

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: $X \rightarrow HH \rightarrow b\bar{b}ZZ^* \rightarrow b\bar{b}\ell\ell\nu\bar{\nu}$. 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.

In the loving memory of my grandmother Raisa, who always believed in me, who showed me how to never lose hope, and Who taught me how to be diligent with my work, humble with people, and tough with myself.

8 July 1939 - 27 July 2018

Table of Contents

List of Figures	ix
List of Tables	xiii
1 Introduction	1
1.1 “All things are made of atoms”	1
1.2 A brief history of particle physics	4
1.3 Fundamental forces	6
1.4 The Brout-Englert-Higgs mechanism	11
2 Theory	18
2.1 Lagrangian formalism of the Standard Model	18
2.2 Double Higgs in Beyond the Standard Model Theories	24
3 The LHC and the CMS experiment	32
3.1 The Large Hadron Collider	32
3.1.1 The history of the LHC	32
3.1.2 The layout of the LHC	33
3.1.3 LHC infrastructure	34
3.1.3.1 Magnets	34

3.1.3.2	Cooling System	38
3.1.3.3	Radio Frequency Cavities	40
3.1.3.4	Vacuum System	41
3.1.4	LHC operations	43
3.1.5	Luminosity	45
3.2	The CMS experiment	46
3.2.1	The CMS coordinate system	51
3.2.2	The Inner Tracker	52
3.2.3	The ECAL	54
3.2.4	The HCAL	57
3.2.5	The Superconducting Solenoid	59
3.2.6	The Muon Tracker	60
3.2.7	The Triggers and DAQ	63
3.2.7.1	The L1 Trigger	63
3.2.7.2	The High-Level Trigger	66
3.2.7.3	The DAQ system	67
3.2.8	The CMS design	68
4	Physics Object Reconstruction in CMS	70
4.1	Track Reconstruction	70
4.1.1	Muon tracking	73
4.1.2	Electron tracking	74
4.1.3	Primary Vertex reconstruction	75
4.2	Particle level objects	77
4.2.1	Particle Flow links and blocks	77
4.2.2	Muons	78

4.2.3	Electrons and isolated photons	79
4.2.4	Hadrons and non-isolated photons	81
4.2.5	Jets and jet corrections	82
4.3	Other important physics quantities and objects	84
4.3.1	The b tagging and secondary vertices	84
4.3.2	Missing transverse momentum	86
4.3.3	Pileup interactions	88
4.3.4	Lepton isolation	88
4.3.5	Data-Monte Carlo corrections	89
4.3.5.1	Lepton efficiencies and the Tag-and-Probe method .	90
4.3.5.2	b tagging efficiency	91
4.4	Datasets and Trigger Paths	92
5	<i>bbZZ Physics Analysis</i>	95
5.1	Physics analysis overview	95
5.2	Data analysis strategy	97
5.3	Data and Triggers	98
5.3.1	Data	98
5.3.2	Triggers	98
5.4	Simulated Samples	103
5.4.1	Signal processes simulation	103
5.4.2	Background processes simulation	104
5.5	Physics Objects Selection	107
5.5.1	Electrons	107
5.5.2	Muons	108
5.5.3	Jets	108

5.5.4	b jets	109
5.5.5	Missing transverse momentum	109
5.6	Event Selection	110
5.6.1	Kinematic selection of physics objects	111
5.6.2	Signal candidate construction and selection	113
5.6.3	Signal and control kinematic regions	115
5.6.4	Signal region candidate selection with a multivariate technique	117
5.6.5	Candidate reweighting in simulated events	117
5.7	Multivariate selection in the signal region	120
5.7.1	Kinematic variables of a candidate	120
5.7.2	Multivariate discriminant: a BDT classifier	137
5.7.3	BDT selection requirement in the signal region	157
5.8	Uncertainties of the measurement	158
5.8.1	Normalisation uncertainties	159
5.8.2	Shape uncertainties	162
5.8.3	Leading uncertainties	164
5.9	Statistical Analysis	166
5.9.1	The likelihood function	166
5.9.2	The fit results	170
5.10	Discussion	179
6	Conclusions	182
Bibliography		184
References		184

List of Figures

1.1	The structure of an atom	3
1.2	All SM interaction and simple vertices	11
1.3	SM particles and force carriers	13
1.4	Coupling of particles to SM Higgs boson	14
1.5	SM Higgs boson production modes	15
1.6	Higgs boson decay channels	17
2.1	SSB Potential form	23
2.2	RS1 branes	25
2.3	SM double Higgs boson production	28
2.4	Double Higgs mass distribution and the total cross-section	29
2.5	BSM Resonant double Higgs decay in the 2 b, 2 lepton, and 2 neutrino final state.	30
2.6	Double Higgs decay channels	31
3.1	Schematic layout of the LHC.	33
3.2	LHC dipoles	36
3.3	LHC quadrupoles	37
3.4	13 kA high-temperature superconducting current lead.	38

3.5	LHC cryogenic states and the temperature scale.	40
3.6	RF cavities module.	41
3.7	Beam screen.	42
3.8	CMS experiment with the main sub-detectors.	48
3.9	The logo of the CMS experiment.	50
3.10	Coordinate system of the CMS detector	52
3.11	The inner tracker	53
3.12	The ECAL	55
3.13	The HCAL	58
3.14	The CMS superconducting solenoid	59
3.15	The CMS muon tracker	61
3.16	The CMS L1 trigger layout	64
4.1	Global combinatorial track finder efficiencies.	72
4.2	MET reconstruction in the CMS.	87
5.1	HLT trigger SFs for electrons.	101
5.2	HLT SFs for muons as a function of p_T and η , measured for eras B to G.	102
5.3	HLT SFs for muons as a function of p_T and η , measured for era H.	102
5.4	SFs of the dZ requirement for muons.	103
5.5	Optimisation of the MET selection for two $bbZZ$ analyses	112
5.6	Analysis phase space with the signal and two control regions.	116
5.7	Data-MC comparison in CRDY.	124
5.8	Data-MC comparison in CRDY.	125
5.9	Data-MC comparison in CRDY.	126
5.10	Data-MC comparison in CRDY.	127
5.11	Data-MC comparison in CRDY.	128

5.12 Data-MC comparison in the BDT sideband of the SR.	129
5.13 Data-MC comparison in the BDT sideband of the SR.	130
5.14 Data-MC comparison in the BDT sideband of the SR.	131
5.15 Data-MC comparison in the BDT sideband of the SR.	132
5.16 Data-MC comparison in CRTT.	133
5.17 Data-MC comparison in CRTT.	134
5.18 Data-MC comparison in CRTT.	135
5.19 Data-MC comparison in CRTT.	136
5.20 Data-MC comparison in CRTT.	137
5.21 BDT trees 1 and 2.	141
5.22 BDT trees 400 and 401.	142
5.23 BDT trees 799 and 800.	143
5.24 Input variable correlations of the di-electron channel, low mass training. .	147
5.25 Input variable correlations of the di-electron channel, high mass training. .	148
5.26 Variables used in the low mass training for di-electron channel.	149
5.27 Variables used in the high mass training for di-electron channel.	150
5.28 BDT discriminants for di-electron channel.	152
5.29 ROC curves for di-electron channel.	154
5.30 BDT distributions for the radion case.	156
5.31 Transverse mass of the reconstructed HH candidates for graviton hypothesis.	169
5.32 Transverse mass of the reconstructed HH candidates for radion hypothesis.	170
5.33 Data-MC comparison in the SR.	172
5.34 Data-MC comparison in the SR.	173
5.35 Data-MC comparison in the SR.	174
5.36 Data-MC comparison in the SR.	175
5.37 Data-MC comparison in the SR.	176

5.38 The expected and observed limits.	178
5.39 Combination of HH channels using 2016 data.	180

List of Tables

4.1	Triggers for dimuon and dielectron channels both at L1 and HLT levels.	93
5.1	List of data sets collected by the CMS in 2016.	99
5.2	Background Monte Carlo samples.	105
5.3	The MET requirements as a function of the mass of the HH candidate. .	113
5.4	Fraction of events surviving the candidate selection and kinematic requirements.	117
5.6	Di-electron channel. Relative importance of the input variables in the high mass BDT training.	145
5.5	Di-electron channel. Relative importance of the input variables in the low mass BDT training.	145
5.8	Di-muon channel. Relative importance of the input variables in the high mass BDT training.	146
5.7	Di-muon channel. Relative importance of the input variables in the low mass BDT training.	146
5.9	The BDT selection values used in this measurement.	157
5.10	Efficiency of the BDT selection.	158
5.11	The effect of leading systematic uncertainties, ee channel.	165

5.12 The effect of leading systematic uncertainties, $\mu\mu$ channel.	165
5.13 Expected event yields and observed data counts in the SR.	166
5.14 The expected and observed limits.	177

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).

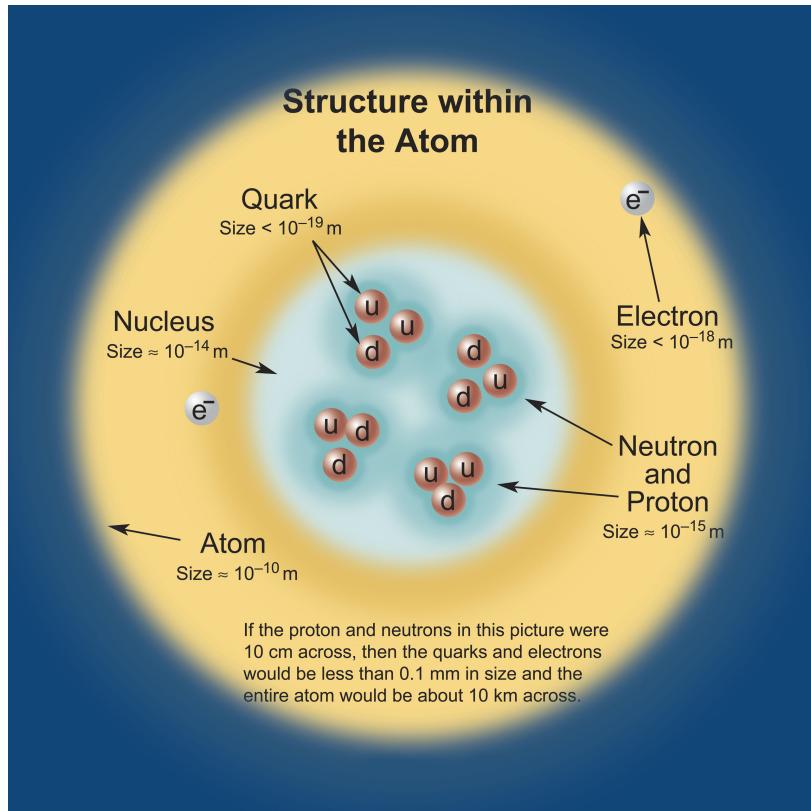


Figure 1.1: The structure of an atom. Approximate scale values are indicated.

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 applicable) 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 exist, 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 $m_Z = 91.189$ GeV².

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 Lagrangian³ 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.

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

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 , WWW , 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.

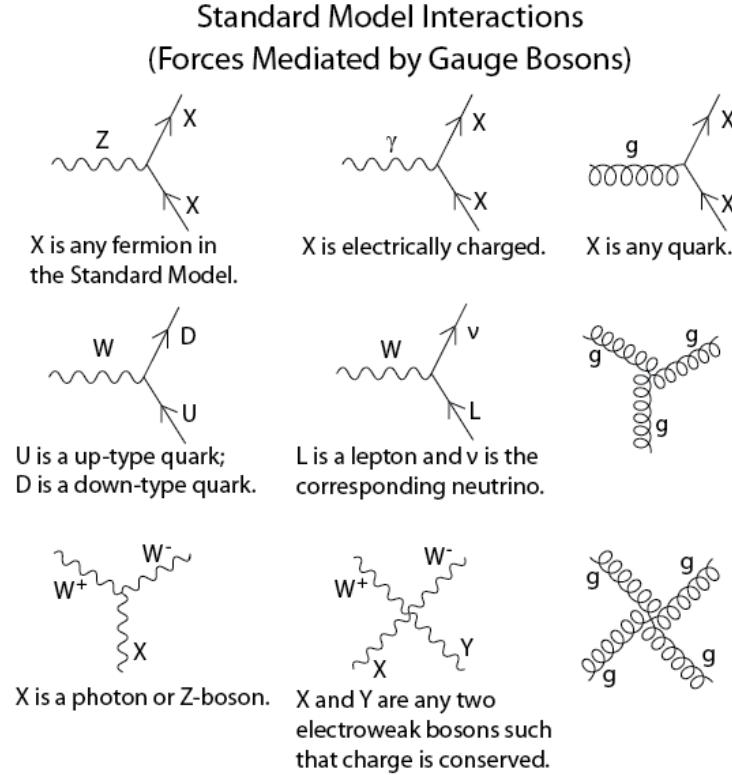


Figure 1.2: All SM interaction and simple vertices.

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]).

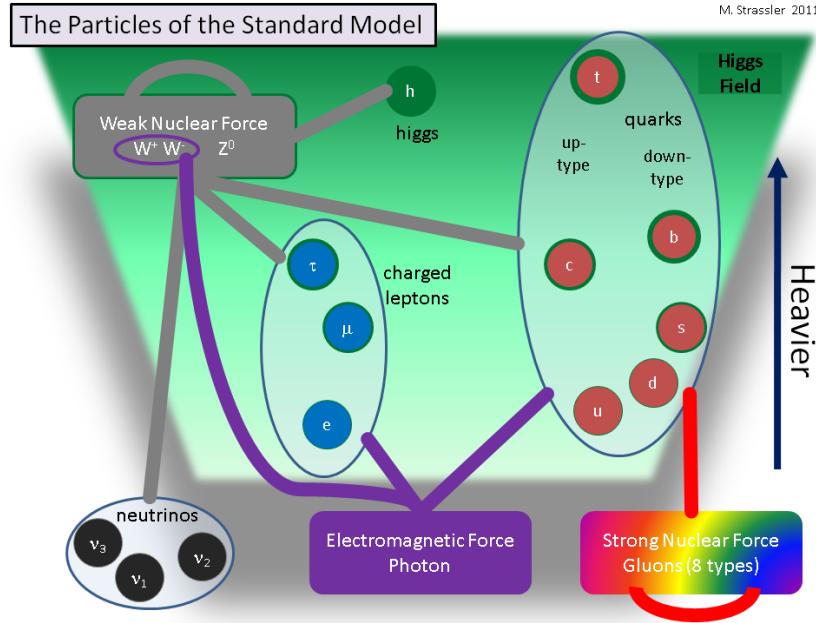


Figure 1.3: SM particles and force carriers. Self-interactions are also shown. The strength of the coupling to the Higgs boson increases from the bottom to the top, which is illustrated by the shades of the green color (the Higgs field).

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.

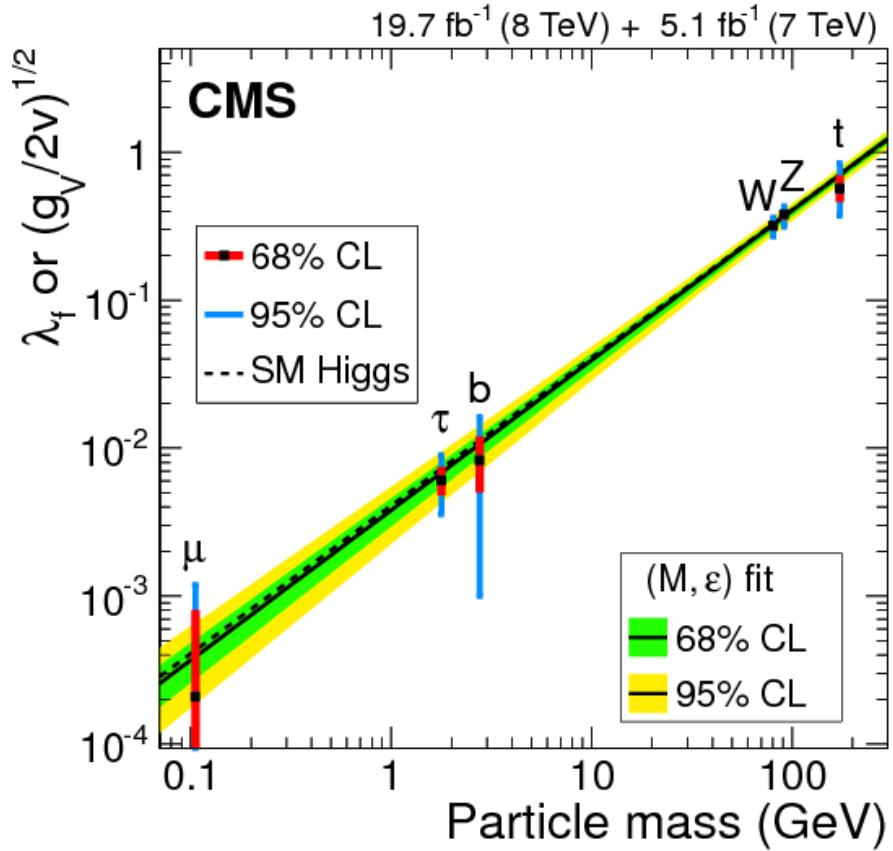


Figure 1.4: Coupling of particles to SM Higgs boson versus the mass of the particle, log-log scale is used. The y axis accommodates the couplings for both fermions and bosons.

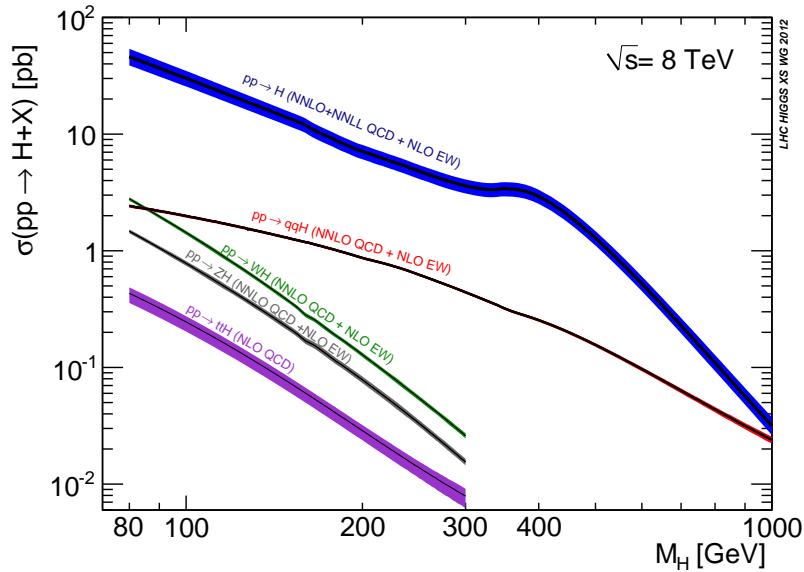
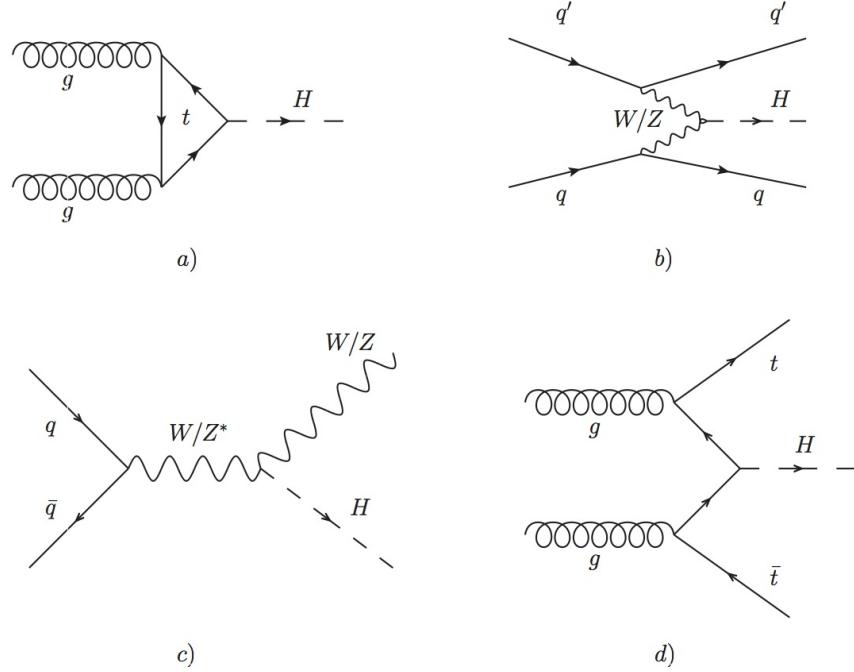


Figure 1.5: Top: SM Higgs boson production modes: a) a gluon fashion (in blue color at the bottom), b) a vector boson scattering (red), c) an associated production with a vector boson (green and brown), d) an associated production with the top anti-top pair (purple). Bottom: Higgs boson production modes 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, $H \rightarrow b\bar{b}$ and $H \rightarrow ZZ$. The first one has the highest branching fraction, while the second one gives a clean signature when subsequent $Z \rightarrow \ell\ell$ 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.

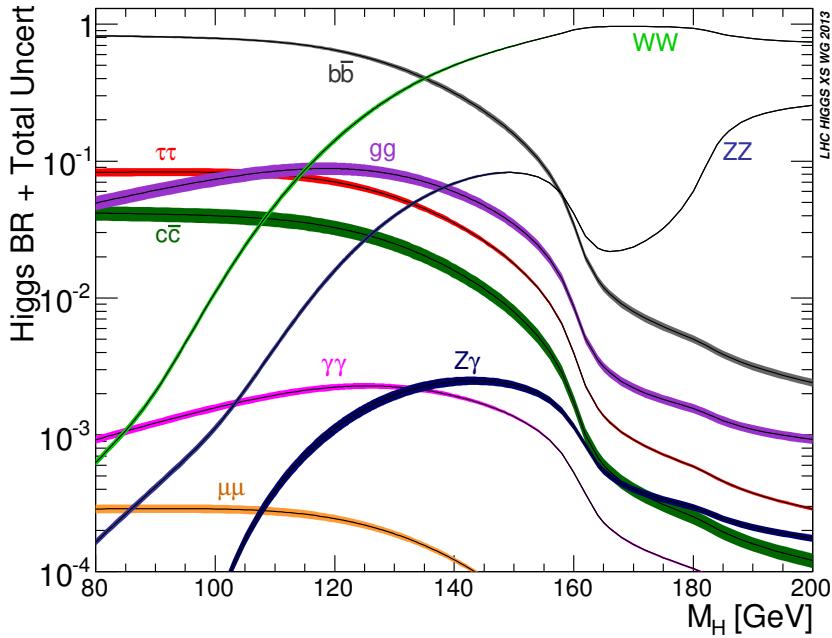


Figure 1.6: Higgs boson decay channels as a function of the Higgs boson mass. At 125 GeV the dominant decay mode is $H \rightarrow b\bar{b}$.

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]:

$$\mathcal{L}_{SM} = \mathcal{L}_{Yang-Mills} + \mathcal{L}_{ferm} + \mathcal{L}_H + \mathcal{L}_{Yuk} \quad (2.1)$$

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:

$$\mathcal{L}_{Yang-Mills} = -\frac{1}{4}W_{\mu\nu}^i(x)W_i^{\mu\nu}(x) - \frac{1}{4}B_{\mu\nu}(x)B^{\mu\nu}(x) - \frac{1}{4}G_{\mu\nu}^a(x)G_a^{\mu\nu}(x) \quad (2.2)$$

where

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

$$W_{\mu\nu}^i(x) \equiv \partial_\mu W_\nu^i(x) - \partial_\nu W_\mu^i(x) - g\varepsilon^{ijk}W_\mu^j W_\nu^k \quad (2.4)$$

$$G_{\mu\nu}^a(x) \equiv \partial_\mu G_\nu^a(x) - \partial_\nu G_\mu^a(x) - g_s f^{abc} G_\mu^b G_\nu^c \quad (2.5)$$

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 $a, b, c = 1, \dots, 8$. 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),

- $W_{\mu\nu}^i$ corresponds to $SU(2)_I$ symmetry of the weak isospin I_w^i . 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.
- $G_{\mu\nu}^a$ 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:

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

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:

$$D_\mu = \partial_\mu + igI_w^i W_\mu^i + ig'Y_w B_\mu + ig_s T_c^a G_\mu^a \quad (2.7)$$

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

fields:

$$\begin{aligned} A_\mu &= B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \\ Z_\mu &= -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W \end{aligned} \quad (2.8)$$

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

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

where

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

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:

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

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

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

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.

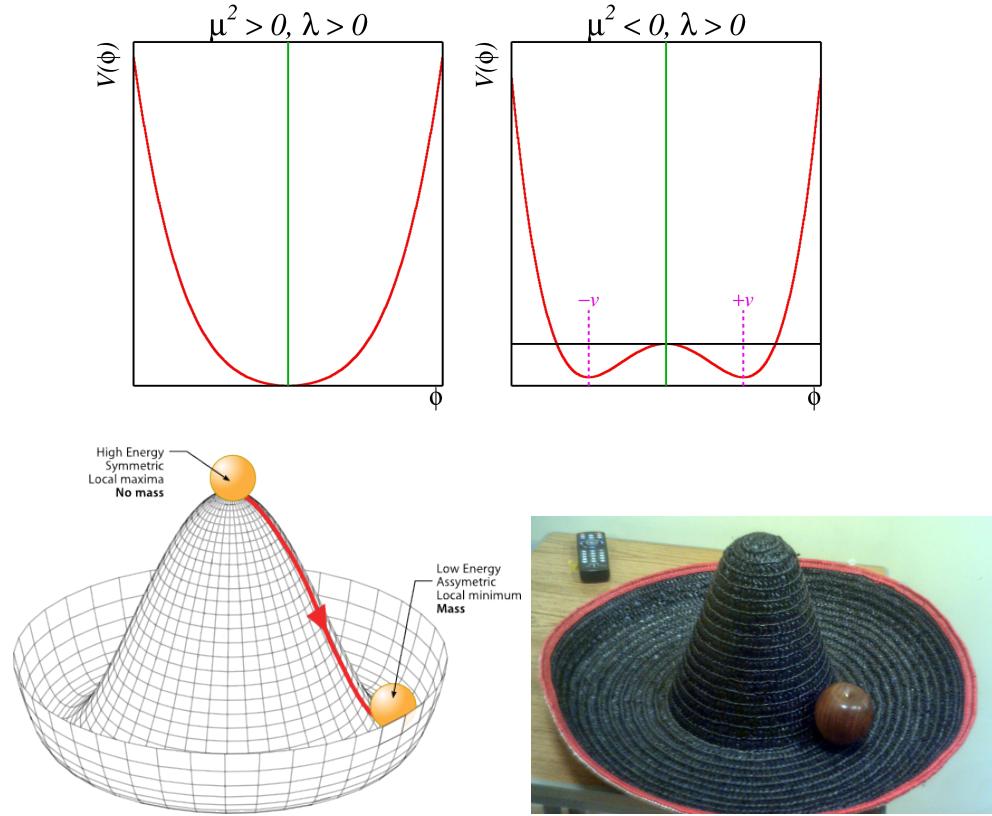


Figure 2.1: Top: Shape of the Higgs potential before and after the SSB that is determined at the leading orders by μ and λ parameters [35]. Bottom Left: Schematic drawing of the Higgs potential (after the SSB) that resembles a Mexican hat. Bottom Right: A real Mexican hat.

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 (RS1) [40], 4D space is an Effective Field Theory (EFT) approximation of the higher dimensional space. Randal

and Sundrum proposed two models. The first of the two models, called RS1 (used in this thesis), has a finite size for the extra dimension. This 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.

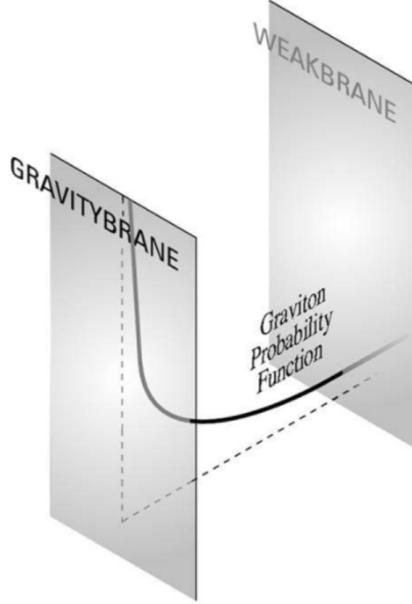


Figure 2.2: 5D space in the RS1 model [42].

The free parameters of the RS1 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 RS1 model matches the observations: the Higgs boson being closer (in the geometric sense in the fifth dimension) to the TeV brane and light fermions being located near the Planck brane,

see Fig. 2.2.

In the RS1 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 RS1 model. The graviton can propagate freely in the full higher-dimensional space of the 5D bulk. The other RS1 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 RS1 particles, the RS1 model considered in this thesis hypothesizes that SM particles are confined to branes. Another explanation of the lack of evidence of the RS1 particles at the LHC could be due to the fact that RS1 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 RS1 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 $\bar{M}_{Pl} = \sqrt{\frac{M_5^3}{k}} \cdot (1 - e^{-2\pi kr_c})$ is a reduced 4D M_{Pl} . Considered in this measurement, the gravi-

ton and radion are RS1 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 n^{th} 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

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

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 Lagrangian is given by:

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

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 RS1 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.

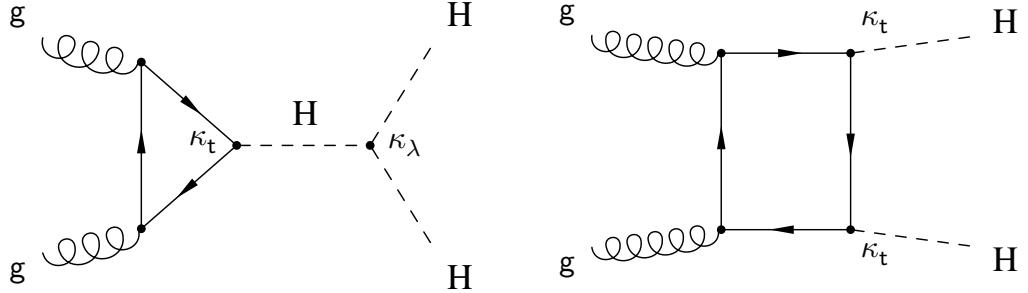


Figure 2.3: SM double Higgs boson production. Left: the triangle diagram with the virtual top quark loop. Right: the box diagram which dominates the overall HH production rate.

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.

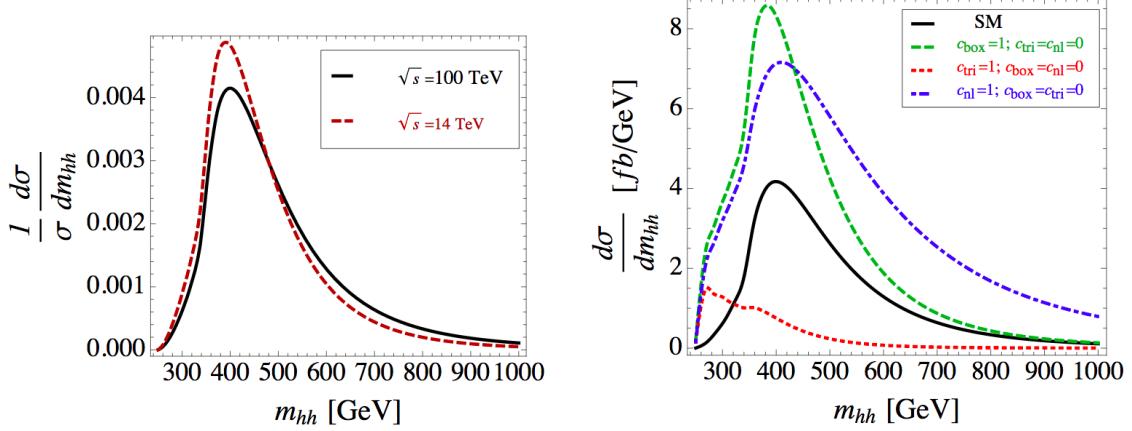


Figure 2.4: Left: comparison of the double Higgs boson mass distribution in the SM at the LO at 14 and 100 TeV center-of-mass energy. Right: the total SM HH cross section and the individual contributions [51]. Green refers to the SM box production, red refers to the SM triangle production, and blue refers 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 RS1 particles that can be observed at the LHC.

