Université catholique de Louvain
Secteur des Sciences et Technologies
Institut de Recherche en Mathématique et Physique
Center for Cosmology, Particle Physics and Phenomenology

Search for 2HDM neutral Higgs bosons through $\ell^+\ell^- b\bar{b}$ final states at CMS in run 2 LHC data.

Doctoral dissertation presented by

Khawla Jaffel

in fulfillment of the requirements for the degree of Doctor in Sciences.

Thesis Jury

Dr. Andrea Giammanco (Secretary)	UCLouvain, Belgium
Prof. Barbara Clerbaux	ULB, Belgium
Prof. Christophe Delaere (Supervisor)	UCLouvain, Belgium
Prof. Fabio Maltoni (President)	UCLouvain, Belgium
Dr. Jan Steggemann	EPFL and ETH Zurich, Switzerland

September 8th, 2023

Acknowledgements

I wish to extend my heartfelt gratitude to all those who have always been by my side in this Ph.D. journey, offering encouragement, help, and persistent support.

First and foremost, I want to express my deep appreciation to my supervisor, Christophe Delaere, for not only granting me this position but also for having faith in my abilities to excel in it. His expertise has been a valuable asset, and our discussions have consistently put me in the right direction and guided me towards success. Despite all the struggles that a Ph.D. brings, these past five years have truly been the best years of my life. Thank you, Christophe.
I would also like to extend my thanks to all my jury members; Jan Steggemann, who has been present since the very beginning, providing helpful feedback and suggestions for both the analysis and manuscript. Additionally, my sincere gratitude goes to Barbara Clerbaux for taking the time to thoroughly review the entire manuscript and offering insightful comments that have greatly contributed to shaping this thesis. Special thanks also go to Andrea Giammanco for the fruitful discussion before the private defense and Fabio Maltoni for all his suggestions and useful comments.

A big thanks to my colleagues turned lifelong friends, Sandhya Jain and Suat Donertas. I am deeply thankful for all the moments of laughter and coffee breaks that made this Ph.D. journey wonderful. I owe a debt of gratefulness to Pieter David for his amazing advice, support, and encouragement during the initial two years of my Ph.D. Without his guidance, this journey would not have been possible, and for that, I am profoundly thankful. I'd also like to thank Martin Delcourt and Sébastien Wertz, for being wonderful office mates.

I thank Pietro Vischia for his support and guidance, and for the outstanding statistics courses that have proven to be incredibly priceless, along with the insightful discussions they sparked. Additionally, I deeply thank Florian Bury for his exceptional support in guiding me through the implementation of machine learning. I'm grateful for the daily interactions filled with laughter and joy in the kitchen, thanks to Antoine Depasse, Disrael Camargo, Federico De Lillo, Jishnu Suresh, Julien Touchèque, Jindrich Lidrych, Ishan Darshana, Karlijn Kruiswijk, Marwa Al Moussawi, Maxime Lagrange, Olivier Mattelaer, and all the other happy faces and loud laughter, that I saw and heard every single day.

I extend my thanks to our secretariats Carine Baras and, in particular, Carinne Mertens and our amazing IT team Andres Tanasijczuk, Pavel Demin, and Jérôme

de Favereau for their wonderful support and their ability to find solutions to every problem I encountered. I would also like to extend my gratitude for the funding assistance generously provided by the Excellence of Science (EOS) program.

Finally, my deepest thanks to my family, especially my wonderful twin sister, Khouloud Jaffel, who supported me and encouraged me every single day. As well as my partner, Cristian Joana, for his unwavering belief in me, encouragement in difficult times, and constant reminders of all that I have accomplished, thank you for being part of this amazing journey.

I would like to end here by saying that without each and every one of you, this path would have been vastly different and much more difficult. For that, I am sincerely thankful and grateful to you all.

Contents

Introduction

The Higgs boson, the last piece of the puzzle of the Standard Model (SM) of particle physics was finally discovered at the Large Hadron Collider. The discovery involved decades of theoretical and experimental work, from the initial proposals in the 1960s to the discovery in 2012. The process included the development and construction of advanced particle accelerators (Large Electron–Positron Collider (LEP), Tevatron, and the Large Hadron Collider(LHC)), the gathering and analysis of massive amounts of data, and the collaboration of scientists from around the world.

The Higgs boson was found at a mass of 125 GeV which turned out to be a remarkable choice of nature. According to the SM, its vacuum sits very close to the border of stable and metastable, meaning that the Higgs field is not in its absolute lowest-energy state but instead in a long-lived with a lifetime much longer than the age of the Universe. This has some interesting implications because if the vacuum of the Higgs potential were absolutely stable, it would have significant consequences for the evolution of the Universe. Understanding the metastable nature of the vacuum and the potential consequences of a stable Higgs potential is crucial for gaining insights into the fundamental properties of particles, the stability of the Universe, and the presence of physics Beyond the Standard Model (BSM).

The particle spectrum of the standard model appears to be complete after the discovery of the Higgs bosons. However, issues ranging from the existence of dark matter to the pattern of the neutrino mass, the SM's inability to explain the matter-antimatter asymmetry in the Universe, nor the ability to include the gravitational interaction within the model, continue to suggest physics beyond the SM. Therefore, following the Higgs boson discovery, present measurements at the LHC are focused on testing the Higgs boson's couplings to other elementary particles, precision measurements of the Higgs boson's properties, and initial investigation of the Higgs boson's self-interaction and shape of the Higgs potential.

The goal is to uncover information about these remaining unsolved problems by probing the SM to the furthest extent and making highly accurate measurements. By doing so, we may be able to discover some ways in which the model begins to break down and thereby find a more complete theory. Many scenarios of physics beyond the standard model were proposed to answer these questions

and to better understand the behavior of nature at high energies, in the hope of finding hints of what may lie ahead of the SM.

Precision measurement is not necessarily the only way to search for new physics. Direct searches for exotic particles or rare decays are also possible methods. Since there are no requirements for the Higgs sector to be minimal, many of the proposed theories come with various extensions of the SM; through supersymmetry, such as the Minimal Supersymmetric Standard Model (MSSM) and Next-to-Minimal Supersymmetric Standard Model (NMSSM), and entirely novel framework like extra dimension or grand unified theory which uses a single theoretical structure to describe the strong, weak, and electromagnetic forces. The Two-Higgs-doublet (2HDM) model is one of the most natural choices for an extended scalar sector and is present in most BSM theories. The addition of the second Higgs doublet leads to a richer phenomenology as there are five physical scalar states after Electroweak Symmetry Breaking (EWSB).

This manuscript summarizes my contribution to the physics community in finding hints of physics beyond the SM. The technical and physical analysis contributions described in this thesis were made using data collected with the Compact Muon Solenoid (CMS) detector. The manuscript is divided as follows;

First, it's inevitable to understand the BSM theory before looking into the physics that transcends it. Therefore, Chap. 1 is intended to give insight into the Standard Model, what the theory cannot explain, and why we need to go beyond it. It also opens the way to Chap. 2 in which I explain one of the simplest extensions of the standard model, the 2HDM, posing the mathematical consequence of adding another doublet to the images of the SM. I explain the allowed regions of the 2HDM parameter space at the start of the LHC full run 2 that remain accessible by direct searches in some possible decay channels. Chap. 3 gives a brief description of the LHC focusing mainly on the main features of the CMS detector and its main components. It also summarizes my personal contributions to b-tagging measurement studies and CMS tracking detector performance results in the Cluster Parameter Estimator (CPE) software development. Chap. 4 is a brief overview of the method of data analysis and the statistical tools needed to anticipate a search for new physics. Chap. 5 reports the search for a 2HDM neutral Higgs boson, through $H/A \to ZA/ZH \to \ell^+\ell^- b\bar{b}$ process, describing each step of this analysis from the event selection procedure and the background modeling to the last step of likelihood fit and statistical test.

Finally, the outlook aims to reflect on what can be improved and what has been difficult to achieve, these are my personal insights that could be continued if ever this analysis were to be repeated with future data.

1. Chapter
The Standard Model

The complete Standard Model (SM) took a long time to build. Physicist J.J. Thomson discovered the electron in 1897, and sixty-eight years of discoveries (highlighted in Fig. 1.1) had elapsed before the first prediction of the Higgs boson in 1964 [1,2], which took no more than two years before being integrated into the SM with the Glashow, Weinberg and Salam unification of the electromagnetic and weak forces [3]. At long last, the 2012 discovery of the Higgs boson at the Large Hadron Collider (LHC) [4, 5] was the culmination of almost fifty years of searching for the final piece of the puzzle.

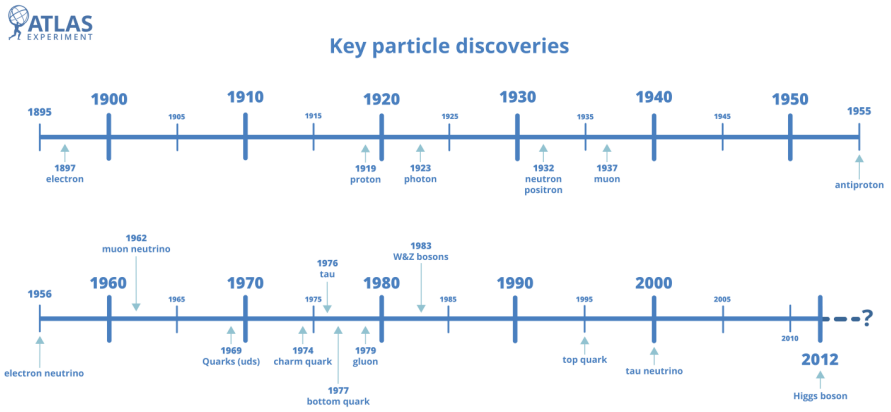


Figure 1.1. | The timeline of important discoveries that led to the Standard Model [6].

1.1. Forces and carrier particles

The Universe exists because the fundamental particles interact. These interactions include attractive and repulsive forces, decay, and annihilation.

The SM explains three of the four fundamental forces that govern the universe: electromagnetism, the strong force, and the weak force. We call the particles that carry the interaction force "carrier particles". Electromagnetism is carried by photons γ and involves the interaction of both electric and magnetic fields. The strong force, which is carried by gluons, g, binds together atomic nuclei to make them stable. The weak force, carried by $W^{\pm}$ and Z bosons, causes nuclear

reactions that have powered our sun and other stars for billions of years. The fourth fundamental force is gravity, which is not adequately explained by the SM. Some physicists expect that the gravitational force may also be associated with a boson particle, often called the "graviton". This hypothetical quanta of gravity is extremely difficult to observe because, at the subatomic level, the gravitational force is several orders of magnitude weaker than the other three elementary forces.

The relative strength of the four types of interaction between two protons in the nucleus is "roughly" listed in Tab. 1.1;

Table 1.1. | Relative strengths for two protons in a nucleus, and their field particles.

Type	Magnitude	Field particle
strong	1	gluons, g
electromagnetic	10^{-2}	photons, γ
weak	10^{-7}	$W^{\pm}$, Z
gravity	10^{-39}	-

$W^{\pm}$, Z, γ and g gauge bosons carry energy and forces throughout the Universe. Fermions obey the Pauli exclusion principle and can in turn be classified into two categories: quarks and leptons. Quarks are sensitive to strong interactions while leptons are not. While the boson's force carriers and the fermions are fundamental in that they cannot be broken down into anything else, there are other composite particles called hadrons that are classified as either fermions or bosons depending on their composition. Hadrons are either mesons which are made up of a quark-antiquark pair or baryons made up of three quark combinations. Mesons are bosons, while baryons are fermions.

The SM is often summarized in a table, similar to the periodic table of elements as illustrated in Fig. 1.2, where it is used to briefly describe particle properties, such as mass, charge, and spin.

- **Electric charge:** The up quark has a charge of $+2/3$ and the down quark has a charge of $-1/3$. The sum of the charges of quarks that make up a nuclear particle determines their electrical charge, but they only form composite particles with integer electric charge. Protons contain two up quarks and one down quark. So that its electric charge is $+1$. Neutrons contain one up quark and two down quarks, which sums to a charge of 0, making neutrons electrically neutral. All particles other than quarks have integer multiples of the electron's charge.
- **Color charge:** Is a property in the theory of Quantum Chromodynamics (QCD), which describes the strong force. It is specific only to gluons and quarks, so all other particles are color-neutral. This color charge is not

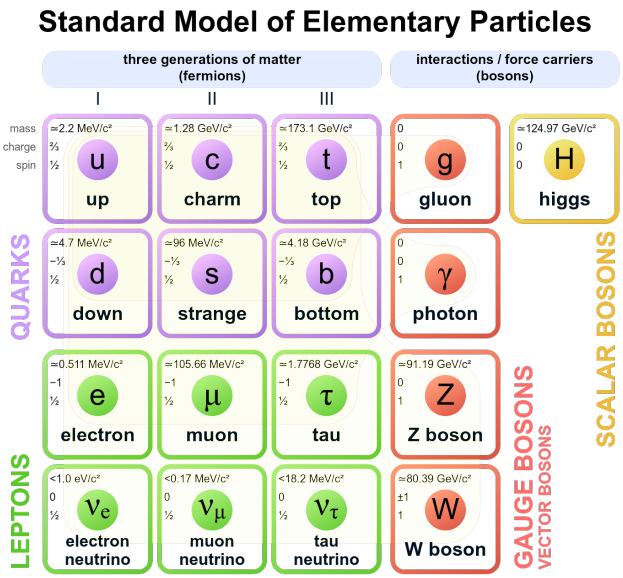


Figure 1.2. The building blocks of matter are six quarks, six leptons, four force carriers, and the Higgs boson which gives mass to all elementary particles, including the W and Z bosons and the fermions (quarks and leptons) [7].

related to the colors we perceive visually but is a concept used to describe the behavior of quarks under the strong force. There are three types (or "flavors") of color charge for quarks, often labeled as red, green, and blue. Antiquarks carry the corresponding "anticolors" – antired, antigreen, and antiblue.

A quark carries one of three color charges, antiquarks carry three types of anticolor and a gluon can be thought of as carrying both color and anticolor. This gives a gluon, eight possible combinations of color, and anticolor [1].

- **Flavor:** Flavor distinguishes quarks (and leptons) from one another.
- **Spin:** Spin is a bizarre but important physical quantity. Large objects like planets or marbles may have angular momentum and a magnetic field because they spin. Since particles also appear to have their own angular momentum and tiny magnetic moments, physicists called this particle property spin. This is a misleading term since particles are not actually "spinning".

[1] The reason there are 9 color-anticolor combinations but only 8 gluons is because in a QCD theory when gluons interact with quarks, they carry both a color and an anticolor charge. Gluons are themselves color-charged. As a result, one of the color-anticolor combinations corresponds to a "colorless" state – a gluon that is a combination of color and an anticolor such that they cancel out, making it effectively color-neutral. This color singlet state is not an active particle and does not participate in strong force interactions.

Spin is quantized to units of 0, 1/2, 1, 3/2 (times Planck's Constant, $\hbar$) and so on. Bosons have an integer value (0, 1, 2 ...) spin quantum numbers while fermions have odd half-integer spin (1/2, 3/2, 5/2...).

The alternative representation to the blocks in Fig. 1.2 is the Lagrangian. Despite the intricacy of appearances, the Lagrangian is one of the easiest, most beautiful, and most compact ways of presenting the theory.

1.2. Foundations of the standard model

1.2.1. Quantum field theory

In particle physics, we study particle interaction at the lowest possible scales where Heisenberg quantum mechanics uncertainties impose higher energies so that most particles involved in a particular scattering event will be moving relativistically. Quantum field theory (QFT) is a theoretical framework based on special relativity and quantum mechanics where it is possible to make predictions about particles scattering cross-sections and decay rates at such short distances. QFT relies on perturbation theory to study such a complicated quantum system thus starting with a mathematical solution to what we know in a simple system and adding an additional disturbance to the system. If those disturbances are weak enough compared to the various physical quantities associated with the perturbed system (e.g. its energy levels and eigenstates) they can be then expressed in terms of "corrections" to those of the simple system. The Hamiltonians for example to which we know the exact solutions of the hydrogen atom, the quantum harmonic oscillator, etc. can be used to describe a wide range of complicated quantum systems using perturbation theory.

The standard model is a quantum field theory meaning that its fundamental particles are quantum fields that are defined at all points in space-time for example an electron is an excitation in the electron field etc. These fields are

- the fermion fields, ψ, which account for "matter particles";
- the electroweak boson fields W^+, W^-, Z, and γ;
- the gluon field, G_a; and
- the Higgs field, ϕ

As a mathematical consequence, they are operator-valued, they act upon a quantum state (ket vector) in Dirac notation.

1.2.2. Renormalisation

In the standard model, when described in quantum field theory, the occurrence of ultraviolet divergence becomes inevitable. Those that arise from the higher-order perturbative calculation that involves summing over all possible intermediate states. In the language of Feynman diagrams, this is represented by a loop and the Feynam rules require integration over all possible momenta of particles

inside the loop which often results in divergence. In modern particle physics, we don't actually interpret this divergence as a genuine infinity but we assume that some new physics enters at some sufficiently high Λ cutoff. Higher-order perturbative corrections are in most cases increased logarithmically. To solve this issue a rescaling is needed to absorb this cutoff dependence. Quantum field theories for which this is possible are called "renormalizable theories" and the criteria for a theory to be renormalizable are:

- The interaction terms must not have a dimension greater than four. Recalling that the dimensions of bosonic fields (both scalar and vector) are 1 whereas the dimension of spin=1/2 fermionic fields is 3/2.
- All propagators must vanish as the momentum of the propagation particle becomes infinite.

Until the 1970's the only renormalizable quantum field theory was quantum electrodynamics (QED) -that is the relativistic quantum field theory of electrodynamics. The QED Lagrangian can be written as:

$$\mathcal{L}_{\mathrm{QED}} = e\overline{\psi}\gamma^{\mu}\psi A_{\mu}, \tag{1.1.}$$

ψ represents the electron and positron fields (collectively referred to as fermions), $\overline{\psi}$ is the Dirac adjoint of the fermion fields, γ^{μ} are the Dirac gamma matrices, A_{μ} is the electromagnetic vector potential (a component of the photon field), and e is the elementary electric charge. The QED Lagrangian density $\mathcal{L}_{\mathrm{QED}}$, has dimension four as required, and for which photon and fermion propagators both vanish as the momentum p tends to infinity. Also, the coupling (i.e. the charge) is small enough that a perturbative expansion in the fine structure can be handled. The interaction term between a charged fermion and a photon is

$$\alpha = \frac{e^2}{4\pi} = \frac{1}{137.06...} \tag{1.2.}$$

1.2.3. Gauge theories

A gauge theory is a theory that is invariant under a set of local transformations. Quantum electrodynamics is an Abelian gauge theory with the symmetry group $U(1)$ and has one gauge field, the electromagnetic four-potential, with the photon being the gauge boson. Gauge theories require vector bosons to be massless known also as gauge bosons, this applies to electromagnetic interactions but not to the weak interactions that involve the exchange of a massive particle. If we still wish to extend the ideas of describing interactions in a gauge theory the symmetry somehow must be broken. Introducing a mass term to the gauge bosons would render the theory non-renormalizable. Another elegant way of doing this is known as "spontaneous symmetry breaking" in which the Lagrangian maintains its symmetry under a set of local gauge transformations, whereas

the ground state (vacuum) is not. When a spontaneous symmetry breaking is applied in a gauge theory a mass manifests itself for the $W^{\pm}$ and Z^0 gauge bosons.

1.2.3.1. Spontaneous symmetry breaking

The Lagrangian of a spontaneously broken symmetry of a scalar field theory with a mass term and quartic self-interaction can be written as.

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - V(\phi), \tag{1.3.}$$

where the potential $V(\phi)$ is given by

$$V(\phi) = \mu^2 \phi^* \phi + \lambda |\phi^* \phi|^2, \tag{1.4.}$$

and the Lagrangian is maintained invariant under global $U(1)$ transformations.

$$\phi \rightarrow e^{i\omega} \phi. \tag{1.5.}$$

If μ^2 is positive, this potential has a minimum at $\phi = 0$ (vacuum), whereas other values of ϕ represent the creation and annihilation operators that populate higher energy states. Suppose now that we reverse the sign of μ^2 so that the potential becomes

$$V(\phi) = -\mu^2 \phi^* \phi + \lambda |\phi^* \phi|^2 . \tag{1.6.}$$

The potential has no longer a minimum at $\phi = 0$ but a maximum and the minimum occurs at

$$\phi = \frac{v}{\sqrt{2}} = e^{i\theta} \sqrt{\frac{\mu^2}{2\lambda}}, \tag{1.7.}$$

with an infinite number of vacuum states where θ can takes any value from 0 to 2π. The symmetry breaking occurs in the choice made for the value θ which represents the true vacuum. If we choose $\theta = 0$ to be this vacuum, $U(1)$ invariance is broken and in quantum field theory, we say that the field ϕ has a non-zero vacuum expectation value

$$\langle \phi \rangle = \frac{v}{\sqrt{2}}. \tag{1.8.}$$

This also means that there are "excitations" with zero energy that can take us from this vacuum to one of the other states that are degenerate in energy. Only massless particles can have zero energy and to prove so we expand ϕ around its vacuum expectation value as

$$\phi = \frac{1}{\sqrt{2}}\left(\frac{\mu}{\sqrt{\lambda}} + H + i\phi\right). \tag{1.9.}$$

The fields H and ϕ have zero vacuum expectation values which are expanded in terms of creations and annihilation operators of the particles that populate excited states. If we insert Eq. 1.9 into Eq .1.6 we find

$$V = \mu^2 H^2 + \mu\sqrt{\lambda}(H^3 + \phi^2 H) + \frac{\lambda}{4}(H^4 + \phi^4 2H^2\pi^2) + \frac{\mu^4}{4\lambda}. \tag{1.10.}$$

Now in Eq. 1.10 there is a mass term for the field H, but no mass term for the field ϕ. Thus ϕ is a field for a massless particle called a "Goldstone boson". Fig. 1.3 is an illustration of the Sombrero potential $V(\phi)$ with a spontaneous symmetry breaking or what we often call the "Mexican hat" in polar coordinates.

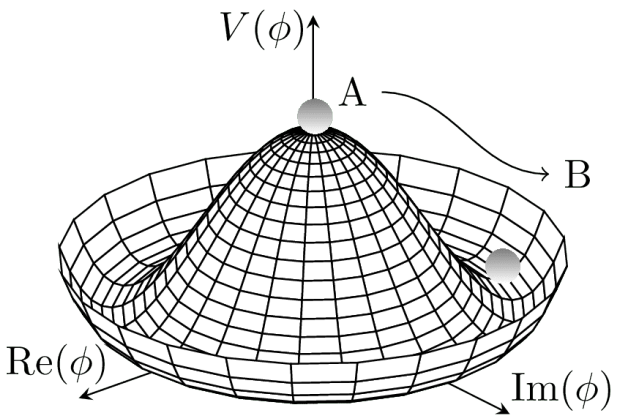


Figure 1.3. | A potential with spontaneous symmetry breaking [8].

1.2.3.2. The Brout-Englert-Higgs mechanism

In Goldstone's theorem, the choice of vacuum is equivalent to choosing a gauge which is necessary to be able to quantize the theory.

A massless vector boson has only two degrees of freedom (the two directions of polarisation of a photon), whereas a massive vector (spin 1) particle has three possible values for the helicity of the particle. In a spontaneous symmetry-breaking gauge theory, the Goldstone Gauge provides the third degree of freedom to the gauge bosons. This means that the gauge bosons become massive. To see how this works we return to $U(1)$ gauge theory but now we promote the symmetry to a local symmetry and we must introduce a gauge boson A_μ. The partial derivative of the field ϕ is replaced by a covariant derivative.

$$\partial_\mu \phi \to D_\mu \phi = \left(\partial_\mu + ieA_\mu\right)\phi. \tag{1.11.}$$

If we include the kinetic term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ for the gauge bosons, the Lagrangian density becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + D_\mu\phi^* D^\mu\phi - V(\phi). \tag{1.12.}$$

Now see what happens if we insert Eq. 1.9 into the term $D_\mu\phi^* D^\mu\phi$

$$\begin{aligned}&\frac{1}{2}\partial_\mu H\partial^\mu H + \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{1}{2}e^2\nu^2 A_\mu A^\mu + \\ &\quad e\nu A^\mu\partial_\mu\phi - eA^\mu(\phi\partial_\mu H - H\partial_\mu\phi) + \frac{1}{2}e^2 A_\mu A^\mu(H^2+\phi^2).\end{aligned} \tag{1.13.}$$

The gauge boson has acquired a mass term $M_A = e\nu$, where $\nu = \mu/\sqrt{\lambda}$ despite the Lagrangian is invariant under local $U(1)$ transformations.

The term $e\nu A^\mu\partial_\mu\phi$ in Eq. 1.13 which can be written as $-M_A\phi\partial_\mu A^\mu$ indicate that the Goldstone boson, ϕ, couples to the longitudinal component of the gauge boson with strength M_A.

So if we separate the gauge-boson field, A_μ, into its transverse and longitudinal components

$$A_\mu = A_\mu^L + A_\mu^T, \tag{1.14.}$$

where $\partial^\mu A_\mu^T = 0$. The longitudinal part of the gauge bosons can be thought of as an oscillation between the Goldstone boson with a mixing term given by $-M_A\phi\partial^\mu A_\mu^L$ so that the physical particle is described by a superposition of these fields. Let's consider two special cases.

1. The unitary gauge:

The first case is when the physical field for the longitudinal component is not simply A_μ^L but is actually a superposition that can be written as

$$A_\mu^{ph} = A_\mu + \frac{1}{M_A}\partial_\mu\phi. \tag{1.15.}$$

When the Lagrangian is written in terms of the physical gauge boson, the Goldstone field ϕ drops out. The equation of motion with the new redefined physical gauge boson field in Eq. 1.15 becomes

$$\left[-g^{\mu\nu}\left(\Box + M_A^2\right) + \partial^\mu\partial^\nu\right]A_\nu^{ph} = 0, \tag{1.16.}$$

leading to a propagator of a massive spin-one particle

$$-i\left(g_{\mu\nu}-\frac{p_\mu p_\nu}{M_A^2}\right)\frac{1}{(p^2+M_A^2)}. \tag{1.17.}$$

Now the only other remaining particle is the scalar, H, with mass $m_H=\sqrt{2}\mu$ that interacts with the gauge boson and also has cubic and quartic self-interactions. It is the Higgs boson particle. The interaction terms involving the Higgs boson are

$$\mathcal{L}_I(H)=\frac{e^2}{2}A_\mu A^\mu H^2+eM_A A_\mu A^\mu H-\frac{\lambda}{4}H^4-m_H\sqrt{2\lambda}H^3. \tag{1.18.}$$

In Feynman's rules, these terms lead to the vertices shown in Fig. 1.4.

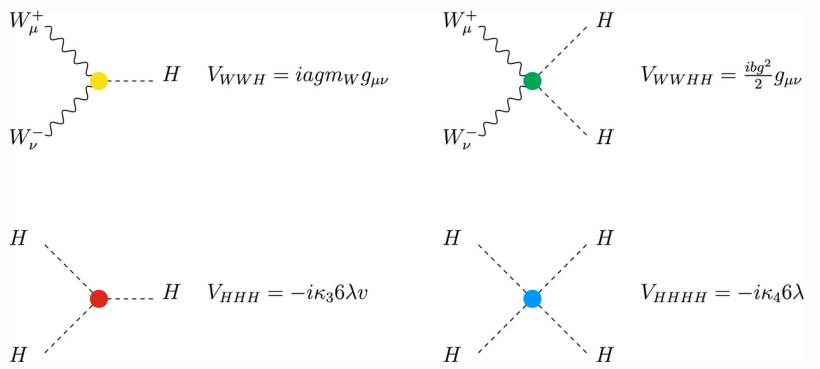


Figure 1.4. | Higgs self-interactions vertices in Feynam rules [9].

2. Feynman gauge:

The second case consists of; instead of fixing the gauge by constraining the gauge field a priori, one can add a gauge-breaking term to the "physical" (gauge invariant) Lagrangian:

$$-\frac{1}{2}(\partial.A+M_A\phi)^2. \tag{1.19.}$$

The equation of motion for the free gauge boson becomes

$$-g^{\mu\nu}(\Box+M_A^2)A_\mu=0, \tag{1.20.}$$

so that the gauge boson propagator simplifies to

$$-i\frac{g_{\mu\nu}}{(p^2+M_A^2)}. \tag{1.21.}$$

For such a choice of gauge fixing. The Goldstone boson is still present and has acquired a mass, M_A, it interacts with the gauge boson, with the Higgs scalar, and itself.

The question now is if the problem of renormalization has been solved. Recall that if we want the theory to be renormalizable, all UV divergences need to be absorbed into the field's masses and coupling. Thus all propagators have to decrease like $1/p^2$ as the momentum tends to infinity.

While the generation of a mass for the gauge bosons by spontaneous symmetry breaking solved the renormalization for massive spin-1 particles. Looking at the propagator in Eq. 1.17 for the gauge boson in unitary gauge we see that this criterion is not valid. Nevertheless, physical quantities, such as energy, mass, etc., do not depend on the gauge, i.e., they are gauge invariant. Therefore there is large freedom in gauge choice to do our calculations. If we look at the propagator in Eq. 1.21 such choice renders the theory renormalizable even if it does not appear in the unitary gauge. It is only when the gauge boson acquires masses through the Brout-Englert-Higgs mechanism that we do have a renormalizable quantum field theory.

The difficulty the gauge fixing brings is that to preserve unitarity ghost fields manifest themselves as they interact not only with the gauge boson but with the Higgs scalar and Goldston boson. Thus resulting in more particles and many more interactions. The ghost fields do not correspond to any real particles in external states: they appear as virtual particles in Feynman diagrams. The Faddeev–Popov ghosts violate the spin–statistics relation, which is another reason why they are often regarded as "non-physical" particles.

1.2.3.3. Non-Abelian gauge theory

Recall that in a quantized gauge theory, gauge bosons are quanta of the gauge fields. Consequently, there are as many gauge bosons as there are generators of the gauge field. In quantum electrodynamics, the gauge group is $U(1)$ and there is only one gauge boson, the photon. In quantum chromodynamics, the gauge group is $SU(3)$ which has eight generators, corresponding to the eight gluons. The electroweak $W^{\pm}$ and Z bosons are described in $SU(2)$ gauge theory.

Let's see how can we extend the gauge fixing in a $SU(2)$ gauge theory. If we consider a complex doublet of a scalar field, ϕ^i, $i = 1, 2$. The Lagrangian density is written as

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + |D_\mu\phi|^2 - V(\phi), \tag{1.22.}$$

where

$$D_\mu\phi = \partial_\mu\phi + igW^a_\mu T^a\phi, \tag{1.23.}$$

and W_μ^a is A_μ^a the gauge boson noted earlier in a different notation, and the potential

$$V(\phi) = -\mu^2 \phi_i^\dagger \phi^i + \lambda(\phi_i^\dagger \phi^i)^2. \tag{1.24.}$$

Now this potential has a minimum at $\phi_i^\dagger = -\frac{1}{2}\mu^2/\lambda$. We can choose the vacuum expectation value to be real in the $T^3 = -\frac{1}{2}$

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{1.25.}$$

where $v = \mu/\sqrt{\lambda}$ in which the vacuum is also invariant under any SU(2) transformations. So that there is no unbroken subgroup and we expect all three gauge bosons and Goldstone bosons to acquire a mass.

If we expand ϕ^i about its vacuum expectation value, often also noted as "*vev*",

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 - i\phi_2\\ v + H + i\phi_0\end{pmatrix}, \tag{1.26.}$$

where ϕ_a, $a = 0...2$ are the three Goldstone bosons and H is the Higgs scalar. All of these fields have zero *vev*. With this expansion of ϕ^i, the potential in Eq. 1.24 returns a mass term for the Higgs field, with value $m_H = \sqrt{2}\mu$.

In the unitary gauge, the covoriant derivative $D_\mu\phi$ can be written as

$$D_\mu\phi = \frac{1}{\sqrt{2}}\left[\partial_\mu\begin{pmatrix}0\\ H\end{pmatrix} + i\frac{g}{2}\begin{pmatrix}W_\mu^0 & \sqrt{2}W_\mu^-\\ \sqrt{2}W_\mu^+ & -W_\mu^0\end{pmatrix}\begin{pmatrix}0\\ v+H\end{pmatrix}\right]. \tag{1.27.}$$

We have introduced $W_\mu^+ = (W_\mu^1 + iW_\mu^2)/\sqrt{2}$ and the explicit form for the generators of $SU(2)$ in the 2×2 matrices. The term $|D_\mu\phi|^2$ then becomes

$$\begin{aligned}|D_\mu\phi|^2 =& \frac{1}{2}\partial_\mu H\partial^\mu H + \frac{1}{4}g^2v^2(W_\mu^+W^{-\mu} + \frac{1}{2}W_\mu^0W^{0\mu}) +\\ & \frac{1}{4}g^2H^2(W_\mu^+W^{-\mu} + \frac{1}{2}W_\mu^0W^{0\mu}) + \frac{1}{2}g^2vH(W_\mu^+W^{-\mu} + \frac{1}{2}W_\mu^0W^{0\mu}).\end{aligned}$$

It can be seen that all the quadratic terms and the three gauge bosons have acquired mass

$$M_W = \frac{gv}{2}. \tag{1.28.}$$

The SM is also a non-abelian gauge theory with the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$ and has a total of twelve gauge bosons; the photon, three weak bosons, and nine gluons. C refers to color charge, L to left-handed fermions and Y to the weak-hypercharge. Each term of the Lagrangian is also invariant under this group of symmetry. Electromagnetic, weak, and strong

interactions are all related to local symmetries and described by Abelian and non-Abelian gauge theories.

The electroweak sector of the Standard Model (SM) consists of the gauge symmetry group $SU(2)L \times U(1)Y$, which combines the weak isospin ($SU(2)L$) and weak hypercharge ($U(1)Y$) symmetries. The $SU(2)_L$ gauge symmetry corresponds to the weak nuclear force and has three gauge bosons associated with it. These are the W^+, W^-, and Z bosons. The $U(1)_Y$ gauge symmetry corresponds to the hypercharge and also has a gauge boson associated with it. This gauge boson is the photon (γ). Therefore, the electroweak sector of the SM has a total of four gauge bosons: W^+, W^-, Z, and γ. In addition to the electroweak sector, the strong nuclear force is described by the $SU(3)_C$ gauge symmetry, corresponding to the color charge of quarks and gluons in quantum chromodynamics (QCD).

The masses of all particles are generated by two mechanisms: confinement and spontaneous symmetry breaking. Spelling out the details reveals a rich theory that accounts for strong and electroweak interactions, confinement and spontaneous symmetry breaking, hadronic and leptonic flavor physics, etc. The SM Lagrangian can be written as:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\overline{\psi} + \gamma\mu D_\mu\psi + D_\mu\phi^+ D^\mu\phi - V(\phi) + \overline{\psi}_L\widetilde{Y}\phi\psi_R + h.c, \quad (1.29.)$$

where $\mu\nu$ indices represent the four space-time dimensions.

The term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ of Eq. 1.29 is the Gauge Lagrangian ($\mathcal{L}_{\text{Gauge}}$) which describes how bosons (strong, electromagnetism and weak force carriers) interact with each other.

$$\mathcal{L}_{\text{Gauge}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu}, \quad (1.30.)$$

The term $i\overline{\psi}\gamma^\mu D_\mu\psi$ of Eq. 1.29 includes how fermions interact with forces described in detail in the $\mathcal{L}_{\text{Fermion}}$.

$$\mathcal{L}_{\text{Fermion}} = i\overline{\psi}\gamma^\mu D_\mu\psi = \sum_{quarks} i\overline{q}\gamma^\mu D_\mu q + \sum_{\psi_L} i\overline{\psi}_L\gamma^\mu D_\mu\psi_L + \sum_{\psi_R} i\overline{\psi}_R\gamma^\mu D_\mu\psi_R, \quad (1.31.)$$

where ϕ is the Higgs field.

$$\mathcal{L}_{\text{Higgs}} = D_\mu\phi^+ D^\mu\phi - V(\phi), \quad (1.32.)$$

The first term of $\mathcal{L}_{\text{Higgs}}$ represents how the force particles interact with the Higgs field while the second part gives us the Higgs self-interactions.

$$\mathcal{L}_{\text{Yukawa}} = \overline{\psi}_L\widetilde{Y}\phi\psi_R = -\overline{L}Y_l\phi l_R - \overline{Q}_{L'}Y_d\phi d_{R'} - Q'_{L'}Y_u\widetilde{\phi}u'_R, \quad (1.33.)$$

where $L(R)$ represents Left(Right)-handed particles and Y is the Yukawa coupling of each particle to the Higgs field. Remember that the stronger the Yukawa coupling is, the heavier the particle interaction with the Higgs field will be.
The first part of $\mathcal{L}_{Yukawa}$ represents the interactions of the leptons to the Higgs field while the second and third terms represent respectively up and down type quarks interaction to the Higgs field.
Finally, $h.c$, with h the Planck's constant, and c the speed of light in vacuum. It represents the Hermitian conjugate (i.e. the same interactions described above but with anti-matter) and is used as a normalization factor to ensure that the Lagrangian has the correct units of energy and can correctly describe the interactions and dynamics of particles in the quantum field theory framework. The sum of all these terms defines the Standard Model Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{Gauge}} + \mathcal{L}_{\text{Fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} + h.c \tag{1.34.}$$

1.3. Testing the Standard Model

The SM has proven incredibly successful at describing a wide variety of phenomena that we observe in our experiments. The most famous example is the agreement of the Standard Model (SM) theory prediction and the experimental measurements of the electron magnetic dipole moment to within about 1 part per 100 billion [10]. Experiments have verified the Standard Model predictions to incredible precision, and all the particles predicted by this theory have been found. To date, all measurements agree reasonably well with SM across nine orders of magnitude as shown in the summary plot of the standard model cross-section predictions and corresponding measurements by the CMS collaboration [11] in Fig. 1.5.

One of the most critical tests to the SM comes down to finding the Higgs bosons, the challenge that a search had to be envisaged not only over a large range of masses but also many possible decay modes: into pairs of photons γ, Z bosons, W bosons, τ leptons, and b quarks. The predicted cross sections and the branching ratios into the various decay modes of the SM Higgs boson as a function of mass are illustrated in Fig. 1.6.

The search for the unknown Higgs mass m_H was one of the objectives of the Large Electron–Positron (LEP) experiment, which established the lower limit:

$$m_H > 114.4 \quad \text{GeV} \tag{1.35.}$$

The consistency of electroweak precision measurements with the predictions of the SM pointed to the presence of a Higgs boson with a mass around the electroweak scale, which was in the range of a few hundred GeV, hinting that the discovery of the Higgs boson might be accessible to experimental searches with the higher energies that the Large Hadron Collider (LHC) would provide. The

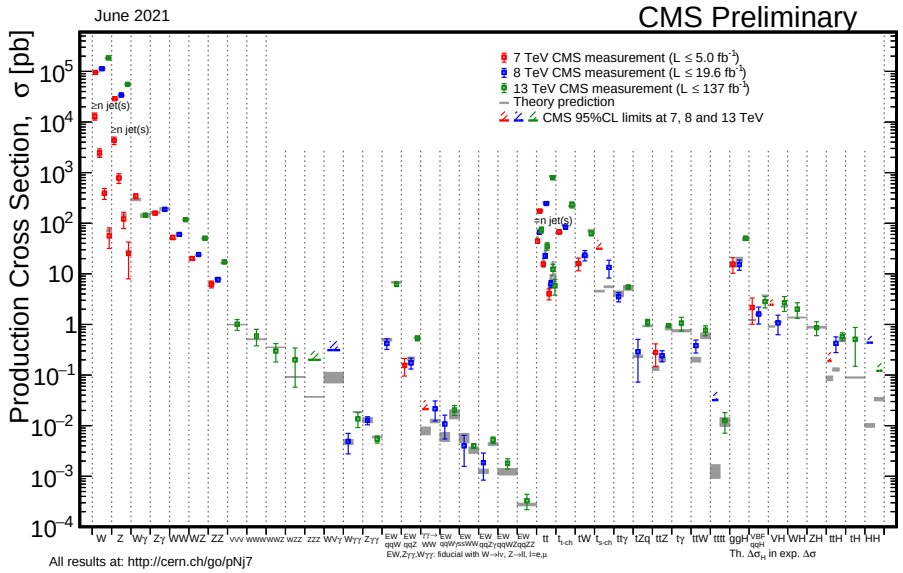


Figure 1.5. | Summary of the cross-section measurements of Standard Model processes [12].

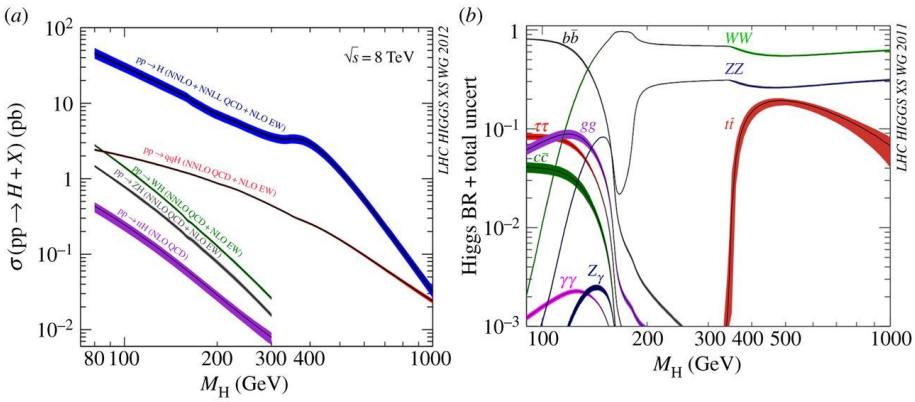


Figure 1.6. | Standard Model Higgs boson decay branching ratios and cross-sections [13].

ATLAS and the CMS experiments embark on the search after LEP experimental runs and indeed *"Lo and behold -there it was"* - detected within just two years of beginning LHC operations. The result from the CMS experiment is shown in Fig 1.7. A clear peak at the diphoton invariant mass of approximately 125 GeV in mass is seen.

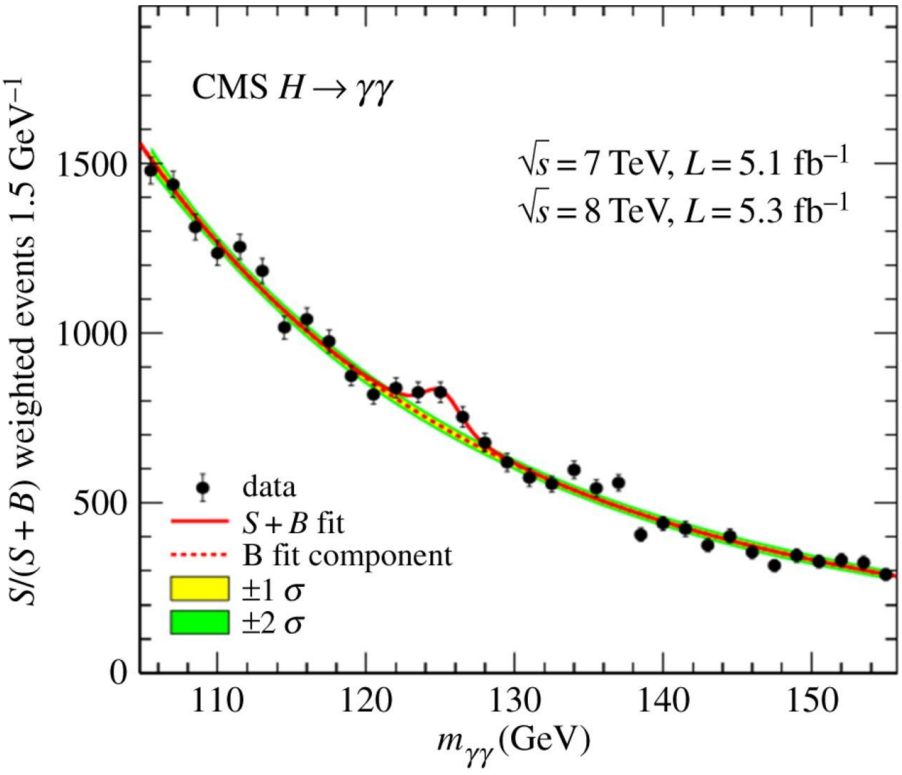


Figure 1.7. | The diphoton invariant mass distribution. Each event is weighted by the $S/(S+B)$ value. The lines represent the fitted background and the expected signal contribution ($m_H = 125$ GeV), and the colored bands represent the ± 1 and ± 2 standard deviation uncertainties in the background estimate [14]. A similar result was obtained by the ATLAS experiment [15].

The expected significance of the SM Higgs boson signal, when the five decay modes (γ, ZZ, WW, $b\bar{b}$, $\tau\tau$) were combined, was 5.6 σ for an integrated luminosity of about 5 to 6 fb^{-1} for each experiment. $\gamma\gamma$ and ZZ achieve the highest sensitivity because of the excellent mass resolution (1-2 GeV). Combining all data taken at $\sqrt{s} = 7$, 8, and 13 TeV, the observed (expected) significance is found to be 6.4 (5.4) standard deviations.

The $W^{\pm}$ and Z^0 bosons are a key building block of the Standard Model, precision measurements of their masses are a perfect test of the SM. The importance of these precision calculations and measurements chock the physic community with the latest 2022 CDF II $W^{\pm}$ mass results come up revealing a mass of 80.433 ± 0.094 GeV [16]. This does not match previous measurements, including those made by CDF II in 2012 (80.387 ± 0.02 GeV) and by ATLAS at CERN in

2023 (80.370 ± 0.019 GeV). The deviation from the standard model prediction of the latest measurements is illustrated in Fig. 1.8.

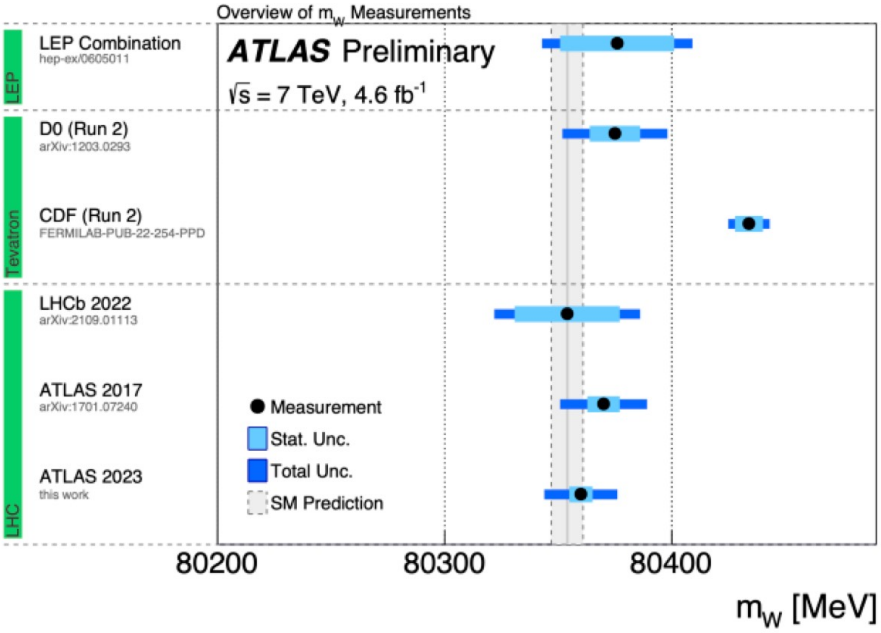


Figure 1.8. | Comparison of the measured value of the W boson mass from different experiments. The vertical bands show the Standard Model prediction, and the horizontal bands and lines show the statistical and total uncertainties of the results [17].

If this ever turns out to be true, this higher mass would be our gateway to new physics but before this extraordinary claim, extraordinary evidence is still yet to be provided.

1.4. Beautiful but flawed

Despite the great predictive power of the SM, it does not answer some crucial questions, which is why, as physicists, we know our job is far from done. It contains at least 19 arbitrary parameters and does not explain the particle quantum numbers, such as the electric charge Q, weak isospin I, hypercharge Y, and color. The need for new physics is not recent. A few examples are presented here, but this selection is by no means exhaustive.

1.4.1. The anomalous magnetic dipole moment

The magnetic moment g also called the magnetic dipole moment, is a measure of the strength of a magnetic source. At the tree level, Dirac equations predict

$g = 2$, the observation shows that this value differs from the observation. For charged leptons ($l = \mu, e, \tau$) the difference is the anomalous magnetic moment, $a_l = (g_l - 2)/2$, in which the precision measurements of this quantity provide a stringent test of the SM and strong evidence for new physics if were to exist.

This anomaly in quantum electrodynamics (QED) is the result of a contribution of quantum effects that can be expressed by loops diagrams in Feynman notations, where it might be a BSM virtual particles or forces providing additional contributions to the Dirac prediction at the tree-level.

The most recent results of this quantity were published in April 2021 [18] by the Muon g-2 experiment where $a_\mu = 0.00116592040(54)$ which exceeds the Standard Model prediction by 4.2 standard deviations. The measurement is precise to 140 parts per billion.

1.4.2. Baryon asymmetry

What's the origin of the matter in the Universe? One of the greatest challenges in physics is to figure out what happened to the antimatter, or why we see an asymmetry between matter and antimatter. According to the SM, the Big Bang should have created equal amounts of matter and antimatter in the early Universe. However, it appears that an unknown mechanism may have interfered causing this symmetry to break. This became an issue over 90 years ago when Paul Dirac realized that his relativistic version of the Schrödinger wave equation for electrons predicted the existence of anti-electrons (i.e. positrons).

In 1967, Andrei Sakharov proposed three conditions for the successful generation of matter-antimatter asymmetry [19]: I) baryon number violation, II) C and CP violation and, III) a deviation from thermal equilibrium. Baryon number violations are predicted to exist in unified theories but have never been observed. C and CP violation as described within the SM is insufficient to generate the amount of matter seen in the Universe today. Finally, interactions are out of thermal equilibrium, meaning that the rate of a reaction that generates baryon-asymmetry must be less than the rate of expansion of the Universe. In this situation, the particles and their corresponding antiparticles do not achieve thermal equilibrium due to rapid expansion decreasing the occurrence of pair annihilation.

1.4.3. Dark energy and dark matter

Observations of the structure of galaxies, the rotation of stars and neutral hydrogen gas in spiral galaxies, the motions of clusters of galaxies, and so on... can not be described in terms of Newton's universal law of gravitation and also the visible ordinary matter within the galactic systems.

An additional gravitational pull provided by a non-relativistic form of invisible

matter, which doesn't bind to ordinary matter, has no electric charge or strong interactions was necessary to account for the rapid movements of the galaxies within the cluster and also to hold the cluster together.

"Dark matter" emerged to clarify the formation of structures within the Universe and their persistence today. Despite its abundance (the Universe's energy budget; 27% dark matter, 68% dark energy, and 5% ordinary matter), until today we are only able to infer the existence of dark matter from the gravitational effect it seems to have on visible matter. We are far more certain to know what dark matter is not than we are what it is. It is predicted to probably be less than about 1 TeV in mass [20], which does not interact with the electromagnetic field, meaning it does not absorb, reflect, or emit electromagnetic radiation and therefore makes it difficult to detect.

This is a cosmological puzzle that needs new physics beyond the Standard Model (BSM) and there are few dark matter possibilities that are viable. One common view is that dark matter is not baryonic at all, but that it is made up of other, more exotic particles like axions or WIMPS (Weakly Interacting Massive Particles) and many other possible scenarios that are supported in BSM physics.

Dark energy on the other hand is a mysterious form of energy that permeates space and is responsible for the accelerated expansion of the Universe. It was first inferred from observations of distant supernovae in the late 1990s, and since then, various cosmological observations have provided strong evidence for its existence.

1.4.4. Neutrinos

Neutrino is one of the building blocks of matter that is most difficult to detect. It is a lepton fermion (i.e. it does not participate in strong interactions) and has zero charges (i.e. it undergoes no electromagnetic interaction) but it interacts via the weak force. There are three types, or "flavors", of neutrinos that are created as a result of beta decay of atomic nuclei or hadrons, nuclear reactions such as those in the sun or nuclear reactors, during supernova, etc. Weak interactions create neutrinos in one of three leptonic flavors: electron neutrinos (ν_e), muon neutrinos (ν_μ), and tau neutrinos (ν_τ); each type also has an antimatter partner, an antineutrino which also has no electric charge and half-integer spin.

In the SM neutrinos were assumed to have zero mass based on the concept of "chirality" or "handedness". However, the discoveries at Super-Kamiokande, Sudbury Neutrino Observatory (SNO), and KamLAND, along with many others have collectively demonstrated that neutrinos undergo flavor oscillations, implying that neutrinos have mass and that the three different flavors (electron, muon, and tau neutrinos) are actually mixtures of three different mass eigenstates. This is because neutrinos of different masses propagate through space at

different speeds due to their differing momenta. As neutrinos travel, the different mass eigenstates (neutrinos with definite masses) composing a neutrino of a particular flavor evolve, leading to a change in the flavor composition observed at a distant location. This discovery opened a door to new physics. The question is -Is that all the SM missed, or is there more we don't know about neutrinos?

The origin of neutrino mass remains a mystery because, in the SM, the mass of an elementary fermion arises from the coupling of Yukawa to the non-zero vacuum expectation value of the Higgs field. Neutrinos are left-handed, so they are unable to interact with the Higgs field and therefore acquire no mass contrary to what observation tells us, if this were required to happen a left-handed neutrino must turn right-handed which does not exist. Ettore Majorana in 1937 proposed that neutrinos are their own antineutrino (also called Majorana fermions), which contrasts with a Dirac fermion which states that fermions are not their own antiparticles. If true, it would explain where neutrinos get their mass from. An observation of neutrinoless double beta decay, which has yet to be seen, would have profound implications for our understanding of neutrinos and particle physics. The decay would involve two neutrons in a nucleus simultaneously transform into two protons without emitting any neutrinos in the final state. This process violates lepton number conservation and would provide evidence for neutrinos being Majorana particles, which are their own antiparticles.

Chapter 2. The extended scalar sector in physics beyond the standard model

Until today, nature has provided us only hints of the existence of hidden physics that can't be described within the standard model (SM) theory, but to make such an extraordinary claim that physics beyond the standard model does exist, extraordinary evidence must be given. Except for the neutrino mass problem, there is no experimental result accepted so far that contradicts the SM at the 5 σ level (considered the threshold of discovery in high-energy physics).

Now that we know that neutrinos have mass, we still have to understand why their mass is so small compared with other particle masses, and for that, we need a new theory Beyond the Standard Model (BSM) to make neutrinos massive. Many BSM physics scenarios have been proposed to explain not only the neutrino mass but all the phenomena that the SM has missed or could not be included in the theory, a few examples have been discussed in Chap. 1. While there are various BSM proposals and theories that don't necessarily involve direct extensions of the SM, such as certain theories related to extra dimensions, and alternative theories of gravity. Many of the emerging BSM scenarios include extensions of the SM via supersymmetry, such as the Minimal Supersymmetric Standard Model (MSSM) and Next-to-Minimal Supersymmetric Standard Model (NMSSM). The goal is to reproduce what has been found and include what has not been found, with the ultimate goal of finding a Theory of Everything (TOF) that fully explains and connects all aspects of the Universe.

While the 2HDM and MSSM introduce additional particles and interactions beyond the Standard Model, they do not directly solve the neutrino problem on their own, they can be part of larger frameworks or extensions that incorporate mechanisms to explain neutrino masses and mixings. For example, the MSSM can be extended to include right-handed neutrinos and a seesaw mechanism, where the smallness of neutrino masses is explained by heavy Majorana neutrinos. This extension, called the "supersymmetric seesaw mechanism", which can account for tiny neutrino masses while maintaining the broader framework of the MSSM.

This chapter will focus on the two-Higgs-doublet model (2HDM) which is considered one of the natural choices for physics beyond the standard model.

2.1. The Two Higgs-doublet model

The two-Higgs-Doublet Model is the simplest extension of the SM which contains two complex doublets of scalar $SU(2)_L$ fields, ϕ_1 and ϕ_2:

$$\phi_i = \begin{pmatrix} \phi_i^+ \\ \phi_i^0 \end{pmatrix}, \tag{2.1.}$$

with $i = 1,2$. The spontaneous electroweak symmetry breaking via the Higgs mechanism is described by the most general $SU(2)_L \times U(1)_Y$ invariant lagrangian for the 2HDM that can be written as

$$\mathcal{L}_{\text{2HDM}} = \mathcal{L}_\phi + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{SM}}, \tag{2.2.}$$

$\mathcal{L}_{SM}$ describes the $SU(2)_L \times U(1)_Y$ standard model interactions of gauge bosons and fermions, $\mathcal{L}_{Yukawa}$ describes the Yukawa interactions of fermions with Higgs scalars as discussed in Chap.1, while the Higgs scalar Lagrangian $\mathcal{L}_\phi$ is given by

$$\mathcal{L}_\phi = \sum_{i=1,2} (D_\mu \phi_i)^+ (D^\mu \phi_i) - V(\phi_1, \phi_2), \tag{2.3.}$$

$V(\phi_1,\phi_2)$ is the 2HDM scalar potential, D_μ is the exact covariant derivative that appears in the Higgs potential of the standard model lagrangian, which can be written as

$$D_\mu = \partial_\mu - igW_\mu^a T_a - ig' \frac{Y}{2} B_\mu, \tag{2.4.}$$

where T_a and Y are the generators of the weak-isospin and weak-hypercharge transformations.

The most general gauge invariant and renormalizable potential $V(\phi_1,\phi_2)$ for the 2HDM is defined in the 8-dimensional space of the Higgs field and is a hermitian combination of the electroweak-invariant combinations of (ϕ_i^+,ϕ_j), $\quad i,j = 1,2$.

$$\begin{aligned} V(\phi_1,\phi_2) =& m_{11}^2 \phi_1^\dagger \phi_1 + m_{22}^2 \phi_2^\dagger \phi_2 - \left[m_{12}^2 \phi_1^\dagger \phi_2 + h.c \right] + \frac{1}{2} (\phi_1^\dagger \phi_1)^2 \\ &+ \frac{1}{2} \lambda_2 (\phi_2^\dagger \phi_2)^2 + \lambda_3 (\phi_1^\dagger \phi_1)(\phi_2^\dagger \phi_2) + \lambda_4 (\phi_1^\dagger \phi_2)(\phi_2^\dagger \phi_1) \\ &+ \frac{1}{2} \lambda_5 (\phi_1^\dagger \phi_2)^2 + \left[\lambda_6 (\phi_1^\dagger \phi_1) + \lambda_7 (\phi_2^\dagger \phi_2) \right] (\phi_1^\dagger \phi_2) + h.c. \end{aligned} \tag{2.5.}$$

The λ are dimensionless coupling constants, λ_{1-4}, m_{11}^2, m_{22}^2 are real numbers, whereas the remaining parameters λ_{5-7} and m_{12}^2 can be complex. The non-zero imaginary parts of the complex parameters that cannot be removed by a rephasing transformation are the ones that give rise to explicit CP-violation in the Higgs sector, which is obviously one of the reasons for introducing an

extended Higgs sector with two Higgs doublets in the first place [21]. The 2HDM scalar potential, V, in Eq. 2.5 can violate CP either explicitly, via the complex phase ϕ in the soft $\mathcal{Z}_2$ breaking term $\left((\phi_1^\dagger\phi_2)^2 + h.c\right)$, or spontaneously due to a relative phase between the *vev* of the two doublets. In this way, with these additional sources of CP violation, the 2HDM model provides a natural explanation for the baryon asymmetry originating in the early Universe. If we want to avoid tree-level Higgs-mediated FCNCs, we can impose a softly broken discrete $\mathcal{Z}_2$ symmetry, $\phi_1 \to +\phi_1$ and $\phi_2 \to -\phi_2$ on the quartic terms of Eq. 2.5, which implies that $\lambda_6 = \lambda_7 = 0$, whereas $m_{12}^2 = 0$ is allowed.

In models of electroweak interactions with spontaneously broken gauge invariance, renormalizability limits the Higgs potential to the degree of four, terms of order greater than four have to be excluded because they are not renormalizable, therefore, the maximum power of the combination $\phi_i^\dagger\phi_j$ is 2. The 2HDM potential with all quadratic and quartic terms contains 14 free parameters and in contrast with the standard model, the potential is not unique, each set of parameters will lead to different mass eigenstates, interactions, Feynman rules, etc.

The two Higgs doublets are constructed analogously to the doublet in the SM, each has four degrees of freedom, and both acquire vacuum expectation values (*vev*) which spontaneously breaks the electroweak symmetry and give mass to the fermions, the $W^\pm$ bosons and the Z boson. The excitations around the degenerate minima of the potential are illustrated in Fig. 2.1. The bottom of the "Mexican hat" corresponds to a massless Goldstone state. Gauge freedom allows us to choose *vev* as the vacuum expectation value of the real part of ϕ with all choices of vacuum states being degenerate as well as physically equivalent. The existence of a large number of scalar degrees of freedom makes the vacuum landscape of a 2HDM more intriguing compared to the SM. This serves as a summary for a reader specifically intrigued by the 2HDM. However, a more elaborate explanation has been provided in Chap. 1 (refer to Sec. 1.2 for additional details). Also, the inclusion of loop diagrams in the potential modifies its shape allowing for three different scenarios:

- The vacuum is stable if there is only one minimum, or if there is another minimum higher than the electroweak minimum.
- If the additional minimum is lower or at the same level, quantum tunneling could make the vacuum decay. If the lifetime of the vacuum is larger than the age of our universe, the vacuum is metastable.
- If the lifetime of the vacuum is smaller than the age of the universe or if the vacuum is not bounded from below, it is unstable.

The 2HDM is governed by the choice of the Higgs potential parameters and also by the choice of the Yukawa couplings of the two scalar doublets to the

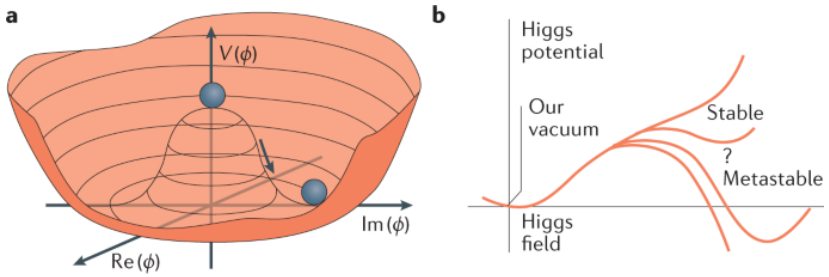


Figure 2.1. The Higgs potential in the 2HDM. Choosing any of the points at the bottom of the potential spontaneously breaks the rotational $U(1)$ symmetry (a). Quantum corrections can change the shape of the potential (b) [22].

three generations of quarks and leptons. Since the potential given by Eq. 2.5 is $U(2)$ invariant, values specified for the parameters m_{ij}^2, λ_i can only have definite (physical) meaning when a particular basis is specified for the scalar fields.

Let's try to express ϕ_1 and ϕ_2 in terms of vacuum expectation values, so a generic basis respecting the gauge symmetry $U(1)_{EM}$ can be written as

$$v_1 = <\phi_1> = \frac{v}{\sqrt{2}}\begin{pmatrix}0\\ \cos\beta\end{pmatrix}, \quad v_2 = <\phi_2> = \frac{v}{\sqrt{2}}\begin{pmatrix}0\\ e^{i\zeta}\sin\beta\end{pmatrix}, \tag{2.6.}$$

where $v = (\sqrt{2G_F})^{-1/2} \simeq 246$ GeV and the parameter β is related to the ratio of the *vev* of the two doublets. By convention $0 \leq \beta \leq \pi/2$ is chosen. A non-zero phase ζ results in a vacuum which breaks CP spontaneously. Note that the value of $\tan\beta = <\phi_2> / <\phi_1> = v_2/v_1$ at this point only determines one particular choice of basis. Since the 2HDM potential is invariant under a change of this basis, $\tan\beta$ can not be a physical parameter of the model in general. Within the set of CP-conserving bases defined above, there exists a special choice called the Higgs basis, in which only one of the two doublets is assigned a non-zero *vev*.

By requiring that v_1 and v_2 minimize the potential of Eq. 2.5, we find two relations between the parameters that must be satisfied:

$$\begin{aligned} m_{11}^2 &= \frac{-\lambda_1 v_1^3 + 2m_{12}^2 v_2^2 - \lambda_4 v_1 v_2^2 - \lambda_5 v_1 v_2^2}{2v_1} \\ m_{22}^2 &= \frac{2m_{12}^2 v_1 - \lambda_3 v_1^2 v_2 - \lambda_4 v_1^2 v_2 - \lambda_5 v_1^2 v_2 - \lambda_2 v_2^3}{2v_2} \end{aligned} \tag{2.7.}$$

It has been reported in [23] that if we can confine ourselves to the situation where the vacuum breaks neither electric charge nor CP, the above conditions can lead to several solutions and at most two non-degenerate minima. In other words, apart from the EW minimum in which the universe currently resides (

$v_1^2+v_2^2=246$ GeV), there exists another minimum somewhere around. Hence, opening for the possibility that the 2HDM might offer parameter points for which a neighbouring vacuum could actually be deeper than the one which corresponds to the observed W and Z boson masses. The EW minimum then loses its status as the global minimum and has been termed the *panic vacuum* in [23]. For the non-EW minimum to happen, the tunnelling from the non-EW minimum to EW one must be lower than the age of the universe.

When the symmetry is broken, the eight degrees of freedom from the $SU(2)$ doublets ϕ_1 and ϕ_2 are usually re-expressed in states with definite physical properties. The spectrum then contains three Goldstone modes: $G^{\pm}$ and G^0, which we say are "absorbed" to give mass to the gauge bosons $W^{\pm}$ and Z, reducing the number of physical Higgs states to five. Three of these states are neutral, of which two (h and H, with $m_h \leq m_H$), are CP-even, and one pseudo-scalar denoted A is CP-odd. The remaining two are a pair of charged Higgs bosons ($H^{\pm}$).

