SEARCH FOR RESONANT DOUBLE HIGGS PRODUCTION WITH bbZZ DECAYS IN THE $b\bar{b}\ell\ell\nu\bar{\nu}$ FINAL STATE IN pp COLLISIONS AT $\sqrt{s} = 13$ TeV

by

Rami Kamalieddin

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Physics and Astronomy

Under the Supervision of Professor Ilya Kravchenko

Lincoln, Nebraska

July, 2019

SEARCH FOR RESONANT DOUBLE HIGGS PRODUCTION WITH bbZZ DECAYS IN THE $b\bar{b}\ell\ell\nu\bar{\nu}$ FINAL STATE IN pp COLLISIONS AT $\sqrt{s} = 13$ TeV

Rami Kamalieddin, Ph.D.

University of Nebraska, 2019

Adviser: Ilya Kravchenko

Since the discovery of the Higgs boson in 2012 by the ATLAS and CMS experiments, most of the quantum mechanical properties that describe the long-awaited Higgs boson have been measured. Due to the outstanding work of the LHC, over a hundred of fb^{-1} of proton collisions data have been delivered to both experiments. Finally, it became sensible for analyses teams to start working with a very low cross section processes involving the Higgs boson, e.g., a recent success in observing ttH and VHbb processes. One of the main remaining untouched topics is a double Higgs boson production. However, an additional hundred of fb^{-1} per year from the HL-LHC will not necessarily help us much with the SM double Higgs physics, as the process may remain unseen even in the most optimistic scenarios. The solution is to work in parallel on new reconstruction and signal extraction methods as well as new analysis techniques to improve the sensitivity of measurements. This thesis is about both approaches: we have used the largest available dataset at the time the analysis has been performed and developed/used the most novel analysis methods. One such method is the new electron identification algorithm that we have developed in the CMS electron identification group, to which I have had a privilege to contribute during several years of my stay at CERN.

The majority of this thesis is devoted to techniques for the first search at the LHC for double Higgs boson production mediated by a heavy narrow-width resonance in

the $b\bar{b}ZZ$ channel: $X \to HH \to b\bar{b}ZZ^* \to b\bar{b}\ell\ell\nu\bar{\nu}$. The measurement searches for the resonant production of a Higgs boson pair in the range of masses of the resonant parent particle from 250 to 1000 GeV using 35.9 fb^{-1} of data taken in 2016 at 13 TeV. Two spin scenarios of the resonance are considered: spin 0 and spin 2. In the absence of the evidence of the resonant double Higgs boson production from the previous searches, we proceed with setting the upper confidence limits.

“Here will be a quote ”

name, year.

ACKNOWLEDGMENTS

This will be a long list!

Table of Contents

List of Figures

List of Tables

CHAPTER *1*

Introduction

This chapter describes the development of particle physics following the historical order of how particles were discovered and parts of the Standard Model (SM) were constructed.

1.1 "All things are made of atoms"

The SM is the theory of particles and their interactions that was built during many decades of intense experimental and theoretical work. Particles in this theory are elementary and have no internal structure, meaning they cannot be divided further. This idea of point-like particles is crucial since it reflects the goal of particle physics to find and describe the smallest and the most fundamental scale at which nature works. For the last several decades the SM has been the most tested theory of elementary particles and forces that is presently generally accepted by the whole physics community.

Formally, all SM elementary particles are split into two classes: fermions and bosons. Particles with a half-integer spin 1/2 (quarks and leptons) are called fermions since they obey Fermi-Dirac statistics [1]. The other class of particles is bosons. They are force carriers, have an integer spin, and are characterised by the Bose-Einstein statistics. A rigorous mathematical description of the SM will be given in the next chapter.

The quest for the smallest scale and the theory that would describe it were the key ideas that ultimately led to the development of the SM. However, particle physics did not start with quarks. Physicists were discovering more and more fundamental scales over the course of hundreds of years; starting first with our macroscopic world that is made of atoms (the atomic theory). Richard Feynman, a Physics Nobel Prize winner, once summarised in a single phrase what he believed to be the most important fact about the world around us: "all things are made of atoms" [2]. Feynman himself was the father of quantum electrodynamics, and in this simple statement - delivered originally to Caltech students and now known to public through his series of physics books - he decided not to go into quantum mechanics principles and instead illustrated at the rather highly abstract that everything can be made of a set of smaller particles, praising the importance of the atomic theory.

Today the physics community knows that atoms are not elementary particles. Instead, they have heavy nuclei and light electrons "orbiting" around the nucleus on the electron shells. The nucleus is positively charged proportionally to the number of protons it contains. To provide the stability of the nuclei of the heavy atoms our world also needs neutrons, which have no electric charge. Going to an even smaller scale, it is now known that protons and neutrons are not elementary; rather, they are composed of point-like constituents that are called quarks (see Figure 1.1).

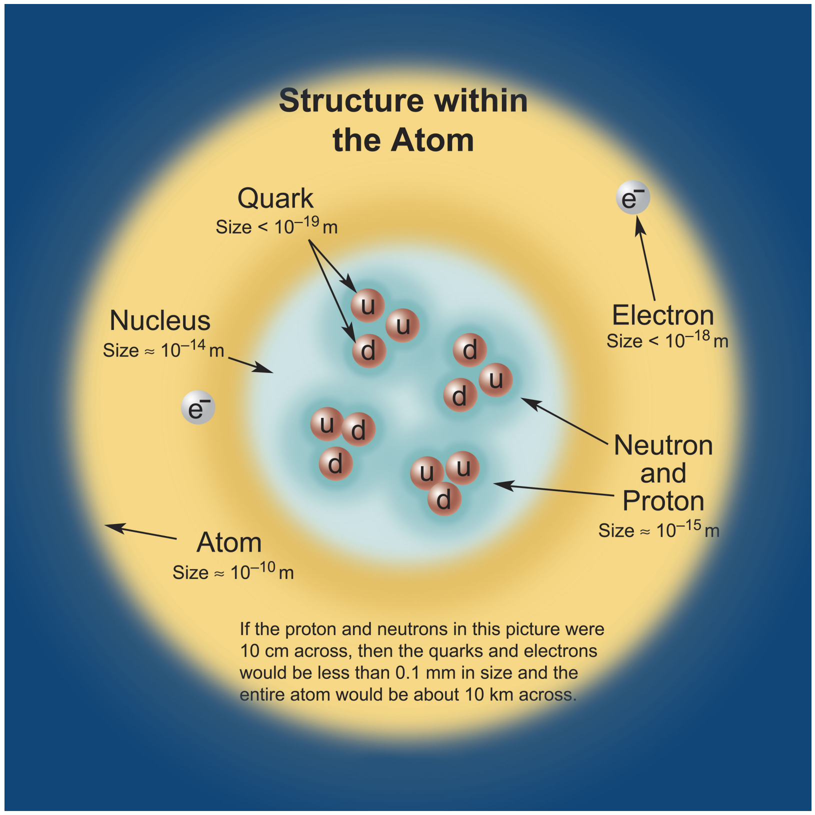


Figure 1.1: The structure of an atom. Approximate scale values are indicated.

Quarks were proposed by Gell-Mann and also by Zweig to explain periodicity in properties of observed subatomic particles [3]. Quarks are arranged in three families or three generations of doublets. A doublet is a mathematical construct that is used to describe a two-value system. For example, in the design of Gell-Mann and Zweig, each quark doublet in their theory is a two quark system that has an "up" quark with electric charge $+1/3$ and a "down" quark with charge $-2/3$. For antiquarks, the signs of the charges are reversed.

The physics world before Gell-Mann and Zweig got used to the fact that particles have integer charges due to an enormous number of observations. The fact that the quark charge values were fractional was so revolutionary to Gell-Mann that he decided not to publish his article in a highly prestigious journal but, expecting a rejection, decided to go with a second tier one [3]. However, with time, out of all the theories that were trying to explain the difference in masses of observed hadrons it was the hypothesis of Gell-Mann and Zweig that was not disproven. Now, the quark theory is generally accepted and is one of the key elements of the SM.

The SM includes six different types of quarks: up, down, charm, strange, top, and bottom. To distinguish one quark from another there is a "flavor" number assigned to them. For instance, a charm quark has +1 unit of "charmness", while a strange quark has -1 unit of "strangeness". All of the other flavor fields are zero for a given quark.

Another important characteristic of quarks was revealed at the e^+e^- colliders when physicists compared production rates of muons and hadrons. At that time it was assumed that quarks differ only by flavor, therefore, when comparing the calculations with the data, the theory was off by a factor of three. This was the motivation to introduce a new quark property - a "colour" - and thus three quark colours: green, blue, and red.

1.2 A brief history of particle physics

The electron is the first particle that was observed in a particle physics experiment. The electron belongs to a family of leptons. A lepton is an elementary particle with a spin 1/2. Charged leptons participate in all but strong interactions. Neutral leptons or neutrinos - interact only weakly, which will be discussed in more detail later in this chapter. The electron was discovered by Thompson [4] in 1897 when he was studying the properties of a cathode ray. Due to this discovery, that year may be considered the beginning of an era of a particle physics: dozens of particles were discovered in the next decades. In 1936, another lepton was observed, the muon [5], in an experiment of Anderson and Neddermeyer who studied cosmic radiation. In essence, a muon has very similar characteristics to those of an electron, but its mass is 207 times heavier. No explanation for this mass difference exists in the SM [1].

Leptons are also arranged in generations, analogously to quark families. Each generation is a doublet that consists of a charged lepton (electron, muon or tau) with the charge -1 and a neutral lepton (corresponding electron, muon, or a tau neutrino). Electron and muon neutrinos had been directly detected in experiments of 1956 and 1962, respectively. The existence of the electron neutrino is deduced from the violation of the conservation of energy in a beta decay. The muon neutrino [7] was discovered by Schwartz, Lederman, and Steinberger during an experiment with a pion beam where leptons from the pion decays arrived to the an aluminum spark chamber after passing a steel wall. Fifty-one events of interest had been observed after running the experiment for several months. Those events could not be initiated by electron neutrinos, since they will interact with the metal and produce electrons. The presence of narrow muon tracks in the chamber in each event, hence muons, was a clear indication that those neutrinos were of a different kind - they were muon neutrinos. Finally, a tau lepton and a tau neutrino were discovered in 1975 and 2000 correspondingly [8, 9]. With that, all three families of the SM leptons were observed: a long-awaited tau neutrino, which decades ago was theoretically speculated to exist, was finally discovered experimentally. In a like manner to families of quarks, lepton masses grow with each generation, where a tau from the third generation is the heaviest lepton. To classify leptons of different families the lepton numbers were reserved: 1 unit of electron number to an electron and an electron neutrino, 1 unit

[1] According to mathematical physicist Carl Bender, who is known for advances in perturbative Quantum Field Theory (QFT) and demonstration of the importance of parity-time (PT) symmetry in quantum theory, there is a story that Feynman was able to derive the mass of the muon starting with the mass of an electron, but the world has never seen that calculation published [6]

of muon number to a muon and a muon neutrino, and 1 unit of tau number to a tau and a tau neutrino.

1.3 Fundamental forces

In nature there are four fundamental forces: gravitational, weak, electromagnetic, and strong forces. This thesis will classify all four forces [10] in terms of the relative strength, the range that they can cover, the spin of the mediator, and whether the force's nature (when appplicable) is attractive, repulsive, or both. This should be taken critically, since this is quite an ambiguous categorisation. It has a deep pedagogical meaning, though, because it helps to illustrate in which regime each of the forces is dominant. According to the world known mathematical physicist Carl Bender, this is of great importance since it is one of the main approaches to solving physics problems: to know which effects are the dominant and which are sub-dominant. This helps to justify what effect can be neglected and what approximation can be used. Thus, it allows the possibility to do calculations for problems where closed-form solutions do not exists, which is almost all the complex phenomena around us [6].

The first force on our list of forces is the gravitational force. This force governs the Universe at the macroscopic level of astronomic objects: planets, solar systems, galaxies. The first theory of gravity was formulated by Newton [11]. Einstein later developed a new theory of gravity (GR). The key difference is that the Newtonian gravity had several "absolutes" that GR does not have: absolute time and space, a preferred separation of spacetime into time and spatial parts, absolute simultaneity, etc. [12]. Butterworth gives a good historical perspective in [13].

It is worth noting that the gravitational force is not included in the SM. Attempts are ongoing to expand the SM, e.g., adding the graviton as a mediator, but no real success so far has been achieved to create a renormalizable theory that would combine both the SM and gravity [14]. Surprisingly, gravity is the weakest force; the only reason why the motion of planets and galaxies is governed by gravity is because those are gigantic objects. Gravity effects become the dominant ones at the macroscopic scale because of an enormous number of particles involved in the interaction. If the strength of the strongest force, which is the strong force, is set to 1, then the strength of the gravity will be about 10^{-41}. In a modern High Energy Physics (HEP) language, which uses the term "coupling constant" to quantify the strength of the interaction between two elementary particles for a given force, a gravitational coupling constant can be considered as a constant characterizing the gravitational attraction between, e.g., a pair of electrons. In this case, $\alpha_G \approx 10^{-39}$. It is contemplated that the gravity mediator, if it exists, would have a charge of zero, zero mass, spin 2, and should be a stable particle. From the observations, the gravitational force is of the infinite range and its nature is purely attractive, while the other three forces can exhibit both an attractive and a repulsive behaviour. Einstein's general relativity theory, though not a quantum theory, is the only working theory of gravity as of now.

The next force, the weak force, is mediated by a charged W (charge +1/-1) boson or a neutral Z boson, thus giving name to charged and neutral weak interactions correspondingly. All SM fermions, quarks and leptons, experience the weak force. All three weak bosons (W^+, W^-, and Z) have spin 1. The relative strength of the weak force is 10^{-16} ($\alpha_W \approx 10^{-6}$) and the range of applicability is 10^{-3} fm. The range of the force can be well approximated by the expression $\frac{\hbar}{mc}$ [15], where m is the mass of the mediator or of the parent particle that decays. The range of applicability of the weak force is relatively short since the bosons are quite massive: $m_{W^\pm} = 80.385$ GeV and $m_Z = 91.189$ GeV [2].

[2] GeV is the unit of the "natural system of units", in which $\hbar = c = 1$. This natural system of

Let us think of an interaction process as initial state particles interacting at the interaction vertex and producing final state particles. For our purposes, we follow this simplified description. In real HEP calculations the Feynman diagram approach is used [17]. In this approach, a set of rules and conventions is developed to substitute the mathematics of a given process by the corresponding diagram. In Feynman diagram formulation of the quantum mechanics, each vertex corresponds to an interaction term in the Lagrangian [3] describing a given process, where both the energy and the momentum of interacting particles have to be conserved. For the details about the actual principles behind the approach of Feynman diagrams, we refer the reader to [3, 15, 18]. In our simplified representation, charged weak interactions are interesting due to the fact that a primitive interaction vertex can be thought of as a point where a charged lepton is converted to a neutral lepton or vice versa. A good example is a muon decay, which is a conversion of the muon to a muon neutrino with the help of the W boson, which further decays to an electron and a corresponding electron antineutrino. It is worth noting that charged weak interactions do not conserve the flavor of quarks. Also, weak interactions do not conserve the generation number, e.g., members of doublets of the third and the second families can be converted into members of the lower family of quarks. This fact is reflected in the Cabibbo-Kobayashi-Maskawa (CKM) matrix [19]. The elements of this matrix are used to quantify the strength of the flavor-changing weak interactions. Since diagonal elements of this matrix are less than one and off-diagonal elements are non-zero, the CKM matrix represents a mismatch of quantum states of quarks when they propagate

units is very popular in the high-energy physics and is widely used in this thesis. Adoption of this system simplifies how many equations look. Using the natural system of units [16], masses, momenta, and energies are measured in electronvolts (eV), with GeV (10^9 eV) and TeV (10^{12} eV) being the most popular units in a modern high-energy physics due to the energy regimes involved

[3] The Lagrangian of the SM will be discussed in the next chapter

freely and when they take part in the weak interactions. In other words, the CKM matrix with non-zero off-diagonal elements means cross-generation interactions are allowed.

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}. \tag{1.1}$$

The third force, the electromagnetic (EM) force, is one of the main forces that we experience in our everyday life. The reason one can sit in the chair and does not fall further down due to gravity is that electrons of the body repel electrons of the chair. The relative strength of the EM force is 10^{-3} ($\alpha_{EM} \approx 1/137$) and the range of applicability is infinite. A photon, as the EM force's mediator, has zero mass and spin 1. The theory that describes photon interaction with leptons and quarks is called quantum electrodynamics (QED) and was developed in 1940s and 1950s by Tomonaga, Schwinger, Feynman, and Dyson [20]. Electric charge is conserved in EM interactions and no single photon-to-fermion vertex is possible; there are always two fermions that must be involved. This force can exhibit both an attractive ($e^{\pm}$ with $e^{\mp}$) and a repulsive behaviour ($e^{\pm}$ with $e^{\pm}$).

In the SM several multi-boson vertices are allowed. W and Z bosons that participate in weak interactions can couple to each other, so $WWZ, WWWW$, and $WWZZ$ vertices are possible in the SM. In addition, W bosons can couple to photons, so γWW, γWWZ, and $WW\gamma\gamma$ vertices are allowed too. Even though Z boson is massive and photon is a massless boson, the Z boson has a neutral charge. This makes it possible that any interaction where the photon is a force carrier, can also be mediated by the Z boson, but not vice versa.

The strong force, the fourth force of nature, is the strongest known force. Gluons are the carriers of this force. They have spin 1 and are massless. The relative strength of the strong force is 1 ($\alpha_s \approx 1$) and the range of applicability is about 1 fm. Each gluon carries one unit of color and one unit of anticolor. There are nine types of gluons but, technically, the ninth gluon is a color invariant and would give rise to an infinite range of the strong force, which contradicts experiments. That is why modern physics assumes that in our world only eight gluons exist [3, 19]. Gluons carry color charge and, thus, can couple to each other. For several high order processes in quantum chromodynamics (QCD), 3- and 4-gluon vertices have to be introduced to restore gauge invariance and no higher order vertices are required [21].

To summarise the knowledge about the SM forces, we show the reference figure of all allowed SM particle interactions and corresponding simple vertices (Fig. 1.2) in the Feynman diagram representation.

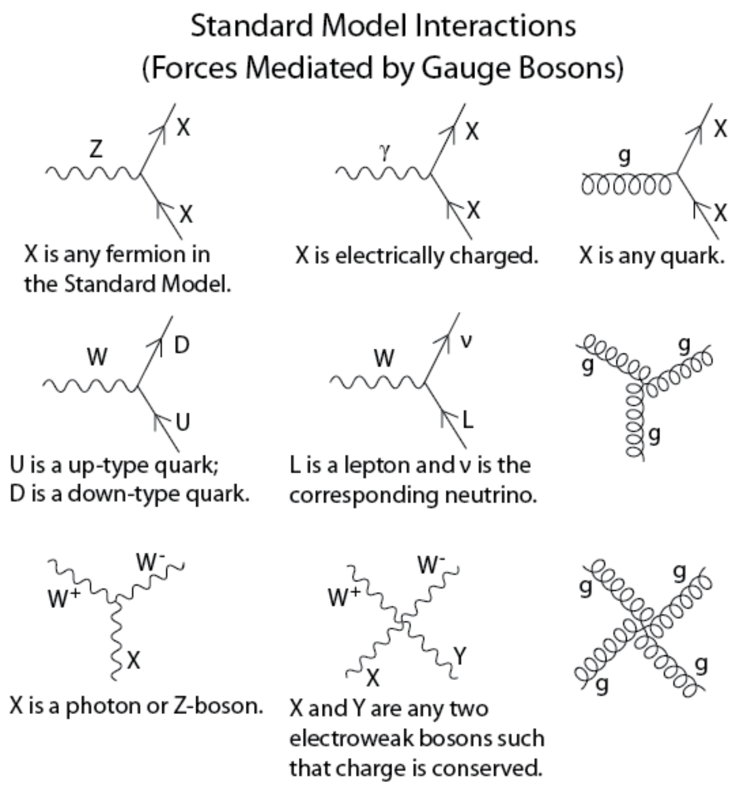


Figure 1.2: All SM interaction and simple vertices.

1.4 The Brout-Englert-Higgs mechanism

The description of the SM picture is not complete without mentioning the main particle - the Higgs boson - that was predicted almost 60 years ago, but was not observed until 2012 (see Fig. 1.3 [22]). After the electroweak (EW) unification by Glashow, Salam, and Weinberg [23], it was still not clear what the origin of the mass of fundamental particles is. In 1964, Robert Brout and François Englert [24], Peter Higgs [25], Gerald Guralnik, C. Richard Hagen, and Tom Kibble [26] (BEHGHK authors), proposed the method by which the particles can acquire mass. This technique consists of three stages and we will discuss them one-by-one:

1. The Brout-Englert-Higgs (BEH) mechanism
2. The BEH field
3. The Higgs boson.

The first stage, the BEH mechanism, is simply a spontaneous symmetry breaking (SSB) mechanism, which is a mathematical trick consisting of rewriting the original scalar fields in the EW Lagrangian, rearranging equations, and requiring that the fields are real. What does this lead to? The BEH authors started with a scalar complex field and a massless vector field and after SSB obtained a single real scalar field (Higgs boson) and a massive vector field. In terms of our physical world this it what gives mass to W and Z bosons.

The second stage is the BEH field. It exists everywhere and has been present almost since the Big Bang [27]. It is a property of our world. All the fundamental particles that interact with the BEH field acquire mass. Those, which do not interact directly (at the tree level) have no mass and all their energy is in the form of the momentum. Such particles can travel with the speed of light. The more the particle interacts with the BEH field, the higher is the coupling to the Higgs boson or simply the higher is the mass of the particle. The coupling of the Higgs boson to fermions is proportional to the mass of the fermions, and for W and Z bosons it is proportional to the squared mass of bosons, making the top quark and the Z boson the most massive fermion and boson respectively (see Fig. 1.4 [28]).

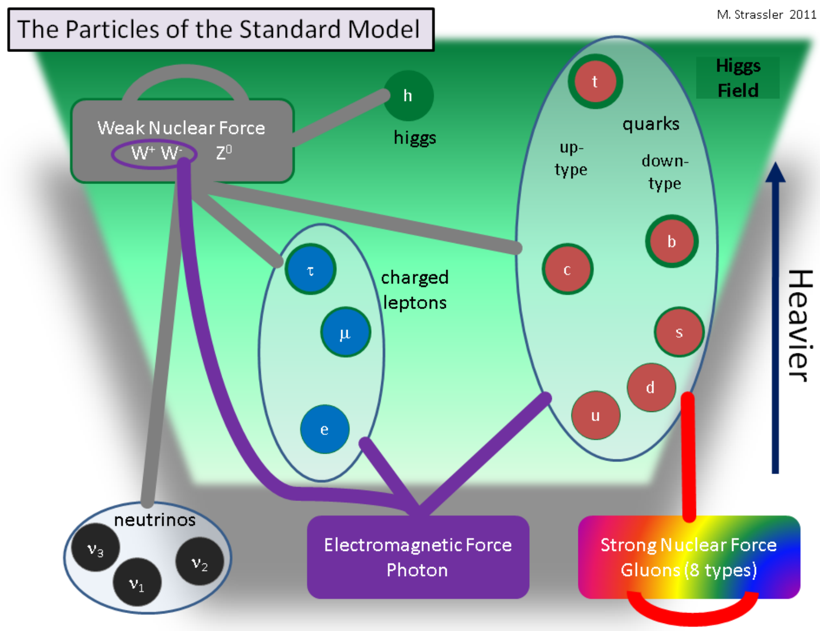


Figure 1.3: SM particles and force carriers. Self-interactions are also shown. The strength of the coupling to the Higgs boson increases from the bottom to the top, which is illustrated by the shades of the green color (the Higgs field).

The third and, arguably the most important stage, is the Higgs boson. The Higgs boson is the excitation of the BEH field. Thus, the Higgs bosons can be produced at colliders by pumping more and more energy in a small space-time region exciting the BEH field to "produce" the Higgs bosons. In reality this happens through making the LHC beams more energetic and thus, during the collision, producing more energetic gluons (and also more energetic quarks). The main Higgs boson production mechanism is called gluon fusion. During this process, two gluons interact through the virtual top quark loop and a Higgs boson is produced as a result. This accounts for about 90% of the overall LHC Higgs production at the 13 TeV energy. The second mechanism is vector boson fusion. The third mechanism is associated production with a weak boson. The smallest contributor to the Higgs boson production is the $t\bar{t}H$ process, which stands for the associated production of the top anti-top quark pair with the Higgs boson. All mentioned Higgs boson production mechanisms are presented in the form of Feynman diagrams in Fig. 1.5. The plot at the bottom is for the 8 TeV centre-of-mass energy. At 13 TeV the exact numbers for curves are different, but the main trends remain.