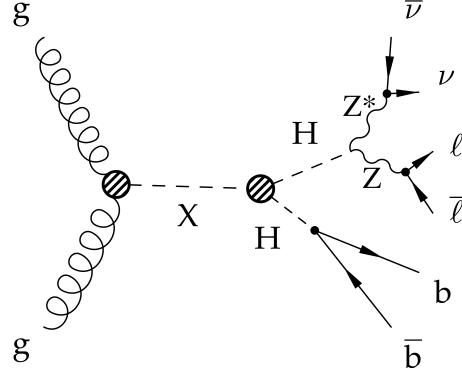


Figure 2.5: BSM Resonant double Higgs decay in the 2 b, 2 lepton, and 2 neutrino final state. X denotes either graviton or radion particles.

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.

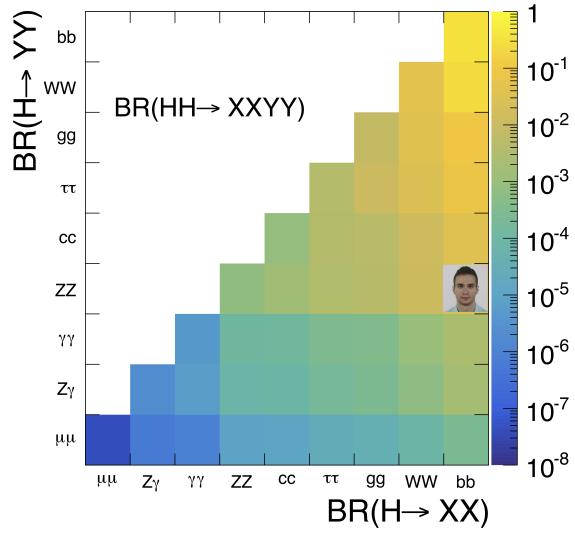


Figure 2.6: Double Higgs decay channels. The SM branching fractions are represented by the color palette. In this measurement $bbZZ$ decays are analysed, which are denoted on the map by the photo of the main $bbZZ$ analyser.

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 $10^{33-34} \frac{1}{\text{cm}^2 \text{s}}$.

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.

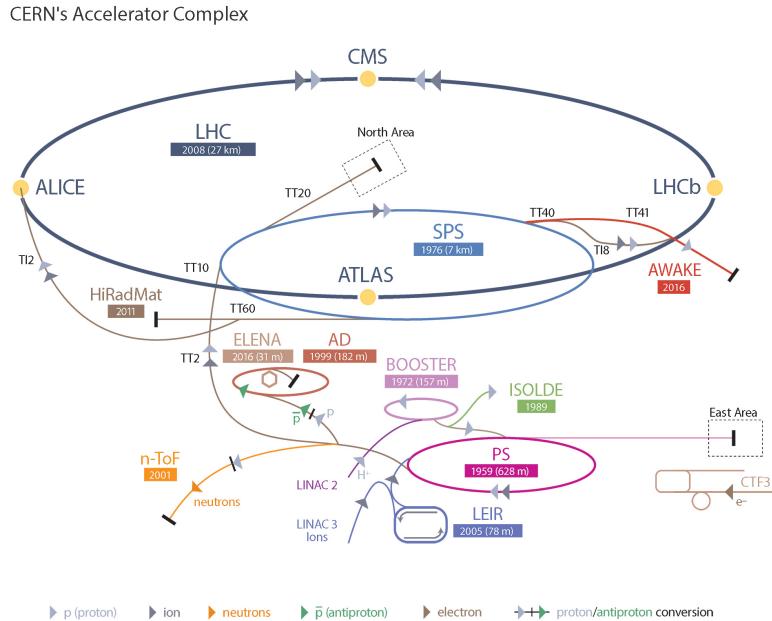


Figure 3.1: Schematic layout of the LHC.

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.

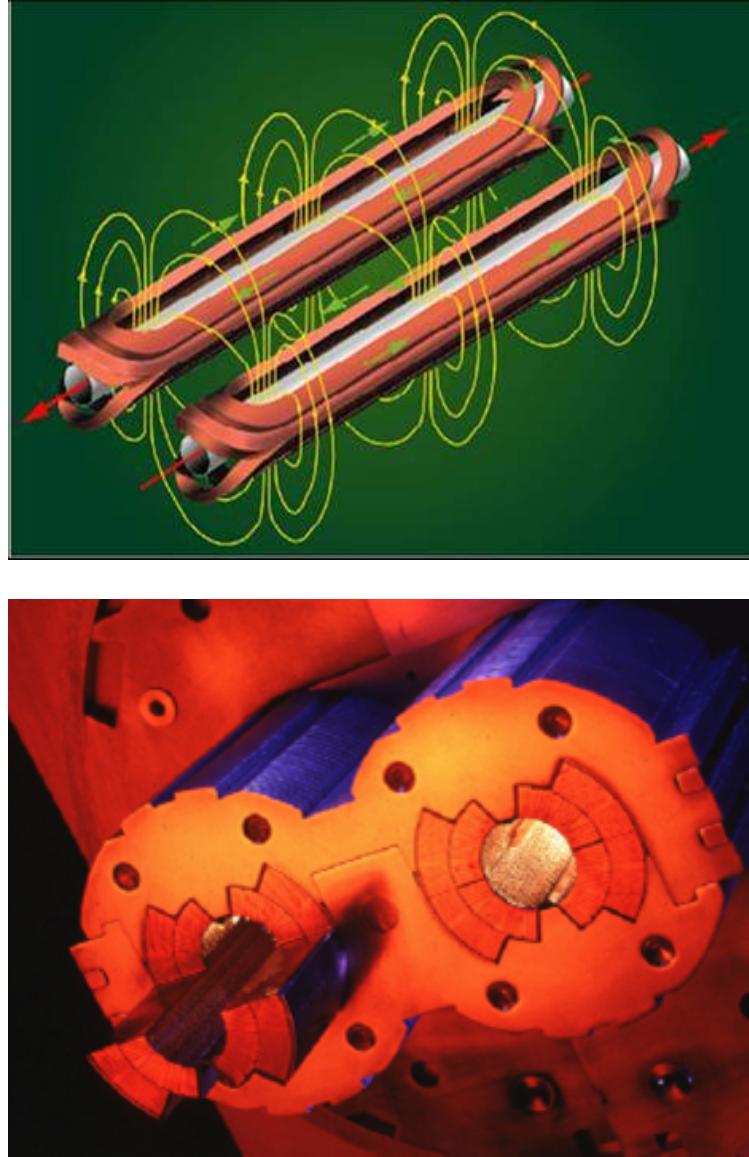


Figure 3.2: LHC dipole magnets. Top: two dipole coils and magnetic field lines. Bottom: two beam pipes with the coils inside of the dipole magnet. Each dipole magnet contains two magnetic configurations to steer the two proton beams travelling in opposite directions.

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.

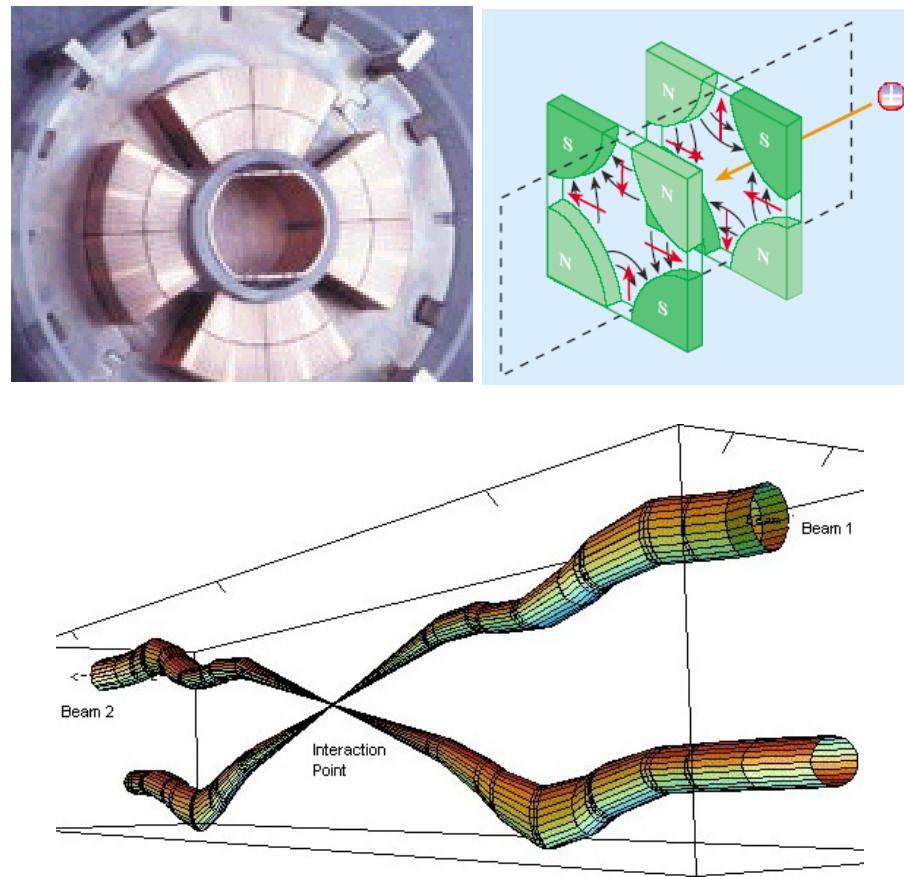


Figure 3.3: LHC quadrupoles. Top left: the coil of the quadrupole magnet. Top right: schematic view of the magnetic fields in the quadrupole. Bottom: two beams and the IP.

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.

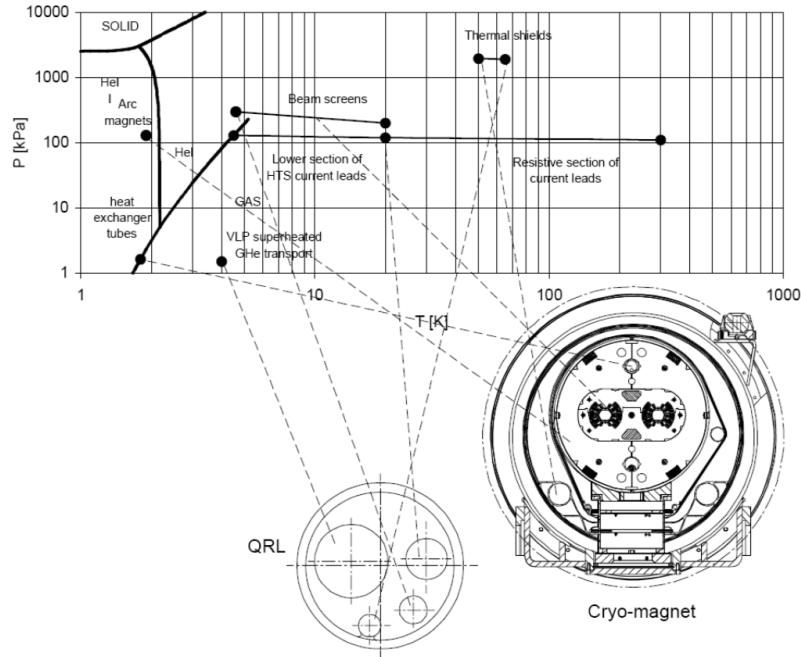


Figure 3.5: LHC cryogenic states and the temperature scale.

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).

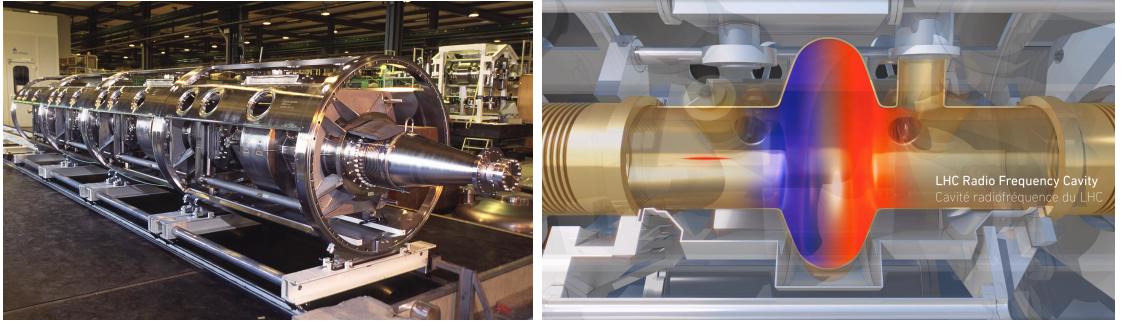


Figure 3.6: LHC RF cavities. Left: a cryomodule with four RF cavities. Right: a schematic drawing of a single RF cavity. The colour field is used to denote positive (red) and negative (blue) polarities. A narrow beam traversing the cavity is coming from the left and is shown in red.

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.

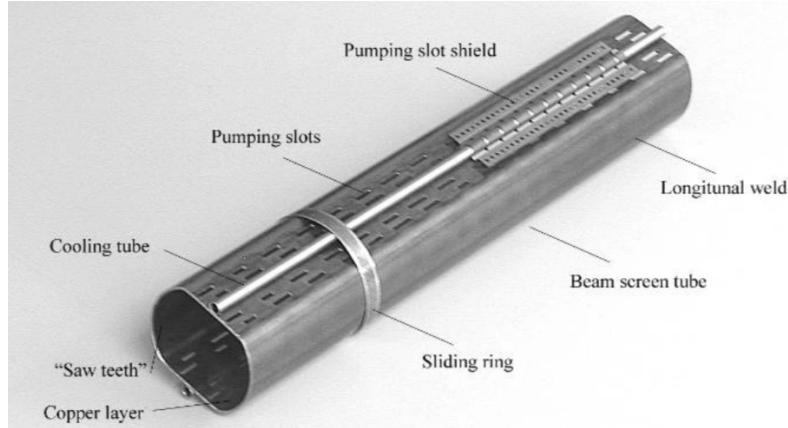


Figure 3.7: Beam screen.

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 $80 - 120 \text{ fb}^{-1}$ (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 $3000\text{ }fb^{-1}$.

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.

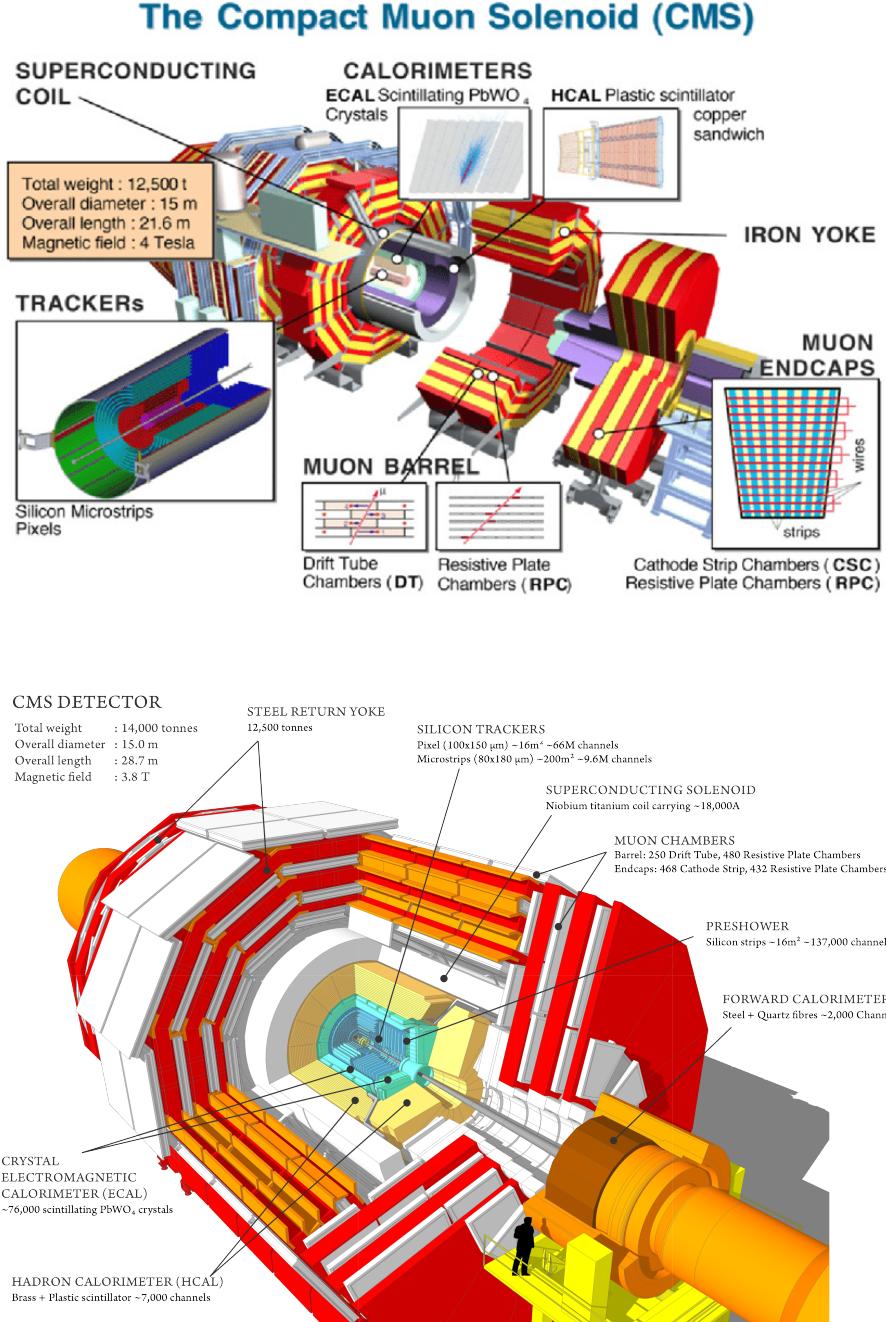


Figure 3.8: CMS experiment with the main sub-detectors.

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).

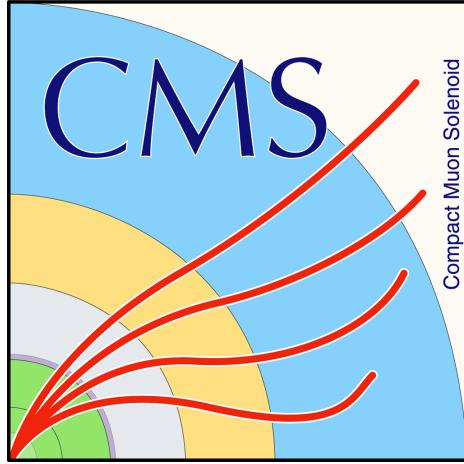


Figure 3.9: The logo of the CMS experiment that is showing curved trajectories of the emerging muons.

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 η 's and φ '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 $E_T = \sqrt{m^2 + p_T^2}$.

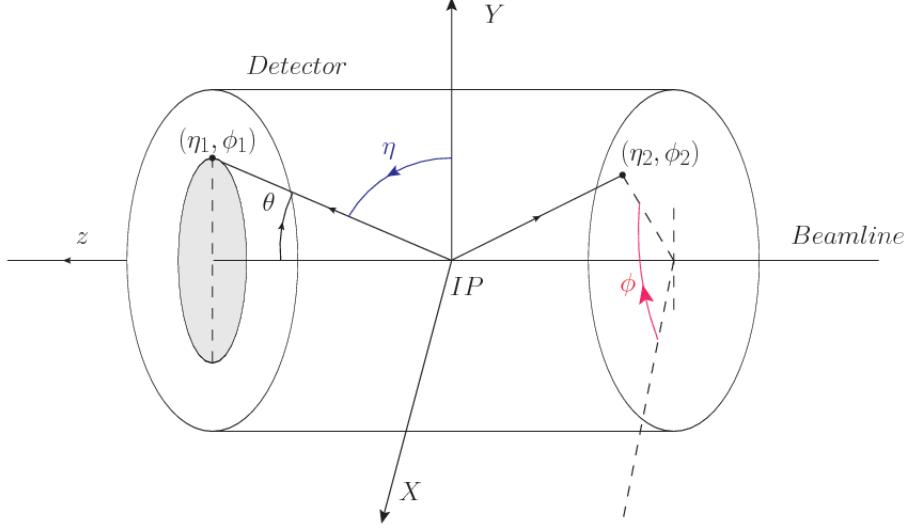


Figure 3.10: Coordinate system of the CMS detector [35]. Two particles (1 and 2) are shown with the corresponding angular variables ($\Delta\eta_1, \Delta\varphi_1$) for the first and ($\Delta\eta_2, \Delta\varphi_2$) for the second particle respectively.

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: $z\text{-}\varphi$ in the barrel and $r\text{-}\varphi$ in the endcaps.

The spatial resolution of each pixel is about 10 μm in the $r\text{-}\varphi$ 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.

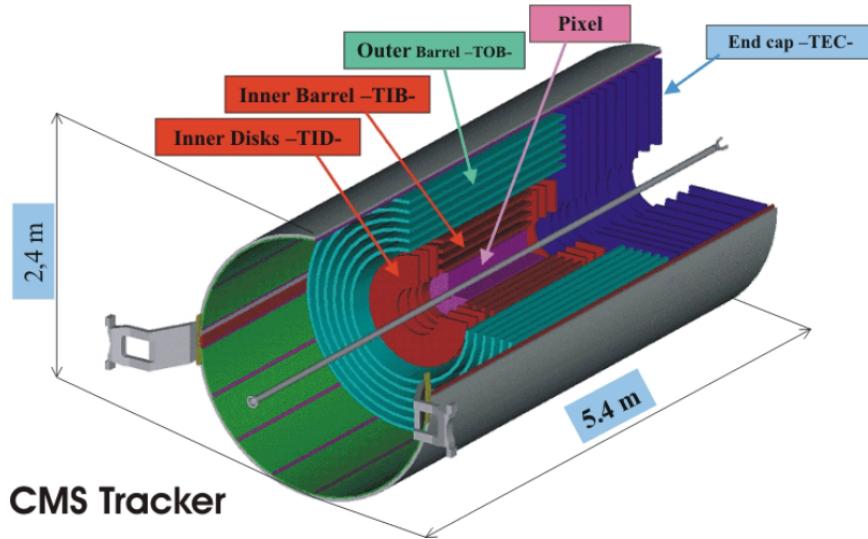


Figure 3.11: The inner tracker. The pixel and strip detectors are shown.

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 ($8.28\text{g}/\text{cm}^3$), a small radiation length ($X_0 = 0.89\text{ cm}$), a short Moliere radius ($R = 2.2\text{ cm}$), 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.

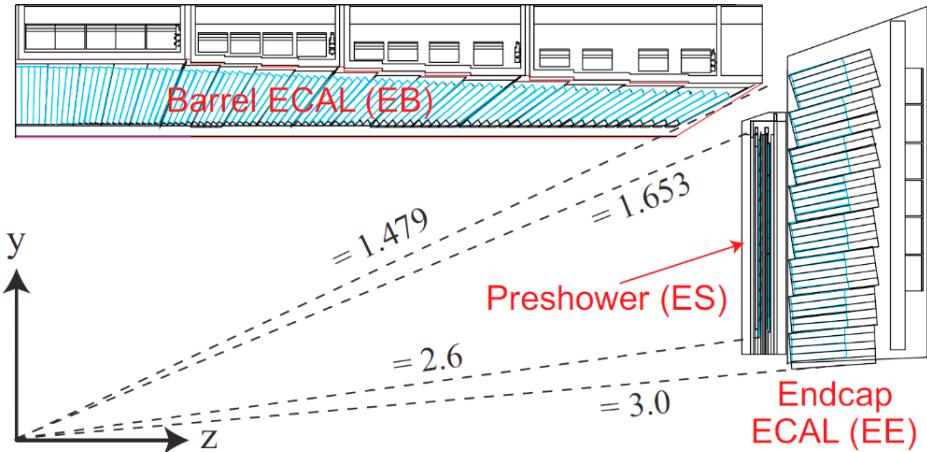


Figure 3.12: The ECAL and the Preshower detectors.

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:

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

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 $\pi^0 \rightarrow \gamma\gamma$. 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 $5.82 \lambda_I$ at $\eta = 0$ and almost $10 \lambda_I$ 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.

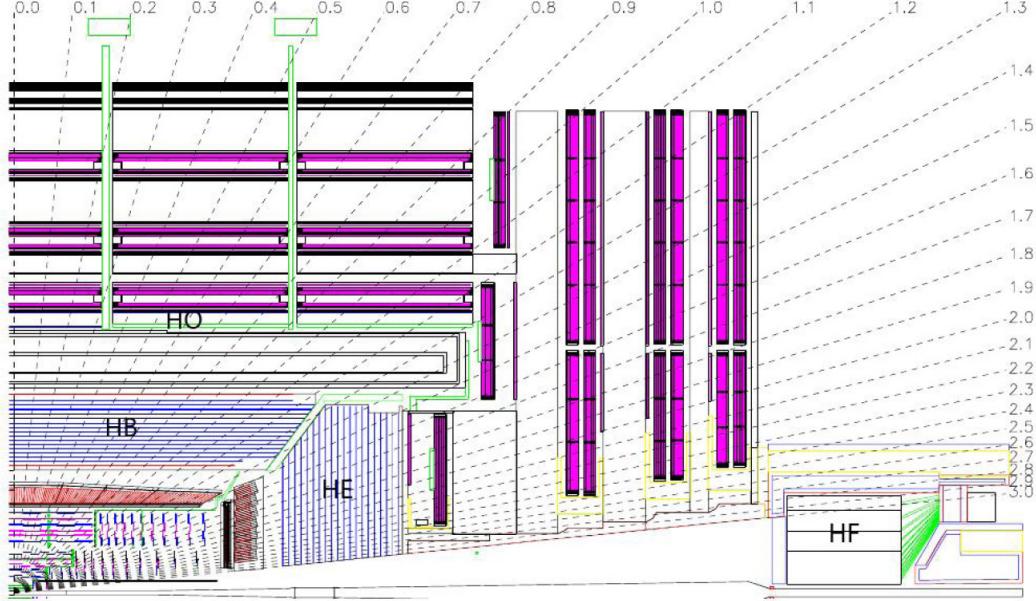


Figure 3.13: The HCAL with the η coverage map.

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 $11.8 \lambda_I$ 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 Superconducting 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.

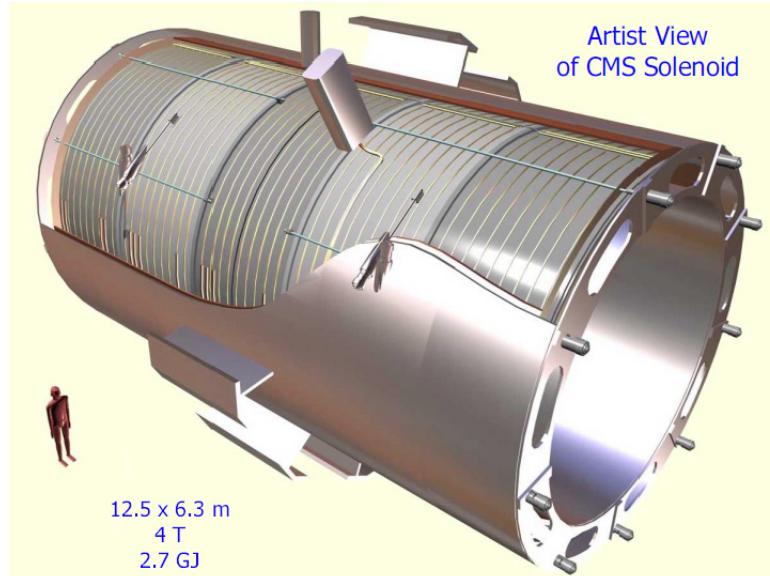


Figure 3.14: The CMS superconducting solenoid. The person on the left is shown to emphasise the size of the magnet.

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.

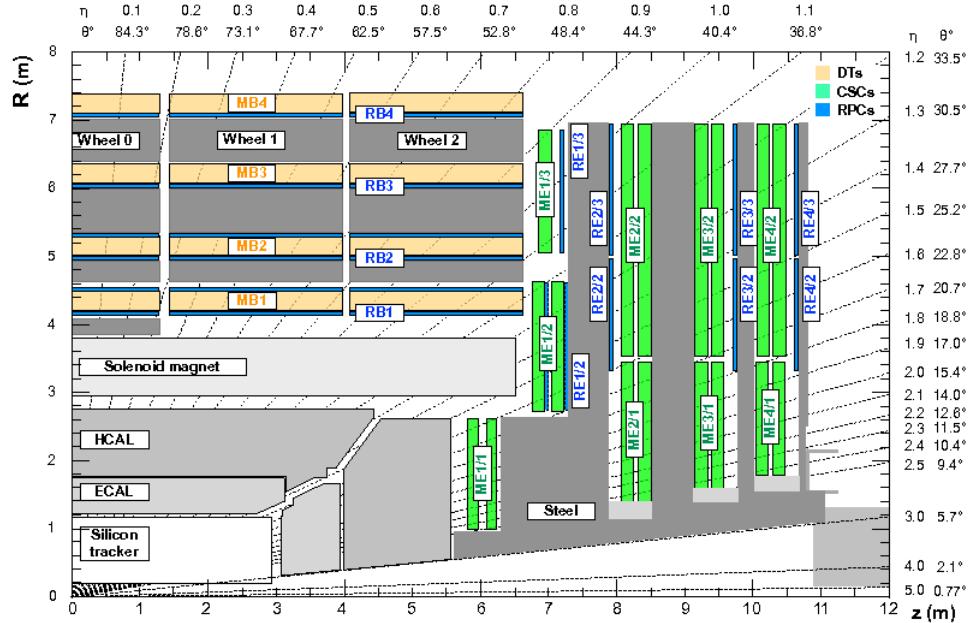


Figure 3.15: The CMS muon tracker. DT, CSC, and RPC detectors are shown in yellow, green, and blue respectively.

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 CO₂ 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, CO₂, and CF₄ 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 C₂H₂F₄, isobutane C₄H₁₀, and SF₆ 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 μ s. 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 $3.8 \mu s$. 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).

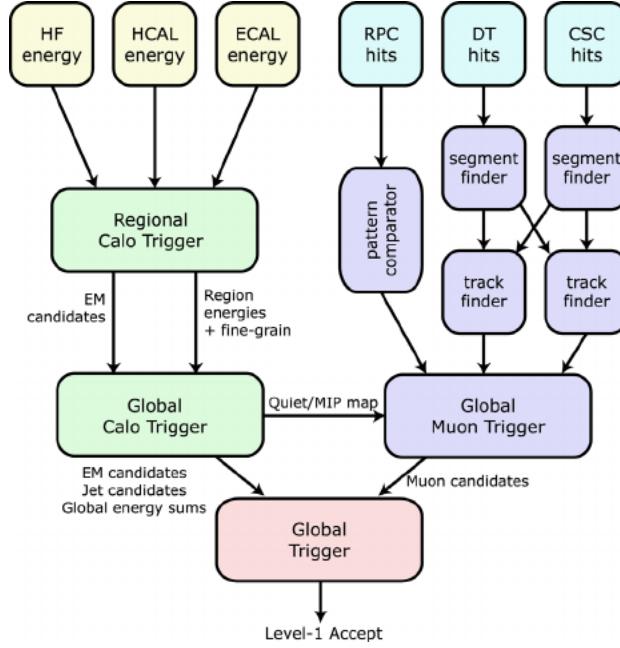


Figure 3.16: The CMS L1 trigger layout.