Let's see when introducing the mixing angle α to diagonalize the mass matrix for the CP-even states, how the doublets can be expressed

$$\phi_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2}(G^+\cos\beta - H^+\sin\beta) \\ v\cos\beta - h\sin\alpha + H\cos\alpha + i(G^0\cos\beta - A\sin\beta)\end{pmatrix},$$

$$\phi_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2}(G^+\sin\beta + H^+\cos\beta) \\ v\sin\beta + h\cos\alpha + H\sin\alpha + i(G^0\sin\beta + A\cos\beta)\end{pmatrix}. \tag{2.8.}$$

For any choice of Higgs potential and basis, we can deduce the Higgs masses from the potential in Eq. 2.5.

$$m_A^2 = \frac{m_{12}^2}{\sin\beta\cos\beta} - \frac{v^2}{2}(2\lambda_5 + \lambda_6\cot\beta + \lambda_7\tan\beta),$$

$$m_{H^\pm}^2 = m_A^2 + \frac{v^2}{2}(\lambda_5 - \lambda_4). \tag{2.9.}$$

Since the CP-even states mix, their mass matrix $\mathcal{M}$ is given by

$$\mathcal{M} = m_A^2 \begin{pmatrix} s_\beta^2 & -s_\beta c_\beta \\ -s_\beta c_\beta & c_\beta^2 \end{pmatrix} + v^2\mathcal{B}^2, \tag{2.10.}$$

where

$$\mathcal{B}^2 = \begin{pmatrix} \lambda_1 c_\beta^2 + 2\lambda_6 s_\beta c_\beta + \lambda_5 s_\beta^2 & (\lambda_3+\lambda_4)s_\beta c_\beta + \lambda_6 c_\beta^2 + \lambda_7 s_\beta^2 \\ (\lambda_3+\lambda_4)s_\beta c_\beta + \lambda_6 c_\beta^2 + \lambda_7 s_\beta^2 & \lambda_2 s_\beta^2 + 2\lambda_7 s_\beta c_\beta + \lambda_5 c_\beta^2 \end{pmatrix}. \tag{2.11.}$$

After performing a rotation of α to diagonalize $\mathcal{M}$, we obtain the masses of the

CP-even states

$$\begin{pmatrix} m_H^2 & 0 \\ 0 & m_h^2 \end{pmatrix} = \mathcal{R}(\alpha)\mathcal{M}^2\mathcal{R}^T(\alpha), \tag{2.12.}$$

with $m_H^2 \geq m_h^2$ and explicity the eigenvalues are the masses given by

$$m_{H,h}^2 = \frac{1}{2}\left[\mathcal{M}_{11}^2 + \mathcal{M}_{22}^2 \pm \sqrt{(\mathcal{M}_{11}^2 - \mathcal{M}_{22}^2)^2 + 4(\mathcal{M}_{12}^2)^2}\right] \tag{2.13.}$$

The gauge invariance requires that all coupling of the Higgs bosons to gauge bosons are determined in terms of the invariants $s_{\beta-\alpha} \equiv \sin(\beta-\alpha)$ and $c_{\beta-\alpha} \equiv \cos(\beta-\alpha)$, with $-\pi/2 \leq \beta-\alpha \leq \pi/2$.

2.2. The Yukawa sector in the 2HDM

The Yukawa sector in the 2HDM is expressed in the physical Higgs mass eigenstates (h, H, A, $H^{\pm}$). The mass eigenstates for the down- and up-type quarks are written as vectors in flavor space, denoted D and U respectively, and similarly for the leptons L and neutrinos ν.

$$\begin{aligned} -\mathcal{L}_{Yukawa} = & \frac{1}{\sqrt{2}}\overline{D}\left(\kappa D s_{\beta-\alpha} + \rho^D c_{\beta-\alpha}\right)Dh + \frac{1}{\sqrt{2}}\overline{D}\left(\kappa D c_{\beta-\alpha} - \rho^D s_{\beta-\alpha}\right)DH \\ & + \frac{1}{\sqrt{2}}\overline{u}\left(\kappa D s_{\beta-\alpha} + \rho^U c_{\beta-\alpha}\right)Uh + \frac{1}{\sqrt{2}}\overline{U}\left(\kappa U c_{\beta-\alpha} - \rho^U s_{\beta-\alpha}\right)UH \\ & + \frac{1}{\sqrt{2}}\overline{L}\left(\kappa L s_{\beta-\alpha} + \rho^L c_{\beta-\alpha}\right)Lh + \frac{1}{\sqrt{2}}\overline{L}\left(\kappa L c_{\beta-\alpha} - \rho^L s_{\beta-\alpha}\right)LH \\ & + \frac{i}{\sqrt{2}}\overline{D}\gamma_5\rho^D DA - \frac{i}{\sqrt{2}}\overline{U}\gamma_5\rho^U UA + \frac{i}{\sqrt{2}}\overline{L}\gamma_5\rho^L LA \\ & + \left[\overline{U}\left(V_{CKM}\rho^D P_R - \rho^U V_{CKM} P_L\right)DH^+ + \overline{\nu}\rho^L P_R L H^+ + h.c\right]. \end{aligned} \tag{2.14.}$$

The diagonal 3×3 matrices κ^F are given by $\kappa^F \equiv \sqrt{2}M^F/v$, where M^F are the corresponding mass matrices for the fermions ($F = D$, U, L), and as usual $P_{R/L} = (1 \pm \gamma_5)/2$. The generality of this Yukawa Lagrangian is present through the freedom to choose the components of each ρ^F arbitrarily, as required by CP conservation, which has six independent components.

CP conservation prerequisite ρ^F to have six components, furthermore, if we want to avoid flavor-changing neutral currents (FCNC) [1] at tree-level we constrain

[1] FCNC refers to processes where a neutral particle (such as a Higgs boson) changes the flavor of a quark without changing its charge. In the SM, FCNC processes are highly

the allowed size of the off-diagonal elements in ρ^F which involves fine-tuning of the individual elements in ρ^F. But how?

One can invoke Glashow and Weinberg's theorem, stating that the tree-level FCNC is absent from any theory where a given fermion does not pair with more than one Higgs doublet. If we impose an appropriate $\mathcal{Z}_2$ symmetry to distinguish ϕ_1 from ϕ_2. By assigning equally defined quantum numbers $\mathcal{Z}_2$ to the right-hand fermions, unwanted couplings can be avoided. When all fermions of a given type (up-quarks, down-quarks, and leptons) have a common quantum number $\mathcal{Z}_2$, the combinatorial possibilities for Yukawa couplings are enumerated in 2HDM language as types I, II, X, Y, III and FCNC-free [1]. The full list of 2HDM models is introduced in Table. 2.1.

Table 2.1. | Couplings of Φ_2 and Φ_1 doublets to charged leptons and up-/down-type quarks which introduce different 2HDM models which may or nor lead to natural flavor conservation. By convention, Φ_2 is the doublet to which up-type quarks couple.

Type	up-type quarks	down-type quarks	leptons
I: Fermiophobic	Φ_2	Φ_2	Φ_2
II: MSSM-like	Φ_2	Φ_1	Φ_1
X: Lepton-specific	Φ_2	Φ_2	Φ_1
Y: Flipped	Φ_2	Φ_1	Φ_2
III: FCNC at tree level	Φ_1,Φ_2	Φ_1,Φ_2	Φ_1,Φ_2
FCNC-free	Φ_1,Φ_2	Φ_1,Φ_2	Φ_1,Φ_2

If the symmetry $\mathcal{Z}_2$ must be explicit in the Yukawa sector, a dependency is introduced on the basis chosen for the Higgs sector. In this way, $\tan\beta$ is promoted to a physical parameter thanks to its appearance in the Yukawa couplings. In Tab. 2.2 I list the relations between ρ^F and κ^F for physically distinct types of Yukawa couplings realized with $\mathcal{Z}_2$ symmetry. The individual components of ρ^F and κ^F, are introduced the interactions of the neutral scalars, namely, h, H, and A, with Standard Model fermions, which are normalized relative to their Standard Model counterparts. These interactions are further detailed in Tab. 2.3.

2.3. Decay widths

At the tree level, the Higgs bosons ($H_1 = h$, $H_2 = H$, $H_3 = A$, $H_4 = H^{\pm}$) may decay into pairs of fermions, pairs of gauge bosons, one gauge boson and another

suppressed and occur only at loop levels.

[1] By finding a set of matrices that can undergo simultaneous diagonalization [24].

Table 2.2. | Relations between Yukawa coupling matrices ρ^F and κ^F, in the four different types of 2HDM Yukawa sectors where all fermions of a given type ($F = U$, D, L) share the same $\mathcal{Z}2$ quantum number.

	Type			
	I	II	III	IV
ρ^D	$\kappa^D \cot\beta$	$-\kappa^D \tan\beta$	$-\kappa^D \tan\beta$	$\kappa^D \cot\beta$
ρ^U	$\kappa^U \cot\beta$	$\kappa^U \cot\beta$	$\kappa^U \cot\beta$	$\kappa^U \cot\beta$
ρ^L	$\kappa^L \cot\beta$	$-\kappa^L \tan\beta$	$\kappa^L \cot\beta$	$-\kappa^L \tan\beta$

Higgs boson, or into pairs of lighter Higgs bosons. The decays $H_i \to \gamma\gamma$ and $H_i \to gg (i = 1...3)$ are induced at the one-loop level. The most relevant decay widths for the research reported in Chap. 5 are introduced.

2.3.1. $H_i \to f\overline{f}'$:

The coupling of the Higgs boson $H_i (i = 1...4)$ to a pair of fermions for the scalar is denoted C_S and C_P for the scalar and pseudoscalar part respectively. The coupling of $h \to d\overline{d}$ are $C_S = \frac{1}{\sqrt{2}}(\kappa^D s_{\beta-\alpha} + \rho^D c_{\beta-\alpha})$ and $C_P = 0$. It is only for the charged Higgs where both C_S and C_P are non-zero at the same time. In general, the decay width of the Higgs bosons to fermions is given by

$$\Gamma(H_i \to f\overline{f}') = \frac{N_c m_{H_i}}{8\pi} \left(\left[1 - (x_1 + x_2)^2\right] |C_S^2| + \left[1 - (x_1 - x_2)^2\right] |C_P|^2 \right) \times \lambda^{1/2}(1, x_1^2, x_2^2), \tag{2.15.}$$

where $x_1 = m_f / m_{H_i}$, $x_2 = m_{\overline{f}'} / m_{H_i}$ and,

$$\lambda(1, x, y) = (1 - x - y)^2 - 4xy. \tag{2.16.}$$

In Eq. 2.15, we refer to N_c as the number of "color" degrees of freedom in a given fermion which is $N_c = 3$ for quarks and $N_c = 1$ for leptons.

2.3.2. $H_i \to V H_j$:

In the case where the Higgs $H_i (i = 1...4)$ decay to another Higgs $H_j (j = 1...4)$ and one massive gauge boson $V =$(W, W^+, or W^-). The width in this case

$$\Gamma(H_i \to V H_j) = \frac{|C_{H_{iVH_j}}|^2 m_V^2}{16\pi^2 m_{H_i}} \lambda(1, \frac{m_{H_i}^2}{m_V^2}, \frac{m_{H_j^2}}{m_V^2}) \lambda^{1/2}(1, \frac{m_V^2}{m_{H_i}^2}, \frac{m_{H_j^2}}{m_H^2}) \tag{2.17.}$$

While the above Eq. 2.17 describes the decay for an on-shell decay of the gauge

Table 2.3. | Tree-level couplings of 2HDM neutral scalars h, H, and A normalized to their SM Yuakaw couplings, $\xi = y/y^{SM}$, where the shorthand notation $c_x = \cos(x)$ and $s_x = \sin(x)$ is used.

2HDM Type	I	II	III	IV
up-type quarks	ϕ_2	ϕ_2	ϕ_2	ϕ_2
ξ_{huu}	c_α/s_β	c_α/s_β	c_α/s_β	c_α/s_β
ξ_{Huu}	s_α/s_β	s_α/s_β	s_α/s_β	s_α/s_β
ξ_{Auu}	$\tan\beta^{-1}$	$\tan\beta^{-1}$	$\tan\beta^{-1}$	$\tan\beta^{-1}$
down-type quarks	ϕ_2	ϕ_2	ϕ_2	ϕ_1
ξ_{huu}	c_α/s_β	$-s_\alpha/c_\beta$	c_α/s_β	$-s_\alpha/c_\beta$
ξ_{Huu}	s_α/s_β	c_α/s_β	s_α/s_β	c_α/c_β
ξ_{Auu}	$-\tan\beta^{-1}$	$\tan\beta$	$-\tan\beta^{-1}$	$\tan\beta$
lepton	ϕ_2	ϕ_1	ϕ_1	ϕ_2
ξ_{hll}	c_α/s_β	$-s_\alpha/c_\beta$	$-s_\alpha/c_\beta$	c_α/s_β
ξ_{Hll}	s_α/s_β	c_α/c_β	c_α/c_β	s_α/s_β
ξ_{All}	$-\tan\beta^{-1}$	$\tan\beta$	$\tan\beta$	$-\tan\beta^{-1}$

boson if we consider off-shell decay, Eq.2.17 becomes

$$\frac{d\Gamma}{d_{x_1} d_{x_2}}(H_i \to H_j V^* \to H_j H_j f_1 \overline{f}_2 = K_{HvH} \frac{(1-x_1)(1-x_2) - k_H}{(1-x_1-x_2-k_H+k_V)^2 + k_{V\gamma V}}), \tag{2.18.}$$

where $k_H = m_{H_j}^2 / m_{H_i}^2$, $k_V = m_V^2 / m_{H_i}^2$, and $\gamma_V = \Gamma_V^2 / m_{H_i}^2$.
The normalization factor is

$$K_{HVH} = \frac{9 G_F}{16\sqrt{2}\pi^3} |C_{H_i V H_j}|^2 m_V^2 \delta_V', \tag{2.19.}$$

where $\delta_W' = 1$ and δ_Z' is given by

$$\delta_Z' = \frac{7}{12} - \frac{10}{9}\sin^2\theta_W + \frac{40}{27}\sin^4\theta_W. \tag{2.20.}$$

2.4. Constraints

This section will focus on the experimental and theoretical constraints on the 2HDM parameters that are considered in the analysis presented in Chap. 5.

2.4.1. Positivity of the potential

For the vacuum configuration to be stable, the Higgs potential of Eq. 2.5 must be positive in all field space (ϕ_1, ϕ_2) directions. This condition implies that

$$\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > -\sqrt{\lambda_1 \lambda_2}. \tag{2.21.}$$

These conditions ensure that the quartic couplings for each individual Higgs doublet are positive, preventing unbounded growth of the potential. When $\lambda_6 = \lambda_7 = 0$ are required in addition to impose $\mathcal{Z}_2$ soft symmetry breaking,

$$\lambda_3 + \lambda_4 - |\lambda_5| > \lambda_1 \lambda_2, \tag{2.22.}$$

becomes a requirement together with those in Eq. 2.21 to guarantee that the quartic couplings involving both Higgs doublets are bounded from below[1]. If this condition is not met, the potential may have regions with negative energy, indicating instability.

Likewise, to ensure that the mass matrix for the Higgs bosons is positive definite, meaning that the masses are real and positive

$$\lambda_1 \lambda_2 - \lambda_3^2 > 0, \tag{2.23.}$$

[1] In the context of the Two Higgs Doublet Model (2HDM), the statement that the energy is bounded from below refers to the requirement that the scalar potential has a global minimum and does not lead to unbounded negative values as the fields go to infinity.

If the energy is not bounded from below, it means that the potential can become infinitely negative for certain field configurations. This would imply an instability in the theory, as the energy can decrease without limit, leading to an unphysical situation.

2.4.2. Extrema of the Higgs potential

The ground states are described by the constant fields which minimize the potential. Thus, to find the excitation spectrum of the system, we need the extrema of the potential to be satisfied under these conditions

$$\left(\frac{\partial}{\partial\phi_1}\right)_{\phi_1=<\phi_1>,\phi_2=<\phi_2>} = 0, \quad \left(\frac{\partial}{\partial\phi_2}\right)_{\phi_1=<\phi_1>,\phi_2=<\phi_2>} = 0, \qquad (2.24.)$$

where ϕ_1, ϕ_2 define the vaccum expectation values $<\phi_1>=v_1$, $<\phi_2>=v_2$. These conditions guarantee that the extremum is a minimum for all directions in the plane (ϕ_1, ϕ_2), except in the direction of the Goldstone modes and this can be achieved in different possible scenarios; I) $v_1\neq 0$ and $v_2=0$, II) $v_1=0$ and $v_2\neq 0$, and III) $v_1\neq 0$ and $v_2\neq 0$. These cases allow for electroweak symmetry breaking since the first occurs only if the minimum of the potential occurs for non-zero expectation values of the scalar fields. In I(II), the potential has two global minima in the direction $\phi_1(\phi_2)$, one for $\phi_1=v_1/\sqrt{2}$ $(\phi_2=v_2/\sqrt{2})$ and one for $\phi_1=-v_1/\sqrt{2}$ $(\phi_2=-v_2/\sqrt{2})$ and it is up to us to decide which of the minima would break the symmetry and thus develop the theory in the vicinity of this minimum of the potential. Scenario III) would occur when the potential acquires four minima, the first two in the direction ϕ_1 corresponding to $\phi_1=v_1/\sqrt{2}$ and $\phi_1=-v_1/\sqrt{2}$. And the second two in the direction ϕ_2 corresponding to $\phi_2=v_2/\sqrt{2}$ and $\phi_2=-v_2/\sqrt{2}$.

Note that the scalar potential in the 2HDM which is responsible for the electroweak symmetry breaking contains several parameters which determine the behaviour of the vacuum. Vacuum instability can occur if the quartic couplings between the Higgs fields take on certain values, causing the potential to develop instability, leading to a lower energy vacuum state. Vacuum instability is characterized by the appearance of a deeper minimum in the scalar potential, which may have lower energy than the electroweak vacuum. If the universe were to tunnel into this low-energy vacuum state, it would cause a catastrophic change in the fundamental properties of particles and interactions. If the stability conditions are violated, it implies that the vacuum is either unstable or metastable which depends on the specific type of 2HDM considered as well as the chosen values of the model parameters.

2.4.3. Perturbativity and tree-level unitarity

Tree-level unitarity ensures that particle scattering amplitudes remain finite and do not violate tree-level unitarity (leading order) in which this involves the self-interactions of the Higgs bosons and the scattering amplitudes of longitudinal vector bosons (W and Z) with Higgs bosons. By examining these amplitudes, one can deduce upper bounds on the masses of the Higgs bosons in the 2HDM and the values of certain Higgs quartic coupling combinations (λ_i). These bounds

help identify regions of the parameter space where the theory remains consistent. It is worth noting that tree-level unitarity provides a necessary condition for the consistency of the theory, but it does not guarantee the absence of unitarity violations at higher orders or the absence of other theoretical issues. One should also consider other constraints, such as vacuum stability and experimental measurements, to fully analyze the 2HDM and its phenomenology.

2.4.4. Electroweak precision

As a consequence of the additional scalars introduced in the theory, the gauge bosons receive extra contributions, also called "oblique" corrections. The difficulty is that some combinations of gauge boson masses and their couplings, whose experimental values are known to agree with the SM prediction to great accuracy start to deviate and the challenge for the 2HDM theory is to predict a deviation that remains within the precision of the experimental measurement. The best example is provided by the ρ parameter,

$$\rho = \frac{m_W^2}{m_Z^2 \cos\theta_w^2}, \tag{2.25.}$$

where m_W and m_Z are the SM W and Z boson masses respectively and θ_W is the Weinberg weak mixing angle.

In the 2HDMs $\rho = 1$ at the tree-level same as in the SM. At the loop level, there are extra contributions concerning the SM that can explicitly be written as

$$\begin{aligned}\Delta\rho^{\text{2HDM}} = &\frac{1}{32\pi^2 vev^2}(F_{H^\pm,A} + \sin^2(\beta-\alpha)(F_{H^\pm,h} - F_{A,h}) \\ &+ \cos^2(\beta-\alpha)(F_{H^\pm,H} - F_{A,H}) \\ &+ 3\sin^2(\beta-\alpha)(F_{Z,H} - F_{Z,h} - F_{W,H} - F_{W,h})),\end{aligned}$$

with

$$F_{x,y} = \frac{m_x^2 + m_y^2}{2} - \frac{m_x^2 m_y^2}{m_x^2 - m_y^2}\ln\left(\frac{m_x^2}{m_y^2}\right). \tag{2.26.}$$

The condition $\rho - \rho_{SM} \simeq 1$ is only satisfied when there is a mass degeneracy between the charged and the neutral 2HDMs scalars, which is related to the limit in which custodial symmetry is recovered [25]. The ρ parameter is only an example of many observables that receive oblique corrections in the 2HDMs extended scalar sector. In a more general extension of the SM preserving the $SU(2)_L \times U(1)_Y$ gauge structure, the oblique corrections often are parametrized by the Peskin-Takeuchi parameters S, T and U [26] and some higher-order extensions of them [27]. The constraints on S and T for a fixed value of U = 0

are shown in Fig. 2.2.

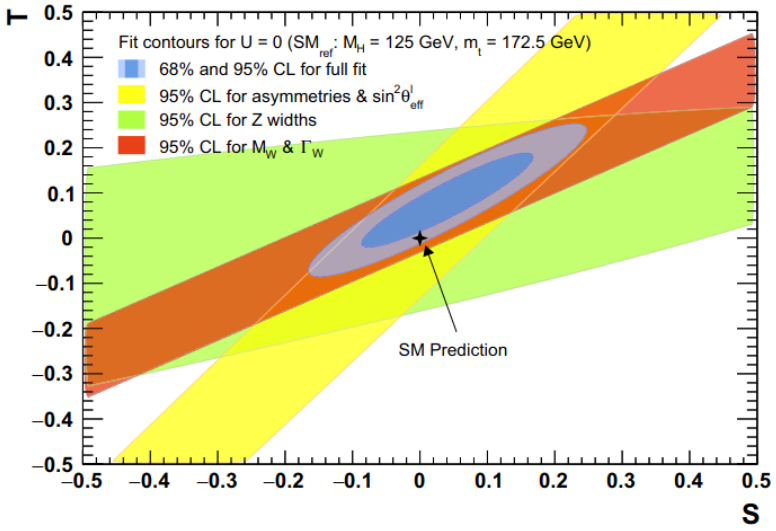


Figure 2.2. | Constraints at 95% CL in the oblique parameters S and T, with the U parameter fixed to zero. Individual constraints are shown from the Tevatron measurements of the effective leptonic electroweak mixing angle ($\sin^2\theta^l_{eff}$) (yellow), LHC Z partial and total widths (green) and W mass and width (red) [28].

2.4.5. The anomalous magnetic moment of the muon

The extended Higgs sector of the 2HDM contributes to the anomalous magnetic moment of the muon, $a_\mu = (g_\mu - 2)/2$. Since the experimentally measured value for a_μ shows a deviation of $\approx 4.2\sigma$ from the SM value [18] which makes the 2HDM with multi-lepton signatures an interesting model to test at the LHC. In general, at the tree level, the extra scalars that appear after the electroweak (EW) symmetry breaking in the 2HDM can have flavor-dependent couplings to quarks and leptons. But if the flavor-violating couplings involving μ and τ, the sizable contribution to δa_μ in the 2HDM can be generated via the 1-loop diagram mediated by the extra neutral Higgs bosons H and A which can be written as follow as given in [29]

$$\begin{aligned}\delta a_\mu &= a_\mu^{\text{2HDM}} - a_\mu^{\text{SM}} \\ &= \frac{m_\mu m_\tau \rho_e^{\mu\tau} \rho_e^{\tau\mu}}{16\pi^2} \left(\frac{\ln \frac{m_H^2}{m_\tau^2} - \frac{3}{2}}{m_H^2} - \frac{\ln \frac{m_A^2}{m_\tau^2} - \frac{3}{2}}{m_A^2} \right)\end{aligned}$$

$$\simeq -\frac{m_\mu m_\tau \rho_e^{\mu\tau}\rho_e^{\tau\mu}}{8\pi^2}\frac{\Delta_{H-A}}{m_A^3}\left(\ln\frac{m_A^2}{m_\tau^2}-\frac{5}{2}\right),$$

where the mass difference between H and A is denoted as $\Delta_{H-A}=m_H-m_A$, ρ_f^{ij} are the Yukawa couplings and a_μ^{2HDM} is the prediction of the magnetic moment of the muon in the 2HDM. If we want to see such deviation at the LHC the $\mu\tau$ couplings need to be $\mathcal{O}(0.1-1)$, but if the other couplings are also sizable and other various flavor processes severely constrain the scalar masses as discussed in [29] this search can be extremely challenging. In [30] it where shown interesting signatures, where a visible rate for 2HDM heavy Higgs pair production is possible to pursue at the LHC to explain the deviation in a_μ. Fig. 2.3 shows the required value of Yukawa coupling products $\rho_e^{\mu\tau}\times\rho_e^{\tau\mu}$ in the (m_A,Δ_{H-A})-plane also showed the constraint on the neutral heavy Higgs masses (H,A) in which 10 GeV $\leq\Delta_{H-A}\leq$ 100 GeV and $m_A\leq 680$ GeV to allow for $\delta a_\mu=2.8\times10^{-9}$.

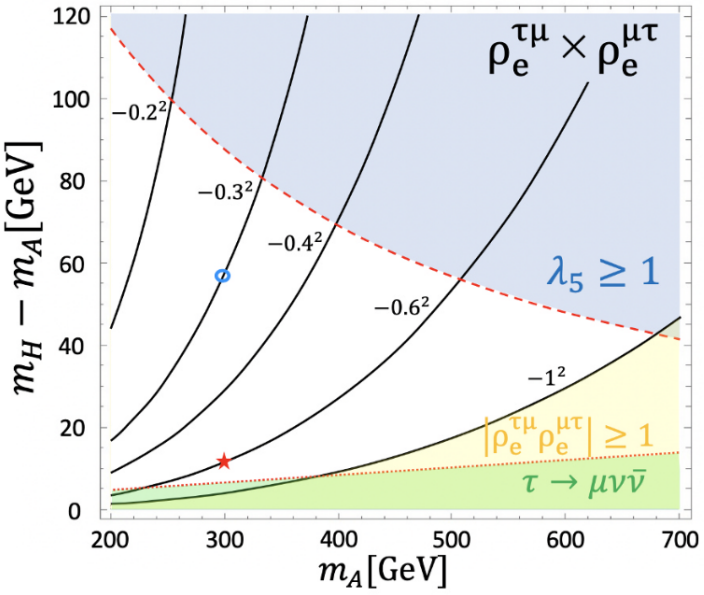


Figure 2.3. The required value of $\rho_e^{\mu\tau}\times\rho_e^{\tau\mu}$ to obtain $\delta a_\mu=2.8\times10^{-9}$. The grey-shaded region corresponds to $\lambda_5\geq1$. The green shaded region is excluded by the $\tau\to\mu\nu\bar{\nu}$ process mediated by the charged Higgs where $m_{H^\pm}=m_H$ is assumed. The yellow-shaded region corresponds to $|\rho_e^{\mu\tau}\times\rho_e^{\tau\mu}|\geq1$. The blue circle and the red star are two theoretically motivated landmarks for LHC search [30].

2.4.6. Higgs boson coupling measurements and flavor physics

Recall that a second Higgs doublet modifies the coupling strengths of the lightest neutral Higgs boson h to SM particles compared to those of the SM Higgs boson. Nevertheless, the size of the modification still depends on the 2HDM scenario and the model-free parameters in particular the angles α and β. Constraints on h are derived from the joint ATLAS and CMS Higgs boson coupling analysis [31] in which measurements sensitive to five Higgs boson production modes (ggF, VBF, WH, ZH, ttH) and five decay modes ($\gamma\gamma$, WW, ZZ, $\tau\tau$, $b\bar{b}$) were combined.

The constraints from the Higgs boson signal strength measurements on the four 2HDM scenarios are shown at 68% and 95% CL allowed regions in the $\tan\beta$ versus $cos(\beta-\alpha)$ plane in Fig. 2.4.

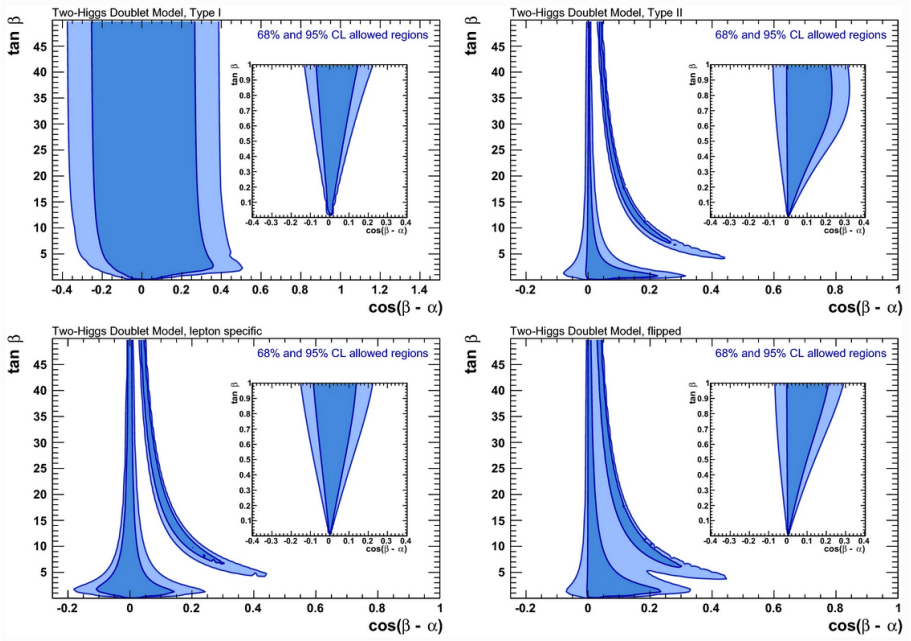


Figure 2.4. | The allowed parameter regions (68 and 95% CL) for the four 2HDM scenarios using the ATLAS and CMS combined Higgs coupling strength measurements. Shown are scans of $\tan\beta$ versus $cos(\beta-\alpha)$: Type-I (top left), Type-II (top right), lepton specific (bottom left) and flipped (bottom right) 2HDMs. The figure insets show a zoom of the region with $\tan\beta<1$ [28].

It can be seen that in all 2HDM scenarios except for Type-I α and β parameters are highly constrained. The allowed parameter regions are concentrated in two bands corresponding to solutions with $\beta\pm\alpha=\pi/2$. For $\beta-\alpha=\pi/2$, the Yukawa

structure of the SM is reproduced also called the alignment limit.While for the case $\beta - \alpha = \pi/2$, the neutral light Higgs bosons will differ from the SM-like Yukawa couplings by a sign flip which is still allowed by the combined coupling strengths measurements.

Some of the constraints on the 2HDM parameters arise from flavor physics which can impose stringent indirect limits on the parameter scan of the physics model far beyond the mass scales reachable in direct searches. The combined flavor constraints are shown in Fig. 2.5 for the four 2HDM Yukawa types.

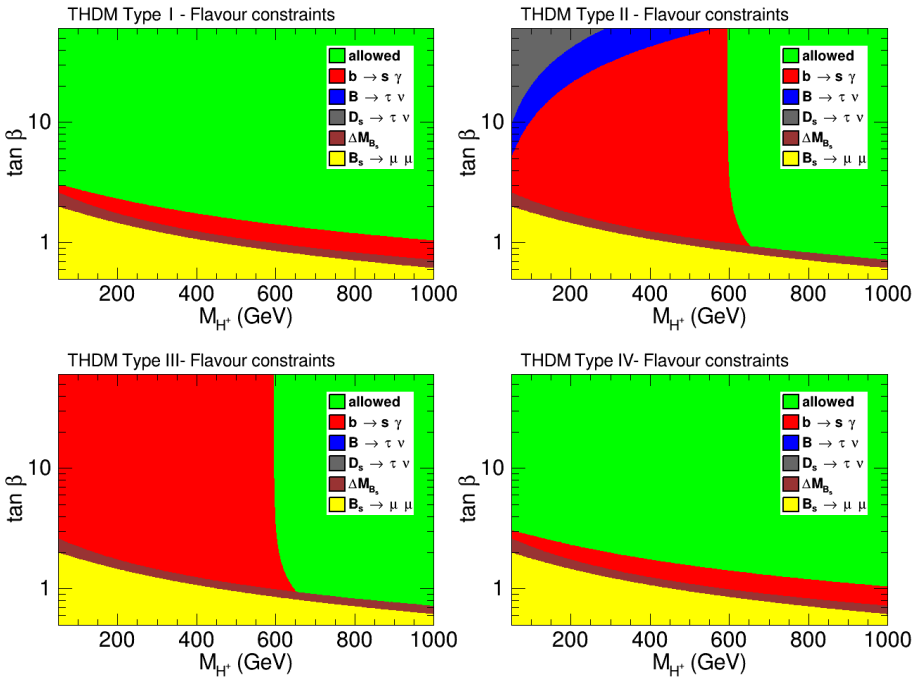


Figure 2.5. The obtained constraints from flavor physics on the 2HDM Yukawa types as a function of (M_{H^+},tanβ) and parameters and independent of the remaining 2HDM parameters. The color coding corresponds to exclusion at 95% CL. The green region is consistent with all constraints [32].

2.5. The 2HDM landscape at the start of LHC run 2

While the constraints discussed earlier indirectly provide an important exclusion in the parameter space, direct searches are still crucial to cover the full picture of the 2HDM. At the alignment limit, we have seen that the light scalar, h, resembles the standard model Higgs boson, making this limit inaccessible with only coupling measurements. A variety of searches for 2HDM neutral scalars have been foreseen during the LHC program and in the following I summarize

the most recent ones that have helped to constrain extended Higgs sectors, focusing mainly on the neutral scalars, H/A that are searched for in Chap. 5.

Most of the existing direct searches for BSM Higgs bosons focus on their conventional decays into a pair of quarks, leptons, or gauge bosons.

1. **Exotic Decays in BSM Sector:**
   - If the two neutral BSM Higgs states A and H have a sufficient mass splitting, $|m_A - m_H| > m_Z$, the exotic decay channel $A/H \to ZH/ZA$ opens up. For H and A masses heavier than two top quarks the decay to $t\bar{t}$ is dominant in the 2HDM [1], at least for moderate values of $\tan\beta$. Unfortunately, a possible signal strongly interferes with the tree-level background process $gg \to t\bar{t}$. Both ATLAS and CMS have performed searches listed in Tab. 2.4.

     Table 2.4. | 2HDM neutral scalars involving Z bosons.

Channel	ATLAS	CMS
$A/H \to ZH/ZA \to llb\bar{b}$	[33] (13TeV)	[34] (13TeV)
$A/H \to ZH/ZA \to ll\tau\tau$	–	[35] (8TeV)

   - Searches for 2HDM neutral scalars with no intermediate light scalar are shown Tab. 2.5.

     Table 2.5. | Conventional 2HDM channels.

Channel	ATLAS	CMS
$A/H \to \mu\mu$	[36]	[37]
$A/H \to b\bar{b}$	[38]	[39]
$A/H \to \tau\tau$	[40]	[41, 42]
$A/H \to \gamma\gamma$	[43–45]	[46], [47]
$A/H \to t\bar{t}$	–	[48]
$H \to ZZ$	[49]	[50]
$H \to WW$	[51]	[52]

2. **Exotic Decays of the SM-like Higgs:**
   - If the 2HDM neutral scalars, H/A, are sufficiently light, $m_{A/H} < m_h/2$, exotic decays of the light SM-like Higgs, h, such as $h \to AA/HH$ open up. In the alignment limit $cos(\beta - \alpha) = 0$, the decay $h \to AA$ is unsuppressed, while the decay $h \to HH$ vanishes. Tab. 2.6 lists the current LHC searches considered for such channels, focusing on masses of $m_A > 4$ GeV.

[1] Given that each 2HDM type is characterized by different coupling structures between the Higgs doublets and fermions. The exact behavior of H and A decays can vary between these types due to the differences in their interactions. In some cases, there might be additional decay channels that become important, especially at larger values of $\tan\beta$.

Table 2.6. | $h \to AA$ searches.

Channel	ATLAS	CMS
$h \to AA \to bb\bar{b}b\bar{b}$	[53]	–
$h \to AA \to bb\bar{b}\tau\tau$	–	[54]
$h \to AA \to bb\bar{b}\mu\mu$	[55]	[56]
$h \to AA \to \tau\tau\tau\tau$	–	[57]
$h \to AA \to \tau\tau\mu\mu$	[58]	[59]
$h \to AA \to \mu\mu\mu\mu$	[60]	[61]

3. **Exotic Decays into the SM Higgs:**
   - If we considered scenarios far from the alignment limit, the heavy CP-odd Higgs can decay into the SM-like Higgs via $A \to Zh$ and the heavy CP-even Higgs can decay into a pair of SM-like Higgses, $H \to hh$. Searchers for such channels from both CMS and ATLAS at 8 and 13 TeV are listed in Tab. 2.7.

Table 2.7. | Searches involve SM-like Higgs bosons.

Channel	ATLAS		CMS	
	8TeV	13TeV	8TeV	13TeV
$A \to Zh \to llb\bar{b}$	[62]	[63]	[64]	[65]
$A \to Zh \to ll\tau\tau$	[62]	–	[66]	[67]
$H \to hh$	[68]	[69]	[70]	[71]

Under the alignment limit, the decay branching fractions of $H/A \to AZ/HZ/H^{\pm}W^{\mp}$ are unsuppressed and the clean identification of the leptons from the Z decays makes searches such as $A/H \to ZH/ZA \to llb\bar{b}$ and $\tau\tau b\bar{b}$ very promising. The constraints in the m_A vs m_H plane of the 2HDM parameter space and for $cos(\beta-\alpha)=0$ is shown in Fig. 2.6. In the gap region along $m_A \sim m_H$ such exotic decay modes are kinematically inaccessible. At low $\tan\beta$ values both type-I and type-II constraints are very similar. At large $\tan\beta$ the type-II 2HDM has an enhanced reach due to the $\tan\beta$ enlargement of the bottom (and τ) Yukawa couplings. At $\tan\beta=30$ (green) the $A/H \to ZH/ZA$ search channel constrains the kinematically allowed region up to Higgs masses of $m_{A/H} \sim 800$ GeV, with the exception of very small daughter particle masses below 125 GeV. Unlike 2HDM type-II, type-I constraints for intermediate values of $\tan\beta$ are weakened due to the suppression of Yukawa couplings $\xi_{A_{ff}/H_{ff}} \sim \tan\beta^{-1}$, while it vanishes for larger values. Assuming an intermediate value of $\tan\beta=1.5$ and $cos(\beta-\alpha)=0$ Fig. 2.6 (bottom) shows the regions excluded by several Higgs searches, to note that constraints from the $A \to Zh$ and $H \to VV$, hh channels in such scenario vanish in the alignment limit, while search results from $A/H \to b\bar{b}$,

$\mu\mu$, and $t\bar{t}$ are too weak to set any constraint. For both 2HDM types I and II, the combination of all channels result in the exclusion of the majority of the mass region in which the light Higgs masses happens always to be below the di-top mass, mA, $mH < 2 \times m_t$.

For Higgs masses above $2 \times m_t$, where $A/H \to t\bar{t}$ decay channel opens up, processes such as $A/H \to ZH/ZA \to t\bar{t}Z$ become more interesting to probe. While at very low masses, m_A, $m_H < m_h/2$, decay channels of the SM-like Higgs $h \to AA$, HH can be produced given that this region is still unconstrained from both direct searches and indirect Higgs width precision measurements.

For non-degenerate masses of H and A, permitting the exotic decay channel of $A/H \to ZH/ZA$, Fig. 2.7 illustrates the constraints of the 2HDM parameter $\tan\beta$ as a function of m_A for this decay channel in type-I (left) and type-II (right) 2HDM. Assuming $m_H = 200$ GeV, the top panel of Fig. 2.7 demonstrate that the region close to the alignment limit $\epsilon[0.,0.2]$ type-I 2HDM can be excluded for $m_A > 290$ GeV and for $\tan\beta$ up to 10 and the reach in $\tan\beta$ reduces the farther this decay channel is probed away from the alignment limit. Smaller $m_A < 110$ GeV $\tan\beta < 2$ can be excluded as well for type-I 2HDM. On the other hand, in type-II the reach extends to large $\tan\beta$ values allowed because of the enhancement of the bbA/H production cross section and branching fraction. It can be seen that by choosing such parameters for type-II 2HDM, almost the entire range of $\tan\beta$ can be constrained by the $A/H \to ZH/ZA$ channel. The lower panel of Fig. 2.7 illustrates the global constraints from a variety of direct search channels of H/A. While for $m_A < 300$ GeV, the strongest conventional search constraints are related to the decay of the pseudoscalar A, at large $m_A > 300$ GeV constraints mainly come from direct searches for H, whose mass is fixed to $m_H = 200$ GeV in this case and therefore, there is no dependence on m_A for the $\tau\tau$, $\gamma\gamma$ and $4t$ exclusion limits in the large m_A region.

The requirements and implications regarding the mass gap between A and H, as well as the charged Higgs mass ($H^{\pm}$), are often subject to experimental and theoretical constraints. These constraints can come from precision measurements of electroweak observables, flavor physics, and direct searches for these new particles. The specific mass gap usually also depends on the model parameters, such as the values of the Higgs potential parameters and the mixing angles between the different Higgs bosons.

In the 2HDM framework, many paths can be taken to search for non-SM Higgs bosons. As can be seen earlier, there are cases where existing constraints based on direct searches and SM Higgs precision measurements result in strong constraints of their phase space. By studying the constraints of type II 2HDM parameter spaces, there is still a window for an exotic decay of $A/H \to ZH/ZA$, which is the most promising channel given the large branching fraction as well as as the clean experimental signal of $\ell^+\ell^- b\bar{b}$. This decay channel is the cornerstone

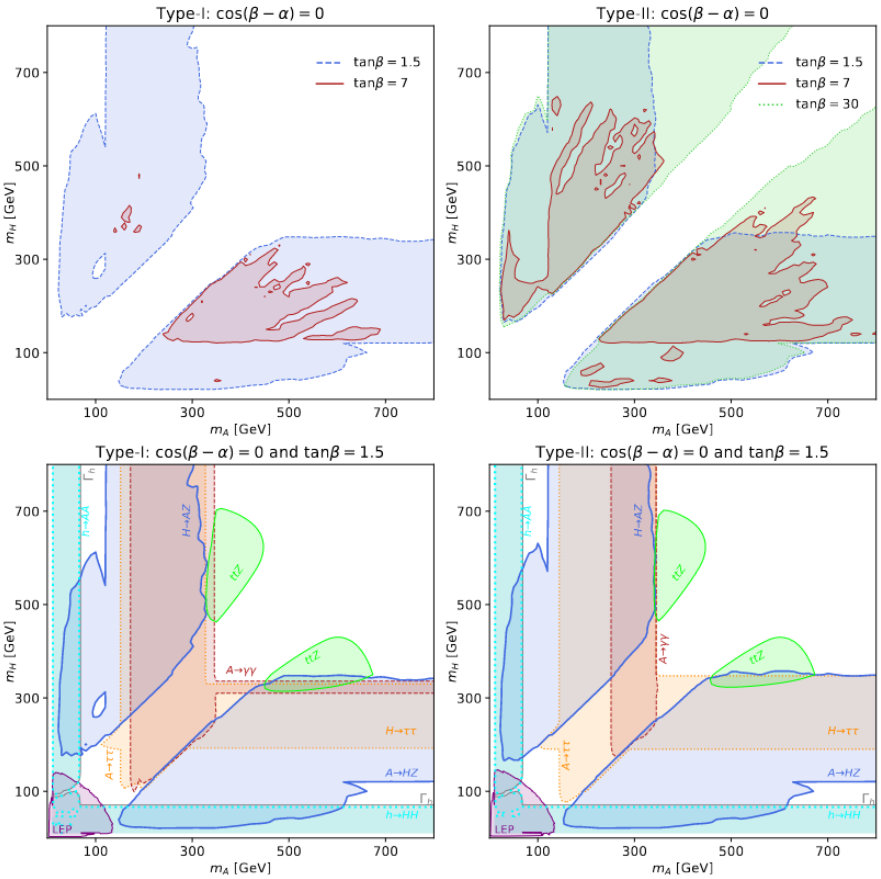


Figure 2.6. | Constraints of type-I (left) and type-II (right) 2HDM in the alignment limit $cos(\beta-\alpha)=0$ given in m_A vs m_H mass plane. **Top:** Parameter space excluded at 95% C.L. by the $A/H \to ZH/ZA$ search for $\tan\beta = 1.5$ (blue), 7 (red) and 30 (green). **Bottom:** Assuming $\tan\beta = 1.5$, the constraints from LHC searches are shown for $A/H \to ZH/ZA$ (blue), $A/H \to \tau\tau$ (dotted orange), $A/H \to \gamma\gamma$(dashed brown), $h \to AA/HH$ (dotted cyan) and $t\bar{t}Z$ production (green) as well as LEP searches (purple) and the Higgs width precision measurement Γ_h (grey). Constraints from $A \to Zh$ and $H \to VV$, hh channels vanish [72].

of this work and will be presented in Chap. 5.

2.6. From theory to experiment: event modeling and generation

If we want to test the prediction of any BSM theory or any SM process, the use of precisely simulated events as they might be produced in nature and observed

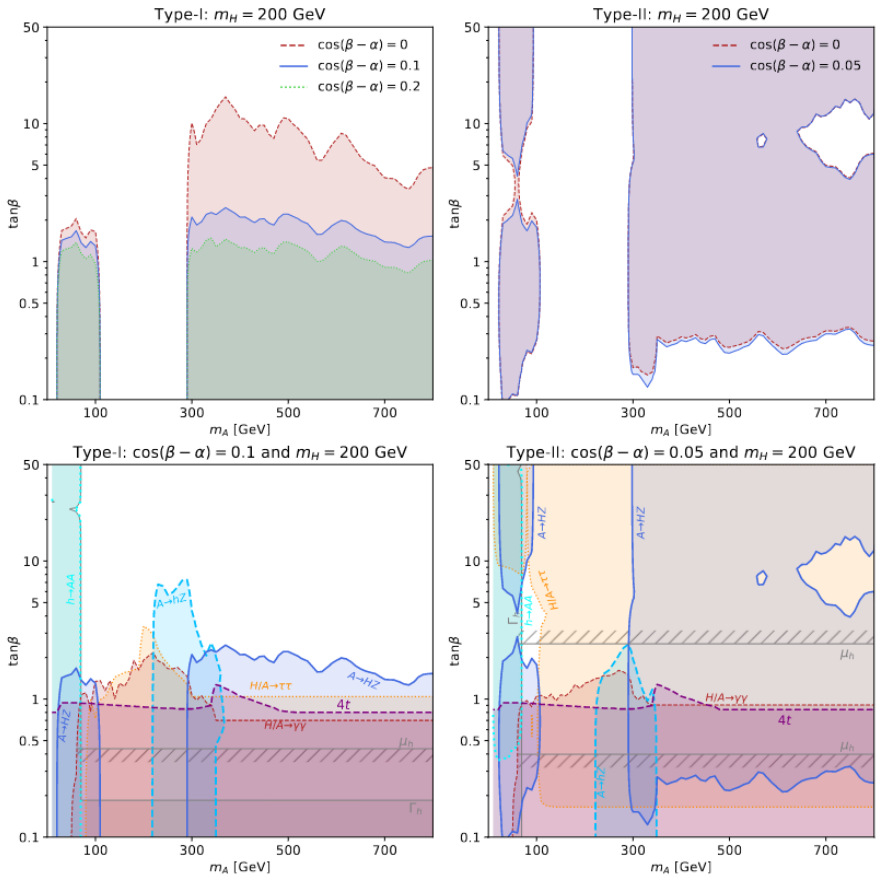


Figure 2.7. | Constraints of type-I (left) and type-II (right) 2HDM in $\tan\beta$ vs m_A mass plane, assuming $m_H = 200$ GeV. **Top:** Parameter space excluded at 95% C.L. by the $A/H \to ZH/ZA$ search for $cos(\beta-\alpha)=0$ (red), 0.1 (blue) and 0.2 (green) in the left panel and $cos(\beta-\alpha)=0$ (red) and 0.05 (blue) in the right panel. **Bottom:** Constraints at 95% C.L. for $cos(\beta-\alpha)=0.1$ from LHC searches for $A \to ZH$ (blue), $A \to Zh$ (dashed dark blue), $H \to AA$ (dashed cyan), $H/A \to \tau\tau$(dotted orange), $H/A \to \gamma\gamma$ (dashed brown) and $4t$ production (dashed magenta) as well as the global fit of SM-like Higgs couplings strength μ_h (grey hatched region) and the Higgs width measurement Γ_h (grey) [72].

by a perfect detector is a key ingredient in any physical data analysis. Therefore, we need (I) higher orders in perturbation theory; (II) and smart algorithms to estimate higher-order perturbative quantum chromodynamics (QCD). CMS relies on Monte Carlo event generators such as `PYTHIA` and `MC@NLO` to produce simulated events that mimic collider-level scattering events with high accuracy.

Event generation starts with the hard-scattering cross section for a given process at a given order in perturbation theory. The production of an N-particle final state in hadron-hadron collisions can be written as follows:

$$\sigma_N = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu^2) f_b(x_2, \mu^2) \hat{\sigma}_N^{ab}, \tag{2.27.}$$

where the hadronic cross section factorizes into universal parton distribution functions (pdf) of the two protons, f_a and f_b, and the partonic cross section $\hat{\sigma}_N^{ab}$ for a specific partonic initial state (ab). The pdf $f_a(x_1,\mu^2)$ is the suitably-regularised hadronic matrix element for the inclusive distribution to find a parton a with energy fraction x_1 in the beam hadron renormalized at scale μ^2.

These incoming protons define the partonic substructure in terms of flavor composition and momentum sharing. The hard scattering process of the two initial hadrons, shown as a red blob in Fig. 2.8 takes place at very large momentum-transfer scales for which the strong coupling constant α_s is small. An event is characterized by the color flow and momentum-transfer scale that outgoing observables in the hard process bring. Obviously, it is desirable to include as many orders as possible to make this event simulation as close as possible to data. But, in practice due to the limitations of computing power, all fixed-order calculations are limited to some low power of the coupling constant α_s.

The complexity of a hadron-collision process is illustrated in Fig. 2.8.

The computation of the partonic cross-section in Eq. 2.27 requires a phase space integration over the differential cross-section thus,

$$\hat{\sigma}_N^{ab} = \int_{cuts} d\hat{\sigma}_N^{ab} = \int_{cuts} \left[\prod_{i=1}^{N} \frac{d^3 q_i}{(2\pi)^3 2E_i} \right] \delta^4(p_1 + p_2 - \sum_i^N q_i) |\mathcal{M}^{ab}_{p_1 p_2 \to \vec{q}}|^2, \tag{2.28.}$$

where $d\hat{\sigma}_N^{ab}$ is the weight associated to a given kinematic configuration $p_1 p_2 \to \vec{q}$ for the process under study. The challenge of this step comes down to the multi-dimensional phase integral and the parton-level matrix element $|\mathcal{M}^{ab}_{p_1 p_2 \to \vec{q}}|^2$ that has to be computed fully differentially. While for $2 \to 2$ and $2 \to 3$ processes, various techniques assist the computation of the matrix element at leading order, as can be seen in Tab. 2.8, in more complicated processes where it becomes necessary to calculate higher-order tree-level matrix elements with many additional partons in the final states, the automated evaluation of Feynman diagrams becomes prohibitively time-consuming, because of the rapid increase in the number of diagrams.

A full event simulation of hadron interactions, like proton-proton collisions within the LHC requires several but there are various "levels" at which we can try to compare our theory prediction to data:

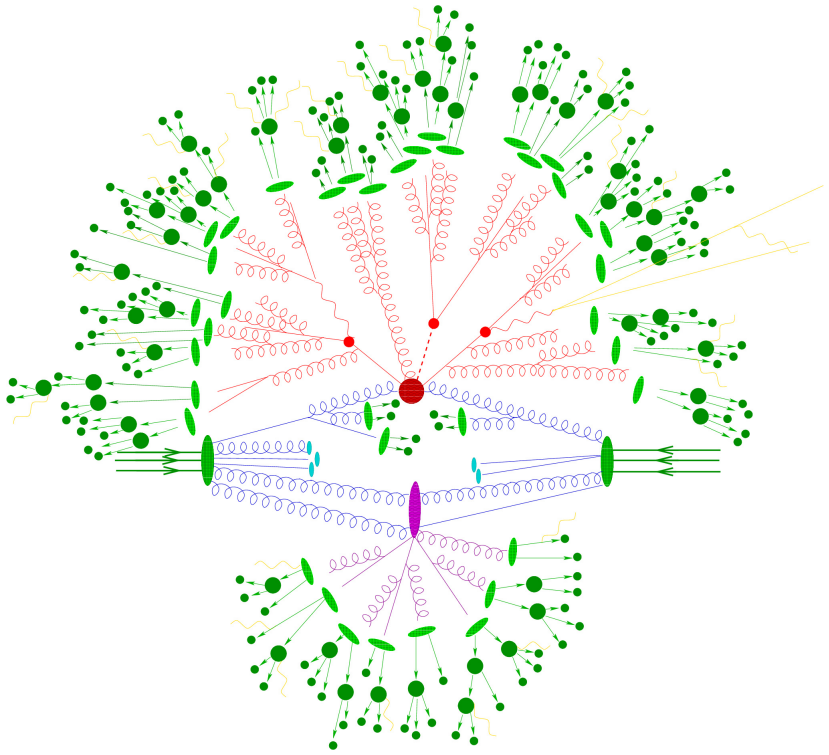
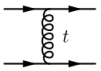
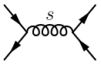
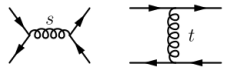
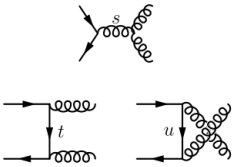
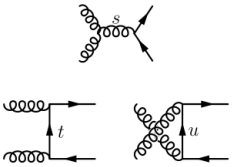
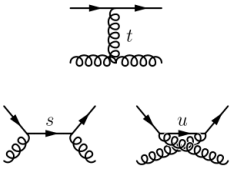
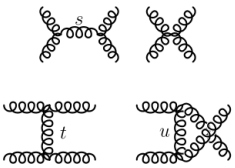


Figure 2.8. | A representation by Monte-Carlo generators of hadron-hadron scattering event. The triple green arrows represent the incoming hadrons. The red blob illustrates the hard interaction. The outgoing red lines are the decays of resonances produced in the collision, as well as the final-state-radiation (FSR) simulated by the parton shower. The initial-state partons and the initial-state-radiation (ISR) are shown as blue lines. Light green ellipses represent the hadronization process, from where hadrons (dark green circles) emerge and (possibly) decay. Finally, multiple partonic interactions (MPI) and Underling Events (UE) are shown in pink and light blue [73].

- **Particle/Parton level:** Define our truth extended *fiducial phase space*[1] based on the kinematics of MC particles at particle level - that is referred to the level in which stable particles are produced as a result of a collision.
- **Particle/hadron level:** Hadron level follows the parton level with an additional step consisting of hadronization.
- **Detector/reconstruction level:** At this stage, the theory results differ from particle level in including an additional detector simulation, that takes

[1]The fiducial definition is a phase space at particle level that is designed to correspond reasonably closely to what is experimentally accessible by the detector and reconstruction/identification algorithms.

Table 2.8. | Diagrams for $2 \to 2$ QCD parton scatterings and the expression for the squared matrix elements when calculated explicitly. $s = (p_1 + p_2)^2$, $t = (p_3 - p_1)^2$ and $u = (p_3 - p_2)^2$ are Mandelstam variables [74].

Process	Amplitudes	$\overline{\sum} \lvert\mathcal{M}\rvert^2 \,/\, (4\pi\alpha_s)^2$
$q\,q' \to q\,q'$		$\frac{4}{9}\frac{s^2+u^2}{t^2}$
$q\,q \to q\,q$		$\frac{4}{9}\frac{s^2+u^2}{t^2} + \frac{4}{9}\frac{s^2+t^2}{u^2} - \frac{8}{27}\frac{s^2}{tu}$
$q\,\bar{q} \to q'\,\bar{q}'$		$\frac{4}{9}\frac{t^2+u^2}{s^2}$
$q\,\bar{q} \to q\,\bar{q}$		$\frac{4}{9}\frac{s^2+u^2}{t^2} + \frac{4}{9}\frac{t^2+u^2}{s^2} - \frac{8}{27}\frac{u^2}{st}$
$q\,\bar{q} \to g\,g$		$\frac{32}{27}\frac{t^2+u^2}{tu} - \frac{8}{3}\frac{t^2+u^2}{s^2}$
$g\,g \to q\,\bar{q}$		$\frac{1}{6}\frac{t^2+u^2}{tu} - \frac{3}{8}\frac{t^2+u^2}{s^2}$
$q\,g \to q\,g$		$-\frac{4}{9}\frac{s^2+u^2}{su} + \frac{s^2+u^2}{t^2}$
$g\,g \to g\,g$		$\frac{9}{2}\left(3 - \frac{tu}{s^2} - \frac{su}{t^2} - \frac{st}{u^2}\right)$

finite resolution effects into account, and possibly also detailed aspects of particular detectors.

Unfortunately, the generation of parton-level events is inadequate for many applications in experimental high-energy physics and so we need to take a step closer to more accurate calculations. At the scale where α_s strong interaction couplings become important and where perturbative methods can no longer be applied, the parton shower takes place. Parton shower event generators such as HERWIG, `PYTHIA`, and SHERPA, perform calculations of cross sections according to an approximation to the standard model or its possible extensions. In a nutshell, these algorithms allow us to simulate the final states of high-energy collisions in full detail down to the level of individual stable particles. Bear in mind that the parton shower grows with decreasing measure of the hardness of interactions and a scattering process appears differently depending on the hardness scale at which one examines it. At the hardest scale, there are just a few partons (typically quarks and gluons) while at the softer scale (low-energy), the partons tend to split, making more partons in what we call a "parton shower".

The goal of these computer programs is to simulate a large number of collision events, each consisting of a list of final-state particles, such that the probability of producing an event with a given list is very close to the probability of an actual event being produced in nature. The first step of the simulation starts with a hard subprocess in which its constituents interact at a high momentum scale to produce a few outgoing fundamental objects, those are SM quarks, leptons and/or gauge or Higgs bosons, or the hypothetical new particle that we are looking for in some new BSM theory. The partons (quarks and gluons) involved, as well as any new particles with color, radiate virtual gluons, which can themselves emit further QCD radiation in the form of gluons or produce quark-antiquark pairs, leading to the formation of parton showers. It is of interest to know what is an evolution variable in the showering algorithm. If we consider a splitting process $i \to j + k$ the process cannot be achieved if all partons are on their mass shells. The dominant contribution will come from a configuration in which evolution variables are understood as virtual masses-squared (the virtualities, q^2) are strongly ordered, with the parton nearest to the hard subprocess farthest from its mass shell and the virtualities falling sharply as the shower evolves away from it. The shower is resumed when the virtualities have fallen to the hadronization scale, $q^2 = Q_0^2 \approx 1\text{GeV}^2$, where Q is the momentum transfer scale of the hard subprocess. Additionally, the parton shower not only takes into account the collinear-enhanced real parton emissions at each order in perturbation theory but also the quantum loop effects of the same order; such effects are included in the probability of not splitting during evolution from scale q_1^2 to q_2^2, which is given by the Sudakov form factor

$$\Delta_i(q_1^2, q_2^2) = exp- \int_{q_1^2}^{q_2^2} \frac{dq^2}{q^2} \frac{\alpha_s}{2\pi} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} dz \int_0^{2\pi} d\phi P_{ji}(z, \phi). \qquad (2.29.)$$

The z integration bounds specify the range of z within which the split can be resolved. An emission outside this range is too soft or at too low an angle to be detected. An example case study is a process $q\overline{q} \to Z$. If we want to simulate the production of Z, then the quark and antiquark at the end of their shower must have exactly the right 4-momenta to combine to form a Z boson mass system. If we decided at the initial low scale to choose their momenta at the beginning of the shower, this would not be ideal because the efficiency of Monte Carlo would be very low. Another way is to use a backward evolution. The way it works is to decide subject to the kinematic constraint $x_1 x_2 S = s$ where S and s are the energies of the center of mass of the collider and the sub-squared process, the longitudinal momentum fractions x_1 and x_2 of the incoming partons from the hard subprocess need to be chosen first. The incoming partons are then evolved backward, gaining energy with each emission, down to the small scale appropriate for the incoming hadron constituents.

We also often use the term initial-state-shower or radiation, which we refer to as the shower that grows on the incoming parton of the hard subprocess, of which there are two at a hadron collider. In such a scenario, the constituent parton from each of the incoming hadrons starts at a high energy and low virtuality and evolves to higher space-like virtuality by emitting partons and losing energy. The showering of these partons terminates when they collide to initiate the hard subprocess, which sets the scale that limits the endpoint virtualities of the showers. On the other hand, an end-state shower or radiation means an outgoing part of the hard sub-process in which the evolution of the shower is very similar to the above; note that the primary parton, in this case, starts at high energy and a large time-like virtuality scale Q^2 set by the subprocess, it continues to lose energy and virtuality until that all of its descendant partons have fallen to scale Q_0^2 at which splitting is complete.

At this stage, our theoretical description of a scattering event is starting to look a lot closer to those occurring in a collider but not exactly, there are additional interactions beyond the main collision under study that we often refer to as the Underlying Event (UE), which can be thought of as the remnants of scattering interaction that is distinct from the hard scattering process, PS, and hadronization. Contents of UE include initial and final state radiation, beam-beam remnants, multiple parton interactions, and pile-up noise.

Chapter 3. The LHC and the CMS experiment at CERN

The absence of evidence of the missing piece of the standard model, the Higgs boson, from both the Tevatron proton-antiproton collider and the Large electron-positron (LEP) collider implied that if it were to exist, its mass should be at higher energies than that provided by each of these experiments. After eleven years of forefront research, LEP was scheduled to close at the end of September 2000 opening the way to the Large Hadron Collider (LHC). Drawing hints to new physics LEP experimental run was only extended by one month. Note that the highest energy achieved by LEP was $\sqrt{s}$=209 GeV at a center of mass energy and that was only accomplished in the last months of the experiment. For an electron-positron pair that annihilated to make a Z and Higgs bosons, the highest energy Higgs boson which could be produced this way would have a mass of 118 GeV since the Z bosons mass is 91 GeV. Some of the energy is always lost as kinetic terms, so in practice, the limit reached was a little lower, around $\sim$ 114 GeV.

As a matter of fact, to date LEP is the most powerful lepton accelerator ever built and still holds the highest-energy lepton accelerator record. LEP experiment could have extended its reach further by running much longer and accumulating statistics. Perhaps the experiment would have achieved 116 GeV, but not 125 GeV which would have to run for quite a long time to detect the Higgs boson that was discovered later by the LHC [75–77].

It is important to point out that leptons are relatively light, their collisions are clean but cannot reach the same energy that can be achieved with heavier particles. On the other hand, hadrons are subatomic composite particles composed of quarks and anti-quarks held together by the strong force, as was discussed in Sec. 1.1 of Chap. 1, and are relatively heavy; protons, for example, have a mass almost 2000 times greater than electrons. Because of their higher mass, they can be accelerated to much higher energies [1], which is the key to directly

[1] This is primarily due to the energy gained by a particle when accelerated which is proportional to its charge and the acceleration voltage. Since hadrons have higher masses and are charged, they can gain much more energy compared to lighter leptons when accelerated through the same voltage. In addition, charged particles when accelerated emit synchrotron radiation. The energy loss due to synchrotron radiation increases with the fourth power of the particle's charge and inversely with the third power of its mass. Since leptons have smaller masses and charges compared to hadrons, they lose more energy to synchrotron radiation

observing new particles or interactions. This is what the LHC is built for, it is able to attain an energy never reached before which can potentially reveal unexpected results that no one has ever thought of!

3.1. The LHC accelerator

The LHC re-used the LEP tunnel. It consists of a 27-kilometer ring of superconducting magnets with a number of accelerating structures to boost the energy of the proton along the way. The proton journey at the LHC descends from a complex succession of machines that accelerate particles to increasingly higher energies. The LINAC4 (Linear accelerator 4) became the source of proton beams for the CERN accelerator in 2020. It accelerates negative hydrogen ions (that is, a hydrogen atom that has captured an extra electron), to 160 MeV to prepare them to enter the Proton Synchrotron Booster (PSB). During injection from LINAC4 into the PSB, the ions are stripped of their two electrons, leaving only protons. These are further accelerated to 2 GeV before the injection into the Proton Synchrotron (PS) that pushes the beam up to 26 GeV. Protons are then sent to the Super Proton Synchrotron (SPS), where they are accelerated further up to 450 GeV.

Finally, in ultrahigh vacuum, the protons make their way to the two beam pipes of the LHC where they travel at close to the speed of light before they are made to collide. The beams travel in opposite directions and in separate beam pipes, guided around the accelerator ring by a strong magnetic field maintained by superconducting electromagnets. The magnets operate at a temperature of 1.9 K (-271.25°C), colder than the 2.7 K (-270.45°C) of outer space. Since we want to avoid collisions with gas molecules, the beam vacuum pressure is about 10^{-7} Pa in the beam pipe at cryogenic temperature ($\sim$ 5 K) and lower than 10^{-9} Pa close to interaction points where collisions take place. It takes 4 minutes and 20 seconds to fill each LHC ring and 20 minutes for the protons to reach their maximum energy of 6.8 TeV. After circulating for many hours inside the LHC beam pipes [1], the two beams are brought into collision inside four detectors – ALICE, ATLAS, CMS and LHCb as shown in Fig. 3.2 – where the total energy at the collision point is equal to 13.6 TeV.

3.1.1. LHC timelines and data taking period

The LHC first went operational for an initial test on 10 September 2008. Its first run was achieved between 2010 and 2013 at an energy of up to 4 TeV per beam which has been very successful and culminated in the discovery of the Higgs boson. Then the LHC was shut down on 13 February 2013 for its first 2-year upgrade called Long Shutdown 1 (LS1), which was to affect many

when accelerated, which limits how much energy they can attain.