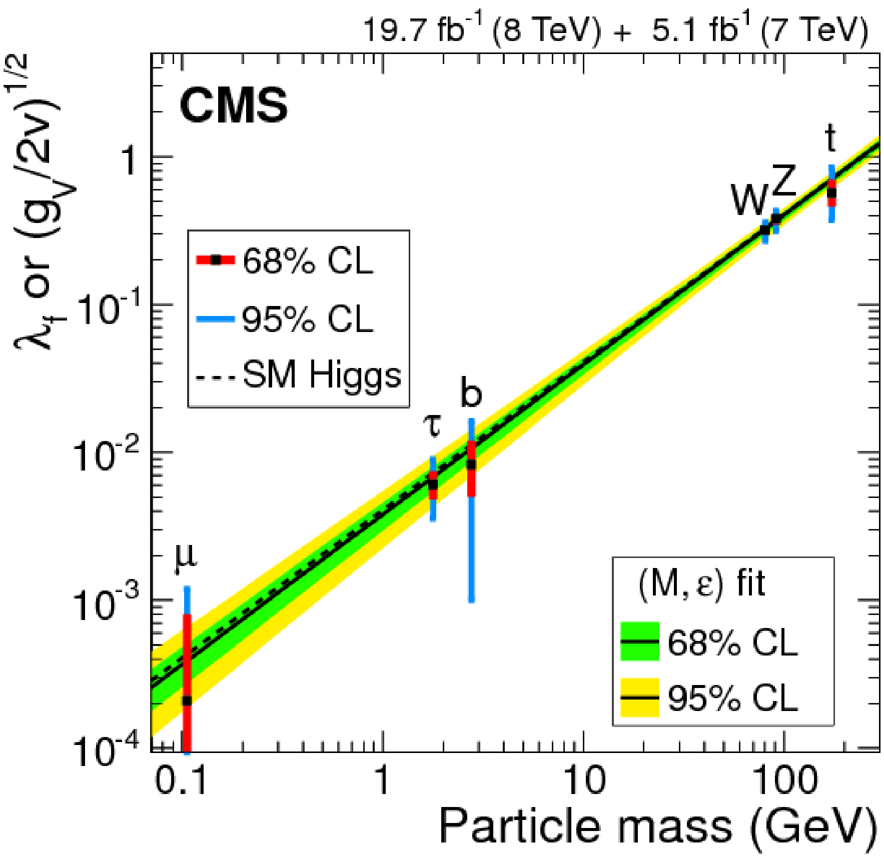


Figure 1.4: Coupling of particles to SM Higgs boson versus the mass of the particle, log-log scale is used. The y axis accommodates the couplings for both fermions and bosons.

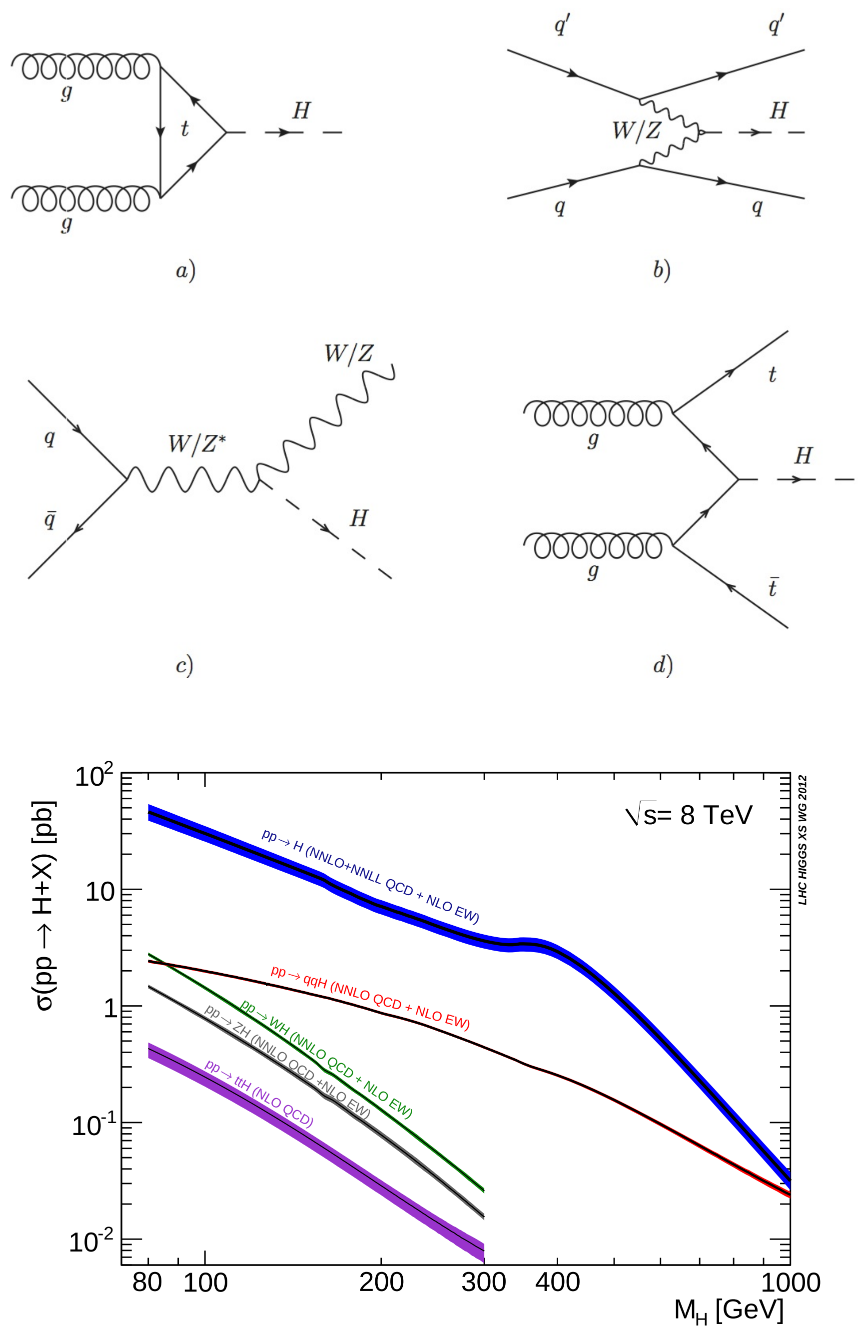


Figure 1.5: Top: SM Higgs boson production modes: a) a gluon fashion (in blue color at the bottom), b) a vector boson scattering (red), c) an associated production with a vector boson (green and brown), d) an associated production with the top anti-top pair (purple). Bottom: Higgs boson production modes as a function of the Higgs boson mass.

When describing Higgs boson physics one cannot avoid mentioning the decay channels of the Higgs boson. In physics, the term "branching fraction" is used to quantify the probabilities with which a parent particle decays to daughter particles (see Fig. 1.6). The picture shows a classic plot produced by theorist before the Higgs boson discovery in 2012. The values of branching fractions are given as a function of the Higgs boson mass. In this thesis we work with the SM Higgs bosons (≈125 GeV) and the measurement focuses on two specific Higgs boson decays, $H \to b\bar{b}$ and $H \to ZZ$. The first one has the highest branching fraction, while the second one gives a clean signature when subsequent $Z \to \ell\ell$ decays are selected.

Before we conclude with the BEHGHK method, a little bit of history, an irony of life, actually. The BEH particle is called the Higgs boson, but Peter Higgs was not the first to publish the article on the BEH mechanism, in fact, he was the last of the BEHGHK authors. His very first article was rejected since it contained no specific predictions or conclusions drawn from his calculations. This is why he was out-published by others. But this rejection made him write another article where he explicitly predicted the existence of the new boson. This is what made all the difference, as he was the first to predict a new boson, and so this boson now is called the Higgs boson.

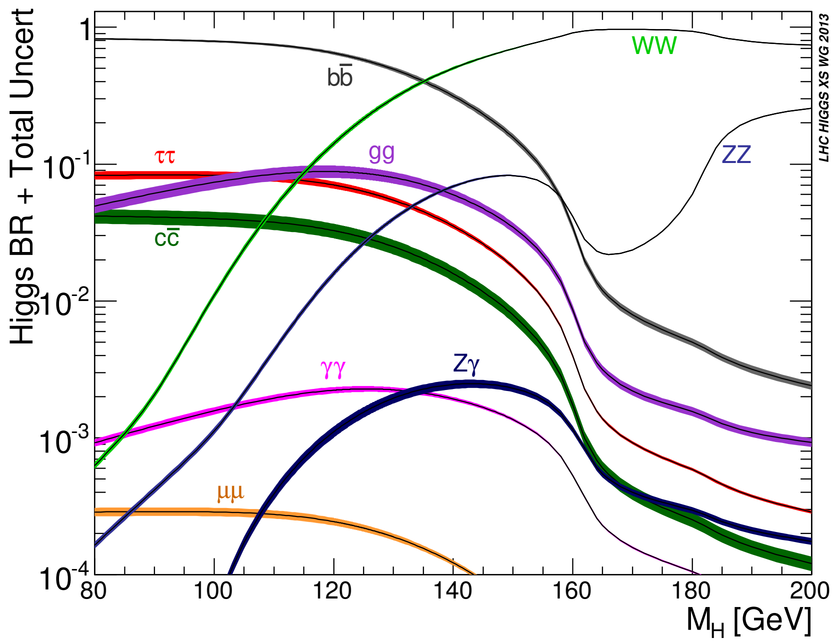


Figure 1.6: Higgs boson decay channels as a function of the Higgs boson mass. At 125 GeV the dominant decay mode is $H \to b\bar{b}$.

Even though the facts above tell us about how great the SM is, the SM is still far from being perfect. Masses of elementary particles are the parameters in this theory; they do not come from SM predictions. It is hypothesized that the SM could be a part of the larger ultimate theory, the so-called "Theory of Everything" (TOE), which is yet to be written (and was a dream of another genius, Einstein [29]). There is hope that the TOE will be able to explain many phenomena, such as the quark mass hierarchy, flavor mixing, etc. Also, in the SM all neutrinos are massless; however, it has been shown that they have a non-zero mass [30]. This fact is one of the main motivations for physicists to look for extensions of the SM.

Chapter 2

Theory

In the previous chapter, we introduced the SM and discussed the particles and their interactions that are described by this theory. In this chapter we will start with the general mathematical formalism of the SM. Then, in the second part we will focus on the double Higgs boson physics described in Beyond the Standard Model (BSM) theories.

2.1 Lagrangian formalism of the Standard Model

The SM uses Lagrangian mechanics as the mathematical approach to describe quantitatively the interactions of elementary particles and fields. The SM Lagrangian can be split into four main contributions [31]:

$$\mathcal{L}_{SM} = \mathcal{L}_{Yang-Mills} + \mathcal{L}_{ferm} + \mathcal{L}_{H} + \mathcal{L}_{Yuk} \tag{2.1}$$

where

- $\mathcal{L}_{Yang-Mills}$ represents gauge bosons and their self-interactions,
- $\mathcal{L}_{ferm}$ describes fermions and their interactions with the gauge bosons,

- $\mathcal{L}_H$ characterises the Higgs boson, its self-interaction, and its interaction with the gauge bosons to give them mass,
- $\mathcal{L}_{Yuk}$ gives details of fermions and their interactions with the Higgs boson, which, through the Yukawa mechanism, give mass to fermions.

The first term in the SM Lagrangian in full can be written as:

$$\mathcal{L}_{Yang-Mills} = -\frac{1}{4}W^i_{\mu\nu}(x)W^{\mu\nu}_i(x) - \frac{1}{4}B_{\mu\nu}(x)B^{\mu\nu}(x) - \frac{1}{4}G^a_{\mu\nu}(x)G^{\mu\nu}_a(x) \tag{2.2}$$

where

$$B_{\mu\nu}(x) \equiv \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.3}$$

$$W^i_{\mu\nu}(x) \equiv \partial_\mu W^i_\nu(x) - \partial_\nu W^i_\mu(x) - g\varepsilon^{ijk}W^j_\mu W^k_\nu \tag{2.4}$$

$$G^a_{\mu\nu}(x) \equiv \partial_\mu G^a_\nu(x) - \partial_\nu G^a_\mu(x) - g_s f^{abc}G^b_\mu G^c_\nu \tag{2.5}$$

with μ and ν indices running from 0 to 3, *SU(2)* indexes $i, j, k = 1, 2, 3$, and *SU(3)* indices given by $a, b, c = 1, ..., 8$. Terms ∂_μ and ∂_ν represent four-vector covariant derivatives. According to the Noether's theorem, each symmetry is intrinsically connected to a conservation law [32]. The invariance of the Lagrangian under certain transformations or, in other words, how the fields in the Lagrangian ($\mathcal{L}_{Yang-Mills}$ in this case) are related to their corresponding underlying symmetries, is explained in the following way:

- $B_{\mu\nu}$ corresponds to $U(1)_Y$ symmetry of the weak hypercharge Y_k with *U(1)* being a unitary one-by-one matrix (a scalar),

- $W^i_{\mu\nu}$ corresponds to $SU(2)_I$ symmetry of the weak isospin I^i_w. Another common representation is $SU(2)_L$, since only left-handed SM fermions are transformed under this symmetry. $SU(2)_L$ is a unitary two-by-two matrix with the determinant equal to one.

- $G^a_{\mu\nu}$ corresponds to $SU(3)_c$ symmetry of the QCD color charge with $SU(3)_c$ being a unitary three-by-three matrix with the determinant equal to one.

The "B" field is a kinematic term, "W" and "G" terms describe interactions among the gauge bosons, g and ε are $SU(2)_L$ coupling and structure constants, g_s and f are coupling and structure constants for $SU(3)_c$.

The second term in the SM Lagrangian is:

$$\mathcal{L}_{ferm} = i\bar{\Psi}_L \not{D} \Psi_L + i\bar{\psi}_{l_R} \not{D} \psi_{l_R} + i\bar{\Psi}_Q \not{D} \Psi_Q + i\bar{\psi}_{u_R} \not{D} \psi_{u_R} + i\bar{\psi}_{d_R} \not{D} \psi_{d_R} \tag{2.6}$$

Notice, that the mass terms are still absent. In Eq. 2.6, Ψ represents a doublet of a charged lepton and a corresponding neutral lepton within the same lepton family of $SU(2)_L$. The subindex Q is reserved for a family of quarks, and ψ_R describes a right-handed leptonic singlet. Gauge boson interactions are present due to the derivative term:

$$D_\mu = \partial_\mu + igI^i_w W^i_\mu + ig' Y_w B_\mu + ig_s T^a_c G^a_\mu \tag{2.7}$$

Physical fields in this notation are represented by a linear combination of W and B fields:

$$
\begin{aligned}
A_\mu =& B_\mu \cos\theta_W + W^3_\mu \sin\theta_W \\
Z_\mu =& - B_\mu \sin\theta_W + W^3_\mu \cos\theta_W
\end{aligned}
\tag{2.8}
$$

where θ_W is known as the *Weinberg angle* [33].

With the first two terms of the SM Lagrangian – $\mathcal{L}_{Yang-Mills}$ and $\mathcal{L}_{ferm}$ – one obtains a valid theory of fermions and bosons; however, these particles are massless in this theory [34], which evidently contradicts reality. However, one cannot simply add mass terms by hand since that would break the Lagrangian gauge invariance. To solve this issue and to ensure that weak bosons are massive, one has to follow a more complex procedure: introduce a Higgs field and an SSB procedure (see section 1.4). During the SSB procedure, the $SU(2)_L \times U(1)_Y$ symmetry needs to be broken to have massive SM particles. The Higgs mechanism enters the SM Lagrangian through the corresponding Higgs Lagrangian term given by

$$
\mathcal{L}_H = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi), \qquad V(\Phi) = -\mu^2(\Phi^\dagger\Phi) + \frac{\lambda}{4}(\Phi^\dagger\Phi)^2 \tag{2.9}
$$

where

$$
\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 = (v + H + i\chi)/\sqrt{2} \end{pmatrix} \quad \text{with} \quad v = 2\sqrt{\frac{\mu^2}{\lambda}} \tag{2.10}
$$

and μ and λ are parameters of the Higgs potential. The Higgs field vacuum expectation value (*vev*) v, after the SSB, can be expressed in terms of μ and λ. The Higgs potential before and after the SSB is shown in Fig. 2.1. The importance of the $\mathcal{L}_H$ in the SM Lagrangian is crucial: after rearranging terms (full derivation is available at [15, 18]), the bosons finally have masses given by:

$$M_W = \frac{gv}{2}, \quad M_Z = \frac{M_W}{\cos\theta_W}, \quad M_H = \sqrt{2\mu^2} \tag{2.11}$$

The final contribution to the SM Lagrangian is the Yukawa term, with Yukawa Lagrangian given by:

$$\mathcal{L}_{Yuk} = -i\bar{\Psi}_L G_l \psi_{l_R}\Phi - i\bar{\Psi}_Q G_u \psi_{u_R}\tilde{\Phi} - i\bar{\Psi}_Q G_d \psi_{d_R}\Phi + h.c. \tag{2.12}$$

where $\tilde{\Phi} = i\sigma^2\Phi^*$. The 3×3 matrices G contain fermion masses, which are free parameters in the SM and have to be determined experimentally. These matrices describe the so-called Yukawa y_f couplings between the single Higgs doublet φ and the fermions. In the case of leptons, matrices G can be diagonalised to provide mass eigenstates of definite generation. Using the mass eigenstates, the strength of the coupling of a fermion y_f to a Higgs boson is given by $y_f = m_f/vev$, where m_f is the mass of a fermion and vev is set by μ and λ parameters, see Eq. 2.10. On contrast, W and Z boson masses are predicted by the SM and are directly related to the weak couplings and the Higgs field parameters, see Eq. 2.11.

The Higgs boson mass is proportional to the μ parameter. In 2012, using precise single Higgs boson mass measurements from both ATLAS and CMS experiments, the value of μ was determined. Additionally, many physics analyses at CERN have been targeting the measurement of the λ parameter, because it is related to the shape of Higgs potential. The simplest potential characterized by μ and λ parameters, sufficient to obtain the SSB phenomenon and give mass to the SM particles, is the so-called "Mexican hat" Higgs potential. This name reflects the fact that the shape of the potential after SSB resembles the Mexican hat, see Fig. 2.1. However, the real shape of the Higgs potential may be more complex or different from the Mexican Hat, thus, direct precise determination of the μ and λ parameters is a sensitive tool to test the limitations of the SM and may open doors to the BSM effects. The simplest interaction suitable for probing the higher order terms of the Higgs potential directly is the one where two Higgs bosons (HH) are present. All this makes HH physics, the topic of this thesis, one of the main goals for the future High Luminosity LHC (HL-LHC) that will start operations in 2026.

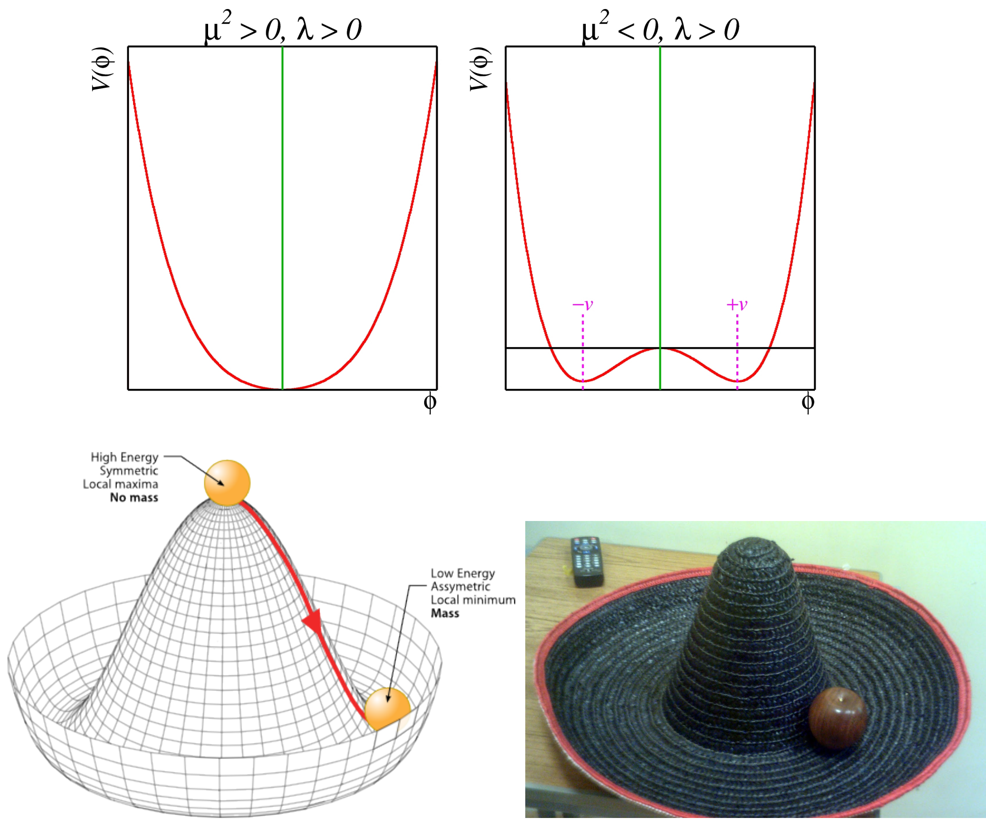


Figure 2.1: Top: Shape of the Higgs potential before and after the SSB that is determined at the leading orders by μ and λ parameters [35]. Bottom Left: Schematic drawing of the Higgs potential (after the SSB) that resembles a Mexican hat. Bottom Right: A real Mexican hat.

2.2 Double Higgs in Beyond the Standard Model Theories

While the mass parameter μ has been measured fairly accurately, λ parameter requires even HL-LHC to run for many years to get enough statistics since HH processes are rare and are of almost three orders of magnitude lower rate than the single Higgs boson production. Technically, the amount of the HL-LHC data is not enough to reach the sensitivity of the SM for HH processes. However, several BSM models predict resonant HH production to which the current LHC data could be sensitive. In these theories, HH is produced through a decay of a heavy resonance, which is not a part of the SM; thus, if such processes are found, a new chapter in HEP will be opened. In this thesis we focus on the resonant production of the HH system, which further decays to leptons and quarks. With the available CMS data, resonant HH analyses are starting to approach the needed sensitivity to many BSM models.

BSM theories such as [36–41] predict a resonant production of double Higgs boson events through a heavy resonance of a narrow width ($\sim O(1-10)$ GeV) [39]. Since the width parameter is proportional to the mass of the particle and its coupling to the Higgs boson, the values of the width larger than the $1-10$ GeV range would correspond to BSM particles too heavy to be produced at the LHC. Additionally, from the perspective of the experimental physicist, the "bump hunt" of a narrow width particle is the well-established technique that led to discoveries of many particles.

In this dissertation data is compared to predictions from the Warped Extra Dimensions theory (WED) [41]. WED theory addresses the hierarchy problem by adding an additional fifth dimension to the conception of 4-dimensional (4D) space-time. In the framework introduced by Randall and Sundrum (RS) [40], 4D space is an Effective Field Theory (EFT) approximation of the higher dimensional space. The extra dimension exists between the gravity (Planck) and weak (TeV) flat 4D branes (see Fig. 2.2) and is called the "bulk". In the bulk, the strength of the gravitational interaction is not uniform. It depends on the coordinate in the 5th dimension and is characterised by the exponentially decaying function.

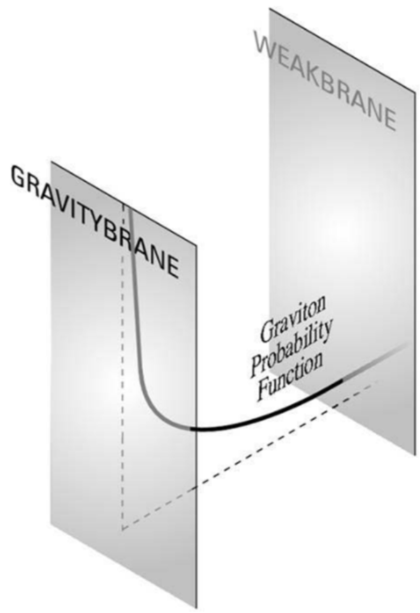


Figure 2.2: 5D space in the RS model [42].