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 E_T or \not{p}_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 \cancel{E}_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 $Z \rightarrow \mu\mu$ and $H \rightarrow \mu\mu$) 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 $H \rightarrow \gamma\gamma$ discovery channel. Also, the ability to reject $\pi^0 \rightarrow \gamma\gamma$, 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:

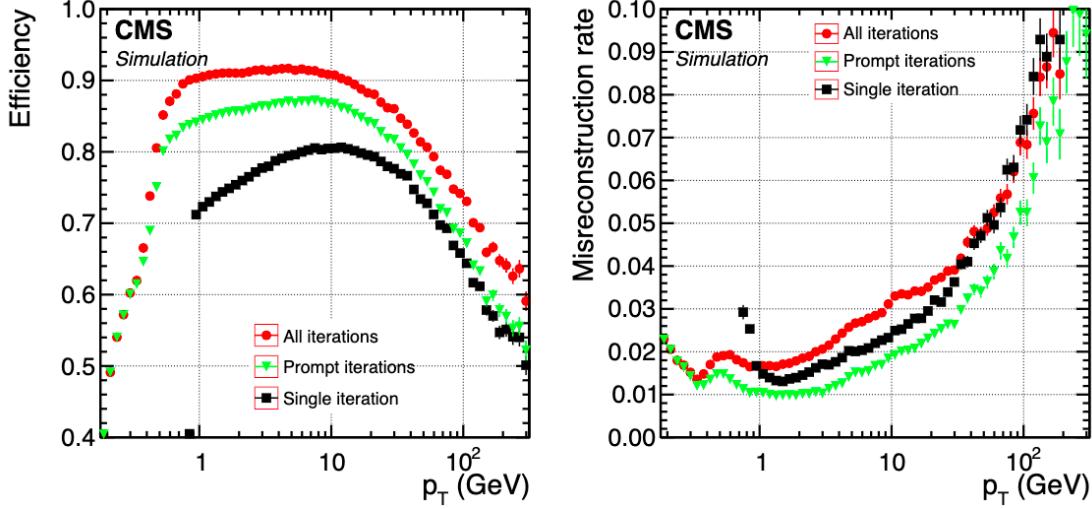


Figure 4.1: Global combinatorial track finder efficiencies as a function of the track p_T . Left: signal efficiency. Right: misidentification rate. Single iteration of the CMT is shown using black squares. The iterative tracking method consisting of 10 iterations is shown using green triangles and red 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 $Z \rightarrow e^-e^+$ 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:

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

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:

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

where $E_{\text{secondary}}$ 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 $c\tau \approx 400 - 500 \mu\text{m}$. 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” \cancel{p}_T (or missing transverse energy \cancel{E}_T (MET)), see Fig. 4.2. MET is constructed using all PF particles in the event and is calculated as:

$$\cancel{p}_T = \vec{p}_T^{miss} = | - \sum_i^N \vec{p}_{Ti} |.$$

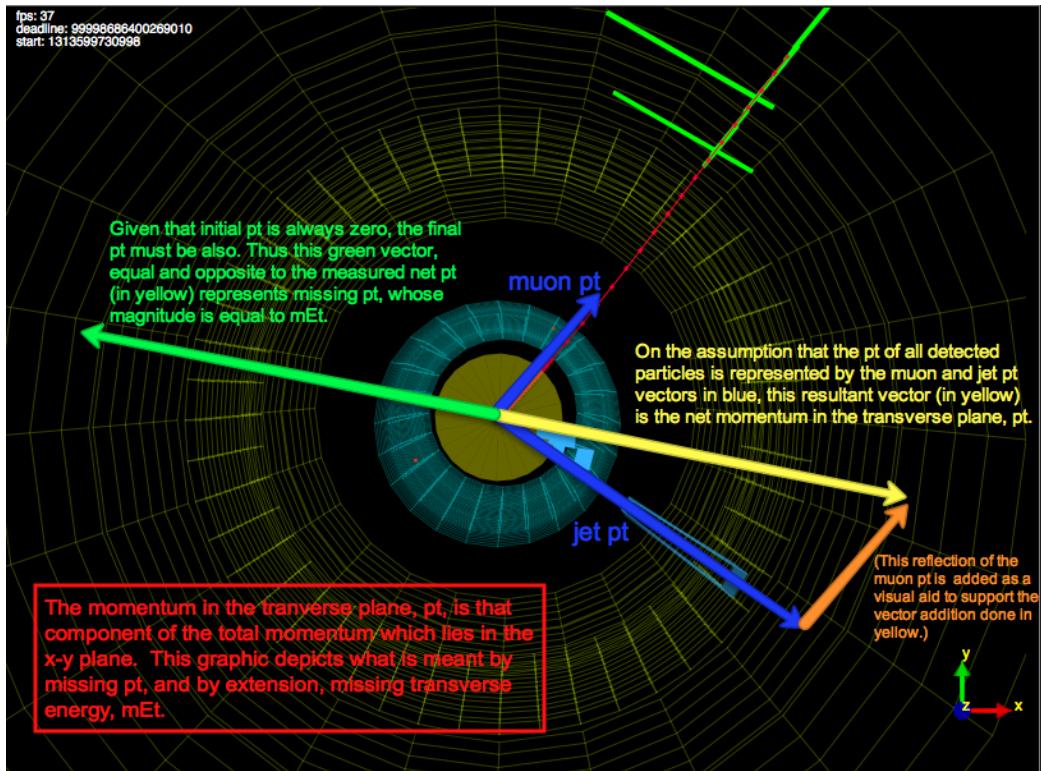


Figure 4.2: MET reconstruction in the CMS. The only detected (visible) particles are muon and jet both shown using blue arrows. The total net momentum “seen” by the CMS detector is represented by the yellow arrow. From the conservation of momentum, the reconstructed MET corresponds to the green arrow.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

To reconstruct $\cancel{\text{pt}}$, 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 $\cancel{\text{pt}}$ is still considered “raw”, since this MET is not yet JEC or JER corrected. After these corrections are applied, one obtains $\cancel{\text{pt}}$ that is called “Type-1 corrected MET”. This is the version of MET that is recommended by the CMS JetMET Physics 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 $\cancel{\text{pt}}$ 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}_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:

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

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 $Z \rightarrow e^-e^+$ or $Z \rightarrow \mu^-\mu^+$ 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 → 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	$\text{HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v}^*$ OR $\text{HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v}^*$ OR $\text{HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v}^*$ OR $\text{HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v}^*$
“Dielectron” $Z(ee)$ $Z(\nu\nu)H(b\bar{b})$	L1_SingleEG30 OR L1_SingleIsoEG22er OR L1_SingleIsoEG24 OR L1_DoubleEG_15_10	$\text{HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ}$

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.

CHAPTER 5

bbZZ Physics Analysis

In this chapter, we report on a search for the resonant production of a double Higgs boson system. We select Higgs boson pairs that subsequently decay through the $bbZZ$ channel. The final state consists of two b jets, two charged leptons, and two neutrinos ($2b2l2\nu$ final state). The analyzed data set was collected in 2016 by the CMS experiment in proton-proton collisions at 13 TeV COM energy, and corresponds to 35.9 fb^{-1} of integrated luminosity.

5.1 Physics analysis overview

We search for di-Higgs production through the gluon fusion mechanism mediated by two types of possible heavy resonances (separately): a spin-2 Randall-Sundrum (RS1) Kaluza-Klein (KK) graviton and a spin-0 RS1 radion [42, 50]. The width of the graviton and radion is assumed to be negligible with respect to the experimental resolution. We look for decays of WED particles (X) to the $bbZZ$ channel with the two b jets, two charged leptons, and two neutrinos in the final state. However, the $bbWW$ intermediate channel can also contribute (0(20)% to the final state of this measurement. Therefore, the full chain of this measurement is $X \rightarrow HH \rightarrow bbZZ/bbWW \rightarrow 2b2l2\nu$. In the above, one of the Higgs bosons decays to a pair of

b quarks and the other Higgs boson decays to ZZ or WW system. For the purposes of this measurement, we consider only leptonic decays of ZZ and WW. Intermediate taus are included when the decay to daughter muons and electrons. We explore the invariant masses of WED particles ranging from 250 to 1000 GeV. The result of the measurement is the upper limit on the production cross section of the resonance multiplied by the branching fraction of its decay into the aforementioned $2b2l2\nu$ final state.

In this data analysis (later referred to as the “analysis”), we will describe first the data sets (often referred to as a “datasets”) and data triggers. The measurement uses the DoubleEG and DoubleMuon PDs defined in the previous chapter. These di-electron and di-muon channels are analyzed separately and the information is combined in the final result. To select the events with two prompt charged leptons, a set of triggers was applied, when the events were recorded. To increase the statistics of the measurement and maximize the number of leptons passing the selection, certain complex trigger strategies were employed.

We then discuss the MC simulation of the signal and background processes (later referred to as “signal” and “background”). The signal MC samples had to be produced for both graviton and radion particles for 16 mass points from 250 to 1000 GeV to cover the whole search range.

The description of the physics object reconstruction and event selection was given in the previous chapter. Physics objects are constructed using the information from the CMS subsystems and the output of the PF algorithm. Then, based on the final state signature, the events containing corresponding physics objects are selected. The construction of the Higgs and Z boson candidates is discussed at length. The characteristics of the signal and background processes are specified and data-MC SFs are discussed.

A signal and background control regions are defined in the kinematic space of the candidates. The MVA technique is used to improve signal-background discrimination. We rely on the BDT classifier (commonly referred to as the BDT for brevity) to reduce the contribution of background processes in the signal region. In all physics measurements, there are sources of systematic and statistical uncertainties that affect the final results. We discuss all major uncertainties at length and specify the size of the effect of individual types of uncertainties on the final result.

We present the statistical analysis used to extract the results of the measurements and discuss the CMS statistical package that has been used to extract the final results. Then, we present the results of the measurement and compare them with the theoretical predictions. Finally, we discuss the “grand HH combination” using measurements of all available HH channels.

The material in this chapter follows the description of two articles to which the author contributed directly [85, 86].

5.2 Data analysis strategy

We form the Lorentz energy-momentum vector associated with the double Higgs system. This vector is constructed as the sum of the Lorentz vectors of the two leptons, two b jets, and the four-vector representing neutrinos (\vec{p}_T^{miss}). As the z component of the neutrinos’ momentum is unknown, we form a pseudo-transverse mass: $\tilde{M}_T(HH) = \sqrt{E^2 - p_z^2}$ (further referred to as “transverse mass” for brevity), where E and p_z are the energy and the z-axis component of the Lorentz energy-momentum vector of the di-Higgs candidate. Because p_z of the neutrinos is unknown and varies from event to event, the variable $\tilde{M}_T(HH)$ has a distribution, the shape of which depends on the mass of the intermediate resonance and how much energy is

carried by the neutrinos.

$\tilde{M}_T(\text{HH})$ distributions are derived for 16 masses and for both spin hypotheses. The $\tilde{M}_T(\text{HH})$ is constructed using signal and background simulations. The distributions are produced separately for the signal-enriched (SR) and control regions (CR), which are defined later in this chapter.

To extract the final results, we perform a simultaneous fit of the sum of the signal and background $\tilde{M}_T(\text{HH})$ distributions in the SR and CRs to the data. The fits are produced for both di-electron and di-muon channels and the results are combined. We obtain 95% CL upper limits on the HH production cross section multiplied by the BFs of the subsequent decay to the final state of this measurement. The employed statistical method is based on the CL_S asymptotic procedure [87].

5.3 Data and Triggers

5.3.1 Data

The search is performed using DoubleMuon and DoubleEG PDs recorded by CMS in 2016. The data were collected in several data taking periods and approximate values of the integrated luminosity of these periods are given in Table 5.1.

5.3.2 Triggers

The data events that are used in this measurement are selected with a set of HLT triggers that required the presence of two muons or two electrons in the event. In the di-electron final state, the trigger requires the leading electron to have p_T above 23 GeV, and the sub-leading electron to have p_T above 12 GeV. In offline selection p_T selection is increased to 25 and 15 GeV correspondingly. The final state that contains

Table 5.1: List of data sets collected by the CMS in 2016. Each era contains a unique letter identifier. Also, the date of the data processing is specified. If re-processing was run, it is labeled as “v2”. Corresponding integrated luminosities are shown in the second column.

Dataset	$\int \mathcal{L} dt (fb^{-1})$
Run2016B-03Feb2017-v2	~ 5.9
Run2016C-03Feb2017-v1	~ 2.7
Run2016D-03Feb2017-v1	~ 4.3
Run2016E-03Feb2017-v1	~ 4.1
Run2016F-03Feb2017-v1	~ 3.2
Run2016G-03Feb2017-v1	~ 3.8
Run2016H-03Feb2017-v1	~ 11.8
Total Luminosity	~ 35.9

two prompt muons is selected with an HLT path which is a combination of several HLT paths chained together using the logical “OR” operation. In other words, the muons are required to pass the selection of at least one of the HLT paths. At the HLT level, the leading and sub-leading muons have to pass the 17 and 8 GeV p_T requirements respectively. The difference among HLT paths has been explained at length in Section 3. The offline analysis increases the p_T threshold values to 20 and 15 GeV correspondingly. For the offline selection, electrons are selected in the range $|\eta| < 2.5$ and muons in the range $|\eta| < 2.4$. For both channels, the $|\eta|$ region in the gap (1.4442 to 1.566) between the barrel and endcap is excluded.

The same trigger selections are applied to MC simulated events to mimic the selection in data. The efficiencies are then derived for MC and data, and SFs are determined using the TnP procedure discussed in Section 4.3.5.1. Trigger SFs have been computed for each trigger leg separately since the selection of each leg varies (Fig. 5.1). Following the recommendations from the CMS Muon POG, scale factors have been calculated separately for two data collecting periods: runs B to G (Fig. 5.2, and run H separately (Fig. 5.3), since the LHC conditions varied significantly for the run H. All eras have slightly different integrated luminosities, see Table 5.1, so the final

SFs are luminosity averaged. Additionally, as was discussed in the chapter on CMS Physics Objects Reconstruction, some triggers did not contain the DZ requirement, while others did. Therefore, scale factors of the DZ selection are also measured, see Fig. 5.4). Prior to measuring the trigger scale factors, the electrons and muons were required to satisfy ID and ISO selections; more details are given in the subsections 5.5.1, 5.5.2.

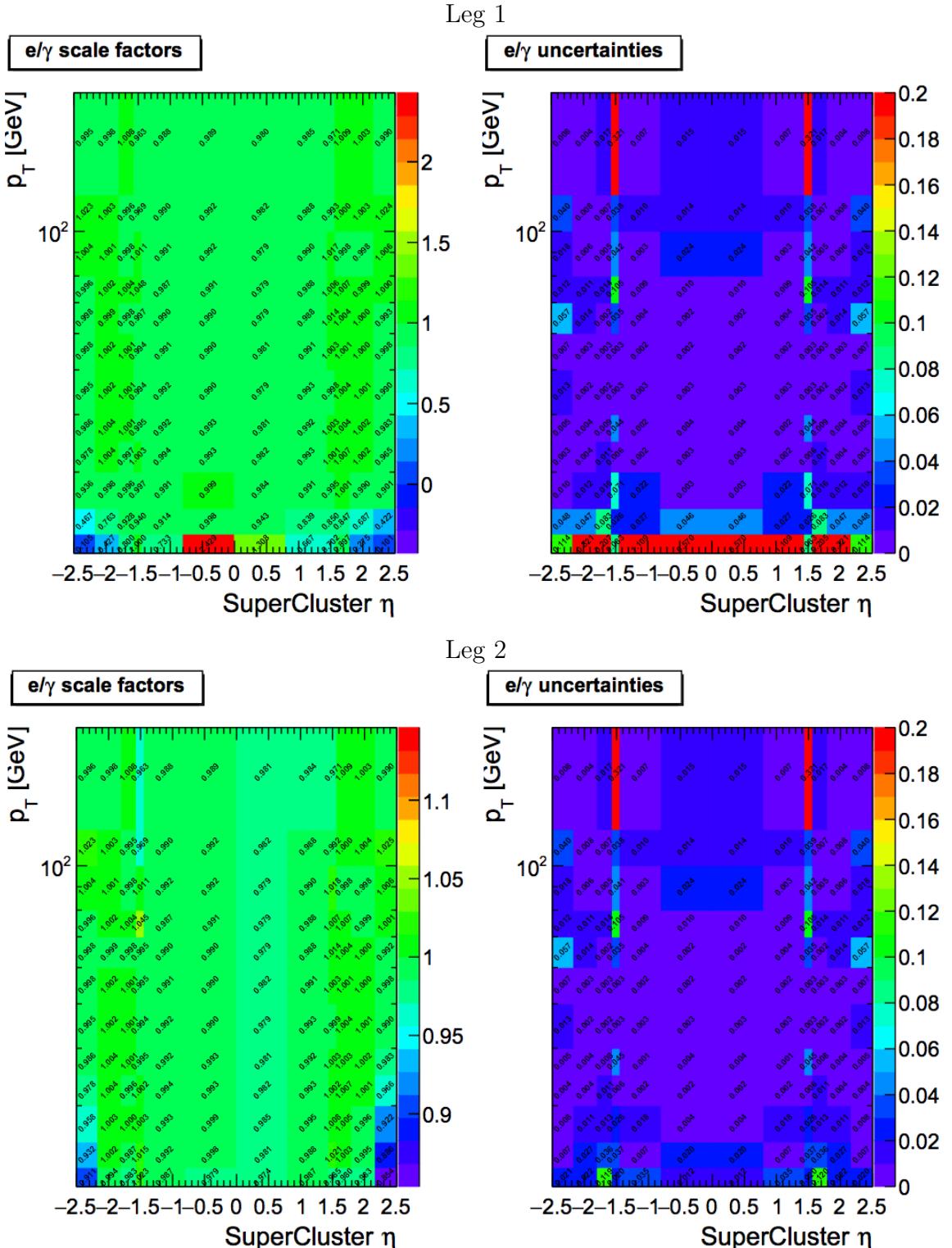


Figure 5.1: HLT trigger SFs for electrons approved by the CMS e/γ POG group. SFs are derived for both legs separately: Leg 1 (top) corresponds to the leading electron and Leg 2 (bottom) corresponds to the sub-leading electron. The values of the SFs are shown on the left, and the associated uncertainties with each value are shown on the right. Taken from [88].

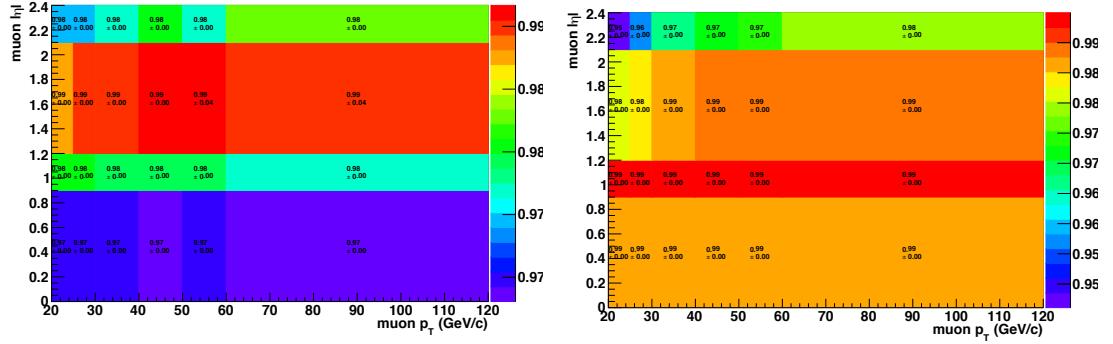


Figure 5.2: Final HLT SFs for muons as a function of p_T and η , measured for eras B to G. Left: Scale factors for 8 GeV leg (sub-leading muon). Right: Scale factors for 17 GeV leg (leading muons), provided that the sub-leading muon already passed the 8 GeV p_T requirement.

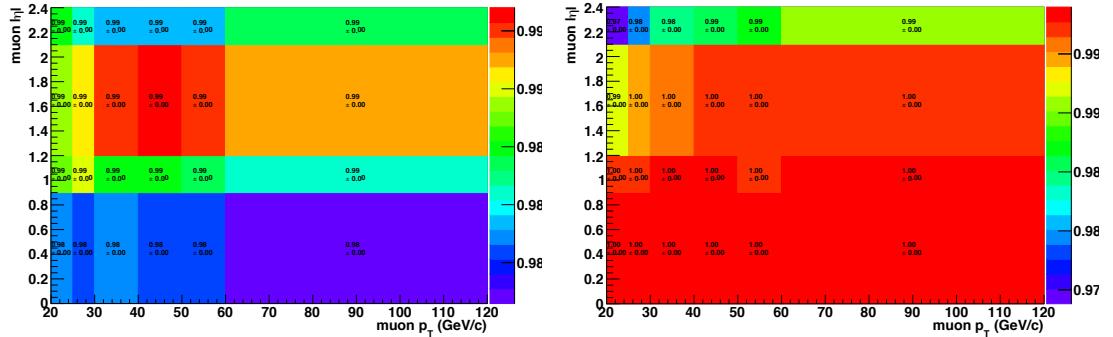


Figure 5.3: Final HLT SFs for muons as a function of p_T and η , measured for era H. Left: Scale factors for 8 GeV leg (sub-leading muon). Right: Scale factors for 17 GeV leg (leading muons), provided that the sub-leading muon already passed the 8 GeV p_T requirement.

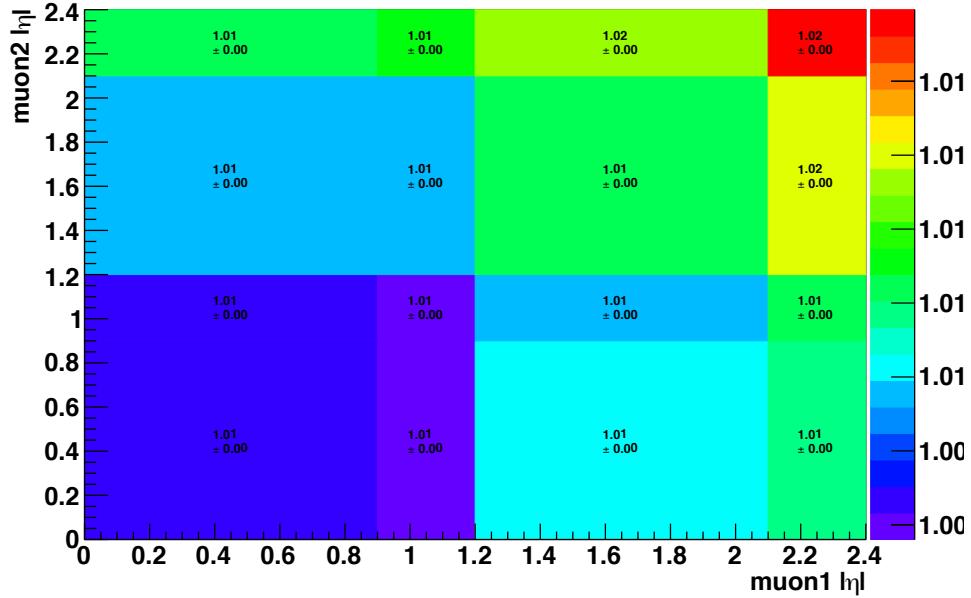


Figure 5.4: SFs of the dZ requirement for muons as a function of the η values of both muons.

5.4 Simulated Samples

The data analysis carried out in this thesis is optimized using the MC simulation. MC samples for signal and background processes have been produced with various HEP software packages (“generators”) that generate the processes of interest: processes that create the final state of this measurement.

5.4.1 Signal processes simulation

The signal Monte Carlo (MC) samples have been generated using the `MADGRAPH5_aMC@NLO` [89] package. In these samples, the process of the gluon fusion production of WED spin-0 and spin-2 narrow resonances is simulated at leading order (LO). The production of resonances is followed by the decay of the resonances in the double Higgs boson system. All Higgs bosons are assumed to be SM Higgs bosons with an invariant

mass of 125 GeV. Samples are generated for two spin hypotheses and 16 mass values covering the range of heavy resonance masses from 250 to 1000 GeV. Two types of signal samples are present: resonance decaying to $2b2l2\nu$ final state through the $X \rightarrow HH \rightarrow bbZZ$ decays and also through the $X \rightarrow HH \rightarrow bbWW$ decays. In both samples, the first Higgs boson decays to a pair of b quarks. However, in the first sample, the other Higgs boson decays to ZZ pair, while in the second sample the other Higgs boson decays to WW pair. Only Z boson decays in a dielectron, a dimuon, or a two neutrino state are selected. For W bosons, the chosen signature is characterized by a W boson decay to an electron and an anti-electron neutrino or a muon accompanied by an anti-muon neutrino. Intermediate decays to tau leptons with the subsequent decays to electrons and muons are also included.

To compare the expected numbers of events in the simulation to the number of observed events in the data for a given integrated luminosity, the signal production cross section has been normalized to 2 pb. This is a typical theoretical value of the production cross section of the WED particle in the 250-300 GeV range, the range to which the current physics analyses are very sensitive with the available LHC data. Additionally, the computed event rates take into account the branching fractions of the corresponding di-Higgs decay chains to the final state: 0.0012 and 0.0266 for $HH \rightarrow bbZZ \rightarrow bb\ell\ell\nu\nu$ and $HH \rightarrow bbWW \rightarrow bb\ell\nu\ell\nu$, respectively [90].

5.4.2 Background processes simulation

In this analysis the main background processes are top-antitop production ($t\bar{t}$) and Drell-Yan production in association with jets. Other background processes that contribute to a lesser degree include single top production, diboson production, and the production of a single Higgs boson in association with a Z boson (“ZH production”),

see Table 5.2. Other background processes are fully rejected in the event selection and thus are neglected.

Table 5.2: Background Monte Carlo samples. First column is the name of the process, the second column is the name of the generator used to simulate the corresponding process.

DY plus 1 Jet	MADGRAPH5_aMC@NLO-PYTHIA
DY plus 2 Jets	MADGRAPH5_aMC@NLO-PYTHIA
DY plus 3 Jets	MADGRAPH5_aMC@NLO-PYTHIA
DY plus 4 Jets	MADGRAPH5_aMC@NLO-PYTHIA
WW	PYTHIA
WZ	PYTHIA
ZZ	PYTHIA
ZH with $H \rightarrow b\bar{b}$ and $Z \rightarrow \ell\ell$	MADGRAPH5_aMC@NLO
$t\bar{t}$	POWHEG-PYTHIA
top quark tW channel	POWHEG-PYTHIA
\bar{t} quark tW channel	POWHEG-PYTHIA
top quark t-channel	POWHEG-PYTHIA
\bar{t} t-channel	POWHEG-PYTHIA
top quark s-channel	MADGRAPH5_aMC@NLO-PYTHIA

The Drell-Yan (DY) process in association with one to four jets is generated at leading order (LO) using MADGRAPH5_aMC@NLO with the MLM matching scheme [91]. To account for the higher order QCD and electroweak effects in V+jets production (following [92]), DY events are further reweighted according to the dilepton transverse momentum, see 5.6.5.

The simulations of the background processes associated with top quark production are generated at next-to-leading order (NLO) in QCD. The POWHEG [93–96] generator was used to generate the samples for top quark pair production and single top quark production in the tW and t channels. For the single top s channel production, the MADGRAPH5_aMC@NLO generator was used. Single top backgrounds have been rescaled to the theoretical values of the next-to-next-to-leading order (NNLO) cross sections [97, 98].

`PYTHIA 8.212` [99,100] was used to generate diboson samples at LO. Diboson background yields are normalized to NLO cross sections [101–103]. The dominant SM Higgs background process, the SM production of a single Higgs boson in the association with a Z boson (ZH), is simulated at NLO using the `MADGRAPH5_aMC@NLO` generator with FxFx merging [104]. The SM Higgs background from the ZH process is scaled to NNLO with the MCFM program [105]. All the final cross section values at the NNLO accuracy in perturbative QCD have been computed with the original generators [106–109].

Normalizations for $t\bar{t}$ and DY background processes are determined from data, as explained later in this chapter.

The NNPDF3.0 [110] Parton Distribution Functions (PDF) set is used for all the LO and NLO samples. `POWHEG` and `MADGRAPH5_aMC@NLO` are interfaced with `PYTHIA8.212` for the parton showering and hadronization stages. The description of the underlying event is done using the tune CUETP8M1 derived in [111]. For the simulation of the CMS detector response, `GEANT4` [112] was used.

For all MC simulations, further reweighting of events is done using the SFs derived to account for the discrepancies between the data and simulation. This includes SFs related to the lepton efficiencies as well as b tagging efficiencies, see Sections 4.3.5.1, 4.3.5.2.

As discussed in Section 4.3.3, multiple overlapping proton-proton interactions occurred in each bunch crossing during data collecting in 2016, with an average of 24 hard scattering vertices per event. To account for this fact, all simulated samples include additional interactions to reproduce the real pileup distribution measured in data. To match the expected pileup in data, pileup reweighting is done for all MC simulated events.

5.5 Physics Objects Selection

During the reconstruction of the collisions at the CMS, data for each event are refined into high-level physics objects that correspond to particles created as a result of a proton-proton interaction. Among the physics objects necessary for this measurement are reconstructed electrons, muons, jets originating from quarks of heavy flavor (b jets), and the missing transverse momentum. The reconstruction details have been given in Section 4. Here we focus on the analysis-specific selection of each of these physics objects.

5.5.1 Electrons

Electrons are reconstructed using the GSF algorithm [72]. The electron candidates are then selected by first applying a loose isolation requirement of 0.4, after which they have to pass the Tight ID criterion defined by the CMS e/γ POG, see 4. At the offline level, the analysis applies additionally the loose WP (WP90) [88]. With this WP, one achieves an electron selection efficiency of 90% by utilizing variables such as the agreement between the position of the ECAL cluster and GSF track that forms the electron, the energy of the 3 by 3 crystal core of the electron’s cluster, the ratio of electron energy measured in ECAL to the electron’s momentum measured in the tracking system, the ratio of the HCAL to ECAL energy, p_T and η of the electron, etc.

This is a justified level of the desired identification efficiency for the $Z \rightarrow \ell\ell$ channel since we select on-shell Z boson decays to charged leptons, where produced prompt leptons are very energetic. Finally, the electrons must be isolated from other particles in the event. Numerically, a selection criterion on the isolation is imposed (see Section 4.3.4), where the expected contribution of particles from pileup inter-

actions is ρ -subtracted using the effective areas method, see Section 4.3.4. In this measurement, the isolation (with the cone size 0.3) for electrons is required to be smaller than 0.06, which means that all other particles together in the isolation cone around the electron cannot have more than 6% of the electron’s energy.

5.5.2 Muons

Muon candidates are reconstructed using tracker muon and global muon tracks identified by the PF algorithm. The selection procedure is similar to the one of electrons: first, a loose relative isolation selection of 0.4 is applied, then muons have to pass the Tight WP of the muon ID selection. At the offline level, a more stringent set of quality requirements (WP Loose) recommended by the CMS Muon POG is applied to the muon object. The ID definition contains the variables such as muon track segments compatibility, the fraction of valid hits in the inner tracker, χ^2 variables of the inner tracker and global tracks, position matching variables, energy deposits in the calorimeters, impact parameter variables, etc. Lastly, a relative $\Delta\beta$ -subtracted PF isolation selection of 0.15 is applied, with the cone size 0.4.

5.5.3 Jets

Jets are reconstructed using the anti- k_T jet clustering algorithm. This algorithm clusters PF candidates in a cone with a radius of 0.4. The energy and the resolution of the produced jets (so-called AK4 jets) are further corrected using JEC and JER corrections. These are used to calibrate the energy of the jets and to smear the resolution of jets to match the resolution of jets in the data.

To reject misreconstructed (or misidentified) jets due to detector noise, pileup, etc., a loose jet identification WP is applied following the recommendations from the

CMS JetMET POG. The ID variables include the jet p_T and η , HCAL and ECAL energy fractions due to charged and neutral hadrons, the jet multiplicity, etc. If jets are found to be overlapping with charged leptons used in the measurement, these jets are not considered by the analysis. We consider jets with a p_T greater than 30 GeV, which are in the range of $|\eta| < 2.4$. At least two jets must be present in the event.

5.5.4 b jets

Each identified jet is further assigned a probability to originate from the b quark using the CMVA b tagging discriminant. This tagger produces as output a continuous discriminant with a value between -1 and 1, that is used to define three WPs depending on the chosen discriminant value. CMS BTag POG provides the CMS with the WP Loose, Medium, and Tight. For our measurement, the WP Medium requirement (>0.4432) has led to the best expected (from the simulation) results (the lowest upper limits as will be explained later). Therefore, the WP Medium is chosen in this analysis for the derivation of the final results. Jets passing the WP Medium requirement are classified as b jets. The threshold is chosen such that the misidentification rate for light-quark and gluon jets is about 1% and the b jet tagging efficiency for this working point is about 66%. Two b jets with the largest scores of the CMVA discriminant are used to reconstruct the $H \rightarrow b\bar{b}$ candidate, which will be explained at length later in this chapter.

5.5.5 Missing transverse momentum

In this measurement two neutrinos are present in the final state. They come from the off-shell Z boson decays to neutrinos and also from leptonic decays of W bosons. Even though the two neutrinos present in the signal events cannot be identified directly

by the CMS detector, their presence can be inferred from the combined transverse momentum vector, using the momentum conservation requirement in the transverse plane for each event. The missing transverse momentum \vec{p}_T^{miss} is computed as the negative vector sum of the transverse momenta of all visible PF objects and is further JEC and JER corrected, see Section 4.3.2. The missing transverse energy \cancel{E}_T is the magnitude of the \vec{p}_T^{miss} vector.

All corrections recommended by the CMS JetMET POG are used in this measurement [113]. Additionally, a set of filters related to the instrumental effects is employed, such as the removal of the misreconstructed signals in the HCAL, noise in the tracker, etc. [114].