[1] To achieve stability of the beam for an optimal collision with more chances of interactions.

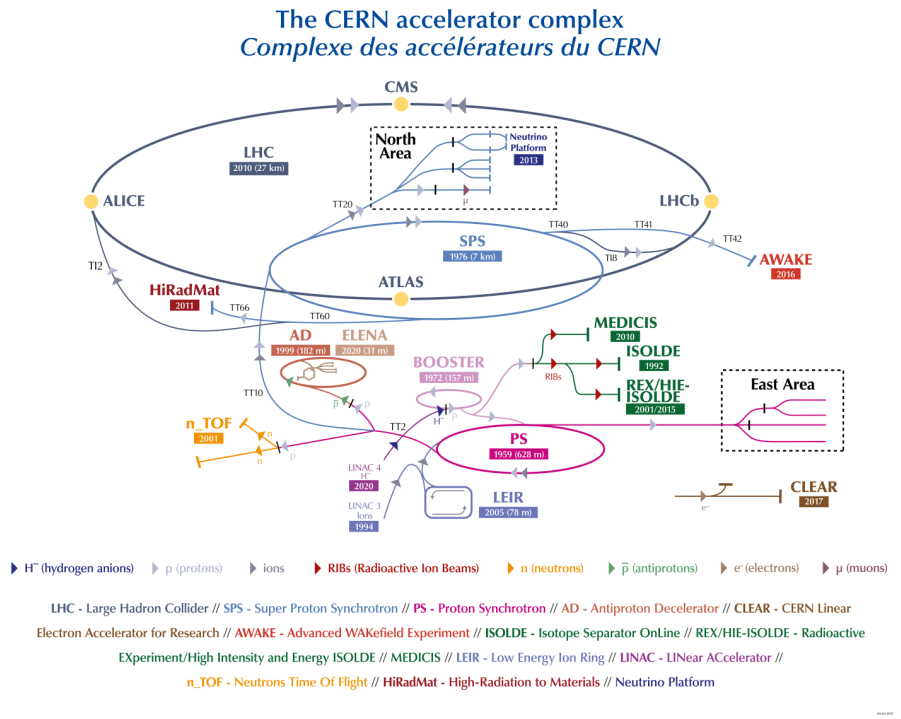


Figure 3.1. The proton journey [78].

aspects of the LHC: enabling collisions at 14 TeV, improving its detectors and pre-accelerators (the PS and the SPS).

The second operational period of the LHC took place between 2015 and 2018 and started with lower energy than initially planned, 6.5 TeV per beam for a total collision energy of 13 TeV. Data accumulated during run 2 was used to search for an extended scalar sector reported in Chap. 5.

The Long Shutdown 2 (LS2) started on December 10, 2018, for a significant upgrade in preparation for the High Luminosity Large Hadron Collider (HL-LHC) project which will increase the luminosity by a factor of 10. The LS2 ended in 2022 allowing LHC's third operational run to begin with an energy of 6.8 TeV per beam and is expected to continue until 2026.

3.1.2. Luminosity and pileup

The luminosity (L) is one of the most important parameters of an accelerator. It measures the number of collisions produced in a detector per cm^2 and per second. The higher the luminosity is, the larger the number of collisions that

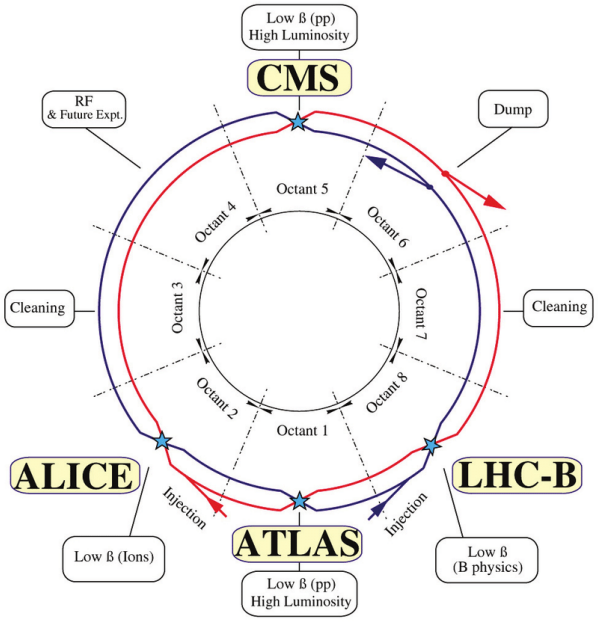


Figure 3.2. | Sector layout of the Large Hadron Collider. Blue stars mark the four interaction points in the experimental halls where the counter-rotating beams cross each other [79].

occur. L can be written semi-qualitatively [1] as follow:

$$L = \frac{N^2}{(tS_{\text{eff}})} = \frac{fN^2}{(4\pi\sigma^2)}, \tag{3.1.}$$

with f the bunch crossing frequency, N number of protons per bunch, t time between bunches, S_{eff} section effective of collision that depends on the cross-section of the bunch, and σ the transversal size of the bunch at the interaction point. Considering different number of protons per bunch crossing, and x and y components for σ separately, the luminosity can be written such that

$$L = \frac{fN_1N_2}{(4\pi\sigma_x\sigma_y)}. \tag{3.2.}$$

Moreover, L is often expressed in terms of ϵ (emittance) and β (amplitude function) as follows:

[1] This formula provides a simplified approximation of the luminosity and the actual luminosity calculation but it can be more complex due to various factors like the shape of the colliding bunches, the focusing of the beams, and other beam dynamics effects.

$$L \approx \frac{f N_1 N_2}{(4\epsilon\beta)}. \tag{3.3.}$$

The emittance ϵ is defined as the smallest opening you can squeeze the beam through as illustrated in Fig.3.3. Keeping the emittance low will result in higher luminosity. The amplitude function, β, is determined by the accelerator magnet configuration and powering, when expressed in terms of σ(cross-sectional size of the bunch) and the transverse emittance, the amplitude function β becomes $\beta = \frac{\pi\sigma^2}{\epsilon}$.

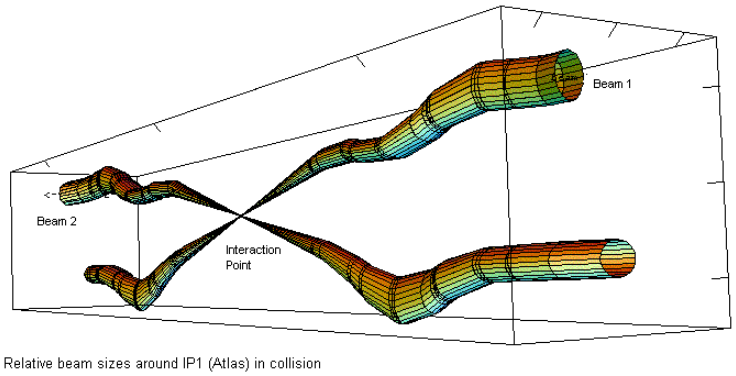


Figure 3.3. | Beam envelopes in the interaction region around point 1 ATLAS showing how the beam sizes are reduced on each side of the interaction point following the squeeze [80].

Thus, to achieve a high luminosity, all one has to do is make high population bunches (N) of low emittance (ϵ) to collide at high frequency at locations where the beam optics provide as low values of the amplitude functions (β) as possible. For any collider experiment reducing these parameters depends on the capability of the hardware to make a near-focus at the interaction point. An increase in the luminosity often requires a redesign of the machine optics in the interaction point and a replacement of the final-focusing quadrupole magnets. All these actions are essentially performed during long shutdowns (LS).

In experimental particle physics, we often use integrated luminosity which is the integral of the delivered luminosity over time. It is a measurement of the collected data size, and it is an important value that characterizes the performance of an accelerator. It has the units of $[L]^{-2}$. Fig. 3.4 represents the delivered luminosity versus time for 2010-2012, 2015-2018, and 2022 proton-proton collisions data.

The accelerator was designed to envisage instantaneous luminosity of $L_{\text{nominal}} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$. But the LHC exceeded all expectations, and that luminosity goal was not only reached but surpassed. The machine is now able to

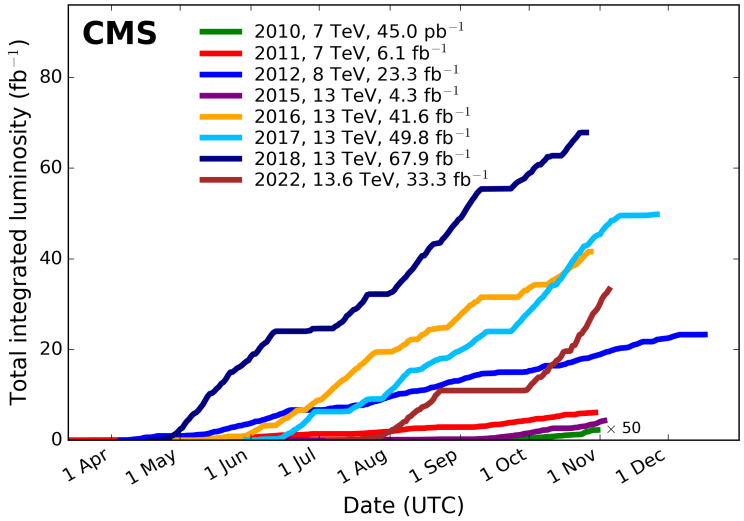


Figure 3.4. | Cumulative luminosity delivered by the LHC to CMS during stable pp-collision beams shown for different data-taking years [81].

deliver luminosities up to twice this nominal value. Since the rate of collisions is proportional to the luminosity while the frequency of beam crossings (f) is constant, the number of collisions per beam crossing will vary with luminosity, we refer to this quantity as "pile-up". While physics analyses gain in sensitivity when integrated luminosity is increased, an increased pile-up leads to more difficult experimental conditions that can decrease the reconstruction performance. The pile-up distribution recorded by the CMS experiment is shown in Fig. 3.5.

3.2. The CMS experiment

The CMS detector operates at the LHC at CERN, located in a cavern 100 m underground at Cessy in France, just across the border from Geneva. Its size is 15 meters high and 21 meters long and for all the detector material it contains, the detector is impressively compact. Muons are expected to be produced in the decay of a number of potential new particles and therefore the CMS detector is designed to detect them very accurately.

The solenoid magnet, giving CMS its name, is shaped from a cylindrical coil of superconducting fibers. The circulation of current through these coils generates a magnetic field of 3.8 TThis field is confined within the detector volume by a steel "yoke", constituting the bulk of the detector's mass. The magnet and its return yoke collectively weigh 12,500 tonnes, making it the heaviest component

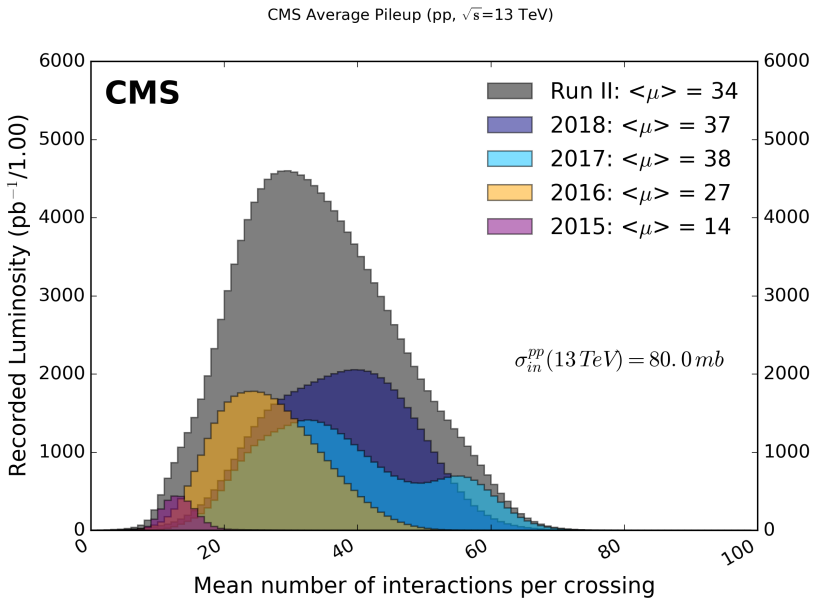


Figure 3.5. | Distribution of the average number of interactions per crossing (pile-up) for pp collisions for 2015-2018 and full run 2. The overall mean values and the minimum bias cross-section at 13 TeV are also shown. These plots use only data that passed the "golden" certification (i.e. all CMS sub-detectors were flagged to be ok for any kind of usage in physics analysis), and the "LHC standard" values for the minimum bias cross sections [82].

of CMS.

Knowing the effects caused by different types of particles as they pass through different materials, the detector is built and designed to tell us which particles were produced in a particular collision, their energy, and their momentum. Therefore, before moving on to the detailed apparatus of the CMS detector, we will first describe the basic approach to measure the outcome of particle collisions. More detailed descriptions will be given in Sec. 3.3.

The knowledge of particle properties such as energy, momentum, interactions, mass, lifetime, decay channels, charge, and magnetic field interactions guides the design choices and construction of both ATLAS and CMS, detectors enabling accurate particle detection, identification, and reconstruction. The inner regions of the detector measure the momentum of charged particles by measuring the tracks that the particles leave behind in various sub-detectors. The outer regions of the detector are designed to measure the energy deposit of particles released as they stop in different layers of the detector. The application of a strong magnetic field bends the trajectory of charged particles, which in addition

serves two purposes; positively and negatively charged particles under the same magnetic field, bend in opposite directions. Also, high-momentum particles bend less compared with low-momentum ones, therefore by knowing the radius of curvature of a certain track we can deduce the particle momentum.

The electrons, positrons, photons, jets, and muons, each of these particles make a different journey throughout the CMS detector and their unique interactions with the sub-detectors material make it possible to tell these particles apart. A photon, because it is electrically neutral will not leave any imprint in the tracker [1]. Instead, it will be stopped in the electromagnetic calorimeter (ECAL). On the other hand, electrons or positrons are stopped in the ECAL, leaving a track behind in the tracking system. Hadrons punch through further and are generally stopped by the hadron calorimeter (HCAL) where jets are confined and only the highest-energy hadrons and muons pass through the superconducting solenoid into the outer regions of the CMS barrel. Quarks and gluons will initiate a jet that leaves a bunch of tracks close together in the tracker, before depositing most of their energy in the hadronic calorimeter. Muons will make their way through the whole detector up to the muon chambers.

It is also possible to deduce tau lepton signatures as they decay promptly to lighter leptons, neutrinos, and/or hadrons. They leave characteristic patterns that one may search for specifically in the detector. Neutrinos, meanwhile, are weakly interacting and escape the CMS detector completely unseen. Nevertheless, we are able to infer their presence from an imbalance of energy in the reconstructed event called "missing transverse energy". Fig. 3.6 is a good illustration of the journey of different particles through the CMS sub-detectors.

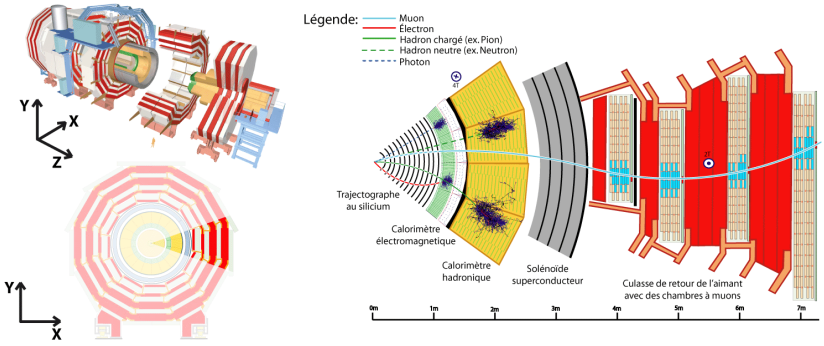


Figure 3.6. | A summary of the particle detection process in CMS [83].

[1] except if it undergoes a photon conversion within the tracker material

3.2.1. CMS coordinate system

Before going into the details of the CMS experiment, let us describe its coordinate system that we will often refer to in this thesis. In a right-handed Cartesian system, the $z-$axis of the CMS coordinate system is always aligned with the beam and directed anti-clockwise, the $y-$axis is in the upward direction, and the $x-$axis points towards the center of the LHC. The plane formed by the x and y directions is referred to as the transverse plane as shown in Fig. 3.7

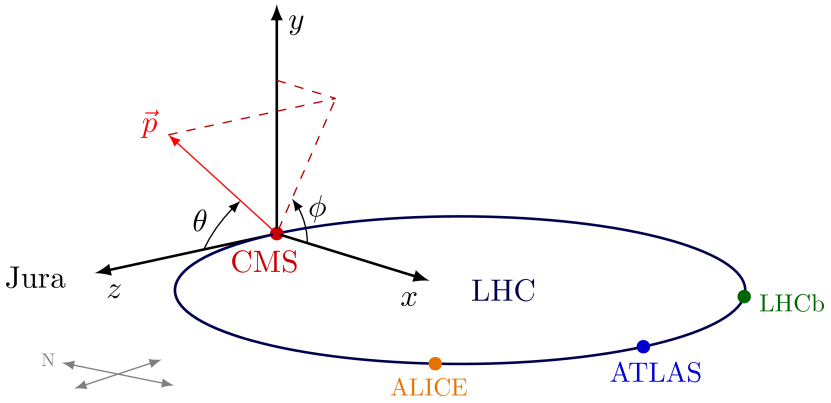


Figure 3.7. | CMS coordinate system definition [84].

In the spherical coordinate system which is most commonly used, a point is defined through its distance to the center of the experiment, its azimuthal angle, ϕ, in the transverse plane with respect to the x-axis, and its angle, θ, in the $z-y$ plane with respect to the z-axis.

We use the pseudorapidity term, $\eta = -ln(\frac{\theta}{2})$, as a spatial coordinate to describe the angle of a particle relative to the beam axis. An angle of zero lies along the beam axis and so particles with high pseudorapidity values ($|\eta| > 2.5$) are usually lost as they are out of detector acceptance.

3.2.2. Tracking system

The tracking system is one of the most ambitious parts of the CMS detector, which required the development of production methods and quality control procedures that were new to the field of particle physics detectors for a few different reasons. Since the tracker is located closest to the interaction point, it also receives the highest dose of radiation damage. Although located in an experimentally difficult region characterized by high-track multiplicity and heavy irradiation, the tracker needed to be reliable for accurate high-resolution trajectory reconstruction. The construction materials were therefore carefully chosen to resist radiation. This part of the detector is also built with fine segmentation accompanied by high power density of the readout electronics

while keeping the amount of material to a minimum, in order to limit multiple scattering, *bremsstrahlung*, photon conversion, and nuclear interactions.

These considerations led to a design entirely based on silicon detector technology due to several reasons:

- **Precision:** Silicon detectors offer high spatial resolution, meaning they can pinpoint the exact location where a charged particle passed through the material. This precision is crucial for accurately reconstructing the paths of particles.
- **Speed:** Silicon detectors can quickly convert the energy deposited by a particle into an electrical signal, allowing for rapid data acquisition.
- **Minimal Noise:** Silicon detectors exhibit low electronic noise levels, which helps in detecting even faint signals from particles.
- **Durability:** Silicon is a robust and durable material that can withstand the harsh conditions of particle collisions and the high radiation environment of a particle accelerator.
- **Modularity:** Silicon detectors can be arranged in layers or modules, allowing for efficient tracking of particle trajectories.

The pixels, at the very core of the detector and the silicon microstrip detectors that surround it which are divided into four parts; the inner barrel (TIB), the inner disks (TID), the outer barrel (TOB), and outer end-caps (TEC). The layout of the tracker substructures is sketched in Fig. 3.8.

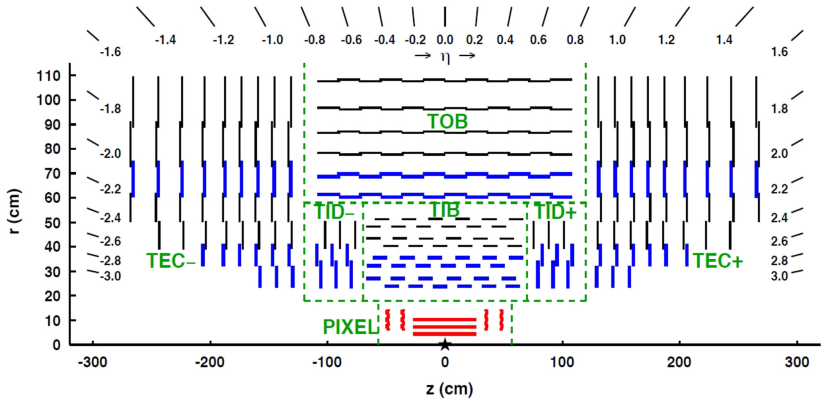


Figure 3.8. | A view of a tracker slice in the $r-z$ plane. Pixel modules are shown in red, single-sided strip modules are depicted as black thin lines, and strip stereo modules are shown in blue thick lines [85].

All parts are built out of a series of distinct modules, which registers the passage of a charged particle through its medium. Measuring a series of hits along the track of a moving particle, we can reconstruct the direction and curvature of the

track and extrapolate a point of origin (i.e. vertex) by finding several trajectories starting from a single point. Importantly, thanks to the high granularity of the modules near the beam pipe, it is also possible to deduce whether particles were genuinely produced at the interaction point (indicating a very prompt decay of some parent particle), or whether particles traveled a short distance before decaying. The trajectories of charged particles are bent in a direction perpendicular to the magnetic field. The direction of bending depends on the charge of the particle, while the amount of bending is a function of particle momentum - this is a central feature of particle identification in modern particle physics detectors.

As particles travel through the tracker modules, electron-hole pairs are created in the detector's bulk material (the details are discussed in Sec. 3.2.2.2) drifting along the field lines resulting in a signal on the corresponding strips. The signals are stored in chip memory for several microseconds and then processed before being sent to a laser to be converted into infrared pulses. These are then transmitted over a 100m fiber optic cable for analysis. The tracker uses 40,000 such fiber optic links providing a low-power, lightweight way of transporting the signal.

3.2.2.1. The CMS tracker

The tracker allows for rapid and precise measurements with temporal and spatial resolutions that fulfill the challenges posed by the high luminosity LHC collisions. It is designed to reconstruct high-p_{T} muons, isolated electrons, and hadrons with high momentum resolution and efficiency greater than 98% in the range $|\eta| < 2.5$. The transverse momentum resolution of a single muon as a function of the parameters of the tracking detector can be written;

$$\sigma(p_{\mathrm{T}}) = \frac{\sigma_{hit}}{0.3BL^2} \cdot \sqrt{\frac{720}{N_{hit}+4}}, \tag{3.4.}$$

where B is the magnetic field strength, L represents the radial extension of the tracker, σ_{hit} is the spatial resolution of a single sensor and N_{hit} corresponds to the number of measured hits for a given track. From Eq 3.4 we can conclude that a larger tracker diameter, stronger magnetic field, or finer granularity results in a better track-momentum resolution. Nevertheless, the formula does not account for the amounts of material that can be placed. In order to optimize the tracking and vertexing resolution, material interactions must be kept to a minimum since large amounts of material lead usually to significant energy losses, multiple scattering of electrons, and photon conversions. The material budget of the CMS tracking detector in a unit of radiation length [1] X_0 varies

[1] mean distance traveled by an electron before its energy is reduced by a factor

from $0.4X_0$ in the very central region to $1.8X_0$ at the pseudorapidity regions that cross both the barrel and the endcap detectors, as shown in Fig. 3.9

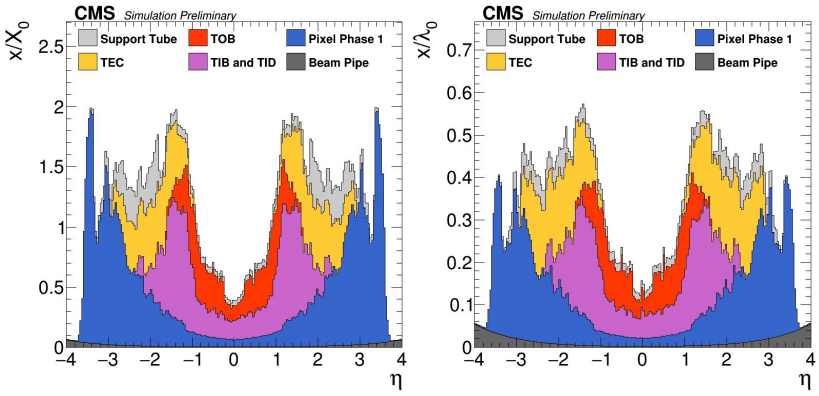


Figure 3.9. | Material budget in unit of radiation lengths (X/X_0, right) of and hadronic interaction length (X/λ_0, left) as a function of the pseudorapidity η. The largest values for radiation length and hadronic interaction length lie outside the tracking acceptance at around $|\eta| = 3.5$ and amount to 1.9 X/X_0 and 0.36 X/λ_0, respectively [86].

3.2.2.2. Tracker layout

The CMS tracker has a length of 5.8 m and a diameter of 2.5 m and consists of 25 684 silicon sensors, which measure the hits from the tracks crossing through them. Two types of sensors are used in the CMS tracker:

- ***Pixel sensors:*** consist of a two-dimensional (2D) array of pixel cells providing 2D hit-position measurements on the surface of the sensor.
- ***Microstrip sensors:*** consist of a set of parallel or quasi-parallel strips providing only 1D hit-position measurements (in a single direction).

A general layout of the tracker with all the subdetectors is shown in Fig. 3.10.

■ The pixel detector

The CMS pixel detector consists of two subdetectors; the barrel pixel detector (BPIX) and the forward pixel detector (FPIX). It is optimized to have four-hit coverage over the pseudorapidity range $|\eta| < 2.5$, accurate pattern recognition and track reconstruction, and added redundancy to cope with hit losses. A schematic view of the pixel-detector sensor layout is shown in Fig. 3.11

The original BPIX is 57 cm long and consists of 3 layers of pixel sensors located at mean radii of 4.4, 7.3, and 10.2 cm. Within the same layer, the modules

$\frac{1}{e}$(characterizes depth of electromagnetic showers)

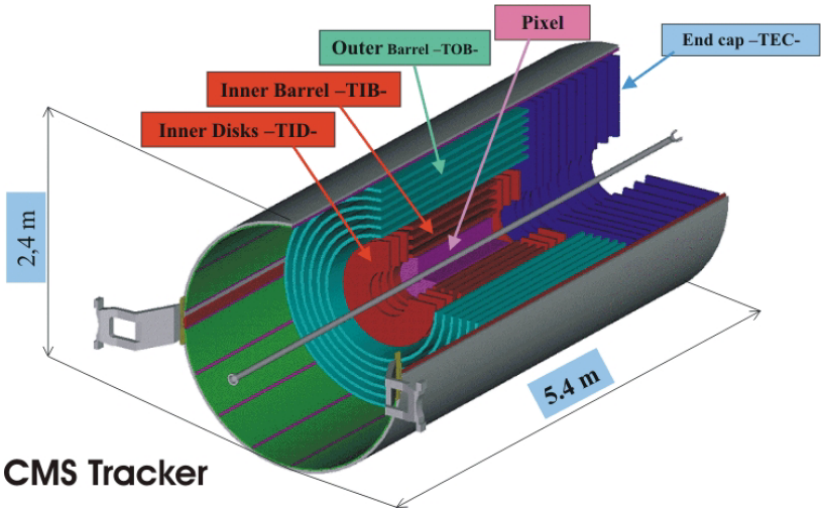


Figure 3.10. | Schematic view of the CMS tracker detector [87].

located at the same ϕ have a common orientation of their sensitive layer, while the following rows of modules in ϕ have an opposite orientation. Thus, a drift direction of charge carriers alternates in adjacent ladders of BPIX modules along global ϕ.

The original FPIX extends from 6 to 15 cm in radius and consists of 2 layers located at $\mid z \mid = \pm 34.5$ cm and $\mid z \mid = \pm 46.5$ cm. The FPIX disk sensors are arranged in blades, which are angled at 20° in a turbine-like geometry to induce charge sharing of particles entering the detector at an average angle of 20° from normal incidence. Tilting of FPIX sensors causes Lorentz drift of charge carriers, which further improves charge sharing between neighboring pixels, leading to a nominal impact position resolution of about 15 μm. This arrangement of 3 barrel layers and 2 forward disks on each z-side guarantees 3 track hits in the pixel tracker over almost the entire pseudorapidity range.

The original pixel was designed for a maximum instantaneous luminosity of $1 \times 10^{34} \mathrm{cm}^{-2} \mathrm{s}^{-1}$ and a maximum average pileup (number of inelastic interactions per bunch crossing) of 25 in LHC operation with 25 ns bunch spacing. With the upgrade of the LHC accelerators during the first long shutdown LS1 in 2013-2014, these parameters have been exceeded compared to the design values. Therefore the original pixel detector has been replaced by a new system, referred to as the CMS Phase-1 pixel detector in order to improve the performance toward higher rate capability, radiation tolerance, and more robust tracking. The installation of the CMS Phase-1 pixel detector took place during the extended year-end technical stop of the LHC in 2016/2017. The CMS Phase-1 pixel detector consists of four concentric barrel layers (L1-L4) at radii of 29, 68, 109, and

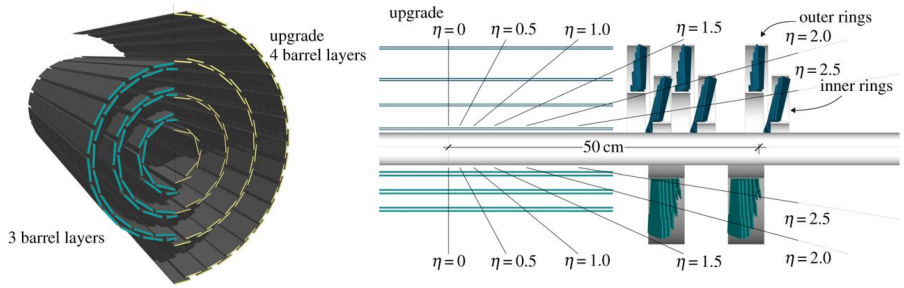


Figure 3.11. Layout of the pixel tracker the lower half shows the original pixel detector, and the upper half represents the newly-installed detector. The barrel part (BPIX) is colored blue while the forward disks (FPIX) are marked with magenta color. The sensors in BPIX are aligned along the global $z-$axis, while in FPIX sensors are tilted at 20° from the plane perpendicular to the $z-$axis [88].

160 mm, and three disks (D1-D3) on each end at distances of 291, 396, and 516 mm from the center of the detector. The upgraded system is able to cope with significantly higher data rates because of the new readout chip design and the increased bandwidth of the digital data transmission via optical links. The CMS Phase-1 pixel detector has been successfully taking data since 2017 and, together with the other CMS subsystems, has delivered high-quality physics data corresponding to an integrated luminosity of more than 138 fb^{-1}.

The silicon strip tracker

The strip tracker surrounds the pixel detector and consists of microstrip sensors that, unlike the pixel provide 1D measurements in a single direction. The sensors are arranged into 4 distinct subdetectors: TIB, TOB, TID, and TEC. Their descriptions were mentioned in Sec. 3.2.2.

The silicon sensors in the strip detector are arranged in cylindrical layers along r in the TIB and TOB, while in the TID and TEC layers, the sensors are arranged in separate disks along z. The TIB and TOB modules located near the $z-$axis have alternate orientations. Unlike FPIX, whose disks are segmented only along ϕ, microstrip endcaps cover a much larger area and therefore have additional segmentation along r into rings. The sensor has different geometries and thicknesses and sensors in the strip tracker are significantly larger than those of the pixel because all regions have to be efficiently equipped.

Since the hit position in microstrip sensors is only measured in one direction perpendicular to the strips, the hit position along the strip is unknown. Therefore, the inner and outer barrel strip modules extend along the global z-axis and provide an accurate measurement of the angle ϕ of the hit as well as its radius, these modules are referred to as $r\phi$ modules. Four layers of the barrel and

three rings of the endcaps (shown in blue in Fig. 3.10) are populated by "stereo modules", mounted back to back and inclined around the normal by a stereo angle of 100 mrad. A schematic overview of the sensor geometries used in the strip tracker is shown in Fig. 3.12.

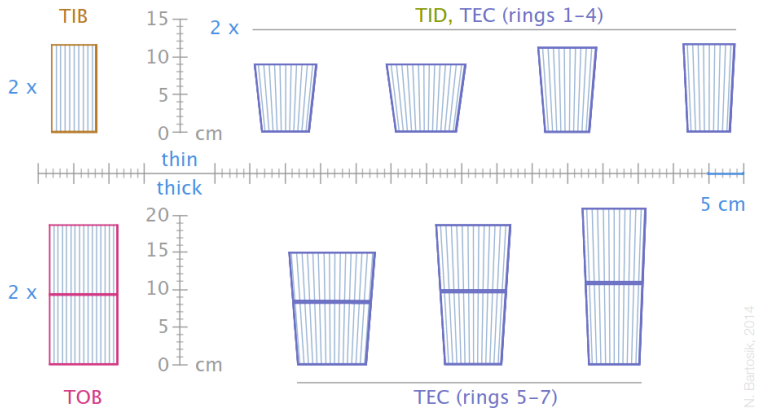


Figure 3.12. | Geometries of the microstrip sensors used in the CMS tracker. TIB and TOB sensors exist in two types, with the same surface but with different pitches. The inner TID and TEC ring sensors come in two different versions [89].

■ Operation of silicon modules

I couldn't finish this section without saying a few words about the physics behind the silicon technology used in the CMS tracker. The sensors are composed of an n-type doped wafer on which the p-type doped pixels/strips are applied, forming a pn junction. Near the boundary between p-type and n-type materials, weakly bound electrons from the donors diffuse into the p-type material and can combine with the holes of the acceptors, while the holes diffuse from the p-type towards the n-type. Therefore, a certain region is depleted of free charge carriers and an electric field builds up between the p-type and n-type material. This electric field is amplified by a bias voltage of the same polarity as the initial potential. The voltage required (V_{depl}) to deplete the entire silicon sensor of thickness d is given by:

$$V_{depl} = \frac{ed^2}{2\epsilon_{Si}\epsilon_0} \mid N_{\text{eff}} \mid, \tag{3.5.}$$

where $\mid N_{\text{eff}} \mid$ is the effective doping concentration, ϵ_{Si} and ϵ_0 represent the dielectric constant of silicon and vacuum respectively, and e is the electron charge. When a charged particle enters the depleted region, it causes ionization which produces electron-hole pairs. Under the applied bias voltage, the charge

carriers drift through the sensor material leading to a small measurable electrical current pulse which is detected by the corresponding readout chip. The principle of operation of a silicon microstrip sensor is illustrated in Fig. 3.13.

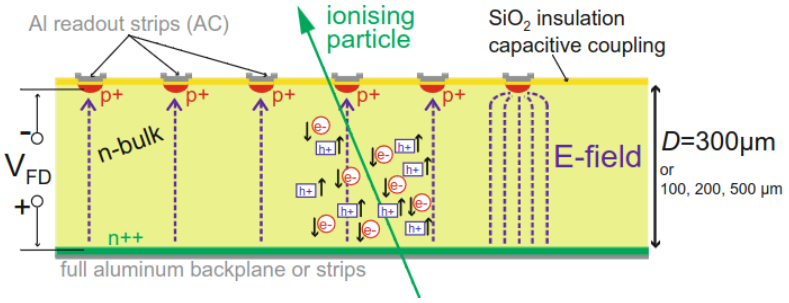


Figure 3.13. | Principle of operation of an AC p-in-n coupled silicon microstrip sensor. The electron-hole pairs, resulting from the ionization of the crossed-charged particle, move towards the electrodes on the planes of the sensor guided by the electric field. The segmentation in the pn junctions allows the charges to be collected on small individual strips, where they capacitively couple to the Al readout strips. These are then connected to the readout electronics, where the intrinsic signal is brought into shape and amplified [90].

3.2.2.3. Hit and track reconstruction

An essential step in track reconstruction involves effectively identifying hits, which indicate the locations in the tracker sensors traversed by charged particles. The intricate process of reconstructing trajectories for charged particles within the CMS tracker is magnified by the substantial particle density. I have contributed to a small part of the software development of this particular phase, thus dedicating this section to elaborate on my contributions. The reconstruction is based on two consecutive steps:

1. Local reconstruction:

Localized hits are constructed using the digitized detector readout of the raw signal from the silicon sensors.

■ **Cluster Parameter Estimator (CPE)**

The local reconstruction phase is also a two-stage process involving clustering and hit conversion. During the clustering step, signals from neighboring channels, those fulfilling certain threshold criteria (specific to signal-to-noise ratio for the individual channels as well as for the whole cluster, etc.) are grouped into clusters. Following the hit conversion step, the clusters are translated into hypothetical hit measurements using a cluster parameter estimator (CPE). To each hit, a position and a corresponding uncertainty are assigned in the local coordinate

frame of the silicon module. The calculated hit coordinates with uncertainties, as well as the charge and the profile of the clusters, form the final output of the local reconstruction. This step is one of the main subjects of the software development contribution to this thesis.

Lorentz angle effect

Not all silicon sensors in the tracker happened their electric field to be aligned to the magnetic field from the solenoid. A combination of the two non-parallel fields induces a Lorentz force on the charge carriers drifting towards the readout channels and the trajectory of the drifting charge carrier is deflected by a certain Lorentz angle, θ_{LA}, which affects the charge sharing between multiple channels as illustrated in Fig. 3.14. In the reconstruction step, this effect must be corrected and is one of the main corrections involved in the development of CPE measurements.

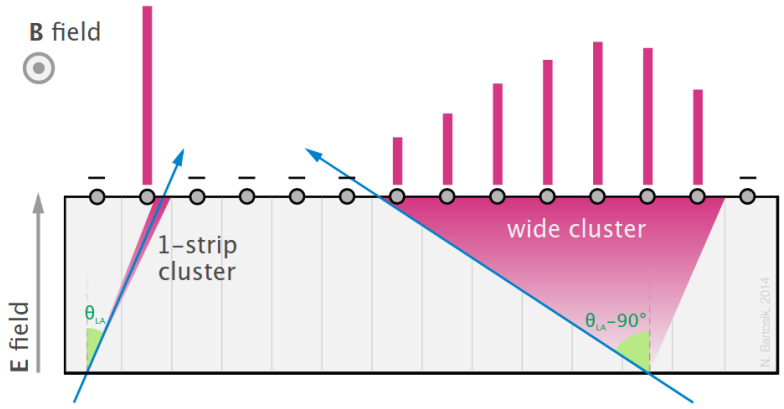


Figure 3.14. | Illustration of a drift of charge carriers in a silicon sensor under the Lorentz force. Assuming the sensor is a barrel strip module with the electric field in the sensor perpendicular to the magnetic field. The black dots mark the strips, and the magenta bars above represent the induced signal on the channel. The charge carrier drift direction is deflected from normal by a Lorentz angle θ_{LA}. A minimal dispersion occurs for small θ_{LA} (left) while larger angles cause a maximum distribution of the read-out signal across multiple strips [89].

Backplane corrections

Although the CMS detector was primarily designed to measure particles originating from proton-proton collisions, it can detect any particles traversing it. For example, high-energetic cosmic muons can penetrate the ground and reach the CMS detector. While such particles are treated as background during hadron collisions and are rejected, they are measured and recorded for alignment purposes when there is no beam in the accelerator. Two different signal

readout schemes are applied to the tracker modules in order to optimize their performance under different conditions:

- ***peak mode:*** is mainly used in cosmic data taking and it is a long readout time that allows measuring the full signal shape from the module.
- ***deconvolution mode:*** is a short readout time limited by the bunch crossing rate. Only the turn-on curve of the signal is measured at three points, with a re-shaping of the analogue pulse shape to one that peaks at 25 ns.

Peak readout aims to capture the full signal shape for accurate energy measurement, while short readout modes might prioritize reducing data size by only recording a subset of information. The choice of readout mode can influence the accuracy of track and hit reconstruction. During hadron collisions, the tracker is more suitable to operate in a deconvolution mode given the 25 ns between two bunch crossings. Given that not all charge carriers reach the channels during the short readout time, the effective depletion area in the silicon sensor becomes thinner, leading to a drastic change in the cluster shape, as shown in Fig. 3.15.

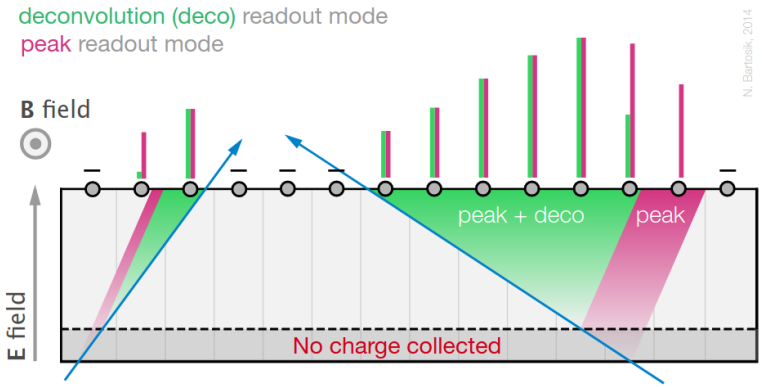


Figure 3.15. Ilustration of the backplane effect in a silicon sensor. In the peak readout mode, all charge from the sensor backplane is collected. The deconvolution short readout mode is a reduction of the cluster size from one side result, which depends on the track incident angle [89].

This effect leads to a shift of the measured hit position depending on the track incident angle, and module location. While the effect on the thin modules used in TIB accounts for only a few microns, the thick modules used in the outermost four layers of the TOB can show a significant difference of about 20 μm. Fig. 3.16 shows an ideal case of a charged particle traversing the silicon sensor where all charges drift to the strip plane of the sensor. The cluster barycentre is extrapolated to the mid-plane of the sensor as the true reconstructed hit position.

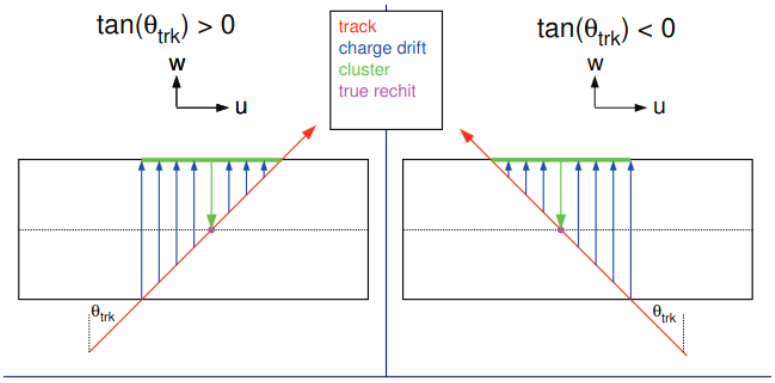


Figure 3.16. | Illustration of an ideal case where all charges drift to the sensor plane are readout, resulting in a perfect match between the reconstructed and measured hit position [91].

In the deconvolution mode, which represents a more realistic scenario, the readout duration is shorter, causing not all of the charge to be captured. Consequently, the derived hit position from the cluster's barycentre is incorrect. However, alignment steps are capable of compensating for this effect by shifting the module in the w direction. This correction ensures that the reconstructed hit aligns once again with the track's trajectory, as illustrated in Fig.3.17. This correction constitutes an important part that had to be taken into account in the development of the CPE software, discussed in Sec.3.2.2.3 of this thesis.

2. Global reconstruction:

This phase includes a search for possible combinations of hits that make up continuous tracks corresponding to potential trajectories from charged particles that are bent by the magnetic field. When a track candidate is found, track fitting, ambiguity resolution, and smoothing are followed. Global reconstruction uses the output of the local reconstruction, as well as important information about the modules and it is based on three successive steps: the seed finding, the pattern recognition, and the final fit, more details on the full tracking procedure can be found in [92].

3.2.3. Calorimeters

The CMS calorimeter consists of two distinct subdetectors: electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). In general, calorimeters are characterized by the radiation length (ECAL) and nuclear interaction length (HCAL) of their active material. ECALs tend to be 15-30 radiation lengths deep while HCALs are 5-8 nuclear interaction lengths deep. The ECAL covers the pseudorapidity region $|\eta| < 3$ of the energy deposit of electromagnetic interacting

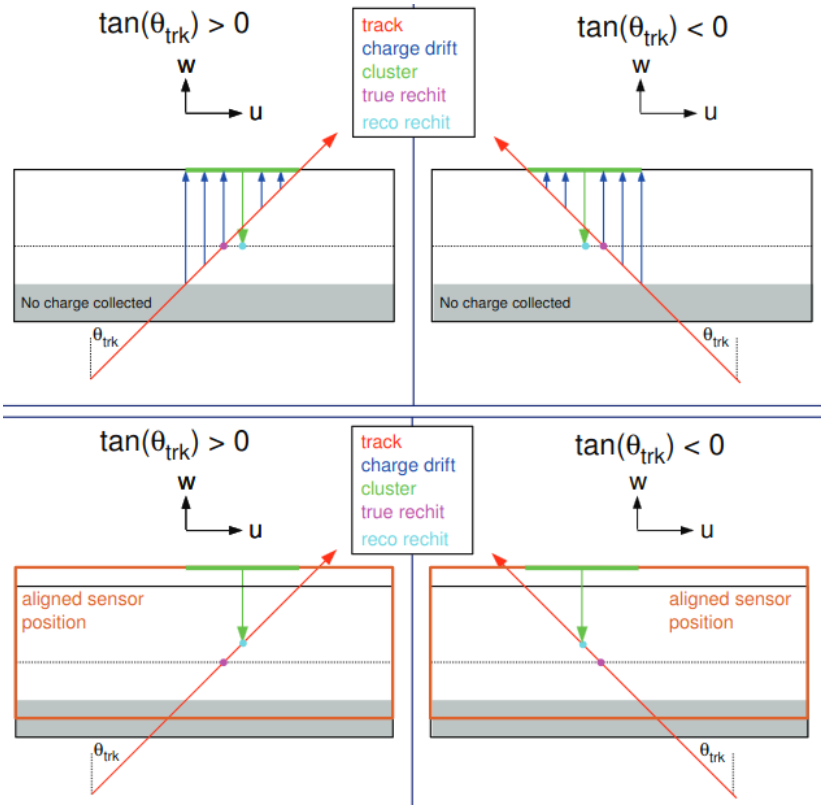


Figure 3.17. | Illustration of the effective collected charge in deconvolution mode which results in a bias of the reconstructed hit position extrapolated from the clusters barycentre [91].

particles such as leptons and photons. Hadrons on the other hand produce hadronic showers via the strong nuclear force that are broader and deeper compared to electromagnetic ones. Their energy depositions are detected in the HCAL which has a larger volume that surrounds the ECAL. The HCAL covers the same pseudorapidity region, while an additional forward calorimeter (HF) extends it to $|\eta| < 5$. In order to provide a reliable measurement of missing transverse energy.

3.2.3.1. Electromagnetic calorimeters

The electromagnetic calorimeter (ECAL), is composed of a barrel (EB) covering the pseudorapidity region below $|\eta| = 1.479$ and two endcaps (EE) reaching up to $|\eta| = 3$. Lead-tungstate ($PbWO_4$) was chosen as the sensitive material for this detector due to its high density (8.28 g cm^{-1}), short radiation length

(0.89 cm) and small Molière radius[1] (2.2 cm), making it ideal for a compact and high-granularity detector.

A pre-shower detector (ES), based on a lead absorber and silicon strip sensors, placed in front of the endcaps at $1.65 < |\eta| < 2.6$, improves the photon-π^0 separation. The total thickness of the ES is ~ 3 radiation lengths.

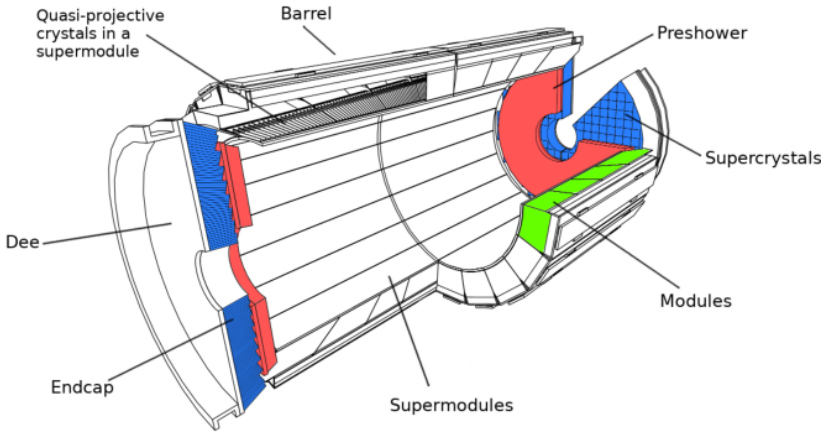


Figure 3.18. | Sketch of the electromagnetic calorimeter [11].

Crystals are projective and positioned slightly off-pointing (~3°) relative to the interaction point to avoid gaps aligned with particle trajectories. The calorimeter has no longitudinal segmentation, and the measurement of the photon angle relies on the primary vertex reconstruction from the silicon tracker. The crystal length in EB is 230 mm (220 mm in EE) corresponding to ~26 (25) radiation lengths. The transverse size of the crystals at the front face is 2.2×2.2 cm^2 in EB (2.86×2.86 cm^2 in EE). The total crystal volume is 11 m^3 and the weight is 92 tons. The barrel calorimeter is organized into 36 supermodules each containing 1,700 crystals while the endcaps consist of two dees, with 3,662 crystals each.

The energy resolution of the ECAL can be parametrized as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{(E)}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{3.6.}$$

where S is a stochastic term, N noise, C is a constant term, and E is the energy in GeV. The values of these terms are obtained from fitted resolution function curve $f(\frac{\sigma}{E})$. The stochastic term, S, is of the order of 2.8% which includes the contributions of the shower confinement, the number of photoelectrons, and

[1] radius of a cylinder that on average contains 90% of the electromagnetic shower's energy deposition (characterizes the width of the shower)

the fluctuations of the gain process. The noise term, N, at 40 MeV becomes non-negligible in the order of 12% which corresponds to a single-channel noise giving 120 MeV in a 3×3 crystal matrix. The constant term, C, (less than 0.3%) dominates the energy resolution for high-energy electron and photon showers and depends on the non-uniformity of longitudinal light collection, energy leakage behind the calorimeter, uniformity, and the stability of the single-channel response.

3.2.3.2. Hadronic calorimeters

The hadronic calorimeter (HCAL) is an essential part of the CMS detector for the measurement of hadronic jets and missing transverse energy $p_{\mathrm{T}}^{\mathrm{miss}}$. The HCAL consists of four subdetectors: barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters. The schematic view of the arrangement of the HCAL subdetector is shown in Fig. 3.19.

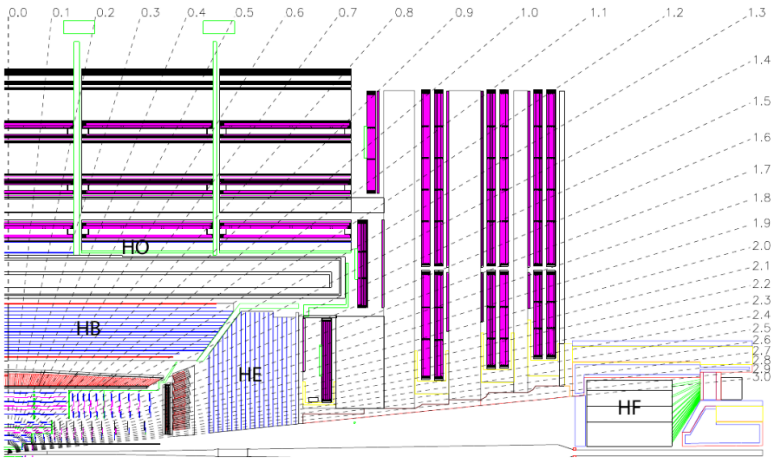


Figure 3.19. | Longitudinal view of the CMS detector showing the HCAL subdetectors [11].

The HB is located between radii of 1775 and 2876.5 mm and covers $|\eta| < 1.39$. The HB is divided into two half-barrels in the direction along the beam (z), each assembled from 18 wedges. Each wedge contains absorber plates made of brass (an alloy with 70% copper and 30% zinc) and subtends 20° in ϕ and extends to 4330 mm from the CMS detector mid-plane. The HB has about 40,000 scintillator tiles. In order to limit the number of individual physical elements, the tiles at the same ϕ and depth are grouped into a single scintillator unit, referred to as a megatile. In 2019, the HB detector moved to phase 1 upgrade to enable higher gain and better photon detection efficiency by replacing the hybrid photodiode transducers (HPD) with silicon photomultiplier devices (SiPM). SiPMs provide

powerful electrical signals that reduce the effect of electrical noise and enable signal splitting for TDC (time-to-digital converter) measurements, reflecting an improvement in calorimeter calibration and monitoring capability.

The HE calorimeter is also made of brass absorber plates with sampling layers of plastic scintillator. The HPDs in HE was replaced by SiPMs during the 2016–17 extended year-end technical stop. The innermost surface of HE is located 4006.5 mm from the interaction point and covers $1.3 < |\eta| < 3.0$. Each endcap has an 18-fold symmetry in ϕ and has 34.5 mm thick sector plates each covering 20° in ϕ. The sector layers are separated by 9 mm thick brass spacers to allow space for the scintillator inserts. Each scintillator insert covers 10° in ϕ. The top edge of the front part of the endcap module has a slope of 53° corresponding to the gap angle between the HB and HE calorimeters.

For the HB and HE calorimeters, clear fibres carry the light to hybrid photodiodes (HPDs), and each HPD signal is digitized in 25 ns time intervals by a charge integrator and encoder (QIE). For the HO calorimeter, light is carried to silicon photomultipliers (SiPMs) and the SiPM signals are digitized by the QIEs.

The HO calorimeter covers $|\eta| < 3$ and consists of one or two layers of scintillator outside the magnet coil. The entire assembly is divided into 5 rings, each having 12 sectors in ϕ. Six trays of scintillators are assembled on a honeycomb structure, which is then mounted in each of these sectors. The central ring0 has two layers of 10 mm thick scintillator on either side of a stainless steel block at radial distances of 3850 and 4097 mm, respectively. All other rings have a single layer at a radial distance of 4097 mm. The $\eta - \phi$ segmentation of the HO calorimeter matches closely that of the HB calorimeter.

The front faces of the HF calorimeters are located 11150 mm away from the interaction point on either side of the CMS detector and cover $2.85 < |\eta| < 5.19$. The detectors covering positive and negative η ranges are referred to as HF+ and HF-. The inner and outer radii of the HF calorimeter are 125 and 1570 mm, respectively. Each HF module is composed of 18 wedges made of steel with quartz fibers embedded along its length. The detection technique utilizes the emission of Cherenkov light by secondary charged particles going through the quartz fibers. The light collected from an HF calorimeter quartz fiber is converted to charge by a photomultiplier tube and digitized by the QIE.

3.2.4. The Muons system

Unlike most particles, muons can travel several meters of material before their decay. The CMS detector is designed so that these particles are not stopped by one of the CMS calorimeters but rather tracked to the outer part of the experiment in the muon chambers, where they are the only particles likely to produce a clear signal. Although the CMS detector was primarily designed to

measure muons originating from proton-proton collisions, high-energetic cosmic muons can penetrate the ground and reach the CMS detector. While such particles are treated as background during hadron collisions and are rejected, they are measured and recorded for alignment purposes when there is no beam in the accelerator.

The muon system is composed of three types of gas ionization chambers which are located outside of the magnet coil, interleaved with iron "return yoke"[1] plates located outside the solenoid.

- ***Cathode Strip Chambers (CSCs):*** are trapezoidal multi-wire proportional chambers with six layers and the gaps are filled with a gas mixture of $Ar(40\%)$, $CO_2(50\%)$ and $CF_4(10\%)$. A muon passing through the gap causes ionization of the gas molecules which results in an avalanche of electrons around the wires. Positive ions move towards the copper cathode, further inducing a charge pulse in the strips. Because the strips and the wires are perpendicular, we get two position coordinates (ϕ, r) for each passing particle.
  The CSCs makes a fast detector suitable for triggering due to the closely spaced wires, precise timing, and spatial information. Moreover, as each CSCs module contains six layers, the muons are precisely identified and their trajectories are well matched to those reproduced in the tracker. The CSCs are arranged in a series of concentric rings centered on the beam line. $\mathbf{ME1/1}$[2] chambers, of all the CSCs, cover the most forward region ($1.6 \leq |\eta| \leq 2.4$) and hence, they are exposed to the highest rate of radiation.
- ***Muon Drift Tubes (DTs):*** The cross-sectional size of the drift cell, which refers to the maximum distance and time taken for drift, was selected to be 21 mm (equivalent to a drift time of 380 ns in a gas mixture comprising 85% Ar and 15% CO2). This value is sufficiently small to maintain a minimal occupancy rate and eliminate the necessity for multi-hit electronics. However, the size of the cell is adequately large to keep the count of active channels within reasonable limits. A tube was settled for the fundamental drift unit, chosen to provide protection against potential damage due to wire breakage. Furthermore, this choice serves to partially isolate adjacent cells from the electromagnetic debris associated with the muon itself. While the CSCs are used in the endcaps, DTs are used for precise trajectory measurements in the central barrel region. Each DT chamber consists of three "superlayers", each comprising four staggered layers of parallel drift cells. The wires in each

[1] The return yoke is a huge steel structure that provides the return route for the magnetic flux and also holds the muon detectors in four concentric layers. The magnetic field associated with the return flux is in the opposite direction to the magnetic field of the main magnet and hence the tracks of the muons in the outer field are also bent in the opposite direction.

[2] The CSCs are labeled according to their position within the CMS Muon Endcap, i.e. ME1/1 refers to the CSCs in the first ring in the first disk.

layer are oriented so that two of the superlayers measure the muon position in the bending plane ($R-\phi$) and one superlayer measures the position in the longitudinal plane ($R-\theta$).

- ***Resistive Plate Chambers (RPCs):*** are installed in both the barrel and the endcaps. They are fast gaseous detectors providing a muon trigger system parallel to those of the DTs and CSCs. The RPC operated in avalanche mode with common read-out strips and their modules consist of two parallel plates [1], a positively-charged anode and a negatively-charged cathode separated by 2 mm gaps of gas volume filled with a $C_2H_2F_4$,iC_4H_{10} and SF_6 mixture. When a muon passes through the chamber, electrons are expelled from the atoms of the gas. These electrons in turn collide with other atoms causing an avalanche of electrons. RPCs combine good spatial resolution with a temporal resolution of only a nanosecond (one billionth of a second).
  An RPC has the capacity to rapidly determine the timing of an ionizing event within a significantly shorter duration than the 25 ns interval between two consecutive LHC bunch crossings (BX). Consequently, an efficient muon trigger system, utilizing dedicated RPCs, can precisely associate a muon track with the pertinent BX even in the presence of the anticipated high rate and background at the LHC. In an environment where rates could reach 103 Hz/cm^2, an RPC-based trigger must not only assign the BX to potential tracks but also effectively estimate the transverse momenta with a high level of efficiency.

A detailed description of these chambers, including gas composition and operating voltage, can be found in [11].

3.2.5. The Trigger and data acquisition system

The trigger system consists of two levels designed to select events of potential physics of interest. The Level-1 (L1) trigger receives information from calorimeter and muon systems generating an initial selection, within a fixed latency of 4 μs the trigger has to provide an accept signal L1 Accept (L1A) deciding the fate for the selected event to be reconstructed in the High-Level Trigger (HLT) and our fate of finding new physics. The maximum output rate of the L1 trigger is 100 kHz, while HLT selection further reduces this rate to about 1 kHz on average.

[1] The parallel plate geometry allows for the creation of a uniform electric field between the plates. This electric field is crucial for the drift of ionized particles produced when a high-energy particle passes through the gas within the RPC. A uniform electric field ensures that the ionized particles experience a consistent acceleration and drift, leading to accurate position and timing measurements.

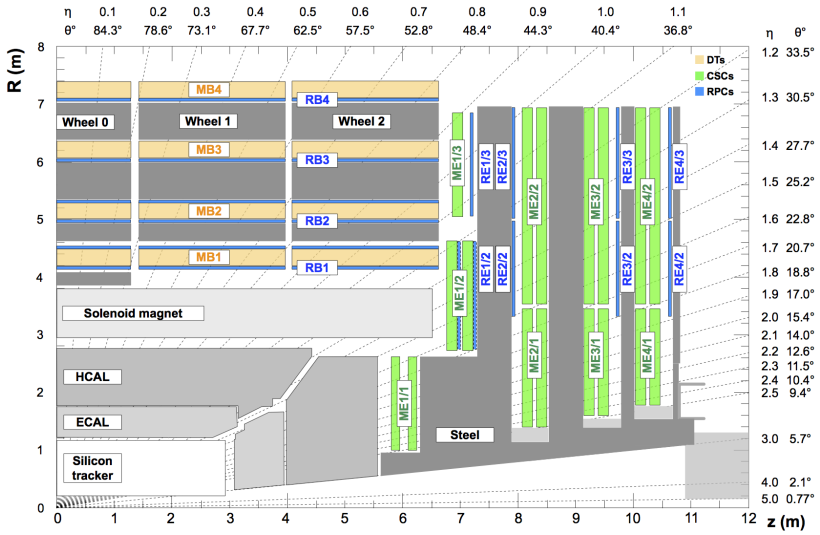


Figure 3.20. | An $r-z$ cross-section of a quadrant of the CMS detector with the axis parallel to the beam (z) showing the current muon system that includes Resistive Plate Chambers RPCs, Drift Tubes DTs, Cathode Strip Chambers CSCs, and the steel flux-return disks (dark areas) [93].

3.2.5.1. Level-1 trigger

The hardware-based Level-1 (L1) trigger checks for each clock cycle of the LHC's 40 MHz clock if the data contains any physics of interest that is worth being processed by the HLT. This rapid verification is based only on a subset of the readout data at a coarse resolution and mainly from muon systems and calorimeters. The ECAL and the HCAL data are combined and processed in the Regional Calorimeter Trigger (RCT) and the Global Calorimeter Trigger (GCT). Data from the three muon detector systems (RPCs, CSCs, and DTs) are combined in the Global Muon Trigger (GMT). All data are further combined in the Global Trigger (GT).

While all of these detector data are still available in the readout pipelines level-1 trigger, a decision needs to be taken in about 4 μs equivalent to 160 clock cycles of the LHC's 40 MHz clock. The Trigger Control System in the GT checks if all parts of the CMS detector are in a position to accept a trigger by using feedback from the Trigger Throttling System (TTS), and checks if some additional conditions are met. In case of a positive verdict, the level-1 trigger sends a Level 1 Acceptance (L1A) signal via the Trigger Timing and Control (TTC) system to read all parts of the CMS detector at full resolution. The architecture of the CMS L1 trigger is illustrated in Fig. 3.21.

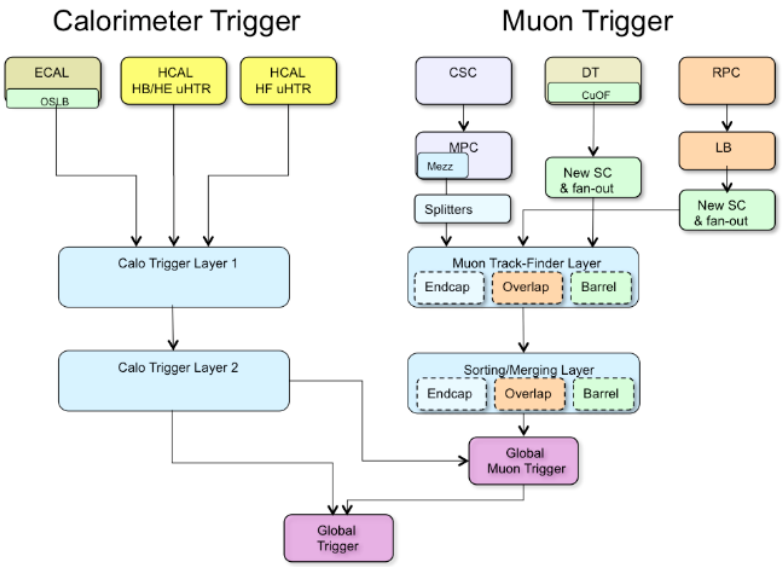


Figure 3.21. | Schematic representation of the CMS Level-1 trigger. Events are progressively processed through different algorithmic layers to provide a final trigger decision [94].

3.2.5.2. High-Level trigger and data acquisition

While collisions take place at a frequency rate of 40 MHz and a hundred million channels are read-out in the experiment, the final distributions included in published papers usually include only a few observables from a very small set of significant events. Once LHC full run 2 finished, much larger data sets had to be processed for physics analyses. Fig. 3.22 is a representation of the HLT rates used within each CMS physics group which clearly indicates the challenge of reducing that rate to only a few Hz for offline analysis, to comply with the performance limitations of the mass storage hardware also ensure that a given set of analyzes could always reach its ultimate sensitivity with a lot of flexibility. At the same time, the physics signals of interest must be retained with high efficiency.