The free parameters of the RS model are the brane separation factor k and the size of the compactified dimension r_c. Another name for the brane separation factor is the curvature factor. This factor is given by $k \approx \sqrt{\frac{\Lambda}{M_5^2}}$, where Λ is the ultraviolet cutoff of the theory and M_5 is the 5D Planck mass. The mass hierarchy of the SM particles between the Planck scale and the electroweak scale can be reproduced when free parameters k and r_c satisfy $k \cdot r_c \approx 11$. In this case, the RS model matches the observations: the Higgs boson being closer (in the geometric sense in the fifth dimenstion) to the TeV brane and light fermions being located near the Planck brane, see Fig. 2.2.

In the RS model under study, two new particles appear: a graviton and a radion.

When the bulk is compactified, the WED theory predicts the existence of the Kaluza-Klein (KK) [43] excitations of the gravitational field, with the zero-th KK mode being a graviton, the mediator of the gravitational force. The graviton (spin 2) is the first WED particle predicted by the RS model. The graviton can propagate freely in the full higher-dimensional space of the 5D bulk. The other RS particle is a radion (spin 0). Its existence is required to stabilise the size of the extra dimension. The WED space necessarily behaves in a quantum way, and, therefore, its size or length is subject to quantum fluctuations. The fluctuations of the length are parametrised by the radion, which is very similar to how the fluctuations of the EM field are parametrized by the photon. Goldberger and Wise [44] wrote down a potential for the radion and showed how *vev* of the radion sets the length of the extra dimension to its desired value. This is the mechanism for stabilizing the WED length. Without this procedure the radion would be massless and would mediate an infinite-range interaction, which is in conflict with cosmological observations.

Since LHC had provided us with no evidence of the SM particles interacting with the RS particles, the RS model considered in this thesis hypothesizes that SM particles are confined to branes. Another explanation of the lack of evidence of the RS particles at the LHC could be due to the fact that RS particles are too massive to be produced at the current LHC energies, but this argument is not addressed in this dissertation.

The theoretical arguments put forward by the authors [45] suggest the RS parameters k and $\bar{M}_{Pl}$ to be constrained by the following range of values: $0.01 \leq k/\bar{M}_{Pl} \leq 1$, where the parameter k is of the order of the Planck scale and $\bar{M}_{Pl} = \sqrt{\frac{M_5^3}{k} \cdot (1 - e^{-2\pi k r_c})}$ is a reduced 4D M_{Pl}. Considered in this measurement, the graviton and radion are RS particles with a KK state mass of the order of TeV [41].

With a part of the KK 5D wave function, often called a profile, expressed as $f_X^{(n)}(\phi)$, where n refers to the nth KK mode, the graviton can be decomposed as

$\sum_{n=0}^{\infty} h_{\mu\nu}^{(n)}(x_\mu) \cdot f_X^{(n)}(\phi)$. Its zero-th mode corresponds to the massless graviton and the first mode corresponds to the lightest KK graviton (later graviton) which has the mass of the O(TeV). The profiles for all the matter fields are described by a combination of Bessel and exponential functions [46–48]. The Lagrangian describing the interaction of the graviton with the SM fields is given then by

$$\mathcal{L}_{graviton} = -\frac{x_1 \tilde{k}}{m_G} h^{\mu\nu(1)} \times d_i T_{\mu\nu}^{(i)}, \tag{2.13}$$

where $x_1 = 3.83$ is the first zero of the Bessel function for a given profile, $\tilde{k} = k/\bar{M}_{Pl}$, $h^{\mu\nu}$ is a symmetric tensor describing the first KK graviton field, m_G is the mass of the graviton of the order of TeV, d_i is an integral of the profiles of the SM fields and KK graviton, and $T_{\mu\nu}^{(i)}$ is a 4D canonical energy-momentum tensor [49] for any SM field i. A free parameter $\tilde{k}$ varies from 0.01 to 1 when m_G is varied from 100 to 1500 GeV.

For the radion, the Lagrandian is given by:

$$\mathcal{L}_{radion} = -\frac{r}{\Lambda_R} \times a_i T_{\mu}^{\mu(i)}, \tag{2.14}$$

where r is a 5D radion field, Λ_R is the scale parameter proportional to $k \cdot \sqrt{(\frac{M_5}{k})^3}$, and a_i is the coupling of the radion to the SM field i. In the studied RS model the profiles of the graviton and radion arise naturally as being localised at the TeV brane for the coupling of a radion and a graviton to the massive SM fields to have the value of the order of one [50].

In the SM, HH production is dominated by two processes, which are shown using Feynman diagram representation in Fig.2.3: the "box" and the "triangle" diagrams. They interfere destructively and the total cross section is thus lowered. The total cross section made of box and triangle contributions is denoted as the “SM” and is shown in black color in Fig. 2.4 on the right. Additionally, this figure includes a BSM contribution - a non-linear ("nl") term $t\bar{t}HH$ that vanishes in SM, but may be present in BSM [51]. The results shown in Fig. 2.4 right, have been produced by theorists [52] for 100 TeV collider. However, the distributions of the double Higgs mass as well as amplitudes remain to a high degree unchanged between 13-14 and 100 TeV (see Fig. 2.4 on the left) - therefore, one assumes that amplitudes would look similarly for 13 TeV.

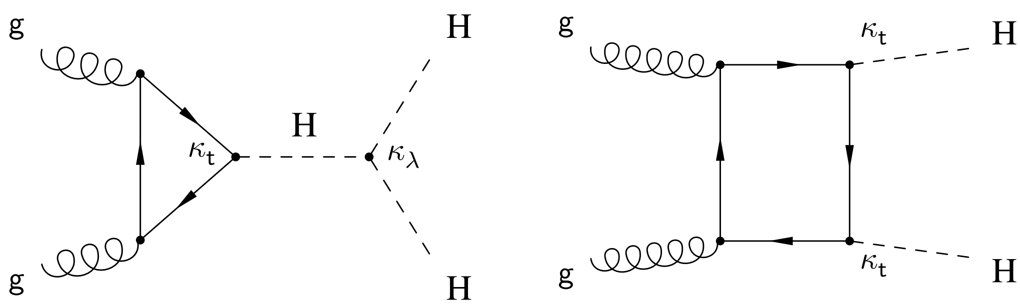


Figure 2.3: SM double Higgs boson production. Left: the triangle diagram with the virtual top quark loop. Right: the box diagram which dominates the overall HH production rate.

The box diagram dominates the double Higgs boson production and peaks near 400 GeV of the di-Higgs mass [52]. Even though the Fig. 2.4 on the right illustrates the SM double Higgs production, which is a non-resonant process, the amplitudes are not flat. Two factors contribute: an amplitude decreases with the COM and, at the same time, the kinematic turn on of the production of the di-Higgs system is always present.

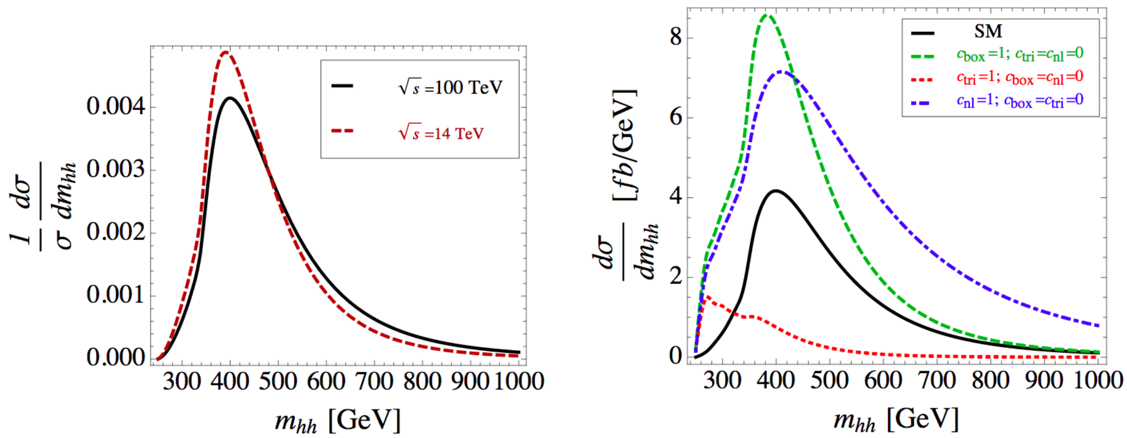


Figure 2.4: Left: comparison of the double Higgs boson mass distribution in the SM at the LO at 14 and 100 TeV center-of-mass energy. Right: the total SM HH cross section and the individual contributions [51]. Green refers to the SM box production, red refers to the SM triangle production, and blue refers to BSM non-linear $t\bar{t}HH$ production, not included in the total SM production in black [52].

In this measurement, the gravitons and radions in the search are expected to be produced by a BSM "contact interaction" Feynman diagram allowed by the WED scenario. This process is shown in Fig. 2.5. A graviton and a radion decays to a pair of Higgs bosons are thoroughly studied theoretically [51–53]. Experimental results produced by this measurement are compared to the theoretical predictions calculated for the WED model with the standard benchmark parameters $\tilde{k} = 0.1$ and $\Lambda_R = 3$ TeV [54, 55]. These reproduce SM observations and allow the production of the RS particles that can be observed at the LHC.

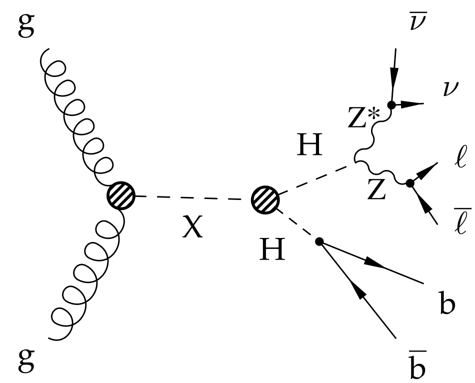


Figure 2.5: BSM Resonant double Higgs decay in the 2 b, 2 lepton, and 2 neutrino final state. X denotes either graviton or radion particles.

This thesis separately addresses both resonant graviton and radion decays into two SM Higgs bosons with the subsequent decays of one Higgs boson to a pair of b quarks, and the other Higgs boson to W or Z boson pairs. We select only leptonic W bosons decays. For Z boson decays, the chosen signature is characterised by the on-shell Z boson decaying into a pair of charged leptons and the off-shell Z boson decaying to neutrinos (see Fig. 2.5). The final state that this thesis focuses on consists of two b quarks, two charged leptons, and two neutrinos. Decays of the double Higgs system to this signature are observed on average in 2.8% of all di-Higgs decays.

To finish this chapter, it is instructive to show all the decay channels of the double Higgs system to the SM particles, which are summarised in the Fig. 2.6. Both the horizontal and the vertical axes show decays of a single Higgs boson to two SM particles. In this representation, each square on the plot specifies a branching fraction of one of the double Higgs boson decays, with the probability of the decay given by the color field map on the right axis. Our signature corresponds to 4 % of all $bbZZ$ decays, which are denoted on the map by the photo of the main $bbZZ$ analyser.

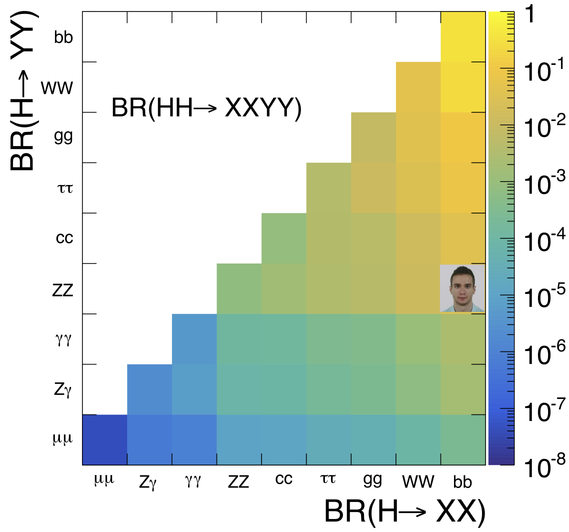


Figure 2.6: Double Higgs decay channels. The SM branching fractions are represented by the color palette. In this measurement $bbZZ$ decays are analysed, which are denoted on the map by the photo of the main $bbZZ$ analyser.

Chapter 3

The LHC and the CMS experiment

The CERN accelerator complex is a sequence of machines that produce and accelerate collections of 10^7 protons, called "bunches", to nearly light speed. In the Large Hadron Collider (LHC) the bunches collide at specific interaction points (IP), where the four main experiments are located: ALICE, ATLAS, CMS, and LHCb. This section will start with a discussion of the LHC machine and then describe the CMS detector.

3.1 The Large Hadron Collider

3.1.1 The history of the LHC

The story of the LHC begins in 1977, when the CERN director general Sir John Adams suggested that the tunnel of the future Large Electron-Positron Collider (LEP) could be reused to accommodate a future hadron collider of more than 3 TeV center-of-mass (COM) energies [56]. At the 1984 ECFA-CERN workshop on a "Large Hadron Collider in the LEP Tunnel" [57], the physics goals of the LHC were stated: confirming BEH mechanism, searching for the Higgs Boson, and exploring of the origin of the masses of W and Z bosons. The parameters of the proposed LHC were very ambitious: the center-of-mass collision energy of 10 to 20 TeV, and a target instantaneous luminosity of $10^{33-34}\frac{1}{\mathrm{cm}^2\mathrm{s}}$.

The Large Hadron Collider (LHC) is the most powerful (in terms of the COM energies) particle accelerator that has ever been built. It is located at the border of France and Switzerland at a depth from 50 to 175 m underground. The LHC ring is 26.7 km in circumference and is the final stage in a sequence of accelerators. In the following section, there is a discussion of the whole sequence of accelerators.

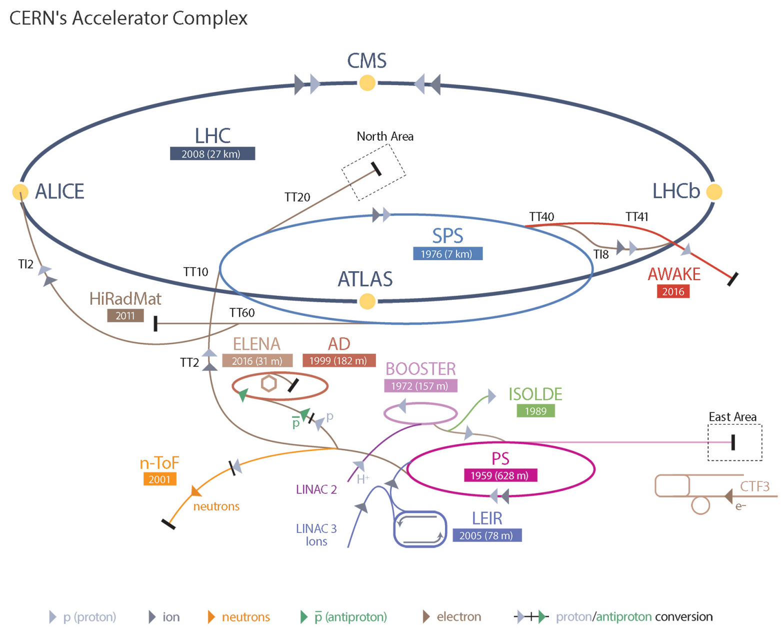


Figure 3.1: Schematic layout of the LHC.

3.1.2 The layout of the LHC

It is a complex process to start proton-proton collisions in the LHC at 13 TeV and, therefore, the process consists of several stages (see Fig. 3.1). The process begins with a bottle of hydrogen. The hydrogen atoms from the bottle are fed into the source chamber of the Linear Accelerator (Linac). In the chamber, the hydrogen is heated up to the plasma state until the electrons are stripped off of the hydrogen atoms. Electrons are then removed and the remaining protons are directed to the first acceleration stage which increases the energy of the protons to 50 MeV. After the Linac, the beam of protons is injected into the Proton Synchrotron Booster (PSB). The PSB contains four rings, which accelerate protons to 1.4 GeV. The third stage is the Proton Synchrotron (PS), which further splits the incoming beam into 72 bunches separated by 7.5 m. The energy of the protons is increased to 25 GeV. After that, the protons are sent to the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV and split into more bunches. The SPS then fills the LHC ring with two beams, each consisting of 2808 bunches of protons with nearly 10^{11} protons in total. It takes SPS about $O(10)$ minutes to fill each LHC ring with bunches. In the LHC, the two beams are circulating in opposite directions in two separate beam pipes. During standard data collecting ("data taking"), the beams circulate for $O(10)$ hours. This number has been found experimentally and leads to the highest possible amount of data per year.

3.1.3 LHC infrastructure

The equipment of the LHC tunnel serves several purposes, the main objective being to keep the colliding beams on a circular orbit, which requires a complex synchronization of bending dipole magnets, cooling systems, accelerating radio frequency cavities, and vacuum insulation systems.

3.1.3.1 Magnets

Most of the LHC's circumference is covered by 1232 superconducting magnets placed evenly around the tunnel to approximate a circular orbit. These are dipole magnets (see Fig. 3.2) that bend the beam and keep it on the circular orbit, hence why they are commonly called "Main Bends" (MB). The proven technology for magnets has existed since Tevatron and relied on NbTi superconductors. This technology also satisfied the LHC cost and performance requirements, thus, the decision was made to reuse the same choice of the alloy for the LHC superconducting dipole magnets that steer the proton beams. The dipoles need to produce the magnetic field of 8.3T. Each dipole is 16.5 m long (with ancillaries) and 570 mm in diameter and is placed inside of the dipole cryostat, called the "Helium bath".

The cryostat is a long cylindrical tube 914 mm in diameter, made of low-carbon steel, and is where the dipole mass is cooled down to 1.9 K. Even though the inner structure of such cryostats is very complex and includes both beam pipes, two sets of coils for two beam pipes, vacuum pipes etc., this compound object is normally called a dipole magnet. The name "dipole" is reserved for MBs as each beam pipe in the magnet consists of two "poles" that provide a vertical magnetic field similarly to a simple dipole system of magnets.

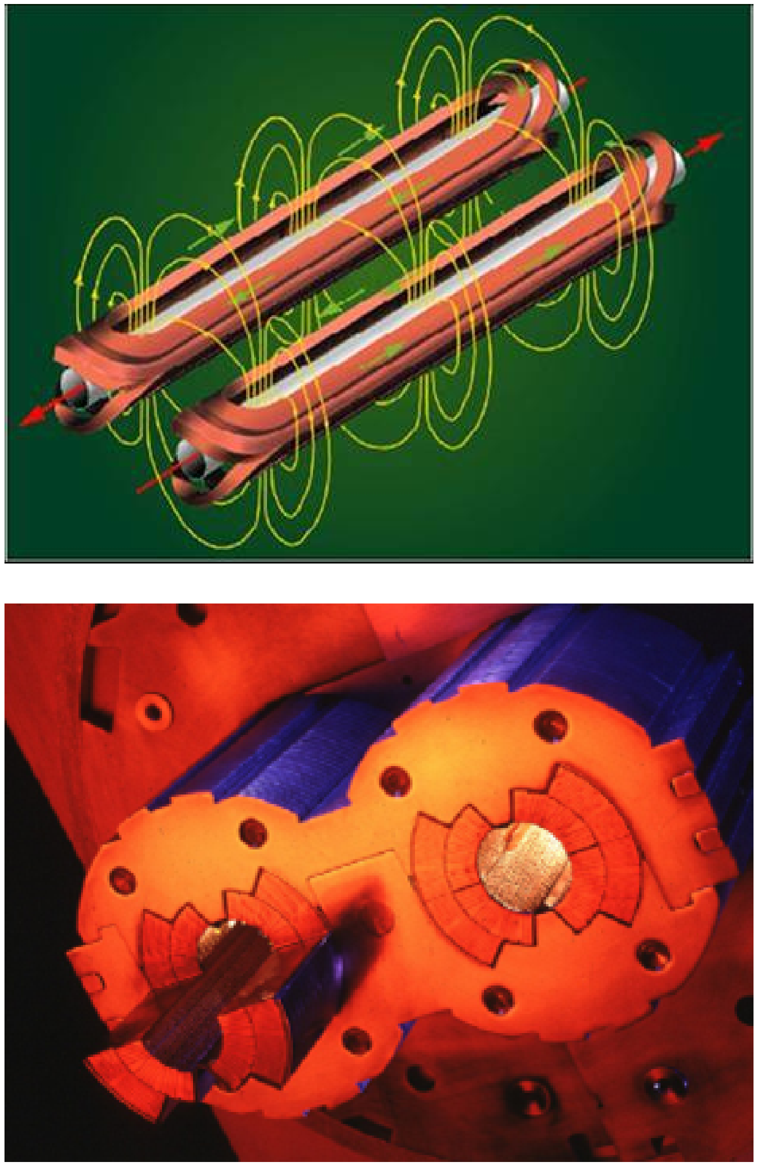


Figure 3.2: LHC dipole magnets. Top: two dipole coils and magnetic field lines. Bottom: two beam pipes with the coils inside of the dipole magnet. Each dipole magnet contains two magnetic configurations to steer the two proton beams travelling in opposite directions.

A dipole magnet must be curved to help a chain of dipoles complete 360 degrees. The curvature is 5.1 *mrad* per dipole, which is equivalent to a sagitta (explained later in this chapter) of about 9 mm, corresponding to a radius of curvature of 2812.36 m.

Other important magnets are quadrupoles. They are used to ensure proper beam dynamics. In total, 392 quadrupole magnets ranging from 5 to 7 metres in length are used to squeeze the beam in transverse directions (with respect to the z axis) and to keep it narrow during the run duration. Additional special quadrupole magnets (SQM) are installed right before the IPs to focus the beams even more. These increase the density of protons in the beam and guarantees the maximum luminosity (explained in 3.1.5). In addition, SQMs help to decrease the chance of the parasitic collisions when bunches from the same beam or bunches outside of the IP center interact (see Fig. 3.3). To further correct the beam path or orbit, about 5000 higher order correcting magnets are used, which are evenly spaced around the circular trajectory of the LHC.

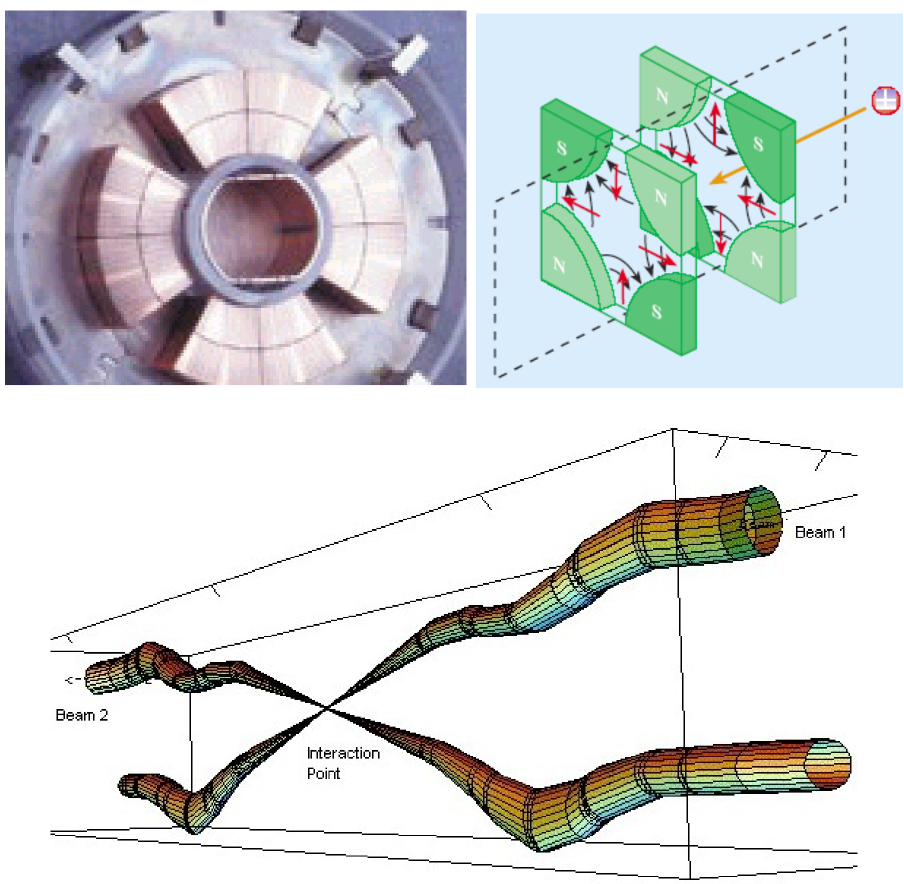


Figure 3.3: LHC quadrupoles. Top left: the coil of the quadrupole magnet. Top right: schematic view of the magnetic fields in the quadrupole. Bottom: two beams and the IP.