5.6 Event Selection

The candidate events are selected in the data sample if they contain the following physics objects: 2 b jets, 2 charged leptons, and missing transverse energy. All these objects pass specific kinematic requirements defined in 4, 5.3.2, and 5.5. Then we reconstruct intermediate particles of the decay chain by combining the observed physics objects. $Z \rightarrow ll$ and $Z^* \rightarrow \nu\nu$ are reconstructed first. Then $H \rightarrow ZZ^*$ and $H \rightarrow b\bar{b}$ are reconstructed. The distributions of these variables will be shown later after the data-MC corrections are discussed. Finally, two Higgs boson candidates are used to form the HH system that could originate from the narrow resonance we are searching for.

In this chapter, we discuss how HH candidates are formed, how the selection criteria are applied to reduce contamination from background processes in the signal-enriched region (discussed later), how kinematic quantities used to compute the final results are defined, and how well simulation represents the collected data.

5.6.1 Kinematic selection of physics objects

We have described the kinematic selection of leptons and jets (and b jets) in the previous chapter. Because there is another $bbZZ$ measurement in the CMS Higgs group, our analysis shares some phase space with this other measurement. Therefore, the selected data samples may contain a certain overlap. The other $bbZZ$ team studies the $2b2l2q$ final state, where q stands for any quark initiating the jet. To ensure that both measurements produce statistically independent (“orthogonal”) results, a selection on the MET is introduced. Since the final results - the production cross sections multiplied by the BFs of the final state - depend on the mass point, the MET selection varies with mass. While the derivation of the final results (called “limits”) will be explained later in this chapter, at this point it is worthwhile to point out that both analyses computed limits for all resonance mass hypotheses applying different MET selection criteria. Using a significance-like figure of merit ($S/(S + B)$), see Fig. 5.5, we derived the set of MET requirements. These values have been checked by the other $bbZZ$ team and further optimized to yield the best simplified combined limits, when the results of both measurements are merged together. The final selection is shown in Table 5.3.

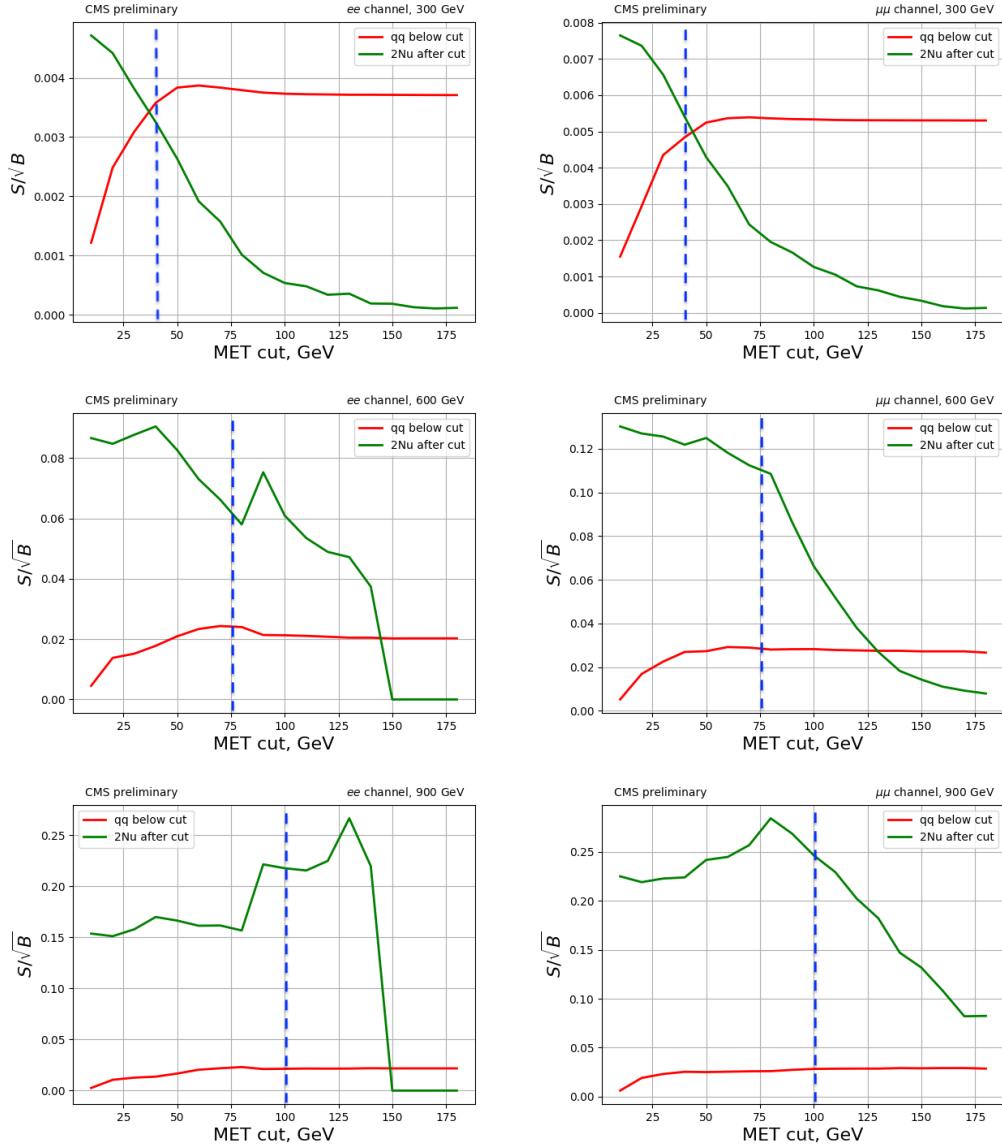


Figure 5.5: Optimisation of the MET selection for two $bbZZ$ analyses. S and B stand for signal and background process yields respectively. A significance-like figure of merit (\sqrt{S}/B) as a function of the MET cut is used to identify the best selection for each given mass hypothesis. The green curve shows the significance for this measurement, where events are kept above a given threshold. The red curve shows the significance for the other $bbZZ$ measurement, where the MET requirement is inverted (events are kept below that threshold). Top: 300 GeV resonance mass. Middle: 600 GeV resonance mass. Bottom: 900 GeV resonance mass. The plots on the left are for the di-electron channel and the plots on the right are for the di-muon channel. The dashed blue lines indicate the values of the MET selection, see Table 5.3.

Table 5.3: The MET requirements as a function of the mass of the HH candidate. Selection values (the second column) are provided for different mass hypotheses (the first column) of the narrow resonance decaying to the HH system.

Signal mass (GeV)	\cancel{E}_T selection (GeV)
260-300	> 40
350-600	> 75
650-1000	> 100

5.6.2 Signal candidate construction and selection

Two leptons of the opposite signs and the same flavor with an invariant mass greater than 76 GeV are selected as Z candidates. This mass requirement helps to reject dilepton candidates not corresponding to real Z bosons. Events with a dilepton mass higher than 76 GeV will be used for SR and $t\bar{t}$ CR. Additional Z mass selection will be discussed later in this chapter. These on-shell Z bosons are assumed to come from the $H \rightarrow ZZ^*$ decays. The other Z boson, the off-shell Z boson, decays to two neutrinos in our signature and is represented by MET. The Lorentz four-vectors of the dilepton candidate and MET are added and the resulting four-vector represents the first Higgs boson candidate.

The other Higgs boson is formed from the pair of b jets with the highest output value of the CMVA algorithm. No requirement is applied on the existence or absence of other b jets in an event because they may be present due to a b jet misidentification or b jets resulting from pileup interactions or the underlying event. The invariant mass of the $H \rightarrow b\bar{b}$ candidate is required to be greater than 20 GeV to remove contamination from background events with low mass resonances that decay to two jets, such as events with J/psi, Upsilon or low energy QCD interactions. There is no upper requirement on the $H \rightarrow b\bar{b}$ invariant mass. Together with the first Higgs boson, the constructed di-Higgs system approximates the double Higgs boson

production that this measurement studies.

The final double Higgs boson candidate (HH candidate) comprises the $Z(\ell\ell)$ candidate, the MET representing the $Z(\nu\nu)$ decay, and the $H \rightarrow b\bar{b}$ candidate. The four-momentum of this HH candidate is defined as the sum of the four-momenta of the two leptons and two b jets in the candidate as well as the four-momentum of the MET defined as $(\cancel{E}_T, \vec{p}_T^{miss})$, see Sec. 4.

HH system decays to a pair of b quarks and a pair of W bosons can also result in the same final state. The expected yields for the $bbWW$ channel with respect to the $bbZZ$ yields are comparatively small (1 to 4) because of the stringent kinematic selection on the dilepton invariant mass. However, the contribution from the HH system decaying through the $bbWW$ intermediate state is still considered to be part of our signal in this measurement. The minimum requirement on the dilepton mass is necessary to ensure that our measurement is orthogonal to the known HH search from CMS in the $bbVV$ channel that focused on $bbWW$ decays [115], where only events with the dilepton mass below 76 GeV were studied.

In collider physics, one of several common definitions of the transverse mass is given by: $M_T = \sqrt{(E^2 - p_z^2)}$. This quantity is used in CMS searches for new particles. We proceed in the same fashion constructing our final variable - $\tilde{M}_T(HH)$, using the di-Higgs four-vector defined in the previous paragraphs.

While up until this moment we had the same requirements for the search of the radion/graviton $\rightarrow HH$ for any heavy resonance mass, it is worthwhile noting that at this point the analysis becomes mass point specific. This is because the selection on the BDT discriminant (which will be explained later) and the MET requirement differ with the mass hypothesis. When we compute limits for different radion or graviton hypotheses we apply different selection criteria.

Finally, the reader should be informed that if an event does not have enough

physics objects that are suitable (as discussed in the objects selection section) for building the $Z(\ell\ell)$ or $H \rightarrow b\bar{b}$ candidates, or if the event does not pass candidate selection requirements on those objects or the full HH system, the event is discarded.

5.6.3 Signal and control kinematic regions

Three regions are defined: signal-enriched or signal region, and two CRs. The signal region is chosen such that the expected signal fraction is the largest there. Two control regions are defined such that they contain predominantly the background events with almost no signal. We define two CRs, one for each of the two main background sources: the CR for $t\bar{t}$ (CRTT) and the CR for Drell-Yan in association with jets (CRDY). The CRDY contains a high fraction of DY events, and the CRTT contains nothing but $t\bar{t}$ (more than 95%). This ensures the correctness of the DY and $t\bar{t}$ normalizations extracted during the fit procedure (explained later).

Two intermediate particles in the decay chain are almost fully reconstructed, $H \rightarrow b\bar{b}$ and $Z(\ell\ell)$. The signal region is defined in the kinematic space of $H \rightarrow b\bar{b}$ and $Z(\ell\ell)$ events. This corresponds to an area determined by the mass of the Higgs boson (125 GeV) and the mass of the Z boson (91.2 GeV). To account for detector smearing effects, the mass windows near the pole masses are defined such that we select events within the $H \rightarrow b\bar{b}$ mass range from 90 to 150 GeV and Z mass range from 76 to 106 GeV, see Fig. 5.6. In CMS, these relatively standard mass windows are chosen to take into account the detector resolution effects on a two b jet system and the dilepton mass (and the natural width for a Z boson), containing approximately 95% of true Higgs and Z boson candidates. The proportion of signal events is further increased with respect to background events by applying an additional requirement, the MVA requirement, which will be explained in the next section.

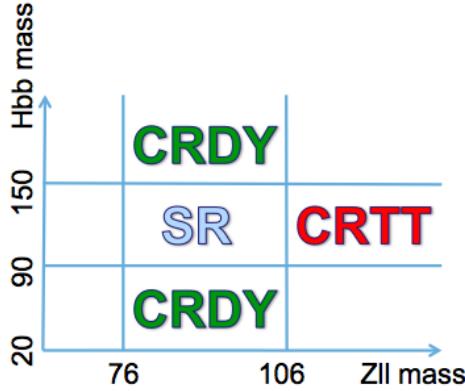


Figure 5.6: Signal region, control region $t\bar{t}$, and control region Drell-Yan in the kinematic space of $Z \rightarrow \ell\ell$ and $H \rightarrow b\bar{b}$ masses.

CRs are defined by inverting the SR selection. The Drell-Yan plus jets (or “DY” later for brevity) background process contains a real Z boson with two jets. Two jets from this process have a broad distribution. We obtain the CRDY by inverting the $H \rightarrow b\bar{b}$ selection. To construct the CRTT, we use the fact that $t\bar{t}$ events do not contain a real Z boson. Thus, to define CRTT, we invert the $Z(\ell\ell)$ selection and use only the events in the upper range of the Z mass, see Fig. 5.6.

CRs are used to fit the event yield of the simulated background processes to the event yield in data and, thus, to determine the relative normalizations of $t\bar{t}$ and Drell-Yan background processes, which is done using CRTT and CRDY correspondingly.

The efficiency of the candidate selection, up to this point, is summarised in Table 5.4, where numbers are provided for the SR. The number of observed data events and expected signal events will be presented in the Section 5.9. The signal contributions are split into two components: $bbZZ$ and $bbWW$, and are shown separately. The corresponding total number of events passing the selection are also given, shown in the fourth column. Note that the number of signal events surviving the BDT selection in the SR is 2 - 5 times greater (depending on the mass point) than the number of

Table 5.4: Fraction of events surviving the candidate selection and kinematic requirements. Efficiencies are given for $bbZZ$ and $bbWW$ contributions in the SR and are normalised to the initial event counts before any selection is applied. A di-muon channel is presented. Numbers for the di-electron channel have the same trend but lower values, because in CMS efficiencies for electrons are lower than for muons.

Process	Mass (GeV)	Efficiency (%)
bbWW	300	0.2
	300	10.4
bbZZ	900	0.1
	900	15.1

all signal events present in the CRDY, while CRTT contains no signal.

5.6.4 Signal region candidate selection with a multivariate technique

After the SR and CRs have been defined, in the final step of the event selection we require the events in the SR to pass the threshold value of the BDT discriminant. An MVA discriminant that uses a boosted decision trees algorithm is trained considering a number of kinematic quantities of a candidate to be above a certain threshold. As this selection step is complex, it is described in a dedicated chapter (see 5.7).

5.6.5 Candidate reweighting in simulated events

Several factors lead to a difference in the behavior of simulation in comparison to real data. These factors include the difference in the kinematics of simulated events from what it is in nature as well as the differences in the simulation of the detector response that affects reconstruction efficiencies for physics objects that form candidates, as well as tagging and mistagging rates for b jets.

These factors are quantified through per-event SFs that are used as event weights

in preparation of all kinematic distributions and computation of the expected event counts. The SFs of all types are produced centrally in CMS by the appropriate POGs. The full per-event weight for a simulated event is a product of all individual SFs discussed below.

- **Lepton efficiency**

Scale factors that reflect the differences between data and simulation in lepton reconstruction, selection, and triggering are computed as the ratios of the corresponding efficiencies between data and simulation. Additionally, since the DZ requirement was absent for some triggers, the efficiency of the DZ selection needs to be measured.

The efficiencies and, subsequently, the scale factors are measured separately for the efficiencies of the reconstruction procedure, lepton identification, isolation requirements, and impact parameter requirements. The TnP procedure discussed in Section 4.3.5.1 is applied to measure all of these efficiencies. The scale factors of all listed types for both leptons are used as multiplicative factors to form the per-candidate lepton efficiency SF.

On average, the muon ID scale factors are equal to 0.992 - 1.004, depending on the p_t and η of leptons. The muon ISO SFs are equal to 0.991-1.004. The electron ID · ISO SFs, computed together, are ranging from 0.832 to 1.073. For the HLT SFs, see Figs. in the subsection 5.3.2.

- **b tagging and light flavor mistagging probability scale factors**

To identify the b jets, a selection on the output of the CMVA tagger is applied. The efficiency of the b jet selection is measured (with the TnP method) in data and in the heavy-flavor enhanced MC sample, the $t\bar{t}$ process containing top and

anti-top quark decays. Since these efficiencies are different, the b tagging SFs are derived. Because this measurement contains two b jets in the final state, a total per event b tagging SF is equal to 0.7 - 1.3 for signal and 0.6 - 1.4 for background processes.

- **Kinematic reweighting of the DY plus jets samples**

DY plus jets samples used in this measurement are produced at LO (see Section 5.3). The samples are corrected to exhibit NLO kinematic behaviour in the EW interaction by assigning a weight to each event. As the biggest discrepancy between data and simulation for such samples is typically found in the di-jet invariant mass distribution, an LO-to-NLO correction as a function of the $\Delta\eta_{b \text{ jets}}$ separation is computed following the procedure described in [116]. As stated there, this correction improves the data-MC agreement in many distributions derived from the di-jet objects or based on the $p_T^{Z(\ell\ell)}$ variable, while has a negligible effect on the remaining kinematic variables. The correction means that a kinematic weight is assigned to each event.

It was found by several Higgs analysis teams that the previous correction does not address the data-MC discrepancies for Z bosons with high p_T . The spectrum of the p_T of the Z boson distribution is harder for DY plus jets MC simulation than for data. This is due to the fact that the simulation does not include higher-order QCD+EW corrections [117]. The DY plus jets simulated samples are reweighted as a function of the $p_T^{Z(\ell\ell)}$ to update the LO samples to the NLO in QCD+EW. The corrections can be as large as 10% for energetic Z bosons with $p_T > 400$ GeV.

5.7 Multivariate selection in the signal region

To reduce the background contamination in the SR, we employ an MVA discriminant. In this section, we will discuss which discriminating variables were used, how the MVA discriminant was constructed, and what the efficiencies of the BDT selection are for all major signal and background processes.

5.7.1 Kinematic variables of a candidate

A number of kinematic quantities can be constructed out of four-vectors that are used to construct candidates of our final state. The distributions of these quantities, in general, differ for candidates originating from different background processes as opposed to for candidates from the hypothetical signal.

In this section, we discuss a set of kinematic quantities that are the most discriminating between the expected signal candidates and background candidates. This set of variables is used as an input for computing the multivariate discriminant that is later discussed in this section. In total, nine variables are chosen for this purpose. About 30 - 40 other kinematic variables were considered at the early stages of this data analysis. However, they were discarded as it was found that they do not improve the results of this measurement in a significant way, while at the same time they considerably increase the complexity of the measurement.

Since the MVA selection requirement will be applied on a quantity that was computed based on these variables, the simulation has to reproduce the variables well. We will define each of the nine variables below. For SR and both CRs, we will show the comparison of distributions from data and simulation for each variable. Since the number of plots would be too large to show for two channels and two spin hypotheses, not to mention for all masses, we will include in this dissertation only di-muon

channel plots for the 300 GeV graviton mass hypothesis. However, all other relevant figures of the nine main variables, including the di-electron channel and other mass points, were examined and documented in the internal CMS notes, and show similar behavior and agreement as those included here.

As we outlined before, even for simple physics objects the distributions from MC simulations and collected data have discrepancies. For more complex objects and candidates, these discrepancies become even larger. In general, MC simulations do not reproduce very narrow regions of the kinematic phase space well, or align with the data perfectly at the high order of the QCD. Therefore, one normally adjusts the event yields (rates) of background processes using the normalization values determined in the fit to data in the control regions. The shape of the distributions are further modified during the bin-by-bin correction procedure. The distributions as they are, without the extra normalizations from the fit being applied to the event yields of MC simulated background processes, are called pre-fit distributions. After the normalizations from the fit are applied to the rates of background processes, post-fit distributions are obtained. In the next Section, we provide both pre-fit and post-fit distributions back-to-back for an illustration how MC simulation performs out of the box and after the adjustment based on the fit to the data. The fitting procedure is discussed at length in the Statistical Analysis chapter 5.9.

The list below summarises the set of variables used as input to the MVA discriminant:

- **ΔR separation between two b jets.** The ΔR (defined in Section 3) separation between the b jets ($\Delta R_{b\ jets}$) gradually decreases as we start considering larger graviton or radion masses. This fact is explained by the Lorentz boost of the $H \rightarrow b\bar{b}$ system. When a heavy resonance produces two SM Higgs bosons,

they are highly boosted. For the Higgs boson, which will further decay to a pair of b quarks ($H \rightarrow b\bar{b}$), the b quarks will come out of the decay almost collinear to each other with a minimal value of $\Delta R_{b\text{ jets}}$. For background processes, this separation does not depend on the graviton or radion mass.

- **ΔR separation between two leptons.** The distance between two leptons in $\eta - \varphi$ space ($\Delta R_{leptons}$) tends to be smaller for the dilepton system coming from the real Z bosons originating from $H \rightarrow ZZ$ decays, in comparison to the $\Delta R_{leptons}$ distance between two random OSSF leptons produced by background processes.
- **Missing transverse momentum.** The \vec{p}_T^{miss} distribution coming from the off-shell Z bosons is constrained by the invariant mass of the parent Higgs boson and, thus, has a more narrow distribution. Conversely, the MET coming from $t\bar{t}$ background process can produce smaller and larger values of the MET outside of the aforementioned constraintment.
- **Invariant mass of the $H \rightarrow b\bar{b}$ candidate.** The mass of a two b jet system coming from a real Higgs boson tends to be close to the SM Higgs boson invariant mass smeared by the b jet energy-momentum resolution, while the background candidates from unrelated b jets can have smaller and larger masses, especially if they come from the top quarks of the $t\bar{t}$ background.
- **Transverse momentum of the $H \rightarrow b\bar{b}$ candidate.** The transverse momentum of the $H \rightarrow b\bar{b}$ candidate (p_T^{Hbb}) tends to have a relatively narrow and peaked distribution for candidates coming from the hypothetical signal in comparison to the broad distributions produced by background processes.

- **Invariant mass of the ZZ^* system.** ZZ^* mass of two Z bosons coming from the Higgs boson decay ($H \rightarrow ZZ$) tends to be close to the SM Higgs boson mass smeared by the \vec{p}_T^{miss} resolution, while the background candidates from DY production in association with jets can have random masses. These can be particularly large if they come from DY production in association with several jets in addition to some instrumental MET present in the event.
- **Transverse momentum of the ZZ^* system.** The transverse momentum of the $H \rightarrow ZZ$ candidate (p_T^{Hzz}) tends to have a relatively narrow and peaked distribution for candidates coming from the hypothetical signal in comparison to broad distributions produced by the background processes.
- **Invariant mass of the Z boson candidate.** Z bosons coming from the Higgs boson have the invariant mass near the pole mass, while $t\bar{t}$ contains no real Z bosons.
- **Transverse momentum of the Z boson candidate.** The transverse momentum of the Z candidate (p_T^Z) tends to have a relatively narrow distribution for candidates coming from a real Z boson produced in the decays of the hypothetical signal in comparison to broad distributions of $p_T^{\ell\ell}$, when dilepton systems are produced by background processes.

At this stage of the analysis description, we refer the reader to the pre-fit plots (figures in the left columns). The figures are shown in the following order: all nine distributions in CRDY (see Figs. 5.7, 5.8, 5.9, 5.10, 5.11), in the BDT sideband¹ of the SR (see Figs. 5.12, 5.13, 5.14, 5.15), and in CRTT (see Figs. 5.16, 5.17, 5.18,

¹ Once the BDT selection is introduced, in the SR the candidates will have to pass the BDT requirement. The region below the BDT selection is called the BDT sideband. In the unblinding scheme employed in this measurement, discussed later, the SR distributions are not examined at this point; data-MC agreement is checked using CRs and the BDT sideband of the SR.

5.19, 5.20). The simulated distributions are found to be in an acceptable agreement ($\chi^2 \sim 2$) with those from the data and are judged to be adequate for use in a multivariate discriminant construction discussed next. The distributions in the SR will be shown in the Section 5.9.2.

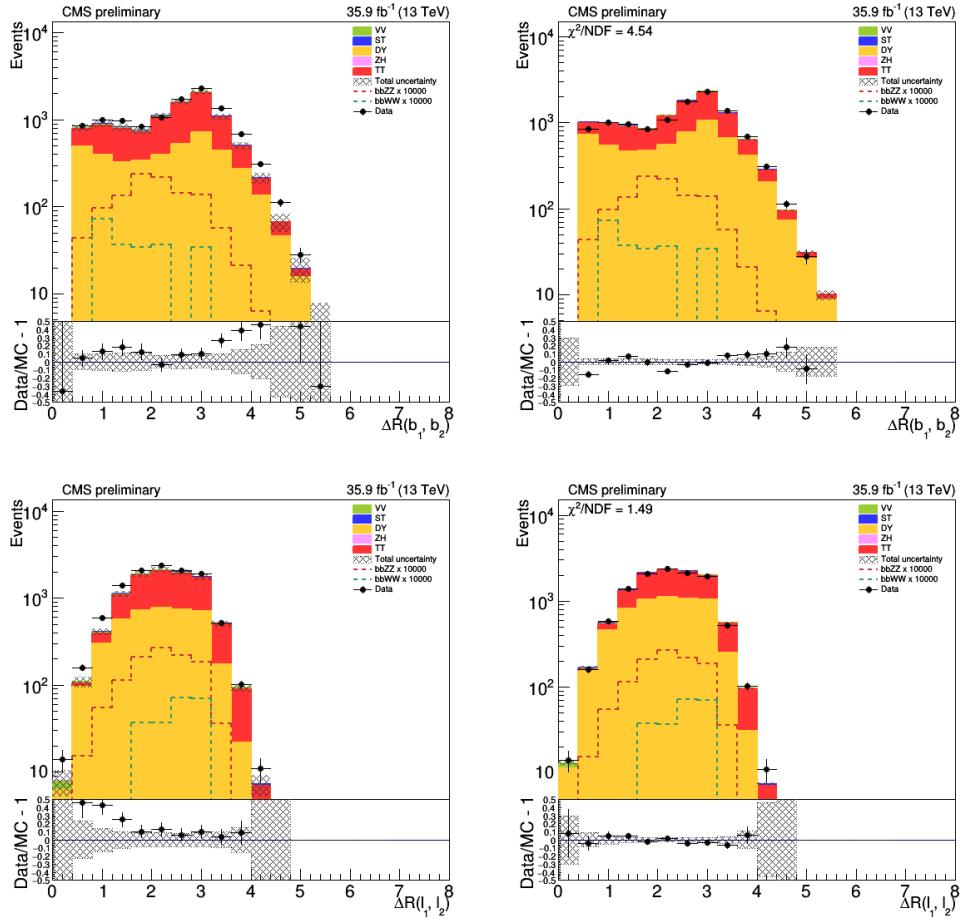


Figure 5.7: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRDY region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

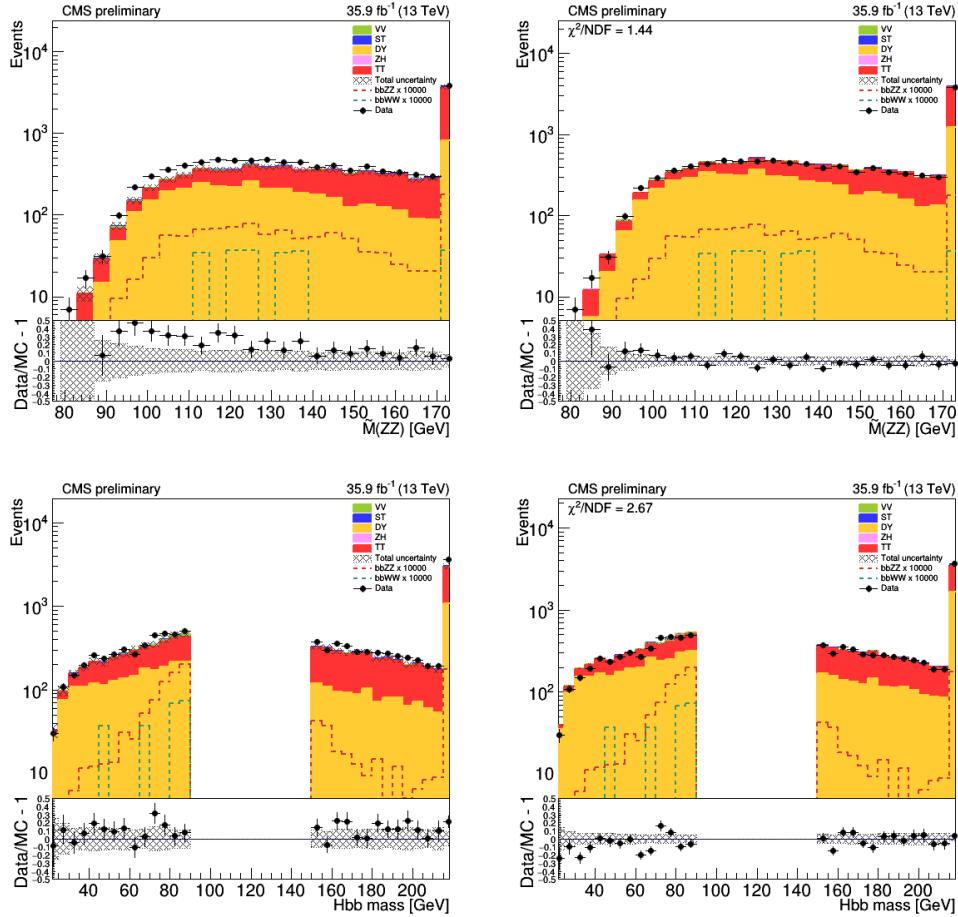


Figure 5.8: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRDY region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

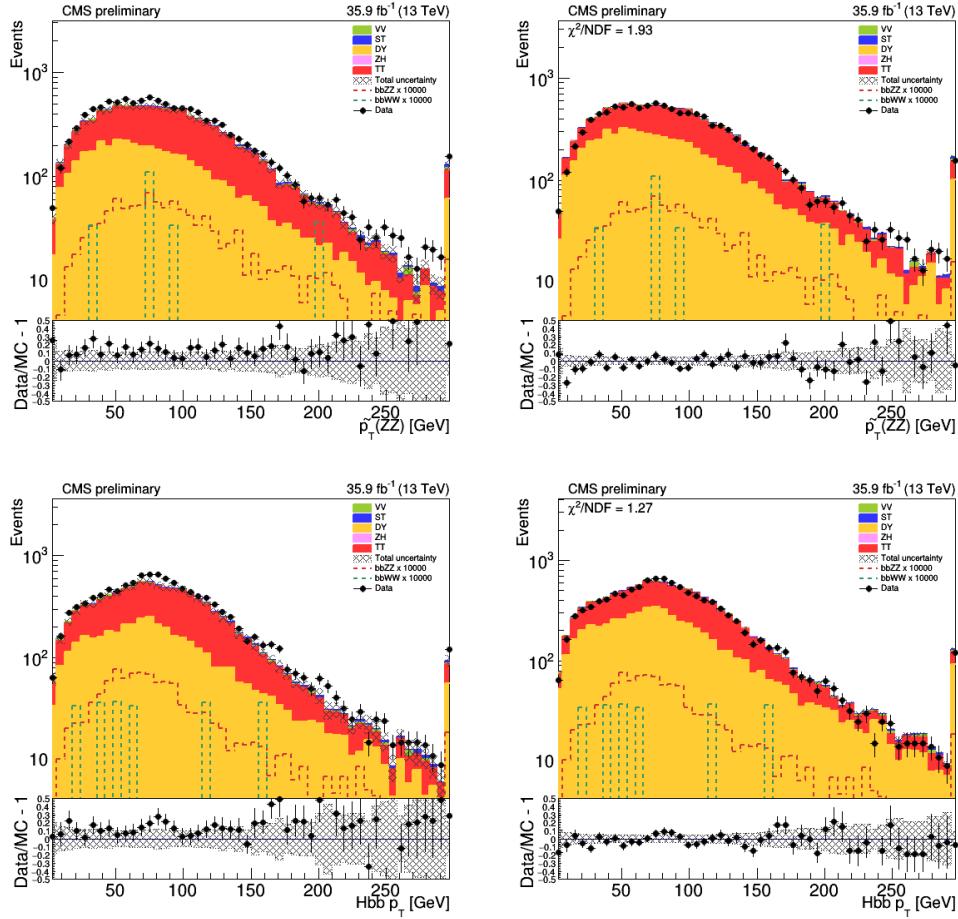


Figure 5.9: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRDY region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