The HLT consists of a farm of computers that runs the full reconstruction software framework of CMS called CMSSW. Unlike Level-1 trigger the HLT has privileged access to the full detector readout. It is designed with more than 400 different paths targeting a broad range of physics signatures, to keep events the most highly energetic, rare or otherwise interesting ones for permanent storage and offline analysis. It is purely software-based and must achieve the remaining rate reduction of Level-1 trigger by executing sophisticated offline-quality algorithms. HLT algorithm sequences are moreover efficiently scheduled:

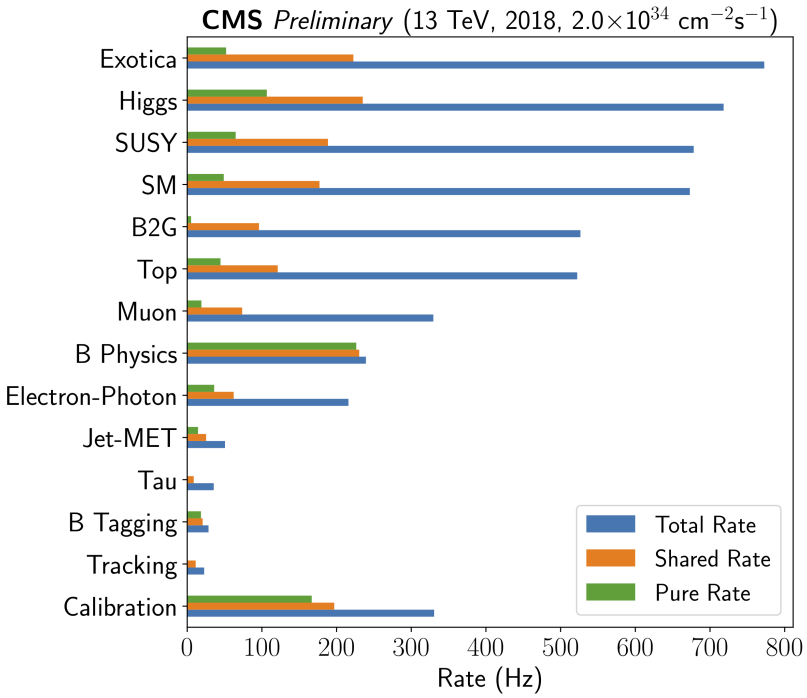


Figure 3.22. | HLT rate allocation per physics group [95].

in each sequence (or "path"), algorithms are executed in order of increasing complexity, and execution of a path is stopped unless evidence for the signal of interest is found. The average computing time is 300 ms, which is fixed by the number of cores (up to 30 000 comprising over 30'000 Central Processing Unit (CPU) cores during the 2018 data-taking) and the input rate. The final decision is based on the detector's full information and only accepted events are sent to the disk for prompt reconstruction. The breakdown of the CPU time spent in the reconstruction code approximately is; 10% in ECAL calorimeter, 10% in HCAL calorimeter, 35% in tracking and 20% in Particle Flow (PF) algorithm. The PF algorithm is used in almost all HLT paths in CMS. It uses the full detector information to describe the overall collision event, identifying individual particles and grouping them into more complex objects which results in improvements in the energy resolution of trigger objects and makes online event reconstruction and selection much closer to that performed in offline reconstruction and analyses. This is described in detail in Sec. 3.3.1.

A flexible and extremely accurate HLT selection together with the stable operation of the DAQ is vital for the success of the CMS physics program. The architecture of the CMS HLT combines the power of the offline reconstruction with the robustness of the online operation of the DAQ. The baseline of the

DAQ design is sketched in Fig. 3.23.

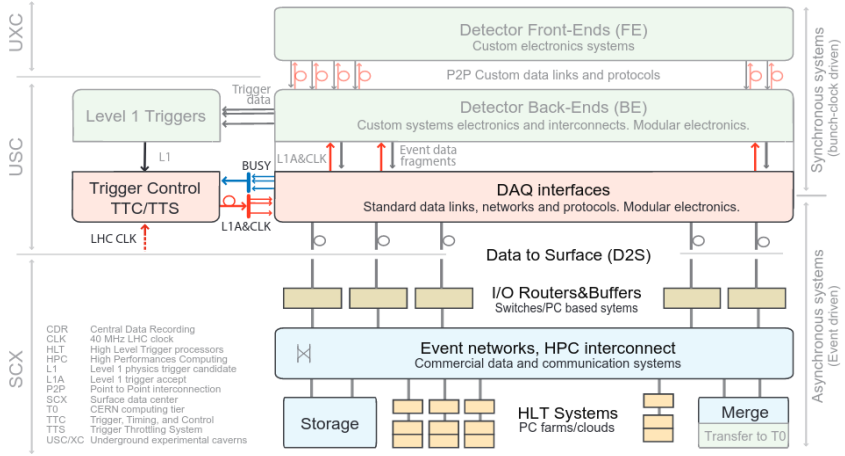


Figure 3.23. | The conceptual design of the overall CMS DAQ. The detector front-end (FE) and back-end (BE) are connected by bidirectional links [96].

Event fragments are read out and stored in "readout buffers" for each event accepted by the Level-1 trigger. The fragments are subsequently assembled into complete events by an Event Builder through a complex of switched networks into "event buffers". The full event content is then handed to one of the HLT Filter Units (FU) for processing, and if events are found to be sufficiently interesting for offline analysis they are forwarded to the storage manager. At the end of Run 1, the DAQ electronics were at the end of their life cycle, also the estimated values of high-level pileup interaction that the detector expected to be facing in Run 2, called the collaboration for a major update to a more powerful DAQ that leverages advances in computing and networking technology. The Advanced Mezzanine Card (AMC) and MicroTCA (short for Micro Telecommunications Computing Architecture, also $\mu-$TCA) were selected for the Phase-1 upgrades. With the CMS Phase-2 upgrade, the number of links from front-end systems and bandwidth requirements increases further. Aggregating many serial input links requires large Field Programmable Gate Arrays (FPGAs), which require more power and board space. simultaneously, a large number of optical connectors must be accommodated on the front panel. These form the cornerstone of CMS's adoption of the Advanced Telecommunications Computing Architecture (ATCA) for CMS's Phase-2 upgrade of modular electronics.

3.2.6. CMS distributed computing system

The computing centers available to CMS around the world are distributed and configured in a tiered architecture that functions as a single coherent system.

Each of the three-tier levels provides different resources and services; Tier-0 CMS-only site at CERN performs the initial processing of the data coming from the detector, those are the data accepted in CMS Online Data Acquisition and Trigger System (TriDAS). It archives the repacked RAW data to tape and distributes a copy among the next-tier stage resources (Tier-1). It runs prompt calibration that proceeds to the reconstruction step, then writes the RECO, Analysis Object Data (AOD), and MiniAOD which will also be distributed to Tier-1. Tier-1 is a set of seven sites, those are CMS's collaborating centers (large national labs, e.g. FNAL, and RAL), which provide substantial CPU power for re-reconstruction, skimming, calibration, and securing a copy of CMS RAW data. Tier-2 centers are based in 55 different countries with substantial CPU resources that provide capacity for user analysis, calibration studies, and Monte Carlo production. Tier-2 centers rely on Tier-1 for large datasets access and for secure storage of Monte Carlo production.

Data is stored as ROOT files. The smallest unit in computational space is the block of files which corresponds to a group of ROOT files that can be accessed together. CMS has a global data catalog called Dataset Aggregation System (DAS) which provides a mapping between the "event collection" and the list of fileblocks corresponding to this abstraction. The locations of these fileblocks within the CMS grid are available for all users and handled recently by the Rucio [97] Distributed Data Management System. Rucio is responsible for transporting data across CMS sites and keeping track of what data exists on which site.

3.2.7. Data quality monitoring

The high quality of data used for the physics analyses is ensured by the dedicated Data Quality Monitoring process (DQM). The purpose of the DQM is to evaluate the detector conditions during the data taking, as well as in the simulated samples. The DQM process in CMS is split into two different stages, with very small latency during the data-taking (online) and after prompt reconstruction of the events (offline).

- ***Online DQM:*** Each detector subsystem is monitored in real-time during the data-taking period, to discard events with low quality of the measured physics objects. The online DQM system receives event data through a storage manager event server and fills histograms at an event rate of 10–15 Hz. In addition, a small number of histograms is filled in the HLT filter units, which process events at up to 100 kHz. These histograms, including alarm states based on quality test results, are made available to a central DQM graphical user interface server (GUI) for visualization in real-time and are stored in a ROOT file periodically during the run. At the end of the run, the final archived results are uploaded to a large disk pool. Eventually,

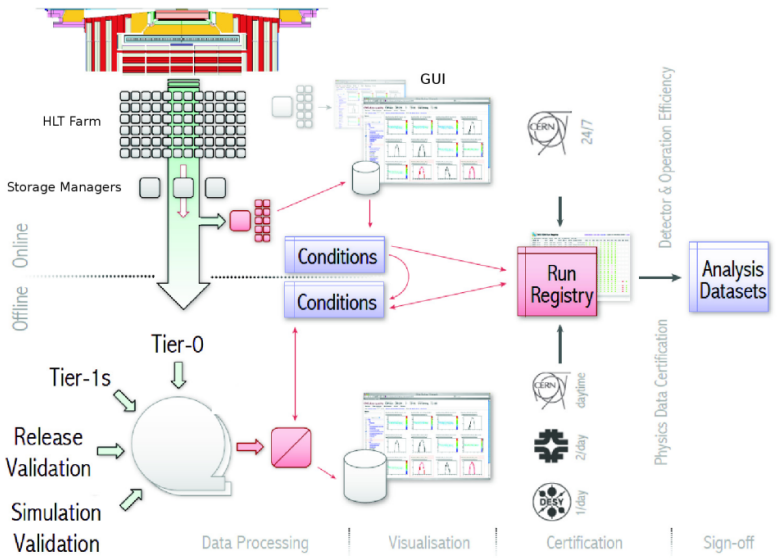


Figure 3.24. | A simplified schematic view of the online and offline DQM workflow [98].

the files are merged and backed up to tape.

- ***Offline DQM:*** During the prompt reconstruction performed at Tier-0, the reconstruction and calibration are certified in the recorded data for a selection of events that can be used in physical analyses, it provides the final quality flags used to perform the data certification. During release validation, which is a central workflow, it assessed the quality of new reconstruction algorithms and high-level physical object definitions that are regularly integrated into the CMS Software Release Framework (CMSSW).

Importantly the online and offline DQM GUI web servers provide a common interface giving the entire worldwide collaboration access to inspection and analysis of all DQM data.

3.3. Event reconstruction

Until now, we've exclusively acquainted ourselves with the raw signatures of SM particles as they pass through the CMS detector. This section is devoted to explaining the use of electronic signatures of particle travel through the different parts of the CMS sub-detectors to reveal their types, charges, mass, momentum, etc.

3.3.1. Particle-flow algorithm

The event reconstruction at the CMS experiment is predominantly based on the Particle-Flow (PF) algorithm [99] which performs a correlation of the basic elements (i.e. tracks and clusters) obtained from all sub-detector systems, in order to identify all particles in the event and measure their properties accurately. Fig. 3.25 is a good illustration of the concept of the PF algorithm. Here is a possible list of combinations considered in particle identification and reconstruction.

- inner detector track, electromagnetic calorimeter cluster;
- inner detector track, hadronic calorimeter cluster;
- electromagnetic calorimeter cluster, hadronic calorimeter cluster;
- inner detector track, muon track;

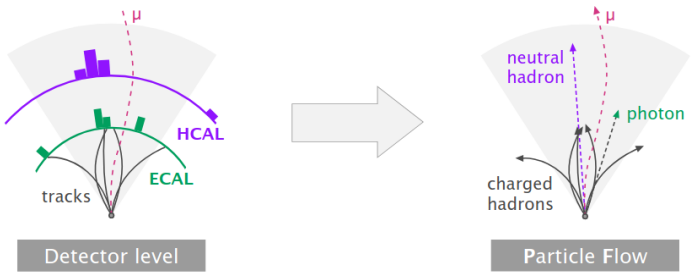


Figure 3.25. | Schematic view of the Particle-Flow PF concept: a combination of signals from different sub-detectors to construct objects, and represent individual particles with defined types [89].

It is important to bear in mind that PF reconstructed particles are not actually the SM particles but rather a defined collection of detector signatures that resemble most to those we know in the SM. The selections used to define each object are carefully counted for fake objects as for example electrons misidentified as jets, or charged pions as electrons.

3.3.2. Object reconstruction

3.3.2.1. Electrons and photons

From Sec. 3.2 we acquired knowledge about how electrons and photons release their energy within the electromagnetic calorimeter. However, the reconstruction of these two objects commonly encounters two primary challenges. First, the ECAL alone it is impossible to tell electrons from photons. Second, the difficulties lie in distinguishing prompt from non-prompt objects since *bremsstrahlung* causes electrons to emit photons, whilst pair-production causes photons to

produce electron–positron pairs. Thus, by the time the electron or photon reaches the ECAL, it may no longer be a single particle, but instead a shower of multiple electrons and photons. Luckily, photons are electrically neutral and do not leave tracks, whilst electrons and positrons will produce tracks that point to their calorimeter deposits. Additionally, recall that the trajectory of an electron losing momentum by emitting *bremsstrahlung* photons changes its curvature in the tracker, this is helpful when we want to tell these particles apart. To estimate the track parameters, a dedicated tracking algorithm, based on the Gaussian sum filter (GSF) is used. For these specific types of particles, PF algorithm is used to combine the clusters from the individual particles into a single object to recover the energy of the primary electron or photon. Here is a brief outline of the reconstruction steps. A more detailed description can be found in [100].

1. The energy reconstruction algorithm starts with the formation of clusters by grouping together crystals with energies exceeding a predefined threshold (typically $\approx$ 80 MeV in EB and $\approx$ 300 MeV in EE), which is generally two or three times bigger than the electronic noise expected for these crystals. A seed cluster is then defined as the one containing most of the energy deposited in any specific region, with a minimum transverse energy (E_T^{seed}) above 1 GeV.
   We define the energy of an object transverse to the beam $E_T^{seed} = \sqrt{m^2 + p_{\mathrm{T}}^2}$ for an object of mass m and transverse momentum p_{T}.
2. A procedure referred to as "superclustering" (SC) then follows which involves clustering within a certain geometric area ("window") around the seed cluster of ECAL clusters to include photon conversions and *bremsstrahlung* losses.
3. Trajectory seeds in the pixel detector that are compatible with the SC position and the trajectory of an electron are used to seed the GSF tracking step.
4. All tracks selected with $p_{\mathrm{T}} > 2$ GeV that are reconstructed from hits in the tracker using an iterative algorithm known as the Kalman filter (KF) are to be checked for compatibility with an electron trajectory hypothesis; if successful, they are also used to seed the previous GSF tracking step.
5. A dedicated algorithm is used to find "generic tracks" that are likely to originate from photons converting into e^+e^- pairs.
6. ECAL clusters, SCs, GSF tracks, and generic tracks associated with electrons, as well as conversion tracks and associated clusters, are all imported into the PF algorithm that links the elements together into blocks of particles.
7. These blocks are resolved into electron and photon (e and γ) objects, starting from either a GSF track or an SC, respectively. At this point, there is no differentiation between electron and photon candidates. The final

list of linked ECAL clusters for each candidate is promoted to a refined supercluster.

8. Electron or photon objects are built from the refined SCs based on loose selection requirements. All objects passing the selection with an associated GSF track are labeled as electrons; without a GSF track, they are labeled as photons. This collection is known as the unbiased e/γ collection and is used as a starting point for the vast majority of analyses involving electrons and photons.
9. To separate electrons and photons from hadrons in the PF framework, a tighter selection is applied to these e/γ objects to decide if they are accepted as an electron or an isolated photon. If the e/γ object passes both the electron and the photon selection criteria, its object type is determined by whether it has a GSF track with a hit in the first layer of the pixel detector. If it fails the electron and photon selection criteria, its basic elements (ECAL clusters and generic tracks) are further considered to form neutral hadrons, charged hadrons, or non-isolated photons in the PF framework.

For physics analysis, we often define several different categories of electron reconstruction; ***loose***, ***medium***, and ***tight*** working point. The ***loose*** electron identification might involve basic requirements such as minimum energy and minimal track quality. It aims to include as many electron-like signatures as possible without overly rejecting potential candidates. While ***medium*** electron identification criteria are more restrictive than the ***loose*** criteria. They might include additional requirements on the shower shape of the energy deposition in the electromagnetic calorimeter, the matching between the track and the calorimeter energy deposits, and the track's impact parameter. ***Tight*** electron identification involves even stricter requirements. Moving from ***loose*** to ***tight***, the electron definitions are set to be more strict to reduce the fake jet background for electron identification, at the cost of reducing the overall electron reconstruction efficiency. This trade-off between purity and efficiency is characteristic of all attempts to reconstruct objects from detector signatures. Fig. 3.26 for example illustrates the electron identification efficiency in data when using a cut based ***loose*** working point.

3.3.2.2. Muons

Muons are, in principle, straightforward to reconstruct, given that only neutrinos can reach the muon system which leaves no trace behind, unlike muons.

In CMS, we define different muon categories as follows:

1. ***Standalone muons:*** These are formed from hits in the drift tubes, cathode strip chambers, or resistive plate chambers of the muon spectrometer, which are combined to form track segments.

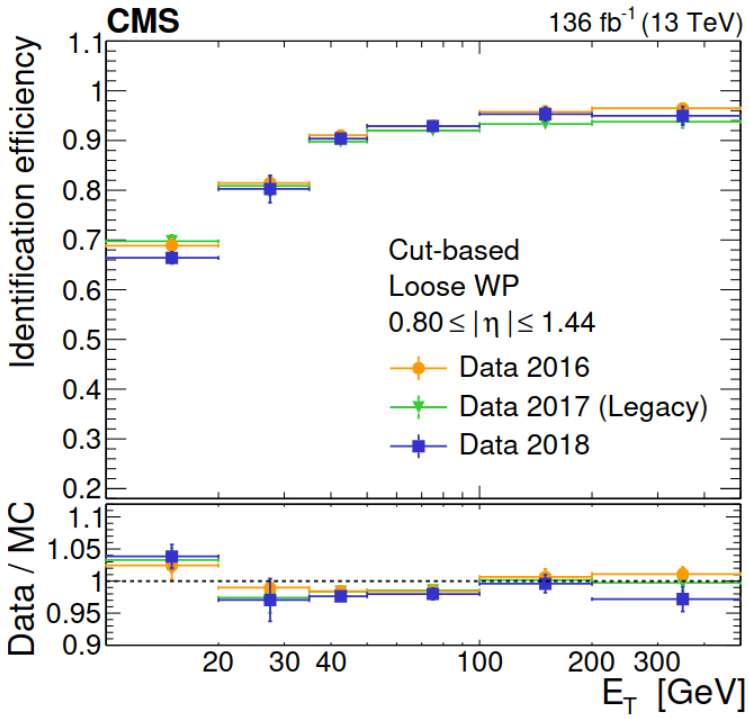


Figure 3.26. | Cut-based loose electron identification efficiency in data as a function of the electron p_{T} shown for 2016, 2017 "Legacy", and 2018 data-taking periods. The vertical bars on the markers represent combined statistical and systematic uncertainties [100].

2. ***Tracker muons:*** All tracks with $p_{\mathrm{T}} > 0.5$ GeV in the inner detector are potential candidates for muon tracks. These are extrapolated to the muon chambers, and if they have at least one match with a track stub [1], the track becomes a tracker muon track.
3. ***Global muons:*** The standalone muon tracks are extrapolated back to the inner detector and, if a match is found, the two tracks are combined to form a global muon. The information from both sub-detectors is then used to obtain the best resolution on the muon track parameters.

[1] The CMS tracker is built from two types of modules named "2S" (double strip) and "PS" (pixel-strip) with the ability to discriminate hits from particles with $p_{\mathrm{T}} > 2-3GeV$ meaning eliminating about 99% of hits with no relevance for the trigger decision. These double hits often called "Stubs", are sent to the back-end system where tracks are reconstructed and then used in the generation of the Level-1 trigger.

3.3.2.3. Jets

Jets are spatially grouped collections of long-lived particles that are produced in the hadronization of quark or gluon, with the following predictable energy composition:

- ∼65% charged hadrons.
- ∼25% photons (from neutral pions).
- ∼10% neutral hadrons.

As these jets of particles propagate through the CMS detector, they leave signals in the subsystems such as the tracker, electromagnetic (ECAL), and hadronic calorimeters (HCAL). These signals are combined to serve as input for the jet reconstruction algorithms. Within CMS, three different algorithms are supported; a calorimeter-based approach which uses the ECAL and HCAL energy deposits in the form of CaloTowers[1] objects to reconstruct what we often call calorimeter jets (CaloJets); the "Jet-Plus-Track" approach, which improves the measurement of CaloJets such as the p_{T} resolution by exploiting the associated tracks; and finally, the "Particle Flow (PF)" approach, which attempts to reconstruct individually each particle in the event, prior to the jet clustering, based on information from all relevant sub-detectors. The significant improvement resulted in jet energy resolutions, especially at low transverse momentum (p_{T}) gives the "Particle Flow" approach advantages over other types of algorithms and makes it by far the most used within CMS physics analyses and for which this section will be devoted.

■ Pileup jets

Inevitably, during each bunch crossing of the LHC, several simultaneous "pile-up" collisions produce particles that will coexist in the list of PF candidates of the same jet. Since they do not originate from the same primary interaction point (or vertex), they must be removed during reconstruction. Charged Hadron Subtraction (CHS) or PileUp Per Particle Identification (PUPPI) are two algorithms used in addition to the reconstruction step of PF for this purpose.

- ***Charged hadron subtraction (CHS) jet:***
  All charged hadron candidates are associated with a track. If the track is not associated with the primary vertex, that charged hadron can be removed from the list. CHS is limited to the region of the detector covered by the inner tracker.
- ***PileUp Per Particle Identification (PUPPI) jet:***
  For each particle, a local shape α is defined which probes the collinear versus soft diffuse structure in the neighborhood of the particle. The former is

[1] A calorimeter tower consists of one or more HCAL cells and the geometrically corresponding ECAL crystals.

indicative of particles originating from the hard scattering and the latter of particles originating from pileup interactions. The distribution of α for charged pileup is used on an event-by-event basis to calculate a weight (w_i) for each particle. The weights describe the degree to which particles are pileup-like and are used to rescale their four-momenta, superseding the need for jet-based corrections. This method is more stable and performant in high pileup scenarios such as the upcoming HL-LHC era.
The α variable for a given particle i is defined as;

$$\alpha_i = log \sum_{j \neq i, \Delta R_{ij} < R_0} (\frac{P_{T,j}}{\Delta R_{ij}})^2, \tag{3.7.}$$

where i refers to the particle in question, j are other particles, $P_{T,j}$ is the transverse momentum of particle j in GeV, and $\Delta R_{ij} = \sqrt{\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2}$ (where ϕ is the azimuthal angle in radians) is the distance between the particles i and j in the $\eta - \phi$ plane.
For $|\eta_i| < 2.5$, where tracking information is available, only charged particles associated with the "leading vertex" (LV) are included as particle j, for $|\eta_i| > 2.5$, j are all kinds of reconstructed particles. Charged particles assigned to PU vertices are used to generate the expected PU distribution in an event. From this expected distribution a median ($\overline{\alpha}_{PU}$) and root-mean-square (RMS) (α_{PU}^{RMS}) of the α values are computed. The α_i of each neutral particle is compared with the computed median and RMS of the α distribution of the charged PU particles using a signed χ^2 approximation

$$\chi^2 = \frac{(\alpha_i - \overline{\alpha}_{PU})|\alpha_i - \overline{\alpha}_{PU}|}{(\alpha_{PU}^{RMS})^2}. \tag{3.8.}$$

The weight for each neutral particle is calculated with a cumulative distribution function and multiplied by the four-momentum of the particle:

$$w_i = F_{\chi^2, NDF=1}(\chi_i^2), \tag{3.9.}$$

where $F_{\chi^2, NDF=1}$ is the cumulative distribution function of the χ^2 distribution with one degree of freedom.

The PF particles are formed into jets using a "clustering algorithm" which iterates over particle pairs and finds the two closest (i and j) in some distance measure and determines whether to combine them or not:

$$d_{ij} = min(k_{T_i}^{2p}, k_{T_j}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \tag{3.10.}$$

where k_T is the transverse momentum and $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the angular distance between the two entities. The choice of the parameter p defines distinct algorithms, for instance; $p = 1$ (k_T), $p = 0$ (Cambridge/Aachen), $p =$

-1 (anti$-\mathrm{k_T}$).

The radius parameter $R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ effectively defines the size of the jet cone. Anti$-\mathrm{k_T}$ jets with cone R = 0.4 (AK4) are used for jet ID studies, quark-gluon jet discrimination, pileup jet identification, and in the event selection for W-tagging and top studies. While anti$-\mathrm{k_T}$ jets with cone R = 0.8 (AK8) are used to reconstruct W-jets and high p_T top jets which are often the results of two or more collimated boosted AK4 jets. Other cases for jets with a size parameter R= 0.5, R= 0.7, and R= 1.5 (the latter is called Cambridge-Aachen jets and it is often used to reconstruct low p_T top jets) are out of the scope of this thesis. All jet substructure observables are computed using PF candidates calibrated prior to jet clustering. Nevertheless, the resulting jets require another small correction to the jet momentum and energy that accounts for tracking inefficiencies.

For this thesis, the building blocks for the jet reconstruction are particle-flow candidates, which are clustered into jets using the algorithm of cone radius R=0.4 and R=0.8. This will be discussed in detail later in section 5.2.

■ Noise jet identification

The fractions of the jet energy carried by certain types of PF candidates clustered into a jet along with its multiplicity are used in order to differentiate genuine jets from "noise jets" which are instances where detector signals mimic the characteristics of jets but do not arise from actual particle interactions of interest. These noise sources can include electronic noise, cosmic rays, detector defects, or other sources of spurious signals that are mistaken for particles. Both jet energy fraction and multiplicity variables are sensitive to different sources of noise from the hadronic (HCAL) and electromagnetic (ECAL) calorimeters.

Within CMS, three PF jet ID working points are defined and their criteria are defined in Tab. 3.1, ***"loose"***, ***"tight"*** and ***"tight lepton veto"***. While ***"loose"*** and ***"tight"*** working points are designed to remove jets originating from calorimetric noise, the ***"tight lepton veto"*** additionally rejects any potential background from mis-reconstructed electron or muon candidates, effectively resolving also the ambiguity between isolated lepton candidates and jets reconstructed from single lepton candidates.

■ Tagging of heavy-flavour jets

Heavy-flavour (HF) jets contain bottom or charm hadron as a result of the fragmentation process that carries a large fraction of the initial parton momentum. Hadrons within HF jets have a sizeable lifetime, with a $c\tau$ of $\sim$ 0.5 mm and $\sim$ 0.3 mm for the bottom and charm, respectively. Most of the HF hadrons produced in the fragmentation process undergo a decay process far enough from

Table 3.1. | The PF jet ID criteria [101].

Jet Variables	$\|\eta\|$ range	Loose	Tight	Tight Lepton Veto
Charged Hadron Fraction	$\|\eta\| < 2.4$	> 0		
Charged Multiplicity	$\|\eta\| < 2.4$	> 0		
Charged EM Fraction	$\|\eta\| < 2.4$	< 0.99	< 0.99	< 0.9
Muon Fraction	$\|\eta\| < 2.4$	< 0.8		
Neutral Hadron Fraction	$\|\eta\| < 2.7$	< 0.99	< 0.9	< 0.9
Neutral EM Fraction	$\|\eta\| < 2.7$	< 0.99	< 0.9	< 0.9
Neutral Multiplicity	$2.7 < \|\eta\| < 5$	< 0.9		
	$2.7 < \|\eta\| < 5$	> 2		
	$3. < \|\eta\| < 5$	> 10		

the primary interaction vertex (PV) to result in displaced tracks. Their decay products often are clustered in secondary vertices (SV) as shown in Fig. 3.27. The displacement of tracks with respect to the primary vertex is characterized by their impact parameter (IP), which is defined as the distance between the primary vertex and the tracks at their points of closest approach.

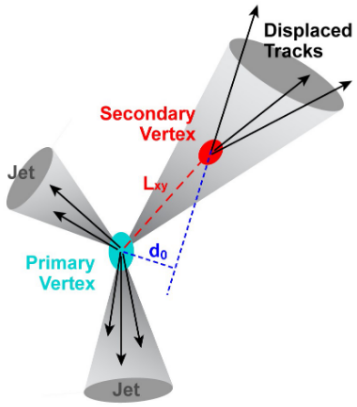
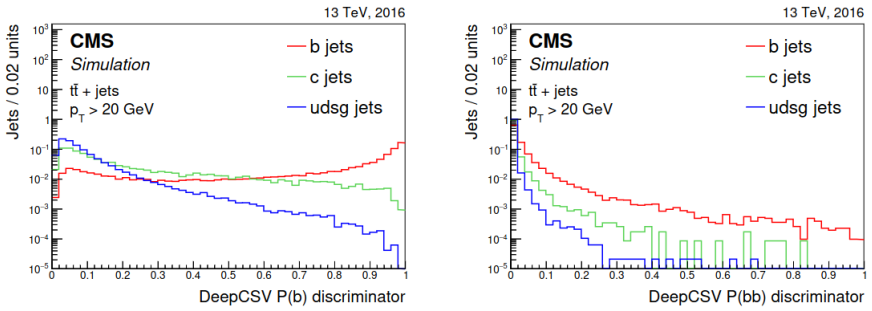


Figure 3.27. | Illustration of heavy-flavour jet production with a secondary vertice [101].

Machine learning has long been used in the field of jet flavor classification. Within CMS, the Combined Secondary Vertex (CSV) algorithm developed in Run-I combines the variables of secondary vertexes in a likelihood-ratio discriminant. In Run-II, the CSV algorithm was updated to a new version, referred to as CSVv2. In addition, another version of the CSV algorithm was developed that uses deep machine learning (DeepCSV). Ultimately, a more sophisticated algorithm,

DeepJet, was developed with a new network architecture different from its predecessors. More details about these taggers can be found in [101].

DeepCSV and DeepJet [102] are the two algorithms used in the search $(H/A \rightarrow ZA/ZH)$ which will be described more in detail in Chap. 5 and therefore this section will be devoted to them. DeepCSV is a fully connected neural network, developed with KERAS [103] and consisting of 5 layers with 100 nodes, each using high-level features of selected tracks and vertices as input. The inputs are a combination of Inclusive Vertex Finder (IVF) secondary vertices and up to the first six track variables, with all jet flavor and vertices categories that are within the tracker accept and p_{T} range from 20 GeV up to 1 TeV. The Fig. 3.28 shows the DeepCSV probability $P(f)$ distributions for b hadron jets and at least two for b hadrons jets in $t\bar{t}$ +jets events.



(a) (b)

Figure 3.28. | Discriminator distributions of (a) DeepCSV P(b), (b) DeepCSV P($b\bar{b}$) [101].

The DeepCSV tagger is used also for c tagging. Combining the probabilities from the five categories, the algorithm defines a *DeepCSVCvsB* discriminant that is used to discriminate c jets from b jets;

$$DeepCSVCvsB = \frac{P(c) + P(c - \bar{c})}{1 - P(udsg)}. \tag{3.11.}$$

In the same way, *DeepCSVCvsL* is defined to discriminate c jets from light jets;

$$DeepCSVCvsL = \frac{P(c) + P(c\bar{c})}{1 - (P(b) + P(b\bar{b}))}. \tag{3.12.}$$

The DeepJet algorithm, on the other hand, is designed in such a way that its architecture simultaneously examines all jet components using a large number of input variables that are classified into four groups: global variables (jet kinematics, the number of tracks in the jet, etc.), charged and neutral PF candidates, and variables of the SVs related to the jet.

Training is performed using the Adam optimizer [103] with a learning rate of 3×10^{-4} for 65 epochs and categorical cross-entropy loss. The Receiver–Operative Characteristic (ROC) curves are shown in Fig. 3.29 for two different p_{T} ranges for the same dataset are reported and compared to the performance of the DeepCSV tagger, such curves indicate the background misidentification efficiency versus the signal efficiency measured in the Monte Carlo simulation.

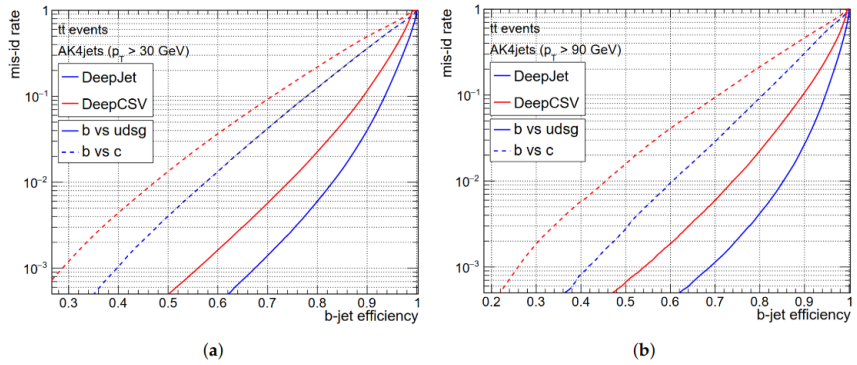


Figure 3.29. | Performance of the DeepJet and DeepCSV b-tagging algorithms on $t\bar{t}$ events with both top quarks decaying hadronically. The jets are required to have $p_{\mathrm{T}} > 30$ GeV (a) and $p_{\mathrm{T}} > 90$ GeV (b). The performance is shown for both b .vs. c hadron classification (dashed lines), and b .vs. light hadron (solid lines) [102].

the heavy flavor algorithm had different working points, categorized as loose, medium, and tight, and they are defined by specific cut values within the discriminator distribution. Each working point is established by assessing the rate at which a non-b jet is incorrectly identified as a b jet: 10% for loose, 1% for medium, and 0.1% for tight. The determination of these working points involves using the b-tagging discriminant on jets that meet the criteria of $|\eta| < 2.5$ and $p_{\mathrm{T}} > 30$GeV, as drawn from QCD multijet Monte Carlo simulations as shown in Fig. 3.30 for DeepJet tagger.

■ Identification of b jets in boosted topologies

In CMS, we have often used algorithms for heavy flavor tagging in boosted topologies different from the ones mentioned earlier for resolved AK4 jets. For boosted jets, heavy flavor tagging can be applied to either the AK8 jet or its sub-jets, as shown by the diagram in Fig. 3.31 (left and middle).

While DeepCSV and CSVv2 taggers can be used as well for boosted jet tagging, a more performant algorithm was developed for Run-II, DeepDoubleB/C. This algorithm can be used for the identification of the decay of a boosted object to a b/c quark pair. It is trained as a binary tagger aiming at rejecting the QCD background (DeepDoubleBvL and DeepDoubleCvL), exploiting not only

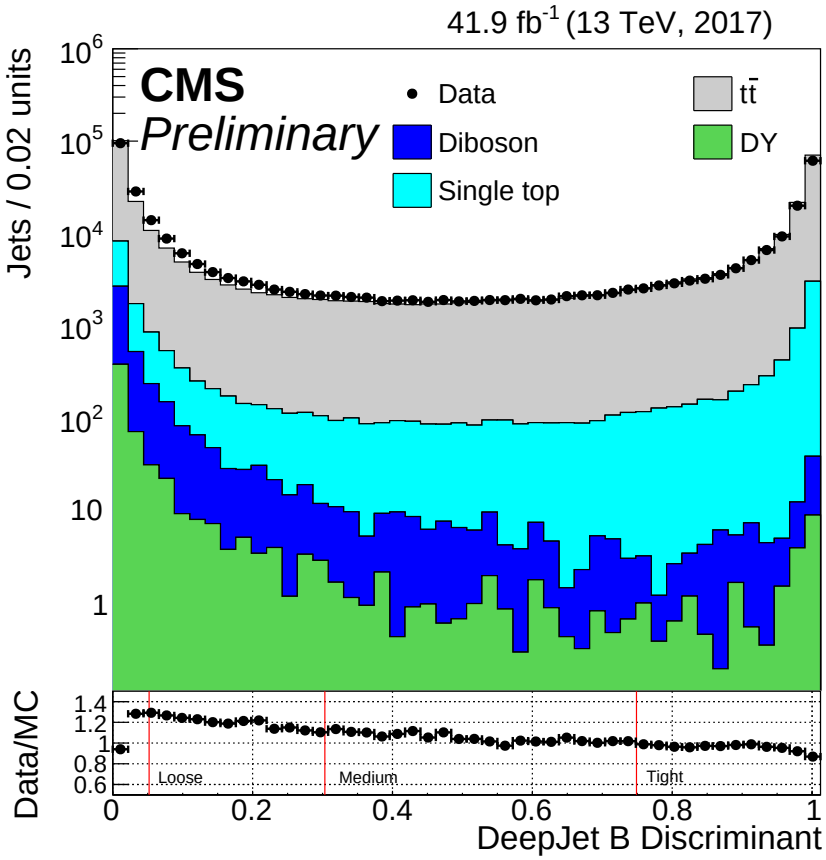


Figure 3.30. | Comparison of the CMS DeepJet b tagging discriminant distribution between simulation and 2017 data in dilepton top pair events. The loose, medium, and tight working points are determined based on distinct thresholds of the btag discriminant [104].

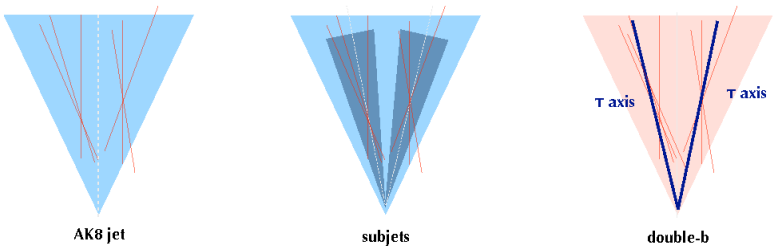


Figure 3.31. | schematic representation of the AK8 jet (left) and subject (middle) b-tagging approaches, and of the double-b tagger approach (right) [54].

the presence of two b/c hadrons inside the AK8 jet but also the correlation between the directions of their momentum. Boosted HF tagger also uses the N-subjettiness axes to associate tracks and vertex to the subjets. A key sensitive variable to the substructure is N-subjettiness, τ_N, which is a jet shape variable that is defined as the weighted distance p_T between each jet constituent and its axis nearest subjet (ΔR):

$$\tau_N = \frac{1}{d_0} \sum_k p_T^k min(\Delta R_{1,k} ... \Delta R_{N,k}), \tag{3.13.}$$

where k runs over all jet constituents, $d_0 = \sum_k p_T^k R_0$ is a renormalisation factor and $R_0 = 0.8$ is the jet distance parameter.

A smaller value of the τ_N variable is consistent with fewer subjets within the jet. The estimation of the subjet axes, also called τ axes can be derived to estimate the directions of the partons giving rise to the subjets. The advantage of DeepDoubleBvL tagger is that allows the discrimination of single-parton jets in multijet events from $H \to b\bar{b}$. It also permits the identification of boosted $Z \to b\bar{b}$ jets or any other boosted $b\bar{b}$ resonance where the kinematics of the decay products are similar. The performance of the double-b discriminator is shown in Fig. 3.32.

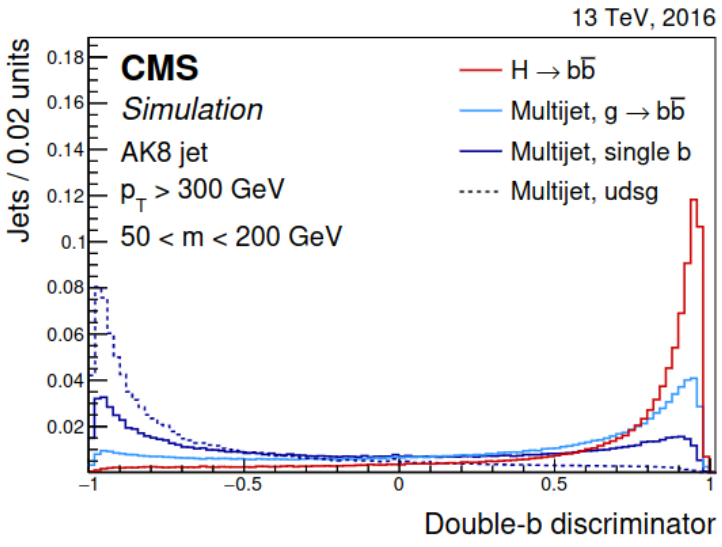


Figure 3.32. The deep double-b tagger performance when discriminating $H \to b\bar{b}$ signal events of a Kaluza–Klein graviton decaying to two Higgs bosons from background jets in an inclusive multijet sample containing zero, one, or two b quarks [54].

Whether the use of subjet b-tagging or double-b tagging is favored in physics data analysis it strongly depends on the flavor composition, the p_{T} distribution of the signal jets and the background processes considered. Fig. 3.33 illustrate the difference in the performance of these two algorithms using a multijet sample, in $g \to b\bar{b}$ jets, and in single $b\bar{b}$ jets.

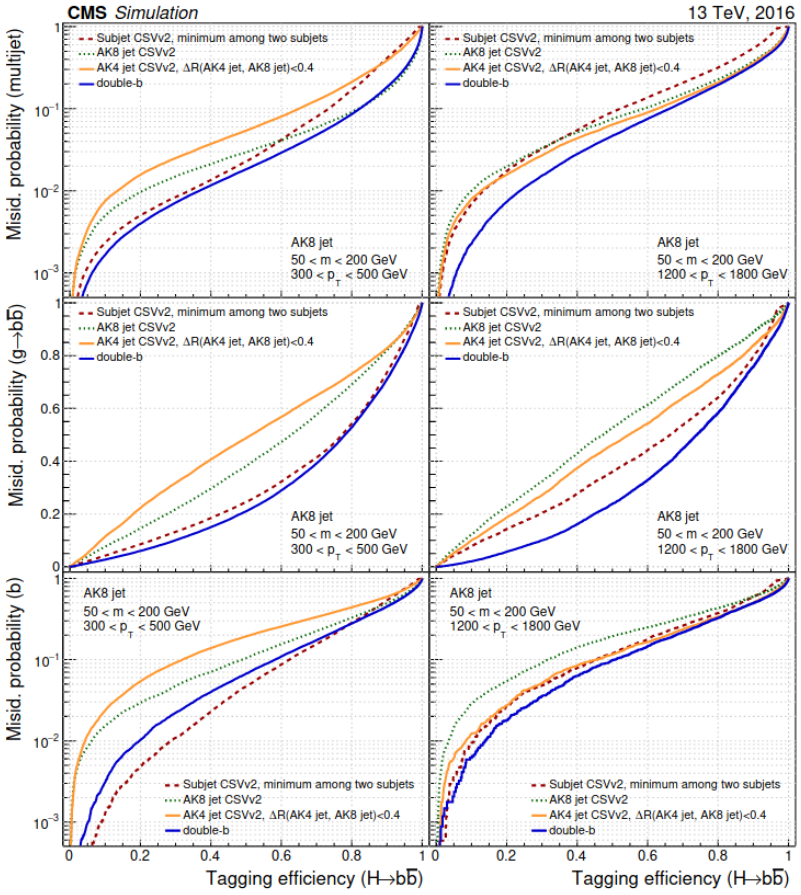


Figure 3.33. | Three scenarios with different simulated events are used to derive the efficiency to correctly tag $H \to b\bar{b}$ jets as a function of the misidentification rate in a multijet sample (upper), in $g \to b\bar{b}$ jets (middle), and in single $b\bar{b}$ jets (lower). Different choices of heavy flavor tagging were checked. The CSVv2 algorithm is used on AK8 jets, their subjets, and AK4 jets matched to AK8 jets. For the subjet b-tagging curves, both subjets are required to be tagged. The double-b tagger is applied to AK8 jets [54].

■ Two-tag counting method

Two-tag counting (2TagCount) is used to derive b tagging scale factors that would be used to correct for any discrepancies between data and simulated

events of any analysis. This is a simple but robust cut and count measurement in a $t\bar{t}$ dilepton event sample. The leptons pair are required to be opposite in charge and their invariant mass must be $M_{ll} > 12$ GeV. If the two same charge leptons have also the same flavor, the contribution of Z+jet events is reduced by applying a veto around the Z boson mass, $|M_{ll} - M_Z| > 10$ GeV, and requiring $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV.

For events that meet the mentioned criteria, it is necessary to have exactly two jets. The efficiency of identifying b jets, labeled as ϵ_b, can be determined by counting how many events with two b-tagged jets are found within the selected group of events.

$$N_{\mathrm{2b-tagged}} - N_{\mathrm{2b-tagged}}^{\mathrm{non-bjet}} = \epsilon_b^2 n_{\mathrm{2bjets}}, \tag{3.14.}$$

where $N_{\mathrm{2b-tagged}}$ is the number of events with two b-tagged jets taken from data,$N_{\mathrm{2b-tagged}}^{\mathrm{non-bjet}}$ is the number of events with two b-tagged jets when one of them at least mistagged c or light, and n_{2bjets} is the number of events with two true b-jets. The unknown terms in Eq. 3.14, $N_{\mathrm{2b-tagged}}^{\mathrm{non-bjet}}$ and n_{2bjets} can be obtained from simulation.

Naturally, one would think that this measurement will be very dependent on Monte-Carlo statistics, so we can reduce the dependency on the production cross-section of $t\bar{t}$, by dividing Eq. 3.14 by the number of events selected, which becomes as follows:

$$\epsilon_b = \sqrt{\frac{F_{\mathrm{2b-tagged}} - F_{\mathrm{2b-tagged}}^{\mathrm{non-bjet}}}{f_{\mathrm{2bjets}}}}, \tag{3.15.}$$

where $F_{\mathrm{2b-tagged}}$ is the fraction of events with two b-tagged jets, $F_{\mathrm{2b-tagged}}$ is the fraction of events with two b-tagged jets of which at least one is a non-b jet, and f_{2bjets} is the fraction of events with two true b jets. Depending on the tagger and working point under study, data and simulated samples passing the event selection criteria are filled in histograms for different jet p_{T} and η ranges similar to Fig. 3.34 (a representation of my own work) which allow the computation of p_{T} and η dependent b-tag efficiency scale factor using Eq. 3.15.

3.3.2.4. Missing Transverse Energy

Particles that do not interact with electromagnetic or strong forces such as neutrinos are not easy to detect. They come out of the detector unseen, but due to the laws of conservation of energy and momentum, we can infer their presence in the detector. The initial momentum of the colliding partons along the beam axis is not known and because the energy of each hadron is divided and constantly exchanged between its constituents we cannot determine the

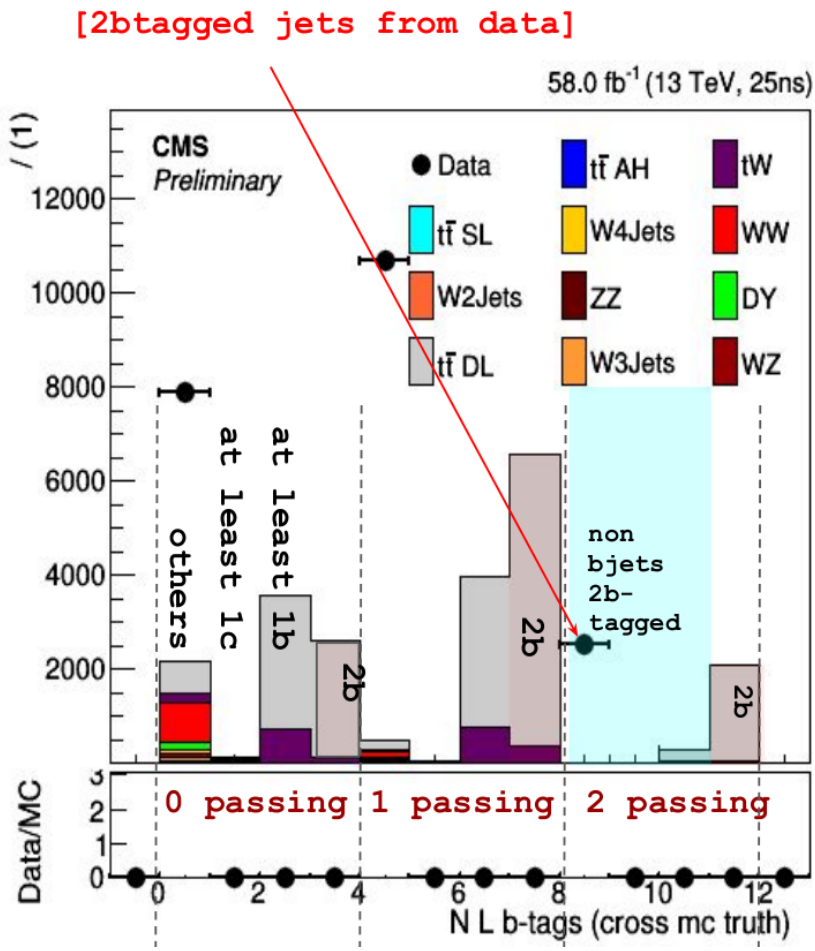


Figure 3.34. | The 2TagCount histogram which illustrates the distribution of data and Monte-Carlo simulated events samples between 0 btag, 1 btag, and 2 btag jets passing DeepJet discriminator loose working point for jet $|\eta| < 2.5$ and 140 GeV $< p_T <$ 200 GeV (own work).

total momentum missing. Nevertheless, the initial energy or particles moving transverse to the detector beam axis are zero, so any sharp pulse in the transverse direction indicates missing transverse energy, also known as $\vec{E}_T^{miss}$ or MET, which can be defined as

$$\vec{E}_T^{miss} = -\sum_i (E_i \sin\theta_i \cos\phi_i \hat{x} + E_i \sin\theta_i \sin\phi_i \hat{y}) = E_x \hat{x} + E_y \hat{y}, \qquad (3.16.)$$

where $\hat{x}$, $\hat{y}$ are the unit vectors in the direction of the x and y axes.

The majority of the searches involving missing energy are more challenging and highly dependent on PF to accurately reconstruct all other particles involved in the collisions, as errors in measuring the energy of any particle may appear like the missing energy was swept away by another elusive particle that does not actually exist.

Chapter 4. Data analysis methods and tools

Statistical analysis plays a major role in experimental physics, it is through statistical concepts that we quantify the correspondence between theoretical predictions and experimental observations. While the statistical analysis of data is often the last step in an experimental physics result, thinking through the requirements of a robust statistical statement is the best way of developing an analysis strategy. This chapter will introduce some important details about statistical models, likelihoods and related concepts so that their meaning is clear when referred to in Chap. 5. The concepts covered in this chapter are based on the following references [105–107].

4.1. From statistical models to likelihoods

Statistical models are the cornerstone of statistical inference. Therefore, before delving into the use of likelihood in high-energy physics, we should explore the concept of "statistical models".

A statistical model is a mathematical representation of a system or process that integrates a set of statistical assumptions concerning the data at hand and the theory model being used. For example, linear regression is a statistical model that describes the relationship between a dependent variable and one or more independent variables. Depending on the number of random and non-random variables, there are three primary types of statistical models:

- **Parametric:** Those type of model requires some pre-defined parameters such as mean, variance, or shape to make predictions. They assume a specific functional form or distribution for the underlying data. Poisson distribution and linear regressions are examples of parametric models. They are often used when there is prior knowledge or theoretical understanding of the data-generating process, as the assumptions can help simplify the analysis and interpretation. However, these assumptions may not hold in all cases, and if violated, the model may provide inaccurate results. In such situations, non-parametric or semi-parametric models can be employed to relax some of these assumptions.
- **Non-parametric:** Those such as Decision Trees, Gaussian kernels and Bayesian Blocks (the latter will be used for signal search reported in Chap. 5) in which they do not make explicit assumptions about the functional form or distribution of the data. They are more flexible than parametric models

and offer more accurate results given their flexibility in capturing nonlinear patterns, interactions, and other complex structures that may not be captured by parametric models. They are data-driven models that rely heavily on the observed data to make inferences or predictions and are often used when the data does not meet the assumptions of parametric models or when there is limited prior knowledge about the data-generating process.

- **Semi-parametric:** Those are the type of statistical model that combines parametric and non-parametric elements aiming to capture complex relationships in the data while leveraging the advantages of parametric modeling. Gaussian mixture models are an example of semi-parametric models.

Statistical models provide a rigorous framework for analyzing and interpreting the data obtained in HEP experiments. If we consider a collection of values x as the outcome of a measurement, which could represent single or multiple values, and its components can be discrete or continuous (e.g. it could be the result of an individual collision event, a set of events, or a collection of multiple datasets such as tracks of particles in the CMS detectors). Let θ be the parameters of the model, it can be the quantities that characterize the response of the detectors or calibration parameters, etc. Assuming θ are kept fixed, the observations x won't be the same given that each measurement is subject to many sources of variability that can lead to different observations. Therefore, statistical models are usually written in the form of $p(x|\theta)$ and are not deterministic equations but rather probability density functions (PDFs) for obtaining various x, given a set of parameters θ.

The Likelihood $L(\theta)$ (often also denoted as $L(x|\theta)$, representing the likelihood of observing a set of data x given a specific parameter or set of parameters θ) on the other hand, is very often confused with probability when it is not. It is a measure of how well a statistical model fits a set of observed data and it is a function that takes as input the model parameters and outputs the probability of observing the data given those parameters. The likelihood function is used to estimate the values of the model parameters that best fit the observed data. In other words, the likelihood is a tool used to assess the fit of a statistical model to a set of data. We say a good statistical model will have a likelihood function that assigns high probabilities to the observed data, indicating that the model is a good fit for the data.

4.2. The maximum likelihood estimation

In statistics, maximum likelihood estimation (MLE) is a method of estimating the parameters of an assumed probability distribution, given some observed data. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable. Wherever there is a parameter space that maximizes the likelihood function, we speak of maximum

likelihood estimation. In many cases, the likelihood is a function of more than one parameter, but in the search for new physics our interest is in estimating just one, this is what we call the parameter of interest or the *intensity of the signal* μ, where $\mu = 1$ corresponds to the presence of a signal, while $\mu = 0$ to its absence.

If we consider a simple case of a single channel with one signal, one background contribution, and no systematic uncertainties, denoting the number of signal events as S and the number of background events as B. Similarly, we denote the signal and background "shapes" by $f_S(x)$, $f_B(x)$. From the probability model, then we obtain n events in the data where the discriminant variable for event e has the value x_e; thus the complete dataset will be denoted $x_1 ... x_n$. We can then write the likelihood function from the product of the Poisson probability of obtaining n events when $\mu S + B$ are expected and the probability density of obtaining x_e based on the relative mixture $f_S(x)$ and $f_B(x)$ for a given value of μ:

$$P(x_1 ... x_n | n) = \text{Pois}(n | \mu S + B) \left[\prod_{e=1}^{n} \frac{\mu S f_S(x_e) + B f_B(x_e)}{\mu S + B} \right]. \tag{4.1.}$$

Because logarithms are strictly increasing functions, maximizing the likelihood is equivalent to maximizing the log-likelihood. But for practical purposes, it is more convenient to work with the log-likelihood function in maximum likelihood estimation, because of desirable mathematical properties that make optimization easier and numerical computation more stable, etc. The logarithm of the likelihood function given $\text{Pois}(n|\nu) = \nu^n e - \nu / n!$ can be written as

$$\begin{aligned} -\ln L(\mu) &= -n \ln(\mu S + B) + (\ln S + B) + \ln n! - \left[\prod_{e=1}^{n} \frac{\mu S f_S(x_e) + B f_B(x_e)}{\mu S + B} \right] \\ &= (\mu S + B) + \ln n! - \sum_{e=1}^{n} \ln[\mu S f_S(x_e) + B f_B(x_e)]. \end{aligned} \tag{4.2.}$$

Since histograms are almost the universal practice for binning data in particle physics, we need to think of the binned equivalent of the model given above in Eq. 4.2. If we denote the signal and background histograms as ν_b^{sig} and ν_b^{bkg} respectively, where b is the bin index. The probability density function can be written as

$$f_S(x_e) = \frac{\nu_b^{sig}}{S \Delta_{b_e}}, \quad f_B(x_e) = \frac{\nu_b^{bkg}}{B \Delta_{b_e}}, \tag{4.3.}$$

where b_e is the index of the bin containing x_e and Δ_{b_e} is the width of that same bins. The probability model becomes a product over Poisson distributions for

each bin of the histogram as follows

$$P(n_b|\mu) = \text{Pois}(n_{tot}|\mu S+B)\left[\prod_{b\epsilon\text{bins}} \frac{\mu\nu_n^{sig}+\nu_b^{bkg}}{\mu S+B}\right]$$
$$= N_{\text{comb}} \prod_{b\epsilon\text{bins}} \text{Pois}(n_b|\mu\nu_b^{sig}+\nu_b^{bkg}), \quad (4.4.)$$

where n_b is the data histogram and N_{comb} is a combinatorial constant factor.

Unfortunately, in practice, the calculation of the MLE is not so simple, and in most cases, it is evaluated with additional parameters that are not of primary interest (i.e. μ) but must be taken into account during inference. These additional parameters are referred to as *nuisance parameters* which may include imperfections in the modeling of the physical processes employed by the simulation (such as the use of a leading order approximation for the hard scattering process), limited precision in the description of detector response (e.g. due to imperfect knowledge of the relevant calibration constants), uncertainty in fundamental physics parameters liable to condition the observations (e.g. the mass of a decaying particle), or simply a consequence of the finiteness of the number of simulated observations in a given region of feature space.

The challenge with nuisance parameters is that they enlarge the model and can complicate the estimation process. They are also the main factors limiting the accuracy and discovery range of high-energy physics analyses, nevertheless, they allow us to model the uncertainties of not having an exact model for the data. One common approach to handling nuisance parameters in MLE is to marginalize them out, we simplify the estimation problem by removing the dependence on parameters that are not of primary interest. This involves integrating the nuisance parameters from the likelihood function to obtain a marginalized likelihood function that depends only on the parameters of primary interest. To understand this concept, let us consider a scenario where we have a likelihood function that depends on both the parameters of primary interest, denoted as μ, and the nuisance parameters, denoted as θ. The joint likelihood function is typically expressed as $L(\mu,\theta|\text{data})$, representing the likelihood of the observed data given the parameters. To marginalize out the nuisance parameters, we integrate the joint likelihood over all possible values of the nuisance parameters. Mathematically, this can be represented as:

$$L_{\text{marginalized}}(\mu|\text{data}) = \int L(\mu,\theta|\text{data})d\theta \quad (4.5.)$$

The integral is taken with respect to the nuisance parameters θ, and it represents the sum or integration over all possible values of the nuisance parameters. The marginalized likelihood, $L_{\text{marginalized}}(\mu|\text{data})$, is the likelihood function that only

depends on the parameters of primary interest, μ. It is obtained by effectively "averaging out" the nuisance parameters, accounting for all possible values they can take. Once the marginalized likelihood is obtained, the estimation of the parameters of primary interest, μ, can be carried out using standard MLE techniques. Nevertheless, the process of marginalization can be analytically challenging or computationally intensive, depending on the complexity of the model and the number of nuisance parameters involved.

In a generalization of the above model, CMS uses HistFactory to build parametrized probability density functions (PDFs) in the RooFit [108]/RooStats [109] frameworks based on simple ROOT histograms that are sufficiently flexible to describe any analyses in the form

$$P(n_c,x_e,a_p|\phi_p,\alpha_p,\gamma_p) = \prod_{c\epsilon\text{channels}} \left[\text{Pois}(n_c|\nu_c)\prod_{e=1}^{n_c} f_c(x_e|\alpha)\right] \times G(L_0\lambda,\Delta_L) \times \prod_{p\epsilon S+\Gamma} f_p(a_p|\alpha_p), \tag{4.6.}$$

where the following convention is used, $e\ \epsilon$ events, $b\ \epsilon$ bins, $c\ \epsilon$ channels, $s\ \epsilon$ samples, and $p\ \epsilon$ parameters. $G(L_0\lambda,\Delta_L)$ denote the value of the goodness-of-fit test statistic G for a specific observed dataset, given a particular set of model parameters λ and the null hypothesis L_0. $f_p(a_p|\alpha_p)$ is a constraint term describing an auxiliary measurement a_p that constrains the nuisance parameter α_p. S are the parameters associated with systematics that have external constraints, and Γ are the bin-by-bin uncertainties with constraints. The channel is a region of the data for a given event selection, in which all channels must have disjoint event selection requirements. Finally, the expected number of events in a given bin is expressed as follows

$$\nu_{cb}(\phi_p,\alpha_p,\gamma_b) = \lambda_{cs}\gamma_{cb}\eta_{cs}(\alpha)\sigma_{csb}(\alpha), \tag{4.7.}$$

which implies the following probability density

$$f_c(x_e|\phi_p,\alpha_p,\gamma_b) = \frac{\nu_{cb_e}}{\nu_c} \quad \text{with} \quad \nu_c = \sum_{b\epsilon\text{bins of channel c}} \nu_{cb}, \tag{4.8.}$$

where λ_{cs} is the luminosity for a given channel and sample, γ_{csb} is the bin-by-bin scale factor used for statistical uncertainties[1], ϕ_{cs} is the product of unconstrained normalization factors for a given sample within a given channel[2], and finally η_{cs} the parameterized normalization uncertainties for a given sample within a

[1] For samples that do not have any bin-by-bin scale factors $\gamma_{csb} = 1$ and the index s can be omitted.

[2] These typically include the parameter of interest, e.g. the signal cross-section or branching ratio.

given channel.

In summary, the generalized form of Eq. 4.6 taking into account the binning of a given histogram, channels, and all systematic uncertainties, can be rewritten as follows

$$P(n_c, x_e, a_p|\phi_p, \alpha_p, \gamma_p) = \prod_{c\epsilon\text{channels}} \prod_{b\epsilon\text{bins}} \text{Pois}(n_{cb}|\nu_{cb}) G(L_0\lambda, \Delta_L) \prod_{p\epsilon S+\Gamma} f_p(a_p|\alpha_p). \tag{4.9.}$$

4.3. Discovery as a statistical test

In particle physics, we frequently quantify the significance of an observed signal by quoting the p-value of the background-only hypothesis. One method for defining the p-value for a hypothesized value of μ is to construct a test statistic (q-value) based on the profile likelihood ratio. The parameter of interest is the overall signal strength factor μ, which acts as a scaling to the total rate of signal events. We often write $\mu = \sigma/\sigma_{SM}$, where σ_{SM} is the standard model production cross-section. The signal strength is called so that $\mu = 0$ corresponds to the background-only model and $\mu = 1$ is the standard model signal. It is convenient to separate the full list of parameters α into the parameter of interest μ and the nuisance parameters θ: $\alpha = (\mu,\theta)$.

For a given data set and values for the global observable, there is an associated likelihood function $L(\mu,\theta)$ derived from the combined model over all the channels including all the constraint terms as was defined earlier in sec. 4.1. Let's assume $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximize the likelihood function $L(\mu,\theta)$ or equivalently minimize $-\ln L(\mu,\theta)$ and $\hat{\hat{\theta}}(\mu)$ is a conditional maximum likelihood estimate of θ that maximizes the likelihood function with μ fixed; this procedure for choosing specific values of the nuisance parameters for a given value of μ, thus the dataset and the global observables, is often referred to as "profiling". Similarly $\hat{\hat{\theta}}(\mu)$ is often referred to as "the profiled value of θ". A profiled likelihood ratio in the form

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})}, \tag{4.10.}$$

allows us to estimate the parameter of interest μ independently from the nuisance parameter θ which have been eliminated via "profiling". Despite the fact that the rate of signal events is non-negative, thus $\mu \geq 0$, it is often convenient to allow $\mu < 0$ (which indicates a deficit of events signal-like with respect to the background only) as long as the PDF, $f_c(x_c|\mu,\theta) \geq 0$ everywhere. Permitting $\mu < 0$ complicates the test statistic at the boundary, therefore we define the test

statistic $\widetilde{q}_\mu$[1] to differentiate the hypothesis of a signal being produced at a rate μ from the alternative hypothesis of signal events being produced at a lower rate:

$$q_\mu = \begin{cases} -2\ln\hat{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{4.11.}$$

The result of the test statistic is a single real number that represents the outcome of the experiment, but if one were to repeat the experiment multiple times, the test statistic would take on different values, so conceptually the test statistic has a distribution. Since the number of expected events $\nu(\mu,\theta)$ and the distributions of the discriminant variables $f_c(x_c|\mu,\theta)$ depend explicitly on θ the distribution of the test statistic will also depend on θ and each distribution of $f_c(\widetilde{q}_\mu|\mu,\theta)$ for a given test statistic will have an analogous expression to what is described in Eq. 4.11.

The p-value $p_{\mu,\theta}$ associated with a specific observation under a given hypothesis (μ,θ) represents the likelihood of obtaining an outcome as extreme or more extreme than the one observed, assuming the validity of that hypothesis. Consequently, a low p-value is interpreted as evidence contradicting the respective hypothesis.

$$p_{\mu,\theta} = \int_{\widetilde{q}_{\mu},obs}^{\infty} f(\widetilde{q}_\mu|\mu,\theta)d\widetilde{q}_\mu. \tag{4.12.}$$

A stumbling block is that we are interested in μ but the p-values depend on both μ and θ and in the frequentist approach the assumption $\mu = \mu_0$ would not be rejected unless the p-values of the θ global parameters are evaluated[2]. Fortunately, if we choose the test statistic to be based on the profile likelihood ratio, Wilks' theorem says that as the sample size approaches ∞, the distribution of the test statistic $-2\log(\lambda)$ asymptotically approaches the chi-squared (χ^2) distribution under the null hypothesis H_0, meaning that the distribution of the

[1] $\widetilde{q}_\mu$ is a test statistic for a one-sided alternative.

[2] μ_0 is often chosen as a reference or null hypothesis, and it serves as a baseline for comparison against other hypotheses or experimental observations. The goal is to assess whether the observed data supports or contradicts the assumption that μ takes on this specific value μ_0.
One can use the supremum p-value for overall θ to base the decision to accept or reject the hypothesis at $\mu = \mu_0$. The supremum p-value, addresses the issues of increased chance of obtaining at least one statistically significant result purely by chance when conducting multiple hypothesis tests, thus controlling the overall type I error rate (the probability of falsely rejecting a true null hypothesis). It is defined as the maximum p-value obtained from all the individual hypothesis tests in which it quantifies the strength of the most compelling evidence against the null hypothesis among all the comparisons made.

profile likelihood ratio $\lambda(\mu = \mu_{\text{true}})$ become independent of the values of the nuisance parameters, θ. Avoiding any residual dependence of the p-values on the choice of θ we would like to know the specific value of θ^{sup} that produces the supremum p-value over θ. Since larger p-values indicate better agreement of the data with the model, $\hat{\hat{\theta}}(\mu)$ is a good estimate of θ^{sup}. The p-value expression quantifying the hypothesis of signal strength μ can be written as follow:

$$p_\mu = \int_{\widetilde{q}_\mu, obs}^{\infty} f(\widetilde{q}_\mu|\mu, \hat{\hat{\theta}}(\mu, obs))d\widetilde{q}_\mu. \tag{4.13.}$$

A 95% confidence-level, one-sided frequentist confidence interval (upper limit) can be obtained by solving for $p'_{\mu_{up}} = 5\%$. The CL_s upper limit can be defined as the ratio $p'_\mu$of p-values

$$p'_\mu = \frac{p_\mu}{1 - p_b} \tag{4.14.}$$

where p_b is the p-value of the test statistic under the background-only hypothesis

$$p_b = 1 - \int_{\widetilde{q}_\mu, obs}^{\infty} f(\widetilde{q}_\mu|0, \hat{\hat{\theta}}(\mu = 0, obs))d\widetilde{q}_\mu. \tag{4.15.}$$

When performing searches, one usually converts the p-value into an equivalent significance Z, defined such that a Gaussian distributed variable, which is found Z standard deviations above its mean, has an upper-tail probability equal to p_0. This quantity can be defined as follows

$$Z = \phi^{-1}(1 - p_0), \tag{4.16.}$$

where ϕ^{-1} is the quantile (inverse of the cumulative distribution) of the standard Gaussian. In the absence of a signal, one would obtain a p-value of 0.5 for data in perfect agreement with the expected background. To establish a discovery, the standard discovery convention is $Z = 5$ (i.e. 5σ), corresponding to p-value $p_0 = 2.87.10^{-7}$. The critical α value and the p-value approach to Hypothesis Testing are illustrated in Fig. 4.1.