To power the LHC, 1612 electrical circuits are used. These circuits are mostly needed to power the dipole and quadrupole magnets, which is done in eight evenly-spaced locations of the LHC. A total of 3286 current leads are needed to connect all the circuits and power cables. More than a thousand of the leads operate between 600 A and 13 kA (see Fig. 3.4). The other leads operate in the range 60 to 120 A.

Figure 3.4: 13 kA high-temperature superconducting current lead.

3.1.3.2 Cooling System

To ensure that the dipoles are in the superconducting state, they have to be cooled to 1.9 K using superfluid helium-4. The cooling (cryogenic) system is needed to keep the superconducting LHC magnets at the appropriate temperature. The choice of cooling gas depends on the magnet type and location. This dictates the required range of temperatures, which differs from system to system by 75 K. The cryogenic system uses a layered design with the temperature becoming progressively colder moving from outside the dipoles to closer towards the beam pipe.

The "coldest" part of the cryogenic system is designed for the inner part of the dipoles. This system (see Fig. 3.5) must cool down 37 Mkg of the LHC magnets within 15 days to the required temperatures, which is done through a system of pipes that transports and directs the flow of the superfluid helium. The cryogenic system must also be able to deal with the fast increases of the pressure flow and flow surges, as it is crucial for the LHC operation to keep dipoles constantly cooled and at the superconducting state.

As the cost to cool the LHC equipment to 1.8-1.9 K temperatures is high, several temperature levels are employed (see Fig. 3.5):

- 50 to 75 K for the thermal shielding used in the dipoles,
- 20 to 300 K for upper ("warm") sections of the high-temperature superconducting current leads,
- 4.6 to 20 K for lower temperature interception,
- 4.5 K for radio frequency cavities and lower ("cold") sections of the high-temperature superconducting current leads,
- 4 K for the transportation system that directs the 1.8 K helium to dipoles,
- 1.9 K for helium in the superfluid state to cool magnet masses.

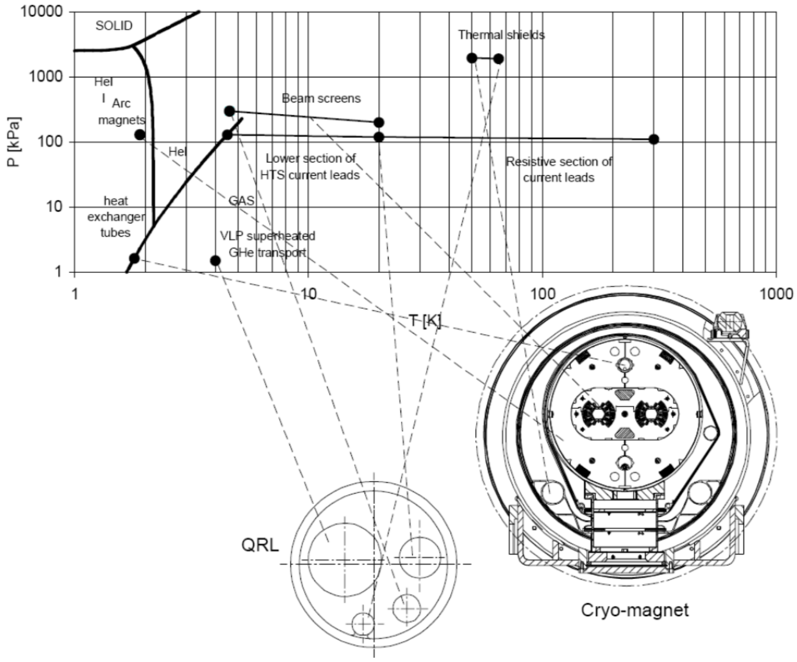
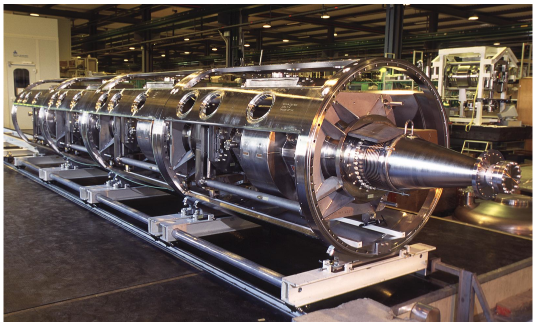


Figure 3.5: LHC cryogenic states and the temperature scale.

3.1.3.3 Radio Frequency Cavities

Proton bunches need to be ramped to 6.5 TeV energies. To achieve this 13 TeV COM energy, eight superconducting radio-frequency cavities (RFC) are used per beam. They are located in front of the IPs of four experiments. Electromagnetic waves of 400 MHz with a peak field strength of 5.5 MV/m adjust the speed of protons in bunches. Each RFC (see Fig. 3.6) increases the energy of protons by 60 keV per revolution and it takes $O(20)$ minutes to reach 6.5 TeV beam energy. The RFC frequencies are gradually increased by 1 kHz to match the acceleration of the protons in the bunch as they gain more energy. When the ramp is completed, the RFCs are used to compensate for small energy losses due to the synchrotron radiation (7 keV per revolution).



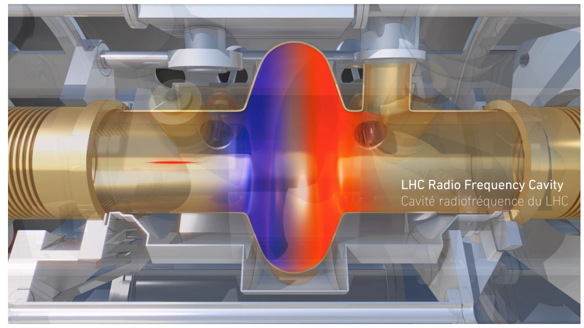


Figure 3.6: LHC RF cavities. Left: a cryomodule with four RF cavities. Right: a schematic drawing of a single RF cavity. The colour field is used to denote positive (red) and negative (blue) polarities. A narrow beam traversing the cavity is coming from the left and is shown in red.

3.1.3.4 Vacuum System

The work of the LHC depends on three vacuum systems [58]. Without them, the dipoles would not achieve the superconducting state, the beams would not be able to circulate, and no stable collisions would occur. With a total of 104 kilometres of vacuum pipes, the LHC owns the largest vacuum system in the world. The main types of vacuum systems are:

- insulation vacuum for cryomagnets,
- insulation vacuum for the helium distribution line,
- beam vacuum.

The insulation vacuum is needed to ensure the operations at both the low temperatures of the magnets and the room temperatures in the tunnel. The insulation vacuum of 10^{-6} mbar is used for a total of 15000 cubic metres. In building this vacuum system, the LHC used 250,000 welded joints and 18,000 vacuum seals.

The vacuum for the helium distribution lines is needed to protect the flow of the helium-4 from the heat. This helium flow is used to cool down the dipole mass.

Cryogenic distribution lines (QRL) of 3.3 km each are connected to eight cryogenic plants that pump the helium-4 into the LHC. The vacuum in these systems is at $10^{-7} - 10^{-10}$ mbar level.

For the beam pipes, the LHC uses ultra-high vacuum of 10^{-10} mbar at a cryogenic temperature of 5 K. The vacuum is getting progressively closer to 10^{-11} mbar near the IPs, because these locations are where collisions take place and any additional gas is highly undesirable. This vacuum is the emptiest space in the Solar System. This ultra-high vacuum is needed to reduce the beam degradation due to the beam-gas interactions in the pipe and parasitic collisions of the proton bunches with the collimators near the IPs.

The vacuum systems are affected by the heat produced from the synchrotron radiation that is emitted by the proton beams when they are bent. To reduce the amount of this heat and to narrow down the beam size in the transverse direction when the beam widens, the LHC uses "beam screens", which operate at between 5 and 20 K.

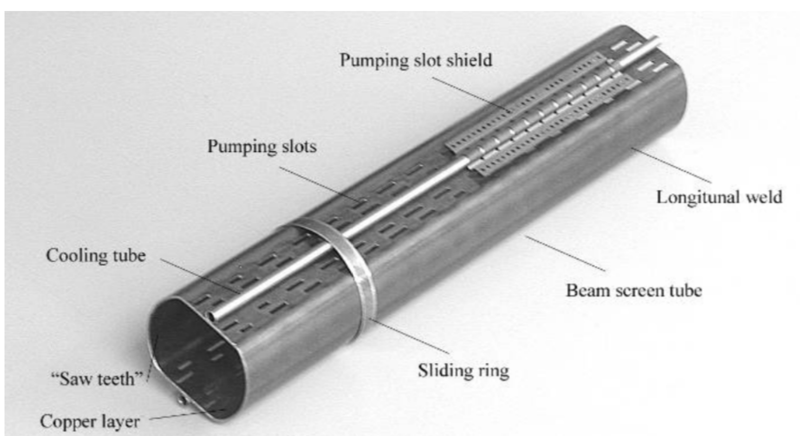


Figure 3.7: Beam screen.

The beam screens are necessary to reduce the number of protons scattering on the residual gas of the beam pipes, which could lead to a magnet quench and even interrupt the machine operation.

The table below summarises the main heat sources that degrade the vacuum quality in the beam pipe, where the vacuum must exist at 1.9 K:

- synchrotron radiation (0.2 W/m per beam),
- energy loss by nuclear scattering (30 mW/m per beam),
- image currents (0.2 W/m per beam),
- electron cloud related effects (vary).

Having discussed the LHC collider, the following section explores the Compact Muon Solenoid (CMS) detector, which was used to collect the data analysed in this thesis.

3.1.4 LHC operations

The first LHC budget plan was finalised in 1996, with the final cost being approved just a few years later. The first proton beam entered the LHC ring in 2008. However, an incident delayed the original LHC plans. It was caused by the mechanical damage of the tunnel equipment due to the release of the helium. Thus, the real data taking period (called LHC Run 1) began only in 2010, and lasted for two years, with 7-8 TeV COM energies used. The recorded dataset contained enough Higgs bosons to claim a discovery of this rarely produced particle. After this achievement, the LHC was closed in 2012 for the first long shutdown (LS1). During this time, necessary upgrades of the main detectors and the LHC were performed. This was an unavoidable and essential step to prepare the LHC for the more challenging environment of COM energies increased to 13 TeV.

If denote the area of 10^{-28} m^2 as barn (b), then the femtobarn (fb) equals to 10^{-43} m^2. When HEP physicists talk about inverse femtobarns, they mean collisions per femtobarn of beam cross-sectional area. Thus, the inverse femtobarn unit can be thought of as the effectiveness of the particle accelerator. In terms of these new units, the LHC can theoretically produce $80 - 120/fb$ (inverse femtobarns) of data a year. In practice, these numbers were lower because the LHC operated at the revolution frequency below the nominal, used fewer proton bunches in the beam, etc. These resulted in lower than the expected instantaneous luminosity, an important term in collider physics that will be explained in the next section.

The LHC Run 2 started in 2015 and the CMS collected 4.2 fb^{-1} of data that year. Over the course of the 2016 data collection, an integrated luminosity of 35.9 fb^{-1} was recorded. This luminosity is the amount of data that was collected by the CMS detector and later approved by the CMS physics coordination for the use in the physics analyses. Note that the CMS collects as much data as possible, but not all data have good quality for the offline physics analyses, so the amount of the approved data is always less than the collected one. The data set of proton-proton collisions collected in 2016 at 13 TeV COM energy will be used in this thesis to analyse double Higgs boson decays. Together with the 2017 and 2018 data taking, almost 150 fb^{-1} have been delivered and recorded by the CMS detector during the whole Run 2 period of four years (2015-2018).

At the moment of writing this thesis, the LHC has entered LS2. Data taking will resume in 2020 and proton-proton collisions will continue for three years with the integrated luminosity expected to be equal to nearly 300 fb^{-1}. This will conclude the LHC Phase 1 program.

The new upgraded LHC, the High-Luminosity LHC (LHC) or the Phase 2, will start operations in 2026 and run until 2035. The COM energy will be increased to 14

TeV (may happen during Run 3) and physicists expect to record an unprecedented dataset of 3000 fb^{-1}.

3.1.5 Luminosity

The instantaneous luminosity $\mathcal{L}$ is the parameter which relates the cross section σ of the process to the number of events N_{events} produced per unit time during the interaction: $N_{events}/dt = \mathcal{L}\sigma$. Luminosity is the parameter controlled by the machine and can be written as:

$\mathcal{L} = \frac{N^2 n_b f_{rev}}{4\pi\sigma_x\sigma_y}$

where N_b is the number of particles in the colliding bunch, n_b is the number of colliding bunches in the beam, f_{rev} is the revolution frequency of the beam, a σ_x and σ_y are the standard deviations of the beam density profile (BDP) in the transverse plane, where it is assumed that the BDP of both beams can be described by a Gaussian distribution.

To maximise the amount of collected data, the luminosity parameter should be as high as possible. It is worth noting that the luminosity is not constant and decays with time due to the degradation of the initial circulating beams. The theoretical decay time (the time to reach $1/e$ level) is approximately 29 h. In practice, taking into account the decrease of protons in the bunch due to collisions, contributions from the intrabeam scattering, scattering on the residual gas, etc., the real luminosity lifetime is about 15 h.

A useful variation of the luminosity parameter is the total integrated luminosity. This is the number normally quoted for the dataset collected over the period T:

$L = \int_0^T \mathcal{L}dt.$

In collider physics the "beam dump" is a process of burning off exhausted low luminosity beams by intentionally directing them towards a target made of concrete and steel. The time from the start of the collisions to the beam dump is usually called the “run” or “fill”.

The amount of data delivered by the LHC can be calculated for a single run period of $O(10)$ h. Performing the integration, this is obtained:

$$L = \mathcal{L}_0 \tau_{\mathcal{L}} \left[1 - e^{\frac{-\tau_{run}}{\tau_{\mathcal{L}}}} \right],$$

where $\mathcal{L}_0$ is the initial peak instantaneous luminosity at the start of the run, τ_{run} is the total duration of a run, and $\tau_{\mathcal{L}}$ is the luminosity lifetime. The optimum run time is 12 hours. Between the runs, the LHC center needs to dump the old beams, fill the rings with the new beams, and increase ("ramp") the energy of new beams to 13 TeV. After that a new run can be started. This restarting process normally takes two to six hours.

3.2 The CMS experiment

The CMS detector (later the CMS for brevity) is a multi-purpose particle detector built to study a variety of complex particle interactions produced by the LHC. The CMS is located in the underground cavern at the Point 5, which is one of the four main IPs of the LHC. The CMS detector, with additional computing infrastructure, is able to detect the produced particles, to measure their main physics parameters, and to send the related data to computing data centers for persistent storage.

The CMS detector has a cylindrical shape and consists of a central cylinder (the "barrel") and two forward (the "endcaps") sections (see Fig. 3.8). The CMS detector is the heaviest detector ever built with the mass of nearly 12500 tons. The mass is explained by the amount of the used superconducting metal, which serves as the magnet, as well as by all the iron on the outside used for the magnetic flux return and shielding. The CMS is 21.6 m long and 14.6 m high. The CMS has an onion-like structure of concentric layers made up of detectors around the IP. Additionally, the outer part has a large superconducting solenoid to produce a homogeneous magnetic field of 3.8 T inside the detector.

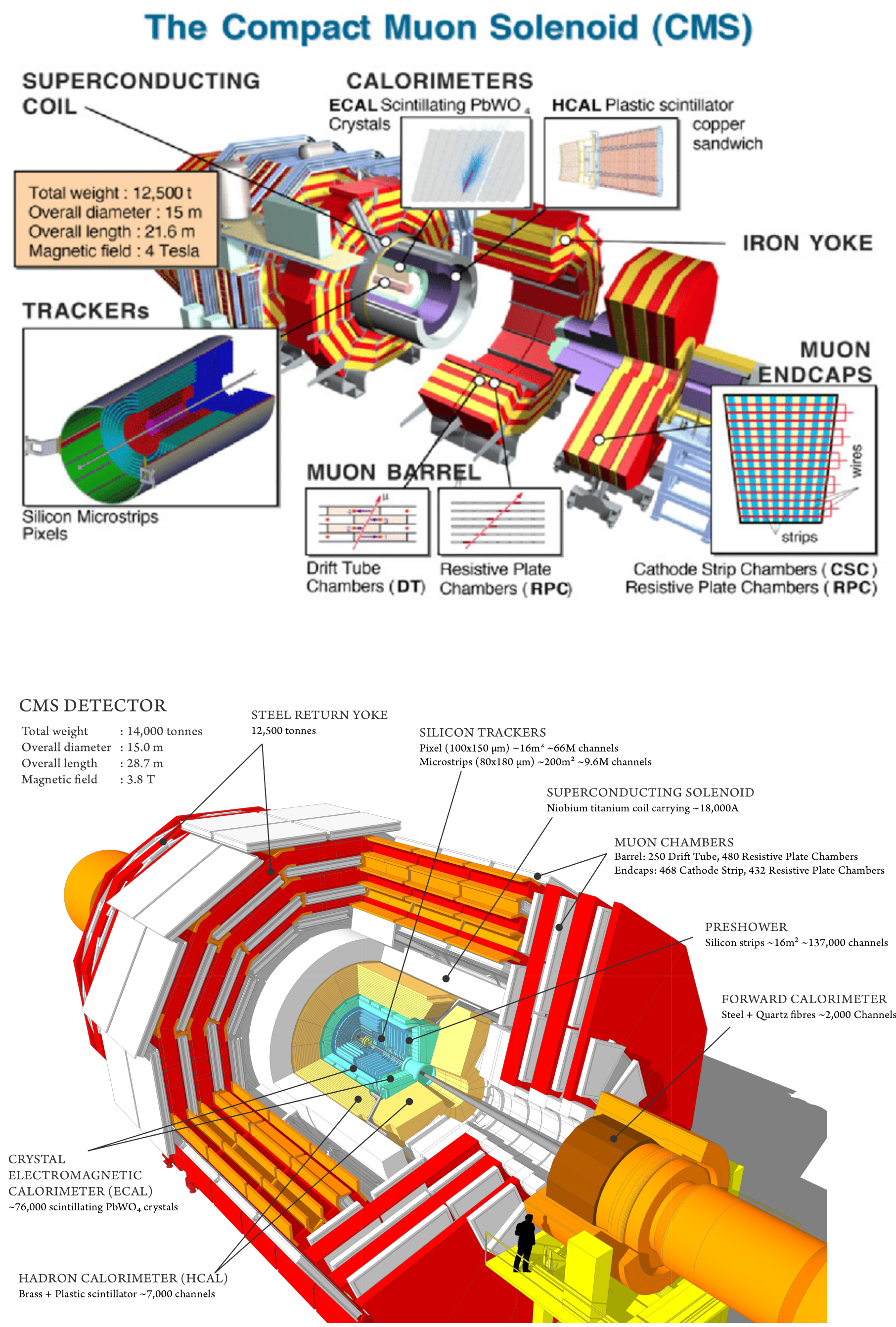


Figure 3.8: CMS experiment with the main sub-detectors.

All sub-detectors can be categorized into trackers and calorimeters [59]. As the particle passes through the material of the tracker, it leaves a path of the emerging particle, called a “track”. Trackers focus on the direction and the track curvature of the charged particles. Tracking information allows the determination of the particle's momentum.

There are two trackers in the CMS detector: an inner tracking system that encloses the IP and an outer tracking system that is located outside of the solenoid magnet. The first system contains the pixel and the strip trackers. The second tracking system, embedded within the steel yoke of the magnet, is dedicated to the detection of muons, and is usually called a muon tracker or a muon system.

The magnet yoke is made of five barrel wheels. Such an arrangement saves the CMS some space and also is used for the magnetic flux return. Additionally, it serves as a support for the embedded muon system, which is located outside of the ECAL and HCAL systems (described later). Prompt muons from the heavy particle decays are energetic enough to traverse the ECAL and leave the detector because they are minimum ionizing particles. The muon system-magnet yoke structure provides a return field of the magnet of about 2 T and is used to measure the momentum of muons. The complex magnetic field causes the muons trajectories to be bent in opposite directions in the inner tracker in contrast to the outer tracker. This important feature of the CMS detector is depicted in the CMS logo (see Fig. 3.9).

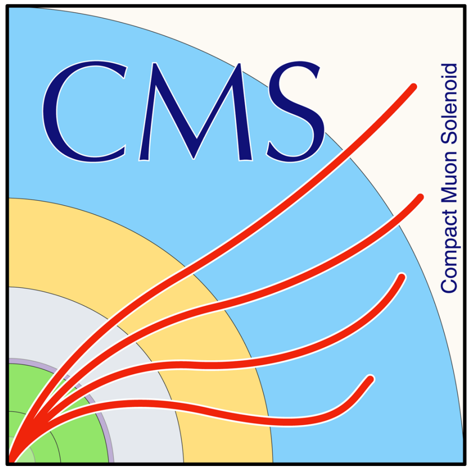


Figure 3.9: The logo of the CMS experiment that is showing curved trajectories of the emerging muons.

The CMS has two calorimeters: the electromagnetic and the hadronic. They both rely on high density materials either to sample or to contain almost all the energy of the incoming particles with their secondary interaction products. However, these two systems focus on two different sets of particles. As will be discussed later, the electromagnetic calorimeter (ECAL) is dedicated to measuring the energy of photons and electrons, while the hadronic calorimeter is targeting the measurement of the energy of hadrons.

The rate of the incoming data at the LHC is 40 MHz, which is related to the 25 ns bunch spacing. This corresponds to almost 70 TB produced every second. It is impossible to store that much data, and, most importantly, most of the information in this data is not pertinent for future physics analyses (the Physics program of the CMS is well defined and specific). To reduce the data rate, the CMS uses a highly efficient system of triggers. The first one, the Level-1 (L1) trigger, reduces the readout rate from 40 MHz to 100 kHz. The subsequent High-Level Trigger (HLT) further decreases the rate to 1 kHz. With the help of the trigger system, the original 70 TB per second rate is transformed into manageable 1 GB per second that is stored for offline analysis use.

3.2.1 The CMS coordinate system

The CMS uses a right-handed Cartesian coordinate system to define the axes of the colliding beams (see Fig. 3.10). The center is located at the IP and the x axis points to the center of the LHC ring. The y axis points upwards, and the z axis points along the proton beam direction. This is defined for one of the proton beam directions, and is opposite for the other one. As the CMS detector has a cylindrical shape, the polar system is used in the x-y plane: a standard set of the azimuthal angle φ and the radial coordinate r. A polar angle θ is defined in the r-z plane and an angular variable η (called pseudorapidity), which is widely used in this thesis, is defined as $\eta = \ln \tan(\theta/2) = \ln(\frac{|\vec{p}|+p_z}{|\vec{p}|-p_z})$. Another useful quantity in the collider physics - the rapidity - is given by $y = 1/2 \ln(\frac{E+p_z}{E-p_z})$. Rapidity is a function of the energy E and longitudinal momentum p_z of the particle (the projection of $\vec{p}$ on the z axis). Note that η converges to y when the mass is negligible and the particle travels with the speed close to the speed of light. Most angular variables that are used currently in modern high-energy physics (HEP) at the LHC are defined in terms of η and φ. For example, a relative distance ΔR in $\eta - \phi$ plane between two particles is given by: $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}$, with $\Delta\eta$ and $\Delta\varphi$ being the absolute values of the relative differences in $\eta's$ and $\varphi's$ of two particles.

Another useful quantity is the projection of the momentum of a particle on the transverse plane, which is called "transverse momentum" p_T. This variation of the momentum is independent of the z axis, and, thus, of the Lorentz boost. Similarly, the transverse energy of a particle is defined as $E_T = \sqrt{m^2 + p_T^2}$.