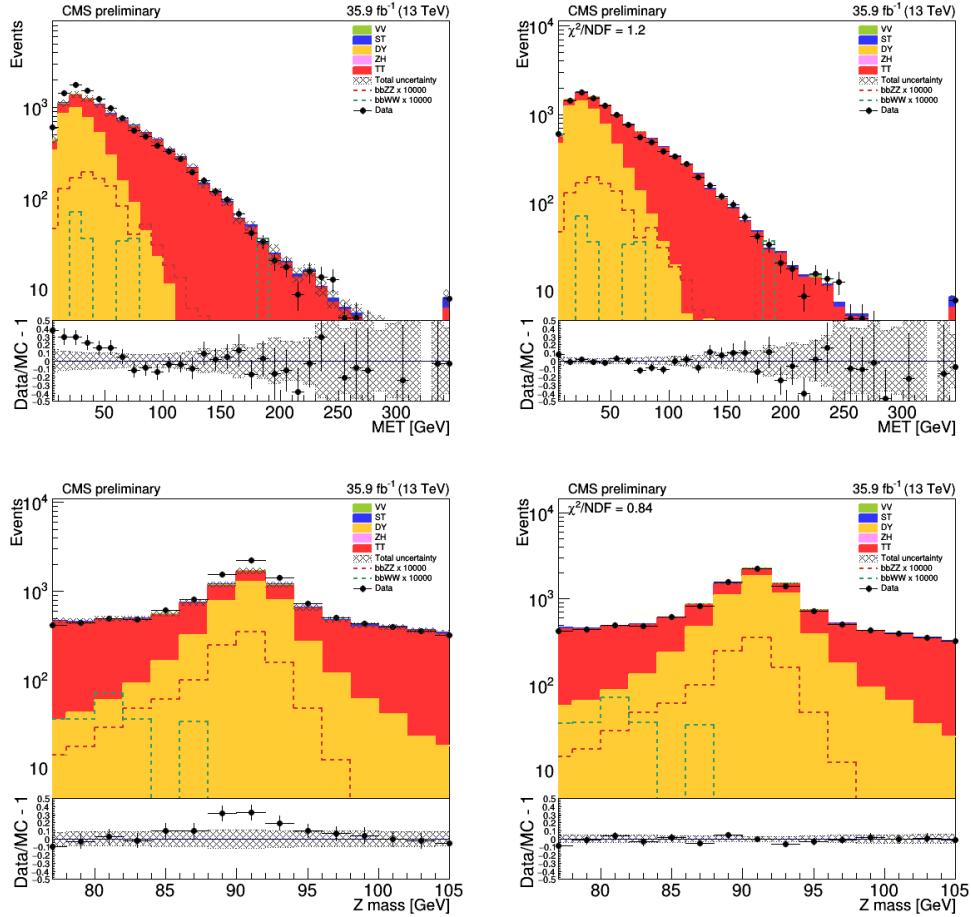


Figure 5.10: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRDY region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

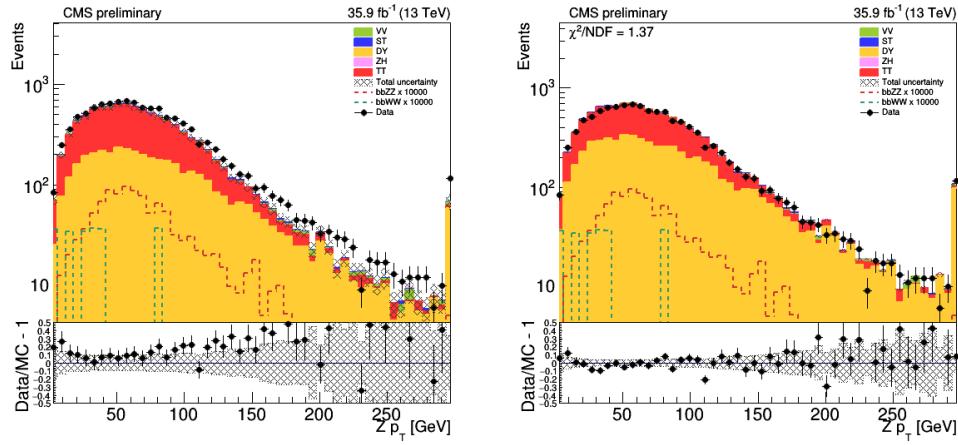


Figure 5.11: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRDY region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

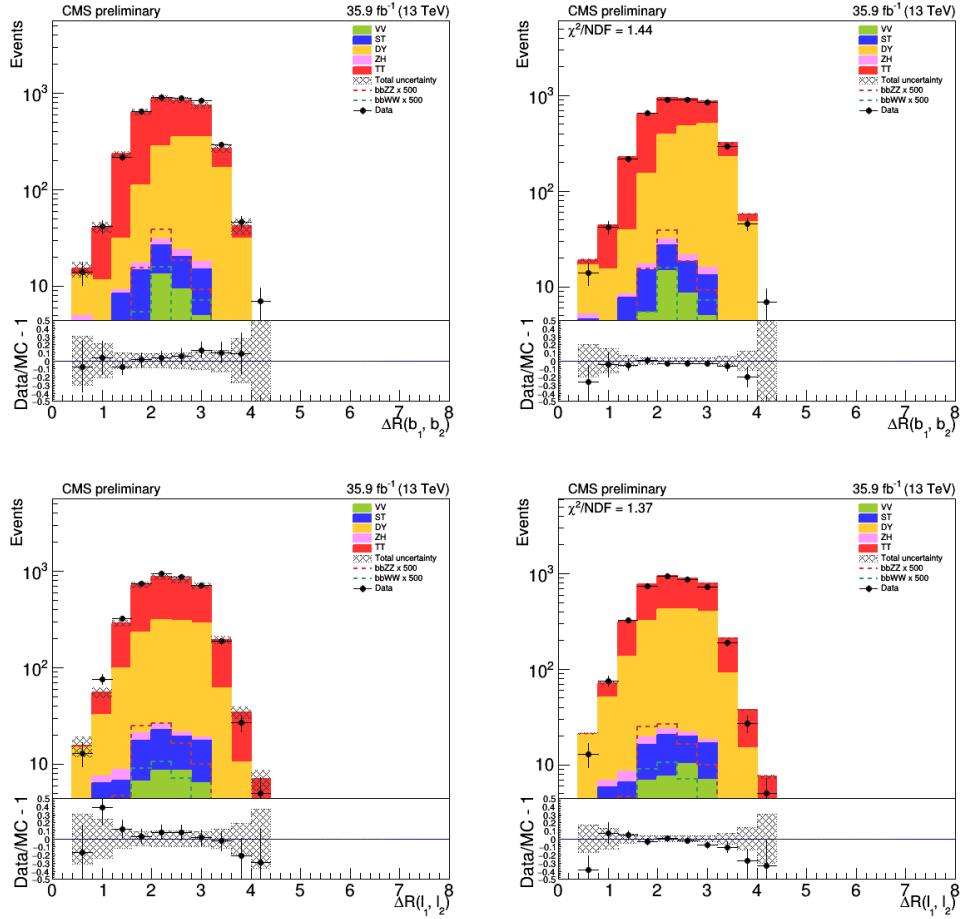


Figure 5.12: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, BDT sideband of the SR, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

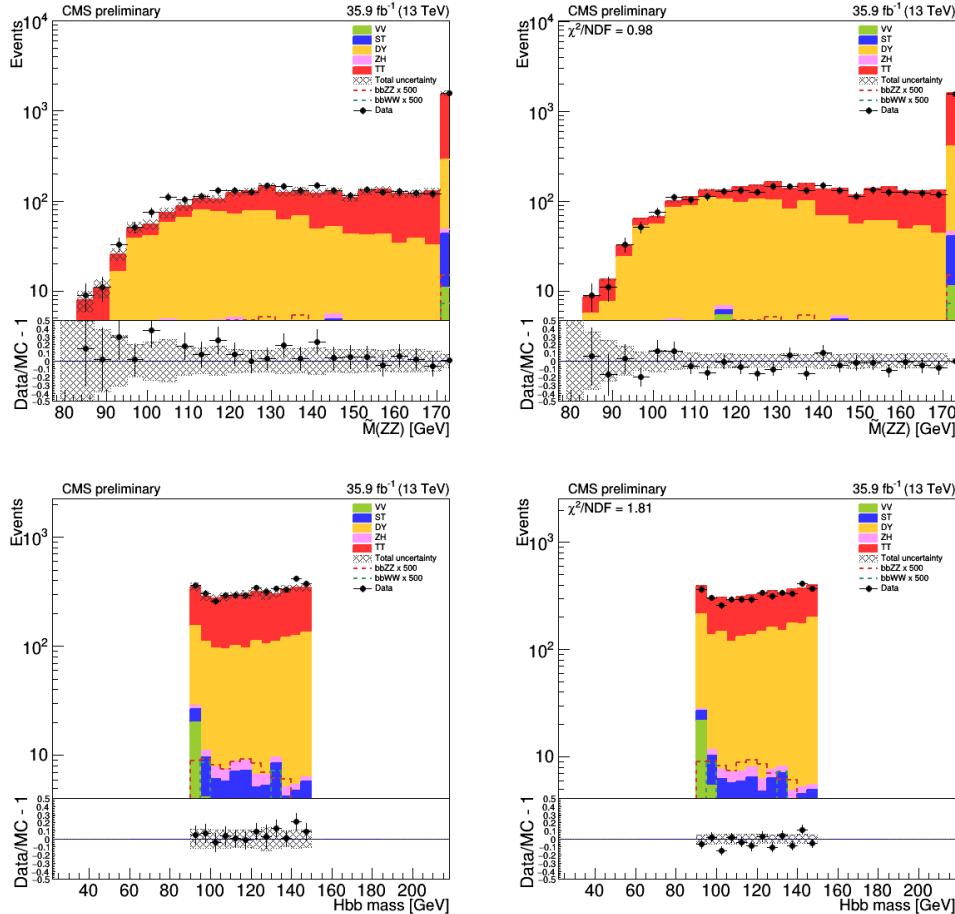


Figure 5.13: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, BDT sideband of the SR, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

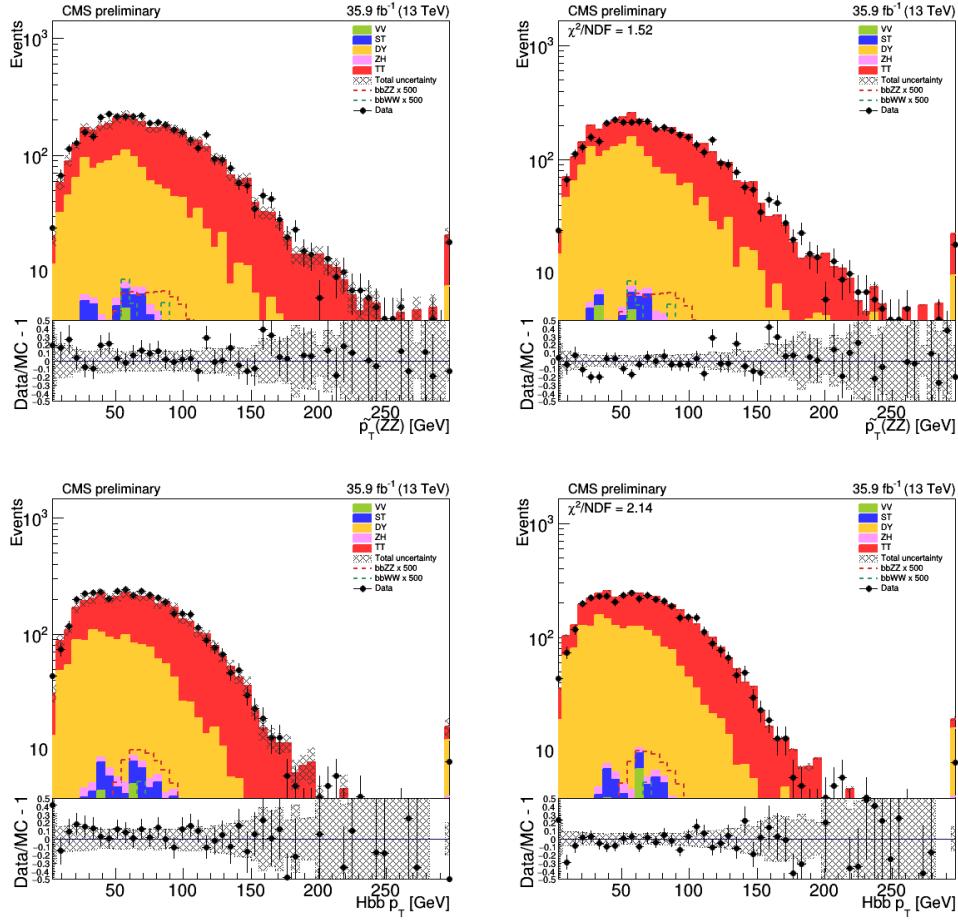


Figure 5.14: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, BDT sideband of the SR, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

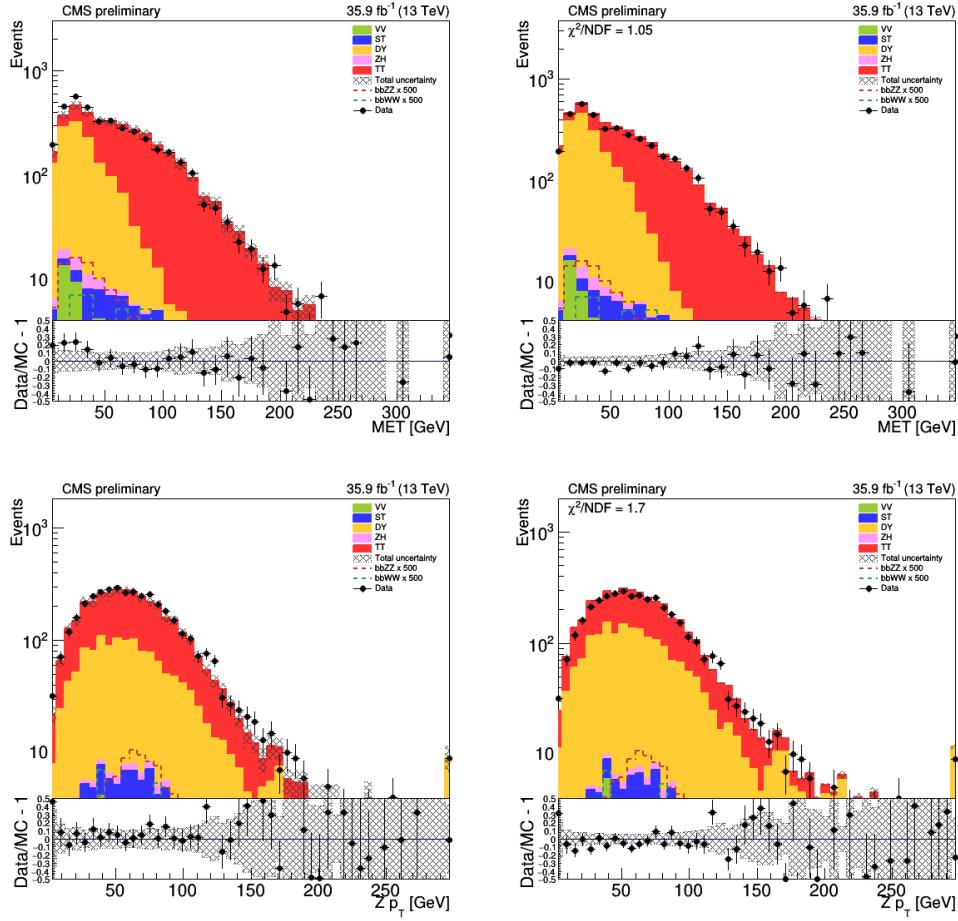


Figure 5.15: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, BDT sideband of the SR, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

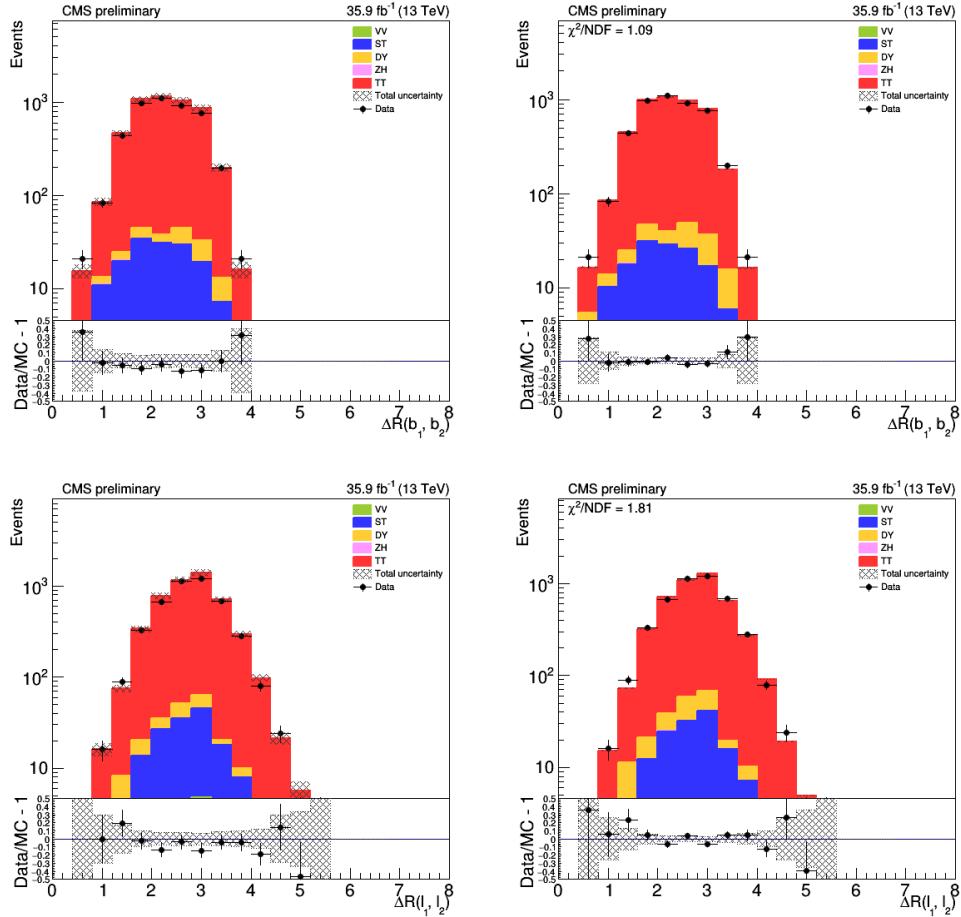


Figure 5.16: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRTT region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

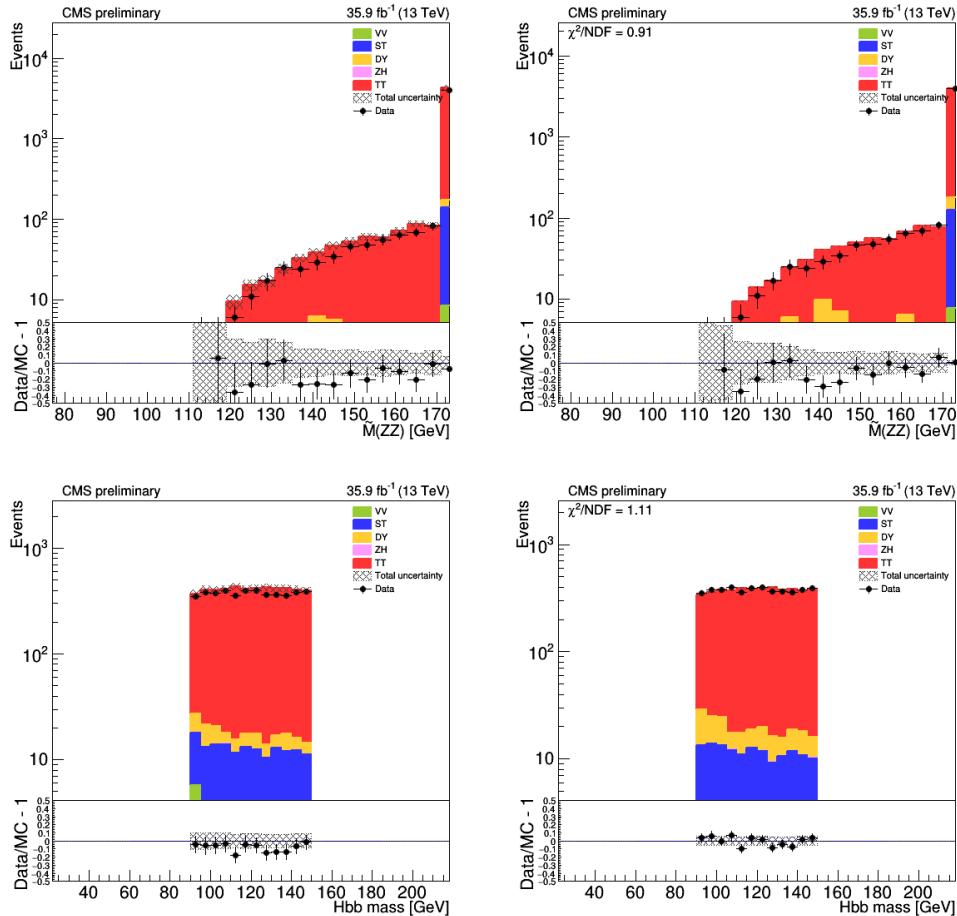


Figure 5.17: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRTT region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

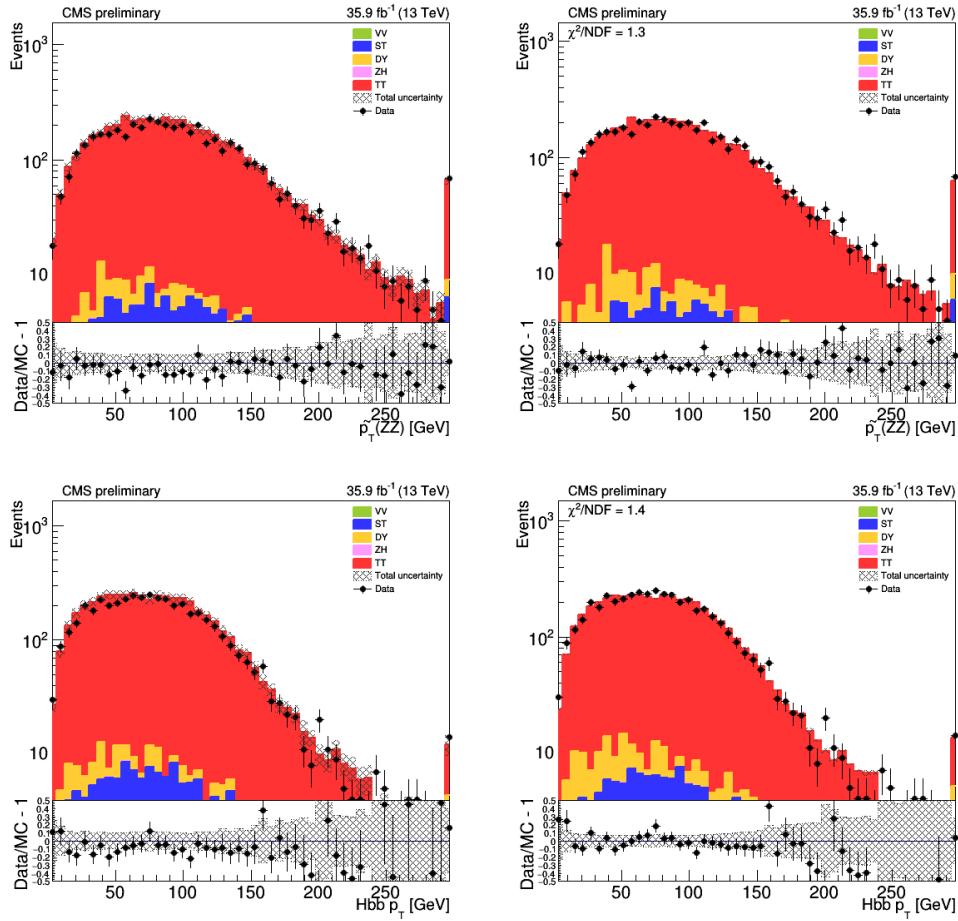


Figure 5.18: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRTT region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

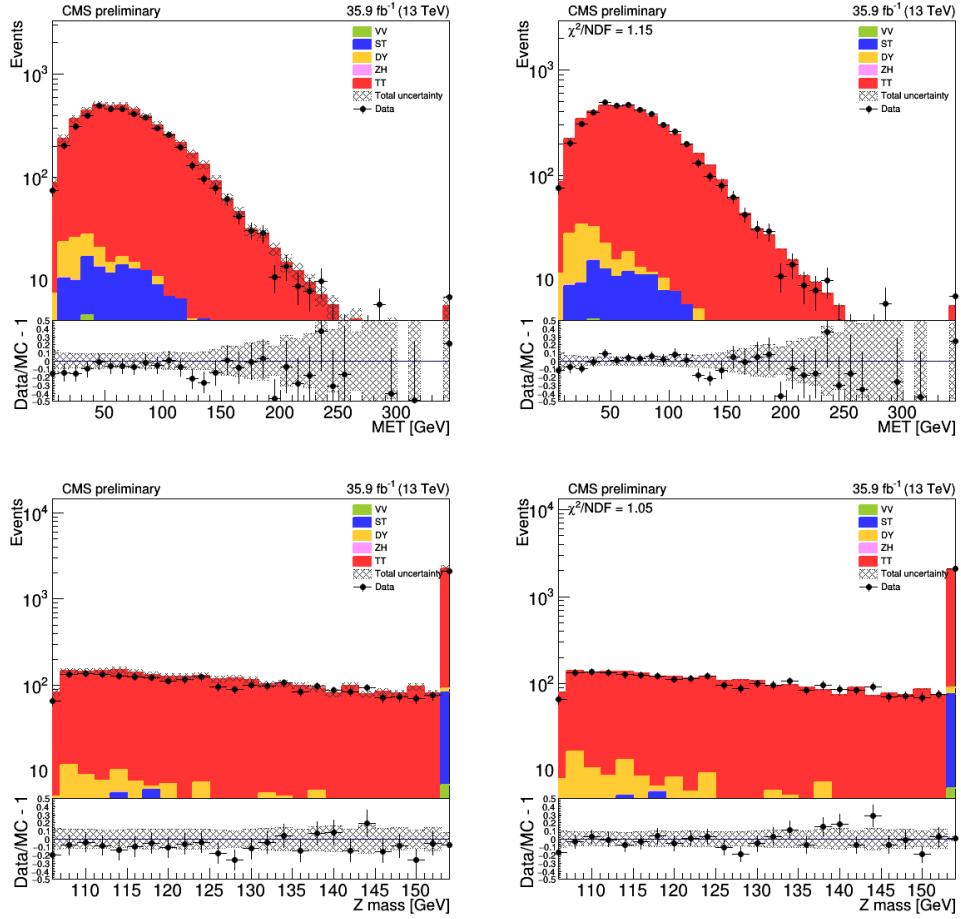


Figure 5.19: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRTT region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

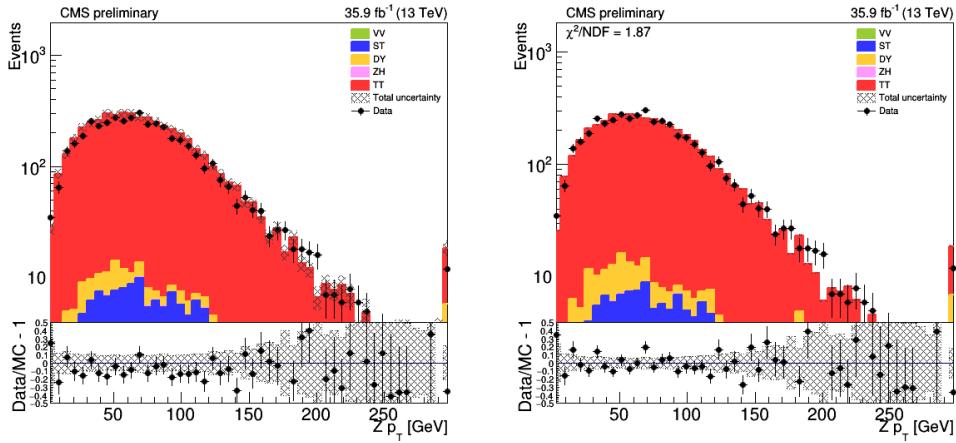


Figure 5.20: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, CRTT region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

5.7.2 Multivariate discriminant: a BDT classifier

In many areas of HEP, such as object identification, a reduction of the background contribution in the signal region, etc., there are needs to discriminate a “signal” from one or several “backgrounds.” The discriminating algorithms can use a single physical variable for signal-background separation, or several variables. The Higgs boson discovery, in part, was possible due to an employed BDT approach, where several variables are combined into an artificial BDT score or output. The BDT output, as a higher-level construct, has no direct physical interpretation. However, due to the demonstrated gain in signal-background separation performance, the BDT has been successfully used for years for the final classification decision.

Multivariate discriminants (or models or classifiers) based on the BDT or other algorithms are constructed using machine learning (ML) techniques. Now, at almost every stage of the HEP data analysis, starting from the physics object reconstruction, through the object identification and the signal region purification, and up to the

construction of the final discriminating MVA variable, physicists rely on the ML methods.

In simple terms, the ML procedure consists of training and testing stages, and sometimes a validation stage is added. First, the input data are split into two non-overlapping sets: train and test data. Then the classifier is trained when the model is learning the patterns in the train data. After that, the performance of the model is tested using independent test data.

After the training procedure, the ML algorithm produces a discriminant, which is a function that is related to the likelihood of an event to belong to a signal or to a background category. Here and in other parts of the thesis, we are describing the BDT procedure, which is a non-parametric supervised learning method used for classification; other MVA methods that are not related to this measurement, are well-explained in [118].

One prefers a classifier with low bias and low variance. A bias is the difference between the average prediction of our model and the correct value, the latter of which we are trying to predict. A classifier with a high bias oversimplifies the model and does not take into account all the nuances of a given train dataset. Therefore, analysts prefer low bias models, which are flexible and perform well both on the train and test data, and always provide a stable separation between the signal and background hypotheses. A variance is the variability of model prediction for a given data point and is related to the spread of our data (a sharp peak vs a broad distribution). A model with high variance learns almost every single pattern in the train data and, therefore, will not generalize well when given new unseen data. As a result, models with a high variance achieve the best scores using the train data but perform poorly on the test data. Thus, low variance models, which are robust when applying the trained classifiers on a dataset that is statistically independent from the one used for

training, are preferable.

One of the most popular ML algorithms in HEP is based on boosted decision trees (BDT). A single decision tree is a sequence of simple binary splits on each of the input variables. The tree can thus be depicted as a multidimensional set of selections on input space. The tree assigns a score of +1 or -1 if a given event is signal or background-like. The key elements in each step of the BDT splits are the variables used for the selection and the positions of the selections. Both are determined using an impurity criterion $I(p_n)$, which is a function of the signal purity p_n in a tree node n . As a definition of the impurity criterion for this measurement, the Gini coefficient is used: $I_G(p_n) = 2p_n \cdot (1 - p_n)$. The signal purity is defined as a ratio of the signal contribution to the total number of signal and background events, $p_n = S/(S + B)$. The impurity criterion is chosen such that it is at the minimum when $p_n=0$ or $p_n=1$. These values would give perfect signal-background discrimination. The criterion is at the maximum, when $p_n=0.5$, which means no discrimination is achieved. Each split maximises the gain function $G = I(p_n) - f_1I(p_{c1}) - f_2I(p_{c2})$. In this formulation of G , the labels c1 and c2 refer to child nodes of the parent node, and f_1 and f_2 are the fractions of events of the parent node that are found in child nodes c1 and c2, respectively. The tree keeps applying the selection on the input variables until either a maximum depth is reached (a number of consecutive splits), or a minimum number of events in the final child nodes are attained. The problem with such a tree growing procedure is that it is prone to overtraining, which means the model will not generalize well and have a high variance.

One way to deal with BDT overtraining is boosting. The idea behind boosting is that signal events from the training sample that end up in a background node or vice versa are given intentionally large weights (a penalty). These weights are normally significantly larger than the weights for events with the correct leaf node prediction.

Such a penalty procedure re-weights the train data and then new decision trees are created. The boosting is applied hundreds of times, which produces a set of new reweighted trees (a forest). Intuitively, boosting is a method in which a large number of shallow trees, which have only a few splits, are combined by taking a weighted average of their output scores. The discriminating power of a single shallow tree is usually poor, but such a tree is less prone to overtraining, therefore when combined, the ensemble of these trees produces a model with high and stable performance.

For the boosting algorithm, we chose the gradient boosting method, as is implemented in the popular MVA package for HEP - TMVA [119]. Gradient boosting can be thought of as a function expansion approach. In this case, each tree corresponds to a summand. In this approach, the parameters for each summand (tree) are determined by minimizing an error function. In the TMVA implementation, the binomial log-likelihood loss function is used. The adopted procedure follows the greedy algorithm approach - only one tree is modified at a time, while the other trees remain invariable.

