s}$ = 13 TeV*, CMS Analysis Note 2016/217, 2016

[83] M. Ahmad *et al. Measurement of the properties of the Higgs boson in the four-lepton final state at $\sqrt{s}$ = 13 TeV*, CMS Analysis Note 2015/277, 2015

[84] G. Miniello *et al. Measurement of the properties of the Higgs boson in the four-lepton final state at $\sqrt{s}$ = 13 TeV*, CMS Analysis Note 2016/442, 2016

[85] G. Miniello *et al. Search for Dark Matter Produced in Association with a Higgs Boson in the four lepton final state at 13 TeV*, CMS Analysis Note 2016/328, 2017

[86] The CMS Collaboration, *Studies of Higgs boson production in the four-lepton final state at $\sqrt{s}$ = 13 TeV*, CMS-PAS-HIG-15-004, 2015

[87] The CMS Collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$*, 10.1088/1748-0221/10/06/P06005, 2015

[88] G. Petrucciani, *Observation of a New State in the Search for the Higgs Boson at CMS*, 10.1007/978-88-7642-482-3, 2013

[89] M. Cacciari, G. P. Salam, G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP 0804:063, 10.1088/1126-6708/2008/04/063, 2008

[90] BtagPOG, *Usage of b/c Tag Objects for 13 TeV Data in 2016 and 80X MC* CMS Twiki,

https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco, 2016

[91] JetMET POG, https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2, 2017

[92] JetMET POG, https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETRun2Corrections, 2017

[93] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN, Geneva, 2016

[94] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, "Nucl. Phys. Proc. Suppl., Vol. 205, 2010, 10.1016/j.nuclphysbps.2010.08.011, 2010

[95] M. Grazzini, S. Kallweit, D. Rathlev, *ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD*, 10.1016/j.physletb.2015.09.055, 2015

[96] M. Bonvini, F. Caola, Fabrizio, S. Forte, K. Melnikov, G. Ridolfi, *Signal-background interference effects in $gg \to H \to WW$ beyond leading order*, Phys. Rev. D, 10.1103/PhysRevD.88.034032, 2013

[97] K. Melnikov and M. Dowling, *Production of two Z-bosons in gluon fusion in the heavy top quark approximation*, Phys. Lett. B, Vol. 744, 2015, 10.1016/j.physletb.2015.03.030, 2015

[98] C. S. Li, H. T. Li, D. Y. Shao, Ding Yu and J, Wang, *Soft gluon resummation in the signal-background interference process of $gg(\to h^*) \to ZZ$ "*, ARXIV:1504.02388, 2015

[99] The LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, CERN Report, 10.5170/CERN-2013-004", 2013

[100] G. Passarino, *Higgs CAT*, Eur. Phys. J. C, Vol. 74, 10.1140/epjc/s10052-014-2866-7, 2014

[101] The ATLAS and CMS Collaboration, The LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011* ATL-PHYS-PUB-2011-11, CMS NOTE-2011/005, 2011

[102] R. D. Cousins and G. J. Feldman, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D, 57(7):3873-3889, 1998