Expected sensitivity and bands

In many experimental analyses, expected sensitivity and bands refer to the anticipated performance and uncertainty estimation and they are used to represent the range of possible outcomes or measurements due to statistical and systematic uncertainties. CMS standard limit plot shows a dark green band corresponding to $\mu \pm 1$ (corresponding to 68% confidence level) defined by

$$\int_0^{\mu \pm 1} f(\mu_{up}|0, \hat{\hat{\theta}}(\mu = 0, obs))d\mu_{up} = \phi^{-1}(\pm 1). \tag{4.17.}$$

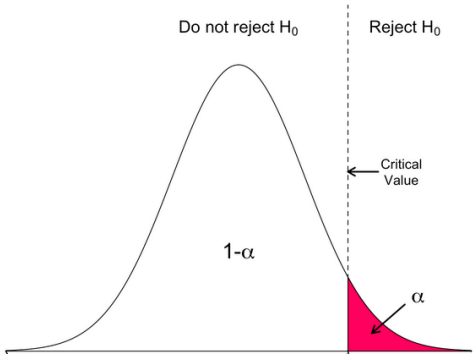


Figure 4.1. | For a right-tailed test, the null hypothesis is rejected if the test statistic is too large. Thus, the rejection region for such a test consists of one part, which is right from the center [110].

and a light yellow band corresponding to $\mu \pm 2$ (corresponding to 95% confidence level) defined by

$$\int_0^{\mu\pm 2} f(\mu_{up}|0, \hat{\hat{\theta}}(\mu = 0, obs))d\mu_{up} = \phi^{-1}(\pm 2). \tag{4.18.}$$

An example of such a plot is shown in Fig. 4.2.

4.3.1. Confidence interval

If we want to exclude the presence of a possible signal at a desired confidence level we often combine the results of several searches, particularly in the scenario where the sensitivity is limited by the collected luminosity, and not by any kinematic boundary. In statistics, confidence intervals and confidence limits are related but distinct concepts. A confidence interval is a range of values that is constructed based on sample data to estimate unknown parameters and quantify uncertainties associated with measurements or model predictions, such as a 95% confidence interval would indicate that if the experiment were repeated many times, 95% of the resulting intervals would contain the true value of the parameter. On the other hand, confidence limits refer to the lower and upper boundaries of the confidence interval.

In the following, I intend to explain how the lower and upper bound is derived with the so-called CL_s method, why this method is used, and how to interpret the result.

The test-statistic Q is constructed to increase monotonically for increasingly signal-like (decreasingly background-like) experiments so that the confidence in the signal+background hypothesis is given by the probability that the test statistic is less than or equal to the value observed in the experiment, Q_{obs}:

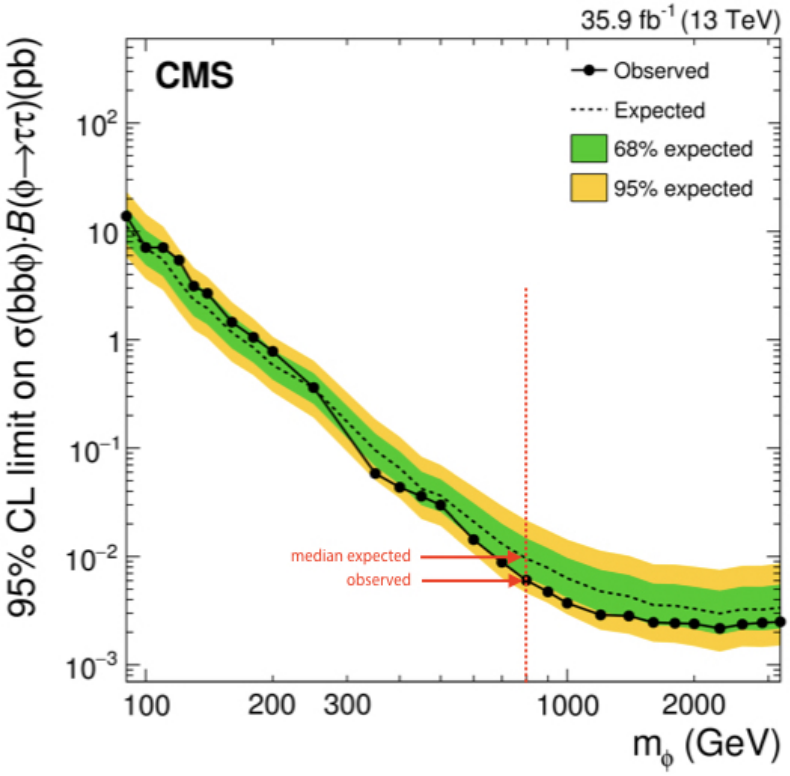


Figure 4.2. The expected limit of $\sigma \times B$ limits for the $bb\phi \to \tau\tau$ process is shown under the background-only hypothesis, with $\phi = (H \text{or} A)$ in the context of the 2HDM. The median value under this hypothesis, as well as the 68% and 95% intervals of $\mu \pm 1$ and $\mu \pm 2\ \sigma$ bands respectively [111].

$$CL_{s+b} = P_{s+b}(Q \leq Q_{obs}) = \int_{-\infty}^{Q_{obs}} \frac{dP_{s+b}}{dQ} dQ, \tag{4.19.}$$

where $\frac{dP_{s+b}}{dQ}$ is the probability distribution function of the test-statistic for signal+background hypothesis, thus smaller values of CL_{s+b} is in favor of the background hypothesis, indicate poor compatibility with the signal+background hypothesis. Similar to Eq. 4.19, the background hypothesis is given by the probability that the test statistic is less than or equal to the observed Q_{obs}.

$$CL_b = P_b(Q \leq Q_{obs}) = \int_{-\infty}^{Q_{obs}} \frac{dP_b}{dQ} dQ, \tag{4.20.}$$

where $\frac{dP_b}{dQ}$ is the PDF of the test-statistic for the background-only hypothesis, thus $CL_b \simeq 1$ is in favor of the signal+background hypothesis and indicates poor

compatibility with the background hypothesis.

During the process of hypothesis testing, there are two potential kinds of errors that can occur, often termed Type I and Type II errors. Type I error (denoted as α and calculated as $P(\text{reject}H_0|H_0 = \text{True})$) signifies the probability of incorrectly rejecting the null hypothesis (H_0) when it is, in fact, true. This is commonly known as a "False Positive". Conversely, Type II error (denoted as β and calculated as $P(\text{accept}H_0|H_0 = \text{False})$) is referred to as a "False Negative". It occurs when the null hypothesis is accepted even though the alternative hypothesis H_{alt} is true. These errors are problematic when striving to find statistical evidence. They can only be considered reasonable if there is substantial evidence against the null hypothesis, with a low probability α of occurrence when the null hypothesis is true, and a significantly higher probability $(1-\beta)$ when the alternative hypothesis is true.

We define the power of the statistical test as $1-\beta \equiv P(\text{reject}H_0|H_{alt} = \text{True})$. Fig. 4.3 show the dependence of power on the true mean of the alternative hypothesis: the closer it is to the mean of the null hypothesis the less power to reject H_0 when it is false. At its limit, when both means coincide, the power becomes identical to the α risk.

In a situation where the two distributions (i.e. H_0 and H_{alt}) are very close, the statistic power is very small and the experiment has no sensitivity to rejecting $s+b$ hypothesis because it almost rejects the background only hypothesis as well. The CL_s [107] procedure was introduced to handle type I and type II errors, preventing excluding or discovering signals that a search is in fact not sensitive to, due to a downward fluctuation of the background for example. It is a modified frequentist to obtain conventional bounds on the signal hypothesis which is defined as follows

$$CL_s \equiv CL_{s+b}/CL_b. \tag{4.21.}$$

This procedure is so conservative as if the experiment was performed in the complete absence of background or with absolute certainty of selected events due to background processes. As a consequence, both the false exclusion rate becomes much lower, and similarly, the use of $(1-CL_b)/(1-CL_{s+b})$ instead of $1-CL_b$ for discovery causes the false discovery rate to be lower than the significance level mentioned in Eq. 4.16.

4.3.2. Goodness of fit

The term goodness-of-fit refers to a statistical test determining how well our data fits a hypothesized distribution. Common statistic tests for goodness of fit are; Chi-square, Kolmogorov-Smirnov, Anderson-Darling, and Shapiro-Wilk, among others. The chi-square is one of the most common tests which is also

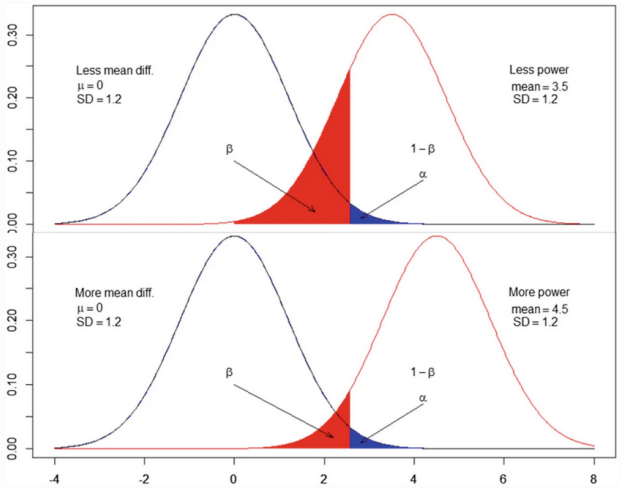


Figure 4.3. | Demonstration of how the means of both the null and alternative hypothesis distributions influence the statistical power $(1-\beta)$. The blue curve on the left represents the mean and standard deviation (SD) of a variable of interest based on the null hypothesis of no difference $(\mu = 0)$; the red curve on the right represents the distribution of a sample mean with SD based on the alternative hypothesis of 3.5 (top) and 4.5 (bottom) difference in the mean from the null. A greater disparity in means corresponds to higher statistical power. Conversely, a larger SD, indicating a broader distribution, leads to increased overlap between the distributions. This overlapping diminishes statistical power [112].

used in Chap. 5. It determines if an observed categorical dataset follows an expected theoretical distribution and compares the observed frequencies in different categories to the expected frequencies specified by the theoretical distribution

$$\chi^2 = \sum_{i}^{k=1} (O_i - E_i)^2 / E_i,$$

where O_i = an observed count for bin i, E_i an expected count for bin i, asserted by the null hypothesis.

For a chi-square goodness-of-fit test, we need to set a desired alpha level of significance such as 0.05 (i.e. 5%) which is the probability of making the wrong decision when the null hypothesis is true. The use of significance levels during hypothesis testing helps us determine which hypothesis the data support. If the p-value is less than the significance level, we can reject the null hypothesis and conclude that the difference between the true population and the null hypothesis value is statistically significant. One common limitation of such a test is that

the sample size should be sufficiently large so that the expected frequencies in each category are not too small. If the expected frequencies are too small, the chi-square test may not be valid and alternative methods such as the ones mentioned above should be considered.

4.4. Machine Learning

Machine learning has become an increasingly crucial part of the High Energy Physics (HEP) program; the vast amounts of data that HEP experiments collect can no longer be analyzed without the help of machine learning algorithms that can efficiently extract meaningful features from high-dimensional data and identify subtle signals buried in noisy backgrounds. Additionally, because signals of interest can be very rare and difficult to detect, machine learning algorithms can be trained to recognize specific signatures of new physical events, allowing for more efficient signal detection and classification. They can help identify rare events and reduce false positives, improving the sensitivity of experiments.

The usage of machine learning involves creating a model that is trained on some training data but can still process additional data to make predictions. Artificial Neural Networks (ANNs) and Boosted Decision Trees (BDT) are such types of models. The focus will be on ANN more specifically on Deep Neural Networks (DNNs) which are a class of ANNs that aim to mimic the information processing of the brain.

Supervised or unsupervised learning is often used in machine learning which represents techniques that differ in how the model learns and makes predictions based on data. Supervised learning involves training a machine learning model using labeled data, where input data points are paired with their corresponding desired outputs or labels. The goal is for the model to learn the underlying mapping or relationship between input data and output labels and it is commonly used in tasks such as classification and regression. On the other hand, unsupervised learning involves training a machine learning model with all prior knowledge of the data using unlabeled data. There are also other types of learning paradigms, such as semi-supervised learning and reinforcement learning, but this section will only focus on DNN supervised learning which is the learning type used to search for $H/A \to ZA/H \to \ell^+\ell^- b\bar{b}$, reported in Chap. 5.

DNNs have more than one hidden layer situated between the input and output layers. Each layer contains a given number of neurons that apply a specific functional transformation to the input. There are different types of neural networks but they always consist of the same components: neurons, synapses, weights, biases, and functions with similar architecture to the one shown in Fig. 4.4.

The following terms collectively form the foundational components of a deep neural network and are essential for understanding the behavior and functionality

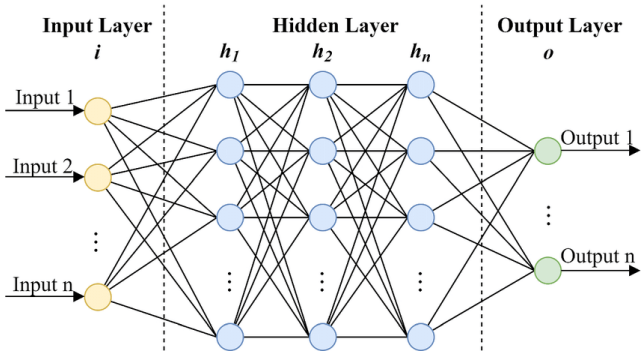


Figure 4.4. General Neural Network Architecture. The output of a unit (i) in a layer (l) is related to the output (o) of the earlier layer with outputs through a set of weights, a bias, and a non-linear activation function [113].

of the network during training and inference.

- **Neurons:** Also known as nodes or units, in neural networks nodes receive inputs, perform computations on them, and produce an output. Neurons are typically organized into layers, such as the input layer, hidden layers, and output layer.
- **Synapses:** Synapses represent the connections between neurons in a neural network. They are analogous to the connections between biological neurons. Synapses transmit signals or information from one neuron to another and they are represented by weights.
- **Weights:** Weights are parameters associated with the synapses in a neural network that play a crucial role in determining the output of a neuron based on its inputs. They determine the strength or influence of the connection between two neurons. Each synapse (connection) is assigned a weight value, which is adjusted during the learning process.
- **Biases:** Biases are additional parameters in a neural network that are associated with each neuron. Biases allow the network to introduce an offset or shift in the activation function, helping in better fitting the data. They also provide flexibility to the model, allowing it to learn and represent different patterns and trends in the data.
- **Activation Functions:** Activation functions are mathematical functions applied to the output of each neuron (or node) based on its weighted sum of inputs in a layer. They introduce nonlinearity into the calculations performed by neurons, allowing the network to model complex relationships between inputs and outputs. It decides whether the input of a neuron in the network is essential or not in the prediction process.
  The choice of the activation function is a critical decision in the design

of a neural network as it affects the learning ability of the network, its convergence speed, and its ability to handle different types of data. There are different types of activation functions that totally depend on the problem to be solved and the network architecture used. The most common are

- **Sigmoid:** The Sigmoid function is defined as $f(x) = 1/(1+e^{-x})$ and illustrated in Fig. 4.5 (a). It squashes the input values between 0 and 1, which makes it useful for binary classification problems. However, it can suffer from the "vanishing gradient" problem during backpropagation.
- **Tanh (Hyperbolic tangent):** The tanh function is defined as $f(x) = (e^{x} - e^{-x})/(e^{x} + e^{-x})$ as illustrated in Fig. 4.5 (b). It squashes the input values between -1 and 1, providing a stronger non-linearity than the sigmoid function, but also can suffer from the vanishing gradient problem.
- **ReLU (Rectified Linear Unit):** The ReLU function is defined as $f(x) = max(0,x)$ as illustrated in Fig. 4.5 (c). It sets all negative input values to zero and keeps the positive values unchanged. ReLU has become very popular in deep learning due to its simplicity and ability to alleviate the vanishing gradient problem.
- **Leaky ReLU:** The Leaky ReLU function is an extension of the ReLU function that introduces a small negative slope for negative input values as illustrated in Fig. 4.5 (d), preventing neurons from completely "dying" during training.
- **Softmax:** The softmax function is used in the output layer for multi-class classification problems. It takes a vector of arbitrary real values as input and converts them into a probability distribution, where the sum of the outputs is equal to 1. It is often used in conjunction with the cross-entropy loss function.

4.4.1. Improving accuracy: The black box problem

Deep neural networks (DNNs) have been always a powerful tool for achieving accuracy in various tasks, but they are also often referred to as black boxes because they can be challenging to interpret and understand. The nonlinearity of a large number of connections between layers makes it difficult and almost impossible for humans to trace back how the input information is processed and transformed into output predictions. DNNs operate in high-dimensional spaces that lack transparency, human-like intuition, and any explicit rule. Nevertheless, DNNs have demonstrated exceptional performance in a wide range of tasks, such as image recognition, natural language processing, and speech recognition. They have surpassed human-level performance in many cases and have demonstrated their ability to learn complex patterns and make accurate predictions. The

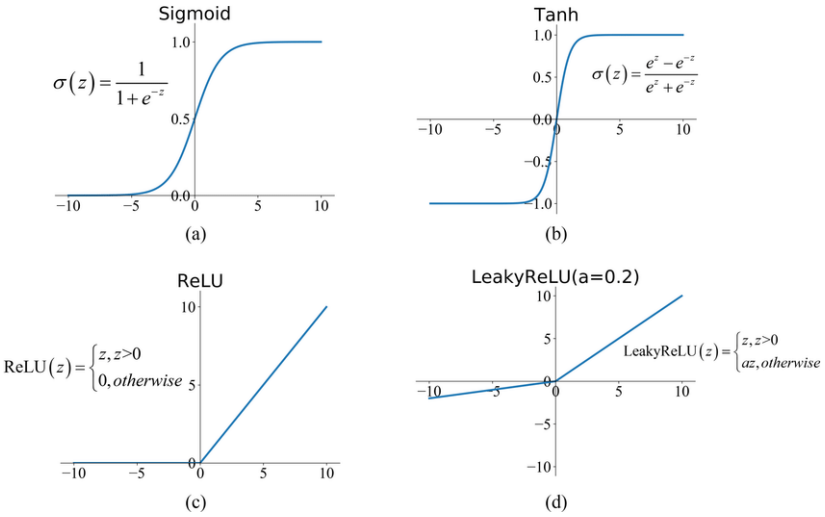


Figure 4.5. | Commonly used activation functions [114]: (a) Sigmoid, (b) Tanh, (c) ReLU, and (d) LReLU.

empirical evidence of their effectiveness and superior performance builds trust in their capabilities.

Deep neural networks undergo rigorous validation and testing procedures to assess their performance and reliability. We can evaluate the performance of a neural network using various metrics and techniques; confusion matrix, ROC and AUC curve, cross-validation, loss function, precision, recall, and F1 score, etc. are such examples. The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) (see Fig. 4.6) are the most commonly used techniques to get insight into the true positive rate against the false positive rate at different classification thresholds. It provides a comprehensive view of the trade-off between sensitivity and specificity. The area under the ROC curve (AUC) is often used as a single metric to quantify the overall performance of a classifier.

The confusion matrix in Fig. 4.7 also provides the counts of true positives, true negatives, false positives, and false negatives. From the confusion matrix, several metrics can be derived, including precision, recall, and F1 score[1].

Cross-validation is another technique used to assess the model's generalization performance by splitting the dataset into multiple subsets (folds). The neural network is trained and evaluated on different combinations of these folds, allowing

[1]Precision measures the proportion of correctly predicted positive instances out of the total predicted positives. Recall, also known as sensitivity, calculates the proportion of correctly predicted positive instances out of the total actual positives. The F1 score is the harmonic mean of precision and recall, providing a balanced measure of a classifier's performance.

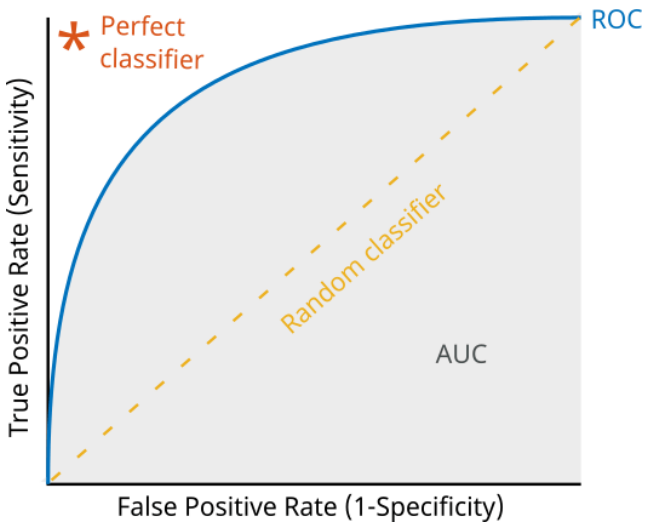


Figure 4.6. | Receiver operating characteristic (ROC) curve. The area under the curve (AUC) reflects the performance of the classifier (i.e. a classifier with optimal performance has an area under the curve close to 1) [115].

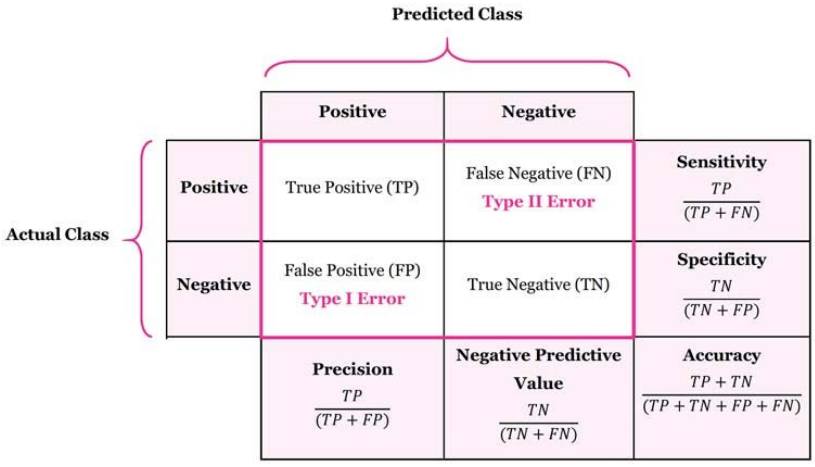


Figure 4.7. | Confusion matrix and performance equations [116].

for more robust performance estimation. During training loss function can be very insightful into the optimization and learning process. Monitoring the training and validation loss can give an indication of how well the model is learning and whether it is overfitting or underfitting the data.

While using one of the above techniques is sufficient to provide insight into the performance of the trained model, it is still useful to consider several of them to

gain an overall understanding of network performance, but all still depend on the nature of the problem, available data set, and specific goals of the neural network application.

Chapter 5. Search for $\ell^+\ell^- b\bar{b}$ signal through $H/A \to ZA/ZH$ process

After the discovery of the Higgs bosons at the Large Hadron Collider (LHC), many questions were left open. One of the remaining questions is - if the discovered scalar particle at 125 GeV mass is part of an extended scalar sector. In the Standard Model (SM), there is only one Higgs doublet, which gives rise to a single physical Higgs boson. However, many extensions of the SM, such as the Two Higgs Doublet Model (2HDM), propose the existence of additional scalar fields. The 2HDM introduces a second Higgs doublet, leading to the presence of five physical Higgs bosons after Electroweak Symmetry Breaking (EWSB), as was discussed in Chap. 2: the two CP-even Higgs bosons h and H, the CP-odd pseudo-scalar A, and the pair of the charged Higgs bosons, $H^{\pm}$.

The 2HDM allows for regions of the parameter space where the light neutral scalar, h, has the same properties as the observed Higgs at the LHC with an equal 125 GeV mass. To confirm or disprove this possibility, an investigation would involve detailed measurements of the properties and couplings of the discovered Higgs boson, as well as searches for the other predicted Higgs bosons in the extended scalar sector.

This work pursues a direct search for the neutral Higgs bosons, H/A, in order to test such a theory. The lightest scalar Higgs boson, h, in this search, is assumed to be SM-like with a mass set to the experimentally measured value of 125 GeV. H/A, are looked for in a signature composed of Z and lighter Higgs bosons, A/H, where the Z bosons decay into a pair of leptons of opposite charge and of the same flavor (muons or electrons), $\ell^+\ell^-$, and the lighter A/H bosons decay into a pair of resolved $b\bar{b}$ jets or a two collimated b tagged jets, if it happened to be reconstructed as a single boosted fat b jet.

Previous searches for $H/A \to ZA/ZH$ and $A \to Zh$ signatures in the same final state ($\ell^+\ell^- b\bar{b}$) have been performed before by both the ATLAS and the CMS experiments [117–119], and those have been reported in detail in Sec. 2.5 of Chap. 2.

This chapter reports the analyzed data using full run 2 LHC proton-proton collision, collected by the CMS experiment at $\sqrt{s} = 13$ TeV, which corresponds to an integrated luminosity of 138 fb^{-1}. This analysis focuses on the invariant mass distributions of the $\ell^+\ell^- b\bar{b}$ and $b\bar{b}$ systems, in both gluon-gluon fusion

and b-associated production mechanisms, searching for a resonant-like excess compatible with the H and A masses kinematically allowed in the 2HDM theory.

The free parameters of the model allow several hierarchies and branching fractions, as shown in Fig. 5.1. The 2HDM theory, predicts that the light CP-even, h can be SM-like, in the phase space where the 2HDM parameter $\cos(\beta-\alpha)=0$, known also as the alignment limit. At low $\tan\beta \approx 1$ the branching fractions of the heavier neutral Higgs boson $H \to ZA$ and $A \to b\bar{b}$ via gluon fusion production are among the largest.

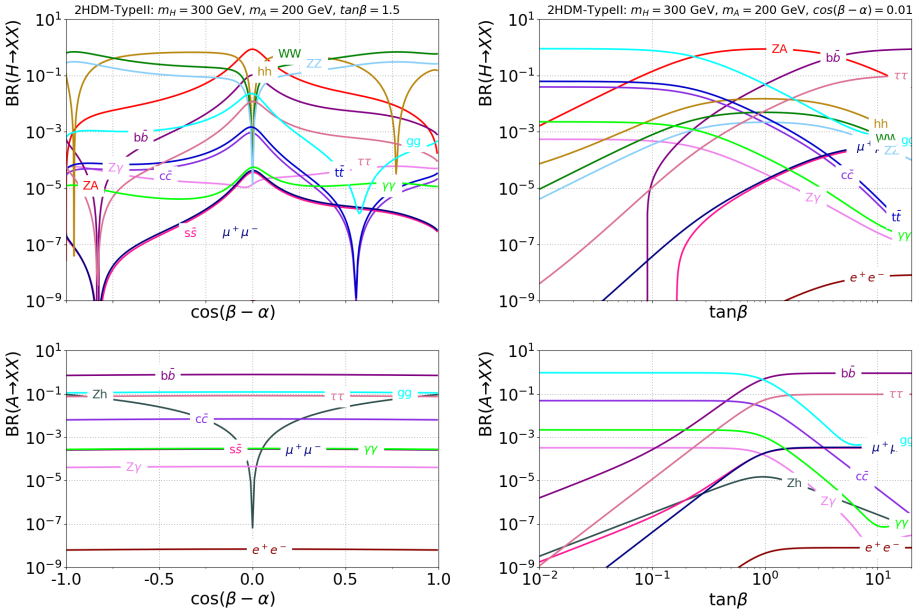


Figure 5.1. | H and A branching fractions as a function of $\cos(\beta-\alpha)$ (left) and $\tan\beta$ (right) for the following set of parameters: $\tan\beta = 1.5$ ($\cos(\beta-\alpha) = 0.01$) left (right), $m_H = 300$ GeV, and $m_A = 200$ GeV. The plots are personal work and were created using the 2HDMC-1.8.0 tool [120].

At low $\tan\beta$ the decay channel $A/H \to Z(\to \ell^+\ell^-)H/A(\to b\bar{b})$ only dominates below $2m_{\text{top}}$ as illustrated in Fig. 5.2. The signal is expected to be observed mainly as a result of the gluon-gluon fusion production.

In type-II 2HDM the most important contribution to the total cross-section is given by resonant production of the heavy neutral Higgs via gluon-gluon fusion (ggH) production and b-associated production (bbH). The latter is subdominant at low values of $\tan\beta$ and becomes the dominant contribution as $\tan\beta$ increases, this is demonstrated for various benchmarks of (m_H, m_A) in Fig. 5.3 and further draw the conclusion about the need to take into account this process

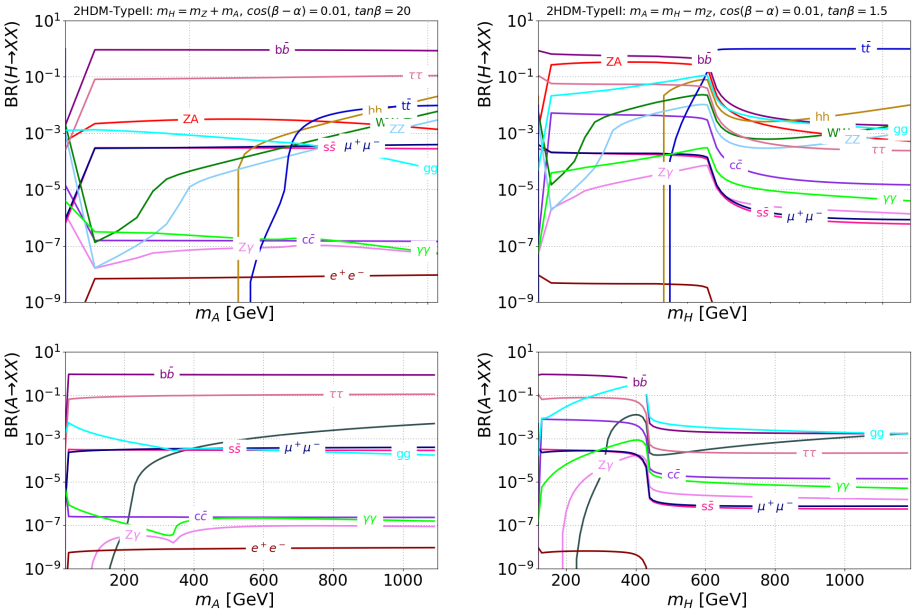


Figure 5.2. | H and A branching fractions as a function of m_A (left) and m_H (right) for the following set of parameters: $\cos(\beta-\alpha)=0.01$, $\tan\beta=20$ (1.5) left (right). The plots are personal work and were created using the 2HDMC-1.8.0 tool [120].

at a large value of $\tan\beta$. Henceforth, these parameters set the phase space of the 2HDM for this search.

5.1. CMS event data for physics analysis

5.1.1. The NanoAOD format

In CMS, the typical data analysis flow usually involves several steps of data reduction and processing. At the end of LHC full run 2, huge datasets needed to be processed putting the majority of the physics analysis group under a big challenge. Thus, data reduction becomes a crucial step to ensure that a given set of analyses can still achieve their ultimate sensitivity with much flexibility. Fortunately, NanoAOD brings CMS to a new era, it achieves a compact data level that is 20 times smaller than its descendant MiniAOD format (was used for run 1). It consists of a Ntuple-like format reduced to 1-2 kB per event, readable with Root [122], and contains the per-event information that is needed in most generic analyses. Despite its small size, it retains all observed particles with a high level of detail on leptons, photons, jets, etc. It supports jet re-clustering, re-calibration, re-computation of b-tagging, jet substructure, and more.

In this search, events are processed up to the NanoAOD level using CMSSW

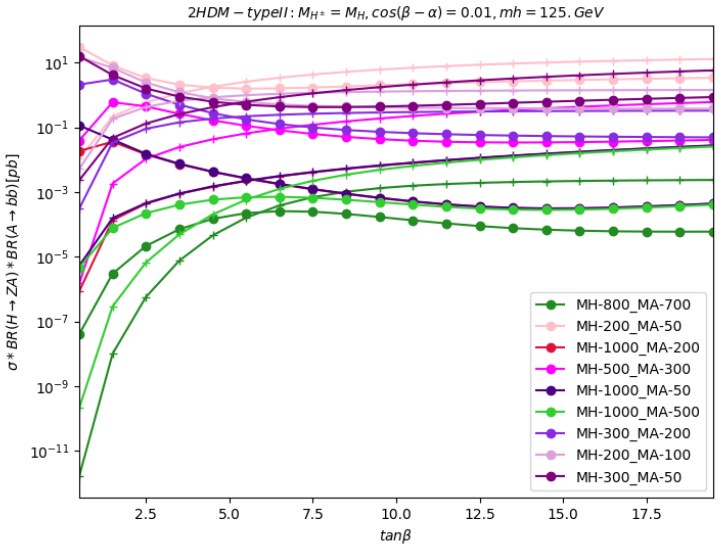


Figure 5.3. Cross section in pb multiplied by $BR(H \to ZA)$ and $BR(A \to b\bar{b})$ is shown for the 2HDM type-II near the alignment limit as a function of tanβ for different (m_H, m_A) benchmarks, (o) curves show the cross-section of the gg-fusion process, while (+) curves show the b-associated production. The plots are personal work, created using the 2HDMC-1.8.0 [120] and SusHi-1.7.0 [121] tools.

software. Events selection and histogram filling are performed using the Bamboo framework [123], an analysis library based on RDataFrame [124] which was developed at the Université Catholique de Louvain in early 2019. I began as Bamboo's initial and sole user, then took on key responsibilities for its maintenance and further development as it evolved. This transition was prompted by the departure of the primary developer of the framework within the CMS collaboration. This shift in responsibility constitutes a noteworthy technical contribution to this thesis.

5.1.2. Collected data and Monte-Carlo simulation

5.1.2.1. Data

The data used in this search were collected during LHC full run 2 of proton-proton collisions at a center of mass energy $\sqrt{s} = 13$ TeV which corresponds to an integrated luminosity of 138 fb^{-1}. The full list of data samples used is summarized in Tab. 5.1.

Table 5.1. | Data samples were used in the full analysis.

Year	Channel	Dataset	lumi [fb^{-1}]
2016	$\mu\mu$	DoubleMuon/Run2016***BCDEF(preVFP)-FGH(postVFP)***-UL2016_*/NANOAOD	19.66 preVFP
	$e\mu$	MuonEG/Run2016***BCDEF(preVFP)-FGH(postVFP)***-UL2016_*/NANOAOD	16.97 postVFP
	ee	DoubleEG/Run2016***BCDEF(preVFP)-FGH(postVFP)***-UL2016_*/NANOAOD	
		SingleElectron/Run2016***BCDEF(preVFP)-FGH(postVFP)***-UL2016_*/NANOAOD	
		SingleMuon/Run2016***BCDEF(preVFP)-FGH(postVFP)***-UL2016_*/NANOAOD	
2017	$\mu\mu$	DoubleMuon/Run2017***BCDEF***-UL2017_*/NANOAOD	41.53
	$e\mu$	MuonEG/Run2017***BCDEF***-UL2017_*/NANOAOD	
	ee	DoubleEG/Run2017***BCDEF***-UL2017_*/NANOAOD	
2018	$\mu\mu$	DoubleMuon/Run2018***ABCD***-UL2018_*/NANOAOD	59.74
	$e\mu$	MuonEG/Run2018***ABCD***-UL2018_*/NANOAOD	
	ee	EGamma/Run2018***ABCD***-UL2018_*/NANOAOD	
		SingleMuon/Run2018***ABCD***-UL2018_*/NANOAOD	

5.1.2.2. Signal

There are two main production mechanisms that were considered in this search: the dominant one is the contribution of gluon-gluon fusion production, where bottom-quark loop amplitudes interfere with top-quark loop amplitudes as illustrated in Fig. 5.4 (a); the second-largest is the associated production of b quarks.

For all mechanisms that feature b quarks at the hard-process level, there are two ways of performing the computation of b-associated production, which are usually called four- and five-flavour schemes (abbreviated with 4FS and 5FS respectively), further displayed in Fig. 5.4 (b) and (c). In 5FS the b quarks are part of the proton, $g \to b\bar{b}$ splitting with two b quarks that could possibly be seen in the forward region of the CMS detectors, while in the 4FS the b quark never appears in the initial state and in this case the Higgs boson responds to a gluon and recoils against it, with the gluon subsequently undergoing a division into a pair of bottom quarks [1]. These two schemes are supposed to address issues that arise in different kinematic regimes, a detailed comparison was addressed in [125].

One of the main objectives of this analysis is to attempt to b tagging the additional b associated quarks that appear in the b-associated process, as will be further discussed in Sec. 5.4. Given that, the mass of the b quark in the 5FS has to be regarded equal to zero ($m_b = 0$) as a technical mean to achieve the resummation of large logarithms, which results in non-consistent of the b-quarks

[1] i.e. the Higgs boson, which is being produced in the collision, can react with a gluon from one of the initial protons. The gluon and the Higgs boson are connected by a virtual interaction. As a result, the Higgs boson "responds" to the presence of the gluon, leading to the emission of a b quark pair from the gluon. The recoil refers to the change in momentum experienced by the Higgs boson due to this interaction.

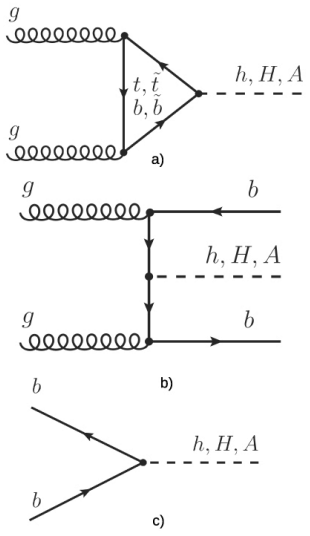


Figure 5.4. | Leading order diagrams of the a) gluon fusion and b-associated Higgs production in the b) four-flavour and c) five-flavour scheme [64].

with the b-tagged objects conversely to the 4FS. The 4FS diagram shown in Fig. 5.4(b) is, therefore, best suited for this search.

Simulated signal events of a heavy Higgs boson, H, decaying to Z bosons and a pseudo-scalar, A, were generated at both leading and next-leading order with `MADGRAPH5aMC@NLO` version 2.6.5, showered using `PYTHIA` 8.210 [126], and produced for both gluon-fusion loop-induced process and b-associated production in the 4 flavor-scheme.

The central values of the renormalization (μ_R) and factorization (μ_F) scales have been taken equal to the reference scale:

$$bb\phi : \mu = \frac{H_T}{4} = \frac{1}{4}\sum_i \sqrt{m_i^2 + p_t^2(i)},$$

$$gg\phi : \mu = \frac{H_T}{2}, \tag{5.1.}$$

where $\phi = (H, A)$ and the sum runs over all final-state particles at the hard-process level. The internal masses are set to their on-shell values m_h= 125 GeV, m_t = 172.5 GeV and m_b= 4.92 GeV. The top-quark Yukawa is renormalized on-shell, while for the bottom-quark Yukawa, we adopt the $\overline{MS}$ scheme, with a four-loop evolution $m_b(\overline{m}_b)$= 4.18 GeV. The parton distribution function (PDF) set used along with the CP5 tune is *NNPDF31_nnlo_as_0118_mc_hessian_pdfas*

next-to-next-to-leading order [127]. The signal samples were generated for H bosons of masses in the range of 125-1000 GeV and for narrow-width pseudoscalars (A) of masses in the range of 30-1000 GeV. In addition, benchmark points were generated for the twisted scenarios $A \to ZH$ and for $A \to ZH(125)$ end up covering a wide range of masses as illustrated in the search map in Fig. 5.5.

The idea behind selecting these specific points can be attributed to the initial focus of the analysis on $H \to ZA$, which was intended to be a continuity of the preceding 2016 CMS search [34]. However, there was a subsequent shift in perspective. Given that by reconstruction an analysis optimized for $H \to ZA \to \ell^+\ell^- b\bar{b}$ could be effectively repurposed for $A \to ZH \to \ell^+\ell^- b\bar{b}$ as well. This realization prompted the exploration of the latter channel.

On the other hand, generating extensive masses to cover the complete two-dimensional mass spectrum comes with considerable computational expenses. As a result, an alternate strategy was adopted. Instead of spanning the entire mass plane, the focus was on testing the hypothesis that the mass degeneracy and the spin-related aspects of H and A permit an extension of the search for the $A \to ZH \to \ell^+\ell^- b\bar{b}$ process. Therefore, a selection of a few mass points for $A \to ZH$, which were equally applicable in the twisted scenario was simulated to investigate this hypothesis. Additionally, some of these chosen points were motivated by the fact that were explored but not excluded by a previous ATLAS search [33].

Recall that in addition to the mass parameters in the Higgs basis, the other parameters of the model are the two mixing angles α and β which diagonalize respectively the sector of the Higgs boson CP-even and CP-odd. $\tan\beta = v_2/v_1$ is the ratio of the two Higgs vacuum expectation values, which is also related to SM *vev* by $\sqrt{(v_1^2 + v_2^2)}$. Also, in order to avoid flavor-changing neutral currents (FCNC) at the tree-level, a $\mathcal{Z}_2$ symmetry is imposed. If this symmetry were allowed to be softly broken in the Higgs potential, an extra 2HDM parameter m_{12} with a dimension of mass is included in the Higgs potential and the parameters $\lambda_{6,7} = 0$ are needed to ensure CP conservation at the tree-level. Four types of the 2HDM can then be realized, depending on how $\mathcal{Z}_2$ symmetry is implemented into the fermion sector: types I-IV (also detailed in Sec. 2.2 of Chap. 2).

The simulated signal samples for this analysis are generated for a type-II 2HDM near the alignment limit where $\cos(\beta - \alpha) = 0.01$, with $\tan\beta = 1.5$ (20.) for gluon-gluon fusion (b-associated production) process, and the mixing parameter [1] $m_{12}^2 = \sqrt{m_{H^\pm}^2 \frac{\tan\beta}{1+\tan\beta^2}}$. The choice of the different $\tan\beta$ values is driven by the

[1] The parameter m_{12} represents the off-diagonal element of the 2×2 mass-squared matrix that governs the mixing between the neutral scalar Higgs bosons (h and H) in the 2HDM. For more comprehensive details, refer to Chap. 2.

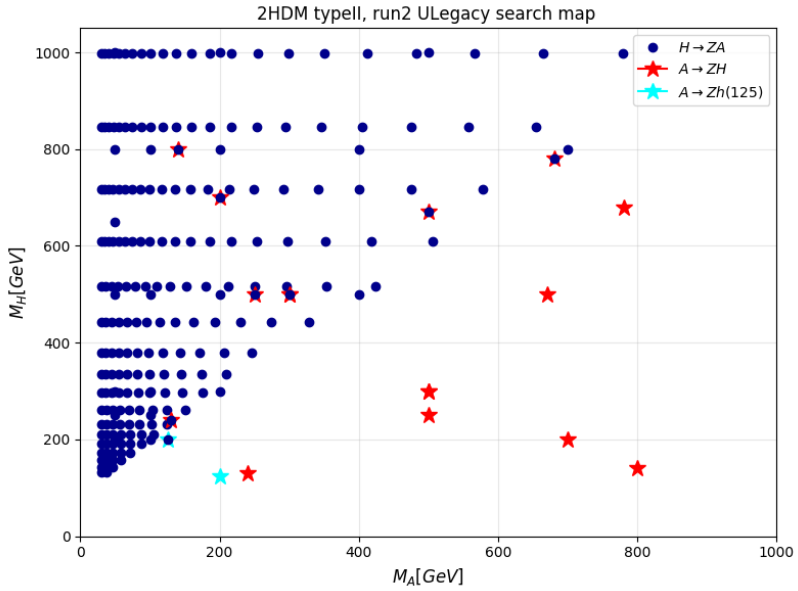


Figure 5.5. A two-dimensional search map for the simulated signal mass points (M_A, M_H) for the three scenarios: $H \to ZA$, $A \to ZH$, and for $A \to ZH(125)$.

fact that the branching fraction of $h_{SM-like} \to b\bar{b}$ start to diverge from the measured branching fraction of $H_{SM} \to b\bar{b}$, this is demonstrated in Fig. 5.6 and Fig. 5.7. As the goal is to conduct the search as close as possible to the alignment limit; $\tan\beta = 20$. seemed to be the optimal choice, to probe the $b\bar{b} \to H/A$ decay channel while keeping the CP even light Higgs h indistinguishable from that of the SM.

The significant predominance of the (bbH) process over (ggH) confirms the theoretical predictions of the prevalence of this process due to the increase in the Higgs Yukawa boson coupling constant for large values of $\tan\beta$. This search is an attempt to extend the previous CMS [34] search by exploiting this production mechanism at large $\tan\beta$ values that were not previously considered.

The mass points were derived to achieve good coverage in the two-dimensional mass plane (m_A, m_H) in terms of experimental mass resolution;
$m_A = m_A \times (1+\sigma_x)$, and $m_H = m_H \times (1+\sigma_y)$, with $m_A < m_H - m_Z$ (to ensure the decay $H \to ZA \to \ell^+\ell^-b\bar{b}$), and where σ_x and σ_y are the resolutions of the reconstructed m_{bb} and m_{llbb} distributions respectively for the signal masses under studies m_A and m_H.

All benchmark points were checked to be consistent with the potential stability, perturbativity, unitarity, and the current experimental limits on Higgs boson masses using 2HDMC-1.8.0 [120]. Similarly, the Narrow Width Approximation

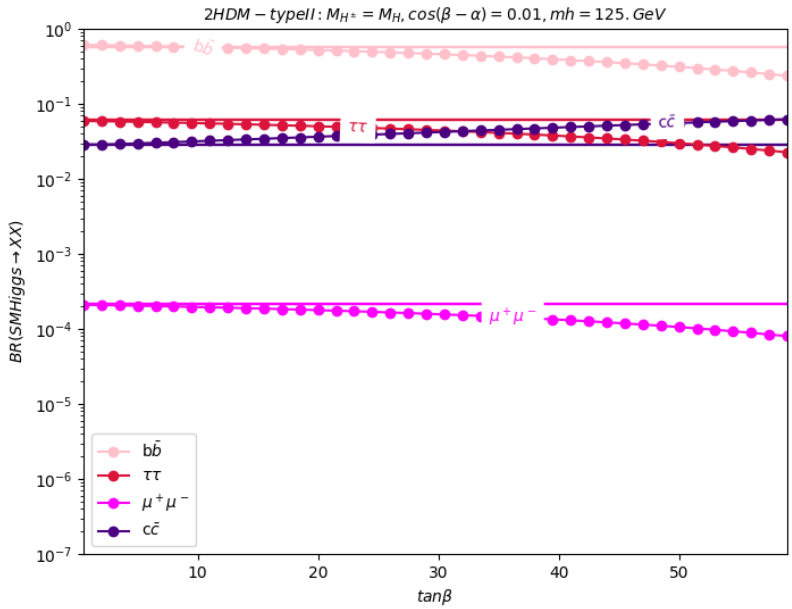


Figure 5.6. The predicted 2HDM SM-like CP even Higgs boson of mass $m_h = 125$ GeV and the observed SM Higgs branching ratios of the decays into pair of fermions ($b\bar{b}$, $\tau\tau$, $\mu^+\mu^-$, and $c\bar{c}$) are compared as a function of $\tan\beta$ and near the alignment limit $\cos(\beta-\alpha) = 0.01$ for $(m_H, m_A) = (500, 300)$ GeV benchmark. The SM-like (SM) Higgs is shown in the o $(-)$ curve.

(NWA) is still valid for all scenarios, and it is within our experimental resolution as shown in Fig. 5.8. The cross-sections for H/A boson production in the 2HDM are calculated using up to NNLO QCD corrections for gluon-gluon fusion and b-associated production in the four- and five-flavour scheme standard matching using SusHi-1.7.0 [121]. This is because, neither the 4FS nor the 5FS alone can accurately describe processes that span a wide range of energies such as those involving the production of a Higgs boson in association with a bottom quark. The standard matching procedure between these schemes is designed to smoothly combine their strengths and benefits in different kinematic regions, ensuring accurate predictions for cross-sections in processes involving bottom quarks [128]. Given that the distinction between the 4FS and 5FS is fundamentally characterized by a logarithmic behavior, the degree to which they influence each other concerning the Higgs boson mass should be governed by a logarithmic term i.e.

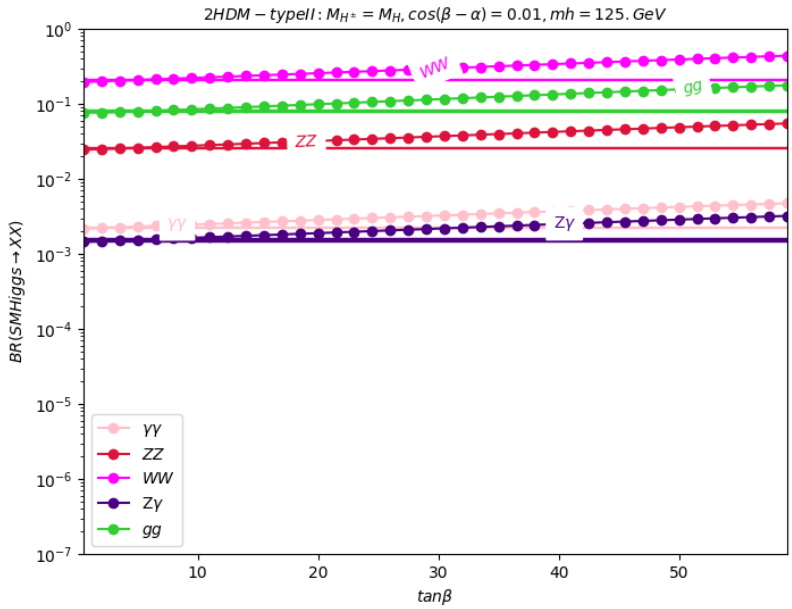


Figure 5.7. The predicted 2HDM SM-like CP even Higgs boson of mass $m_h = 125$ GeV and the observed SM Higgs branching ratios of the decays into pair of gauge bosons ($\gamma\gamma$, ZZ, WW, $Z\gamma$, and gg) are compared as a function of $\tan\beta$ and near the alignment limit $\cos(\beta-\alpha) = 0.01$ for $(m_H, m_A) = (500, 300)$ GeV benchmark. The SM-like (SM) Higgs is shown in the o $(-)$ curve.

$$\sigma_{\text{matched}} = \frac{\sigma_{4FS} + w\sigma_{5FS}}{1+w}, \tag{5.2.}$$

with the weight w defined as $w = \ln\frac{m_H}{m_A} - 2$. The theoretical errors are defined as follows

$$\Delta\sigma^{\pm}_{\text{matched}} = \frac{\Delta\sigma^{\pm}_{4FS} + w\Delta\sigma^{\pm}_{5FS}}{1+w}, \tag{5.3.}$$

The width of the 2HDM particles can vary significantly depending on the masses of the particles they can decay into and the couplings involved. In general, larger masses for the decay products will result in wider widths as illustrated in Fig. 5.9. Additionally, the specific 2HDM type and the values of the model parameters, such as the mixing angles and couplings, can also affect the widths.

When simulating signal events in the 2HDM, it was crucial to consider the relevant decay channels and the corresponding widths of the heavy neutral

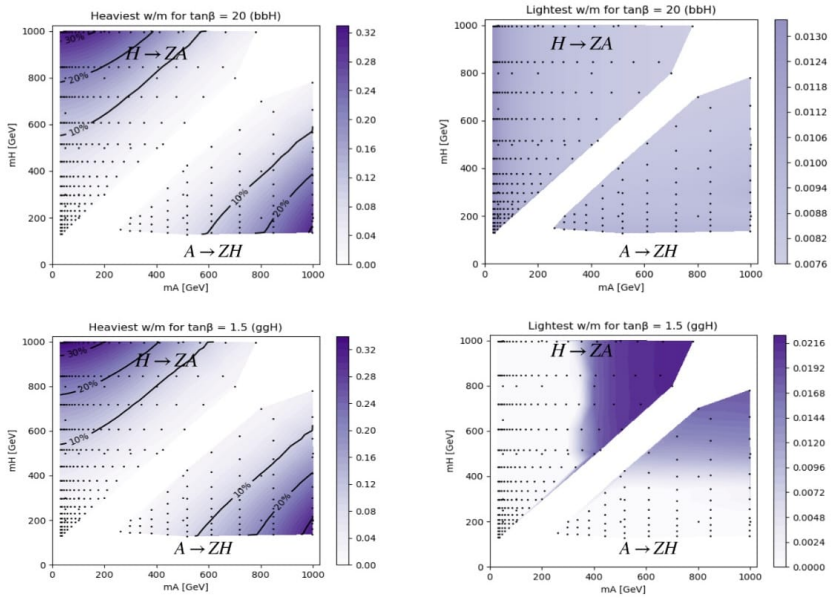


Figure 5.8. | Width/Mass plotted for the most massive (left) and least massive (right) among the (H, A) bosons throughout the mass spectrum. The (upper) plots represent b-associated production at a value of $\tan\beta = 20$, whereas the (lower) plots show gluon-gluon fusion with $\tan\beta = 1.5$. In both scenarios of $H \to ZA$ and $A \to ZH$, the regions where the relative width surpasses 20% are enclosed by the black contour area, confined to resonance masses greater than 800 GeV.

Higgs boson (H) and pseudoscalar (A). These widths influence the branching ratios of the decays and, consequently, the expected event rates for different final states.

5.1.2.3. Background

Several Standard Model processes will produce signatures in the detector very close to our signal and here is a list of the most dominant and irreducible ones.

Drell-Yann plus jets

The production of one Z/γ-boson in addition to at least two jets will be the most important background. At leading order, two production mechanisms are possible, the gluon-gluon case where a gluon from one proton interacts with a gluon from the other proton and the quark-antiquark case where a quark from one proton interacts with an anti-quark from the other proton. In this search, the most dominant, irreducible backgrounds are the production of Drell-Yan in association with heavy-flavor jets and the production of a top quark pair in the fully leptonic final state. The $gg \to Z(\ell^{\pm}\ell^{\mp})jj$ has the biggest cross-section

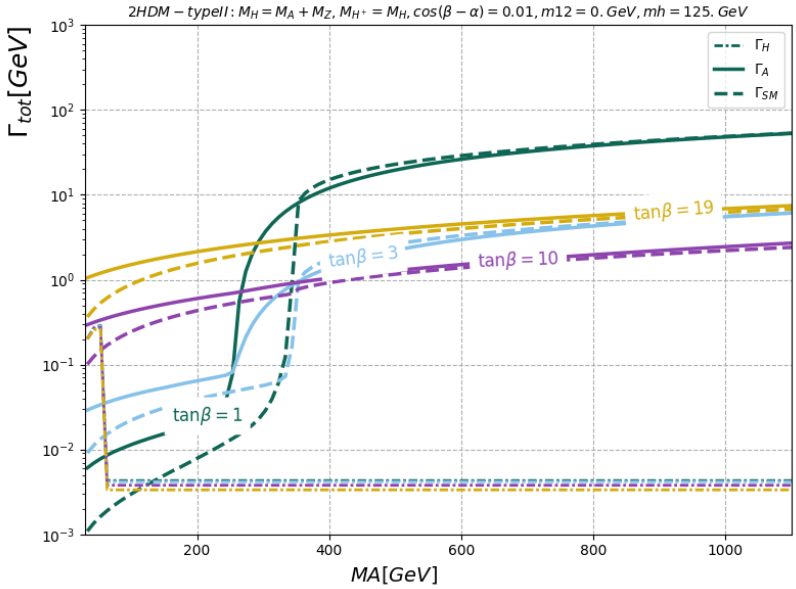


Figure 5.9. The total decay width of H, A, $h_{SM-like}$ as function of m_A for different values of tanβ for 2HDM type-II signal parameters close to the alignment limit.

being the carriers of the strong force, having higher probabilities to appear in the initial state of the protons. Gluon-gluon interactions are also more frequent than quark-antiquark interactions $q\bar{q} \to Z(\ell^\pm\ell^\mp)jj$ which is reduced by the parton density function of the antiquark. As shown in the Feynman diagrams in Fig. 5.10, these two processes will be irreducible when two b-jets are in the final state because there will be no way to separate them from the signal process of interest. However, when only one or no b quark is produced in the acceptance, the background becomes reducible and all will depend on the performance of b tagging.

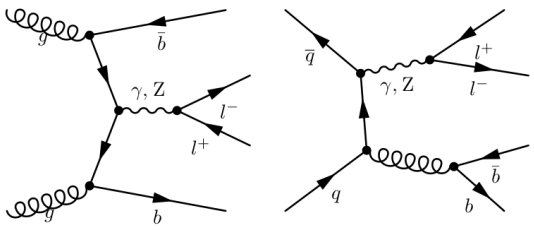


Figure 5.10. Example of leading order diagrams of the $gg \to \ell^\pm\ell^\mp b\bar{b}$ process [129].

The top-pair production

The production of a top-antitop pair also provides two leptons when both $W^{\pm}$ decay leptonically. One possible scenario that could possibly mimic our signal appears when both W decay to electrons or muons another one is when at least one of the W decays to a τ lepton. The τ can itself decay leptonically, creating an extra possibility to observe two same-flavor leptons. The top-pair production channels are shown in Fig. 5.11.

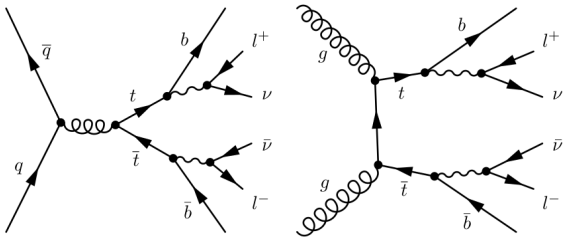


Figure 5.11. | Example of production of $t\bar{t}$ process in the fully-leptonic final state at leading order [129].

ZZ and Zh

The production of two $Z-$bosons arises at a much lower rate because of the extra weak couplings. A sample of $ZZ \to llq\bar{q}$ is generated with `MADGRAPH5aMC@NLO` and a cross-section of 3.22 pb. The associate production of a Higgs boson with a Z where the Higgs particle decays to two b-quarks is a similar background with its cross-section computed at NLO of 0.87 pb. Leading order diagrams of these two processes are drawn on Fig. 5.12.

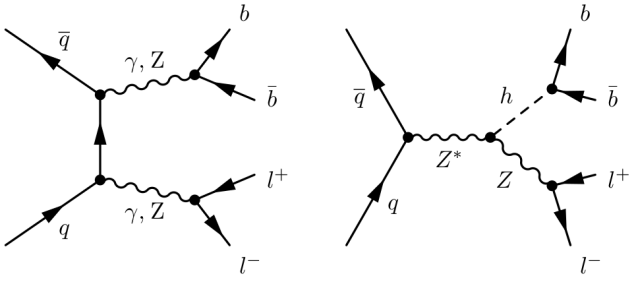


Figure 5.12. | Leading order ZZ (left) and Zh (right) process diagram [129].

Single Top

The production of only one top quark is also a source of background. The most important production mode is the associated production of an (anti-)top quark and a W boson as shown in Fig. 5.13.

There are two other mechanisms of single top production with a W in the t- and s-channel as shown in Fig. 5.14.

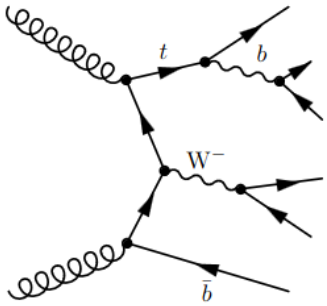


Figure 5.13. | Single top production in addition to a W^- boson, leading diagram [129].

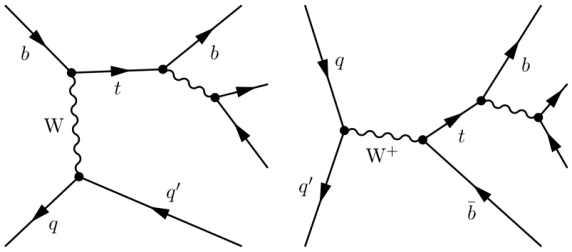


Figure 5.14. | Feynman diagram of the production of a single top with a W boson in the t-channel (left) and s-channel (right) [129].

W plus heavy flavor jets

The production of one W boson in addition to at least two jets is also a background. The signal's final state can be mimicked when one or more gluons are radiated. Its production is sketched in Fig. 5.15.

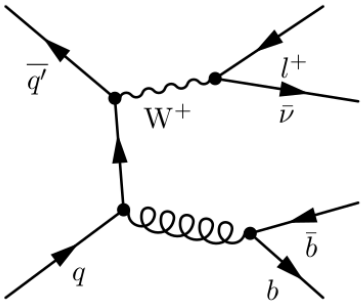


Figure 5.15. | Example of leading order diagrams of the $gg \to W^+ b\bar{b}$ process [129].

The full list of background samples used in the analysis, along with their cross sections is shown in Tab. 5.2.

Table 5.2. | Background simulated samples used in the full run2 analysis and their cross sections in pb.

Process	Dataset	σ [pb]
Drell-Yan	DYJetsToLL_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8	4757.0
	DYJetsToLL_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	859.589
	DYJetsToLL_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	361.4
$t\bar{t}$	TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	365.35
	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	88.28
	TTToHadronic_TuneCP5_13TeV-powheg-pythia8	377.96
Single-Top	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
	ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	136.02
	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	80.95
	ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8	3.36
VV	ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8	3.22
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	0.564
	ZZTo4L_13TeV-amcatnloFXFX-pythia8	1.256
	WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	49.997
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	12.178
	WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	5.595
	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.033
	WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	10.71
VVV	WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.2086
	WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.1651
	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05565
	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01398
$W+$ jets	WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8	61526.7
$t\bar{t}+V$	TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
	TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
	TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8	0.5297
	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2529
$t\bar{t}h(h \to b\bar{b})$	ttHTobb_M125_TuneCP5_13TeV-powheg-pythia8	0.2934
$t\bar{t}h(h \to \text{non-}b\bar{b})$	ttHToNonbb_M125_TuneCP5_13TeV-powheg-pythia8	0.2151
$Zh(h \to WW)$	HZJ_HToWW_M125_13TeV_powheg_jhugen714_pythia8_TuneCP5	0.0406
$Zh(h \to b\bar{b})$	ZH_HToBB_ZToLL_M125_13TeV_powheg_pythia8	0.173
	ggZH_HToBB_ZToLL_M125_13TeV_powheg_pythia8	0.00695
	ggZH_HToBB_ZToNuNu_M125_13TeV_powheg_pythia8	0.00695
$ggh(h \to ZZ)$	GluGluHToZZTo2L2Q_M125_TuneCP5_13TeV_powheg2_JHUGenV7011_pythia8	0.00695

5.2. Objects reconstruction

The reconstruction of particle objects in CMS has been discussed in detail in Sec 3.3 of Chap. 3. In this section, the selection requirements for a global event reconstruction which is performed using the particle-flow (PF) algorithm [130] are briefly summarized. The algorithm aims to identify and reconstruct individually all the particles produced in a collision, starting from the digital raw information recorded in various CMS sub-detectors.

Particle-flow electrons are reconstructed from GSF tracks pointing towards an ECAL cluster that is not linked with more than two tracks. Additional ECAL clusters aligned with tangents of the tracks are also considered to have an estimate of bremsstrahlung and a better energy resolution.

Muons are reconstructed from tracks found in the muon system, associated with tracks in the silicon tracking detectors. They are identified based on the quality of the track fit and the number of associated hits in the different tracking detectors.

The lepton isolation cut is typically expressed in terms of transverse momentum (p_{T}) as the sum of the p_{T} of all the particles that lie within a certain isolation cone size (ΔR) around the lepton of interest divided by the p_{T} of the lepton. The cut in dimensionless units is required to be < 0.06 for electrons (with a cone of radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$) and < 0.15 for muons (with a cone of radius $\Delta R = 0.4$).

Jets are reconstructed by clustering PF Charge Hadron Subtracted (CHS) particles using the anti-kT (AK) jet clustering algorithm, with a distance parameter $\Delta\mathrm{R} = 0.4$ (AK4 or resolved jet). For the boosted topologies, jets are clustered with a larger opening angle corresponding to $\Delta R = 0.8$ (AK8 or boosted fatjet). When clustering the particles in jets, isolated electrons and muons as well as charged particles associated with other interaction vertices are removed. Resolved jets are required to have $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.4$, and be separated from identified leptons by a distance of $\Delta R > 0.4$. While boosted fat jets are required to have $p_{\mathrm{T}} > 200$ GeV, $|\eta| < 2.4$, and be separated from identified leptons by a distance of $\Delta R > 0.7$.

A summary of the object definitions and selection cuts is shown in Tab. 5.3. For more comprehensive information about each specific cut and its associated parameters[1], readers are referred to the detailed documentation provided in

[1] *dz* (Longitudinal Impact Parameter): It measures the distance between the point of closest approach of the electron's track and the z-coordinate of the primary vertex along the beamline direction. A smaller *dz* implies that the electron's trajectory aligns well with the position of the primary vertex along the beamline.

dxy (Transverse Impact Parameter): It measures the distance between the point of closest approach of the electron's track to the beamline and the position of the primary vertex projected onto the transverse plane. A smaller *dxy* indicates that the electron's trajectory is

Sec. 3.3.2 of Chap. 3.