Several decision trees of the BDT procedure which were employed in the analysis are shown below. This training was done in the di-electron channel. In all figures produced by TMVA, index 0 refers to $H \rightarrow ZZ$ candidate and index 1 refers to $H \rightarrow b\bar{b}$ candidate. For example, “hmass1” denotes the mass of the $H \rightarrow ZZ$ candidate, “hpt0” denotes the p_T of the $H \rightarrow b\bar{b}$ candidate, etc. For the Z boson candidate, the on-shell Z boson decaying to charged leptons, $Z(\ell\ell)$, has either an index 0 or the index is dropped. The closer the node color is to blue, the more the event is classified as signal-like. The same idea is illustrated with the red color for background events.

Di-electron and di-muon channels have been trained separately and the BDT metrics (discussed later in this chapter) show similar performance. To save space, in this chapter we show mostly figures related to the di-electron channel.

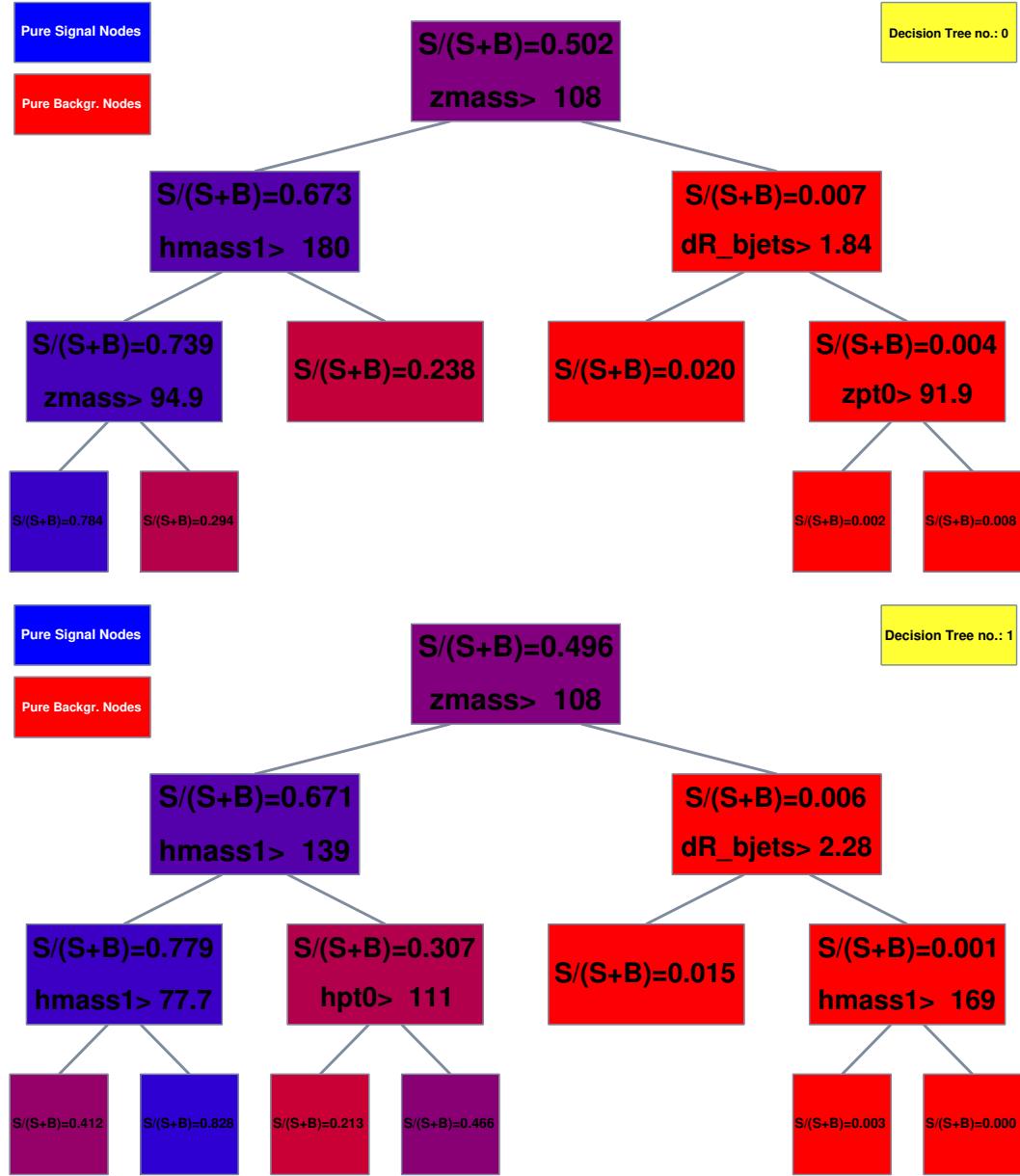


Figure 5.21: BDT trees 1 and 2 (indexing starts from 0). These trees relied on: the mass and the p_T of the $Z(\ell\ell)$ candidate, ΔR separation between two b jets, the mass of the $H \rightarrow b\bar{b}$ candidate, and the p_T of the ZZ^* system. The TMVA notation is explained in 5.7.2.

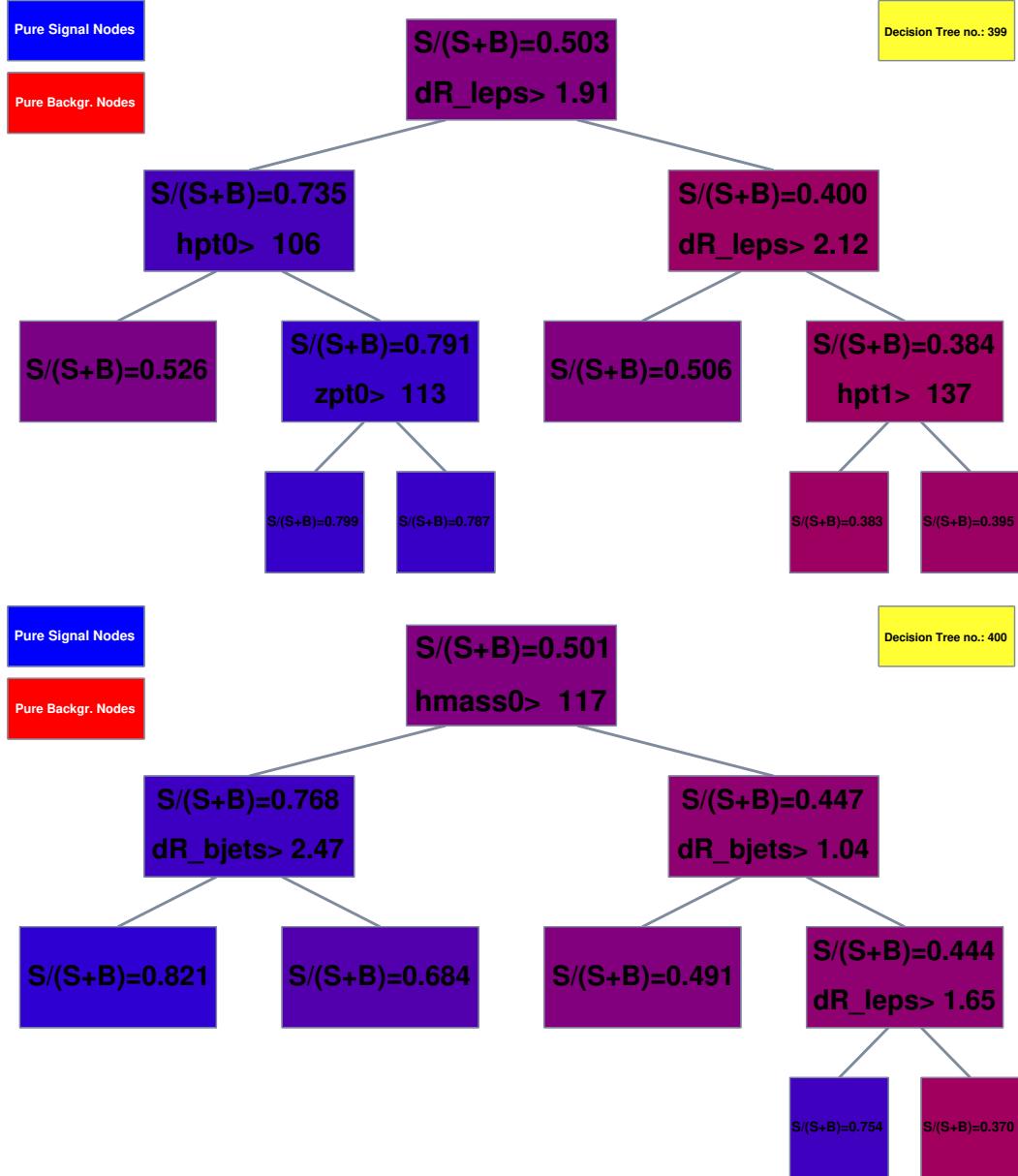


Figure 5.22: BDT trees 400 and 401 (indexing starts from 0). These trees relied on: the p_T of the $Z(\ell\ell)$ candidate, ΔR separation between two b jets, ΔR separation between two charged leptons, the p_T of the $H \rightarrow b\bar{b}$ candidate, and the mass and the p_T of the ZZ^* system. The TMVA notation is explained in 5.7.2.

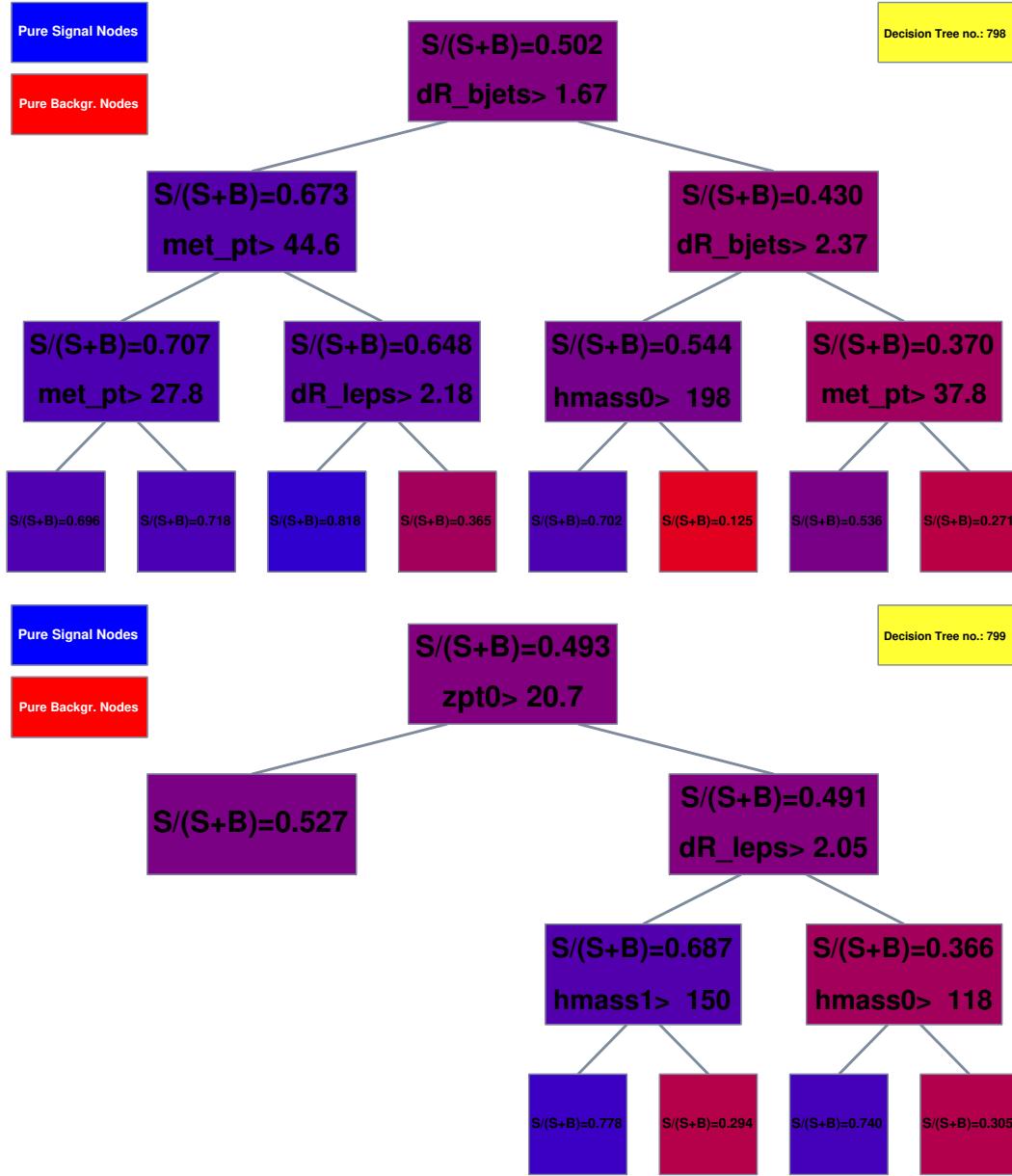


Figure 5.23: BDT trees 799 and 800 (indexing starts from 0). These trees relied on: the p_T of the $Z(\ell\ell)$ candidate, ΔR separation between two b jets, ΔR separation between two charged leptons, the mass of the $H \rightarrow b\bar{b}$ candidate, and the mass of the ZZ^* system. The TMVA notation is explained in 5.7.2.

The BDT discriminant is trained on a pure sample of signal and background events from simulation, using the train data. The properties of the discriminant are studied

on an independent testing sample of pure signal and background events, using the test data. The Deep Neural Network (DNN) approach has also been studied - however, due to the lack of event statistics, the DNN approach could not offer a performance higher than the one achieved with the BDT, and thus was abandoned. Training 16 BDTs per spin hypothesis seemed impractical since our data analysis is dominated by statistical uncertainties. Instead, we follow the approach of other HH analyses [120]: we split the range from 250 to 1000 into two regions: low mass (250 to 450 GeV) and high mass (above 450 GeV up to 1 TeV) regions. This separation boundary at 450 GeV was optimized first using ROC curves (described later in this section), then by running the whole analysis chain all the way to the final limits.

During the training procedure some variables provide more discriminating power and on average are used more often than others. The more often the variable is used, the higher its importance is. The tables below show the ranking and average importance of each of the BDT training. For the low and high mass regions, and di-electron and di-muon channels, see Tables 5.5, 5.6, 5.7, 5.8. As can be seen from these Tables, all nine selected variables are equally important, and one cannot use 1-2 single variables to perform signal-background discrimination. The ranking of variables is compatible between di-electron and di-muon channels within the statistical uncertainty of $O(\text{sub percent})$.

Rank	Variable	Importance (%)
1	$\Delta R_{leptons}$	14.1
2	Mass of $H \rightarrow b\bar{b}$ candidate	13.7
3	$\Delta R_{b\ jets}$	13.2
4	p_T of $H \rightarrow b\bar{b}$ candidate	12.1
5	p_T of Z boson candidate	11.5
6	$p_T^{ZZ^*}$	11.3
7	MET	10.3
8	Mass of ZZ^* system	7.7
9	Mass of Z boson candidate	6.1

Table 5.6: Di-electron channel. Relative importance of the input variables in the high mass BDT training.

Rank	Variable	Importance (%)
1	$\Delta R_{b\ jets}$	13.9
2	MET	12.1
3	Mass of $H \rightarrow b\bar{b}$ candidate	11.9
4	$p_T^{ZZ^*}$	11.0
5	$\Delta R_{leptons}$	10.9
6	p_T of $H \rightarrow b\bar{b}$ candidate	10.7
7	p_T of Z boson candidate	10.2
8	Mass of ZZ^* system	10.1
9	Mass of Z boson candidate	9.26

Table 5.5: Di-electron channel. Relative importance of the input variables in the low mass BDT training.

Rank	Variable	Importance (%)
1	Mass of $H \rightarrow b\bar{b}$ candidate	13.8
2	$\Delta R_{b\text{ jets}}$	13.1
3	$\Delta R_{leptons}$	12.9
4	p_T of $H \rightarrow b\bar{b}$ candidate	11.7
5	$p_T^{ZZ^*}$	11.3
6	p_T of Z boson candidate	11.1
7	MET	11.0
8	Mass of ZZ^* system	8.8
9	Mass of Z boson candidate	6.2

Table 5.8: Di-muon channel. Relative importance of the input variables in the high mass BDT training.

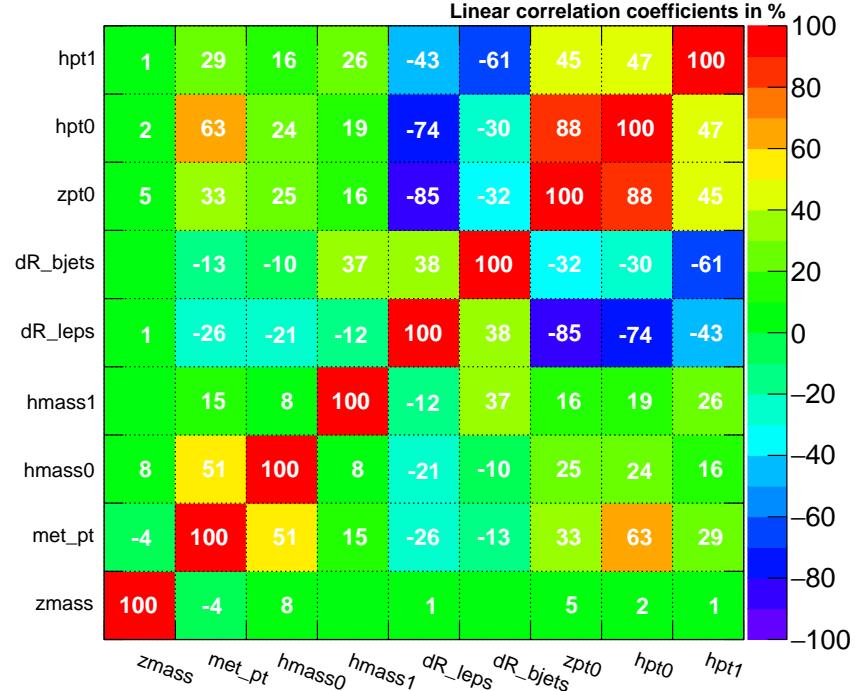
Rank	Variable	Importance (%)
1	$\Delta R_{b\text{ jets}}$	13.0
2	MET	12.2
3	Mass of $H \rightarrow b\bar{b}$ candidate	11.9
4	p_T of $H \rightarrow b\bar{b}$ candidate	11.3
5	p_T of Z boson candidate	11.1
6	$p_T^{ZZ^*}$	10.9
7	$\Delta R_{leptons}$	10.5
8	Mass of ZZ^* system	9.7
9	Mass of Z boson candidate	9.5

Table 5.7: Di-muon channel. Relative importance of the input variables in the low mass BDT training.

The correlations among the input variables are shown in the Figs. 5.24, 5.25. For low mass training, signal samples for mass hypotheses from 250 to 450 GeV are combined together and represent the “signal”. This is done similarly for high mass training - it is a mix of signal samples for mass hypotheses from 500 to 1000 GeV. For background samples, the “background” is a mix of samples of two main background processes: DY in association with jets and $t\bar{t}$ production, weighted by the cross section value. The full selection is applied except for the Z boson and $H \rightarrow b\bar{b}$ candidate invariant mass requirements.

Below we include figures of the BDT input variables. They are the same variables

Correlation Matrix (signal)



Correlation Matrix (background)

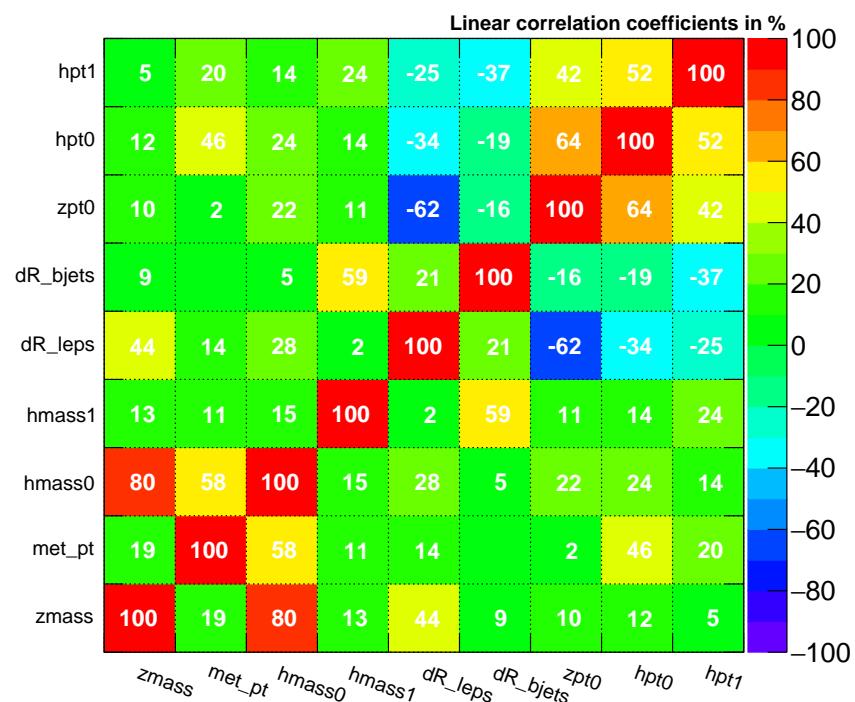
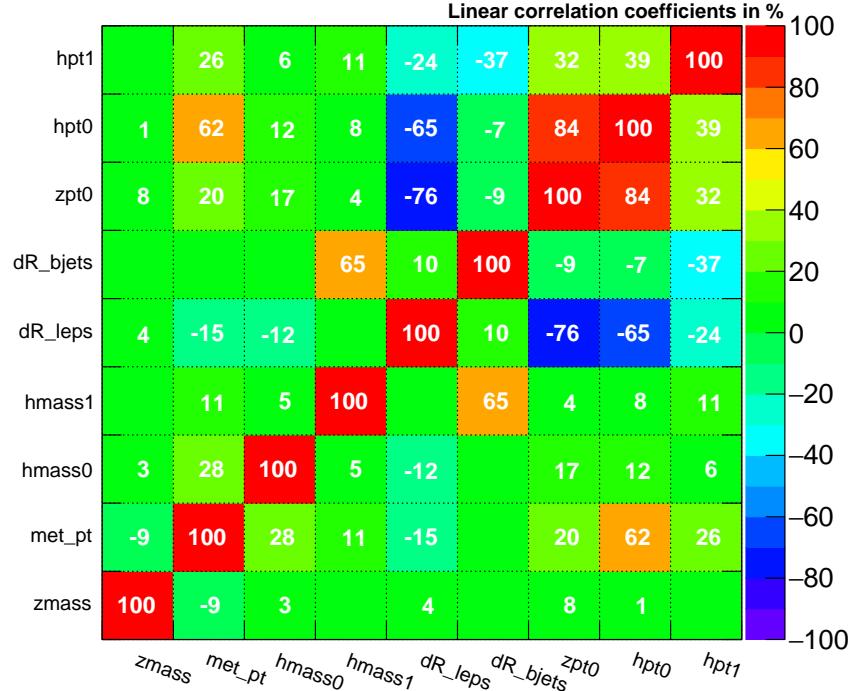


Figure 5.24: Input variable correlations of the di-electron channel, low mass training. Top: signal sample mix. Bottom: background sample mix. The TMVA notation is explained in 5.7.2.

Correlation Matrix (signal)



Correlation Matrix (background)

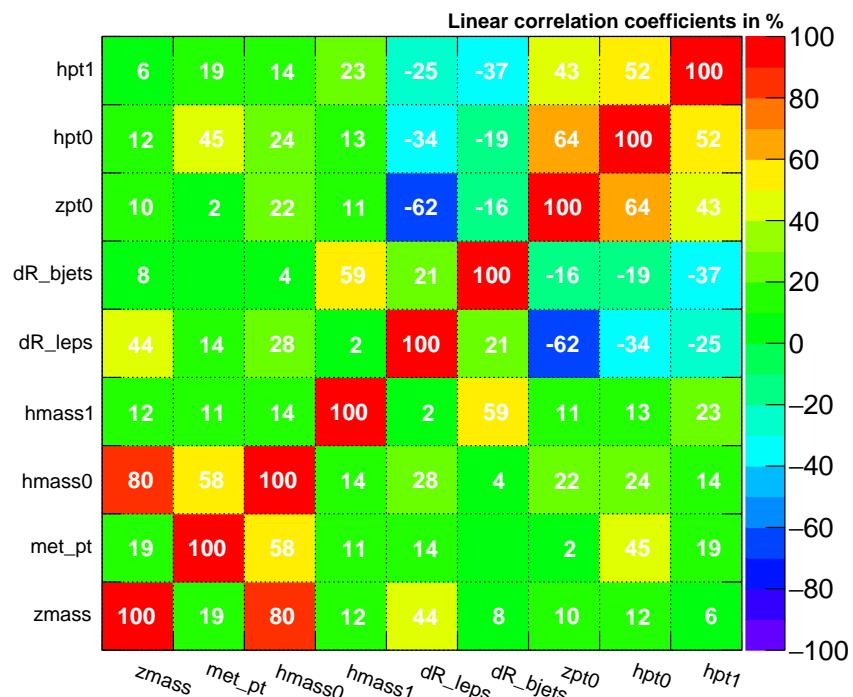


Figure 5.25: Input variable correlations of the di-electron channel, high mass training. Top: signal sample mix. Bottom: background sample mix. The TMVA notation is explained in 5.7.2.

previously shown, but are produced by the TMVA package. In this style they allow for an easier visual judgment of which variables behave with greater difference between the signal and the background, see Figs. 5.26, 5.27.

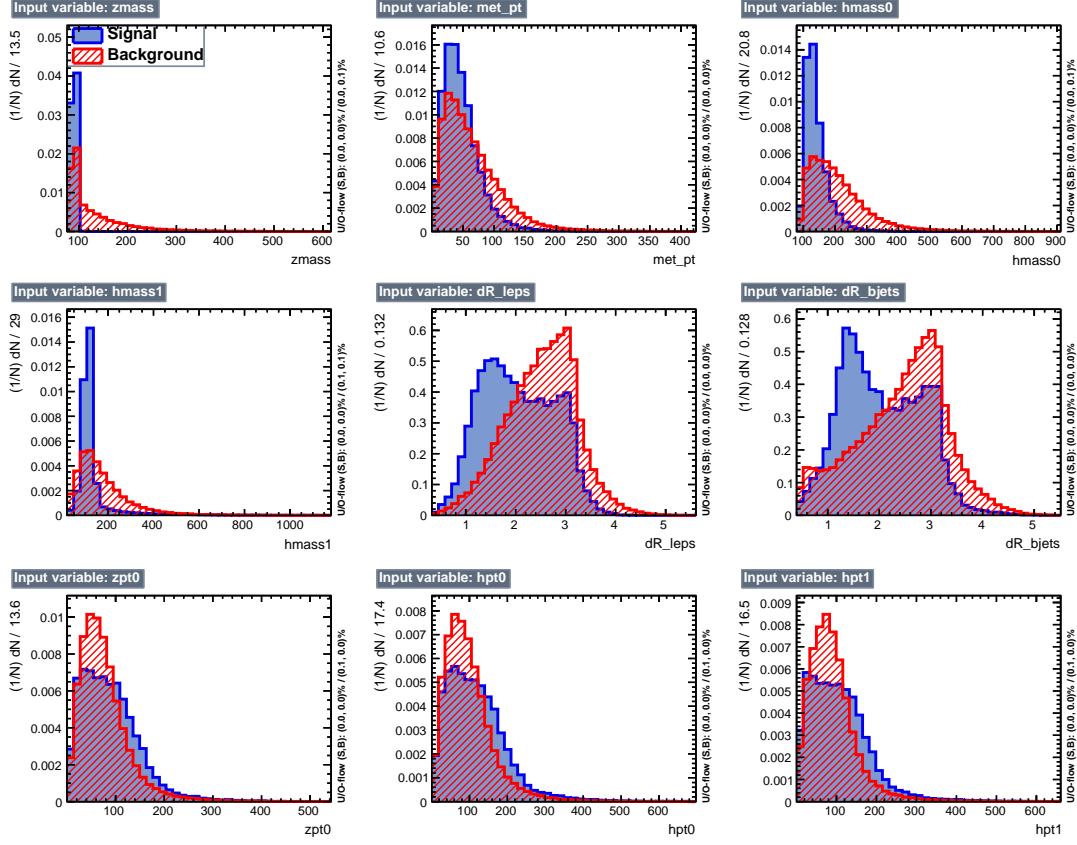


Figure 5.26: Variables used in the low mass training for di-electron channel. Index '1' refers to $H \rightarrow b\bar{b}$ decay and index '0' refers to $H \rightarrow ZZ$ decay. The TMVA notation is explained in 5.7.2.

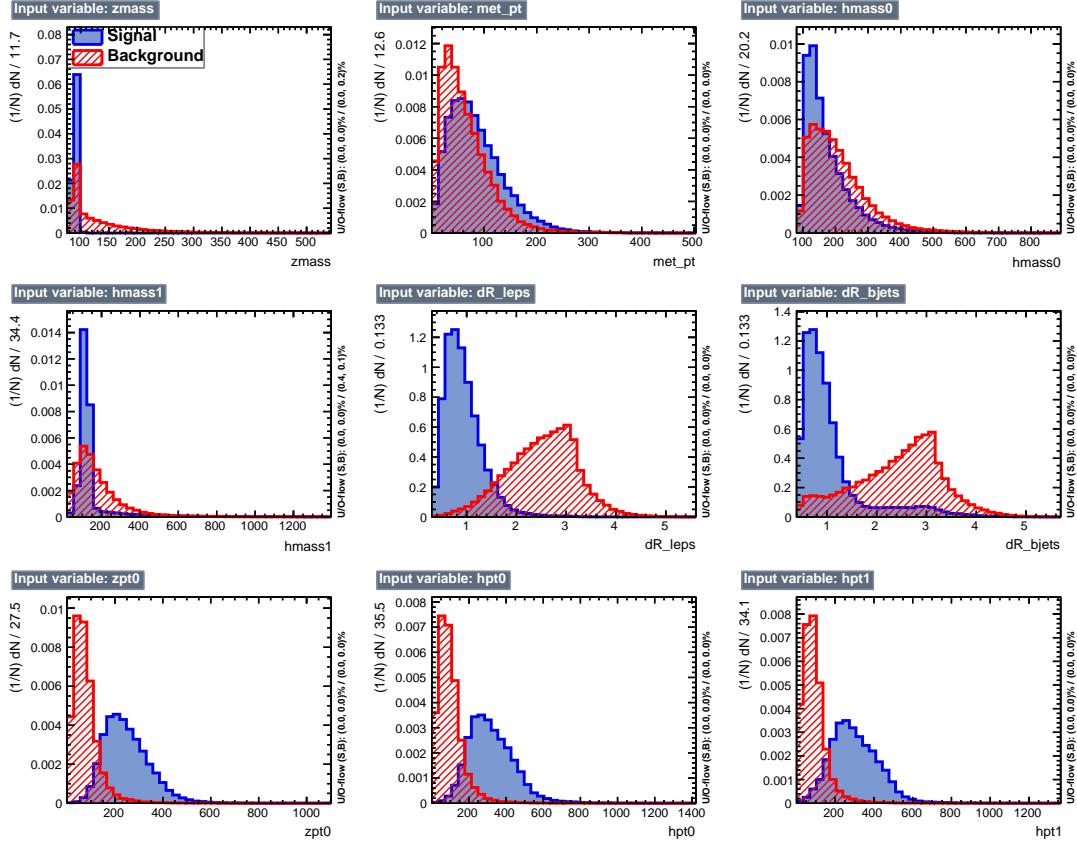


Figure 5.27: Variables used in the high mass training for di-electron channel. Index '1' refers to $H \rightarrow b\bar{b}$ decay and index '0' refers to $H \rightarrow ZZ$ decay. The TMVA notation is explained in 5.7.2.

After the training, the BDT model is created. This model produces the BDT distribution (BDT response), see Fig. 5.28. This figure overlays train and test parameters of the BDT training of the di-electron channel. The parameters of the model, such as the number of trees, the allowed number of splits per variable, the

maximum tree depth and others, have been thoroughly studied. The default TMVA parameters are found to provide a good performance with this measurement.

It is difficult to get high performance during low mass training, since in this region kinematic properties of the background processes are similar to those of the signal, see Figs. 5.26, 5.27. The event yield of background in this region is very high and most variables have similar distributions for signal and background processes. The BDT performance is noticeably better than what can be achieved using a simple linear discriminant method (LD), see Figs. 5.28, 5.29.