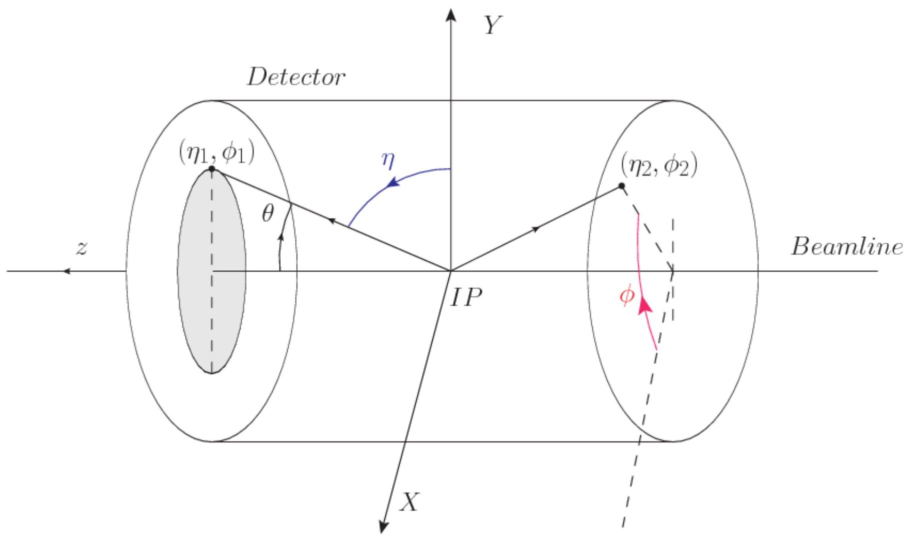


Figure 3.10: Coordinate system of the CMS detector [35]. Two particles (1 and 2) are shown with the corresponding angular variables ($\Delta\eta_1$, $\Delta\varphi_1$) for the first and ($\Delta\eta_2$,$\Delta\varphi_2$) for the second particle respectively.

3.2.2 The Inner Tracker

The inner tracker [60] (see Fig. 3.11) is the closest subdetector to the IP. Using the tracker, the experiment measures the trajectories of charged particles and reconstructs decay vertices. Because this system is constantly under the radiation coming from the interactions with the particle flux of nearly 100 MHz/cm at r = 4 cm, the design of the tracker is focused on two main requirements: high granularity for precise determination of the vertices and tracks, and robustness against the radiation-hard environment with the operational time of at least ten years. As a solution to both challenges, the CMS relies on the silicon technology that provides the tracker with the large surface of thin but highly granular active detectors. The tracking system has a diameter of 2.4 m and a length of 5.4 m covering the detector space of $|\eta| < 2.5$.

The innermost part of the tracker - the pixel detector ("pixel")- consists of three layers in the barrel at the radii of 4.4 cm, 7.3 cm, and 10.2 cm respectively. The pixel also has two active disks in forward regions. They are positioned 34.5 and 46.6 cm away from the IP. The pixel is made of 1440 modules which contain 66 million pixel cells. Each cell is 100 by 150 μm with 285 μm thickness, which allows the determination of "hit" positions (the passage of the particle through the pixel cells) in two directions: z-φ in the barrel and r-φ in the endcaps.

The spatial resolution of each pixel is about 10 μm in the r-φ plane and 20 μm along the z direction. The spatial information that comes from the tracker is used to determine the main interaction point of the hard scattering (the primary vertex) and also additional interaction vertices (the pileup). The tracker also helps to reconstruct the displaced vertices (the secondary vertices) of the particles that decay relatively quickly, e.g., b-jets, which will be discussed later in this chapter.

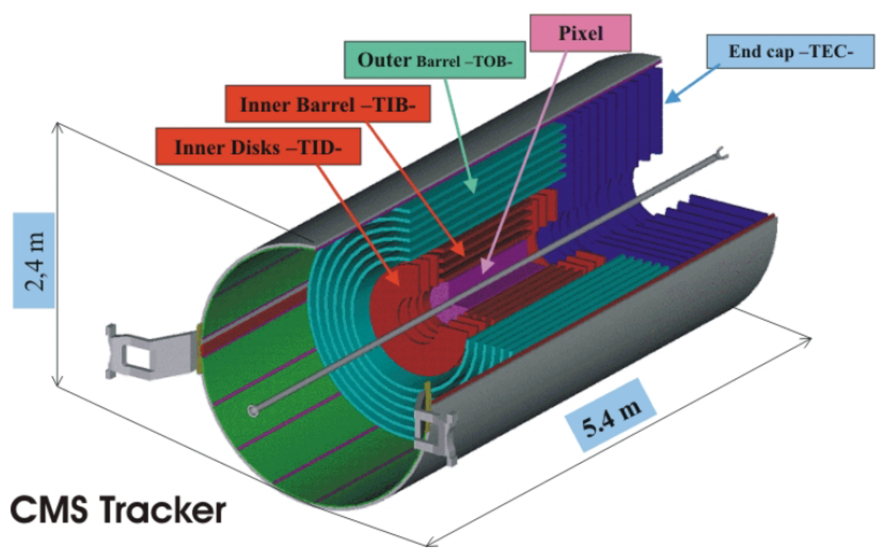


Figure 3.11: The inner tracker. The pixel and strip detectors are shown.

The outer part of the inner tracker is the strip tracker. It contains several subsystems and is made of almost 9.3 million strips arranged in different configurations in 15148 modules. The first subsystem is the tracker inner barrel (TIB), which consists of the four barrel layers of strip modules. The second subsystem is the tracker inner disks (TIDs), which is made of three disks of strip modules. At a radius of about 60 cm, the tracker outer barrel (TOB) starts. TOB is made of six layers of strips. Finally, to cover high η regions, the tracker endcaps (TECs) are used, which are made of two sets of nine disks of strips.

Each strip is O(20) cm long. Its thickness varies from 320 μm for TIB and TID, to 320 μm - 500 μm for TOB and TEC, respectively. Also the width changes from 80 μm - 141 μm for TIB and TID, to 97 μm - 184 μm for TOB and TEC, correspondingly. The resolution on the single point in the radial direction is 20 - 50 μm, and in the z direction it varies from 200 to 500 μm, depending on the value of r.

All subsystems of the inner tracker have to be cooled down to about -20°C. This requirement is needed to minimize the damage to the tracker caused by the radiation from the collisions and to reduce overheating of the electronics. The material of the inner tracker has 0.4 to 1.8 radiation lengths (X_0), which corresponds to 0.1 to 0.5 nuclear interaction lengths (λ_i). Numbers vary with the η. The inner tracker is surrounded by the ECAL described in the following subsection.

3.2.3 The ECAL

The inner tracker and the ECAL provide the detector with complementary measurements. The tracker focuses on the direction and the momentum of the particle and identifies only charged particles. The ECAL [61] (see Fig. 3.12), on the other hand, determines the energy of the particles and detects all particles that interact electromagnetically, including photons and neutral pions. However, the ECAL is primarily designed to make precise measurements of the energy of electrons and photons by selecting the proper type of crystals and their length to contain enough X_0's to stop incoming electrons and photons.

The ECAL is a highly granular detector that relies on the lead tungstate crystal ($PbWO_4$) technology. Electrons and photon passing through the crystal interact with its material and their energy is converted into the produced electromagnetic shower. The scintillation light, produced by $PbWO_4$ scintillators, is further read out by the electronics, more details in [62]. The $PbWO_4$ crystals have a high density (8.28g/cm^3), a small radiation length (X_0 = 0.89 cm), a short Moliere radius (R = 2.2 cm), and a fast response (80% of its scintillation light is produced within 25 ns). These characteristics make $PbWO_4$ crystals ideal candidates for the ECAL, because they guarantee an excellent containment of the electromagnetic shower within the crystals.

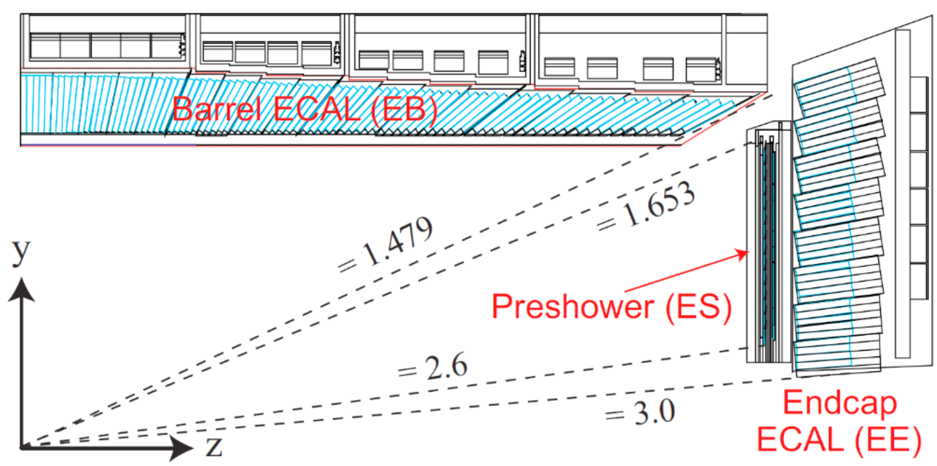


Figure 3.12: The ECAL and the Preshower detectors.

The ECAL has a barrel part (EB), covering the $|\eta| < 1.479$, and two endcaps (EE) covering $1.479 < |\eta| < 3.0$. In the barrel, the ECAL is made of 61 200 crystals. Each crystal is 22 by 22 mm with a length of 23 cm. The endcaps of the ECAL have 7324 crystals. There, each crystal is 28.62 by 28.62 mm with a length of 22 cm. The crystals' layout follows a quasi-geometric projection where the axes of the crystals are slightly tilted to ensure that particle trajectories are never aligned with intercrystal cracks. This layout is optimized for the best particle shower containment with respect to the position of the interaction point.

The resolution of the ECAL is a function of the energy of the incident particle E and can be decomposed into three terms. The first term is a stochastic term that is inversely proportional to the square root of the number n of scintillation photons produced in the interaction. In the main formula n is replaced by E, since n is proportional to E. The second term is a "noise" term that describes the noise in the detector. The dependence of the resolution on the energy component in the noise term is inverse. The third term is related to detector imperfections and is represented by a constant C. The final dependence of the ECAL energy resolution σ on the particle energy E is given by:

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

From dedicated calibration studies, the parameters in the formula above are found to be equal to S = 2.8%, N = 12%, and C = 0.3%. As a "standard" procedure, the CMS often optimizes the performance of the subdetectors for 45 GeV electrons, since they correspond to a classical Drell-Yan decay of Z boson to two electrons. In this case, a typical energy resolution for 45 GeV electrons is about 2% in EB and 2-5% for EE. The constant terms dominate the resolution near the Z peak (91 GeV). The ECAL is operated at a temperature of 18°C and the "active thickness" of the ECAL material corresponds to 25 X_0.

An additional subdetector, called the Preshower, is installed right in front of the EE and covers $1.653 < |\eta| < 2.6$. The Preshower is designed to improve the discrimination of photons from diphoton decays of neutral pions $\pi^0 \to \gamma\gamma$. This is a sampling calorimeter in which the material that produces the particle shower is distinct from the material that measures the deposited energy. Typically, the two materials alternate. The Preshower has two lead layers which launch the electromagnetic showers. This samples the energy of the particles traversing the Preshower material. After these layers, 2 mm-wide silicon strips are placed, and these measure the deposited energy and transverse profile of the shower shape initiated by the lead layers. The thickness of the Preshower material corresponds to 3 X_0.

3.2.4 The HCAL

Hadrons normally go through the ECAL layers without being stopped since ECAL crystals are optimised to contain mostly electron and photon showers. To absorb these particles, the HCAL [63] (see Fig. 3.13) is placed around the ECAL. The HCAL focuses on particles that hadronize. This is the process of the formation of hadrons out of quarks and gluons. The HCAL detects charged and neutral hadrons such as pions, kaons, protons, and neutrons. Hadrons also produce collimated streams of secondary particles (jets) and these jets are identified by the HCAL. Additionally, the HCAL is used to measure the transverse energy of neutrinos indirectly, through the momentum imbalance technique, which will be discussed later in this chapter.

The HCAL is split into the HCAL barrel (HB) and the HCAL endcap (HE) sections. They cover $|\eta|$ <1.3 and 1.3 < $|\eta|$ <3.0 respectively. The HB and HE are sampling calorimeters. They are made of a brass absorber and of active plastic scintillating tiles. The brass plates in HB have thickness of 56.5 mm and in HE the thickness is increased to 79 mm. The absorber material corresponds to 5.82 λ_I at $\eta = 0$ and almost 10 λ_I at $0 < |\eta| < 1.3$.

The gaps in the absorber of the HCAL are filled with an active medium of 70000 plastic scintillator tiles. The scintillation light is guided by wavelength shifting fibers

(WLSs) to hybrid photodiodes (HPDs). The scintillator is quite fast with the 68 % of the light being produced within 25 ns.

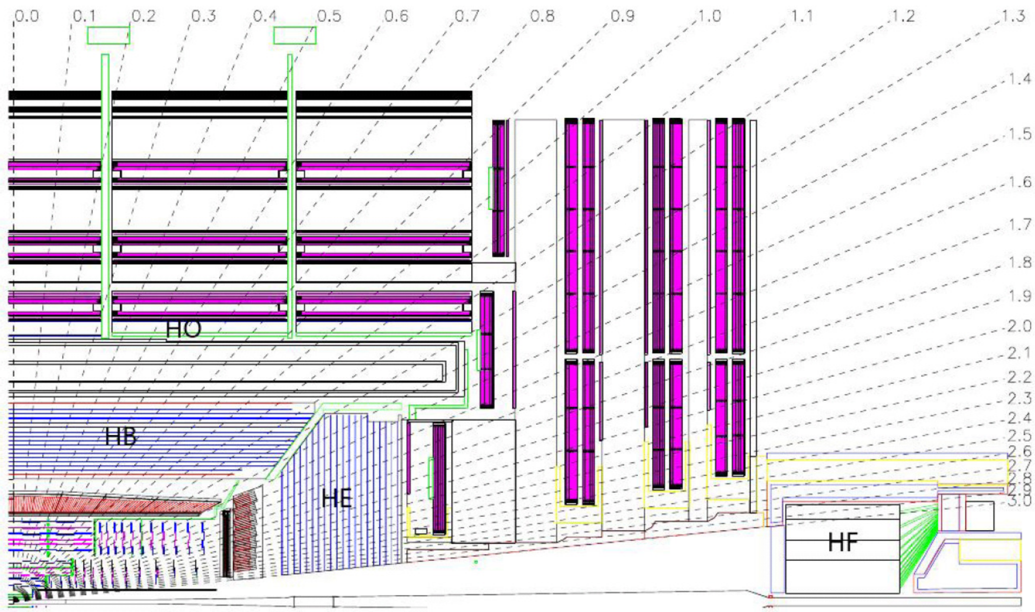


Figure 3.13: The HCAL with the η coverage map.

The CMS also has an outer calorimeter (HO) placed above the HB outside the solenoid. The HO is called a tail catcher system and increases the total calorimeter thickness to 11.8 λ_I in the barrel, with the magnet coil working as an extra absorption layer. The HO consists of five rings of scintillator tiles. A supplementary 19.5 cm-thick iron plate and a second layer of sensitive material are placed around $\eta = 0$ to enhance the absorber depth in the HO.

In the forward directions, two forward calorimeters (HF) extend the coverage to $|\eta| = 5.2$. The HF is composed of steel absorbers and quartz fibres that produce Cherenkov light when the particle in the material travels faster than the light in that medium. The light is then collected by photomultiplier tubes (PMTs).

Since the HCAL is located between the ECAL and the internal surface of the solenoid, the space allocated for the HCAL has not been enough for the HCAL to fully absorb the hadronic showers. This imperfect containment of the hadronic shower limits the performance of the HCAL. Comparing with the formula 3.1, from the calibration using the single pions the values are given by: by S = 115 %, N = 52 %, and C = 5.5 % [64]. This underperformance of the HCAL is perhaps the greatest weakness of the entire detector, which is mitigated by the strengths of the other subdetectors. This motivated the CMS to employ the particle flow algorithm (described in the next chapter), which uses the information from all subdetectors.

3.2.5 The Superconding Solenoid

The NbTi superconducting solenoid (see Fig. 3.14) of 6 m in diameter is the core of the CMS experiment. The magnet operates at a temperature of 4.5K. The bulk of the CMS detector weight (90 %) comes from the magnet steel return yoke and structural supports, which together weigh 12500 tons.

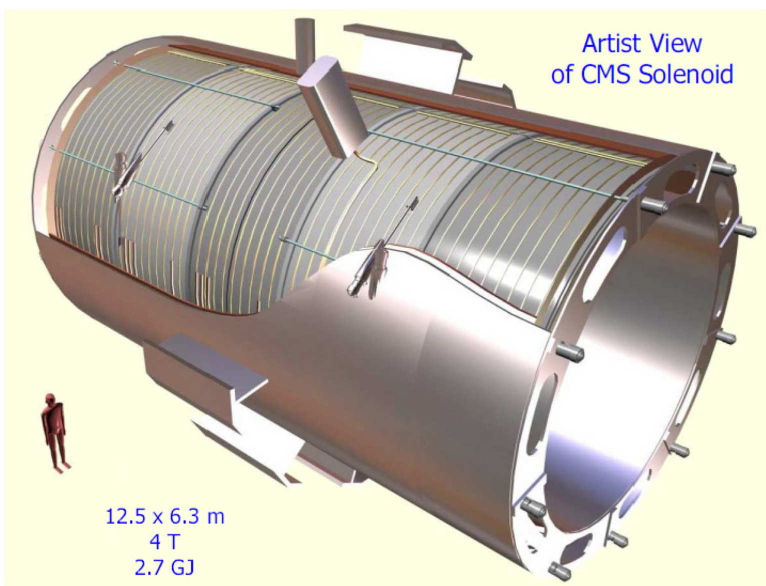


Figure 3.14: The CMS superconducting solenoid. The person on the left is shown to emphasise the size of the magnet.

The solenoid is a central part in the CMS detector design. The idea was to have a uniform magnetic field capable of bending the trajectories of charged particles as they traverse the detector. When a low energy particle is produced ($O(0.5)$ GeV), it curls back and has a helical path that is fully contained within the detector. However, when a highly energised particle is produced, the trajectory is seen as a "straight" incomplete arc. Both situations lead to an imperfect measurement of momentum.

The primary physics quantities measured by the tracking system are presumed to be Gaussian distributed; but there are exceptions, e.g., momentum of the particle is not Gaussian distributed. However, the sagitta, which is inversely proportional to momentum, is Gaussian distributed. When a particle in the magnetic field passes through the material of the detector, the path deviates from the ideal circular line due to multiple scattering effects. The sagitta term is used to quantify the depth of the circular arc and is equal to the distance from the center of the arc to the center of its base. Since the sagitta is following a Gaussian distribution, it may be approximated by simpler expressions in many calculations of the momentum resolution. Hence, the sagitta is widely used in particle physics.

The magnetic field strength B and the length of the track L are dictated by the design of the detector. Because the momentum resolution is given by $\sigma_p/p^2 \approx \sigma_x/BL^2$ (see [59]) and improves linearly with magnetic field B, the CMS decided to invest much of the detector space and budget in the magnet, since the tracking material is more expensive. For a track of the length of O(1) m in the magnetic field of O(3) T, the sagitta is equal to 1 mm, which can be measured with the precision of O(10) μm.

3.2.6 The Muon Tracker

Many physics analyses in the CMS rely on precise measurements of the muons in the detector. Although muons are detected by the inner tracker, that information cannot be used by the trigger (which will be discussed in the following subsection). Therefore, the CMS has an outer tracker or muon tracker [65] (see Fig. 3.15) located outside the calorimeters and the solenoid. Typical high energy muons that are produced in collisions at the LHC traverse the detector material with the minimal energy losses since their EM interaction cross section is much smaller than for electrons. To measure the energy of muons, the CMS uses the muon tracker, which relies on various gaseous detector technologies. The muon tracker is inserted into the gaps of the flux-return yoke. Tracks in the muon system are used to reconstruct standalone muons (see section 4) and, in combination with the inner tracker, to reconstruct the global muons, more in section 4.

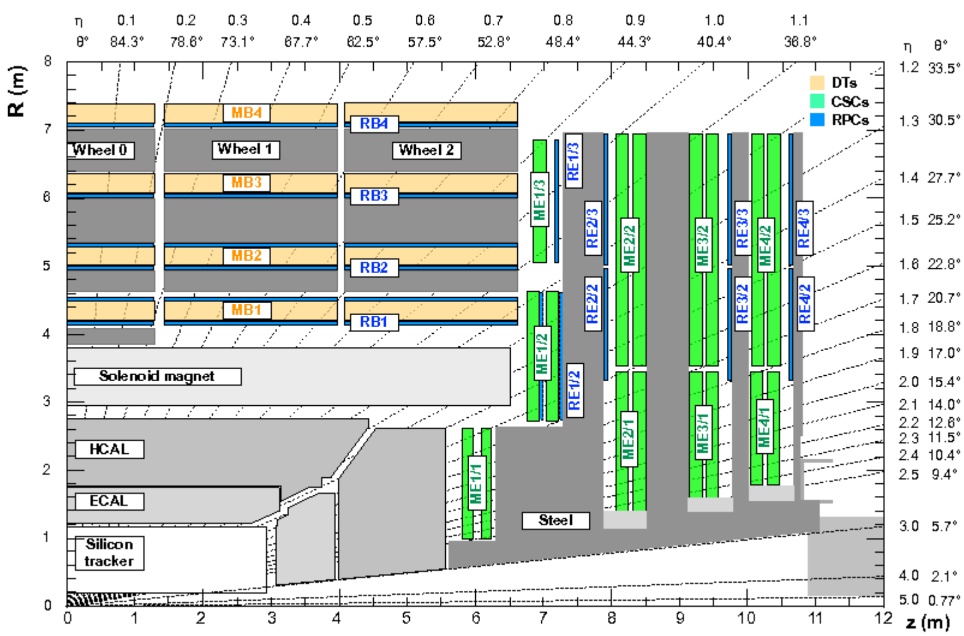


Figure 3.15: The CMS muon tracker. DT, CSC, and RPC detectors are shown in yellow, green, and blue respectively.

CMS muon system has three subdetectors: the drift tubes (DTs), the cathode strip chamber detectors (CSCs), and the resistive plate chambers (RPCs).

In the barrel region, the CMS is equipped with the DT system, consisting of 250 drift tubes arranged into five barrel sections ("wheels"). The working elements of the DT system - cylindrical cells with the rectangular base of 4.2 by 1.3 cm - are tubes with an anode wire in a mix of argon and CO_2 gases. DT cells are 2.4 m long and are organised in three groups of four elements (three "super-layers"). When the muon passes through super-layers, it ionizes the gas in the cells and the released electrons start moving to anodes. In simple terms, the muon position and direction can be determined from the time it takes for electrons to reach the anodes. DT resolution of single-cell hit positions ranges from 200 μm in the r-φ plane to 200-600 μm for forward directions.

CSCs are used in the forward direction to cover the region of $0.9 < |\eta| < 2.4$. The CSCs are multi-wire chambers made of cells that have a trapezoidal shape. The chambers contain radial copper cathode strips and, perpendicular to those, gold-plated tungsten anode wires. Each cell is filled with the mix of argon, CO_2, and CF_4 gases. The strip cells have a single-layer resolution of 300-900 μm. A CSC provides a spatial resolution of 40 -150 μm.

To improve the performance of DTs and CSCs, RPCs are used and cover the barrel and endcaps in the range of $|\eta| < 1.9$. RPCs are double-gap chambers consisting of two resistive 2 mm in thickness Bakelite layers separated by a 2 mm layers filled with a mix of $C_2H_2F_4$, isobutane C_4H_{10}, and SF_6 gases. RPCs operate in avalanche mode, producing an avalanche when the muon traverses the gas of the cell. RPCs have a spatial resolution of 0.8 - 1.2 cm, which is not as good as the ones provided by other muon subsystems, but RPCs have an advantage in terms of an excellent time resolution - just 3 ns. The barrel and the endcaps contain in total 10 RPC stations.

3.2.7 The Triggers and DAQ

The CMS trigger [66] is a system responsible for selecting events of interest and storing them for the offline analysis. The trigger has two stages: the L1 trigger (see Fig. 3.16), which reduces the event rate from 40 MHz to 100 kHz, and the HLT, which further decreases the rate to nearly 1 kHz. The L1 trigger consists of custom hardware that processes a part of the information from calorimeters and the outer tracker systems. The HLT is a part of the detector readout system (DRS) and uses the full detector information for event reconstruction. The HLT is a computing farm consisting of 22000 CPU cores that produce a decision on whether to save or to skip the event with an average time of about 220 μs. The DRS is integrated in the higher level data acquisition (DAQ) system [67]. The events selected by the HLT, which is very similar to offline event selection, are collected and sent by the DAQ to the tapes of the main CERN computing centre (Tier-0) for the persistent storage.