Table 5.3. | Definitions for reconstructing objects, including their identification, isolation, and criteria for selection, are outlined.

Object	Definition	Selection cuts
leading (sub-leading) e	cut-based ID Fall17 V2, medium working point **Barrel:** dz w.r.t first PV < 1 mm, dxy w.r.t the beam spot < 0.5 mm **Endcap:** dz w.r.t first PV < 2 mm, dxy w.r.t the beam spot < 1 mm	$p_{\mathrm{T}} > 25(15)$ GeV, $\lvert\eta\rvert < 2.5$
leading (sub-leading) μ	cut-based ID, medium working point PF relative isolation ΔR=0.4 ($\Delta\beta$-corr. < 0.15) 3D impact parameter w.r.t first PV sipd3D < 4.	$p_{\mathrm{T}} > 25(10)$ GeV, $\lvert\eta\rvert < 2.4$
AK4 jet(resolved)	PF anti-k_T 0.4 Jet ID flagged tight-lepton-veto (ΔR(jet, lepton) > 0.4)	**2016:** $p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.4$ **2017 and 2018:** $p_{\mathrm{T}} > 30$ GeV, $\lvert\eta\rvert < 2.5$
AK8 jet(boosted)	PF anti-k_T 0.8 FatJet ID flagged tight ΔR(jet, lepton) > 0.7 with 2 valid subJetIdx N-subjettiness ratio $\tau_2/\tau_1 < 0.7$	$p_{\mathrm{T}} > 200$ GeV, $\lvert\eta\rvert < 2.5$
b-jet(resolved)	DeepJet, medium working point	**2016:** $p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.4$ **2017 and 2018:** $p_{\mathrm{T}} > 30$ GeV, $\lvert\eta\rvert < 2.5$
b-jet(boosted)	≥ 1 subjet pass DeepCSV, medium working point	$p_{\mathrm{T}} > 200$ GeV, $\lvert\eta\rvert < 2.5$

5.3. Applied corrections

Lepton reconstruction, identification, and isolation efficiencies (detailed earlier in Sec. 5.2) in the simulation are corrected for residual differences with respect to data. These corrections are measured in a data sample, enriched in $Z \to \ell\ell$ events, using a "Tag-and-probe" method and are parameterized as a function of lepton p_{T} and η. Muon momentum scale and resolution are further corrected using the *Rochester corrections* [131] which remove a bias of the muon momentum from detector misalignment or any possible error of the magnetic field.

The reconstructed jets are further calibrated in data and simulated events using jet energy scale (JES) and jet energy resolution (JER) to account successively

consistent with originating from the primary vertex.

The 3D impact parameter refers to a measure of how far a charged particle's trajectory deviates from the primary interaction point of a collision in three dimensions (x, y, and z).

for the effects of the pileup, uniformity of the detector response, and residual data-simulation jet energy scale differences.

The DeepCSV and DeepJet algorithms [132] are used to identify jets originating from b quarks. AK4-jets are considered as b tagged if they pass the medium working point of the DeepJet algorithm, which provides around 70% efficiency with a mistag rate of less than 1%. AK8-jets are considered b tagged if at least one of the subjet pass the medium working point of the DeepCSV algorithm. Correction factors are applied in the simulation to the selected jets to account for the different responses of the algorithms between data and simulation, internally referred to as "fixed working point" b-tagging scale factors, as these only correct the b-tagging discriminator shapes in bins of the working points. The scale factors are calculated on an event-by-event basis, using "method 1a" [133] applied to the medium working point.

In 2016 and 2017, the gradual timing shift of ECAL was not properly propagated to L1 trigger primitives (TP) resulting in a significant fraction of high eta TP being mistakenly associated with the previous bunch crossing (BX) [134]. The effect is strong η and p_{T} dependent and prefiring rates well above 50% can be reached for jets with p_{T} of several hundreds of GeV and $2.75 < |\eta| < 3$. A similar effect is present in the muon system, where the BX assignment of the muon candidates can be wrong due to the limited time resolution of the muon detectors [134, 135]. These effects are not described by the simulations. A recipe was derived to correct the efficiency loss of all objects. The centrally calculated prefiring weight is made available in the NanoAODs format and is applied to all 2016 and 2017 simulated events in the analysis.

The missing transverse momentum E_T^{miss}, is reconstructed as the negative of the vector sum of the transverse momenta of all final-state particles reconstructed in the detector [136]. Its magnitude is denoted by $p_{\mathrm{T}}^{\mathrm{miss}}$. A three-step correction is pieced together to remove any bias due to the nonlinearity of the response of the calorimeter for neutral and charged hadrons, caused by event pile-up, large bending of low p_{T} tracks due to strong magnetic fields in CMS, etc. The jet energy scale corrections are propagated to E_T^{miss} using the so-called "type-I" correction, a "type-II" correction applied, also, to correct for the soft jets below the threshold used for "type-I" correction and energy deposits not clustered in any jet. Subsequently, to reduce the dependency of E_T^{miss} on event pile-up, a so-called "type-0" correction which is a mitigation for the degradation of the PF-MET reconstruction due to the pile-up interactions. In addition, $xy-$shift correction is applied to reduce the MET ϕ modulation by shifting the origin of the coordinate in the transverse momentum plane.

A number of per-event flags are recommended by the MET POG to ensure the events are taken under good data-taking conditions, where the MET is well understood [137]. Events that do not pass all of the required conditions are

vetoed. The full list of flags that are required in the analysis is given in Tab. 5.4.

Table 5.4. | MET flag requirements in the event selection.

	2016PreVFP	2016PostVFP	2017	2018
Flag_goodVertices	✓	✓	✓	✓
Flag_globalSuperTightHalo2016Filter	✓	✓	✓	✓
Flag_HBHENoiseFilter	✓	✓	✓	✓
Flag_HBHENoiseIsoFilter	✓	✓	✓	✓
Flag_EcalDeadCellTriggerPrimitiveFilter	✓	✓	✓	✓
Flag_BadPFMuonFilter	✓	✓	✓	✓
Flag_BadPFMuonDzFilter	✓	✓	✓	✓
Flag_eeBadScFilter	✓	✓	✓	✓
Flag_ecalBadCalibFilter	x	x	✓	✓

Top quarks are the heaviest known elementary particles, and their production mechanisms are governed by quantum chromodynamics (QCD), which describes the strong force interactions among quarks and gluons as was described in Chap. 1. Because of the large mass of top quarks, they are sensitive to higher-order QCD effects. Theoretical predictions for top quark production are usually calculated to certain orders in perturbation theory (such as leading order, next-to-leading order, etc.), but these calculations might not perfectly match the data due to missing higher-order terms or complex QCD effects. During Run 1 and Run 2 of the LHC experiments, it was noticed that the observed p_T distribution of top quarks in data was softer (meaning the p_T values were generally lower) compared to the predictions obtained from various MC simulations based on LO or NLO matrix elements interfaced with parton showers. This discrepancy suggested that the theoretical predictions didn't fully capture all the intricacies of top quark production [138].

To mitigate this discrepancy and align the simulated p_T distribution of top quarks more closely with the data, the top p_T reweighting technique was used in this search. The idea is to modify the p_T distribution in the MC simulations such that it better matches the observed data. This reweighting is typically done based on a scale factor that depends on the p_T of the top quarks. By applying these scale factors to the simulated events, the distribution of top p_T in the MC simulation is adjusted to reproduce the softer spectrum seen in the data.

Finally, the reconstructed mass resolution $\sigma_{m_{bb}}$ and $\sigma_{m_{llbb}}$ are further corrected by applying multivariate regression techniques. This correction has the advantage of further improving detector response and corrects for semi-leptonic b-decays that usually lead to mis-measured energy due to escaping neutrinos worsening the b-jets energy scale and resolution. The correction is used before in previous $H \to b\bar{b}$ CMS searches [139, 140] and it is available for run2 analysis in

both MiniAOD and NanoAOD formats. The correction was derived using a deep neural network trained on simulated top quark events with inputs that provide information about the kinematics, shape, and composition of reconstructed jets such as; jet kinematics, the distance ΔR of clustered electrons or muons within a jet, jet composition, and information about the secondary vertex. Further details can be found in [141].

The applied correction is furthermore validated on $H/A \to b\bar{b}$ mass resolution measured on simulated signal samples. The performance is shown in Fig. 5.16 for simulated samples of $H \to Z(\to ll)A(\to b\bar{b})$ events. Averaging over all resolved categories, the improvement in mass resolution is approximately 16%, resulting in an increase of about 10% in the sensitivity of the analysis.

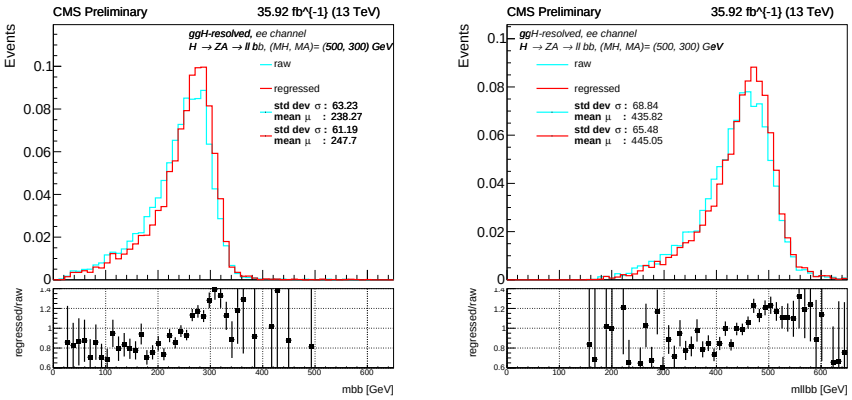


Figure 5.16. m_{bb} and m_{llbb} distributions for simulated 2016 pre-/post-VFP signal samples of $H \to Z(\to \ell^+\ell^-)A(\to b\bar{b})$ events ($m_H = 500$, $m_H = 300$) GeV, before (blue) and after (red) the jet energy regression correction. The displayed mass resolutions are reconstructed in the $nb = 2$ -resolved category in the ee lepton flavor.

5.4. Events selection

5.4.1. Trigger selection

Recorded events are collected with a set of dilepton triggers. The p_T thresholds applied by the triggers are asymmetric, channel-dependent, and vary from 17 to 23 GeV for the leading leptons and from 8 to 12 GeV for the subleading leptons. Trigger efficiencies are measured with the "Tag-and-Probe" technique [142] as a function of lepton p_T and η in a data control region consisting of $Z \to \ell\ell$ events. The derived corrections are summarized in [143, 144].

The list of trigger paths at the HLT level is shown in Table 5.5.

Table 5.5. | Trigger paths for full run2 data taking.

Year	Channel	HLT path	Excluded for run
2016	DoubleMuon	$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ$	
		$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL$	
		$HLT\_Mu17\_TrkIsoVVL\_TkMu8\_TrkIsoVVL\_DZ$	
		$HLT\_Mu17\_TrkIsoVVL\_TkMu8\_TrkIsoVVL$	
	DoubleEG	$HLT\_Ele23\_Ele12\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	
	MuonEG	$HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	B, C, D, E, F-preVFP
		$HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL$	
		$HLT\_Mu8\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	
		$HLT\_Mu8\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL$	H
2017	DoubleMuon	$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ$	
		$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ\_Mass8$	
		$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ\_Mass3p8$	B
	DoubleEG	$HLT\_Ele23\_Ele12\_CaloIdL\_TrackIdL\_IsoVL$	
		$HLT\_DiEle27\_WPTightCaloOnly\_L1DoubleEG$	B
	MuonEG	$HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	
		$HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL$	B
		$HLT\_Mu12\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	
		$HLT\_Mu12\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL$	B
		$HLT\_Mu8\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	
2018	DoubleMuon	$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ$	
		$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL$	
		$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ\_Mass8$	
		$HLT\_Mu17\_TrkIsoVVL\_Mu8\_TrkIsoVVL\_DZ\_Mass3p8$	
	SingleMuon	$HLT\_IsoMu24$	
		HLT_Mu50	
		$HLT\_OldMu100$	
		$HLT\_TkMu100$	
	EGamma	$HLT\_Ele30\_eta2p1\_WPTight\_Gsf\_CentralPFJet35\_EleCleaned$	
		$HLT\_DiEle27\_WPTightCaloOnly_L1DoubleEG$	
		$HLT\_Ele23\_Ele12\_CaloIdL\_TrackIdL\_IsoVL$	
		$HLT\_Ele32\_WPTight\_Gsf$	
		$HLT\_DoubleEle25\_CaloIdL\_MW$	
		$HLT\_Ele50\_CaloIdVT\_GsfTrkIdT\_PFJet165$	
		$HLT\_Ele115\_CaloIdVT\_GsfTrkIdT$	
		$HLT\_Ele28\_WPTight\_Gsf$	A, B, C
		$HLT\_Ele30\_WPTight\_Gsf$	D
		$HLT\_Photon200$	
	MuonEG	$HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL$	
		$HLT\_Mu12\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	
		$HLT\_Mu8\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL\_DZ$	
		$HLT\_Mu27\_Ele37\_CaloIdL\_MW$	
		$HLT\_Mu37\_Ele27\_CaloIdL\_MW$	

5.4.2. Signal selection

The $H/A \to ZA/ZH$ signal is distinguished by two jets originating from b quarks and a pair of light leptons, electrons, or muons of opposite sign. Events with two oppositely charged leptons (e^+e^-, $\mu^+\mu^-$) are selected using asymmetric p_{T} requirements, chosen to be above the corresponding trigger thresholds, for leading and subleading leptons of 25 GeV and 15 GeV for e^+e^- events and 25 GeV and 10 GeV for $\mu^+\mu^-$ events. Electrons in the pseudorapidity range $|\eta| < 2.5$ and muons in the range $|\eta| < 2.4$ are considered.

The different production mechanisms and topologies with different Lorentz boost give distinct jet signatures to the signal events. Therefore the events are divided into four different mutually exclusive categories ($nb = 2$ -resolved, $nb = 2$ -boosted, $nb = 3$ -resolved, $nb = 3$ -boosted).

Events must have exactly two AK4 jets that pass the DeepJet algorithm, medium working point, and meet all of the requirements mentioned in Sec. 5.2 to be suitable candidates to reconstruct a light Higgs boson $H/A \to b\bar{b}$ via gluon-gluon fusion production in the resolved category ($nb = 2$ -resolved). Events with exactly one "fat" AK8 jet with two subjets constituent, where both of them pass the medium b-tagging working point of the DeepCSV algorithm are selected as candidates to reconstruct ggH in boosted event topology ($nb = 2$ -boosted). Events with at least three AK4 jets that pass the medium working point of the DeepJet algorithm are classified as signal events produced via b-associated production (bbH), and the category is denoted ($nb = 3$ -resolved). Finally, the last category ($nb = 3$ -boosted) is reconstructed from events that have at least one b-tagged "fat" AK8 jet and at least one b-tagged AK4-jet, these events are selected as candidates to reconstruct bbH in boosted event topology.

The signal is expected to mostly concentrate at low E_T^{miss} as demonstrated in Fig. 5.17 (left) and around the Z boson peak as demonstrated in the m_{ll} mass distribution in Fig. 5.17 (right). Therefore, tighter selection requirements are customized to roughly keep the same number of signal events while reducing the number of background events, thus improving the sensitivity of the search.

For events with two oppositely charged same-flavor leptons, a requirement of $70 < m_{ll} < 110$ GeV is applied to enhance the presence of $Z \to \ell\ell$ events. Both requirements have a negligible impact on signal events and they are intended to work for all signal hypotheses.

In particle physics, the simple quantity $s/\sqrt{b}$ has been widely used as a measure of expected discovery significance. The rationale behind this formula is that a Poisson distributed quantity n with a large mean value $s+b$ can be approximated by a Gaussian distributed variable x with mean $s+b$ and standard deviation $\sqrt{s}+b$. The intuitive explanation of this formula is that the standard deviation of n assuming background only is $\sqrt{b}$ and therefore the ratio $s/\sqrt{b}$ represents

the size of the signal divided by the statistical error on n expected assuming the signal is absent.

A more precise expression of the formula above $\sqrt{2((S+B)\ln(1+S/B)-S)}$ known as Asimov's estimate [145] is used instead as an expectation of our search power. Both formalisms are valid but the latest is very close to the true significance even for very small statistics, opposite to $s/\sqrt{b}$ estimator as it often gives overestimated values.

We use this mathematical formalism as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$, where $S(B)$ is the number of signals (background) events to introduce a cut to suppress background with real missing energy. Fig. 5.18 shows the signal significance as a function of the $p_{\mathrm{T}}^{\mathrm{miss}}$ cut. For each produced signal sample, we compute this variable for different values of $p_{\mathrm{T}}^{\mathrm{miss}}$ and eventually pick the cut that yields the highest value. Events in the $nb=2$ -resolved, $nb=2$ -boosted, $nb=3$ -resolved, and $nb=3$ -boosted categories are required to have a $p_{\mathrm{T}}^{\mathrm{miss}}<80$ GeV.

It is well known that the best value of the $p_{\mathrm{T}}^{\mathrm{miss}}$ cut slightly changes with respect to the signal hypothesis being considered. However, these two requirements overall achieve a sensible choice to ensure high sensitivity throughout all signal mass hypotheses.

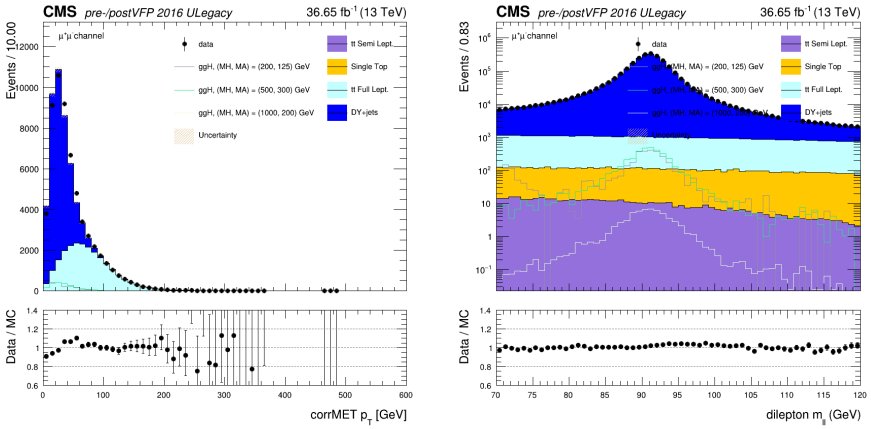


Figure 5.17. 2016 ULegacy pre-/post-VFP corrected $p_{\mathrm{T}}^{\mathrm{miss}}$ for $xy-\phi$ modulation (left) and m_{ll} (right) in the $nb=2$ -resolved ($\mu\mu$) category.

The signal region selection and the defined categories are summarised in Table. 5.7 (see Sec .5.6).

For the gluon-gluon fusion (ggH) production, depending on the assumed m_A and m_H masses, 94%-97% of the events passing the selection detailed in Sec. 5.4.2 fall into the $nb=2$ category. On the other hand, for the b-associated production (bbH), even though more than two b-jets are expected, due to the relatively

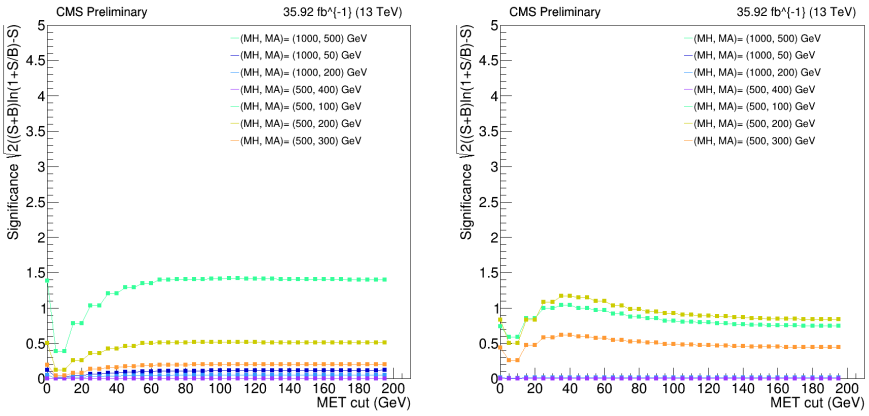


Figure 5.18. | Signal significance as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$ cut, evaluated for different 2016 ULegacy pre-/post-VFP signal benchmarks in the ee lepton flavor and for the two categories $nb = 2$ -boosted (left) and $nb = 2$ -resolved (right). The identical approach was examined across three years of data taking and within both lepton flavors.

soft p_{T} spectrum of the associated b-jets and the geometric acceptance of the tracker, only 27%-36% fall into the $nb = 3$ category. In addition, the overall signal efficiency of the $nb = 3$ category for ggH is very small making this category irrelevant to use for hypothesis testing. A schematic view of the efficiency across the defined signal regions is shown in Fig 5.19.

5.5. Background modelling

Backgrounds originating from various sources are considered, including Drell-Yan events associated with the production of heavy-flavour jets, fully leptonic top quark pair decays, diboson processes, W boson production combined with heavy flavor quarks, Higgs boson production, as well as all other backgrounds outlined in Sec. 5.1.2.3. Their shapes are taken from simulation, whereas they are normalized using precise inclusive cross-sections calculated from theory. The diboson samples are normalized using next-to-next-to-leading-order (NNLO) cross-sections. Single-top-quark production and top-quark-pair production in association with vector bosons are normalized to next-to-leading-order (NLO) cross-sections. The normalization of the Higgs boson production in association with a vector boson using NNLO QCD and NLO electroweak corrections. All cross-sections mentioned above follow the recommendations in Ref. [146].

5.5.1. Background control regions

To help determine the normalization of the main background processes, and to validate how well, the simulated samples model the distributions of variables

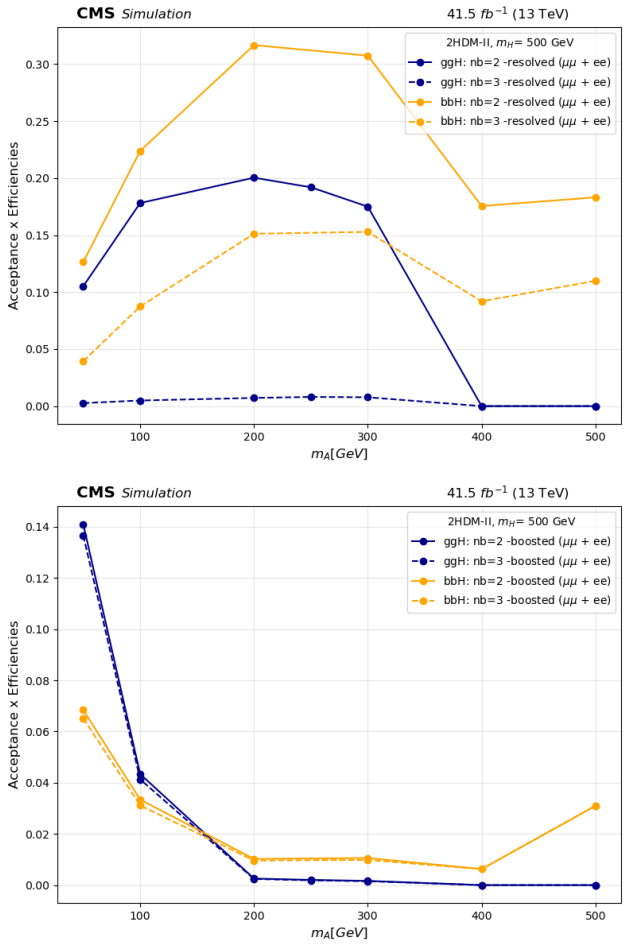


Figure 5.19. | ggH and bbH signal acceptance times efficiency as function of m_A for a Heavy Higgs bosons H mass m_H =500 GeV. The results are shown for the events passing the signal region selections $nb = 2$ and $nb = 3$ in the resolved (top) category, similarly boosted (bottom).

most relevant to the analysis, four control regions are selected in the data. Table. 5.6 lists the selection criteria used to define these regions for the three lepton channels. Proper normalization of background processes is important because the $\mu^{\pm}e^{\mp}$ event category is additionally used to constrain the systematics uncertainties in the maximum likelihood fit.

The dominant backgrounds after the listed selections are from Z+jets and top-quark production. The top control region is defined by keeping the same selection as Z+jets control region, apart from an opposite-flavor lepton criterion, i.e. an opposite-charge μe pair is required instead of an ee or $\mu\mu$ pair (see also

Table 5.6. | Definition of control regions before btag requirements

		Control region (Before btag)	
		resolved ($nj(AK4jets) \geq 2$ and $nj(AK8jets) == 0$)	**boosted** ($nj(AK8jets) \geq 1$ and $nj(AK4jets) \geq 0$)
TTbar Control region	μe	Leading lepton $p_T > 25$ GeV, Sub-leading e (μ) $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.	Leading lepton $p_T > 25$ GeV, Sub-leading e (μ) $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.
Z+jets Control region	$\mu\mu$ or ee pair	Leading lepton $p_T > 25$ GeV, Sub-leading e (μ) $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.	Leading lepton $p_T > 25$ GeV, Sub-leading e (μ) $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.

Table. 5.6). A very pure control region of top-quark pair production is obtained using $p_{\mathrm{T}}^{\mathrm{miss}}$ and the region is further used to determine the normalization of the background control region, whereas its shape in the signal region is taken from the simulation.

5.5.2. Drell-Yan reweighting

In this section, the technique used for optimizing Drell-Yan+jets background modeling is described. The method employs simple event weights which are functions of the di-jets system invariant mass m_{jj}. The event-by-event weight is extracted from m_{jj} event templates which are constructed from 0-btag + $p_{\mathrm{T}}^{\mathrm{miss}}$ cut applied on Z+jets control region done exclusively for resolved and boosted categories [1]. This region is a distinct subset of data and the selected events are entirely separate from the region of interest where the signal is expected to manifest. This ensures that any observed effects or discrepancies in the control region are not influenced by the presence of the signal being investigated.

The weights are derived separately for each year and lepton flavor. An example of these templates (i.e. used to derive DY weights) is shown in Fig. 5.20 are for 2017 ULegacy data and MC simulation in ee channel.

The DY weight is applied as follows

$$W_{DY} = \begin{cases} \text{resolved:} & f_{poly}^{n^{th}}(m_{jj})_{[0,150]GeV} + f_{poly}^{4^{th}}(m_{jj})_{]150,600]GeV} \\ & \qquad + W_{bin]650,1200]GeV}, \\ \text{boosted:} & f_{poly}^{n^{th}}(m_{jj})_{[0,150]GeV} + W_{bin]150,1200]GeV}, \end{cases} \tag{5.4.}$$

where n equals 6 in 2016 and 2018 and 7 in 2017. W_{bin} is the estimated weight

[1] The applied criteria match those specified in Table 5.7 (refer to Section 5.6), but with the exclusion of any requirements related to b-tagging.

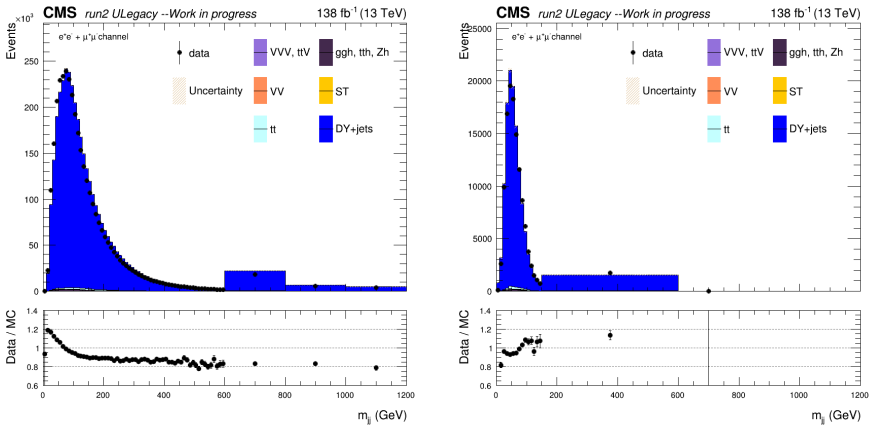


Figure 5.20. Invariant mass distributions of the di-jets system m_{jj} resolved (left) and boosted (right) shown in 0-btag $Z(\mu\mu + ee)$+jets control region which further used to derive the shape uncertainties corrections of DY+jets. Error bars indicate statistical uncertainty.

of a bin in the specified mass range. The polynomial degree fits are extracted from the ratio Data/MC of the m_{jj} distributions of Fig. 5.20 as illustrated in Fig. 5.21.

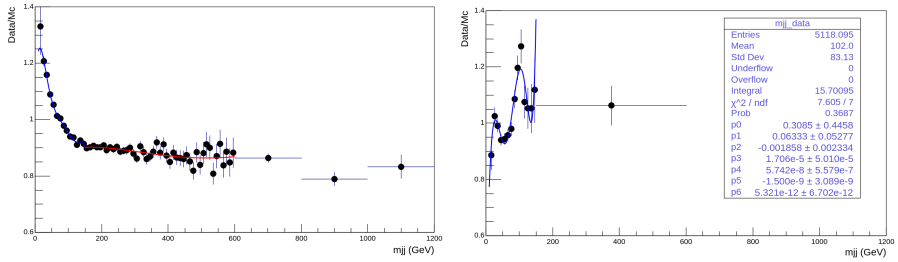


Figure 5.21. The fitted polynomial functions in the resolved (left) and boosted (right) regions taken from Data/MC ratio of the di-jets invariant mass, m_{jj}, of Fig. 5.20. Events are those passing 0-btag selection criteria for 2018 data and Monte-Carlo events.

As a cross-check of the measured DY weights, m_{jj} and m_{lljj} distributions are shown in Fig. 5.22 as cross-validation.

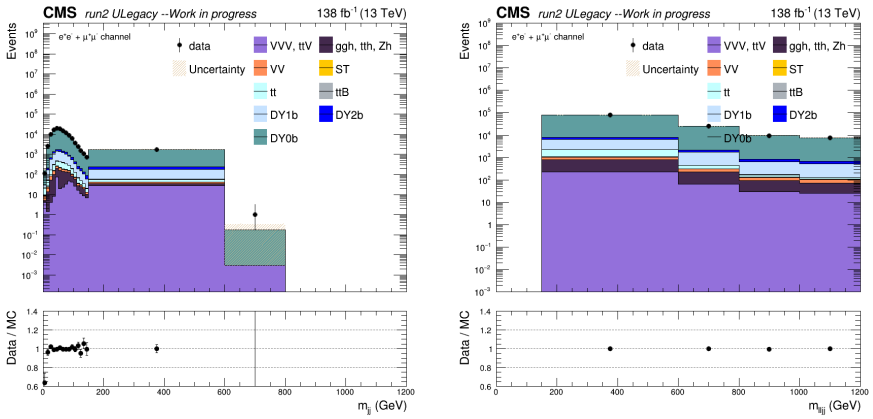


Figure 5.22. | m_{jj} (left), m_{lljj} (right) distributions are shown in $ee+\mu\mu$ lepton flavor for ULegacy 2018 data, simulated signal and Monte-Carlo events after applying DY weights in boosted category. Error bars indicate statistical uncertainties while shaded bands indicate systematic uncertainties.

5.5.3. Kinematics distributions

The following section showcases quantities such as m_{jj}, m_{lljj} prior to the application of b-tagging criteria (refer to Fig. 5.23), as well as m_{bb} and m_{llbb} after b-tagging requirements (refer to Fig. 5.24, and Fig. 5.25). These investigations were carried out in $t\bar{t}$ and Z+jets control regions examining both boosted and resolved categories across all the invariant mass systems detailed in this report, as well as all other possible quantities including the transverse momentum (p_{T}), azimuthal angle (ϕ), and pseudorapidity angle (η), for leading leptons and jets. This careful step in the analysis serves the purpose of validating the consistency between data and simulation after the application of all necessary corrections. It serves as a baseline for understanding the known physics processes and ensures that our understanding of the Standard Model (SM) is reliable and provides confidence in the predictions of background processes. By accurately modeling and reproducing the known physics, any deviations observed in the data can potentially be attributed to new physics.

- Control Region: Invariant mass systems (m_{jj},m_{lljj}) prior to b-tagging:

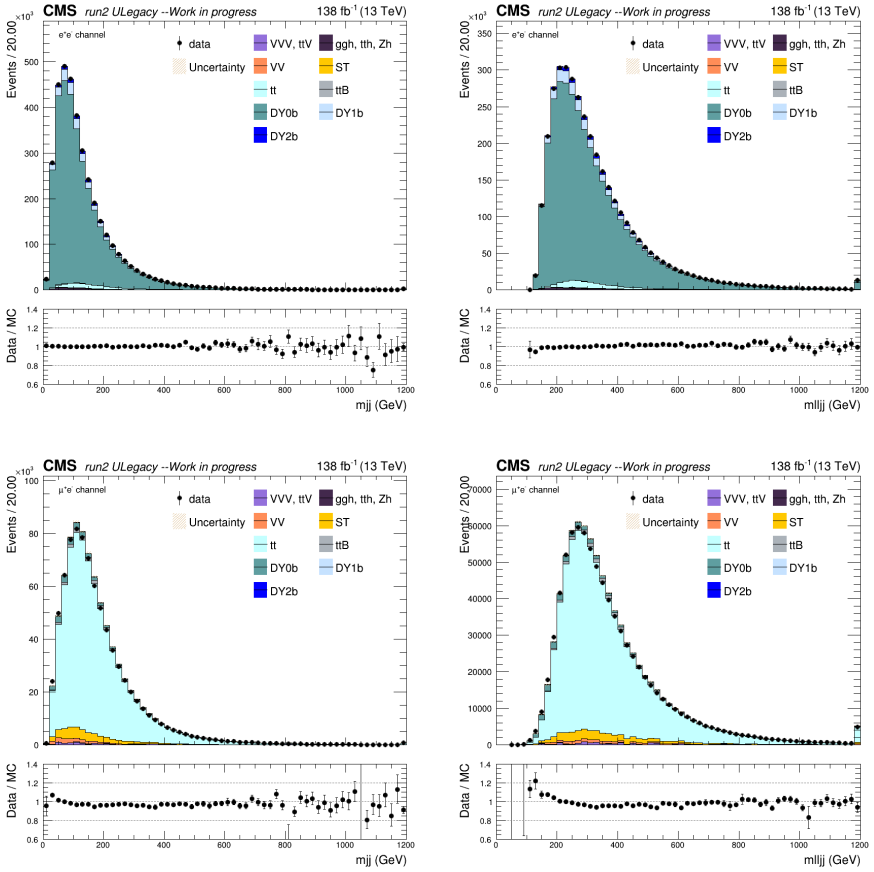


Figure 5.23. | The m_{jj} (left), and m_{lljj} (right) invariant mass distributions are shown in the two lepton flavors (ee (top) and μe (bottom)). The distributions are illustrated for ULegcay full run 2 data, simulated signal, and Monte-Carlo events after requiring two opposite charge same flavor leptons and at least two resolved AK4 jets. Error bars indicate statistical uncertainties.

- Signal Region: Invariant mass systems (m_{bb},m_{llbb}) in the presence of $nb2$ b-tag:

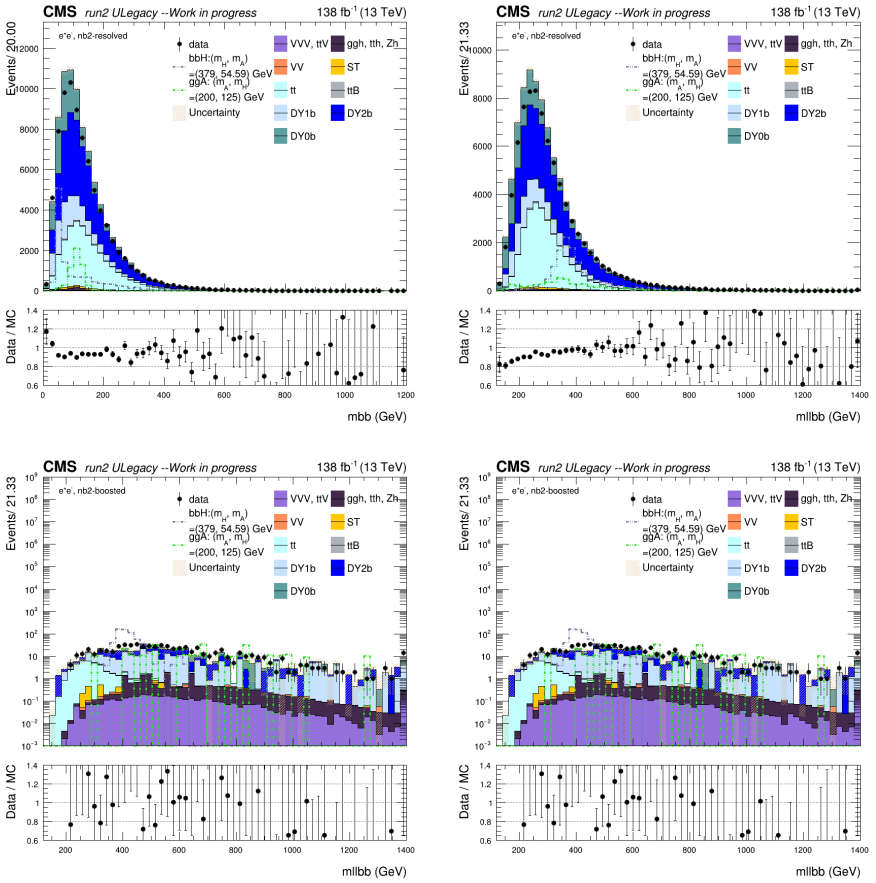


Figure 5.24. The m_{bb} (left), and m_{llbb} (right) invariant mass distributions are shown in $nb2$ -resolved (top) and -boosted (bottom) category. The distributions are illustrated for ULegcay full run 2 data, simulated signal, and Monte-Carlo events. The signal scaled to the cross-section of the 2HDM excluding branching fractions. The event requirements are two opposite charge same flavor leptons (ee) and exactly two b-tagged AK4 jets that pass the DeepJet algorithm medium working point in the resolved category, while two subjet within the AK8 jet to pass DeepCSV tagger medium working point. Error bars indicate statistical uncertainties.

- Signal Region: Invariant mass systems (m_{bb},m_{llbb}) in the presence of $nb3$ b-tag:

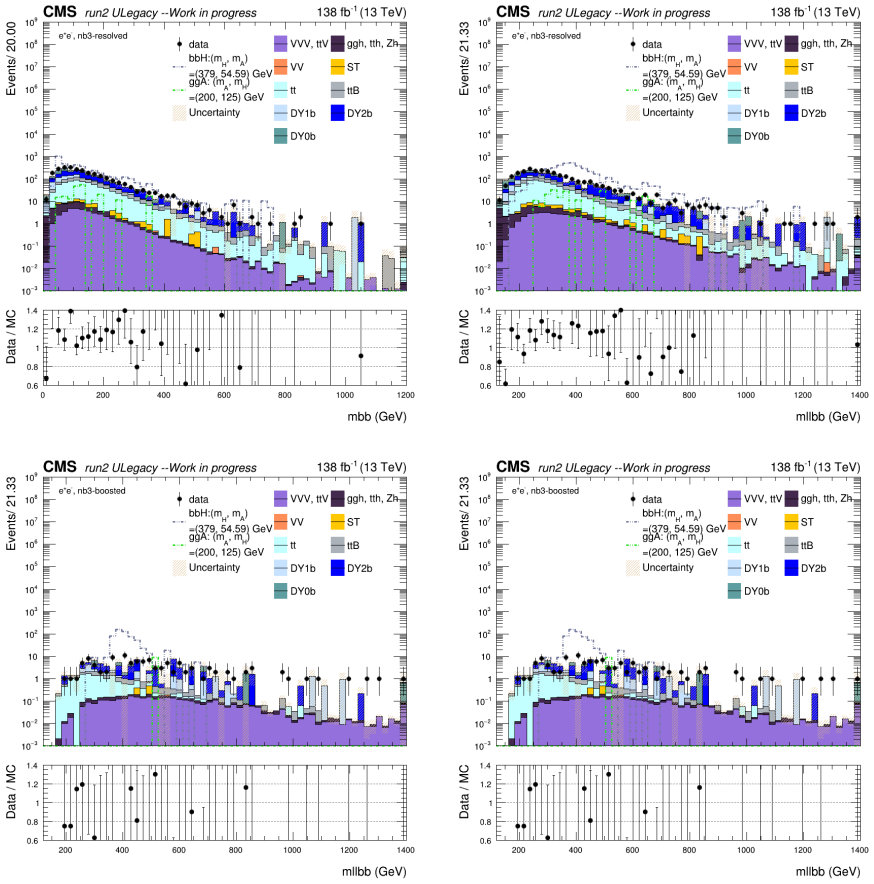


Figure 5.25. | The m_{bb} (left), and m_{llbb} (right) invariant mass distributions are shown in $nb3$ -resolved (top) and -boosted (bottom) category. The distributions are illustrated for ULegcay full run 2 data, simulated signal, and Monte-Carlo events. The signal scaled to the cross-section of the 2HDM excluding branching fractions. The event requirements are two opposite charge same flavor leptons (ee) and at least three b-tagged AK4 jets that pass the DeepJet algorithm medium working point in the resolved category, while at least one AK4 and two subjet within the AK8 jet to pass DeepCSV tagger medium working point. Error bars indicate statistical uncertainties.

Table 5.7. | Summary of the event selection for signal and control regions.

Channel		Control region categories (Before btag)	
		resolved ($nj(AK4jets) \geq 2$ **and** $nj(AK8jets) == 0$)	**boosted** ($nj(AK8jets) \geq 1$ **and** $nj(AK4jets) \geq 0$)
$\mu\mu$ / ee	**Z+jets Control region:**	Exactly 2 opposite charge same flavour leptons (ee or $\mu\mu$), with the leading one having $p_T > 25$ GeV and sub-leading e (μ) having $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.	
μe	**TTbar Control region:**	Exactly 2 opposite charge opposite flavour leptons ($\mu^{\pm}e^{\mp}$), with the leading one having $p_T > 25$ GeV and sub-leading e (μ) having $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.	
		Signal region categories (After btag)	
		resolved DeepJet tagger, medium working point.	**boosted DeepCSV tagger, medium working point.**
		nb=2 -resolved ***(nb(AK4bjets) == 2)***	***nb=3 -resolved*** ***(nb(AK4bjets) $\geq$ 3)***	***nb=2 -boosted*** ***(nb(AK8bjets) == 1)***	***nb=3 -boosted*** ***(nb(AK8bjets) $\geq$ 1 and nb(AK4bjets) $\geq$ 1)***
		Events have exactly 2 b-tagged AK4 jets plus $p_T^{miss} < 80$ GeV cut.	Events have at least 3 b-tagged AK4 jets plus $p_T^{miss} < 80$ GeV cut.	Events have exactly 1 fat AK8 jets and both subjets are required to be btagged plus $p_T^{miss} < 80$ GeV.	The remaining events with AK4 b-jets multiplicity greater than 1 and at least 1 b-tagged fat AK8 jet fall in this category. In addition, $p_T^{miss} < 80$ GeV cut is applied.
$\mu\mu$ / ee	**Z+jets Signal region:**	Exactly 2 opposite charge same flavour leptons (ee or $\mu\mu$), with the leading one having $p_T > 25$ GeV and sub-leading e (μ) having $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.	
μe	**TTbar btagged Control region:**	Exactly 2 opposite charge opposite flavour leptons ($\mu^{\pm}e^{\mp}$) with the leading one having $p_T > 25$ GeV and sub-leading e (μ) having $p_T > 15(10)$ GeV, $70 < m_{ll} < 110$ GeV.	

5.6. Analysis strategy

5.6.1. Signal mass splitting

The first attempt to distribute the generated signal mass points (m_H, m_A) into the categories that yield the largest significance using the Asimov's estimate formula $\sqrt{2((S+B)\ln(1+S/B)-S)}$ defined earlier in Sec. 5.4.2 is shown in Fig. 5.26. This preliminary attempt has revealed that for the two signal processes ggH and bbH, it is difficult and almost impossible to optimize a cut that can split the signal mass points between the resolved and boosted categories based on their estimated significance. Nevertheless, this attempt has also shown the unique sensitivity of the bbH production mechanism in comparison to ggH. For low values of $\tan\beta$, the branching factions of $H \to ZA$ and $A \to b\bar{b}$ are no longer dominant above twice the top quark's mass. The most significant process will be $A \to ZH(\to t\bar{t})$ and this explains why the signal categories for the ggH process are no longer significant for pseudo-scalar masses A greater than 360 GeV.

By studying the expected upper limits of each category and the improvement that each of them brings to the total combination, the mass points generated are distributed between the categories, as shown in Fig.5.27.

5.6.2. Auxiliary measurements

Auxiliary measurements can be used to estimate or reduce the effect of systematic uncertainties and improve the precision or efficiency of the estimation in hypothesis testing. It can be a control region or variable. Overall, the inclusion of auxiliary measurements in statistical tests aims to improve the validity, efficiency, and precision of the analysis. By leveraging additional information provided by these variables, researchers can gain a more comprehensive understanding of the relationships and draw more reliable conclusions from their statistical analyses. In this search, μe control region is used for such purpose.

5.6.3. Deep Neural Network

In this search, a parametric Deep Neural Network (DNN) is used to learn in an unbiased way the signal shape distribution. The DNN output is evaluated for each signal mass hypothesis under study in the four-event category ($nb=2$ -resolved, $nb=2$ -boosted, $nb=3$ -resolved, and $nb=3$ -boosted) in both ee and $\mu\mu$ channels, in addition to μe channel which further used as a control region to help constrained $t\bar{t}$ background in the statistical test. The DNN distributions are combined in the three lepton flavors in the maximum likelihood fit.

We use all available signal samples in the training phase, the network is then able to infer the dependence of the signal behavior on these parameters, but only the leading irreducible backgrounds that could possibly mimic the signal's final state; $t\bar{t}$ and Single-Top processes where both top quarks decay leptonically and

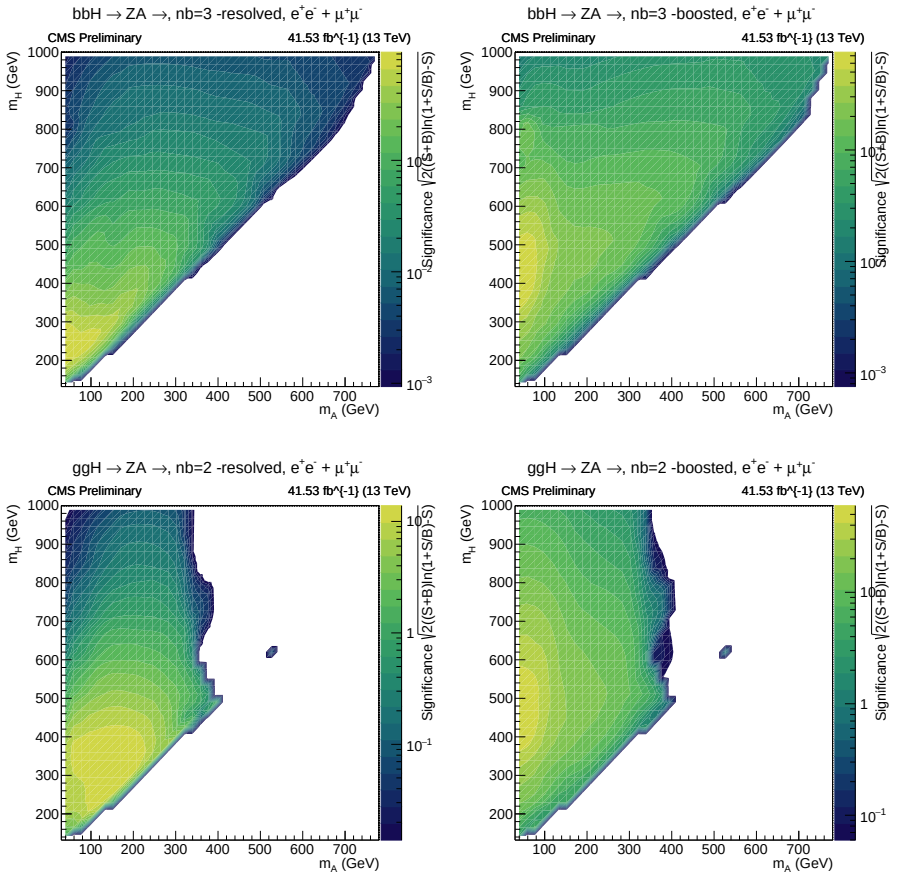


Figure 5.26. | Estimation of the significance across categories and signal processes. Shown bbH (top) and ggH (bottom) in $ee + \mu\mu$ combined lepton flavor for both resolved (left) and boosted (right) categories.

Drell-Yan process when is produced with the additional presence of two b-jets. In order to replicate the best available emulation of how events are distributed and reconstructed, the output of the network is a multiclass output with the softmax activation applied, thus each event produces 3 classes output (DY, $t\bar{t}$, and ZA) values $\epsilon[0, 1]$ with their sum equal to 1. Consequently, it can be interpreted as the probability that a given input belongs to the corresponding class, that is process or group of processes. During training, each event is weighted the same way as they are histogrammed, in particular, this includes the various scale factors introduced during the data-to-simulation corrections.

The ZA in the case of $(H \to ZA)$ and ZH in the case of $(A \to ZH)$ output nodes are the output of interest and they are used to evaluate the DNN model on

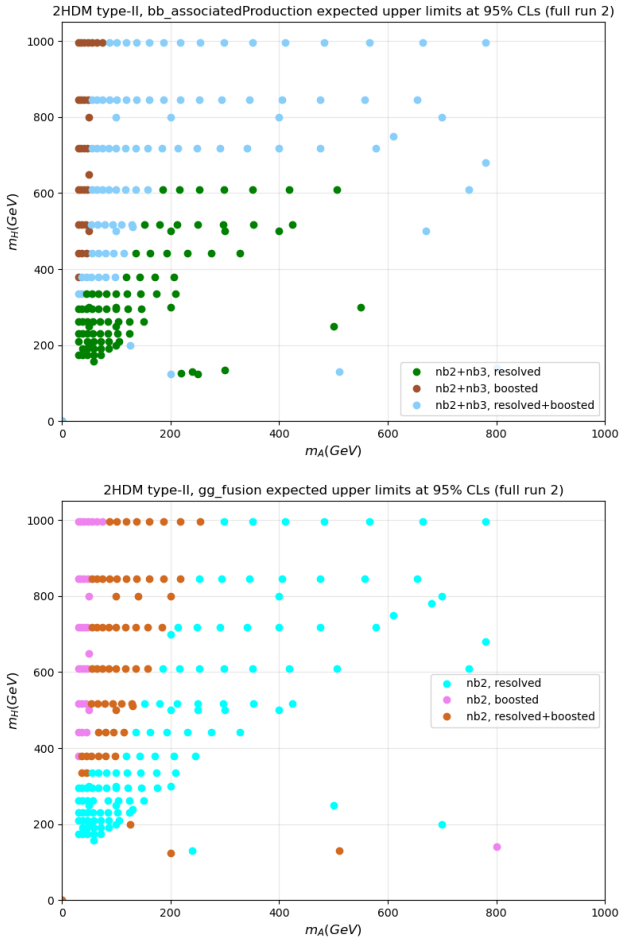


Figure 5.27. | Division of the generated mass points between signal regions categories (resolved, boosted, and resolved+boosted). The strategy for bbH (ggH) signals is displayed on the top (bottom).

both data and simulated events to derive what we referred to as "DNN output $ZA(ZH)$" which will be used in the hypothesis testing. In the evaluation of the performance of the DNN model, the signal parameters are frozen to a specific value, and only the signal sample corresponding to that value is considered. This procedure is repeated for every parameter value for which a signal sample is available.

For this analysis, we have constructed a single parameterized neural network and among all architectures, we tried, the following DNN structure was found to perform satisfyingly well:

- The neural network implementation is done using the Keras and TensorFlow frameworks.
- The input layer has 13 nodes:
  - The reconstructed Higgs masses and the masses squared [1]: m_{llbb}, m_{bb}, m^2_{llbb}, m^2_{bb} and, $m_{llbb} \times m_{bb}$.
  - One hot encoder on the era and lepton PDGID to perform "binarization" of the category and include it as a feature to train the model
  - The categorization is done in Sec. 5.4 is implemented with a simple use of binary logic with the core algorithm of "Switching Neural Networks" between resolved, boosted, ggH and, bbH.
  - The signal mass hypothesis m_H and m_A; these parameters are the values used for the signal sample generation, and for the background, they are assigned randomly.
- 256 neurons, six hidden layers with SeLU, ReLU activation functions, and dropout is applied with a dropping probability of 0.5 if no further regularisation is imposed.
- The output layer has a softmax activation function, and we apply batch normalization between all layers.
- The weights are optimized using Adam with cross-entropy loss and an initial learning rate of $1e^{-5}$.
- The learning rate is reduced by a factor of 5. If there is no improvement of the loss on the validation data for 5 consecutive epochs.
- The network is trained for 200 epochs or until no improvement is observed, whichever occurs first.

The only physical quantities used as input are the reconstructed masses m_{bb} and m_{llbb}, however, the DNN demonstrated high prediction in the whole region of the parameter space (m_A,m_H). Fig. 5.28 illustrates the performance of the DNN outside of its training set (i.e. evaluating the model for parameters on which it was not trained). Continuity is obtained by using parameters not seen by the network, further proving the correct interpolation of the DNN.

In the low-mass region, the resonance is a sharp Breit–Wigner distribution, so the signal region is small. It can easily be seen that the DNN has converged to an ellipse, which was used as an analysis strategy in the previous CMS search [34].

The boosted region, on the other hand, where $m_H >> m_A$ means that the decay products of the light pseudo-scalar A are boosted in the same direction, which often makes them impossible to distinguish. Events in this region are

[1] Derived features (i.e. m^2_{llbb}, m^2_{bb} and, $m_{llbb} \times m_{bb}$) are created to capture additional information, patterns, or relationships that may not be evident or adequately represented by the original input variables alone. By introducing derived features, the DNN can have access to a richer representation of the data, potentially improving its ability to learn complex patterns and make accurate predictions.

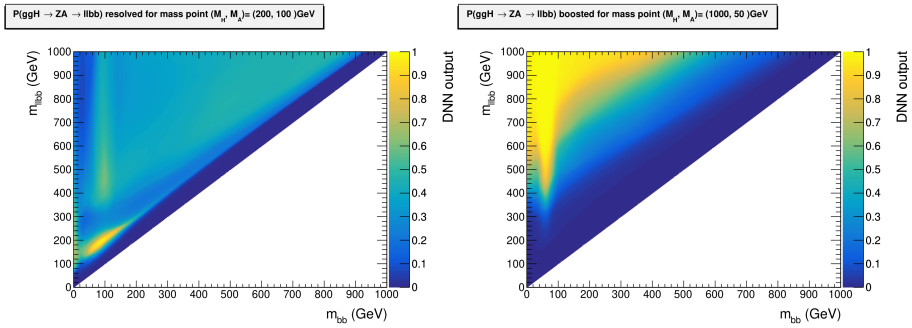


Figure 5.28. The DNN output score distributions are shown for two distinct ggH signal samples, corresponding to the given (m_H, m_A) GeV values of (1000, 50) (on the left) and (200, 100) (on the right), showcasing variations with respect to the invariant masses m_{bb} and m_{llbb}.

difficult to select and in this case, the yellow area that reflects a high DNN score discriminant becomes much larger and far in shape from an ellipse. This is understandable because this region has very little background contamination and the DNN learns that it can incorporate events from this area with minimal purity cost. This newly DNN-based approach, developed within this study, is therefore anticipated to surpass previous studies, particularly excelling in this boosted region.

As was mentioned in Sec. 4.4 in Chap. 4 evaluating the performance of a trained model is very important as it allows for informed decision-making, model selection, and iterative improvement. It also ensures that the model meets the desired performance standards, generalizes well to unseen data, and provides reliable predictions. The performance of the network is illustrated in Fig. 5.29 with the so-called Receiver operating characteristic curve (ROC). Each curve shows the efficiency of correctly identifying signal or background events as a function of the misidentification rate, when applying a sliding requirement on the score of the DNN. A high signal efficiency for a fixed background rejection indicates good performance.

The Area under the ROC Curve (AUC) is also shown which provides us with an aggregate measure of performance across all classification thresholds. AUC ranges in value from 0 to 1. Therefore, a model whose predictions are 100% wrong has an AUC of 0.; one whose predictions are 100% correct has an AUC of 1.

Additionally, it is extremely useful to see whether the classifier is confusing classes. Therefore, the number of correct and incorrect predictions are summarized with count values and broken down by each class in the confusion matrices in Fig. 5.30.

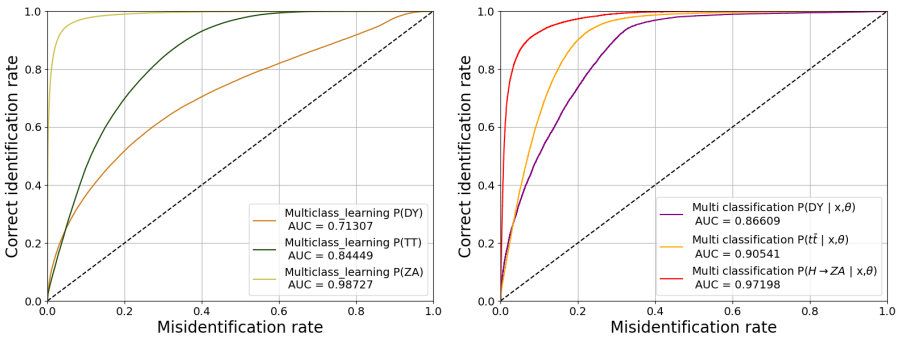


Figure 5.29. | Visualization of the performance of the parametrized DNN, using ROC ummed over full run 2 trained ggH signal samples. On the left, a multiclass ROC curve is shown with class weights applied, while on the right, the same curve is shown without the influence of class weights. These visualizations provide insight into how the inclusion of class weights impacts the model's capacity to differentiate between different classes.

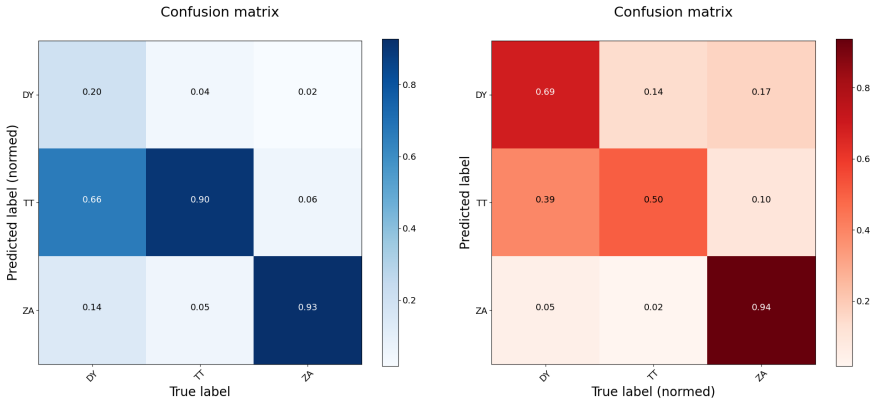


Figure 5.30. | Confusion matrices show the network's multi-class classification performance. Both plots demonstrate for each true class their contributions to the different predicted labels. The values in these figures are normalized over the predicted conditions (left) and over the true conditions (right). Values on the diagonal represent the counting of the true prediction so the higher these values the better the predictive ability of the model.

5.6.4. Bayesian Blocks binning

The almost universal practice of binning event data uses global methods, essentially averaging over the observation interval or sub-segments of it that are large enough to provide a good statistical sample. Although they are good for some problems, global methods dilute small local signals, throw away a considerable

amount of information, and introduce a dependency on bin sizes and locations.

In physics data analysis, it's common to handle a set of discrete measurements, requiring the grouping of these data points into finite-sized intervals, each having a midpoint and a total width. The practice of "binning" is extensively employed in statistics, frequently applied when constructing histograms. While histograms serve as valuable tools for visualizing data distributions, they come with certain drawbacks and constraints that merit attention when utilizing histogram-based approaches in statistical testing:

- ***Bin Selection:*** The choice of bin width and the number of bins can have a significant impact on the interpretation of the histogram. Different binning schemes can lead to different visual representations of the data distribution. If the bin width is too wide, important features or patterns in the data may be obscured. Conversely, if the bin width is too narrow, the histogram may exhibit excessive noise or variability. Determining an appropriate binning strategy is not an easy task and is often subjective and can introduce bias into the analysis.
- ***Loss of Information:*** Histograms summarize data by grouping values into bins. This process inevitably leads to a loss of information. By discretizing the data, the histogram fails to capture the exact values and nuances of the original dataset. This loss of information can impact the accuracy and precision of statistical testing, especially when dealing with small sample sizes or when subtle differences need to be detected.
- ***Sensitivity to Bin Width:*** The shape and appearance of a histogram can vary significantly depending on the bin width chosen. A different bin width may result in a different interpretation of the data distribution, potentially leading to different conclusions in statistical testing. This sensitivity to bin width makes histogramming less robust when compared to other non-parametric techniques like kernel density estimation.
- ***Arbitrary Cut-Off Points:*** In statistical testing, hypothesis tests often involve comparing distributions or assessing whether a value falls within a certain range. When using histograms, the determination of cut-off points or thresholds becomes arbitrary, as it depends on the chosen binning scheme. Different choices of binning can lead to different conclusions about the statistical significance of a result, making the interpretation subjective.
- ***Smoothing and Parameter Selection:*** Histograms can exhibit noise and fluctuations, especially with small sample sizes. Applying smoothing techniques, such as kernel density estimation, can alleviate this issue to some extent. However, these techniques require selecting appropriate parameters, such as the bandwidth, which can introduce bias if chosen suboptimally.

With modern statistical methods, there is usually no need to engage in binning,

since anything that can be done on "binned" discrete data can usually be done on the underlying continuous values. It is desirable to avoid analytical techniques that introduce arbitrary assumptions, especially in cases where alternative techniques are available to easily avoid these assumptions. In the next section, the Bayesian block (BB) approach will be introduced as an alternative method to detect localized structures, reveal pulse signal shapes, and generally characterize intensity variations avoiding any possible arbitrariness. The algorithm was originally developed for applications in astronomy [147] and is not commonly used in high-energy physics. We will see how this algorithm is used in this analysis and its interest in particle physics to reveal small signals without the need to make a rough parametric estimation, randomness, or smoothing which very often results in a degradation of signal sensitivity.

5.6.4.1. Approach

The Bayesian approach makes explicit use of the correct statistical distribution avoiding arbitrary binning. A sketch of standard Bayesian model fitting will set the notation and the context. Let's assume we have some data D, and a model M containing a parameter θ; if there are several parameters, simply we can interpret θ as a vector. We want to estimate how probable it is that the model is correct, and we want to learn something about the likely values of the parameter -all based on the data and any prior information that we might have.

The basic relation quantifying parameter inference is Bayes's theorem, one form of which is;

$$P(\theta|D,M)P(D|M) = P(D|\theta,M)P(\theta|M), \tag{5.5.}$$

where $P(\theta|D,M)$ is the posterior probability density of θ given the data, $P(D|M)$ the prior predictive probability for the data (often also called the global or marginal likelihood for the model), $P(D|\theta,M)$ the likelihood for the parameter, and $P(\theta|M)$ the prior probability of the parameter.

Prior predictive probability represents what is originally believed before new evidence is introduced, and posterior probability takes this new information into account. Posterior probability distributions should be a better reflection of the underlying truth of a data-generating process than the prior probability since the posterior includes more information. This decomposition can provide simple estimates of the width, location, and amplitude of pulses-assuming their overlap is neglectable -and of the background level, without invoking parametric or other explicit pulse-shape models.

In astronomical time series studies which was the inspiration for the use of the BB algorithm in this thesis, the objective was to set bin edges, called "change points", at times when the light flux from an astrophysical object suddenly changed.

The flux is represented by the arrival times t_i of photons in a telescope; given a set of event data t_i for $i = 1,...,N$. Readers unfamiliar with Bayesian time series analysis might consult [148], or the overview, with a specific discussion of the change-point problem. In our case, the change points play the role of histogram bin boundaries and the formed bin (block) is consistent with a probability density function (pdf), and the entire dataset is represented by the collection of finite pdfs.

$$F_{Total} = \sum_{k=1}^{N_{blocks}} f(B_k), \tag{5.6.}$$

where F_{Total} is the total fitness and $f(B_k)$ is the fitness of block k. The latter can be any convenient measure of how well a constant signal represents the data within the block. We define the fitness function as an unbinned log-likelihood [149]:

$$f(B_i) = ln(L_i(\lambda)) = N_i ln\lambda - \lambda T_i, \tag{5.7.}$$

The number of blocks and the edges of the blocks are determined through the optimization of the fitness function, which depends only on the size of the dataset and a regularization parameter P_0. To achieve a reasonable binning it is desirable to impose a prior that assigns a smaller probability to a large number of blocks and we adopt the following geometric prior [147]:

$$P(N_{blocks}) = P_0 \gamma^{N_{blocks}}, 0 \leq N_{blocks} \leq N, \tag{5.8.}$$

where P_0 is the false positive probability that is the relative frequency with which the algorithm falsely reports the detection of a change-point in data with no signal present. In general, the larger the number of bins, the small the P_0 should be to prevent the creation of spurious, jagged bins. The normalization constant P_0 is given by:

$$P_0 = \frac{1-\gamma}{1-\gamma^{N+1}}, \tag{5.9.}$$

where γ characterizes the assumed prior distribution for the number of blocks. An overall conservative value will suppress the detection of true change points, while a too-liberal value will lead to spurious change points (eventually reaching the limit of N_{blocks}=N). In the tail of the DNN scores distribution, it is expected from the Bayesian Blocks to misidentify the location of the beginning of the signal, and instead treat that area as uniform, due to immense falling background and the non-significant rising signal. Therefore, I tried to adapt a hybrid-binning strategy to isolate the signal into a few bins. It works by calculating the change points (bins) of the "background-only" and the "signal-only" separately (denoted

"B-only" and "S-only" respectively for simplicity), results of the rebinning are displayed in Fig. 5.31, and then combining the two sets of binning along with the removal of any "B-only" change points that are located within the bounds of the "S-only" region, the results are further illustrated in Fig. 5.32.