Performance of the high mass training can be considered perfect, see Fig. 5.27. High mass resonances produce boosted Higgs bosons whose kinematical properties are different from those of the background processes. This is clearly seen in variables related to the boost, such as $\Delta R_{b\text{ jets}}$, $\Delta R_{leptons}$, p_T of $H \rightarrow b\bar{b}$ candidate, and p_T of $H \rightarrow ZZ$ candidate. Therefore, for the high mass region even the linear discriminant is performing well, see Figs. 5.28, 5.29.

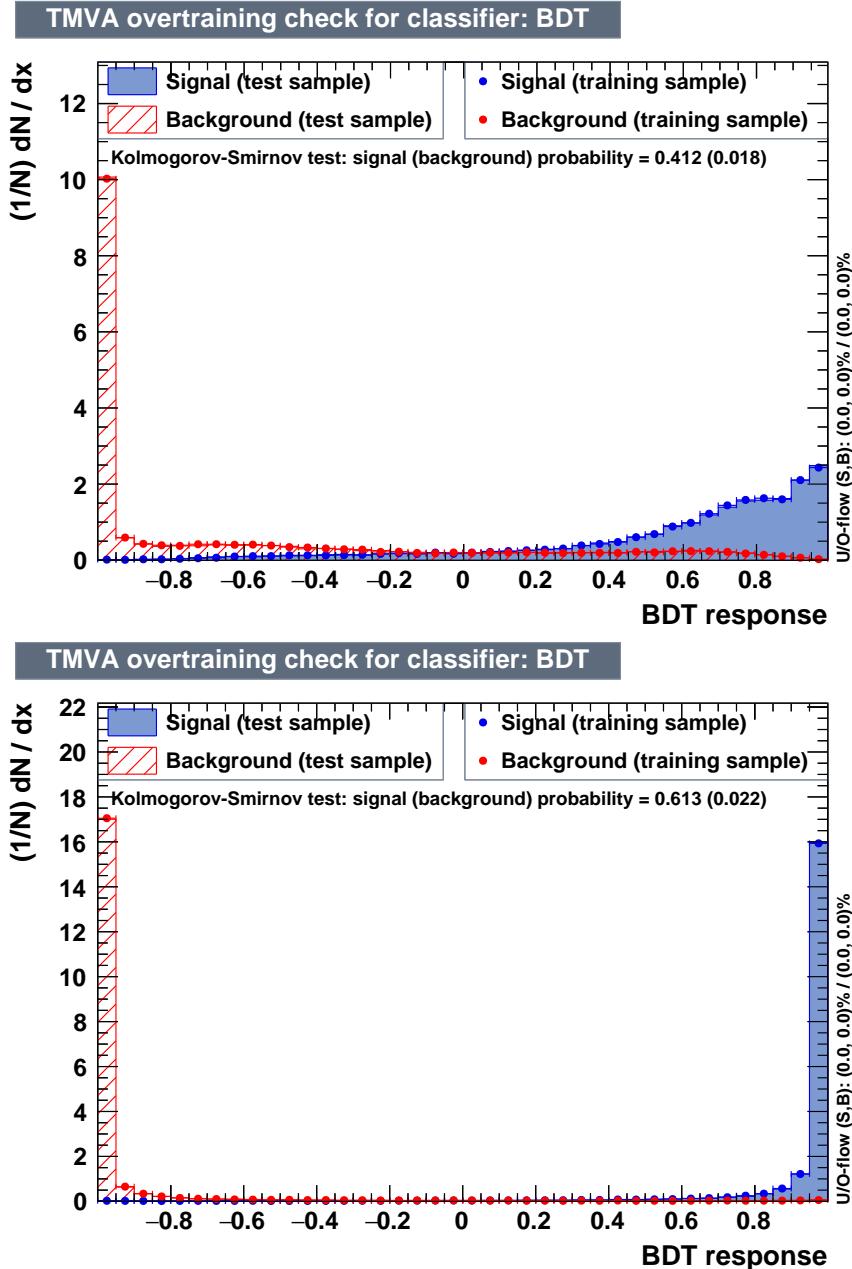


Figure 5.28: BDT discriminants for di-electron channel. Top: low mass training. Bottom: high mass training.

In machine learning, performance measures are needed to compare several models. Usually for a classification problem, one relies on the Receiver Operating Characteristic curve (ROC) and its Area Under The Curve (AUC). These two are the most im-

portant evaluation metrics for checking the performance for any classification model. The ROC is a curve plotted in the space of signal efficiency and background rejection efficiency, where background rejection is equal to one minus the background efficiency. On the 0 to 1 ranges for axes (values of efficiencies), a diagonal line represents the efficiency of the random guess, which is 0.5. A perfect ROC would bend very close to the top right corner. Such a ROC has an AUC almost equal to 1. The ROC curves for di-electron training in low and high mass regions are shown in Fig. 5.29.

The ROC can be regarded as a probability curve and, in this case, the AUC represents the degree or measure of signal-background separability. It tells how much the model is capable of distinguishing between two given classes. The higher the AUC is, the better the model is at predicting signal as a signal-like event and background as a background-like event.

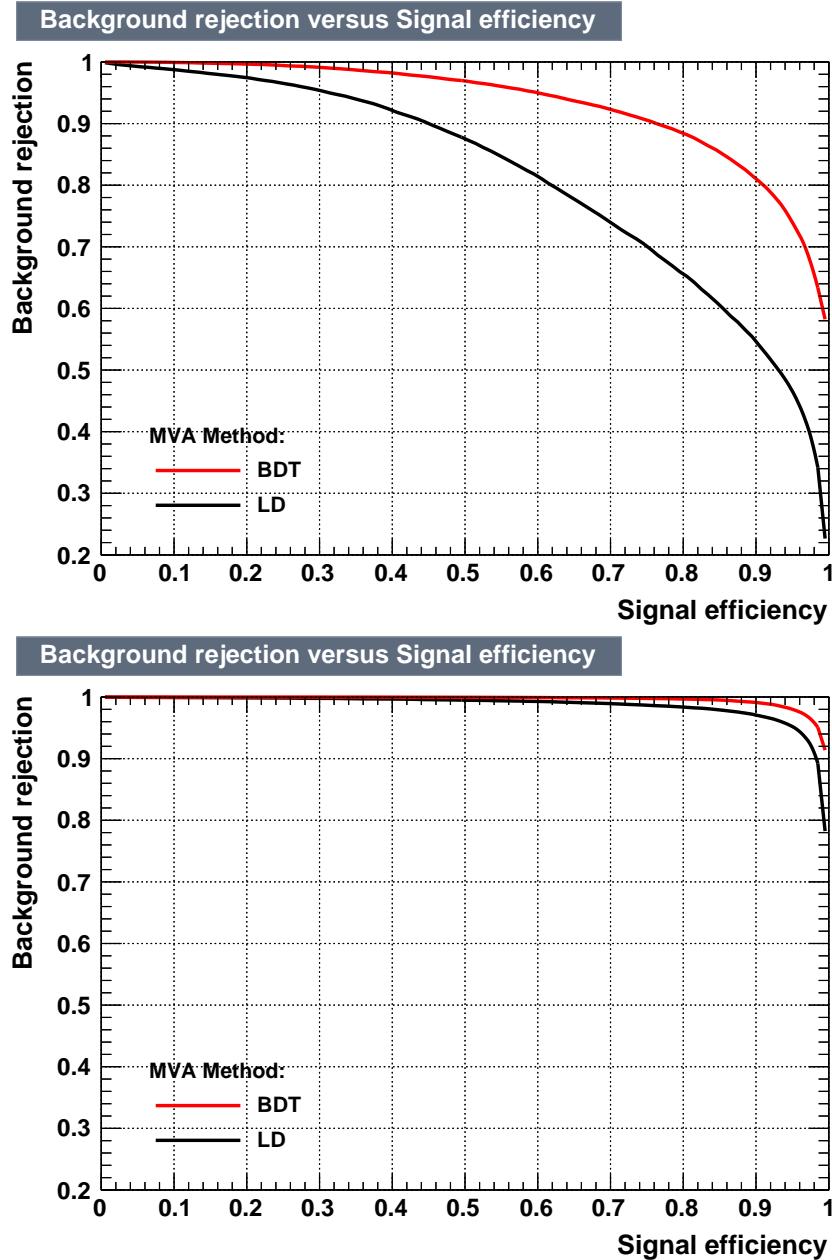


Figure 5.29: ROC curves for di-electron channel. Top: low mass training. Bottom: high mass training.

For two given signal efficiencies, one would prefer the model that has a lower background efficiency; or in other words, a higher background rejection efficiency. In practice, when AUCs are close to 1 and ROCs are similar, one compares AUCs

directly and decides which model to use based on the AUC value.

Finally, we provide the figures of the BDT discriminant in the SR. The plots are shown for the 300 GeV radion decaying in the di-electron and di-muon channels.

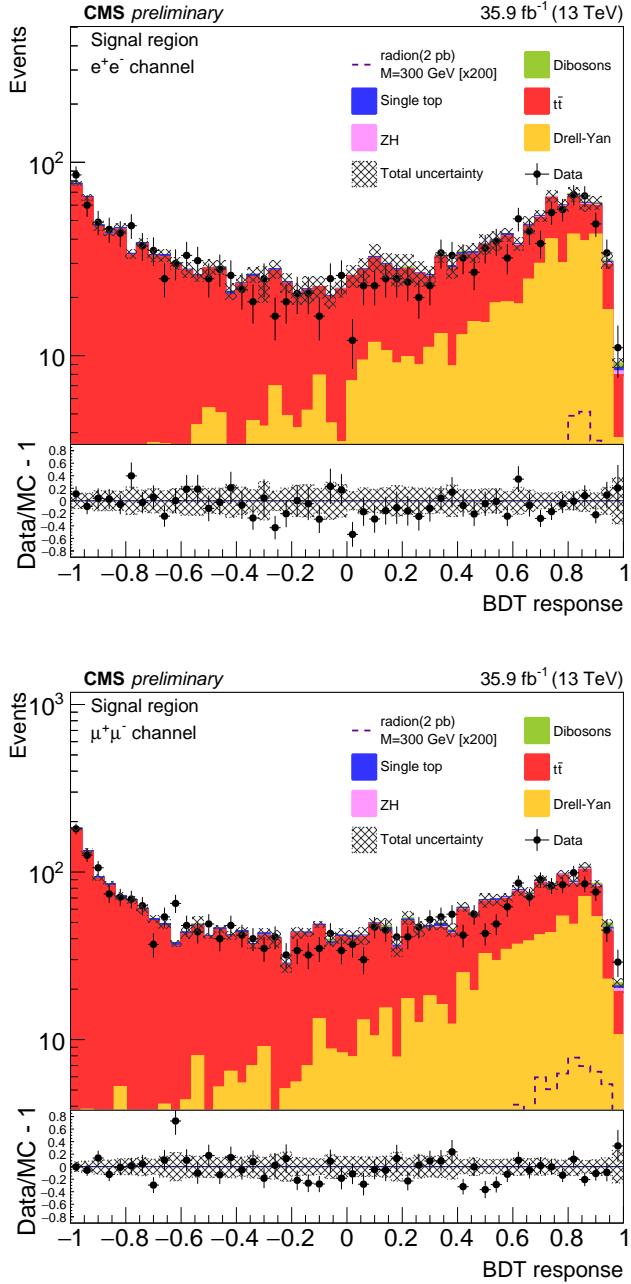


Figure 5.30: BDT distributions (post-fit) for the radion case, electron(muon) channel is shown at the top (bottom). Signal region is presented for the 300 GeV mass hypothesis. For electrons the selection is at 0.4, for muons at 0.7, as is described in Section 5.7.3.

Table 5.9: The BDT selection values used in this measurement.

Channel	260 and 270 GeV	300 and 350 GeV	400 and 450 GeV	500 to 1000 GeV
Di-muon	0.1	0.7	0.7	0.99
Di-electron	0.4	0.4	0.925	0.99

5.7.3 BDT selection requirement in the signal region

To remove the background contribution in the SR, we apply a BDT requirement to di-Higgs candidates. The requirement is specific to the mass hypothesis and specific to the lepton channel; however, the requirement does not depend on the resonance nature. A simple grid search method was adopted to find the best BDT selection for each mass point and channel. During the optimization procedure, we produced the final limits for each value of the considered BDT requirement. Then, we determine the value of the BDT requirement that corresponds to the best expected limit. The optimized values of the BDT selection that are used in this physics analysis are summarised in the Table 5.9. The BDT selection is not applied in the CRs. The BDT selection achieves about a 90 % efficiency for the signal while reducing the background contribution by a factor of few at the low mass region to more than 100 for the high mass region.

The corresponding efficiencies of the BDT selection are presented in Table 5.10. These values are derived for the graviton signal hypothesis and the main background processes for two channels separately.

Table 5.10: Efficiency of the BDT selection in two mass regions and for two channels: ee channel (top) and $\mu\mu$ channel (bottom).

sample	Efficiency at 300 GeV (%)	Efficiency at 900 GeV (%)
signal (bbZZ)	89.2	94.9
signal (bbWW)	75.0	88.4
$t\bar{t}$	28.8	0.2
Drell-Yan	74.2	1.2
Single top	33.1	1.1
ZH	88.8	10.7
Dibosons	90.0	5.0

sample	Efficiency at 300 GeV (%)	Efficiency at 900 GeV (%)
signal (bbZZ)	58.1	91.1
signal (bbWW)	25.9	96.3
$t\bar{t}$	13.6	0.2
Drell-Yan	39.0	0.8
Single top	13.0	0.2
ZH	56.0	8.4
Dibosons	51.4	6.2

5.8 Uncertainties of the measurement

The outcome of the statistical analysis is the limit on the HH production cross section multiplied by the BF of the final state (or just the limit). The statistical analysis uses the predicted event counts for different processes as well as the shape information of the final kinematic variable, the pseudo-transverse mass of the di-Higgs system $\tilde{M}_T(HH)$, for those events. The predicted event counts are computed by multiplying the number of events in simulated samples of different processes by the most accurately predicted cross sections for those processes, while also assigning event weights based on several other factors. However, the results of the analysis are highly affected by statistical uncertainties, and to a smaller degree, by systematic uncertainties. We first discuss the types of uncertainties, how they affect the final limits, and also provide the size of this impact for dominant sources of the uncertainties. Then, we will

discuss the fit procedure and the variable that was used to extract the limits.

Uncertainties that affect the sensitivity of the search come from a variety of sources: statistical uncertainties due to the limited size of the sample statistics, experimental uncertainties related to the accuracy of the detector response, the amount of collected data (uncertainty on the luminosity), differences of simulated samples from real data, and the theoretical uncertainties on cross sections or proton structure. Following CMS practices, below we will divide all uncertainties into two broad classes: the “normalization” and the “shape” uncertainties. The former modifies only the event yields (or “normalizations” for brevity) of selected events from different processes. The latter may also distort the shape of the distribution of the final kinematic variable that is used in the extraction of the limits.

To compute the effect of a particular source of the systematic uncertainty on the final limit, we use the up/down method. This is a procedure of varying a factor affecting the expected yields up or down by the amount of the estimated uncertainty on the effect and examining how the yields change.

5.8.1 Normalisation uncertainties

The sources of systematic uncertainties that affect normalization are discussed below. The sizes of certain types of uncertainties vary depending on the resonance mass hypothesis and the type of the channel (di-electron or di-muon). In this case, ranges of the uncertainty values are listed. Normalization uncertainties listed in this section modify all background processes but $t\bar{t}$ and DY. The levels of those backgrounds are determined from data during the statistical analysis that uses two control regions dominated by these two background processes.

- **Luminosity.** CMS estimated the uncertainty on the integrated luminosity of the 2016 data set to be 2.5%, see [121]. This uncertainty directly affects the expected event yields for both signal and background processes whose yields are directly normalized to the integrated luminosity (excluding two dominant ones: DY and $t\bar{t}$).
- **Pileup.** How well the PU interactions are replicated in the MC simulation affects the signal and background event yields. The effect of this uncertainty on normalizations is approximately 6%. The effect is estimated by varying the number of pileup interactions in simulated samples up and down within the range, reflecting the limited knowledge of the total inelastic proton-proton interaction cross section at 13 TeV.
- **Proton PDF.** The impossibility to know precisely the structure of the interacting protons is evaluated using an ensemble of PDF replicas from the NNPDF set [110] following the PDF4LHC prescription [122,123]. The root mean square value of the expected event yields in simulated samples, computed with the PDF replicas from this ensemble, is taken as a measure of the uncertainty. The size of the effect is found to be of order 5%.
- **Theoretical uncertainties of the QCD scales.** The uncertainty associated with the theoretical uncertainties in the QCD factorization and renormalization scales is evaluated. This uncertainty affects the expected yield of the signal and background events (except the $t\bar{t}$ and DY processes, which normalizations are fixed to the values determined during the fit). This uncertainty is computed by independently varying the factorization and renormalization scales, used as parameters in the Pythia event generator. The variation is done in the MC simulation following a standard CMS prescription given by the PDF4LHC [124].

The scales in signal samples are varied by a factor of 2 up and down - multiplying the original nominal scale values by two corresponding factors: 0.5 and 2. The unphysical cases, when one of the scales fluctuates up, while the other fluctuates down, are not considered. In each bin of the final kinematic variable that is used in the statistical analysis, the maximum and minimum variations of the yields are used to build an uncertainty region around the nominal shape. The size of the envelope indicates that these scales affect the yields of the processes at a level of 4-6%.

- **Theoretical values of the cross sections.** The uncertainty on the theoretical cross section value of the single top, ZH, and diboson production is propagated to the background yields of these background processes and used to normalize their event yield. The uncertainty is 5–7%.
- **Missing transverse momentum.** During the clustering of the jets and a subsequent application of JEC and JER corrections, neutral hadrons and photons that do not belong to any jet (“unclustered energy”) and jets with transverse momenta below 10 GeV - lack such corrections. This affects the MET reconstruction and results in a small systematic uncertainty. The effect of JEC on the unclustered energy and subsequently on the magnitude of the \vec{p}_T^{miss} is studied. The effect is propagated to the final kinematic variable, which contains the \vec{p}_T^{miss} . The effect of this uncertainty is studied by shifting the energy of each particle not contained in jets or contained in low- p_T jets by its uncertainty, which is estimated during particle’s reconstruction. These variations are found to affect the event yields of signal and background processes at the level of 3%; however, they do not show a visible effect on the shape of the final kinematic variable. This uncertainty is categorized as a “normalization” systematic source.

5.8.2 Shape uncertainties

Certain types of systematic uncertainties (later referred to as “shape uncertainties”) distort the shapes of some kinematic distributions and BDT discriminant distributions. As a consequence, the shape of the final discriminating variable is modified. The parameters defining each source of the uncertainty are varied within one standard deviation up and down. The effect is propagated through all variables included in the construction of the final kinematic variable used in the statistical analysis procedure. This method prepares three shapes: in addition to the nominal shape of the final variable, it also produces two modified shapes corresponding to the up and down variations of the parameters. All these shapes are then used in the statistical analysis of the data, see 5.9. Below we discuss each source of the shape uncertainty individually.

- **Lepton efficiency.** Data-MC simulation differences in the lepton reconstruction in the tracker, identification and isolation selection criteria, and in the efficiencies of the HLT trigger requirements are computed and used to correct the MC samples. The final statistical analysis relies on the MC correctly simulating the efficiency for an event to be reconstructed and to pass online and offline selection requirements. The lepton scale factors are derived from high-statistics samples of Z boson decays. This simulation is not perfect, thus, corrections are applied to candidates as per-candidate (or per-event) weights, as explained later in this chapter. These corrections, though, are also known with limited accuracy. The accuracy as well as the size of the correction depend on p_T and η of the leptons. The uncertainty of lepton efficiency corrections as a function of lepton p_T and η is propagated to the final kinematic distribution. The effect of these uncertainties is sub-percent for the muon channel and up to 6% for some

bins of the electron channel $\tilde{M}_T(\text{HH})$ distributions.

- **Jet energy scale.** The JEC directly affects the invariant mass and p_T of the $H \rightarrow b\bar{b}$ candidates and the magnitude of the \vec{p}_T^{miss} . Both objects are used in the construction of the final kinematic variable; therefore the JEC effect is thoroughly studied. The JEC is varied within one standard deviation of its uncertainty as a function of jet p_T and η , and the effect on the b jet kinematics and on the \vec{p}_T^{miss} is calculated. This effect is propagated through the steps of the measurement yielding the variation of the final variable shape. The JEC uncertainty has an effect on the yields of the signal and background processes at a level of 5 to 10%.
- **Jet energy resolution.** The shapes of the final kinematic distributions are affected by the difference between the jet energy resolution in data and simulation, because this bias modifies the mass and the p_T of $H \rightarrow b\bar{b}$ candidates and the magnitude of the \vec{p}_T^{miss} . The JER is varied in simulation by one standard deviation as a function of jet p_T and η , and the effect is propagated through the steps of the measurement. This uncertainty affects the final distribution at the order of 0.5%.
- **b tagging.** The data-mc difference in the efficiency to tag a b jet (and the probability to misidentify a light flavor or a gluon jet as a b jet) is calculated and MC samples are corrected by the corresponding SFs, see Section 5.5.3. They are derived using heavy-flavor enhanced jet MC samples. The uncertainties on these SFs are propagated through the measurement steps to the final variable. The effect of the b-tagging efficiency (flavor misidentification) is the highest for DY process, at about 5% (7–10%), and is at the sub-percent level for other processes (7–10%).

- **Bin-by-bin uncertainties (BBB).** This is a statistical uncertainty. Out of hundreds of uncertainty sources in this measurement, the most dominant ones are related to the limited size of simulated samples. In certain cases, these statistical uncertainties can produce sizeable fluctuations of the bin content of the shape of the final variable. Because of this, we add an individual nuisance parameter for each bin of the final variable distribution. For each bin in the bins of the final variable, the corresponding bin-by-bin uncertainty, defined as a single Gaussian-constrained Barlow-Beeston-lite parameter [125,126], is created and used to scale the total yield in a given bin.

5.8.3 Leading uncertainties

The summary of the sources of systematic uncertainties is given in the Tables 5.11, 5.12. As can be seen there, the leading systematic uncertainties are associated with jets: b-tagging and b jet mistagging, and jet energy resolution and scale. It is worthwhile noting, that out of 300 hundred uncertainties in this measurement, the most dominant ones - the first 30 - 40 - are predominantly statistical uncertainties (BBB) 5.8.2 on the bin content of the $\tilde{M}_T(\text{HH})$ distributions that are used in the fit.

Table 5.11: The effect of leading systematic uncertainties on the total event yield for main signal and background processes, ee channel, 300 GeV graviton mass hypothesis.

Sample	b-tagging	mistagging	electron ID and ISO	electron tracker eff.	electron trigger eff.	JER	JEC
DY	4.3	7.4	5.4	1.1	2.1	0.2	5.3
$t\bar{t}$	0.5	7.4	4.7	1.1	1.9	0.0	0.5
Signal (bbZZ)	0.2	7.6	5.0	1.1	2.0	0.7	5.8
Signal (bbWW)	0.0	7.6	6.7	1.1	2.9	0.0	1.6

Table 5.12: The effect of leading systematic uncertainties on the total event yield for main signal and background processes, $\mu\mu$ channel, 300 GeV graviton mass hypothesis.

Sample	b-tagging	mistagging	muon ID	muon ISO	muon tracker eff.	muon trigger eff.	JER	JEC
DY	4.9	7.0	0.2	0.1	0.1	0.4	0.2	9.4
$t\bar{t}$	0.9	7.2	0.2	0.1	0.0	0.4	0.6	0.7
Signal (bbZZ)	0.3	7.7	0.2	0.1	0.0	0.4	0.5	4.4
Signal (bbWW)	0.0	9.2	0.2	0.1	0.0	0.1	0.0	8.5

5.9 Statistical Analysis

As was mentioned in Section 5.2, the final variable ($\tilde{M}_T(\text{HH})$) is constructed as the sum of the Lorentz vectors of the $Z(\ell\ell)$ candidate, the $H \rightarrow b\bar{b}$ candidate, and \vec{p}_T^{miss} vector. This variable, a pseudo-transverse mass of the double Higgs system, is used to extract the results in this measurement. In this section, we discuss how the distributions of the $\tilde{M}_T(\text{HH})$ variable are used in the statistical analysis to obtain the limits on HH production multiplied by the BFs of the final state.

Below we present the expected event yields and observed data counts in the SR, see Table 5.13.

Table 5.13: Expected event yields and observed data counts in the SR, split by channel. 300 GeV graviton mass hypothesis. Signal is normalized to 2 pb cross section. BDT selection is 0.4 for the ee channel, and 0.7 for $\mu\mu$ channel.

Process	ee channel	$\mu\mu$ channel
Signal	0.09	0.12
DY	341.55	318.32
$t\bar{t}$	261.42	216.85
ZH	7.27	10.62
Single top	6.79	5.99
Dibosons	7.64	9.08
Data	598	513

5.9.1 The likelihood function

The results in this measurement are obtained with a maximum likelihood fit. We perform a simultaneous fit of the SR and both CRs for both di-electron and di-muon channels using the likelihood function constructed as a product of Poisson terms over all bins of the input $\tilde{M}_T(\text{HH})$ distributions in the three regions (SR, CRDY, CRTT) with Gaussian terms to constrain the nuisance parameters:

$$L(r_{\text{signal}}, r_k | \text{data}) = \prod_{i=1}^{N_{\text{bins}}} \frac{\mu_i^{n_i} \cdot e^{-\mu_i}}{n_i!} \cdot \prod_{j=1}^{N_{\text{nuisances}}} e^{-\frac{1}{2}\theta_j^2}$$

where the product index i refers to the bin of the input distributions, the product index j refers to uncertainties accounted for by the fit model, and n_i is the number of observed data events in the bin i . The mean value for each of the Poisson distributions is computed as:

$$\mu_i = r_{\text{signal}} \cdot S_i + \sum_k r_k \cdot B_{k,i},$$

where k refers to the background process k , and $B_{k,i}$ is the content of the bin i of the background shape for a process k , while S_i is the content of the bin i of the signal shape. The S and B_k distributions are prepared using MC simulations with the best-known cross section values. This is done to have the number of expected MC events equal to the number of events in the analyzed data set (with the corresponding integrated luminosity). When filling the distributions for S and B_k , we also use a set of weights composed of scale factors of various types that depend on kinematic parameters of the candidates.

The parameter r_k sets the normalization of the background process k while r_{signal} is the signal strength parameter. All r parameters float freely in the fit. Two values of the signal strength parameter are of special interest: $r_{\text{signal}} = 0$ describes the background-only scenario, while $r_{\text{signal}} = 1$ corresponds to the case where the HH cross section matches the cross section used for the initial signal normalization inspired by BSM models, 2 pb in our case, which is a typical value for predictions of WED models (e.g.,

at 300 GeV). The terms θ_j represent the set of nuisance parameters that are introduced into the likelihood function as Gaussian constraints. In the final likelihood function, we have a nuisance parameter for each source of systematic uncertainty associated with each physics object and for each normalization and shape uncertainty for a given process. Each normalization nuisance parameter describes a relative difference in the event yield of the process before and after the fit to data; and each shape nuisance parameter quantifies the difference in shape of the $\tilde{M}_T(\text{HH})$ distribution obtained with interpolation or extrapolation of the nominal, up, and down shapes (see Section 5.8.2) corresponding to the variation of a given systematic effect to achieve the best fit to the data.

The fit is performed separately for each of the 16 resonance mass points in the range 250 to 1000 GeV for a radion and a graviton resonance hypothesis individually. During the fit, the values of the r_{signal} and r_k , along with the uncertainties, are determined.

The implementation of the maximum likelihood fit is done in the Higgs Combine package, a software framework supported by the CMS experiment [127]. The framework is based on the RooStats package [128] that is widely used in the HEP community.

Figure 5.31 shows the result of the fit for one mass point - 300 GeV. These HH transverse mass distributions for the graviton hypothesis are of the post-fit type - the normalizations and shapes of all components were adjusted according to the best-fit values. A similar figure in the case of the radion hypothesis is shown in Fig. 5.32. The signal component is very close to zero, it would not be visible on the figures. Therefore, the signal event yield is further scaled by a certain high factor for visual purposes, to make the signal contribution clearly visible.

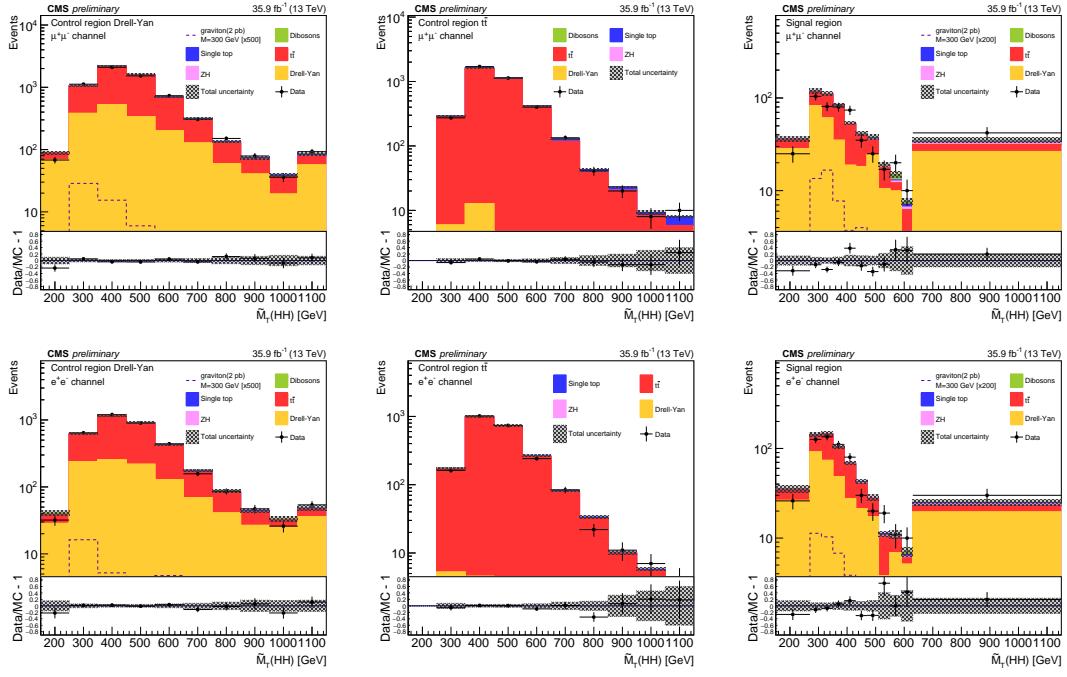


Figure 5.31: Transverse mass of the reconstructed HH candidates for data, the simulated signal for the 300 GeV mass hypothesis of the graviton, and simulated background processes. Event yields of the MC simulations are scaled according to the fit results. The top row shows the figures for the di-muon channel, while the bottom row is for the di-electron channel. For each row, the left plot is for the CRDY, the middle is for the CRTT, and the right is for the SR. The signal scaling choice is discussed in the text. The crosshatched area represents the sum of statistical and systematic uncertainties.

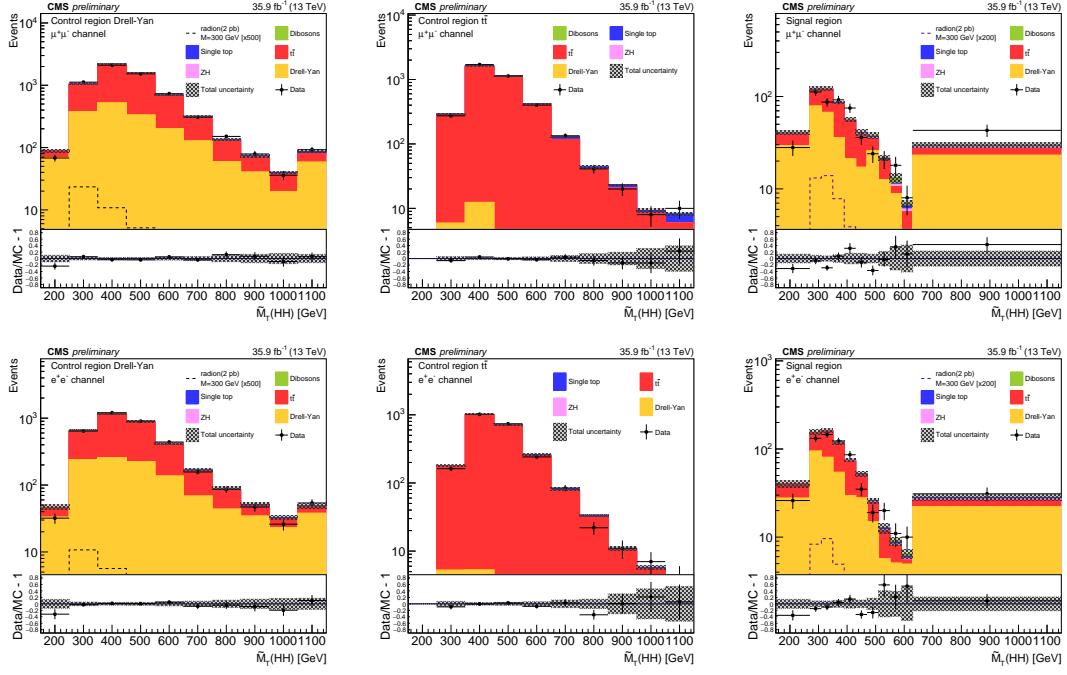


Figure 5.32: Transverse mass of the reconstructed HH candidates for data, the simulated signal for the 300 GeV mass hypothesis of the radion, and simulated background processes. Event yields of the MC simulations are scaled according to the fit results. The top row shows the figures for the di-muon channel, while the bottom row is for the di-electron channel. For each row, the left plot is for the CRDY, the middle is for the CRTT, and the right is for the SR. The signal scaling choice is discussed in the text. The crosshatched area represents the sum of statistical and systematic uncertainties.