The L1 and HLT systems have differences and similarities. They operate at different time scales and the volumes of data they are processing are completely different. However, the goals of these systems are similar: to identify and reconstruct physics objects and combine their properties to produce an acceptance/rejection decision for each event.

3.2.7.1 The L1 Trigger

The L1 system [62] contains a so-called menu of 500 algorithms, called "seeds," designed to identify useful physics events. These seeds include trigger criteria varying from basic single-object identification to complicated selections that require certain topological conditions to be met. Each seed has a set of assigned "prescale" factors f that reduce the rate of events accepted by a particular trigger algorithm from 100% to

$100/f\%$. One of the purposes to introduce prescale factors is that the luminosity level decreases during the run period and they adjust the trigger rate to keep it constant during the data taking time.

Since the processing time of the L1 system is very important for the whole CMS operation, the L1 is built using FPGAs and ASICs custom hardware. L1 produces decisions within 3.8 μs. Data from all the calorimeters are first processed by the L1 regional calorimeter trigger (RCT) and then by a more selective global calorimeter trigger (GCT).

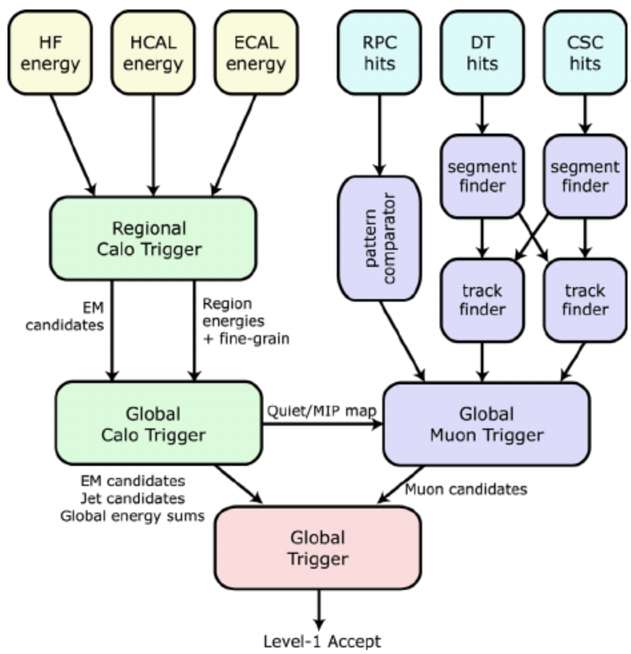


Figure 3.16: The CMS L1 trigger layout.

The RCT receives the information about energy deposits from all calorimeters and covers the range $|\eta| < 5$. The RCT processes this information in parallel and produces e/γ candidates as an output. The information from the inner tracker is not available, therefore, L1 identifies both electrons or photons, but cannot distinguish them. L1 also detects jets, taus, missing transverse energy (MET or $\not\!\!E_{\mathrm{T}}$ or $\not\!p_{\mathrm{T}}$), and muons. L1 RCT is responsible for determining the first estimates of several main parameters of interest: p_T, isolation (described later in this chapter), etc. The CMS has a limited resolution on any of these quantities at L1 and, therefore, the HLT is needed (discussed in the next subsection).

First stage reconstruction uses particle hits in the muon detectors and analyzes them using track finder algorithms. All three muon detectors of the CMS are used by the L1 muon trigger. Using DT and CSC systems, track segments from the hit information are identified. The pattern recognition algorithms are applied to these segments to reconstruct muon candidates and measure their momenta.

More complex, but slower, algorithms then reuse hits for more precise particle identification using a global muon trigger (GMT). The hits from the RPCs are used directly by the pattern comparator trigger (PACT) which reconstructs muon candidates at the high radii. Then several regional track finder algorithms sort the identified muon candidates and send this information to the GMT. Each candidate contains p_T and angular information. The GMT then combines the muon information from different subsystems to avoid duplicating the candidates. The GMT also performs more precise quality checks and may discard a portion of the input candidates it receives.

Finally, the information from the GCT and GMT is combined by a global trigger (GT). The GCT sorts the created e/γ candidates, identifies jets, and calculates $\not{E}_{\mathrm{T}}$. The final decision of the GT is to store or to skip the event. If the event satisfies the acceptance requirements and is going to be retained, the L1 accept signal (LAS) is generated and propagated by the trigger control and distribution system (TCDS) to all subdetectors. The information from this signal is then refined by the HLT.

The GT is the final step of the CMS L1 trigger system and implements a menu of triggers. The output of this system is used as an input to the HLT algorithms.

3.2.7.2 The High-Level Trigger

The selection done by the HLT mimics the offline analysis. For all reconstructed objects in the event, including electrons, muons, and jets, the identification criteria are applied to select only events of interest (defined by the physics program of the CMS). Each offline analysis defines what these events are in a different way, but to name a few, almost all analyses need true prompt leptons, well-reconstructed jets, or some other commonly used objects.

The HLT computing farm has an event filter farm, which consists of filter-builder units (FBU). In the FBU the parts of the events and information from different detector subsystems is combined to produce complete events. Then the filter unit unfolds the raw detector data into an experiment specific data structure and performs the event reconstruction and trigger filtering.

The whole event processing procedure of the HLT is centered around the HLT path. The HLT path is a set of algorithmic instructions that in a sequential manner reconstructs physics objects and performs the object selection. The instructions include kinematic and isolation requirements on the objects. The complexity of the steps in the path sequence increases and the quality of the physics objects (the probability to have a correct label) improves too (with respect to the L1). After this step is completed, selected events are sent to another software processing farm. In this storage-manager farm, the date is archived, stored locally on disk, and later sent to the CMS Tier-0 for offline use.

Most data enters the queue for processing and is ready to be sent to Tier-0 very soon. In some cases, the CMS prefers not to process immediately the special data, also known as the parked data. These data are collected and kept until the run is finished. In these situations, the data is sent for parking and the CMS tape is used. This mostly includes physics analyses such as vector boson fusion or parton distribution studies with Drell-Yan processes.

The output of the HLT is limited by capacities of the Tier-0. This includes the bandwidth of the data transfer, how quickly we want to process and store the data, and, to a degree, the amount of the available tape. All these factors complicate the work of the DAQ, because in addition to physics data streams, the calibration streams also need to be processed and stored. These streams, though, use information only from few subdetectors.

3.2.7.3 The DAQ system

The DAQ systems in modern high energy physics are responsible for many tasks. The challenges are well known: high data rates and volumes, limited tape space, and limited transfer bandwidth. CMS DAQ is based on homogeneous architecture, scales well with the different beam energy regimes and data rates, and has a stable performance in a variety of operating conditions.

To illustrate an example of a complex computing task that is elegantly solved by the DAQ system, discussing the aforementioned FBUs system of the HLT in more detail is needed. The FBU relies on a single multi-core machine in which the communication with other units is done via shared memory. The data from the full detector is used for the filtering process. Complicated offline-like reconstruction algorithms are then used for the full precision event selection. With more CPU cores available for Run 2, the per-event time budget to run HLT-style reconstruction is increased to a "comfortable" 175 ms per event, which is a long enough time to run most of the CMS reconstruction algorithms.

The current CMS DAQ was developed to address these core requirements:

- The data from one or several data transfer lines are available for other lines,
- the event building is done in parallel profiting from multiple processing units,
- almost real-time process monitoring.

Proper design patterns are used in the software for the DAQ, which decouple the user interface from the implementation. The design also allows for the remote control. The software system can be run on a number of different operating systems and hardware platforms. Finally, the memory management tools of the underlying system are not linked directly to the applications, it is done using a dedicated abstract addressing scheme.

3.2.8 The CMS design

We discussed all CMS subdetectors and DAQ. Now, one can summarize in one list the requirements on the design of the CMS detector to successfully complete its physics program. We refer to the CMS Technical Design Report [62]:

- good muon momentum resolution over the momentum scale covering almost a TeV range with the focus on dimuon resolution (mostly $Z \to \mu\mu$ and $H \to \mu\mu$) at the O(100) GeV. The capability to determine correctly the charge of the highly energetic muon all the way up to 1 TeV,
- good momentum resolution of all charged particles in the inner tracker,
- good diphoton mass resolution with the focus on the $H \to \gamma\gamma$ discovery channel. Also, the ability to reject $\pi^0 \to \gamma\gamma$, which is one of the main background processes to many physics analyses. This requirement mostly concern the performance of the ECAL,

- good resolution of the missing transverse energy (discussed in the next section) and of the mass of the two-jet system. This task depends heavily on the performance of the HCAL.

Since in this dissertation we study boosted Higgs and Z bosons, as well as leptons from $Z(\ell\ell)$decays and b jets from the H $\rightarrow b\bar{b}$decays, all the CMS subdetectors are equally important for the double Higgs boson measurement.

Chapter 4

Physics Object Reconstruction in CMS

The excellent spatial resolution of the CMS trackers, high granularity of the calorimeters, and almost 4π coverage of the detector, allowed the CMS to introduce the particle flow (PF) algorithm [68] for a global event reconstruction. The PF takes the input from all subdetectors, analyzes the redundant information and removes the duplicate one. The PF procedure starts with identification of tracks and calorimeter clusters, then reconstructs the physics objects, such as muons, electrons, and jets. In this section, first we will discuss the tracking procedure, then the elements of the PF algorithm in greater detail will be described.

4.1 Track Reconstruction

The reconstruction starts with the clusters of signals ("hits") in the inner tracker. The information from these clusters in the Pixel and Strip subdetectors is aggregated based on their signal-to-noise ratios. A charge-weighted averaging is performed (for different particle charge hypothesis). Then, a set of corrections is further applied to identify the real hit positions.

Since the number of tracks at any given event can exceed O(100), a pattern recognition technique (PRT) is employed before any further reconstruction starts. The

PRT identifies relevant tracks compatible with the candidates emerging from or near the IP. Then, the track reconstruction is applied to the selected set of tracks.

The helix trajectory that a particle follows in the magnetic field inside of the detector is characterised by five parameters: the direction in η, the 3D position (of a particular point on the helix trajectory) with respect to the reference point, which is the center of the IP, and the curvature of the track (the inverse of the helix radius R). The radius of the trajectory is related to the momentum of the particle p travelling in the magnetic field B by the formula $R \approx p/B$. This information is enough to compute estimates of basic physics quantities; however, this task is complicated by the presence of high particle multiplicity in the event (the number of charged particles produced in the same event) and also by a physics aspect of the electron propagation in matter: an electron traversing the detector has nearly 85 % probability to emit a bremsstrahlung photon. Hadronic effects also need to be taken into account; a hadron has a 20% probability to experience multiple scattering on the nuclei of the detector before reaching the HCAL.

To keep the track finding efficiency high, while maintaining low efficiency of misidentified tracks, track reconstruction is performed sequentially using the combinatorial track finder (CTF) [69], see Fig. 4.1. First the "purest" tracks are reconstructed; they have high p_T and the hits of the particle candidate point towards the beam spot. Then the hits associated with these pure tracks are removed from the collection of considered hits and another round of the track reconstruction starts. Applying this procedure several times reduces the combinatorial factor and also simplifies the identification of tracks with low p_T or those which do not point to the beam spot. During each iteration, the reconstruction follows these steps:

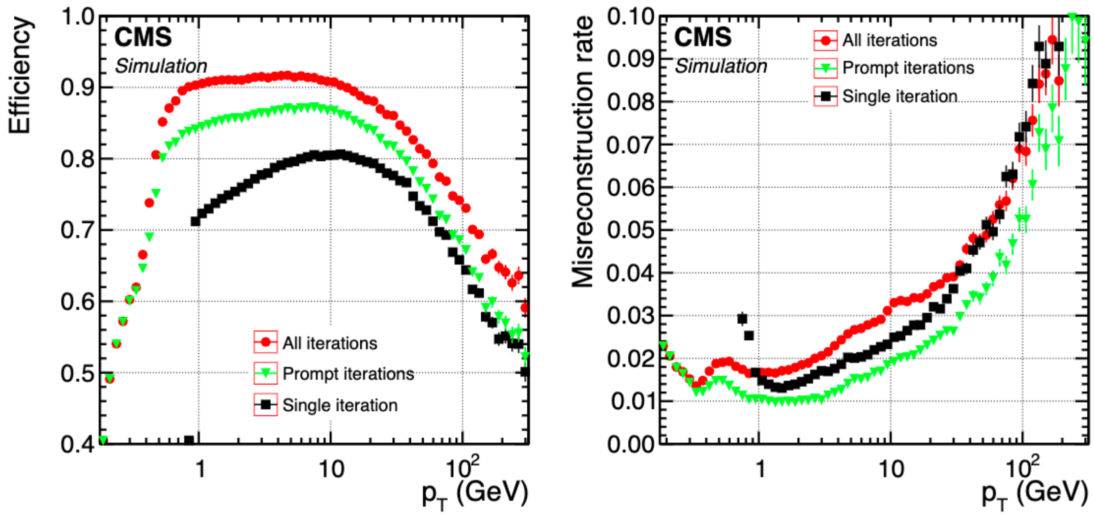


Figure 4.1: Global combinatorial track finder efficiencies as a function of the track p_T. Left: signal efficiency. Right: misidentification rate. Single iteration of the CMT is shown using black squares. The iterative tracking method consisting of 10 iterations is shown using green triangles and red circles. Prompt iterations based on seeds with at least one hit in the pixel detector are shown in green. All iterations are shown in red, including iterations when displaced seeds are present [70].

- Seed generation. Rough estimates of the particle trajectories ("seeds") are produced using either three hits or two hits and a beam spot constraint. Based on which iteration the algorithm is on, some additional constraints are applied, e.g., a minimal p_T selection, the requirement for the seed to originate close to the beam spot, etc.

- Trajectory building. Initial seeds are projected towards the compatible hits in the next layers. This approach is based on the Kalman Filter (KF) procedure [71]. The extrapolation is done until the outermost layer of the tracker is reached. Each obtained trajectory is updated using a KF approach based on the compatibility of hits to form a better track candidate. The procedure is complicated by the fact that the same initial seed can give rise to several track candidates or vice versa, the same track candidate may be compatible with different seeds. Additionally, the trajectory building step should take into account energy loses of the particle due to multiple scattering on the detector material, inhomogeneities of the tracker material, and the effects of the regions of nonconstant magnetic field.

- Track fitting. After the track candidate has been built, the track parameters are refitted by a KF and by the "smoother". This step uses the full available information about the track and gives optimal estimates of the track parameters. To remove a large number of fake tracks, which are present due to a complicated nature of the problem and the high track multiplicity in the event, a multivariate (MVA) selection is applied. The MVA discriminant incorporates variables that discriminate real tracks from the fakes: the signed transverse curvature and impact parameters (with respect to the beam spot), the polar and azimuthal angles, number of missing hits, the fit quality variables, etc.

The CTF procedure runs for 10 iterations. In 2016, the proton-proton collisions had a mean pileup (PU or additional hard scattering vertices) of 24, the CTF efficiency to identify real tracks varied from 80 to 95% with the mis-identification efficiency of 5 - 10 % depending on the p_T and η of the tracks.

4.1.1 Muon tracking

Muons are detected not only by the inner tracker but also by the outer (muon) tracker. This greatly improves muon track reconstruction, in comparison to using just the inner tracker information. Also, it motivated the development of dedicated muon reconstruction algorithms. Therefore, the ninth and the tenth iterations of the CTF are focused on the muon reconstruction. These iterations use three separate algorithms to identify:

- Standalone muons. This algorithm uses only muon tracker information: DTs, CSCs, RPCs. Hits from the inner chambers are used as seeds and are projected to hits in the outer chambers. Then, a standard KF procedure is used to identify track candidates, which are called standalone muons.

- Tracker muons. Only the inner tracking information is used to form tracks. Tracks are further projected to muon subsystems, where a compatibility with at least one muon hit is required. This algorithm works with low momentum muons: tracks with p_T above 0.5 GeV and a total momentum greater than 2.5 GeV.

- Global muons. Tracks of the tracker and standalone muons are projected to the outermost layer of the muon system, checking the compatibility between two approaches. The resulting combined set of track hits is refitted to produce a global muon track. Mostly high momentum muons with $p_T > 200$ GeV profit from this algorithm.

4.1.2 Electron tracking

Electrons are also detected by the inner tracker. However, their reconstruction is complicated by the fact that they emit bremsstrahlung photons and the trajectory becomes more complex. As a result, the clustering algorithms also need to identify the bremsstrahlung photons and account for the fact that the energy clusters corresponding to these photons may be located outside of the main electron trajectory when extrapolated to the ECAL.

In the KF approach, it is assumed that energy losses are Gaussian, and this is not the case for electron. A dedicated procedure is developed - a modified KF -

the Gaussian Sum Filter (GSF) [72]. In this method, the radiated energy losses are approximated by the sum of Gaussian distributions.

The electron seeds for the GSF are built using the ECAL information. Two different approaches are developed for the track reconstruction:

- Super Cluster (SC) based electrons. Clusters of energy in the ECAL are grouped together to form SCs. Using the information of the energy spread among the clusters, the curvature of the electron's path is estimated and tracker seeds are formed, using the SC position as a constraint.
- Track based electrons. Tracks from the inner tracker are projected to ECAL clusters, checking the compatibility using quality variables, such as χ^2, number of missing hits (absent hits along the path of the track), etc.

A typical momentum resolution for electrons in $Z \to e^-e^+$ decays is in the range of 1.7 - 4.5%.

4.1.3 Primary Vertex reconstruction

At the level of the LHC luminosity, several hard scattering vertices are produced in each collision. The location of these vertices is reconstructed using the tracks. However, normally only the primary vertex (PV) of these vertices produces interesting physics interactions and is the actual point of origin of the produced prompt particle. Other vertices are referred to as additional vertices or PU.

All the vertices are important because they are reused in the feedback loop of the track reconstruction procedure. Furthermore, a precise identification of primary vertices is important for determining the effect of the PU on all physics objects in general and b-tagging in particular (this will be discussed later in this chapter). The

PV identification consists of three steps: the tracks are selected, then the tracks from the same PV are combined in clusters, and, finally, the position of the PV is determined from the fits to tracks.

When selecting the tracks, only those consistent with the locations of the hard scattering vertices are considered. Three variables are used to improve the quality of the track selection procedure: the transverse impact parameter (the relative distance in the vertical plane with respect to the center of the beam spot), the number of strip and pixel hits associated with a track, and the χ^2 of the fit.

Clustering of the tracks is based on their z position in conjunction with the beam spot. A deterministic annealing (DA) algorithm is used to find the global minimum for this problem with many degrees of freedom. Deterministic annealing is a heuristic algorithm which comes from information theory. The idea of this algorithm is based on the thermodynamic system in physics, approaching a state of minimal energy through a series of temperature reductions. More details can be found in [73].

Once the clustering is completed and hard scattering vertices are identified, the candidates with more than two tracks are fitted using an adaptive vertex fitter [74]. The result of this procedure is a set of probabilities assigned to tracks. Each probability can be conceptualized as the likelihood that the track originated from a given vertex.

At the final stage, the vertices are ordered by the sum of the squared transverse momenta ($\sum p_T^2$) of the corresponding tracks and the vertex with the largest value of the $\sum p_T^2$ is called the main hard scattering vertex, or the PV. The resolution of the PV varies between 10 to 100 μm and depends on track qualities.

4.2 Particle level objects

4.2.1 Particle Flow links and blocks

The signals a particle leaves (in several CMS subdetectors) are stored as PF [75] elements. The link algorithm (LA) was developed to connect these PF elements. The LA can test any pair of elements in the event. To speed up the calculations, only pairs of elements that are "neighbors" in the given $\eta - \varphi$ plane are considered. When the pair of elements are linked, the LA determines the distance measure between the elements which is related to the quality of the link. This procedure produces PF blocks of elements, where the blocks are connected either by a direct or an indirect link (through the common elements).

The procedure produces inner tracker-calorimeter, HCAL-ECAL, ECAL-ECAL cluster-to-cluster, and ECAL-Preshower links. In most cases the link distance is defined in the $\eta - \varphi$ or $x - y$ planes between the two cluster positions. In case of the ambiguity when, e.g., several HCAL clusters are linked to the same ECAL cluster, only one link is kept, the one with the smallest distance. The last stage of the link algorithm is dedicated to formation of the inner tracker-muon system links.

Once all links are established and PF blocks are formed, the PF algorithm proceeds reconstructing objects in a specific sequence. Muon candidates with corresponding PF tracks and clusters are reconstructed and then removed from the PF block. After that electron candidates with the identified bremsstrahlung photons, and high momentum isolated photons with related tracks and clusters are reconstructed and removed, in that order. Finally, charged hadrons and non-isolated photons are processed (all main stages of the PF sequence are discussed separately below).

The elements that are still left in the PF blocks are then re-considered for another round of identification of other objects: charged hadrons and neutral hadrons, photons from parton fragmentation and hadronization, and jet decays. Lastly, when all PF blocks have been sorted and identified, the global event description is completed and the reconstructed event is re-processed by a post-processing step (PP step), addressing the possible particle misidentification and improper reconstruction during the previous steps.

4.2.2 Muons

Muon reconstruction is the first stage of the PF algorithm. It identifies muons using global and inner tracker muon properties. Global muon candidates are selected at this step and the isolation requirement (explained later in this chapter) is applied. This isolation requirement is efficient enough to reject hadrons mis-identified as muons. The muons inside jets or secondary muons from hadron decays complicate the identification of prompt muons, such as those originating from Higgs, W, or Z boson decays. Therefore, an additional more stringent selection is further applied.

For non-isolated global muons, in addition to the tight muon selection (a set of requirements to remove fake muons) [76], the following extra selection is applied: more than two matching track segments should be present in the muon detectors or the calorimeter energy deposits must be compatible with the energy of the muon candidate. This selection discriminates against high p_T hadrons. If muons fail Tight WP, there are additional recovery procedures that use the relaxed selection criteria to search for lower quality muons.

Even at this stage in the PF procedure, the muon identification and reconstruction is not finished. Charged hadrons reconstructed during the next PF stages can be reconsidered as muon candidates. Only after the whole PF sequence is completed, including the PP step, does the PF algorithm terminate.

All muons considered for this measurement are global muons that satisfy the Tight WP requirements (as an initial selection) with extra requirements on the number of hits in the tracker and muon system, on the impact parameter, and the quality of the global track.

The efficiency ϵ to successfully identify a prompt isolated lepton can be decomposed as:

$$\epsilon = \epsilon_{tracker} \cdot \epsilon_{ID|tracker} \cdot \epsilon_{ISO|ID}$$

where the first term refers to the tracker efficiency, the next term is the Bayesian term which refers to the identification (ID) efficiency given that the lepton already passed the tracker requirements, and the final term refers to the isolation (ISO) efficiency given that the lepton already satisfied identification criteria. All the efficiencies are well optimised in the CMS and are in the range from 85 % to more than 99 % depending on the p_T and η of the lepton (muon in this case). The muon momentum resolution for 20 to 100 GeV momentum range varies from 1 % in barrel to 5 % in endcaps.

4.2.3 Electrons and isolated photons

The reconstruction of electrons is complicated by the fact that they lose energy emitting bremsstrahlung photons. As the trajectory of the particle change, the curvature of the helix trajectory is also changing. This makes the electron's trajectory more complex than the one of a muon. Additionally, bremsstrahlung photons often convert to e^+e^- pairs. This is a recursive process, as daughter electrons also emit photons. Due to this complication, it was decided to use almost the same procedure to reconstruct electrons and photons. First a GSF track plays a role of the seed for the electron candidate. For the photon candidate, an ECAL supercluster with no links to the GSF track is used as a seed. For both electron and photon candidates energy deposits in the HCAL must not exceed 10% of the ECAL energy.

ECAL energy deposits, that are above a certain threshold and located close to the most energetic deposit, are grouped into the SC and may be linked to one of the GSF track candidates. During the linking procedure, the energy of the collected ECAL energy deposits is corrected for the additional energy losses. An electron candidate is formed from a combination of the corrected ECAL energy and the electron direction given by the GSF track. An additional MVA discriminator of O(20) variables is applied to improve electron identification efficiency. The MVA approach, based on the Boosted Decision Trees (BDT) classifier [77], profits from the following highly discriminating variables: the amount of energy radiated off the GSF track, the distances between the ECAL SC position and the position given by the tangent to the GSF track, track-cluster linking variables, KF and GSF track quality variables, etc.

Photon candidates are kept if the photons are isolated and the corresponding configuration of ECAL energy deposits is compatible with those expected from a given photon shower. All identified electron and photon tracks and clusters in the PF block are masked before the algorithm starts processing hadrons.