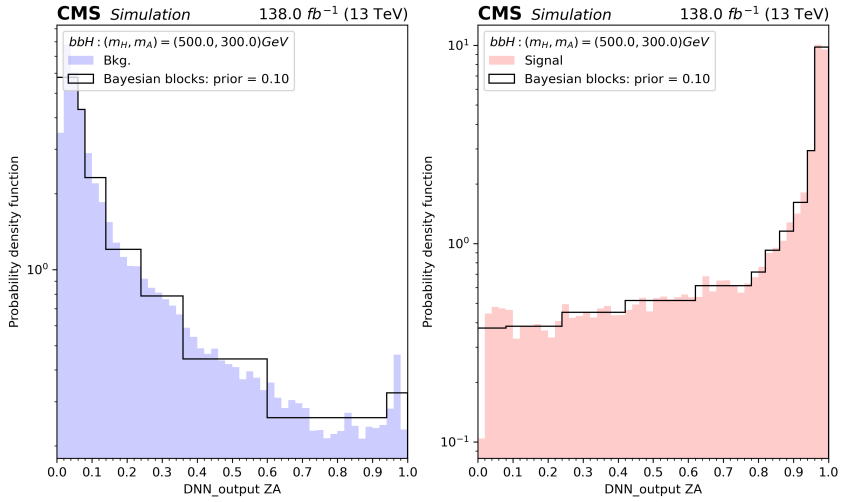


Figure 5.31. The Bayesian blocks binning is shown in black using "background-only" (left), "signal-only" (right) DNN scores template for signal mass hypothesis consisting of m_H =500 GeV and m_A =300 GeV. Events are constructed in $nb = 2$ -resolved category, $ee + \mu\mu$ channel. The Bayesian blocks are defined from a uniform DNN template of 50 bins shown in (red) and (blue) for signal and background events respectively.

"Background-only", "signal-only", and "hybrid" refer to the distribution used in the Bayesian blocks algorithm to define the binning, not the content of the histogram. In the "S-only", the signal histogram is used alone, leading to small bins in the signal regions and wide bins in the background region, while in the "B-only" a sum of background histograms is used, leading to the reverse scenario of "S-only". In the "hybrid" case, background histograms and signal histograms are used exclusively from the DNN template to set the binning, from [0,0.6] DNN scores, only background histograms are used while from [0.6,1] the signal histogram alone is used.

5.6.4.2. The final templates

Given the parametric approach of the used DNN, this will result in a different BB binning strategy for each signal mass hypothesis, category, and production mechanism. The final template obtained from the BB algorithm for a given mass

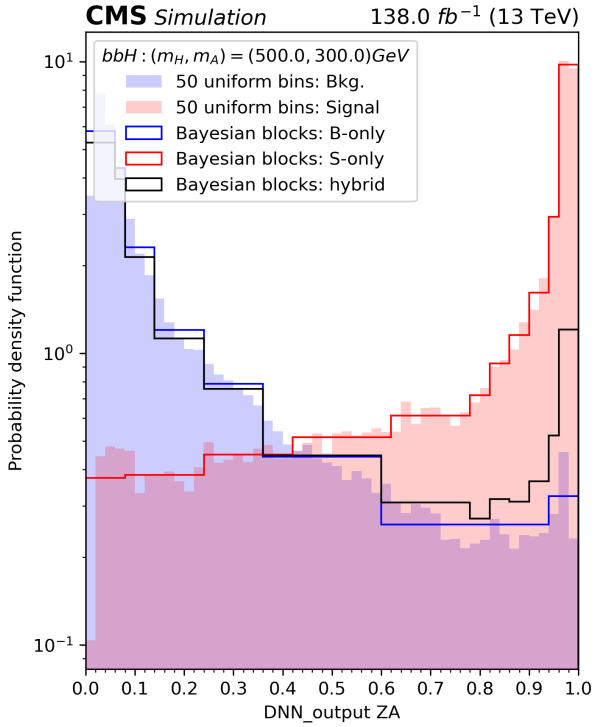


Figure 5.32. Different scenarios using the Bayesian blocks binning is shown; "background-only" (blue), "signal-only" (red) and "hybrid" (black). The DNN scores template is used for the signal mass hypothesis consisting of $m_H = 500$ GeV and $m_A = 300$ GeV. Events are constructed in $nb = 2$ -resolved category, $ee + \mu\mu$ channel. The Bayesian blocks are defined from a uniform DNN template of 50 bins shown in (red) and (blue) for signal and background events respectively.

is shown in Fig. 5.33. These templates are then used in the maximum likelihood fit. The fit is applied simultaneously to both the signal region categories (i.e. those with $\mu\mu$, ee, or $\mu\mu + ee$ lepton flavors) and the μe control region [1].

[1] This essentially means performing a multi-channel (each channel is a DNN output score within the range of [0, 1]) fit where the different categories are treated as separate channels. The observed data and the expected signal and background distributions in each category are taken into account collectively in the fit. The main idea is to gain more statistical power by exploiting the information from different categories simultaneously.

5.7. Systematic uncertainties

The impacts on the statistical interpretation of the result of systematic uncertainties affecting normalization and the shape of the kinematic distributions were studied. Experimental, theoretical, statistical, and modeling uncertainties on the normalization and shape of the signal and background estimates give a non-negligible contribution to the total uncertainty and are thus evaluated for use in the statistical model; these are described below.

1. **Theoretical uncertainties**
   - **Background theoretical cross sections**: Theoretical uncertainties in the cross sections of backgrounds estimated using simulation are considered systematic uncertainties in the yield predictions.
   - **QCD scale uncertainty**: This uncertainty is estimated by varying the renormalization (μ_R) and the factorization (μ_F) scales, used during the MC generation of the sample, independently by a factor 0.5, 1 or 2. Unphysical cases, where one scale fluctuates up while the other fluctuates down, are not considered. An envelope is built from all 6 possible variations by taking in each bin of the distribution the maximum (minimum) variation and is used as an estimate of the QCD scale uncertainties for all the background and signal samples.
   - **Parton Distribution Functions (PDFs) uncertainty**: The magnitude of the uncertainties related to the parton distribution functions and the variation of the strong coupling constant for each simulated background and signal process is obtained using the replicas of the NNPDF3.1. set [127].
2. **Experimental uncertainties**
   - **Luminosity**: Uncertainties on the estimation of the luminosity follow the recommendations of the LUMI POG [150]. The uncertainties are applied as rate-only log-normal uncertainties. The uncertainty values per era and correlations between years are summarised in Table 5.8.

Table 5.8. | Uncertainties on the luminosity estimation in percent.

	2016	2017	2018
Uncorrelated 2016	1.0%		
Uncorrelated 2017		2.0%	
Uncorrelated 2018			1.5%
Correlated 2016/2017/2018	0.6%	0.9%	2.0%
Correlated 2017/2018		0.6%	0.2%

   - **Drell-Yan reweighting**: The derived weights in Sec. 5.5.2 are considered as a source of systematic uncertainty for the DY+jets background

during the statistic test. Uncertainties are applied as uncorrelated shapes across the three-year, resolved, and boosted categories and lepton flavors. The systematics variations are summarised in Table. 5.9.

Table 5.9. | Summary of the event weight used to reweight DY+jets events together with its up and down variations, assigned as a source of systematic uncertainty.

Nominal	w
Up	2 × (w -1) +1
Down	(w -1)/2 + 1

- **Lepton identification, reconstruction and isolation**: Uncertainties on the electrons and muons identification, reconstruction, and isolation are determined and provided centrally, extracted with a "tag-and-probe" analysis on $Z \to ll$ events. The impacts of this uncertainty on the shape and normalization of the $t\bar{t}$ process in the DNN score template is shown in Fig. 5.34.
- **Jet energy scale and resolution**: Uncertainties in the jet energy scale are of the order of a few percent as a function of jet p_{T} and η. We vary the jet energy scale by $\pm 1\sigma$. A difference in the jet energy resolution of about 10% between data and simulation is accounted for by worsening the jet energy resolution in the simulation by η-dependent factors. The uncertainty due to this is estimated by a variation of the factors applied by $\pm 1\sigma$. We treat both JER/JES as one combined scheme of uncertainty sources, uncorrelated across all years.
  In 2018 an additional uncertainty source was applied to cover the effects of the HEM1516 issue [1]. This additional uncertainty consists of a 20% downward variation of the jet energy for jets with $-1.57 < \phi < -0.87$ and $-2.5 < \eta < -1.3$.
- **Trigger efficiency**: Trigger efficiencies are evaluated using the "tag-and-probe" technique on $Z \to \ell\ell$ events. Uncertainties on this measurement are considered a source of systematic uncertainties.
- **b-tagging**: The b-tag and light-flavor mistag rate uncertainties are determined as a function of the jet p_{T} and η. Effect on the analysis is estimated by varying these corrections respectively by one standard deviation, separately for heavy- and light-flavour jets.
- **Pileup**: The measured minimum-bias cross section is varied by ±5% to produce different expected pileup distributions.

[1] The endcaps of the hadron calorimeter failed to cover the phase space at $-3 < \eta < -1.3$ and $-1.57 < \phi < -0.87$ during the 2018 data-taking period (referred to as the hadronic calorimeter endcaps minus (HEM) issue) which causes a significant increase in the multijet background. The transverse momentum of the jets in this region is typically under-measured, which results in over-measured E_T^{miss}

3. **MC statistics**: Our estimates of the backgrounds come mainly from MC simulation and in this case, these estimates are subject to a statistical uncertainty on the number of simulated events. In principle, we should include an independent statistical uncertainty for every bin of every process in our model. We take the uncertainty from statistical fluctuations in the predictions into account following a method similar to the one proposed by Barlow and Beeston [151]. We use a simplification of their method, often called the Barlow-Beeston-lite approach, which uses a single nuisance parameter per bin, rather than one for every sample in every bin [152].
   In the Barlow-Beeston method, each sample receives a nuisance parameter in each bin which multiplies the bin yield and is constrained according to the pdf of the number of expected events [1]. While, in the full method, the expected pdf is taken to be a Poisson distribution, based on the number of events in the sample. Additionally, rather than constraining the nuisance parameter using a Poisson distribution, it is constrained using a Gaussian. The effective number of events in each bin, n_{eff}, defines the variance of the Gaussian and is calculated as:

$$n_{\text{eff}} = \frac{n^2}{\sum_{\text{proc}} e^2_{\text{proc}}}, \tag{5.10.}$$

   Where n is the sum of events from all background sources, and e_{proc} is the statistical uncertainty on a given process.
   The largest source of uncertainty in this search arises from the statistical uncertainty of the data in the $nb3$ signal region. Nonetheless, experimental, theoretical, and modeling uncertainties626 on the normalization and shape of the signal and background estimates give a non-negligible contribution to the total uncertainty. It is important to note that since the probed mass range is large and the parametric nature of the DNN, we expect nuisance parameters to yield different impacts depending on the considered mass hypothesis.

Correlation matrices are useful when identifying relationships and dependencies between nuisance parameters, indicating that changes in one parameter are associated with predictable changes in another. This knowledge is essential for understanding the behavior of the model and the impact of the uncertainties on the final results[2].

It's important to note that the specific calculation and interpretation of the correlation matrix may vary depending on the statistical fit procedure and the

[1] Meaning only a single parameter is applied for the yield in each bin, summed over all processes, including the signal.

[2] Correlation matrices quantify the linear relationship between pairs of nuisance parameters, ranging from -1 to 1. A correlation coefficient of 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates no correlation.

Table 5.10. | Correlations of systematic uncertainties source.

Uncertainty source	Type	# sources	Year correlation	Comment
μ_R scale	shape	1/year	✓	Separately for each process
μ_F scale	shape	1/year	✓	Separately for each process
PDF	shape	1/year	✓	Separately for each process
PS scale ISR	shape	1/year	✓	Separately for each process
PS scale FSR	shape	1/year	✓	Separately for each process
$t\bar{t}$ cross-section	rate	1	✓	Only for $t\bar{t}$
Drell-Yan cross-section	rate	1	✓	Only for DY
Single-Top cross-section	rate	1	✓	Only for Single-Top
Luminosity	rate	1/year + 2	partially	See Table 5.8
Lepton scale factors	shape	4/year	—	Reco and ID for electrons, ID and iso for muons
Trigger efficiency	shape	3/year	—	One for each pair of ee, $\mu\mu$, and μe
L1 prefire correction	shape	1	✓	only for 2016 and 2017
Pileup	shape	1	✓	
Jet energy scale	shape	1/year	—	combined scheme
Jet energy resolution	shape	1/year	—	combined scheme
HEM issue	shape	1	—	Only for 2018
b tagging light flavour	shape	1/year + 1	partially	
b tagging heavy flavour	shape	1/year + 9	partially	Pileup correlated with other pileup variations
Top pt re-weighting	shape	1	✓	
DY re-weighting	shape	1/year/category	—	
Size of MC samples	shape	1/bin	—	

assumed mass parameter of the signal.

The analyses conducted on various masses and the signal region categories for the two production modes (ggH and bbH) required a separation of the DY+jet samples outlined in Table 5.2 of Section 5.1.2.3 into subprocesses associated with DY+(0, 1, 2) B hadrons. Similarly, a comparable approach was applied to $t\bar{t}$, splitting it into distinct samples with a minimum of (0, 1, 2) B hadrons within a jet. As a consequence, the decorrelation of all interdependencies among the theoretical uncertainties associated with these processes. The following allows for model reparameterization and simplification, transforming possible anti-correlated parameters into independent or weakly correlated parameters. The effects on the total yields in the signal region are summarised in Tab. 5.11 for a given mass of bbH signal.

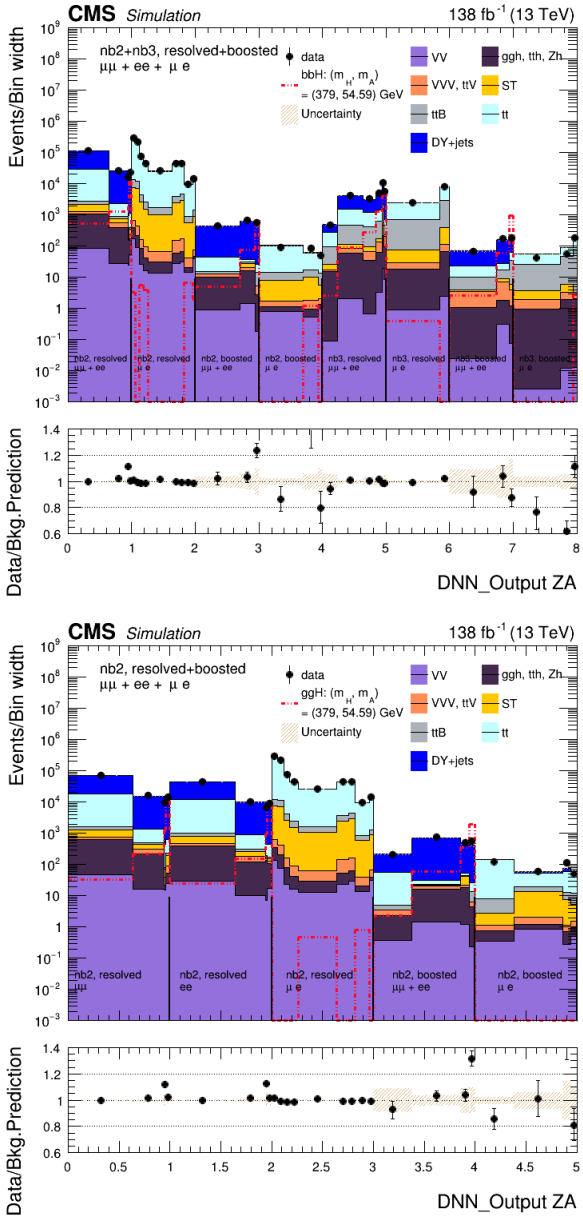


Figure 5.33. Unblinded post-fit distributions. The final templates are built using the "S-only" Bayesian block binning in the signal region category (i.e. ee,$\mu\mu$, $ee+\mu\mu$) flavors of lepton) and "B-only" in the control region (i.e. μe). The fit is performed on full run 2 data assuming a background-only hypothesis ($\hat{r}=0$). The signal mass is defined $(m_H, m_A) = (379, 54.59)$ GeV for signal bbH (top) and signal ggH (bottom). Error bars indicate statistical uncertainties, while shaded bands show post-fit systematic uncertainties. The signal (dashed red) is normalized to $H \to ZA \to \ell^+\ell^- b\bar{b}$ cross-section and branching ratio in the context of 2HDM type-II for $cos(\beta-\alpha)$= 0.01 and tanβ= 5.

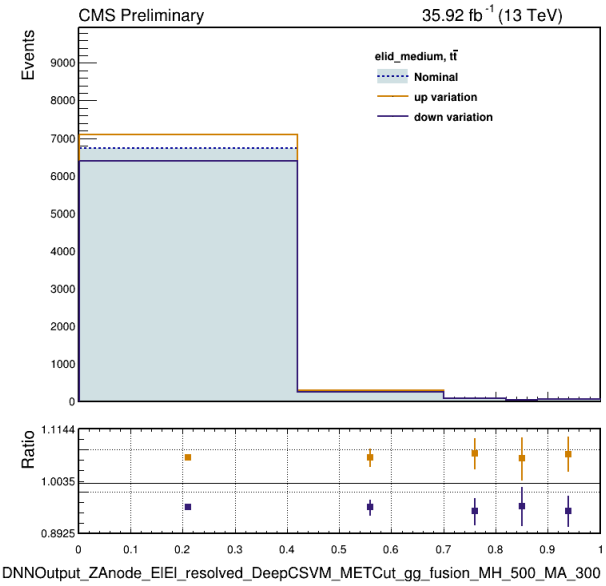


Figure 5.34. | Effect of the electron identification systematic variations on the DNN scores template after adopting the Bayesian blocks rebinning "signal-only" strategy. The DNN is evaluated on 2016 ULegacy pre-/post-VFP $t\bar{t}$ background processes and on the simulated signal samples for mass parameters $m_H = 500$ GeV, $m_A = 300$ GeV. Events are constructed in $nb = 2$ -resolved category, ee lepton flavor.

Table 5.11. | Summary of the systematic uncertainties and their impact range on total yields in the signal region $nb2+nb3$ -resolved category, $(ee+\mu\mu)$ combined leptons flavor for background and for bbH signal mass hypothesis $m_H = 379$ GeV, $m_A = 44.72$ GeV of the final DNN template.

Source	Background yield variation	Signal yield variation
Light flavor jet b-tagging (DeepJetM)	4.0%	0.3%
Jet energy scale FlavorQCD uncertainty	2.8%	1.2%
Electron identification	2.4%	2.1%
Jet energy scale Absolute uncertainty	1.9%	0.6%
Heavy flavor jet b-tagging (DeepJetM)	1.8%	1.8%
Correlated luminosity 2016,2017,2018	1.2%	1.3%
Jet energy scale BBEC1 uncertainty	1.1%	0.4%
Jet energy scale 2018 uncertainty	1.1%	0.5%
Light flavour jet b-tagging (DeepJetM)	0.9%	< 0.1%
Jet energy scale RelativeBal uncertainty	0.9%	0.3%
MC statistics	< 0.1%	< 0.1%
Affecting only DY0b (18.3% of the total bkg.)		
QCD renormalisation scale (μ_R) uncertainty	4.7%	
QCD factorization scale (μ_F) uncertainty	2.6%	
QCDMuRF_DY0b	1.7%	
Parton shower initial state (ISR) uncertainty	1.4%	
Parton distribution functions (PDFs) uncertainty	0.6%	
Parton shower final state (FSR) uncertainty	0.3%	
MC statistics	< 0.1%	
Affecting only DY1b (13.3% of the total bkg.)		
QCD renormalisation scale (μ_R) uncertainty	7.2%	
QCD factorization scale (μ_F) uncertainty	4.8%	
Parton shower initial state (ISR) uncertainty	2.8%	
QCDMuRF_DY1b	2.0%	
Parton distribution functions (PDFs) uncertainty	0.7%	
Parton shower final state (FSR) uncertainty	0.4%	
MC statistics	< 0.1%	
Affecting only DY2b (36.6% of the total bkg.)		
QCD renormalisation scale (μ_R) uncertainty	10.1%	
QCDMuRF_DY2b	8.8%	
Parton shower initial state (ISR) uncertainty	1.6%	
QCD factorization scale (μ_F) uncertainty	1.4%	
Parton shower final state (FSR) uncertainty	1.3%	
Parton distribution functions (PDFs) uncertainty	0.6%	
MC statistics	< 0.1%	
Affecting only tt (29.0% of the total bkg.)		
QCDMuRF_tt	11.4%	
QCD renormalisation scale (μ_R) uncertainty	9.6%	
Parton shower final state (FSR) uncertainty	2.8%	
QCD factorization scale (μ_F) uncertainty	1.4%	
Parton distribution functions (PDFs) uncertainty	0.4%	
TopPt_reweighting	0.4%	
Parton shower initial state (ISR) uncertainty	0.2%	
MC statistics	< 0.1%	
Affecting only ttB (0.9% of the total bkg.)		
QCDMuRF_ttB	12.3%	
QCD renormalisation scale (μ_R) uncertainty	10.6%	
Parton shower final state (FSR) uncertainty	9.2%	
Parton shower initial state (ISR) uncertainty	1.8%	
QCD factorization scale (μ_F) uncertainty	1.6%	
Parton distribution functions (PDFs) uncertainty	0.5%	

Continue on the next page

Table 5.11. | Summary of the systematic uncertainties and their impact range on total yields in the signal region $nb2+nb3$ -resolved category, $(ee+\mu\mu)$ combined leptons flavor for background and for bbH signal mass hypothesis $m_H = 379$ GeV, $m_A = 44.72$ GeV of the final DNN template (continuation).

Source	Background yield variation	Signal yield variation
TopPt_reweighting	0.2%	
MC statistics	< 0.1%	
Affecting only Single top (0.8% of the total bkg.)		
QCD renormalisation scale (μ_R) uncertainty	5.1%	
QCD factorization scale (μ_F) uncertainty	2.6%	
QCDMuRF_SingleTop	2.2%	
Parton shower final state (FSR) uncertainty	2.0%	
Parton shower initial state (ISR) uncertainty	1.3%	
Parton distribution functions (PDFs) uncertainty	0.5%	
MC statistics	< 0.1%	
Affecting only bbH signal		
QCD renormalisation scale (μ_R) uncertainty	22.5%	
QCD factorization scale (μ_F) uncertainty	11.1%	
Parton distribution functions (PDFs) uncertainty	1.6%	
Parton shower final state (FSR) uncertainty	1.6%	
Parton shower initial state (ISR) uncertainty	< 0.1%	
QCDMuRF_bbH	< 0.1%	
MC statistics	< 0.1%	

In addition, if we want to evaluate the behavior of the profile likelihood, we very often look at the nuisance parameters pulls and impacts; Pulls quantify how far we have to "pull" the nuisance parameter from its expected value to find the MLE. While the impact gives a measure of the "impact" of the variation of the nuisance parameter on the parameter of interest. Fig. 5.35 illustrate the main impacts of nuisance parameters on signal mass parameters ggH $(m_H, m_A) = (500, 250)$ GeV.

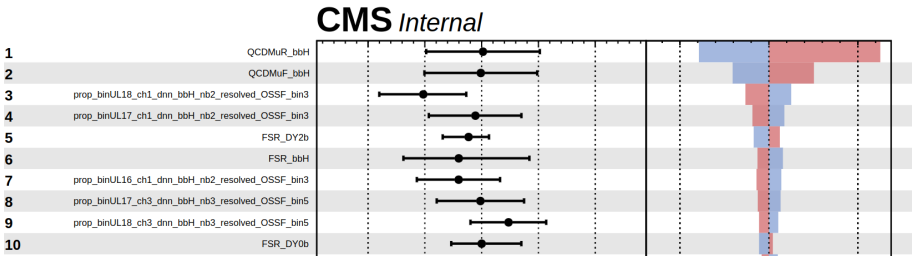


Figure 5.35. The first ten leading nuisance parameters with the largest impacts on the fitted signal strength $\hat{r}$ assuming no signal ($\hat{r}$= 0) ranked from highest impact to lowest. The fitted DNN templates corresponding to full run2 data, $nb2+nb3$ - resolved category, ($ee+\mu\mu+\mu e$) combined lepton flavours, and for bbH signal mass parameters (m_H, m_A)= (379, 44.72) GeV. The left panel illustrate the shifts between post-/pre-fit values ($\widehat{\Theta}-\Theta_0$) and uncertainties of these parameters, relative to their pre-fit uncertainty $\Delta\Theta$, while the right panel shows the impact of each parameter on the signal strength when varied ±1 standard deviation.

5.8. Results and interpretation

5.8.1. Significance and p-value scan

A posteriori expected significance is computed for full run 2 data from an Asimov dataset of signal+background and illustrated in Fig. 5.37. The signal is normalized to ggH production cross-section and branching ratio BR($H \to ZA$) $\times$ BR($Z \to \ell^+\ell^-$) $\times$ BR($A \to b\bar{b}$) in the context of type-II 2HDM. The expected p-value scan as was described in Chap. 4 indicates how likely it is that our data would have occurred by random chance (i.e. that the null hypothesis is true). The result from the p-value scan displayed in Fig. 5.36 and Fig. 5.37 indicate that lighter pseudo-scalar masses ($m_A \leq 200$ GeV), which correspond to heavier Higgs masses ($m_H \leq 500$ GeV) in the 2HDM type-II, might impose the most stringent constraints on the 2HDM parameters and potentially lead to the rejection of the null hypothesis.

The computed local p-value in the (m_A, m_H) plane was obtained by evaluating the parametric DNN output at each simulated mass point. These points corresponded to signal benchmarks with their respective cross-sections and branching fractions for type II 2HDM, specifically for $cos(\beta-\alpha)$=0.01 and tanβ=1.5 (20.) for ggH (bbH) signal. Mass values falling between these points were extrapolated.

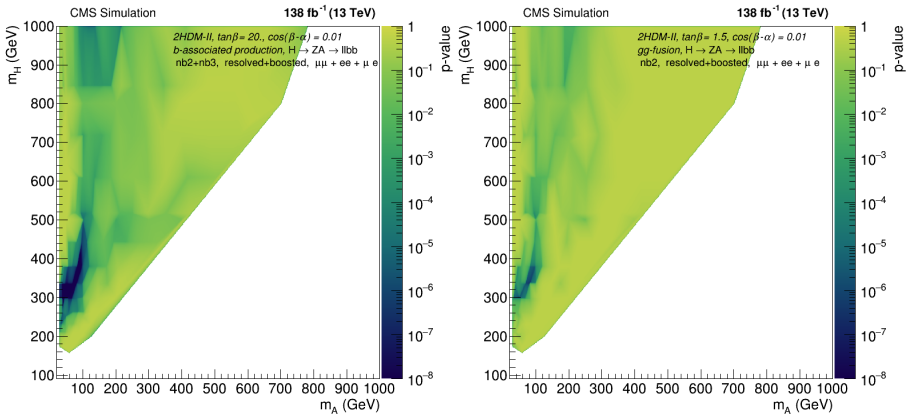


Figure 5.36. Expected local p-values for a 2HDM Higgs boson mass (m_H) as a function of the light pseudo-scalar (m_A) are presented, considering the decay mode $H \to ZA$ and two production mechanisms: b-associated production (left) and gluon-gluon fusion production (right). These values are derived from the combination of $nb2+nb3$ (left) and $nb2$ (right) in both resolved and boosted signal region categories. The mass splitting procedure outlined in Sec.5.6.1 is applied, and the results are obtained for a combined lepton flavor ($ee+\mu\mu+\mu e$).

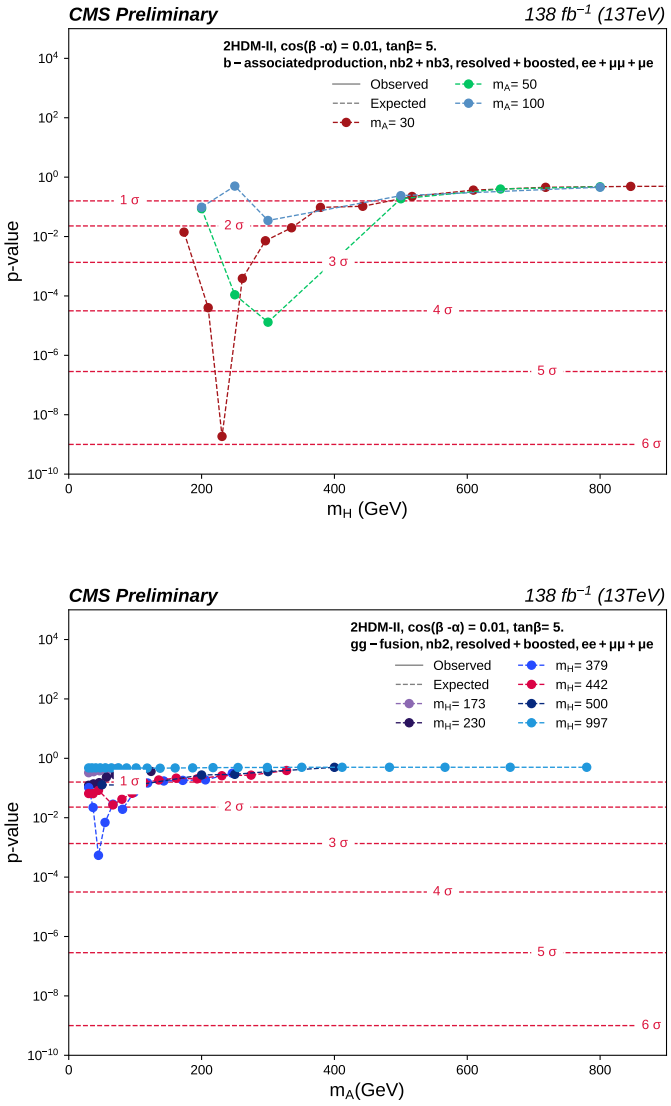


Figure 5.37. The local p-value for various benchmark signal hypothesis. Top: bbH expected significance is shown for a fixed value of the pseudo-scalar A mass, m_A, and as a function of the heavy Higgs H mass, m_H. Bottom: ggH expected significance is shown for a fixed value of the heavy Higgs H mass, m_H, and as a function of the pseudo-scalar A mass, m_A. The dashed curves show the median expected local p-value while the horizontal dashed lines indicate the p-values corresponding to the significance of 1 to 6 σ.

5.8.2. 95% confidence level upper limit scan

5.8.2.1. 1D scan

- $H \to ZA$

Assuming the absence of any signal surpassing the anticipated background levels, a confirmation or rejection of this assumption will be found in the upcoming publication [153]. We can set the upper limits at 95% confidence level on the product of the production cross section and branching fraction for $H \to ZA \to \ell^+\ell^- b\bar{b}$ using the asymptotic modified frequentist method (asymptotic CL_s) [154] as a function of H and A mass hypothesis. 1D slices are shown in this section for a fixed mass of m_H for clarity and illustration. Fig. 5.38 shows the expected limits from full run 2 data for given values of the heavy Higgs mass.

Fig. 5.39 effectively illustrates that this analysis has attained sensitivity in the low-mass regions through the utilization of boosted categories. In summary, when $m_H > 4 \times m_A$, the anticipated upper limits on σ are superior in the boosted categories compared to the resolved ones. This plot also highlights the complementary nature of both production mechanisms, i.e., ggH and bbH. For instance, gluon-gluon fusion predominates at lower values of $\tan\beta$, meaning that certain mass points such as the one shown on the plot can only be excluded when focusing on the lower end of this 2HDM parameter space. Conversely, for b-associated production, which dominates at higher values of $\tan\beta$ due to the enhanced bottom Yukawa coupling proportional to $\tan\beta$, such masses cannot be excluded unless we consider the higher range of $\tan\beta$.

The enhancement achieved by incorporating the boosted signal category is evident in Tab. 5.12. This improvement is particularly noticeable for a signal mass parameter of $(m_H, m_A) = (650, 50)$ GeV, where the signal events are expected to be significant in the boosted region due to the substantial mass difference between the H and A bosons.

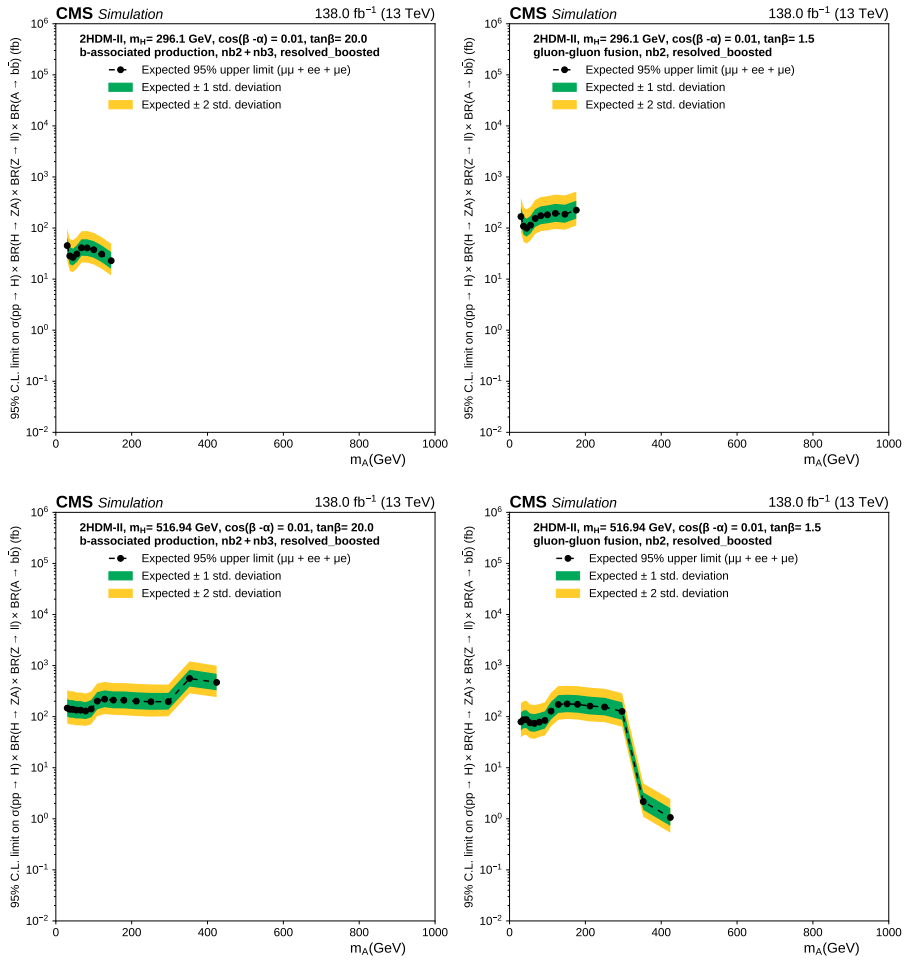


Figure 5.38. Upper limits on $H \to ZA \to \ell^+\ell^- b\bar{b}$ production cross section times branching ratio for bbH signal (left) and ggH (right) using the $nb2+nb3$ category for bbH and $nb2$ for ggH. The limits are shown for a combination of resolved+boosted regions, following the mass splitting detailed in Sec. 5.6.1 for two fixed values of the Heavy Higgs, H, mass, $m_H = 296.1$ GeV (top) and $m_H = 516.94$ GeV (bottom). The dashed line represents the expected limit, computed for full run 2 data using the asymptotic CL_s method. The green and yellow bands represent the 1 and 2 standard deviations for the expected limit.

- ***$A \to ZH$***

Recall that the idea behind generating only a limited set of mass points in the context of $A \to ZH$, which are intentionally identically produced for $H \to ZA$, is to validate the interchangeability of these masses. This approach was employed

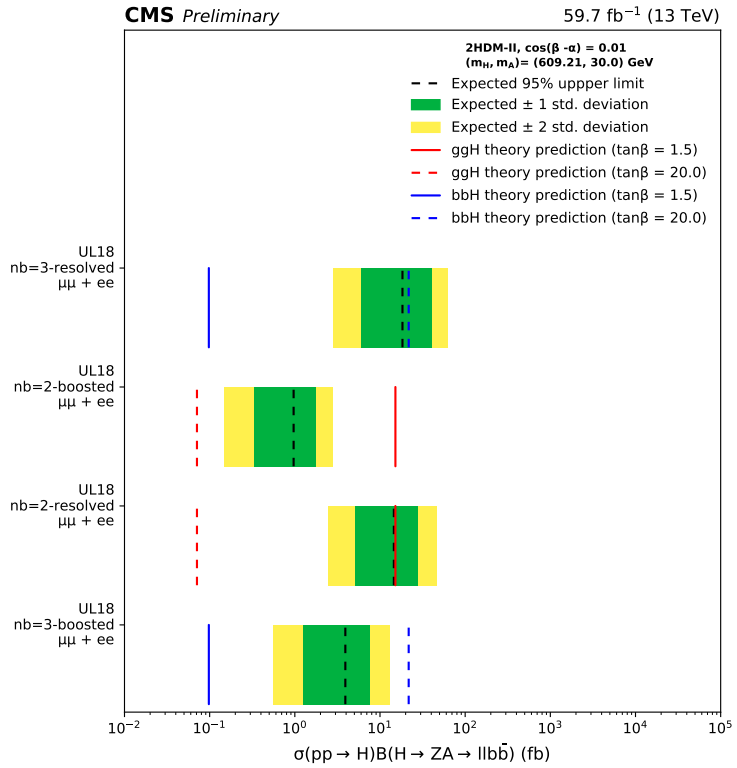


Figure 5.39. | Expected 95% CL upper limits on $H \to ZA \to \ell^+\ell^- b\bar{b}$ production cross section times branching ratio for $(m_H, m_A) = (609.21, 30)$ GeV ggH limits are shown in $nb = 2$ -resolved, $nb = 2$ -boosted, while bbH limits are shown in $nb = 3$ -resolved, and $nb = 3$ -boosted. Expected upper limits are calculated for ULegacy 2018 data using the asymptotic CL_{s+b} method.

in a prior search [34]. Although both processes culminate in the same final state ($\ell^+\ell^- b\bar{b}$), distinctions can emerge in terms of event characteristics, signatures, couplings, kinematics, and branching ratios. If these discrepancies prove to be negligible, it would obviate the need for an extensive Monte-Carlo simulation in both scenarios, as it becomes apparent that $H \to ZA$ and $A \to ZH$ are essentially equivalent through the reconstruction process.

The expected limits for the $A \to ZH$ scenario are presented for two representative mass points in Tab. 5.13 and Tab. 5.14. For a more comprehensive and detailed set of results, please consult the forthcoming publication in [153].

Table 5.12. | Observed and expected limits for 2HDM type II signal mass $(m_H, m_A) = (500, 50)$ GeV signal mass hypothesis. The outcomes are derived within the context of the type II 2HDM, with parameters set at $cos(\beta - \alpha) = 0.01$ and $\tan\beta = 1.5$ (20.) for gluon-gluon fusion (b-associated production) mechanism and have been obtained for CMS full run 2 datasets.

Cat.	Observed (fb)	Expected (fb)	±1 Standard deviation (fb)	±2 Standard deviations (fb)
$bbH: nb2, boosted,\ \mu\mu + ee$	-	14.16	11.568 ± 9.334	27.646 ± 11.948
$bbH: nb2, resolved,\ \mu\mu + ee$	-	17.09	15.26 ± 11.057	36.217 ± 14.153
$bbH: nb3, boosted,\ \mu\mu + ee$	-	15.137	12.941 ± 9.978	30.47 ± 12.772
$bbH: nb3, resolved,\ \mu\mu + ee$	-	18.066	15.102 ± 11.688	35.375 ± 14.961
$bbH: nb2 + nb3, boosted,\ \mu\mu + ee$	-	10.254	8.961 ± 6.509	21.441 ± 8.331
$bbH: nb2 + nb3, resolved,\ \mu\mu + ee$	-	10.254	8.182 ± 6.759	19.438 ± 8.652
$bbH: nb2 + nb3, resolved + boosted,\ \mu\mu + ee$	-	7.324	5.288 ± 4.835	12.724 ± 5.951
$ggH: nb2, boosted,\ \mu\mu + ee$	-	8.301	5.993 ± 5.691	14.42 ± 7.004
$ggH: nb2, resolved,\ \mu\mu + ee$	-	8.301	5.993 ± 5.691	14.42 ± 7.004
$ggH: nb3, boosted,\ \mu\mu + ee$	-	8.301	7.254 ± 5.48	15.741 ± 6.744
$ggH: nb3, resolved,\ \mu\mu + ee$	-	105.957	92.6 ± 68.551	216.402 ± 87.746
$ggH: nb2 + nb3, boosted,\ \mu\mu + ee$	-	6.348	4.341 ± 4.19	10.236 ± 5.157
$ggH: nb2 + nb3, resolved,\ \mu\mu + ee$	-	7.324	5.844 ± 4.835	12.95 ± 5.951
$ggH: nb2 + nb3, resolved + boosted,\ \mu\mu + ee$	-	4.395	3.674 ± 3.124	8.495 ± 3.571

Table 5.13. | 95% CL_s upper limits on $A \to ZH \to \ell^+\ell^- b\bar{b}$ production cross section times branching ratio $B(A \to ZH) \times B(Z \to \ell^+\ell^-) \times B(H \to b\bar{b})$ for $(m_A, m_H) = (500, 250)$ GeV signal mass hypothesis. The outcomes are derived within the context of the type II 2HDM, with parameters set at $cos(\beta - \alpha) = 0.01$ and $\tan\beta = 1.5$ (20.) for production mechanisms of gluon-gluon fusion (b-associated) and have been obtained for CMS full run 2 datasets.

Cat.	Observed (fb)	Expected (fb)	±1 Standard deviation (fb)	±2 Standard deviations (fb)
$bbA: nb2 + nb3, boosted,\ \mu\mu + ee + \mu e$	-	52.189	28.916 ± 16.771	76.806 ± 26.502
$bbA: nb2 + nb3, resolved,\ \mu\mu + ee + \mu e$	-	10.529	5.75 ± 3.358	14.48 ± 5.306
$bbA: nb2 + nb3, resolved + boosted,\ \mu\mu + ee + \mu e$	-	10.255	5.6 ± 3.295	14.102 ± 5.207
$ggA: nb2, boosted,\ \mu\mu + ee + \mu e$	-	265.012	155.284 ± 88.203	411.325 ± 137.682
$ggA: nb2, resolved,\ \mu\mu + ee + \mu e$	-	3.976	2.061 ± 1.258	5.112 ± 1.988
$ggA: nb2, resolved + boosted,\ \mu\mu + ee + \mu e$	-	3.932	2.037 ± 1.244	5.054 ± 1.966

- ***H/A → ZA/H(125)***

A comparison of the expected upper limits of $H \to ZA(125) \to \ell^+\ell^- b\bar{b}$ and $A \to ZH(125) \to \ell^+\ell^- b\bar{b}$ at 95% CL can be seen in Tab. 5.15.

A was demonstrated in Chap. 2 H and A bosons have distinct properties and differ in their couplings to fermions and gauge bosons. Their specific properties depend on the choice of the 2HDM model and its parameter space. The H and

Table 5.14. | 95% CL_s upper limits on $A \to ZH \to \ell^+\ell^- b\bar{b}$ production cross section times branching ratio $B(A \to ZH) \times B(Z \to \ell^+\ell^-) \times B(H \to b\bar{b})$ for $(m_A, m_H) = (780, 680)$ GeV signal mass hypothesis. The outcomes are derived within the context of the type II 2HDM, with parameters set at $cos(\beta - \alpha) = 0.01$ and $\tan\beta = 1.5$ (20.) for production mechanisms of gluon-gluon fusion (b-associated) and have been obtained for CMS full run 2 datasets.

Cat.	Observed (fb)	Expected (fb)	±1 Standard deviation (fb)	±2 Standard deviations (fb)
$bbA: nb2+nb3$,boosted, $\mu\mu+ee+\mu e$	-	8.022	4.413 ± 2.594	11.233 ± 4.074
$bbA: nb2+nb3$,resolved, $\mu\mu+ee+\mu e$	-	3.419	1.922 ± 1.121	4.964 ± 1.75
$bbA: nb2+nb3$,resolved+boosted, $\mu\mu+ee+\mu e$	-	3.036	1.694 ± 0.988	4.332 ± 1.542
$ggA: nb2$,boosted, $\mu\mu+ee+\mu e$	-	0.178	0.107 ± 0.061	0.28 ± 0.095
$ggA: nb2$,resolved, $\mu\mu+ee+\mu e$	-	0.003	0.001 ± 0.001	0.004 ± 0.001
$ggA: nb2$,resolved+boosted, $\mu\mu+ee+\mu e$	-	0.003	0.001 ± 0.001	0.004 ± 0.001

A bosons correspond to the two neutral scalar fields in the 2HDM. The H boson is typically referred to as the "CP-even" scalar, while the A boson is known as the "CP-odd" scalar or pseudo-scalar [1].

While conducting an analysis focused on the $\ell^+\ell^- b\bar{b}$ final state, it was initially considered challenging to discern discrepancies between the H and A bosons. Surprisingly, the expected upper limits displayed in Tab. 5.15 revealed a substantial disparity. It was previously believed that only precise measurements, including production rates, decay branching ratios, and angular distributions, would be instrumental in distinguishing between the H and A bosons. The initial investigation suggests that the DNN training, which was applied simultaneously to both production processes, managed to successfully differentiate between these two processes. However, for a more comprehensive understanding, it is recommended to refer to the forthcoming publication [153].

5.8.2.2. 2D scan

Similar to the one-dimensional limit scans we can also set upper bounds in the two-dimensional plane of 2HDM parameters. Upper limits at 95% CL on the product of the production cross-section and branching ratio for $H \to ZA \to \ell^+\ell^- b\bar{b}$ are set in the mass plane as a function of the mass hypothesis m_H and m_A. These results are shown for different signal points in Fig. 5.40 and are interpreted in the context of the type-II 2HDM for the theoretical benchmark $cos(\beta - \alpha)$ = 0.01 and $\tan\beta$ = 1.5 (20.) for ggH (bbH). In order to describe the entire mass plane, the outcomes for intermediate mass values are interpolated, as demonstrated in Fig. 5.41.

Certain results may exhibit the utilization of the $nb2+nb3$ category for gluon-gluon fusion or the inclusion of all mass points in both resolved and boosted

[1] The CP (charge-parity) quantum number describes the symmetry properties of a particle under charge conjugation (C) and parity (P) transformations.

Table 5.15. | 95% CL_s upper limits on $H/A \rightarrow ZA/H(125) \rightarrow \ell^+\ell^-$ b-associated production cross section times branching ratio $B(H/A \rightarrow ZA/H) \times B(Z \rightarrow \ell^+\ell^-) \times B(A/H \rightarrow b\bar{b})$ for $(m_{H/A}, m_{A/H}) = (200, 125)$ GeV signal mass hypothesis. The outcomes are derived within the context of the type II 2HDM, with parameters set at $cos(\beta - \alpha) = 0.01$ and $\tan\beta = 1.5$ (20.) for production mechanisms of gluon-gluon fusion (b-associated) and have been obtained for CMS full run 2 datasets.

Cat.	Observed (fb)	Expected (fb)	±1 Standard deviation (fb)	±2 Standard deviations (fb)
$bbH: nb2+nb3$,boosted, $\mu\mu+ee$	-	23.715	15.125 ± 8.049	43.461 ± 12.414
$bbH: nb2+nb3$,boosted, $\mu\mu+ee+\mu e$	-	22.669	13.916 ± 7.545	39.441 ± 11.777
$bbH: nb2+nb3$,resolved, $\mu\mu+ee$	-	38.084	19.431 ± 11.715	47.895 ± 18.745
$bbH: nb2+nb3$,resolved, $\mu\mu+ee+\mu e$	-	28.738	13.975 ± 8.77	33.52 ± 14.032
$bbH: nb2+nb3$,resolved+boosted, $\mu\mu+ee$	-	19.182	10.169 ± 6.069	24.998 ± 9.591
$bbH: nb2+nb3$,resolved+boosted, $\mu\mu+ee+\mu e$	-	17.019	8.48 ± 5.301	20.376 ± 8.377
$bbA: nb2+nb3$,boosted, $\mu\mu+ee$	-	110.306	67.711 ± 37.492	177.158 ± 58.169
$bbA: nb2+nb3$,boosted, $\mu\mu+ee+\mu e$	-	102.785	60.636 ± 34.467	157.339 ± 53.801
$bbA: nb2+nb3$,resolved, $\mu\mu+ee$	-	44.847	24.133 ± 14.365	60.461 ± 22.423
$bbA: nb2+nb3$,resolved, $\mu\mu+ee+\mu e$	-	30.989	16.552 ± 9.805	41.539 ± 15.494
$bbA: nb2+nb3$,resolved+boosted, $\mu\mu+ee$	-	41.365	22.094 ± 13.25	55.448 ± 20.682
$bbA: nb2+nb3$,resolved+boosted, $\mu\mu+ee+\mu e$	-	29.387	15.579 ± 9.298	39.338 ± 14.694

categories, as detailed in Sec. 5.6.1 a strategy was implemented to distribute the generated signal mass hypotheses between resolved and boosted categories, and the $nb2$ category was also eliminated for gluon-gluon fusion. Therefore, for a comprehensive presentation of the results, including the observed outcomes, please consult the forthcoming publication [153].

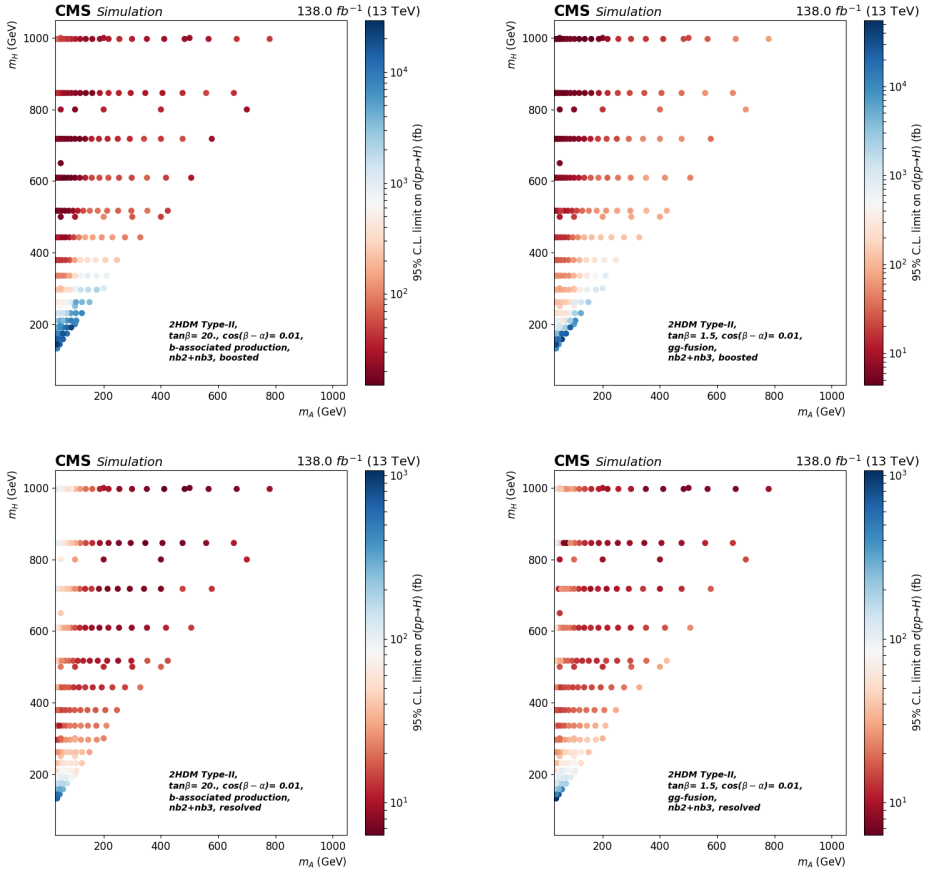


Figure 5.40. | Expected 95% CL upper limits on $\sigma \times BR(H \to ZA \to \ell^+\ell^- \to b\bar{b})$ in fb for bbH signal (left) and ggH (right). The limits are shown for full run 2 data taking, computed using the asymptotic CL_s method, combining ee, $\mu\mu$, and μe lepton flavors in boosted (top) and resolved (bottom) signal region category.

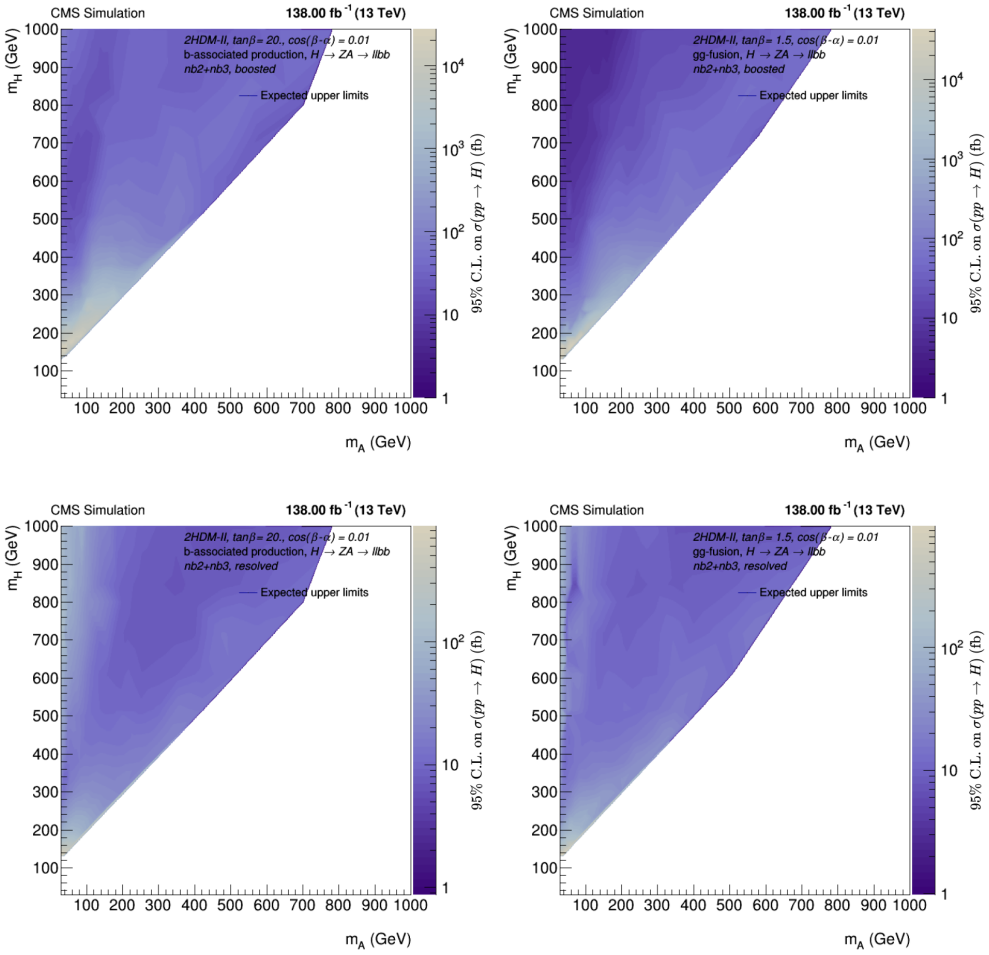


Figure 5.41. | Expected 95% CL upper limits on σ in fb scaled to the branching fractions $BR(H \to ZA \to \ell^+\ell^- \to b\bar{b})$ for the type-II 2HDM benchmark $\tan\beta = 1.5$ (left), $\tan\beta = 20.$ (right), and $\cos(\beta-\alpha) = 0.01$. The limits are computed using the asymptotic CL_s method in the $nb2+nb3$, boosted (top) and resolved (bottom) signal region categories combining both the e^+e^-, $\mu^+\mu^-$ and μ^+e^- channels.

5.8.3. 2HDM interpretation

Results for one particular signal mass point (m_H, m_A) can be scanned in $cos(\beta-\alpha)$ vs. $\tan\beta$ plane, as illustrated in Fig. 5.42 for the benchmark $m_H = 500$ GeV, $m_A = 300$ GeV. The expected excluded region is delimited by the red dashed area: values of $|cos(\beta-\alpha)| \leq 0.1$ and of $\tan\beta$ from approximately 1.1 to 3.0 are expected to be excluded at 95% CL for ggH signal process of type-II 2HDM. On the other hand, bbH allows the exclusion of high values of $\tan\beta > 10$. due to the enhancement of bottom Yukawa coupling as was demonstrated in Chap. 2.

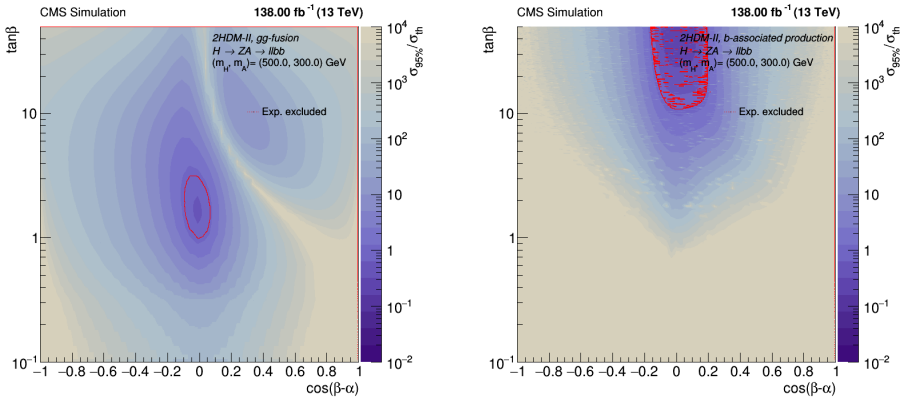


Figure 5.42. | Expected exclusion contours at 95% C.L. on $\frac{\sigma(pp\to H)}{\sigma_{theory}}$ for 2HDM type II benchmark (m_H, m_A)= (500, 300) GeV as a function of $\tan\beta$ and $\cos(\beta-\alpha)$ (right). Limits are calculated using the asymptotic CL_s for gluon-gluon fusion (left) and b-associated production (right) combining the two categories $nb=2$ and $nb=3$ in the signal regions resolved and boosted as well as e^+e^- and $\mu^+\mu^-$ lepton flavors.

The upper limit results will be presented using two distinct approaches. The first approach involves treating the two processes (i.e. ggH and bbH) independently. The second approach, on the other hand, treats both processes as a single parameter of interest, denoted as "r". To achieve this, a new model is constructed, taking into account the predicted cross sections of both processes, for every point examined within the 2HDM parameter space. Upper limits are then determined for $\sigma \times BR(H/A \to ZA/H) \to BR(A/H \to b\bar{b})$, with σ encompassing contributions from both processes. However, it's important to note that the first approach will necessitate appropriate scaling for a given $\tan\beta$ when conducting the statistical test. Achieving exclusion contours for each $\tan\beta$ value will demand substantial computational power and extensive CPU usage. For an overview of these results, please consult the forthcoming publication in [153].

5.8.4. Comparison with previous searches

This search has analyzed the complete CMS run 2 datasets, exploring both the resolved and boosted regions, and stands out as the first prior to previous ATLAS and CMS searches in the same final state (i.e. $\ell^+\ell^- b\bar{b}$), to probe as well as two production mechanisms: gluon-gluon fusion and b-associated production representing a significant leap forward. This has led to the need to create distinct categories based on the b-jets kinematics and multiplicity.

In the past, the ATLAS search employed (ExpGaussExp (EGE) and Crystal Ball (DSCB)) to examine and match the invariant masses of m_{llbb} for the purpose of identifying the target signal region [155]. Conversely, in an earlier approach by CMS during the initial run, they relied on defining a specific area, commonly termed a "box", within the two-dimensional mass plane of (m_{bb},m_{lbb}) to isolate the desired signal region [129]. However, it was subsequently found that using ellipses as an extraction technique yielded superior results. This shift in methodology resulted in a reduction of background interference and an enhancement of the signal's clarity [34]. For more detailed information on these different approaches, you can refer to the respective papers associated with each of these studies. Although previous methods effectively fulfilled their roles in their respective timeframes, they proved insufficient in capturing the full range of intricate signal features across the entire mass spectrum. The adoption of a parametric Deep Neural Network in this study has emerged as the ultimate and highly efficient solution in this search context.

The major benefit of the DNN is illustrated in Fig. 5.43 in a comparison with the ellipses strategy used in the previous 2016 analysis [34].

In the previous 2016 analysis, the observed limits for the production cross section times branching ratio ($\sigma \times \mathrm{BR}$) of ggH ($H \to ZA \to \ell^+\ell^- b\bar{b}$) with a mass configuration of (m_H, m_A) = (200, 50) GeV were determined to be 157 fb. The limit has seen a notable enhancement, by a factor of 1.57. It's worth highlighting that the degree of improvement varies based on the assumed mass hypothesis. This improvement is not solely attributable to the increased luminosity but also comes from the advancements in the analysis strategy and techniques.

As an example, consider a different signal mass hypothesis shown in Tab. 5.16, which highlights that the enhancements can result in a significant improvement, reaching up to a factor of 4.94 times better for ggH compared to the previous 2016 analysis [34]. In the earlier analysis, the observed limit for the same mass point was 74.8 fb. The table also serves as a demonstration to underscore the significance of combining categories to enhance the limits.

A full comparison of the mass plan with the 2016 analysis in Fig. 5.44 reveals enhancements spanning the entire mass spectrum. These improvements become evident when considering the incorporation of the previously highlighted yellow-

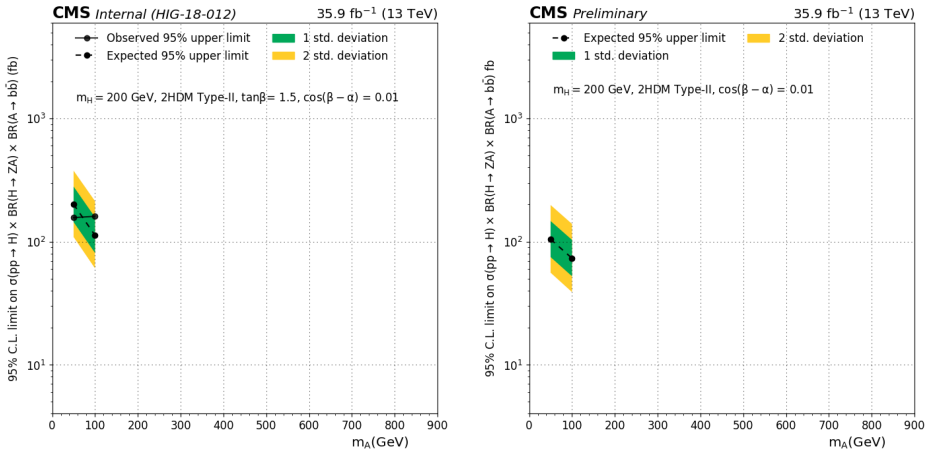


Figure 5.43. The expected 95% C.L upper limits on σ ggH production cross section times branching ratio ($H \to ZA \to \ell^+\ell^- b\bar{b}$) for $m_H = 200$ GeV as a function of m_A for the gluon-gluon fusion process in the $nb = 2$ -resolved category. The left plot shows both the expected and observed limits hypothesis using the ellipses strategy used in the 2016 search [34], while the right plot shows the expected limits using the DNN score as a template for signal extraction. The limits are computed using the asymptotic CL_s method, combining both the e^+e^- and $\mu^+\mu^-$ channels.

Table 5.16. A comparaison of the ggH and bbH $\sigma \times BR(H \to ZA) \times BR(A \to b\bar{b})$ using CL_s method for (m_H, m_A)= (500, 400) GeV.

Cat.	Observed (fb)	Expected (fb)	1 Standard deviation (fb)	2 Standard deviations (fb)
bbH : $nb2$, boosted, $ee + \mu\mu$	-	44.922	39.259 ± 29.063	92.293 ± 37.201
bbH : $nb2$, resolved, $ee + \mu\mu$	-	21.973	17.95 ± 14.216	42.622 ± 18.196
bbH : $nb3$, boosted, $ee + \mu\mu$	-	51.758	45.725 ± 34.459	107.083 ± 43.671
bbH : $nb3$, resolved, $ee + \mu\mu$	-	15.137	13.516 ± 9.978	32.078 ± 12.772
bbH : $nb2 + nb3$, boosted, $ee + \mu\mu$	-	36.621	32.352 ± 23.693	76.652 ± 30.327
bbH : $nb2 + nb3$, resolved, $ee + \mu\mu$	-	12.207	11.596 ± 7.725	26.011 ± 10.3
bbH : $nb2 + nb3$, resolved+boosted, $ee + \mu\mu$	-	12.207	10.668 ± 8.047	25.823 ± 10.3
ggH : $nb2$, boosted, $ee + \mu\mu$	-	76.172	64.399 ± 49.281	150.307 ± 63.08
ggH : $nb2$, resolved, $ee + \mu\mu$	-	16.113	13.776 ± 10.228	30.847 ± 13.092
ggH : $nb3$, boosted, $ee + \mu\mu$	-	81.055	70.067 ± 54.419	162.372 ± 69.656
ggH : $nb3$, resolved, $ee + \mu\mu$	-	83.984	69.009 ± 53.249	158.808 ± 69.55
ggH : $nb2 + nb3$, boosted, $ee + \mu\mu$	-	58.105	49.953 ± 37.536	116.324 ± 49.026
ggH : $nb2 + nb3$, resolved, $ee + \mu\mu$	-	15.137	12.941 ± 9.608	28.978 ± 12.299
ggH : $nb2 + nb3$, resolved + boosted, $ee + \mu\mu$	-	15.137	12.366 ± 9.608	28.789 ± 12.299

boosted regions, which have undergone significant enhancement in the current analysis.