5.9.2 The fit results

Based on fit results, we proceed setting upper 95 % confidence level limits on the HH production cross section multiplied by the BF of the final state, using the modified frequentist CL_s approach (asymptotic CL_s) [?, 129–131].

It is worthwhile letting the reader know that before the real data is used to derive observed limits, the toy data set is first used to optimize the expected limits. This toy data set, called the Asimov data set, is commonly used by CMS and ATLAS experiments for statistical analysis procedures. The Asimov data set (see [131]) represents

the hypothesis, which in most cases is the expected background with no fluctuations. The signal strength is set to zero, which, in a simple counting experiment [132] corresponds to the total number of events equal (if there are no fluctuations) to the number of background events (no signal events).

In this physics analysis, we follow a standard CMS practice that helps experimenters avoid bias in the measurement - a blinding procedure. First, the measurement is performed only with a simulation. The expected sensitivity and the expected limits (with the Asimov data set) on the quantity being measured are determined, and the procedure is frozen. Second, the likelihood fit is performed and a set of diagnostic figures is prepared with real data, including all kinematic distributions shown in Section 5.7.1, the BDT outputs 5.30, and $\tilde{M}_T(\text{HH})$ distributions in the signal region 5.31, 5.32. The figures are examined for data and simulation agreement. At this stage, the results are kept blind and no real data are used.

Finally, the real data is used and all outputs of the fit are examined. We look at the distributions in the SR and determine r_{signal} . The data-MC agreement is once again verified for the input variables to the BDT. This is an additional way to check that fit results are correct, see Figs. 5.33, 5.34, 5.35, 5.36, 5.37.

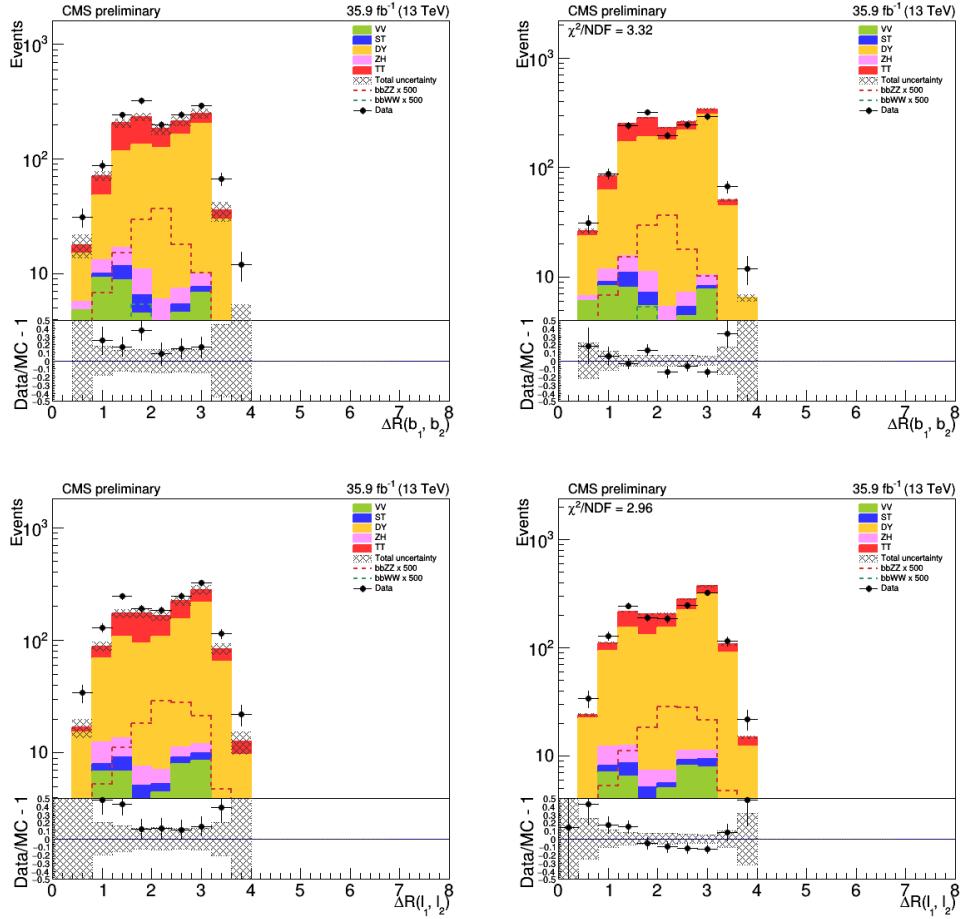


Figure 5.33: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, SR region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

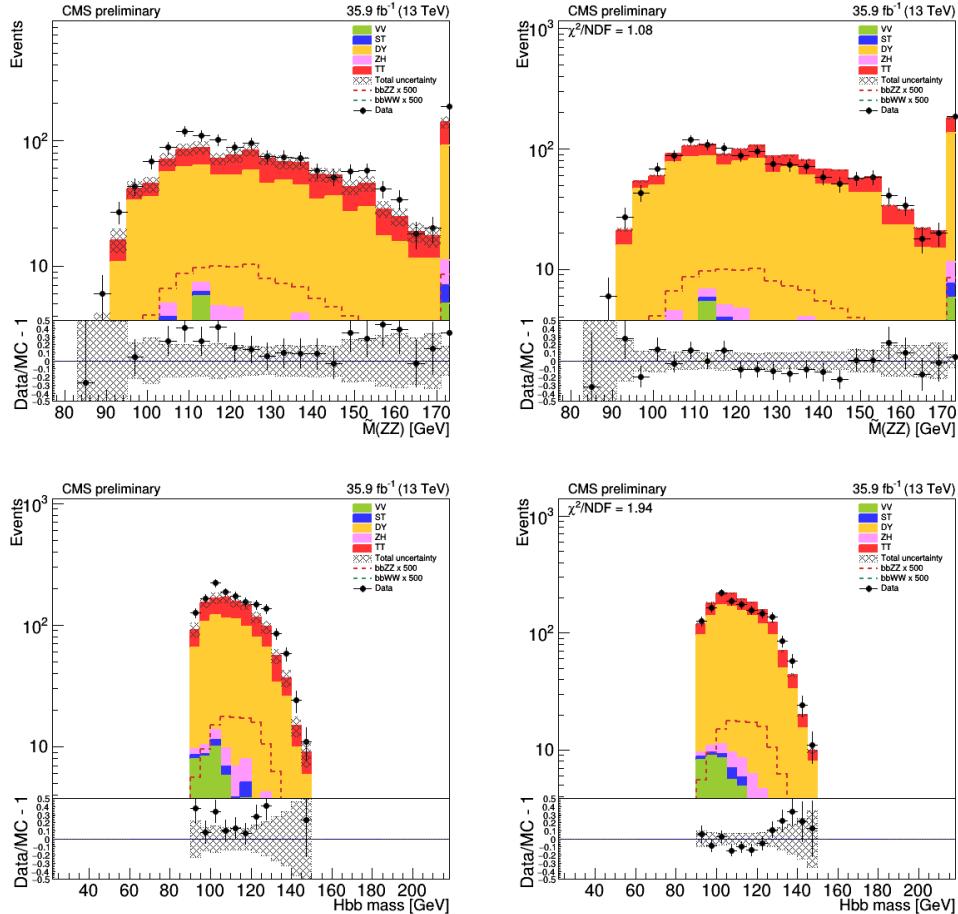


Figure 5.34: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, SR region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

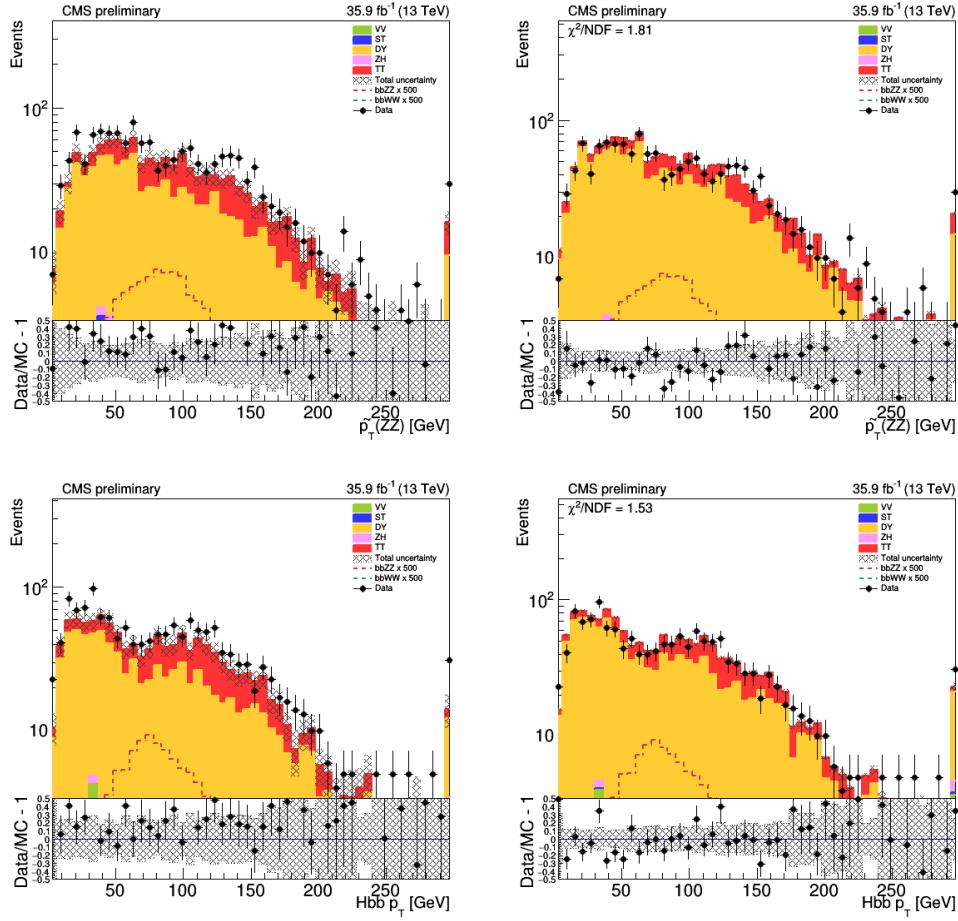


Figure 5.35: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, SR region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

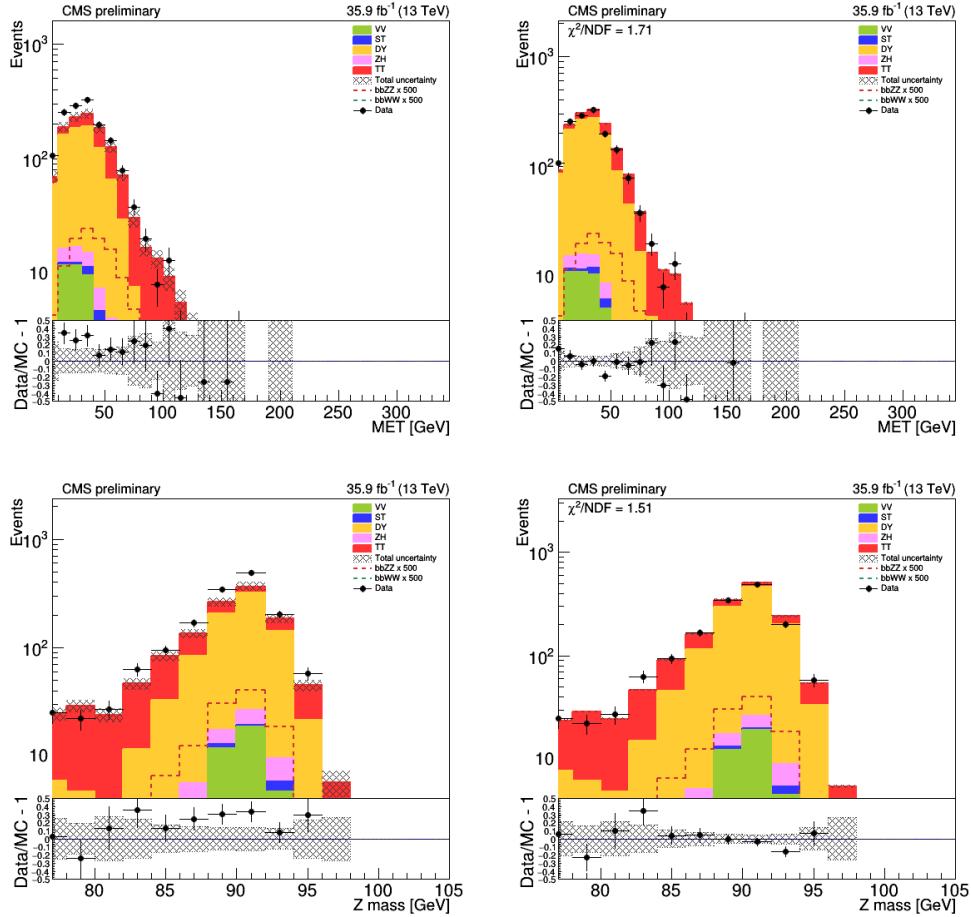


Figure 5.36: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, SR region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

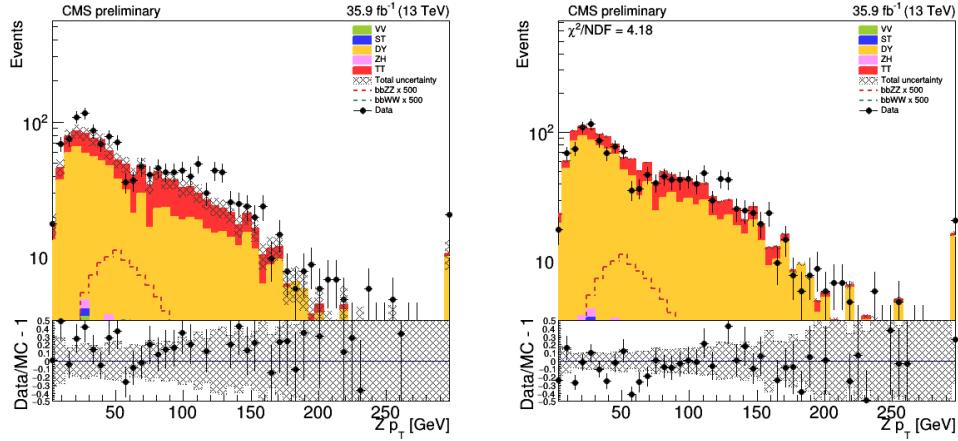


Figure 5.37: Comparison of data and MC samples. 300 GeV graviton mass hypothesis, SR region, $\mu\mu$ channel. Pre-fit plot on the left; Post-fit plot on the right.

The observed and expected 95% upper CL limits for the full mass range and both resonances are listed in Table 5.14. Produced figures of limits are shown in Fig. 5.38. The green and yellow bands correspond to one and two standard deviations around the expected limit respectively. Since 450 GeV is the separation boundary between two mass regions, low mass and high mass - in which different BDT discriminants are applied - the limit calculation is performed with both of the BDTs at 450 GeV, where the discontinuity is seen in the figure. The figure also shows the expected production cross section for an RS1 KK graviton and RS1 radion in WED models. These cross sections are computed in [41]. As seen from the figure, this measurement is not sensitive by itself to probe the chosen model of HH production. The combination of this measurement with other similar measurements is necessary, as discussed in Section 5.10.

Table 5.14: The expected and observed HH production cross section upper limits at 95% CL for different narrow resonance graviton (top) and radion (bottom) mass hypotheses for both dielectron and dimuon channels combined.

Mass, GeV	Observed Limit (pb)	Expected Limit (pb)
250	253.5	589.1
260	272.2	585.9
270	274.4	537.5
300	380.0	434.4
350	330.6	309.4
400	90.4	119.9
450	59.8	63.3
500	31.0	36.6
550	14.5	20.2
600	9.8	12.7
650	18.5	11.1
700	16.1	10.1
750	13.7	8.8
800	10.1	6.5
900	8.1	4.8
1000	5.8	4.2

Mass, GeV	Observed Limit (pb)	Expected Limit (pb)
250	107.3	297.7
260	170.8	410.9
270	207.0	470.3
300	451.7	496.9
350	532.6	496.9
400	155.7	171.1
450	89.3	82.0
500	36.0	54.4
550	18.7	28.5
600	13.2	19.6
650	24.6	17.2
700	16.4	12.0
750	13.9	10.4
800	12.6	9.8
900	6.9	5.6
1000	5.7	4.5

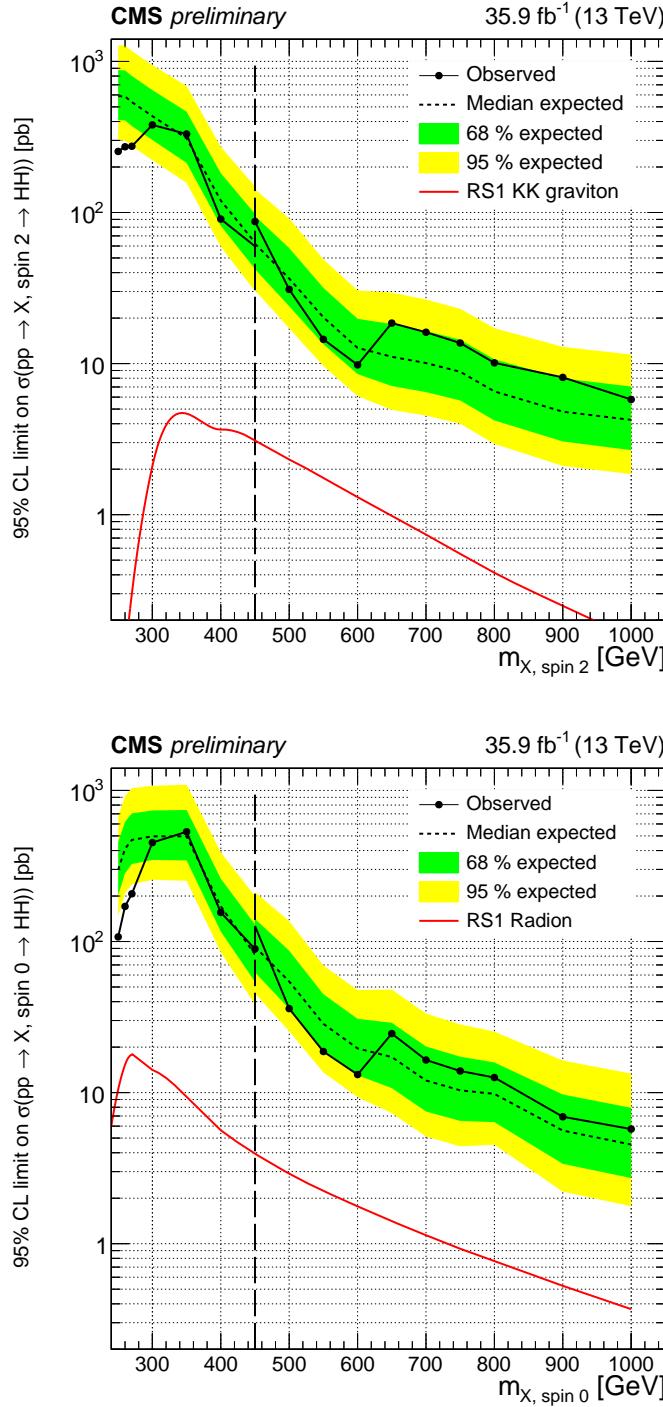


Figure 5.38: Expected (dashed line) and observed (solid line) limits on the cross section of a resonant HH production as a function of the mass of the narrow resonance for both leptonic channels combined. The graviton case is shown at the top and radion case at the bottom. The red line shows a theoretical prediction for the production of a WED particle with certain model assumptions [41].

5.10 Discussion

In searches for a process, when intermediate particles can decay through a variety of channels, a common experimental approach is to search for multiple final states in parallel. Then the results are combined into a single result. This is a common method in the measurement of limits on a cross section of a physics process such as the production of a heavy resonance.

At CMS, the Higgs boson program searches for HH production, both of the SM and BSM types. These searches cover multiple channels, including this measurement, and are published as individual papers. For di-Higgs measurements, the individual channels include many of the channels seen in Fig. 2.6. This is usually followed by the final paper that presents the combined result of all individual channels, the “grand combination”.

This measurement, the search for a resonant production of the HH system decaying through the intermediate $bbZZ$ state into $2b2\ell2\nu$ final state is completed, approved by the CMS collaboration [86], and is shown at conferences [133]. This was the first measurement of the HH production in the $bbZZ$ intermediate state that was performed at CERN using the LHC data.

This data analysis, at the time of this writing, is being combined with a similar channel where the HH system decays through the same intermediate state $bbZZ$ into the final state of $2b2\ell2q$, and a paper will be submitted to the Physical Review D shortly after the defence of this dissertation.

While this measurement by itself is not able to exclude the RS1 model in search, this physics analysis combined with the other HH channels is able to exclude a part of the parameter space of the WED theory. As of this writing, the best available grand combination from the CMS experiment is based on a partial 13 TeV data set

analysis, see Fig. 5.39, published in [134] that includes a subset of all channels. That paper does not include the present measurement, which would be found just above the $bbWW$ graph, which is the curve shown in orange. As can be seen from the figure, the most sensitive channels are the $bb\gamma\gamma$ for the mass region below 500 GeV and the $bbbb$ channel for higher masses.

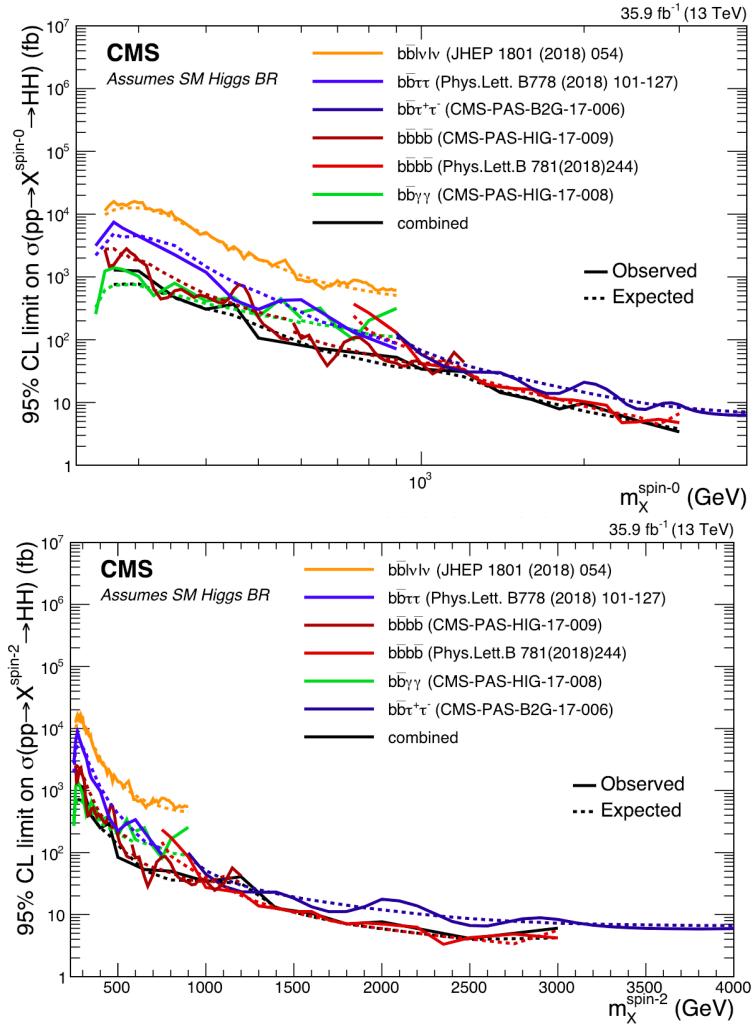


Figure 5.39: Combination of HH channels using 2016 data. Expected (dashed) and observed (solid line) 95% CL exclusion limits are shown. The results describe the production cross section of a narrow width spin 0 (top) and spin 2 (bottom) resonance decaying into a pair of SM Higgs bosons.

While no individual channel measurement is sensitive enough to see resonant HH production at the level predicted by the WED theory, the grand combination for the full data set may be able to see the process, if it exists, and approach a sensitivity to cross sections of order picobarns. With more data taken in Run 3 and the HL-LHC period, the BSM theories discussed here will be probed more stringently. However, the ability of a particle collider experiment to detect the SM production of the HH system is further off into the future.

A projected sensitivity for a combined HH measurement based on the full data set expected to be collected with the HL-LHC (almost 3 ab^{-1} of data) shows that this amount of data is not enough to reach the discovery sensitivity [135]: “the statistical combination of the five decay channels results in an expected significance for the standard model HH signal of 2.6σ ”. This is a clear sign that more data are needed for both BSM and SM searches of the HH production. This is a sign that new statistical and MVA tools should be developed and employed. Thus, the next versions of the HH analyses will most likely use a sophisticated neural network not only for signal-background separation but also for physics object reconstruction and more. At the moment, there is a strong effort in the CMS community to develop such tools and methods.

This concludes the discussion of analysis details. In the next chapter, we will summarise the main ideas that have been covered throughout this thesis.

CHAPTER 6

Conclusions

In this dissertation, a search for double Higgs boson production (HH) mediated by a KK graviton or a radion is presented. The double Higgs system subsequently decays through the intermediate $bbZZ$ state - one of the Higgs bosons decays to two b quarks, while the other decays to a pair of Z bosons, which, in turn, decay to a pair of neutrinos and a pair of electrons or muons. The search is performed in data samples corresponding to an integrated luminosity of 35.9 fb^{-1} at $\sqrt{s} = 13 \text{ TeV}$, recorded by the CMS experiment at the LHC in 2016.

This is the world-first search for double Higgs boson production in the $bbZZ$ intermediate channel. A measurement of double Higgs production process is an important test of the standard model (SM) and can open doors to beyond the standard model physics. The HH process, observed in a resonant production mode, would mean new physics, while the HH process observed in a non-resonant mode at the SM predicted level allows one to study Higgs boson self-coupling.

No statistically significant deviations from the SM predictions for background processes were observed in this measurement, and 95% upper-level confidence limits are reported for the production cross section of a KK graviton or a radion multiplied by the branching fraction of its subsequent decay into a di-Higgs system, and further to the $2b2\ell2\nu$ final state. The limits are derived for resonance masses in the 250 GeV

to 1 TeV range.

This measurement was released to the public in November 2018. The measurement is part of an intensive program of searching for a di-Higgs production in CMS in a wide range of channels. Current projections suggest that if no new physics is present, the first observation of this process is not expected in any individual HH channel in the near future. For best sensitivity, CMS plans to combine several HH channels. Present CMS searches in $bbZZ$ channels with different final states will be combined, and then added to a combination of the CMS HH searches in other channels.

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. Goulian, 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*. 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 $\gamma\gamma b\bar{b}$ 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] Kunihiro 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 $b\bar{b}\ell\nu\ell\nu$ 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 Pobladuev, 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 Emelianchik, 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 Fröhwirth, 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 Fröhwirth, 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 $\text{sqrt}(s) = 7 \text{ TeV}$. *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] 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.
- [86] Search for resonant double Higgs production with $b\bar{b}ZZ$ decays in the $b\bar{b}\ell\ell\nu\nu$ final state. Technical Report CMS-PAS-HIG-17-032, CERN, Geneva, 2018.
- [87] Gunter Zech. Upper Limits in Experiments with Background Or Measurement Errors. *Nucl. Instrum. Meth.*, A277:608, 1989.
- [88] 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.
- [89] 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.
- [90] 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
- [91] 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.

- [92] Stefan Kallweit, Jonas M. Lindert, Philipp Maierhofer, Stefano Pozzorini, and Marek Schonherr. NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging. *JHEP*, 04:021, 2016.
- [93] 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)].
- [94] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [95] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [96] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [97] Nikolaos Kidonakis. Differential and total cross sections for top pair and single top production. Technical Report arXiv:1205.3453, Kennesaw State University, May 2012. Comments: 4 pages, 10 figures; presented at XX International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2012), Bonn, Germany, March 26-30, 2012.
- [98] Michal Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$. *Phys. Rev. Lett.*, 110:252004, 2013.

- [99] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.
- [100] Torbjorn Sjostrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.
- [101] Measurements of the $\text{pp} \rightarrow \text{WZ}$ inclusive and differential production cross section and constraints on charged anomalous triple gauge couplings at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-SMP-18-002, CERN, Geneva, 2018.
- [102] Measurement of the WW cross section pp collisions at $\text{sqrt}(s)=13$ TeV. Technical Report CMS-PAS-SMP-16-006, CERN, Geneva, 2016.
- [103] Vardan Khachatryan et al. Measurement of the ZZ production cross section and $\text{Z} \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13$ TeV. *Phys. Lett. B*, 763:280, 2016. [Erratum: *Phys. Lett.B*772,884(2017)].
- [104] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012.
- [105] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205-206:10, 2010.
- [106] SM Higgs production cross sections at $\text{sqrt}(s) = 13$ TeV.
<https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV#ZHllH>
- [107] NNLO+NNLL top-quark-pair cross sections.
https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO#Top_quark_pair_cross_se

- [108] Single Top Cross sections. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopSigma>.
- [109] CMS GEN XSEC Task Force. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/GenXsecTaskForce>
- [110] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [111] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C*, 76(3):155, 2016.
- [112] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [113] CMS MET group. MET Corrections and Uncertainties for Run-II. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETRun2Corrections>.
- [114] CMS MET group. MET Filter Recommendations for Run II. <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2>.
- [115] Albert M Sirunyan et al. Search for resonant and nonresonant Higgs boson pair production in the bblnln final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2017.
- [116] A.M. Sirunyan, R. Kamalieddin, and et al I. Kravchenko. Evidence for the higgs boson decay to a bottom quark-antiquark pair. *Physics Letters B*, 780:501 – 532, 2018.
- [117] 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.

- [118] Geoffrey I. Webb Claude Sammut. *Encyclopedia of Machine Learning*. Springer, 1st edition. edition, 2011.
- [119] 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 Techniques in Physics Research (ACAT)*, page 40, 2007.
- [120] 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.
- [121] CMS Collaboration. CMS Luminosity Measurements for the 2016 Data Taking Period. CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
- [122] Michiel Botje et al. The PDF4LHC Working Group Interim Recommendations. 2011.
- [123] Sergey Alekhin et al. The PDF4LHC Working Group Interim Report. 2011.
- [124] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys.*, G43:023001, 2016.
- [125] Roger Barlow and Christine Beeston. Fitting using finite monte carlo samples. *Computer Physics Communications*, 77(2):219 – 228, 1993.
- [126] 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.

- [127] CMS Higgs WG. Documentation of the RooStats -based statistics tools for Higgs PAG. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit>.
- [128] 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.
- [129] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl.Instrum.Meth.*, A434:435, 1999.
- [130] A. L. Read. Presentation of search results: the CLs technique. *J. Phys. G: Nucl. Part. Phys.*, 28, 2002.
- [131] 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)].
- [132] Eilam Gross. Practical statistics for high energy physics. *CERN Yellow Reports: School Proceedings*, 4(0):165, 2017.
- [133] Andres Tiko. HH resonance searches at ATLAS and CMS. Tokyo, Japan, November 2018. "Higgs Couplings 2018" conference.
- [134] 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.
- [135] Prospects for HH measurements at the HL-LHC. Technical Report CMS-PAS-FTR-18-019, CERN, Geneva, 2018.