As offline physics analyses apply different selection for electrons and photons, PF selection is relatively loose and the full electron and photon reconstruction information is saved in case a different re-interpretation must be run. This offers the saving of the computing time in the future, since re-running the electron track reconstruction would not require re-running the complete PF algorithm again.

4.2.4 Hadrons and non-isolated photons

After the muons, electrons, and isolated photons are reconstructed, they are removed from the PF blocks. The next PF algorithm iterations proceed with hadrons from jet fragmentation and hadronization. These particles can be "seen" by the detector as charged pions, kaons or protons, neutral pions and kaons, and non-isolated photons from neutral pion decays. During the reconstruction, precedence is given in the ECAL to photons over neutral hadrons. This priority does not hold above $|\eta| > 2.5$. In that region, ECAL clusters linked to a given HCAL cluster are identified as hadrons and, only if ECAL clusters are without such a link, then they are classified as photons.

What is left in the PF block may be considered a neutral hadron if there are "absent" tracks in the tracker. If tracks are present and compatible with the energy deposits, a charged pion candidate is formed.

In situations where the energy deposits in the calorimeters do not match the energy hypothesis from the tracker well, and this discrepancy is larger than three standard deviations, a new muon reconstruction starts, with the relaxed muon selection. This approach allows to improve the muon identification efficiency without increasing the rate of mis-identified muons.

At times the energy associated with the track momentum sum may be found to be significantly larger than the calorimetric energy. Usually this excess in momentum is found to arise from residual mis-reconstructed tracks with a p_T larger than 1 GeV. These tracks are sorted in decreasing order of their p_T and are removed one-by-one from the PF block until no such tracks are left in the initial track collection.

The hadron traversing the material of the tracker interacts with the nuclei of the tracker material and often produces secondary hadrons. These secondary hadrons are produced outside of the PV - at a secondary (intermediate) interaction vertex.

When the tracks of the charged particles corresponding to these secondary particles are linked together, the resulting secondary particle candidates can be replaced in the list of reconstructed particles by a single (original) charged hadron.

The estimate of the energy of the primary charged hadron is then given by:

$E = E_{secondary} + f \cdot p_{primary}$

where $E_{secondary}$ is the vectorial sum of the momenta of the secondary charged particles, $p_{primary}$ is the momentum of the incoming track, and f is a factor that needs to be determined from the simulations. This factor is used to estimate the energy of undetected secondary particles, which are often reconstructed neither as secondary charged particles nor as neutral particles.

4.2.5 Jets and jet corrections

Jets are collimated streams of particles created during the processes of the fragmentation and hadronization of the original parton, quark, or gluon. As jets propagate through the CMS detector, they leave tracks in the tracking system and interaction showers in the calorimeter crystals.

Several jet reconstruction algorithms have been developed. In the Higgs boson group of the CMS, most measurements are using anti-k_T algorithm [78]. If the jet clustering uses PF particles, then PF jets are reconstructed. If only the ECAL and HCAL information is used, calorimeter jets are identified ("calo jets"). When all stable particles (in case of the simulation, at the generator level) excluding neutrinos are used - reference jets are reconstructed ("Ref jets"). In this measurement PF anti-k_T jets are used.

The anti-k_T algorithm is one of the “cone” algorithms that takes as input a collection of PF objects inside of a cone of the radius R, which is the parameter that determines the final size of the jet and is usually between 0.4 - 0.7. The algorithm defines the distance parameter $d_{ij} = \min(\frac{1}{p_{T_i}^2}, \frac{1}{p_{T_j}^2}) \times \frac{R_{ij}^2}{R}$, where p_{T_i} and p_{T_j} refer to the transverse momenta of PF particles i and j, R_{ij} is the distance in the $\eta - \varphi$ plane between particles i and j. An additional d_{iB} parameter is defined as the distance between the particle i and the beam spot position: $d_{iB} = \frac{1}{p_{Ti}^2}$.

This algorithm iteratively finds the minimum distance selecting at each step the minimal value for each (d_{ij}, d_{iB}) pair, using a collection of the PF particles as an input. The algorithm then seeks out the smallest d for a given input. If the minimum distance is d_{ij}, then the four-vectors of i and j particles are summed to form a new particle. Particles i and j are removed from the initial input collection. If the minimum distance is d_{iB}, then the particle i (a PF particle candidate) is considered a jet. This particle is also removed from the set of particles and the algorithm continues until all initial particles have been combined into jets. The mechanics of the algorithm is such that first jet candidates with the hardest particles are clustered, producing a perfect cone-shaped jets, then more complex jets are reconstructed.

PF jets were used to conduct the data analysis in this thesis. They are superior to calo jets because the former have a better angular resolution. The PF algorithm allows the precise determination of the charged hadron direction and momentum, while in calorimeters, the energy deposits of charged hadrons are spread along the φ direction in the presence of the magnetic field, which leads to an extra degradation of the azimuthal angular resolution of jets.

On average, the relative contributions to jet energy are: 65% from charged hadrons, 25% from photons, and 10% from neutral hadrons. The possibility to identify the contributors to the total jet energy during the jet reconstruction is one of the reasons to use the PF algorithm for jet reconstruction. In practice, the identification of particles inside jets is done by comparing the jet energy fractions measured in PF jets to those of the corresponding Ref jets.

To remove the jet energy dependence on p_T and η (JE map), and make the corresponding two dimensional JE map uniform, the jet energy correction (JEC) procedure is introduced. The jet energy resolution (JER) correction is also necessary. The latter is defined as the Gaussian width of the ratio of the energies of the corrected PF jets to Ref jets. Both corrections improve the angular resolution, energy response, energy resolution of jets, and make the simulation match the data.

The JEC scales the four-momentum of jets. Various detector effects are addressed. This correction is often split into separate components, which are applied individually in a sequence. The most important individual corrections remove energy contributions due to PU, effects of the calorimeter response, residual data-Monte Carlo (MC) discrepancies, and effects of the jet flavor.

JER smears the four-momenta of reconstructed jets to match the energy resolution observed in data. The smearing procedure derives the correction factors which scale the momentum of the reconstructed jet with respect to the the momentum of the "gen" jet, which is same jet, but clustered at the MC generator level.

4.3 Other important physics quantities and objects

4.3.1 The b tagging and secondary vertices

In this measurement, one of the Higgs bosons will decay into b quarks. Jets produced during the hadronization of b quarks are called b jets. A dedicated b tagging is necessary since $H \rightarrow b\bar{b}$ decay has the highest branching fraction of almost 58%. Therefore, measuring precisely this BF is a great test of the SM.

Bottom quarks will produce jets that contain B mesons or baryons containing a b quark, which have a relatively long lifetime $c\tau \approx 400-500\mu$m. This distance, traveled at almost the light speed, would correspond to a dislocation of a few mm from the PV. The positions of B meson decays (or baryons containing a b quark) will be clearly seen in the detector. Each such position with a corresponding displaced vertex, the secondary vertex (SV), is a unique signature of b quark and is used to identify the b quark decay. Sometimes the vertex cannot be unambiguously reconstructed, but even in these cases the properties of tracks within b jets are different from the ones originating from gluons or light quarks (light jets).

After passing the selection criteria, tracks are considered for b tagging. The selection requirements, which include kinematic and impact parameter properties of tracks, are needed to reject fake tracks, tracks coming from PU vertices, and tracks from the long-lived hadrons such as K_S^0 or Λ_b^0.

Two main approaches are used to reconstruct the secondary vertices (SV). The first method is an adaptive vertex reconstruction (AVR) algorithm. AVR is based on the adaptive vertex fitter; it uses the tracks associated with jets and finds PV and SVs. The other algorithm is the inclusive vertex finder (IVF). IVF uses all the tracks in the event and is implemented with the selection looser than for the AVR.

As for the b tagging itself [79, 80], multiple algorithms (taggers) have been developed and successfully used over the last decades. The most known ones are:

- the jet probability (JP) and the jet b probability (JBP) taggers. Both are based on the probability of a jet candidate to be compatible (or incompatible) with the PV using impact parameter significance (IPS) variables.

- The soft electron tagger (SET) and soft muon tagger (SMT). These taggers are based on the presence of soft leptons within jets, focusing on leptonic decays of

B hadrons, because B hadrons have a larger BF to leptons than other hadrons do.

- The combined secondary vertex (CSVv2) tagger. This is a more complex tagger based on an MVA technique. It uses displaced tracks and secondary vertices to tag b jets and takes as input IPS, decay length, SV parameters, number of SVs, etc. This tagger can use both AVR and IVF vertices.

In this physics analysis a new MVA based tagger (cMVAv2) was used. This superior tagger uses the outputs from all the aforementioned "fundamental" taggers: JP and JBP, SET and SMT, CSVv2 using both AVR and IVF vertices. These ensemble learning procedure [81] of combining outputs from fundamental taggers into one complex MVA based tagger that produces the final output is a popular machine learning technique that has been proven to lead to better results than the ones achieved by individual algorithms separately.

4.3.2 Missing transverse momentum

When neutrinos are present in the event, they cannot be directly detected by the CMS; specific neutrino detectors would be needed in this case. However, using the CMS detector one can indirectly estimate the momentum of neutrinos. This procedure relies on the method of the “missing transverse momentum“ $\not{p}_{\mathrm{T}}$ (or missing transverse energy $\not{E}_{\mathrm{T}}$ (MET)), see Fig. 4.2. MET is constructed using all PF particles in the event and is calculated as:

$\not{p}_{\mathrm{T}} = \vec{p}_T^{miss} = |-\sum_i^N \vec{p_{Ti}}|$.

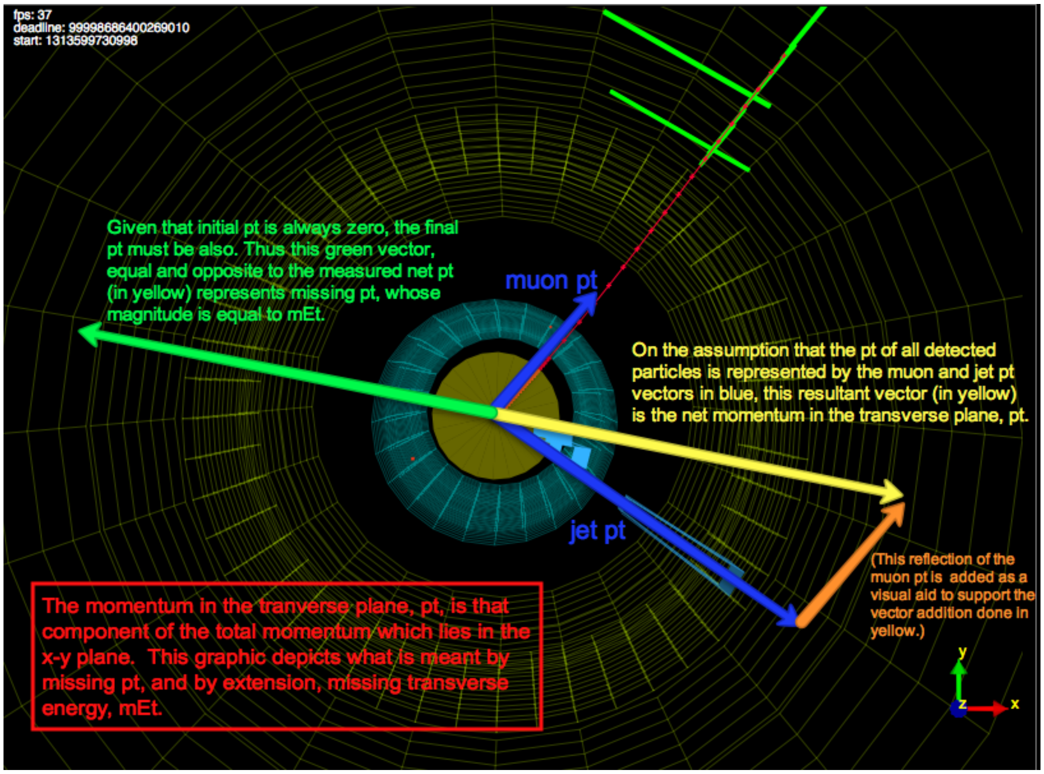


Figure 4.2: MET reconstruction in the CMS. The only detected (visible) particles are muon and jet both shown using blue arrows. The total net momentum “seen” by the CMS detector is represented by the yellow arrow. From the conservation of momentum, the reconstructed MET corresponds to the green arrow. Taken from [82].

To reconstruct $\not{p}_{\mathrm{T}}$, the CMS relies on almost 4π coverage of the detector and precise measurement of the particle properties using PF algorithm that takes information from all subsystems. The resulting $\not{p}_{\mathrm{T}}$ is still considered "raw", since this MET is not yet JEC or JER corrected. After these corrections are applied, one obtains $\not{p}_{\mathrm{T}}$ that is called "Type-1 corrected MET". This is the version of MET that is recommended by the CMS JetMET particle Object Group (POG) group and is used in this measurement. An additional set of filters and corrections is further applied to reject events with artificially large $\not{p}_{\mathrm{T}}$ due to the presence of several noise sources, such as ECAL dead cells, poor quality muon candidates, and HCAL noise.

4.3.3 Pileup interactions

In 2016, a typical collision event contained on average 24 interaction vertices. Some events had nearly 40 inelastic proton-proton interactions. All the vertices excluding the PV can be referred to as the soft vertices. These soft vertices or PU do not produce interesting physics interactions, however, they contribute considerably to the total number of particles produced in the event. PU creates additional hadrons and photons and this affects the PF reconstruction of jets and $\not{p}_{\mathrm{T}}$, and also the lepton isolation calculation.

Charged hadrons which have originated from PU are identified by backtracking the origin of the tracks and checking the compatibility with the PU vertices. These charged hadrons are removed from the collection of particles used to reconstruct physics objects. This procedure is called the charged-hadron subtraction (CHS). Because neutral hadrons and photons leave no tracks in the inner tracker, their presence needs to be identified and addressed differently. Instead, an average density ρ of PU interactions in the given $\eta - \varphi$ slice is calculated. Assigning the area of the candidate to the value A_{eff} ("effective area"), the expected PU initiated contribution that needs to be subtracted, is calculated as $\rho \cdot A$. Several other techniques are available, with one of the most simple relying on the calculation of the ratio of the neutral to the charged energy coming from PU (around a given lepton). From the studies, this value (called $\Delta\beta$) is determined to be very close to 0.5 [83].

4.3.4 Lepton isolation

As mentioned before, lepton isolation is the preferred method to remove clear fakes and select real prompt muons and electrons produced by Higgs boson decay or by the weak decays of Z or W bosons. The isolation quantifies the activity of other particles around the particle of interest. The lepton isolation is defined as the scalar sum of the p_T's of all charged and neutral hadrons and photons inside a cone of the radius $\Delta R < 0.3 - 0.4$(depending on the WP and a lepton flavor). The sum is normalised by the p_T of the lepton of interest:

$$I_{PF} = \frac{1}{p_{T_\ell}}(\sum^{\gamma} p_T^{\gamma} + \sum^{h\pm} p_T^{h\pm} + \sum^{h^0} p_T^{h^0})$$

There are other physics analysis-specific definitions of isolation and they are applied offline. For the purpose of this thesis, the two isolation requirements listed below are used.

For electrons, the isolation selection relies on the notion of the effective area defined in the previous subsection. The effective areas are proportional to ΔR cone size around the electron (0.3 in this case), and for electrons the isolation is given by:

$$I_{PF}^{electron} = \sum^{h^{\pm}} p_T + max(0, \sum^{h^0} p_T + \sum^{\gamma} p_T - \rho \cdot A_{eff}).$$

The $\Delta\beta$ method mentioned previously is used for muon isolation selection. With the cone size of 0.4, the isolation is given by:

$$I_{PF}^{muon} = \sum^{h^{\pm}} p_T + max(0, \sum^{h^0} p_T + \sum^{\gamma} p_T - \Delta\beta \cdot \sum^{h^{\pm}_{PU}} p_T).$$

The last term is the $\Delta\beta$ correction multiplied by the sum of p_T's of the charged hadrons originated due to PU.

Both isolations are constructed from the collection of particles containing charged and neutral hadrons, photons, and charged hadrons from the PU (in case of muons).

4.3.5 Data-Monte Carlo corrections

Even though particle interactions in the detector, the detector response, and the work of subsystems are simulated at a level of high precision, there is still disagreement present between the data and the simulation. The approximations used to speed up the subsystem responses, slight detector misalignments, the impossibility to know the exact parton distribution function (PDF) of the interacting particles at the interaction vertex, fluctuation of the LHC parameters and other factors contribute to the data-MC discrepancy. To reduce this disagreement, corrections factors are introduced. Efficiencies of different selections are measured both in data and in the MC, and their ratio is applied to the MC to make the MC similar to observed data. These ratios (corrections), also called "scale factors" (SF), are derived for all physics objects and the most important ones are discussed below.

4.3.5.1 Lepton efficiencies and the Tag-and-Probe method

Several steps are involved in the process of selecting a prompt lepton: tracking, identification, isolation, and trigger stages. A popular technique to measure the efficiency of these stages is called the Tag-and-Probe ($T\&P$ or TnP) method. Decays of $Z \to e^-e^+$ or $Z \to \mu^-\mu^+$ are used in this technique. The procedure first picks one lepton that has to pass a relatively tight ID selection; this lepton is called the "tag". Tags are often referred to as "golden" electrons or muons and have a low fake rate of the order of 1% or less. Then the other lepton, called the "probe", is selected to make a pair with the tag. This step results in total of P_{total} pairs. This pairing procedure includes some very basic selection: probes should be of the opposite sign and the same lepton flavor (OSSF). We need an unbiased source of leptons, and Z boson decay is a good source of such leptons. Therefore, the consistency with the Z boson pole mass is further checked. The exact definition of the probe object varies depending on the specifics of the selection of interest or the WP. In this framework, the efficiency is defined as a ratio of the number of probes P_{pass} that pass a relatively "tight" ID WP to the total number of probes P_{total} formed by the pairing procedure:

$\epsilon_{WP} = \frac{P_{pass}^{WP}}{P_{total}}$

As the efficiency is not be flat for all p_T and η ranges, a set of efficiencies is derived for different p_T and η slices. The procedure also has an uncertainty associated with the method. Also, the TnP procedure allows for the removal of the combinatorial backgrounds by kinematic fitting or sideband subtraction methods [84].

The TnP method applied independently to data and MC, produces scale factors given by: $SF_{WP} = \frac{\epsilon_{WP}^{data}}{\epsilon_{WP}^{MC}}$.

The L1 trigger did not have a proper simulation for 2016 data taking settings, so only HLT trigger SFs are measured for the trigger. Therefore, for L1, simulated events were weighted by the efficiency measured directly in data.

In addition, some HLT paths contained the DZ requirement, while others did not. Therefore, one needs to estimate additional efficiency related to the DZ selection and derive the corresponding SFs. The DZ scale factor calculation is very similar to estimation of efficiencies of all other sorts: the numerator contains events that pass the DZ requirement and the denominator is equal to the number of events that pass the selection without DZ requirement. Exactly the same procedure is applied to derive tracker, ID, and ISO SFs. Analysis specific figures will be shown in the data analysis chapter.

4.3.5.2 b tagging efficiency

Scale factors also need to be derived for b jets. Tracker misalignment and the imperfect knowledge of the hadronization process of the b quark are the factors that lead to data-MC discrepancies. Additionally, the Strip tracker had known inefficiencies during 2016 data taking, which resulted in poor b tagging performance: lower efficiency to tag real b jets and higher fake rate (incorrect mistagging of light jets or gluons as b jets). The SFs have been derived by the b tagging CMS POG for the use of the whole collaboration. These correction factors are measured using the “true” or “generator” (gen) flavor of the original quark in the MC. Based on that, the b tagging weight is assigned to the particle level jet that is matched to the reconstruction level jet. SFs are provided in p_T and η slices to make the b tagging efficiency more uniform across the whole kinematic phase space. Further, MC events are reweighed using the combined weights from all jets present in the events.

In this measurement, the data analysis is performed with b jets tagged by the cMVAv2 (CMVA) algorithm. The $t\bar{t}$ process containing top and anti-top quark decays is used to determine the b tagging weights. The CMVA discriminant should be above a certain threshold for b jets to be considered originating from the b quarks. The threshold is chosen to correspond to the medium working point of the algorithm defined such that the misidentification rate for light-quark and gluon jets is about 1%. The b jet tagging efficiency for this WP is about 66%.

4.4 Datasets and Trigger Paths

The proton-proton collision data recorded by CMS is split into "eras", which are labelled using alphabetic letters A $\rightarrow$ H. Period A was dedicated to commissioning of the LHC for 2016 data taking. Periods from B to H were used for physics. Each era corresponds to a relatively stable period of the LHC conditions, such as the collision rate, the set of trigger menus, etc. Eras B to G were re-reconstructed at the end of 2016 to take advantage of the updated calibration of subsystems and the detector alignment. The data from the last era (H) was re-reconstructed during data-taking itself.

Dozens of datasets (Primary Datasets or PDs) are recorded and stored by the CMS and its computing centers. The data that is analysed in this thesis belongs to “Dimuon”(“DoubleMuon”) or “Dielectron”(“DoubleEG”) datasets. This naming practice comes from the fact that to select these PDs, the HLT paths with two prompt muons or electrons were used.

The name of the trigger paths (L1 in this case) reflects the number of leptons selected by a given trigger, the type of the lepton(s), following by the minimal p_T requirement(s) on the lepton candidate(s). If two leptons are present, their p_T's are referred to as p_T's of the leading and subleading(trailing) lepton (another common naming is two "legs"). If the suffixes "Iso" or "Id" (or "ID") are present, it indicates that the isolation or identification requirements respectively have been also applied. The label "DZ" or "dz" is an additional requirement on the spatial compatibility along the z axis between the lepton candidates and the PV location. Abbreviations VVL and VL refer to "very very loose" and "very loose" selections respectively. Their exact definition may vary, but the important point is that this selection does reject some clear fakes, while is still loose enough to leave enough statistics for offline analysers.

Table 4.1: Triggers for dimuon and dielectron channels both at L1 and HLT levels.

Channel	L1 Paths	HLT Paths
"Dimuon" Z($\mu\mu$) Z($\nu\nu$)H($b\bar{b}$)	L1_SingleMu20	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v* OR HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v* OR HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* OR HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
"Dielectron" Z(ee) Z($\nu\nu$)H($b\bar{b}$)	L1_SingleEG30 OR L1_SingleIsoEG22er OR L1_SingleIsoEG24 OR L1_DoubleEG_15_10	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ

A simplified version of the PF event reconstruction sequence is performed at the HLT level. For the HLT paths used in this measurement, this sequence is based only on the regional track finding and fitting, relying on the muon and e/γ candidates found by L1 trigger. The HLT paths refer to triggers that select events where one or two leptons that pass certain selection criteria are present. HLT muons (abbreviated to “Mu”) are formed by propagating the L1 track inwards to the inner tracker, or by starting from an inner track and projecting it to the outer tracker. HLT tracker muons (abbreviated to “TkMu”) are reconstructed using the muon tracking procedure discussed in the Section 4.1.1. The HLT electrons are reconstructed using L1 seeds and checking their compatibility with the energy deposits in the ECAL. If the ID selection is applied on the track, the candidate name contains the suffix “TrackId”. If the ID is applied on the ECAL energy cluster parameters, the name of the path contains “CaloId”.

In the next chapter, there will be a thorough discussion of the physics analysis of the data produced by the LHC and collected with the CMS detector that have been used to perform the measurement of the double Higgs boson decays.

References

[1] Erwin Schrödinger. *Statistical thermodynamics; 2nd ed.* Cambridge Univ. Press, Cambridge, 1952.

[2] Richard Phillips Feynman, Robert Benjamin Leighton, and Matthew Sands. *The Feynman lectures on physics; New millennium ed.* Basic Books, New York, NY, 2010. Originally published 1963-1965.

[3] David J Griffiths. *Introduction to elementary particles; 2nd rev. version.* Physics textbook. Wiley, New York, NY, 2008.

[4] E A Davis and Isabel Falconer. *J.J. Thompson and the discovery of the electron.* Taylor and Francis, Hoboken, NJ, 2002.

[5] Oreste Piccioni. *The Discovery of the Muon*, pages 143–162. Springer US, Boston, MA, 1996.

[6] Carl Bender. Mathematical physics.

[7] G. Danby, J-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Phys. Rev. Lett.*, 9:36–44, Jul 1962.

[8] M. L. Perl, G. S. Abrams, A. M. Boyarski, et al. Evidence for anomalous lepton production in $e^+ - e^-$ annihilation. *Phys. Rev. Lett.*, 35:1489–1492, Dec 1975.

[9] K. Kodama et al. Observation of tau neutrino interactions. *Phys. Lett.*, B504:218–224, 2001.

[10] Eric W. Weisstein. Fundamental forces.

[11] S Chandrasekhar. *Newton's principia for the common reader.* Oxford Univ., Oxford, 2003. The book can be consulted by contacting: PH-AID: Wallet, Lionel.

[12] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation / Charles W. Misner, Kip S. Thorne, John Archibald Wheeler.* W. H. Freeman San Francisco, 1973.

[13] Hanoch Gutfreund and Jurgen Renn. *The road to relativity: the history and meaning of Einstein's "The foundation of general relativity" : featuring the original manuscript of Einstein's masterpiece.* Princeton University Press, Princeton, NJ, Apr 2015.

[14] J. Butterworth. *Smashing Physics.* Headline Publishing Group, 2014.

[15] A Zee. *Quantum Field Theory in a Nutshell.* Nutshell handbook. Princeton Univ. Press, Princeton, NJ, 2003.

[16] W N Cottingham and D A Greenwood. *An Introduction to the Standard Model of Particle Physics; 2nd ed.* Cambridge Univ. Press, Cambridge, 2007.

[17] R. P. Feynman. The theory of positrons. *Phys. Rev.*, 76:749–759, Sep 1949.

[18] Francis Halzen and Alan Douglas Martin. *Quarks and leptons: an introductory course in modern particle physics.* Wiley, New York, NY, 1984.

[19] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[20] Andrew Wayne. QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga by Silvan S. Schweber. *The British Journal for the Philosophy of Science*, 46(4):624–627, 1995.

[21] Michelangelo L Mangano. Introduction to QCD. (CERN-OPEN-2000-255), 1999.

[22] Matt Strassler. Of particular significance: Conversations about science with theoretical physicist matt strassler.

[23] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

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

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

[26] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[27] Pauline Gagnon. *Who cares about particle physics? : making sense of the Higgs boson, the Large Hadron Collider and CERN.* Oxford University Press, 2016.

[28] Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model. Technical Report CMS-PAS-HIG-14-009, CERN, Geneva, 2014.

[29] Jennifer Ouellette. Einstein's quest for a unified theory. *APS*, 2015.

[30] S. M. Bilenky. Neutrino in Standard Model and beyond. *Phys. Part. Nucl.*, 46(4):475–496, 2015.

[31] Matthias U. Mozer. Electroweak Physics at the LHC. *Springer Tracts Mod. Phys.*, 267:1–115, 2016.

[32] Gennadi Sardanashvily. *Noether's theorems: applications in mechanics and field theory*. Atlantis studies in variational geometry. Springer, Paris, 2016.

[33] Steven Weinberg. The Making of the Standard Model. *Eur. Phys. J. C*, 34(hep-ph/0401010):5–13. 21 p. ; streaming video, 2003.

[34] Roger Wolf. *The Higgs Boson Discovery at the Large Hadron Collider*, volume 264. Springer, 2015.

[35] Jose Andres Monroy Montanez, Kenneth Bloom, and Aaron Dominguez. Search for production of a Higgs boson and a single Top quark in multilepton final states in pp collisions at $\sqrt{s} = 13$ TeV, Jul 2018. Presented 23 Jul 2018.

[36] Peisi Huang, Aniket Joglekar, Min Li, and Carlos E. M. Wagner. Corrections to di-Higgs boson production with light stops and modified Higgs couplings. *Phys. Rev.*, D97(7):075001, 2018.

[37] Matthew J. Dolan, Christoph Englert, and Michael Spannowsky. New Physics in LHC Higgs boson pair production. *Phys. Rev.*, D87(5):055002, 2013.

[38] Shinya Kanemura, Kunio Kaneta, Naoki Machida, Shinya Odori, and Tetsuo Shindou. Single and double production of the Higgs boson at hadron and lepton colliders in minimal composite Higgs models. *Phys. Rev.*, D94(1):015028, 2016.

[39] Albert M Sirunyan et al. Search for Higgs boson pair production in the $\gamma\gamma\mathrm{b}\overline{\mathrm{b}}$ final state in pp collisions at $\sqrt{s} = 13$ TeV. 2018.

[40] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999.

[41] Alexandra Oliveira. Gravity particles from Warped Extra Dimensions, predictions for LHC. 2014.

[42] Alexandra Oliveira. Gravity particles from Warped Extra Dimensions, predictions for LHC. 2014.

[43] Kunihito Uzawa, Yoshiyuki Morisawa, and Shinji Mukohyama. Excitation of Kaluza-Klein gravitational mode. *Phys. Rev.*, D62:064011, 2000.

[44] Walter D. Goldberger and Mark B. Wise. Modulus stabilization with bulk fields. *Phys. Rev. Lett.*, 83:4922–4925, 1999.

[45] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo. Phenomenology of the Randall-Sundrum Gauge Hierarchy Model. *Phys. Rev. Lett.*, 84:2080, 2000.

[46] Piotr Traczyk and Grzegorz Wrochna. Search for Randall-Sundrum graviton excitations in the CMS experiment. 2002.

[47] Walter D. Goldberger and Mark B. Wise. Bulk fields in the Randall-Sundrum compactification scenario. *Phys. Rev.*, D60:107505, 1999.

[48] Sreerup Raychaudhuri and K Sridhar. *Particle physics of brane worlds and extra dimensions*. Cambridge monographs on mathematical physics. Cambridge University Press, Cambridge, 2016.

[49] Michael Forger and Hartmann Romer. Currents and the energy momentum tensor in classical field theory: A Fresh look at an old problem. *Annals Phys.*, 309:306–389, 2004.

[50] Lisa Randall and Raman Sundrum. Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, Oct 1999.

[51] Roberto Contino, Margherita Ghezzi, Mauro Moretti, Giuliano Panico, Fulvio Piccinini, and Andrea Wulzer. Anomalous Couplings in Double Higgs Production. *JHEP*, 08:154, 2012.

[52] Chuan-Ren Chen and Ian Low. Double take on new physics in double Higgs boson production. *Phys. Rev.*, D90(1):013018, 2014.

[53] Giuliano Panico. Prospects for double Higgs production. *Frascati Phys. Ser.*, 61:102, 2016.

[54] Sebastien Wertz and Vincent Lemaitre. Search for Higgs boson pair production in the $\mathrm{b\bar{b}}\ell\nu\ell\nu$ final state with the CMS detector, 2018.

[55] Luca Cadamuro, Yves Sirois, and Roberto Salerno. Search for Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel with the CMS detector at the LHC. Recherche de la production de paires de bosons de Higgs dans le canal de désintégration $b\bar{b}\tau^+\tau^-$ avec le détecteur CMS auprès du LHC, Sep 2017. Presented 05 Oct 2017.

[56] Thomas Schörner-Sadenius. *The Large Hadron Collider: harvest of run 1.* Springer, Cham, 2015.

[57] CERN. *Large Hadron Collider in the LEP Tunnel*, Geneva, 1984. CERN.

[58] Cecile Noels. Literature in focus - The Large Hadron Collider: A Marvel of Technology. Literature in focus - The Large Hadron Collider: A Marvel of Technology. (BUL-NA-2009-414. 51/2009):11, Dec 2009.

[59] John Hauptman. *Particle physics experiments at high energy colliders.* 2011.

[60] CMS Collaboration. The Phase-2 Upgrade of the CMS Tracker. Technical Report CERN-LHCC-2017-009. CMS-TDR-014, CERN, Geneva, Jun 2017.

[61] Philippe Bloch, Robert Brown, Paul Lecoq, and Hans Rykaczewski. *Changes to CMS ECAL electronics: addendum to the Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 2002.

[62] G L Bayatian, S Chatrchyan, H Hmayakyan, A Poblaguev, M E Zeller, and B S Yuldashev. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software.* Technical Design Report CMS. CERN, Geneva, 2006. There is an error on cover due to a technical problem for some items.

[63] *The CMS hadron calorimeter project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.

[64] G Baiatian, Albert M Sirunyan, Virgil E Emeliantchik, Igor Barnes, Alvin T Laasanen, and Arnold Pompos. Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges. Technical Report CMS-NOTE-2006-138. 1, CERN, Geneva, May 2007.

[65] CMS Collaboration. The Phase-2 Upgrade of the CMS Muon Detectors. Technical Report CERN-LHCC-2017-012. CMS-TDR-016, CERN, Geneva, Sep 2017. This is the final version, approved by the LHCC.

[66] Vardan Khachatryan et al. The CMS trigger system. *JINST*, 12(01):P01020, 2017.

[67] G. Bauer et al. The CMS data acquisition system software. *J. Phys. Conf. Ser.*, 219:022011, 2010.

[68] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017.

[69] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014.

[70] Albert M Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017.

[71] Giuseppe Cerati et al. Parallelized and Vectorized Tracking Using Kalman Filters with CMS Detector Geometry and Events. 2018.

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

[73] E Chabanat and N Estre. Deterministic Annealing for Vertex Finding at CMS. 2005.

[74] R Frühwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Adaptive Vertex Fitting. Technical Report CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.

[75] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017.

[76] The CMS collaboration. Performance of CMS muon reconstruction in pp collision events at sqrt(s) = 7 TeV. *Journal of Instrumentation*, 7(10):P10002–P10002, oct 2012.

[77] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[78] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[79] Identification of b quark jets at the CMS Experiment in the LHC Run 2. Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[80] A. M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13(05):P05011, 2018.

[81] Lior Rokach. Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1):1–39, Feb 2010.

[82] CMS e-Lab. Energy, Mass and Momentum in High-Energy Physics. `https://www.i2u2.org/elab/cms/library/ref-mass-energy-momentum.jsp`.

[83] Gregory Soyez. Pileup mitigation at the LHC: a theorist's view. *Phys. Rept.*, 803:1–158, 2019.

[84] N. Adam, J. Berryhill, V. Halyo, A. Hunt, K. Mishra. Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data . `http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2009_111_v1.pdf`.

[85]

[86] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[87] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl.Instrum.Meth.*, A434:435, 1999.

[88] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[89] SM Higgs Combination. Technical Report CMS-PAS-HIG-11-011, CERN, Geneva, 2011.

[90] S. Frixione, P. Nason, and C. Oleari. Matching nlo qcd computations with parton shower simulations: the powheg method. *JHEP*, 11:070, 2007.

[91] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[92] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. $HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO. *JHEP*, 10:083, 2013.

[93] Comparison of nuisances for background only case, 350 GeV mass hypothesis. http://rkamalie.web.cern.ch/rkamalie/feb12/Comparison_of_nuisances_expectedSignal0_3

[94] Comparison of nuisances for s+b case, 350 GeV mass hypothesis. http://rkamalie.web.cern.ch/rkamalie/feb12/Comparison_of_nuisances_expectedSignal1_3

[95] k-factor for DY/Z. https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns#D

[96] The LHC Higgs cross-section working group. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG.

[97] Standard Model Cross Sections for CMS at 13 TeV. https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSectionsat13TeVInclusi

[98] SM Higgs production cross sections at sqrt(s) = 13 TeV. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV#ZHllH

[99] NNLO+NNLL top-quark-pair cross sections. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO#Top_quark_pair_cross_se

[100] Single Top Cross sections. https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopSigma.

[101] CMS GEN XSEC Task Force. https://twiki.cern.ch/twiki/bin/viewauth/CMS/GenXsecTaskF

[102] SM Higgs Branching Ratios and Total Decay Widths (update in CERN Report4 2016). https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR#Higgs_2_g

[103] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. W Physics at the LHC with FEWZ 2.1. *Comput. Phys. Commun.*, 184:208, 2013.

[104] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP*, 09:126, 2007.

[105] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-

level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[106] Gunter Zech. Upper Limits in Experiments with Background Or Measurement Errors. *Nucl. Instrum. Meth.*, A277:608, 1989.

[107] A. L. Read. Presentation of search results: the CLs technique. *J. Phys. G: Nucl. Part. Phys.*, 28, 2002.

[108] Rikkert Frederix, Emanuele Re, and Paolo Torrielli. Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO. *JHEP*, 09:130, 2012.

[109] Johan Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J. C*, 53:473–500, 2008.

[110] Serguei Chatrchyan et al. Determination of jet energy calibration and transverse momentum resolution in CMS. *JINST*, 6:P11002, 2011.

[111] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205-206:10, 2010.

[112] Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J.*, C71:1547, 2011.

[113] Gael L. Perrin, Pedro Fernandez Manteca. Muon Identification and Isolation Scale-Factors on 2016 Dataset. `https://indico.cern.ch/event/611558/contributions/2465881/attachments/1407735/2151747/TnP_06_02_2017.pdf`.

[114] CMS Muon POG. Tracking SFs on the full 2016 data. https://twiki.cern.ch/twiki/bin/view/CMS/MuonWorkInProgressAndPagResults#Results_on_the_full_2016_data.

[115] Gael L. Perrin. Double Muon trigger efficiency per-leg approach. https://indico.cern.ch/event/636555/contributions/2577291/attachments/1453162/2241537/TnP_DoubleMuSF_03_05_17.pdf.

[116] Grace Dupuis. Collider Constraints and Prospects of a Scalar Singlet Extension to Higgs Portal Dark Matter. *JHEP*, 07:008, 2016.

[117] Oleg Antipin, David Atwood, and Amarjit Soni. Search for RS gravitons via W(L)W(L) decays. *Phys. Lett.*, B666:155–161, 2008.

[118] A. Liam Fitzpatrick, Jared Kaplan, Lisa Randall, and Lian-Tao Wang. Searching for the Kaluza-Klein Graviton in Bulk RS Models. *JHEP*, 09:013, 2007.

[119] Kaustubh Agashe, Hooman Davoudiasl, Gilad Perez, and Amarjit Soni. Warped Gravitons at the LHC and Beyond. *Phys. Rev.*, D76:036006, 2007.

[120] CMS Higgs PAG. List of question for the preapproval checks. https://twiki.cern.ch/twiki/bin/viewauth/CMS/HiggsWG/HiggsPAGPreapprovalChecks.

[121] CMS BTV POG. Supported Algorithms and Operating Points. https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco#Supported_Algorithms_and_Operati.

[122] Albert M Sirunyan et al. Search for resonant and nonresonant Higgs boson pair production in the bblnulnu final state in proton-proton collisions at sqrt(s) = 13 TeV. 2017.

[123] Abdus Salam and John Clive Ward. On a gauge theory of elementary interactions. *Nuovo Cim.*, 19:165–170, 1961.

[124] Steven Weinberg. A model of leptons. *Phys.Rev.Lett.*, 19:1264–1266, 1967.

[125] Albert M Sirunyan et al. Evidence for the Higgs boson decay to a bottom quark-antiquark pair. 2017.

[126] Aruna Kumar Nayak. Reconstruction of physics objects in the CMS detector. *PoS*, CHARGED2012:010, 2012.

[127] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[128] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[129] I. Belotelov, I. Golutvin, D. Bourilkov, A. Lanyov, E. Rogalev, M. Savina, and S. Shmatov. Search for ADD extra dimensional gravity in di-muon channel with the CMS detector. CMS Note 2006/076, 2006.

[130] M. Aldaya, P. Arce, J. Caballero, B. de la Cruz, P. Garcia-Abia, J. M. Hernandez, M. I. Josa, and E. Ruiz. Discovery potential and search strategy for the standard model Higgs boson in the $H \to ZZ^\star \to 4\mu$ decay channel using a mass-independent analysis. CMS Note 2006/106, 2006.

[131] A. Brandt et al. Measurements of single diffraction at $\sqrt{s} = 630$ GeV: Evidence for a non-linear $\alpha(t)$ of the pomeron. *Nucl. Phys. B*, 514:3, 1998.

[132] W. Buchmüller and D. Wyler. Constraints on SU(5)-type leptoquarks. *Phys. Lett. B*, 177:377, 1986.

[133] CMS Collaboration. CMS technical design report, volume II: Physics performance. *J. Phys. G*, 34:995, 2007.

[134] CMS Collaboration. Jet performance in pp collisions at $\sqrt{s}$=7 TeV. CMS Physics Analysis Summary CMS-PAS-JME-10-003, 2010.

[135] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[136] Particle Data Group, J. Beringer, et al. Review of Particle Physics. *Phys. Rev. D*, 86:010001, 2012.

[137] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavour Groups. Precision electroweak measurements and constraints on the Standard Model. 2010.

[138] I. Bertram, G. Landsberg, J. Linnemann, R. Partridge, M. Paterno, and H. B. Prosper. A recipe for the construction of confidence limits. Technical Report TM-2104, Fermilab, 2000.

[139] L. Moneta, K. Belasco, K. S. Cranmer, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf. The RooStats Project. In *13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010)*. SISSA, 2010. PoS(ACAT2010)057.

[140] Vardan Khachatryan et al. Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$. *Phys. Rev.*, D92(3):032008, 2015.

[141] Performance of muon identification in pp collisions at s**0.5 = 7 TeV. Technical Report CMS-PAS-MUO-10-002, CERN, Geneva, 2010.

[142] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.

[143] CMS COLLABORATION. Particle–flow event reconstruction in CMS and performance for jets, taus, and E_T^{miss}. CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CERN, 2009.

[144] CMS COLLABORATION. Commissioning of the particle-flow event reconstruction with the first lhc collisions recorded in the cms detector. CMS Physics Analysis Summary CMS-PAS-PFT-10-001, CERN, 2010.

[145] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06):P06005, 2015.

[146] CMS Collaboration. Search for pair production of first-generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV. Submitted to *Phys. Rev. Lett.*, 2010.

[147] CMS Collaboration. Performance of cms muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. Submitted to *J. Inst.*, 2012.

[148] ATLAS Collaboration. Search for the Higgs boson in the $H \rightarrow$WW(*)$\rightarrow \ell^+\nu\ell^-\bar{\nu}$ decay channel in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. Submitted to *Phys. Rev. Lett.*, 2011.

[149] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett. B*, 641:57, 2006.

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

[151] CMS Muon POG. Reference muon id, isolation and trigger efficiencies for Run-II. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2`.

[152] John M. Campbell, R. Keith Ellis, Paolo Nason, and Emanuele Re. Top-pair production and decay at NLO matched with parton showers. *JHEP*, 04:114, 2015.

[153] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order. *Comput. Phys. Commun.*, 182:2388, 2011.

[154] Ye Li and Frank Petriello. Combining QCD and electroweak corrections to dilepton production in FEWZ. *Phys. Rev. D*, 86:094034, 2012.

[155] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C*, 76(3):155, 2016.

[156] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[157] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012.

[158] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO single-top production matched with shower in POWHEG: s- and t-channel contributions. *JHEP*, 09:111, 2009. [Erratum: JHEP02,011(2010)].

[159] Michele de Gruttola, Caterina Vernieri, Pierluigi Bortignon, David Curry, Ivan Furic, Jacobo Konigsberg, Sean-Jiun Wang, Paolo Azzurri, Tommaso Boccali, Andrea Rizzi, Silvio Donato, Stephane Brunet Cooperstein , James Olsen, Christopher Palmer, Lorenzo Bianchini, Christoph Grab, Gael Ludovic Perrin, and Luca Perrozzi. Search for the Standard Model Higgs Boson Produced in Association with W and Z and Decaying to Bottom Quarks. http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2015/168.

[160] Michele de Gruttola, Rami Kamalieddin, Ilya Kravchenko, Lesya Shchutska. Search for resonant diHiggs production with bbZZ decays with the 2b2l2nu signature using 35.9/fb data of 2016 pp collisions at the LHC. http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2017_198_v17.pdf.

[161] Chris Palmer. VHbb Electron Trigger and ID+ISO SFs for 2016 data. https://indico.cern.ch/event/604949/contributions/2543520/attachments/1439974/2216426/VHbb_TnP_SFs_egamma_april.pdf#search=vhbb%20AND%20cerntaxonomy%3A%22Indico%2FExperiments%2FCMS%20meetings%2FPH%20%2D%20Physics%2FEgamma%22.

[162] CMS JetMET group. Jet Energy Resolution. https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution.

[163] CMS MET group. MET Corrections and Uncertainties for Run-II. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETRun2Corrections.

[164] CMS MET group. MET Filter Recommendations for Run II. https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2.

[165] CMS EGM POG. Multivariate Electron Identification for Run2. https://twiki.cern.ch/twiki/bin/viewauth/CMS/Multivari\ateElectronIdentificationRun2.

[166] Helge Voss, Andreas Höcker, Jörg Stelzer, and Frerik Tegenfeldt. TMVA, the toolkit for multivariate data analysis with ROOT. In *XIth International Workshop on Advanced Computing and Analysis Techniq ues in Physics Research (ACAT)*, page 40, 2007.

[167] CMS Higgs WG. Documentation of the RooStats -based statistics tools for Higgs PAG. https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit.

[168] CMS Higgs WG. Binned shape analysis with the Higgs Combination Tool. https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit#Binned_shape_analysis.

[169] bbbb team. Search for resonant pair production of Higgs bosons decaying to bottom quark-antiquark pairs in proton-proton collisions at 13 TeV. http://cms.cern.ch/iCMS/analysisadmin/get?analysis=HIG-17-009-pas-v5.pdf.

[170] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[171] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.

[172] C. Giunti and M. Laveder. Neutrino mixing. In F. Columbus and V. Krasnoholovets, editors, *Developments in Quantum Physics*. Nova Science Publishers, Inc., 2004.

[173] Savas Dimopoulos, Stuart Raby, and Frank Wilczek. Proton decay in supersymmetric models. *Physics Letters B*, 112(2):133 – 136, 1982.

[174] M. Della Negra, P. Jenni, and T. S. Virdee. Journey in the search for the higgs boson: The atlas and cms experiments at the large hadron collider. *Science*, 338(6114):1560–1568, 2012.

[175] Lyndon R Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001. 164 p, 2008. This report is an abridged version of the LHC Design Report (CERN-2004-003).

[176] Karsten Eggert, K Honkavaara, and Andreas Morsch. Luminosity considerations for the LHC. Technical Report CERN-AT-94-04-DI. CERN-LHC-Note-263. LHC-NOTE-263, CERN, Geneva, Feb 1994.

[177] Oswald Gröbner. The LHC Vacuum System. (LHC-Project-Report-181. CERN-LHC-Project-Report-181):5 p, May 1998.

[178] Thomas Lenzi. Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors. Master's thesis, U. Brussels (main), 2013.

[179] Prospects for HH measurements at the HL-LHC. Technical Report CMS-PAS-FTR-18-019, CERN, Geneva, 2018.

[180] Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-HIG-17-030, CERN, Geneva, 2018.

[181] The CMS collaboration. Missing transverse energy performance of the cms detector. *Journal of Instrumentation*, 6(09):P09001, 2011.

[182] Search for resonant double Higgs production with $bbZZ$ decays in the $b\bar{b}\ell\ell\nu\nu$ final state. Technical Report CMS-PAS-HIG-17-032, CERN, Geneva, 2018.

[183] Werner Herr and B Muratori. Concept of luminosity. 2006.

[184] Charles W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973.

[185] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014.

[186] Geoffrey I. Webb Claude Sammut. *Encyclopedia of Machine Learning*. Springer, 1st edition. edition, 2011.

[187] Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s}$ = 13 TeV. Technical Report CMS-PAS-HIG-17-030, CERN, Geneva, 2018.

[188] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys.*, G43:023001, 2016.

[189] Roger Barlow and Christine Beeston. Fitting using finite monte carlo samples. *Computer Physics Communications*, 77(2):219 – 228, 1993.

[190] J. S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra. In *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, pages 115–120, 2011.

[191] Eilam Gross. Practical statistics for high energy physics. *CERN Yellow Reports: School Proceedings*, 4(0):165, 2017.

[192] Andres Tiko. HH resonance searches at ATLAS and CMS. Tokyo, Japan, November 2018. "Higgs Couplings 2018" conference.

[193] A. M. Sirunyan, A. Tumasyan, and W. et al. Adam. Combination of searches for higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 122:121803, Mar 2019.

[194] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr. Nlo qcd+ew predictions for v + jets including off-shell vector-boson decays and multijet merging. *Journal of High Energy Physics*, 2016(4):21, Apr 2016.

[195] A.M. Sirunyan, R. Kamalieddin, and et al I. Kravchenko. Evidence for the higgs boson decay to a bottom quarkâĂ-antiquark pair. *Physics Letters B*, 780:501 – 532, 2018.