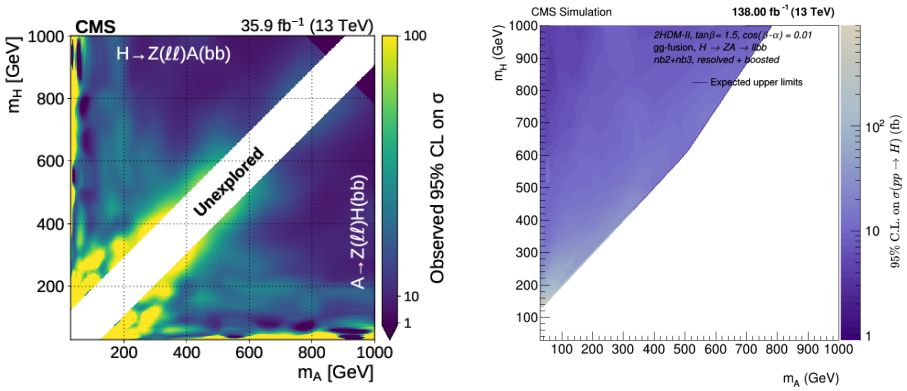


Figure 5.44. Upper bounds at 95% CL on the production cross-section times the branching ratio $B(H \to ZA) \times B(A \to b\bar{b})$ in fb. These limits are illustrated for gluon-gluon fusion, previous CMS 2016 search [34] (left) and our own results (right). An updated version of our results can be found in [153].

This analysis has undertaken a novel binning strategy using the Bayesian Blocks (described in Sec. 5.6.4) that is not predetermined but instead relies entirely on the characteristics of the data itself. This has introduced complexity to the analysis because of the parametric nature of the DNN output which implies that the binning will vary based on the assumed signal mass hypothesis. Furthermore, the optimization process using Bayesian Blocks is not static; it varies according to the number of b-jets, regions, and lepton flavors, adding an additional dimension to the intricacy of the approach. Nevertheless, this complexity has yielded advantageous results by demonstrating sensitivity across the entire (m_A, m_H) mass spectrum.

The ATLAS collaboration looked for the pseudoscalar particle A using 2016 data corresponding to an integrated luminosity of 36.1 fb^{-1}. Their primary focus was on the cascade decay process $A \to Z(\to \ell^+\ell^-)H(\to b\bar{b})$, concentrating on A masses below 800 GeV. In contrast, this investigation extends its scope to the inverse scenario, exploring mass ranges up to 1 TeV. Furthermore, their search looked for H with values exceeding 130 GeV, while our search focused on the lighter Higgs bosons, with masses that extend into a much lower range, as low as 30 GeV.

To account for contributions from both gluon-gluon fusion (when $nb = 2$) and b-associated production (when $nb \leq 3$), two distinct categories are established based on the number of b-jets, in the same fashion as what was undertaken in this study. However, it's important to note that this analysis differs in that it specifically examines the resolved and boosted regions, as opposed to the approach in the previous analysis, which looked solely at the resolved region. A

comparison of the results of the previous search to the current study is illustrated in Fig. 5.45. To facilitate the comparison with previous ATLAS results, consider the region highlighted in green/yellow of Fig. 5.45, approximately around (m_H, m_A) = (300, 50) GeV. In the ATLAS analysis, the expected limits on $\sigma \times B(A \to ZH) \times B(H \to b\bar{b})$ were estimated to be around 0.3 pb. In contrast, our current limits on the cross-section are approximately 50 fb, which is equivalent to 0.05 pb (the branching fractions $B(H \to ZA) \times B(A \to b\bar{b}) \times B(Z \to ee + \mu\mu)$ was already accounted for) [1].

This suggests that our independent limits on $\sigma \times B(H \to ZA) \times B(A \to b\bar{b})$ are approximately 0.045 pb and 0.032 pb for gluon-gluon fusion and b-associated production, respectively. Consequently, this analysis demonstrates an approximate 6.6 factor improvement for gluon-gluon fusion and a 9.3 factor improvement for b-associated production compared to the 2016 ATLAS results.

[1] For (m_H, m_A) = (300, 50) GeV in the context of a type-II 2HDM with an assumed $\tan\beta$ value of 1.5 for gluon-gluon fusion production, the calculated branching fraction is approximately 0.0613. This result is based on the individual contributions from $B(H \to ZA)$ (0.996), $B(A \to b\bar{b})$ (0.9158), and $B(Z \to ee + \mu\mu)$ (0.0672).

In the case of b-associated production, assuming $\tan\beta = 20$, the branching fraction is approximately 0.0431. This calculation takes into account the contributions from $B(H \to ZA)$ (0.6912), $B(A \to b\bar{b})$ (0.9269), and $B(Z \to ee + \mu\mu)$ (0.0672).

These branching fractions have been computed using the 2HDMC-1.8.0 framework [120].

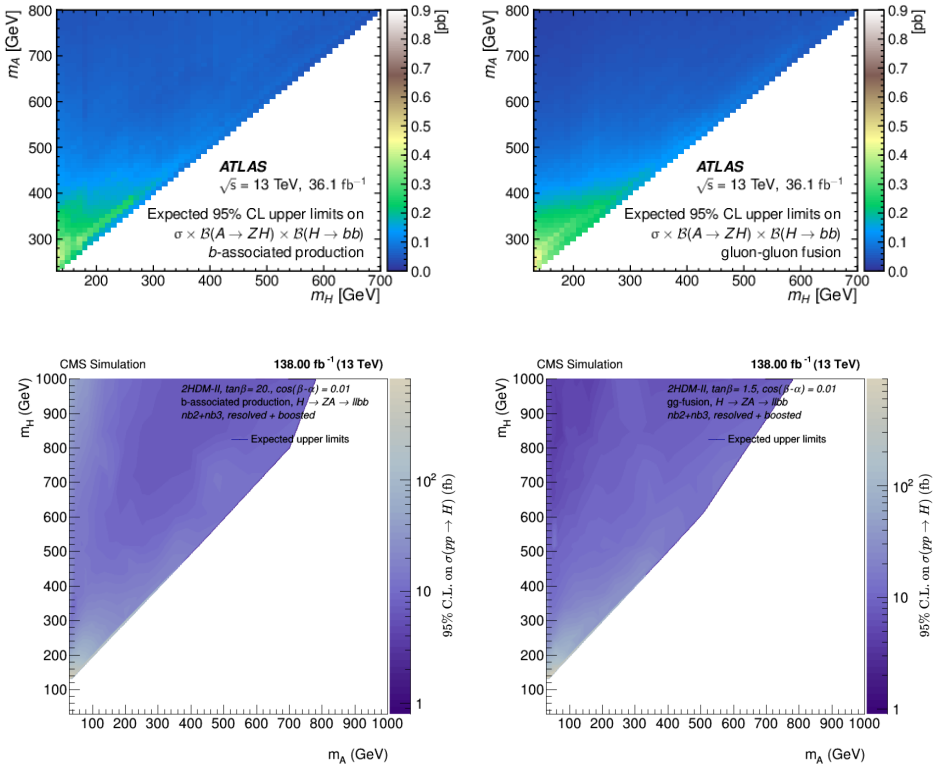


Figure 5.45. | Upper bounds at 95% CL on the production cross-section times the branching ratio $B(A \to ZH) \times B(H \to b\bar{b})$ in pb (ATLAS) and on the cross-section in fb (CMS). These limits are determined separately for b-associated production (left) and gluon-gluon fusion (right). The top plots display the expected upper limits for ATLAS results [155], while our own results are presented at the bottom. An updated version of our results can be found in [153].

Conclusions

The 2HDM remains a promising and well-studied theoretical framework for physics beyond the Standard Model (SM). It provides a rich phenomenology with additional Higgs bosons and can answer major open questions, such as the nature of dark matter, matter-antimatter asymmetry, and the hierarchy problem.

While the discovery of the Higgs boson at the LHC confirmed the existence of a fundamental scalar particle, there are still open questions regarding the complete nature of the Higgs sector. The search for additional Higgs bosons predicted in the Two Higgs Doublet Model (2HDM), along with exploring their properties and interactions, remains an important avenue for ongoing and future research.

The work presented here shows the complete analysis strategy of one of many possible decay channels for 2HDM neutral Higgs bosons. The reported search analyzes 138 fb^{-1} of proton-proton collisions data collected by the CMS experiment at a center-of-mass energy of $\sqrt{s} = 13$ TeV during LHC full run2 in 2016–2018, searching for a heavy scalar Higgs (H) and pseudoscalar (A) incompatible with the standard model prediction.

The search is performed in two decay modes, $H \to ZA$, $A \to ZH$ and to enhance the sensitivity of the analysis, boosted and resolved candidate events are separated into mutually exclusive categories of different expected signal-to-background ratios. Further splitting based on the properties and the multiplicity of the reconstructed b-tagged jets ($nb = 2$ and $nb = 3$) aimed at selecting events in which a Higgs boson or the pseudo-scalar is produced through b-associated (bbH) or gluon-gluon fusion (ggH) production mechanisms. In this search, we use machine learning techniques for the selection and classification of the events which improves this search by a factor of three in comparison to the previous 2016 search [34].

The analysis is currently in the process of unblinding, and while it's not yet fully disclosed, the preliminary findings indicate promising expected outcomes. For the most up-to-date results, please consult the latest version available in the reference [153]. It provides stronger constraints on the 2HDM parameters $\tan\beta$ and $\cos(\beta - \alpha)$ this is due to the optimization of the analysis strategy using a Deep Neural Network to separate signal from background processes, and different binning strategy using the Bayesian Blocks approach. Additionally,

the inclusion of b-associated production and boosted category improves the sensitivity of the analysis in a region of the 2HDM parameter where the resolved category and gluon-gluon fusion production were no longer relevant.

Upper limits on the extended (pseudo-)scalar (A) H sector cross-section times branching fraction are obtained in the mass range (30-1000 GeV) 125-1000 GeV at 95% confidence level. The results are interpreted in type-II 2HDM for the b-associated production and gluon-gluon fusion production mechanisms. The largest sensitivity considering can be attributed to the 2HDM benchmarks at $(m_H, m_A) = (379, 54.59)$ GeV for the gluon-gluon fusion process and $(m_H, m_A) = (230, 30)$ GeV for the b-associated production mechanism.

This analysis has significant enhancements compared to previous investigations carried out by ATLAS and CMS. These improvements are primarily attributed to the increased luminosity and the use of novel analysis strategies and techniques. Specifically, the DNN-based signal extraction and the Bayesian blocks for the binning strategy have been instrumental in advancing the quality of this analysis.

6. Chapter
Outlook and future perspectives

In this Chapter, I briefly summarize some possibilities for improvements that can be considered as prospects for future analysis.

6.1. 2HDM benchmarks

The 2HDM parameter space can be vast and complex. Therefore, when studying new physics models it is very important to establish benchmark scenarios. Choosing benchmark scenarios helps simplify the analysis and focus on specific regions of interest. They serve as guidelines about the latest experimental measurements and theoretical constraints. Additionally, they help identify the most promising regions of parameter space to explore and provide target values for specific observables, such as H and A masses in the context of the 2HDM.

There are some interesting benchmarks for 2HDM scenarios that were discussed in Sec. 2.4 of Chap. 2 that I think are worth investigating in future analyses. They can provide valuable insights into which 2HDM regions are currently favored or excluded by existing data, allowing us to refine the theory model when generating Monte-Carlo signal events. Nevertheless, it should be noted that although benchmark scenarios are interesting approaches, they are not meant to be complete representations of the entire parameter space. It is encouraged to select representative, well-motivated points that cover a range of phenomenologically interesting scenarios, allowing to optimize the capabilities to probe new physics.

6.2. Improving DNN robustness

I suspect that there is still a possible improvement in the performance of the DNN used in this analysis. To begin with, input variables can be added such as b-tagged jet multiplicity, and the angular distance (ΔR) between the two b-tagged jets, which could possibly further improve the sensitivity of the boosted category.

Also, the DNN performance can be enhanced using the Gradient Boosting Machine (GBM) technique which combines the predictions of multiple individual models, typically decision trees, to create a more accurate and powerful predictive model. It is a type of ensemble learning method where the models are trained sequentially, with each subsequent model focusing on correcting the mistakes made by the previous models. A visualization of this technique is illustrated in

Fig. 6.1.

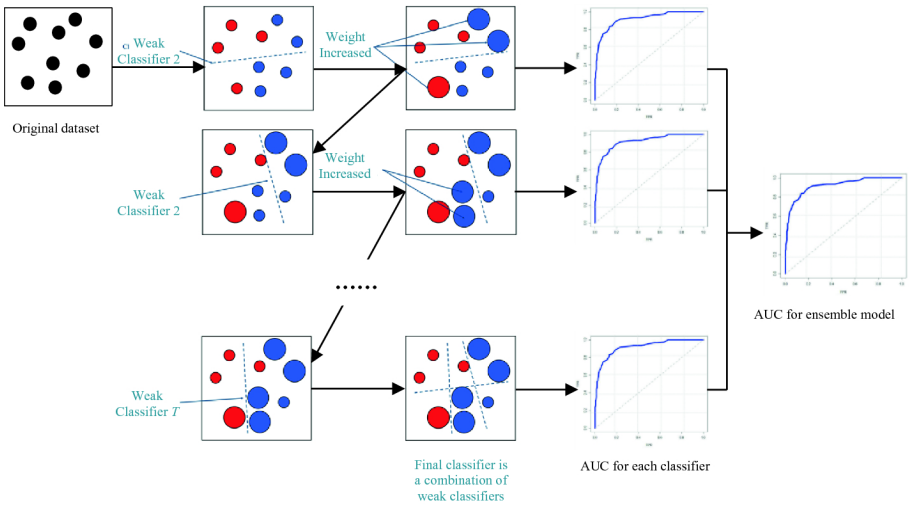


Figure 6.1. | Visualisation of the gradient boosting technique [156].

Popular implementations of gradient boosting include XGBoost, LightGBM, and CatBoost, which provide optimized algorithms for building accurate predictive models by combining weak learners into a strong ensemble, effectively leveraging their complementary strengths.

6.3. Analysis strategy

I would argue that the DNN background nodes (i.e. $t\bar{t}$ and DY) of the multi-classification presented in Sec. 5.6.3 of Chap. 5 would have been a valuable asset for the statistic analysis if were used instead of μe $t\bar{t}$ control region. These additional measurements would have provided information that can help constrain and improve the accuracy of the statistical analysis.

In this analysis, we have faced a few challenges with anti-correlated nuisance parameters that cause issues with the stability of the fit which required the procedure mentioned in Sec. 5.7 of Chap. 5. In my opinion, these additional nodes would have helped to constrain these nuisance parameters by providing additional information about their values or correlations, reducing the uncertainties associated with them.

In addition, given that DY and $t\bar{t}$ are the main sources of background in this analysis, including such nodes would have enhanced the statistical power and discrimination of the likelihood fit. By combining multiple measurements, the overall sensitivity to different parameters can be improved, leading to more precise and robust parameter estimates.

6.4. The 2HDM at HL-LHC and future colliders

The High-Luminosity LHC (HL-LHC), with its increased luminosity and improved experimental capabilities, will significantly enhance the sensitivity to new physics, including the 2HDM. We will be able to probe a wider range of parameter space, allowing for a more comprehensive search for additional Higgs bosons and other new phenomena.

The LHC detector upgrade for the HL-LHC project is currently in progress and physics data-taking is expected to start at the earliest in 2029. The HL-LHC is expected to deliver proton-proton collisions at 14 TeV with an integrated luminosity of 3 ab^{-1} for both ATLAS and CMS experiments, 50 fb^{-1} for LHCb, and 5 fb^{-1} for ALICE. For the heavy-ion sector, the integrated luminosity of 13 nb^{-1} and 50 nb^{-1} should be delivered for lead-lead and proton-lead collisions respectively.

Overall, the expected luminosity peak of $5.10^{34} cm^{-2} s^{-1}$ of the HL-LHC will lead to an integrated luminosity of 250 fb^{-1} for each operating year for a total integrated luminosity target of 3000 fb^{-1}, which is estimated to be a factor of 10 far from the LHC's design value.

A timeline of the LHC-related upgrades is shown in Fig. 6.2.

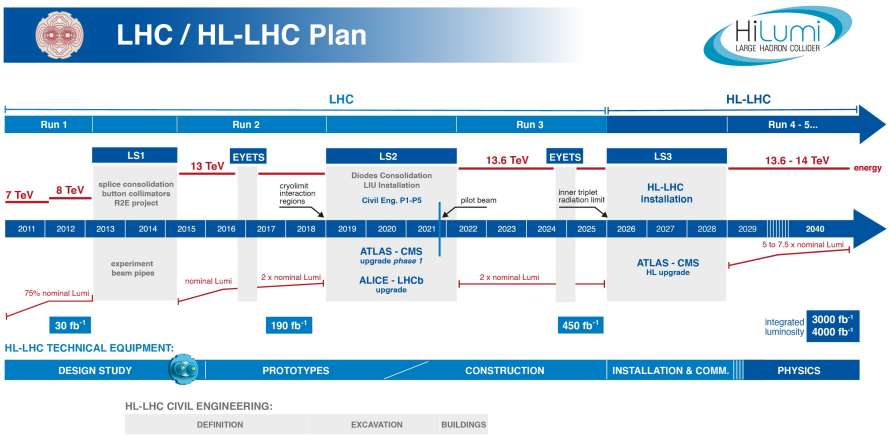


Figure 6.2. | Overseen timeline of LHC and HL-LHC operations [157].

The HL-LHC is expected to produce at least 15 million Higgs bosons per year, compared to around three million from the LHC in 2017. This will allow studies of the Higgs boson in great depth, the search for the H and A bosons, as well as other new physics signatures, rare decays might reveal themselves, and much more...

Furthermore, proposed future colliders, such as the International Linear Collider (ILC), Circular Electron Positron Collider (CEPC), and the Future Circular Collider (FCC), also have the potential to provide additional sensitivity and extend the exploration of the 2HDM parameter space. Recent paper [158] has shown global fit results of type II 2HDM which include the latest measurements and theoretical constraints (unitarity, perturbativity, and vacuum stability) as well as future precision measurements from future colliders, these hypothetical results are shown in Fig. 6.3. The left panels display the results of a global fit analysis using both current data and theoretical constraints. There are minimum thresholds for the masses of the 2HDM heavy scalars $m_{A,H,H^{\pm}}$, approximately around 400 GeV. These limits are primarily derived from precise measurements at the Z-pole and considerations of flavor physics. Additionally, there exists a lower boundary for the parameter $\tan\beta$, predominantly coming from flavor physics.

Moving to the right panels, they provide a comparison between the current 2σ confidence regions depicted in black and those resulting from the integration of forthcoming precision measurements. This integration encompasses data from the HL-LHC (in orange), HL-LHC combined with CEPC (in red), HL-LHC paired with ILC (in blue), and HL-LHC coupled with FCC-ee (in green). The inclusion of precision measurements from the HL-LHC serves to elevate the lower limit on the scalar mass to 500 GeV. Subsequently, these limits escalate even further, reaching up to 700 GeV when additional constraints from future lepton colliders are considered.

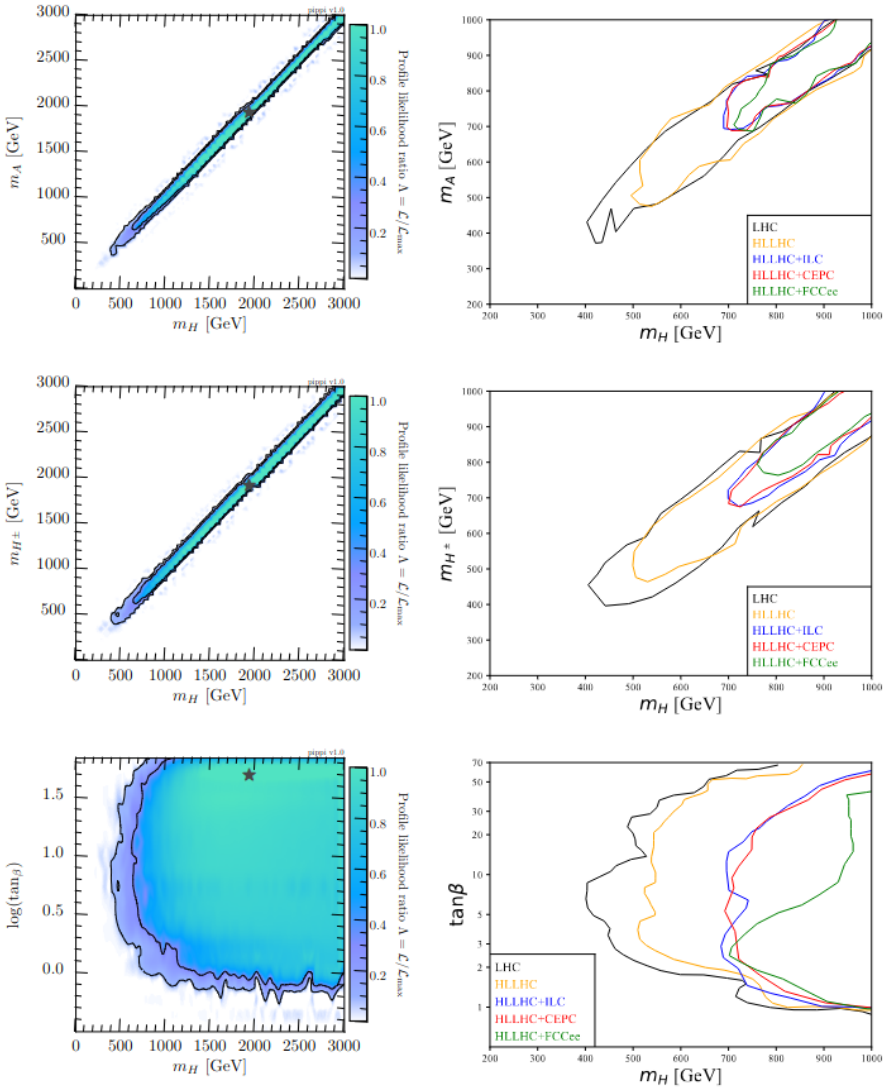


Figure 6.3. | Left panel: 1σ and 2σ global fit results in the $m_A - m_H$ (top), $m_H - m_H^\pm$ (middle), and $m_H - \tan\beta$ (bottom), based on current measurements and theoretical constraints. Right panel: Comparison of current 2σ constraints (black) with future precision measurements from future colliders [158].

Appendix A.

Additional Material: B-tagging efficiencies measurements

This appendix reports the derived b-tagging efficiencies for DeepJet and DeepCSV algorithms, for the medium working point. The calculated b-tagging efficiencies are applied as event-by-event scale factors to correct the differences between data and simulation in the b-tagging algorithms. The scale factor weight w for a single event is calculated for the medium working point to correct the response of DeepJet and DeepCSV taggers respectively in the resolved, boosted categories for all jet flavors as;

$$w = \frac{P(\text{DATA})}{P(\text{MC})}, \tag{A.1.}$$

where the probability of a given configuration of jets in MC simulation and data is defined as:

$$P(\text{MC}) = \prod_{i=\text{tagged M}} \epsilon_i^M \prod_{j=\text{not tagged M}} \left(1-\epsilon_j^M\right), \tag{A.2.}$$

$$P(\text{DATA}) = \prod_{i=\text{tagged M}} \text{SF}_i^M \epsilon_i^M \prod_{j=\text{not tagged M}} \left(1-\text{SF}_j^M \epsilon_j^M\right), \tag{A.3.}$$

where SF^M are the scale factors provided by the BTV POG and ϵ^M denote the efficiencies in simulation for the medium working point. The values of these efficiencies for different flavors, p_T and η values of jets are documented in Figures 1.1 - 1.8 for all the years.

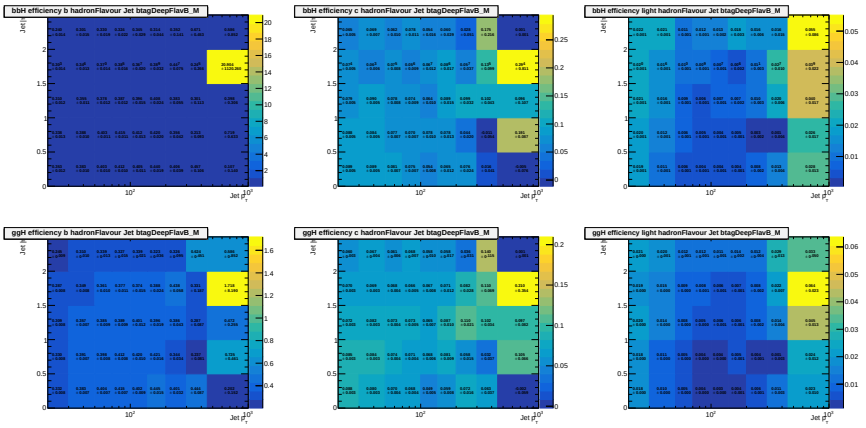


Figure 1.1. | 2016 Ulegacy postVFP b-tagging efficiencies for bbH resolved (top) and ggH resolved (bottom) derived for DeepJet tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

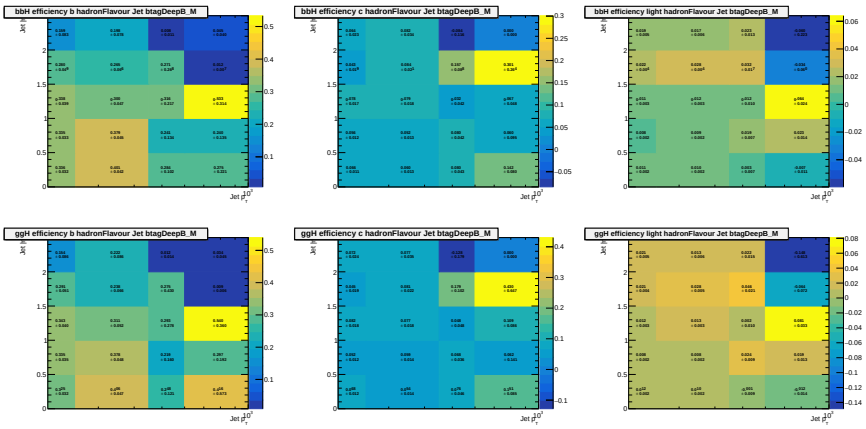


Figure 1.2. | 2016 Ulegacy postVFP b-tagging efficiencies for bbH boosted (top) and ggH boosted (bottom) derived for DeepCSV tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

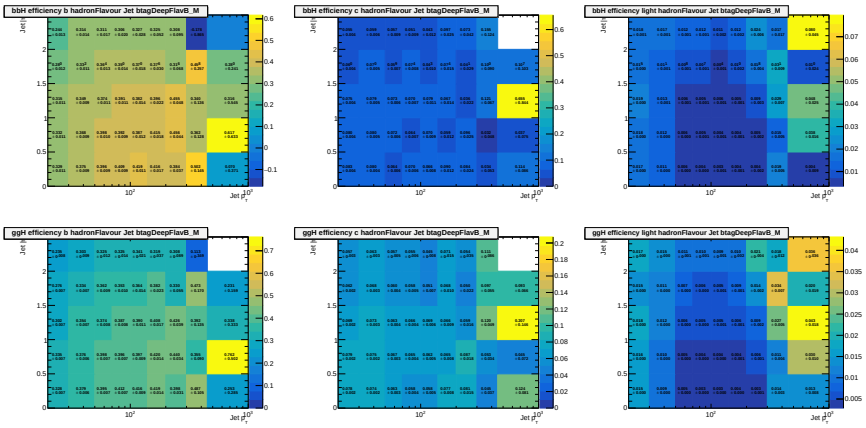


Figure 1.3. | 2016 Ulegacy preVFP b-tagging efficiencies for bbH resolved(top) and ggH resolved (bottom) derived for DeepJet tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

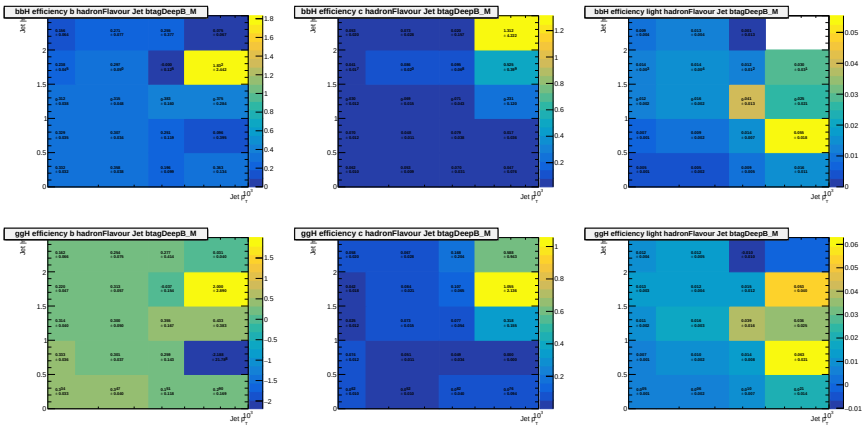


Figure 1.4. | 2016 Ulegacy preVFP b-tagging efficiencies for bbH boosted (top) and ggH boosted (bottom) derived for DeepCSV tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

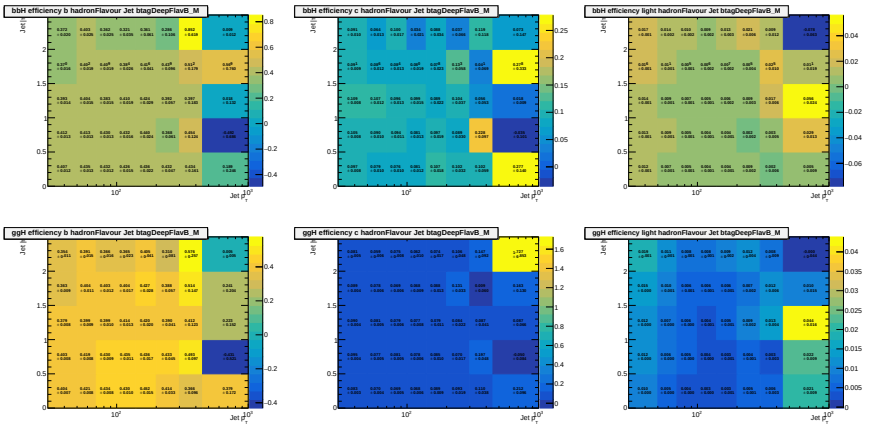


Figure 1.5. | 2017 Ulegacy b-tagging efficiencies for bbH resolved(top) and ggH resolved (bottom) derived for DeepJet tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

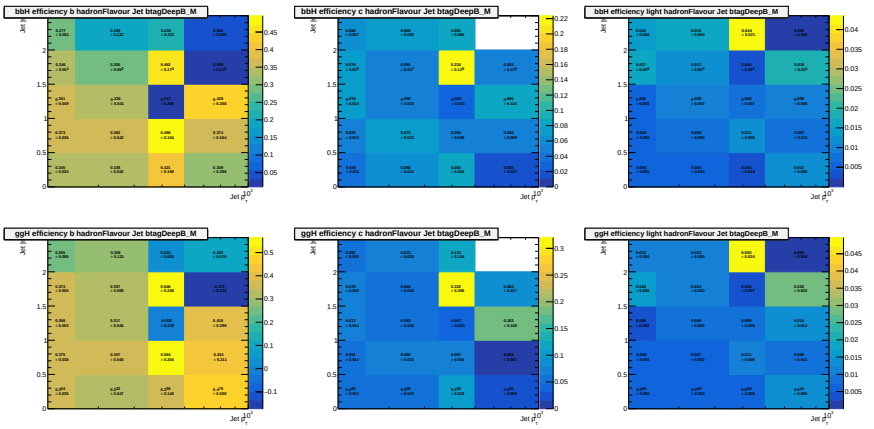


Figure 1.6. | 2017 Ulegacy b-tagging efficiencies for bbH boosted (top) and ggH boosted (bottom) derived for DeepCSV tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

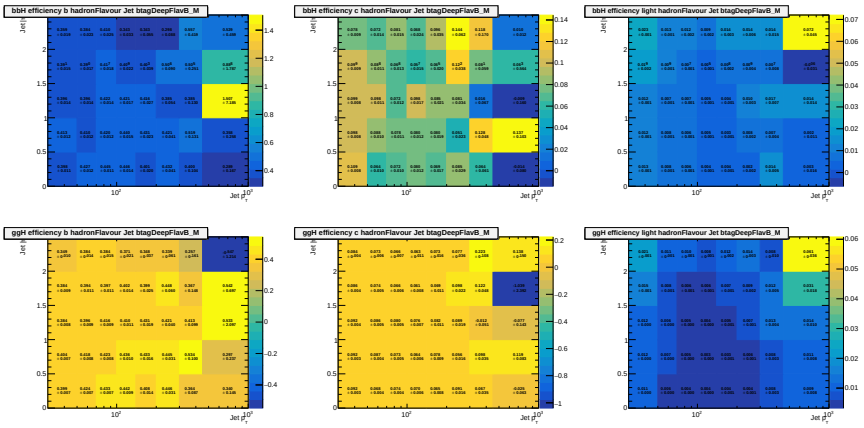


Figure 1.7. | 2018 Ulegacy b-tagging efficiencies for bbH resolved(top) and ggH resolved (bottom) derived for DeepJet tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

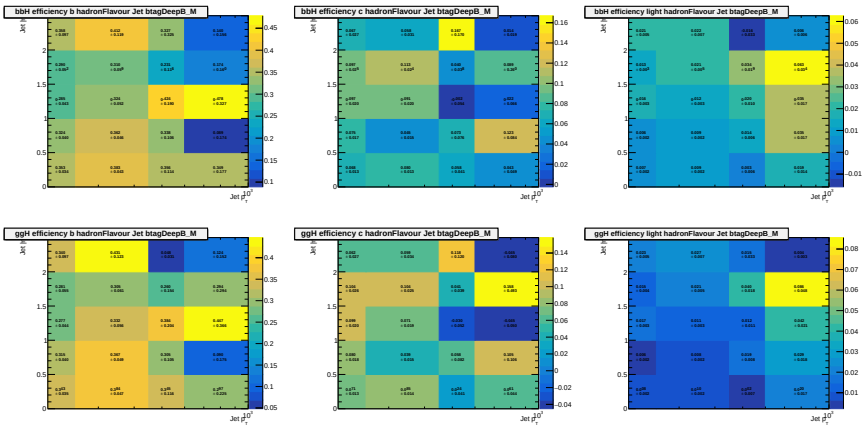


Figure 1.8. | 2018 Ulegacy b-tagging efficiencies for bbH boosted (top) and ggH boosted (bottom) derived for DeepCSV tagger, medium b-tagging working point of b-jets (left), c-jets (middle) and light-jets (right).

B. Appendix

Kinematical differences between leading orderand next-to-leading order b-associated production

The matrix element (ME) is composed of a finite number of fixed-order terms, expanded in the strong-force coupling α_s:

$$d\sigma_{hard} = d\sigma^{LO}_{hard} + \alpha_s(Q) d\sigma^{NLO}_{hard} + \alpha^2_s(Q) d\sigma^{NNLO}_{hard} + ... \qquad \text{(B.1.)}$$

where Q is the renormalization scale of the hard-process ME.

The mode-squaring of the matrix element in Eq. B.1 mixes amplitudes of different orders in α_s so that e.g. the $d\sigma_{NLO}$ cross-section correction includes both the square of a one-emission ("real") amplitude and the cross-term composed of interference between a one-loop ("virtual") amplitude and the Born amplitude. This few-body process is then systematically improved toward a realistic event as would be observed by an ideal collider experiment. A comparison of the Monte Carlo kinematic distributions in the LO and NLO calculation for the b-associated production in the 4F-scheme in the decay channel $b\bar{b}H \to ZA \to \ell\ell b\bar{b}$ as displayed in Fig. 2.1.

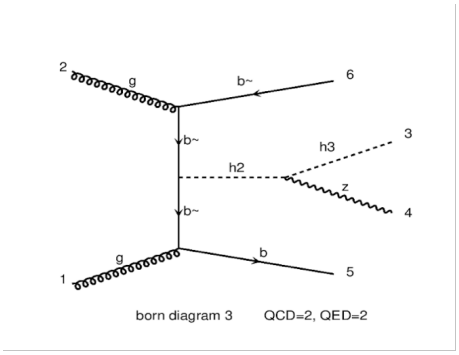


Figure 2.1. | The leading Feynman diagram of the decay in the 2HDM of a heavy neutral Higgs boson H(= h2) to a pseudoscalar A(= h3) with the association of two b quarks.

Although events in the 4FS always include two forward bottom jets at the

generator level, they must still be accurately detected and reconstructed. To be classified as originating from bbH production, a reconstructed event must meet specific selection criteria. It is anticipated that the two forward b jets arising from the hard scattering process should be reconstructed with a broad pseudo-rapidity window and exhibit characteristics similar to jets formed with the PF anti-k_T algorithm within a cone radius of $\Delta R = 0.4$. This investigation is illustrated in Fig. 2.2- 2.4.

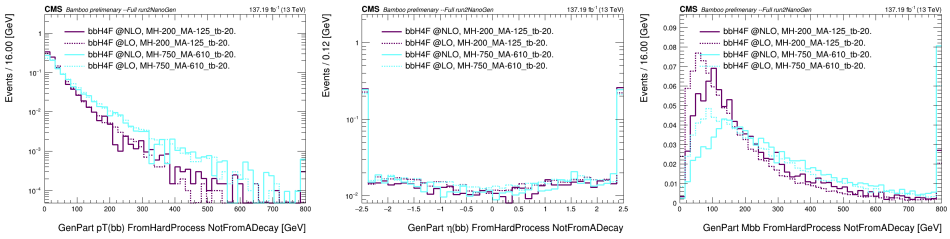


Figure 2.2. | Transverse momentum (left), pseudo-rapidity(middle), and the invariant mass (right) of the two hardest B hadrons at both LO+PS (dashed line) and NLO+PS (solid line).

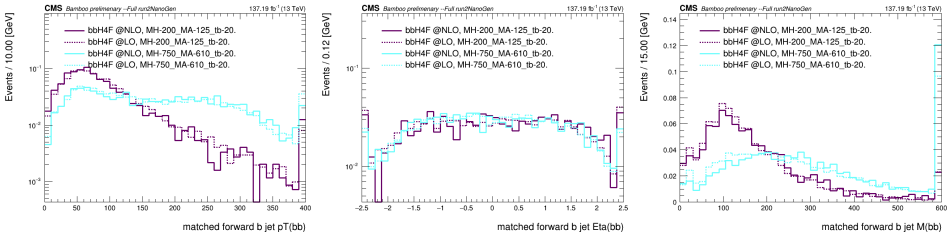


Figure 2.3. | Transverse momentum (left), pseudo-rapidity(middle), and the invariant mass (right) of the two forward-matched b-jets at both LO+PS (dashed line) and NLO+PS (solid line).

Fig. 2.5 demonstrates the difference between the production modes at different stages of selection for signal masses separate in the mass plane.

The consistent treatment of loop amplitudes in NLO MC generation additionally introduces problematic rates of negative weights. Negative weights essentially count double toward generation inefficiencies because not only do they not add statistical convergence but they actively subtract it, cf. the variance formula $\sigma^2 = <\sum w^2> - <\sum w>^2$ in which the first term always increases but the second will be reduced by a mix of positive and negative weights. Generated bbH signal samples, after parton showering show roughly 40% negative events fractions as demonstrated in Table. 2.1.

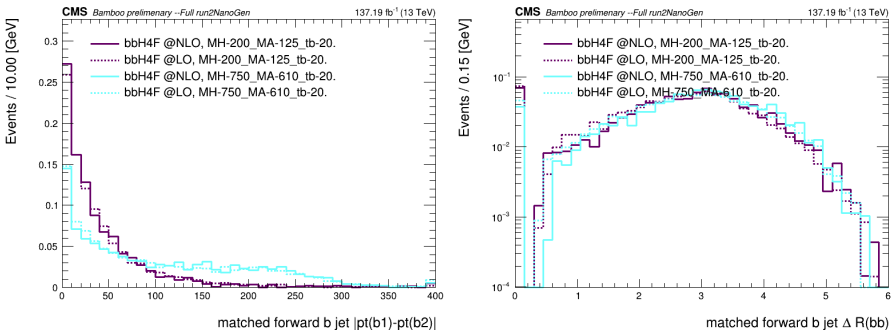


Figure 2.4. | $\Delta p_{\mathrm{T}}(b,b)$ (left), ΔR (right) of the two forward-matched b-jets at both LO+PS (dashed line) and NLO+PS (solid line).

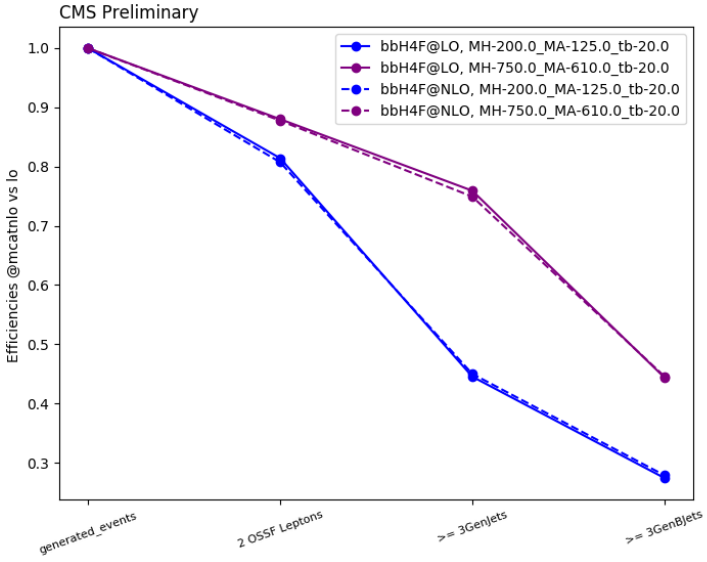


Figure 2.5. | The Efficiencies of the LO and NLO bbH production at different stages of selections for two benchmarks of (m_H, m_A); (200, 125)GeV and (750, 610)GeV.

For the upcoming BSM searches that aim to probe the $b\bar{b}H$ decay channel at the next leading order, it is worth pointing to POWHEG BoxV2 [159] and the new `MC@NLO` $-\Delta$ [160] matching procedure which reduces the number of negative-weight events with respect to the standard `MC@NLO`. Nevertheless, the LO in practice is still the optimal solution for experimental purposes as it is both more approximate and much cheaper computationally and as is shown in the studied distributions the impact of NLO corrections is rather small. Apart

Table 2.1. | Fraction of events with negative weights on bbH signal samples.

(m_H, m_A, tanβ)	Negative weights fractions
(200, 125, 20)	$4.005e^{-1} \pm 4.388e^{-3}$
(250, 50 , 20)	$3.890e^{-1} \pm 4.240e^{-3}$
(500, 50 , 20)	$4.090e^{-1} \pm 4.496e^{-3}$
(700, 200, 20)	$3.950e^{-1} \pm 4.318e^{-3}$

from an overall global correction to the cross sections that could be compensated using the K-factor, the corresponding observables of the $b\bar{b}H$ process can be used in the analysis without taking into account the NLO corrections.

Acronyms

ALICE A Large Ion Collider Experiment 58, 197

ATLAS A Toroidal LHC Apparatus 58, 61, 197

CMS Compact Muon Solenoid 58, 62–65, 68, 71, 78, 82, 83, 87, 88, 94, 97, 197

CSCs Cathode Strip Chambers 80, 82

DTs Drift Tubes 80, 82

ECAL electromagnetic calorimeter 64, 75, 76, 82, 88, 89, 92, 94

GSF Gaussian Sum Filter 89, 90

HCAL hadron calorimeter 64, 75, 76, 78, 82, 92, 94

HERWIG HERWIG 55

HLT High-Level Trigger 81, 82

L1 Level 1 Trigger 81, 82

LEP Large Electron–Positron Collider 23, 24, 57, 58

LHC Large Hadron Collider 11, 57, 58, 61, 62, 69, 82

LHCb Large Hadron Collider beauty 58, 197

LINAC4 Linear accelerator 4 58

PF Particle-Flow 84, 88–90, 92–96, 103

PS Proton Synchrotron 58, 59

PSB Proton Synchrotron Booster 58

RPCs Resistive Plate Chambers 82

SHERPA SHERPA 55

SM Standard Model 11, 12

SPS Super Proton Synchrotron 58, 59

References

[1] F. Englert et al., "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323. doi:10.1103/PhysRevLett.13.321.

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

[3] S. L. Glashow, "Towards a unified theory: Threads in a tapestry", *Rev. Mod. Phys.* **52** (Jul, 1980) 539–543. doi:10.1103/RevModPhys.52.539.

[4] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1–29. doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[5] CMS Collaboration, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61. doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[6] E. Ward, "Key Particle Discoveries Timeline", (2019). https://cds.cern.ch/record/2665175. General Photo.

[7] CMS, "What is the world made of?". https://www.fnal.gov/pub/inquiring/matter/madeof/index.html.

[8] J. Riebesell, "Mexican Hat". https://tikz.net/mexican-hat/.

[9] M. Gonzalez-Lopez et al., "Testing anomalous $H - W$ couplings and Higgs self-couplings via double and triple Higgs production at e^+e^- colliders", *Eur. Phys. J. C* **81** (2021), no. 3, 260. doi:10.1140/epjc/s10052-021-09048-1, arXiv:2011.13915.

[10] X. Fan et al., "Measurement of the Electron Magnetic Moment", *Phys. Rev. Lett.* **130** (Feb, 2023) 071801. doi:10.1103/PhysRevLett.130.071801.

[11] The CMS collaboration, "The CMS experiment at the CERN LHC", *Journal of Instrumentation* **3** (aug, 2008) S08004. doi:10.1088/1748-0221/3/08/S08004.

[12] "Summaries of CMS cross section measurements". https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined.

[13] LHC Higgs Cross Section Working Group et al., "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", *CERN-2013-004* (CERN, Geneva, 2013). doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[14] The CMS collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Physics Letters B* **716** (2012), no. 1, 30–61. doi:https://doi.org/10.1016/j.physletb.2012.08.021.

[15] G. Aad et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* **716** (2012), no. 1, 1–29. doi:https://doi.org/10.1016/j.physletb.2012.08.020.

[16] CDF Collaboration, "High-precision measurement of the W boson mass with the CDF II detector", *Science* **376** (2022), no. 6589, 170–176. doi:10.1126/science.abk1781.

[17] The ATLAS collaboration, "New ATLAS result weighs in on the W boson". https://atlas.cern/Updates/Briefing/2023-W-Mass-Measurement.

[18] Muon $g-2$ Collaboration Collaboration, "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm", *Phys. Rev. Lett.* **126** (Apr, 2021) 141801. doi:10.1103/PhysRevLett.126.141801.

[19] L. Canetti et al., "Matter and antimatter in the universe", *New Journal of Physics* **14** (sep, 2012) 095012. doi:10.1088/1367-2630/14/9/095012.

[20] J. Smirnov et al., "Tev-scale thermal WIMPs: Unitarity and its consequences", *Physical Review D* **100** (aug, 2019). doi:10.1103/physrevd.100.043029.

[21] T. D. Lee, "A Theory of Spontaneous T Violation", *Phys. Rev. D* **8** (Aug, 1973) 1226–1239. doi:10.1103/PhysRevD.8.1226.

[22] S. D. Bass et al., "The Higgs boson implications and prospects for future discoveries", *Nature Rev. Phys.* **3** (2021), no. 9, 608–624. doi:10.1038/s42254-021-00341-2, arXiv:2104.06821.

[23] A. Barroso et al., "Metastability bounds on the two Higgs doublet model", *JHEP* **06** (2013) 045. doi:10.1007/JHEP06(2013)045, arXiv:1303.5098.

[24] F. J. Botella et al., "Flavor conservation in two-Higgs-doublet models", *Phys. Rev. D* **98** (Aug, 2018) 035046. doi:10.1103/PhysRevD.98.035046.

[25] H. E. Haber et al., "Basis-independent methods for the two-Higgs-doublet model III: The CP-conserving limit, custodial symmetry, and the oblique parameters S, T, U", *Physical Review D* **83** (mar, 2011). doi:10.1103/physrevd.83.055017.

[26] M. E. Peskin et al., "Estimation of oblique electroweak corrections", *Phys. Rev. D* **46** (Jul, 1992) 381–409. doi:10.1103/PhysRevD.46.381.

[27] I. Maksymyk et al., "Beyond S, T, and U", *Physical Review D* **50** (jul, 1994) 529–535. doi:10.1103/physrevd.50.529.

[28] J. Haller et al., "Update of the global electroweak fit and constraints on two-Higgs-doublet models", *Eur. Phys. J. C* **78** (2018), no. 8, 675. doi:10.1140/epjc/s10052-018-6131-3, arXiv:1803.01853. 33 pages, 13 figures, submitted to EPJC.

[29] Y. Omura et al., "$\tau-$ and $\mu-$physics in a general two Higgs doublet model with $\mu-\tau$ flavor violation", *Physical Review D* **94** (sep, 2016). doi:10.1103/physrevd.94.055019.

[30] S. Iguro et al., "Testing the 2HDM explanation of the muon g – 2 anomaly at the LHC", *Journal of High Energy Physics* **2019** (nov, 2019). doi:10.1007/jhep11(2019)130.

[31] The ATLAS and the CMS collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}=7$ and 8 TeV", *Journal of High Energy Physics* **2016** (aug, 2016). doi:10.1007/jhep08(2016)045.

[32] A. Arbey et al., "Status of the Charged Higgs Boson in Two Higgs Doublet Models. Status of the Charged Higgs Boson in Two Higgs Doublet Models", *Eur. Phys. J. C* **78** (2018), no. 3, 182. doi:10.1140/epjc/s10052-018-5651-1, arXiv:1706.07414. 36 pages, 10 figures. v2: comments about oblique parameters and references added.

[33] ATLAS Collaboration, "Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the $\ell\ell b\bar{b}$ and $\ell\ell WW$ final states in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector", *Eur. Phys. J. C* **81** (2021), no. 5, 396. doi:10.1140/epjc/s10052-021-09117-5, arXiv:2011.05639.

[34] CMS Collaboration, "Search for new neutral Higgs bosons through the H $\to$ ZA $\to \ell^+\ell^-\ b\bar{b}$ process in proton-proton collisions at $\sqrt{s}=13$TeV", *J. High Energ. Phys.* **03** (2020) 055. doi:10.1007/JHEP03(2020)055.

[35] ALEPH, DELPHI, L3, OPAL, LEP Working Group for Higgs Boson Searches Collaboration, "Search for neutral MSSM Higgs bosons at LEP", *Eur. Phys. J. C* **47** (2006) 547–587. doi:10.1140/epjc/s2006-02569-7, arXiv:hep-ex/0602042.

[36] The ATLAS collaboration, "Search for scalar resonances decaying into $\mu^+\mu^-$ in events with and without b-tagged jets produced in proton-proton collisions at $\sqrt{s}=13$ TeV with the ATLAS detector", *Journal of High Energy Physics* **2019** (jul, 2019). doi:10.1007/jhep07(2019)117.

[37] CMS Collaboration, "Search for MSSM Higgs bosons decaying to $\mu\mu$ in proton-proton collisions at $\sqrt{s}=$ 13 TeV", *Physics Letters B* **798** (nov, 2019) 134992. doi:10.1016/j.physletb.2019.134992.

[38] ATLAS Collaboration, "Search for heavy neutral Higgs bosons produced in association with b-quarks and decaying into b-quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. D* **102** (2020), no. 3, 032004. doi:10.1103/PhysRevD.102.032004, arXiv:1907.02749.

[39] The CMS collaboration, "Search for beyond the standard model Higgs bosons decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics* **2018** (aug, 2018). doi:10.1007/jhep08(2018)113.

[40] ATLAS Collaboration, "Search for heavy Higgs bosons decaying into two tau leptons with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **125** (2020), no. 5, 051801. doi:10.1103/PhysRevLett.125.051801, arXiv:2002.12223.

[41] The CMS collaboration, "Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics* **2018** (sep, 2018). doi:10.1007/jhep09(2018)007.

[42] The CMS collaboration, "Search for a low-mass τ-τ+ resonance in association with a bottom quark in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *Journal of High Energy Physics* **2019** (may, 2019). doi:10.1007/jhep05(2019)210.

[43] ATLAS Collaboration, "Search for Scalar Diphoton Resonances in the Mass Range $65 - 600$ GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV", *Phys. Rev. Lett.* **113** (2014), no. 17, 171801. doi:10.1103/PhysRevLett.113.171801, arXiv:1407.6583.

[44] ATLAS Collaboration, "Search for new phenomena in high-mass diphoton final states using 37 fb^{-1} of proton–proton collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett. B* **775** (2017) 105–125. doi:10.1016/j.physletb.2017.10.039, arXiv:1707.04147.

[45] ATLAS Collaboration, "Search for resonances in the 65 to 110 GeV diphoton invariant mass range using 80 fb^{-1} of pp collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector", technical report, 2018. https://cds.cern.ch/record/2628760. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-025.

[46] The CMS collaboration, "Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s}$ 8 and 13 TeV", *Physics Letters B* **793** (jun, 2019) 320–347. doi:10.1016/j.physletb.2019.03.064.

[47] The CMS collaboration, "Search for physics beyond the standard model in high-mass diphoton events from proton-proton collisions at 13 TeV", *Physical Review D* **98** (nov, 2018). doi:10.1103/physrevd.98.092001.

[48] The CMS collaboration, "Search for heavy Higgs bosons decaying to a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics* **2020** (apr, 2020). doi:10.1007/jhep04(2020)171.

[49] ATLAS Collaboration, "Search for heavy ZZ resonances in the $\ell^+\ell^-\ell^+\ell^-$ and $\ell^+\ell^-\nu\overline{\nu}$ final states using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Eur. Phys. J. C* **78** (2018), no. 4, 293. doi:10.1140/epjc/s10052-018-5686-3, arXiv:1712.06386.

[50] The CMS collaboration, "Search for a new scalar resonance decaying to a pair of Z bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics* **2018** (jun, 2018). doi:10.1007/jhep06(2018)127.

[51] ATLAS Collaboration, "Search for heavy resonances decaying into WW in the $e\nu\mu\nu$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Eur. Phys. J. C* **78** (2018), no. 1, 24. doi:10.1140/epjc/s10052-017-5491-4, arXiv:1710.01123.

[52] The CMS collaboration, "Search for a heavy Higgs boson decaying to a pair of W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics* **2020** (mar, 2020). doi:10.1007/jhep03(2020)034.

[53] A. Collaboration, "Search for the Higgs boson produced in association with a vector boson and decaying into two spin-zero particles in the H $\rightarrow$ aa $\rightarrow$ 4b channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Journal of High Energy Physics* **2018** (oct, 2018). doi:10.1007/jhep10(2018)031.

[54] The CMS collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *Journal of Instrumentation* **13** (may, 2018) P05011–P05011. doi:10.1088/1748-0221/13/05/p05011.

[55] C. Collaboration, "Search for Higgs boson decays into a pair of light bosons in the $b\overline{b}$ μμ final state in pp collision at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Physics Letters B* **790** (mar, 2019) 1–21. doi:10.1016/j.physletb.2018.10.073.

[56] C. Collaboration, "Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two muons and two b quarks in pp collisions at 13 TeV", *Physics Letters B* **795** (aug, 2019) 398–423. doi:10.1016/j.physletb.2019.06.021.

[57] C. Collaboration, "Search for light pseudoscalar boson pairs produced from decays of the 125 GeV Higgs boson in final states with two muons and two nearby tracks in pp collisions at 13 TeV", *Physics Letters B* **800** (jan, 2020) 135087. doi:10.1016/j.physletb.2019.135087.

[58] A. Collaboration, "Search for Higgs bosons decaying to aa in the $\mu\mu\tau\tau$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS experiment", 2015. arXiv:1505.01609.

[59] CMS Collaboration, "Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state of two muons and two τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics* **2018** (nov, 2018). doi:10.1007/jhep11(2018)018.

[60] ATLAS Collaboration, "Search for Higgs boson decays to beyond-the-Standard-Model light bosons in four-lepton events with the ATLAS detector at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics* **2018** (jun, 2018). doi:10.1007/jhep06(2018)166.

[61] CMS Collaboration, "A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 TeV", *Physics Letters B* **796** (sep, 2019) 131–154. doi:10.1016/j.physletb.2019.07.013.

[62] ATLAS Collaboration, "Search for a CP-odd Higgs boson decaying to Zh in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Lett. B* **744** (2015) 163–183. doi:10.1016/j.physletb.2015.03.054, arXiv:1502.04478.

[63] ATLAS Collaboration, "Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector", *JHEP* **03** (2018) 174. doi:10.1007/JHEP03(2018)174, arXiv:1712.06518. [Erratum: JHEP 11, 051 (2018)].

[64] CMS Collaboration, "Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in $l^+l^-b\bar{b}$ final states", *Phys. Lett. B* **748** (2015) 221–243. doi:10.1016/j.physletb.2015.07.010, arXiv:1504.04710.

[65] CMS Collaboration, "Search for a heavy pseudoscalar boson decaying to a Z and a Higgs boson at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **79** (2019), no. 7, 564. doi:10.1140/epjc/s10052-019-7058-z, arXiv:1903.00941.

[66] CMS Collaboration, "Searches for a heavy scalar boson H decaying to a pair of 125 GeV Higgs bosons hh or for a heavy pseudoscalar boson A decaying to Zh, in the final states with $h \to \tau\tau$", *Phys. Lett. B* **755** (2016) 217–244. doi:10.1016/j.physletb.2016.01.056, arXiv:1510.01181.

[67] CMS Collaboration, "Search for a heavy pseudoscalar Higgs boson decaying into a 125 GeV Higgs boson and a Z boson in final states with two tau and two light leptons at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2020) 065. doi:10.1007/JHEP03(2020)065, arXiv:1910.11634.

[68] ATLAS Collaboration, "Searches for Higgs boson pair production in the $hh \to b\bar{b}\tau\tau, \gamma\gamma WW^*, \gamma\gamma b\bar{b}, b\bar{b}b\bar{b}$ channels with the ATLAS detector", *Phys. Rev. D* **92** (2015) 092004. doi:10.1103/PhysRevD.92.092004, arXiv:1509.04670.

[69] ATLAS Collaboration, "Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector", *Phys. Lett. B* **800** (2020) 135103. doi:10.1016/j.physletb.2019.135103, arXiv:1906.02025.

[70] CMS Collaboration, "Search for Higgs boson pair production in the $b\bar{b}\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **96** (2017), no. 7, 072004. doi:10.1103/PhysRevD.96.072004, arXiv:1707.00350.

[71] CMS Collaboration, "Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **122** (2019), no. 12, 121803. doi:10.1103/PhysRevLett.122.121803, arXiv:1811.09689.

[72] F. Kling et al., "2HDM Neutral Scalars under the LHC", *JHEP* **06** (2020) 163. doi:10.1007/JHEP06(2020)163, arXiv:2004.04172.

[73] S. Höche, "Introduction to parton-shower event generators", 2015. arXiv:1411.4085.

[74] R. K. Ellis et al., "QCD and collider physics",. doi:10.1017/CBO9780511628788. Photography by S. Vascotto.

[75] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1. doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[76] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30. doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[77] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **06** (2013) 081. doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[78] E. Mobs, "The CERN accelerator complex. Complexe des accélérateurs du CERN". https://cds.cern.ch/record/2197559.

[79] LHCStudyGroup CollaborationT. S. Pettersson, et al., "The Large Hadron Collider: conceptual design". http://cds.cern.ch/record/291782, 1995.

[80] CMS, "LHC MACHINE OUTREACH". http://lhc-machine-outreach.web.cern.ch/collisions.htm.

[81] Z. Hu et al., "Review of bottomonium measurements from CMS. Review of bottomonium measurements from CMS", *Int. J. Mod. Phys. A* **32** (2017), no. 19n20, 1730015. doi:10.1142/S0217751X17300150, arXiv:1708.02913. 32 pages.

[82] CMS Collaboration, "Highlights from the Compact Muon Solenoid (CMS) Experiment", *Universe* **5** (2019) 28. doi:10.3390/universe5010028, arXiv:1901.05340.

[83] CMS Collaboration, "Interactive Slice of the CMS detector",.

[84] I. Neutelings, "CMS coordinate system". https://tikz.net/axis3d_cms/.

[85] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker",. doi:10.17181/CERN.QZ28.FLHW.

[86] CMS, "Tracker Material Budget plots". https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackerMaterialBudgetplots.

[87] M. Krammer et al., "Tracking Detectors", pp. 317–352. Springer International Publishing, Cham, 2021. doi:10.1007/978-3-319-93785-4_12.

[88] CMS Collaboration, "The CMS Phase-1 Pixel Detector Upgrade", technical report, 2020. https://cds.cern.ch/record/2745805.

[89] N. Bartosik, "Associated Top-Quark-Pair and b-Jet Production in the Dilepton Channel at $\sqrt{s} = 8$ TeV as Test of QCD and Background to tt+Higgs Production", PhD thesis, Hamburg : DESY, *CERN Document Server (CDS)* 312, 2015. https://cds.cern.ch/record/2047049. Presented 03 Jul 2015.

[90] F. Hartmann, "The Design of the CMS Upgrade Tracker and the CMS High Granularity Forward Calorimeter Equipped with Silicon Sensors for the HL-LHC", pp. 291–329. Springer International Publishing, Cham, 2017. doi:10.1007/978-3-319-64436-3_7.

[91] J. Draeger, "Track based alignment of the CMS silicon tracker and its implication on physics performance", PhD thesis, 10.3204/DESY-THESIS-2011-026, Hamburg U., CERN-THESIS-2011-391, *CERN Document Server (CDS)* 391, 2011.

[92] P. Lenzi et al., "Track reconstruction of real cosmic muon events with CMS tracker detector", *Journal of Physics: Conference Series* **119** (jul, 2008) 032030. doi:10.1088/1742-6596/119/3/032030.

[93] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018), no. 06, P06015. doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/MUO-16-001 (CMS Public Pages).

[94] CMS Collaboration, "CMS Technical Design Report for the Level-1 Trigger Upgrade", technical report, 2013. https://cds.cern.ch/record/1556311. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch and Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[95] CMS, "Early 2018 High-Level Trigger rates",.

[96] CMS Collaboration, "The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report", technical report, 2017. https://cds.cern.ch/record/2283193. This is the CMS Interim TDR devoted to the upgrade of the CMS DAQ in view of the HL-LHC running, as approved by the LHCC.

[97] V. Garonne et al., "Rucio - The next generation of large scale distributed system for ATLAS Data Management", technical report, https://cds.cern.ch/record/1627590.

[98] "CERN Newsletter of the HEP department: Data Quality Monitoring- An important step towards new physics". https://ep-news.web.cern.ch/content/data-quality-monitoring-important-step-towards-new-physics.

[99] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017), no. 10, P10003. doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[100] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021), no. 05, P05014. doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

[101] The CMS collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018), no. 05, P05011. doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/BTV-16-002 (CMS Public Pages).

[102] E. Bols et al., "Jet flavour classification using DeepJet", *Journal of Instrumentation* **15** (dec, 2020) P12012. doi:10.1088/1748-0221/15/12/P12012.

[103] F. Chollet et al., "Keras". https://github.com/fchollet/keras, 2015.

[104] CMS, "Performance of the CMS DeepJet b tagging algorithm using $41.9 fb^{-1}$ of data from proton-proton collisions at 13TeV with Phase 1 CMS detector". https://twiki.cern.ch/twiki/bin/view/CMSPublic/BTV13TeV2017DeepJet.

[105] K. Cranmer, "Practical Statistics for the LHC", 2015. arXiv:1503.07622.

[106] ROOT Collaboration, "HistFactory: A tool for creating statistical models for use with RooFit and RooStats", technical report, 2012. https://cds.cern.ch/record/1456844.

[107] A. L. Read, "Modified frequentist analysis of search results (the CL_s method)",. doi:10.5170/CERN-2000-005.81.

[108] W. Verkerke et al., "The RooFit toolkit for data modeling", 2003. arXiv:physics/0306116.

[109] L. Moneta et al., "The RooStats project", in *Proceedings of 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research — PoS(ACAT2010)*, volume 093, p. 057.

[110] Hartmann, K., Krois, J., Rudolph, A. (2023):, "Statistics and Geodata Analysis using R (SOGA-R). Department of Earth Sciences, Freie Universitaet Berlin.". https://www.geo.fu-berlin.de/en/v/soga-r/Basics-of-statistics/Hypothesis-Tests/Introduction-to-Hypothesis-Testing/Critical-Value-and-the-p-Value-Approach/index.html.

[111] CMS Collaboration, "Searches for additional Higgs bosons and for vector leptoquarks in $\tau\tau$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV", 2022. arXiv:2208.02717.

[112] X. Chen, "Sample Size, Statistical Power, and Power Analysis", pp. 301–328. Springer International Publishing, Cham, 2021. doi:10.1007/978-3-030-83852-2_10.

[113] F. Bre et al., "Prediction of wind pressure coefficients on building surfaces using artificial neural networks", *Energy and Buildings* **158** (2018) 1429–1441.

[114] J. Feng et al., "Reconstruction of porous media from extremely limited information using conditional generative adversarial networks", *Physical Review E* **100** (09, 2019). doi:10.1103/PhysRevE.100.033308.

[115] MATHWorks, "Compare Deep Learning Models Using ROC Curves". https://se.mathworks.com/help/deeplearning/ug/compare-deep-learning-models-using-ROC-curves.html.

[116] Data Science and Machine Learning, "What is Confusion Matrix and Advanced Classification Metrics?". https://se.mathworks.com/help/deeplearning/ug/compare-deep-learning-models-using-ROC-curves.html.

[117] CMS Collaboration, "Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in $\ell^+\ell^- b\bar{b}$ final states", *Phys. Lett.* **B748** (2015) 221–243. doi:10.1016/j.physletb.2015.07.010, arXiv:1504.04710.

[118] ATLAS Collaboration, "Search for a CP-odd Higgs boson decaying to Zh in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Lett.* **B744** (2015) 163–183. doi:10.1016/j.physletb.2015.03.054, arXiv:1502.04478.

[119] CMS Collaboration, "Search for neutral resonances decaying into a Z boson and a pair of b jets or τ leptons", *Phys. Lett.* **B759** (2016) 369–394. doi:10.1016/j.physletb.2016.05.087, arXiv:1603.02991.

[120] D. Eriksson et al., "2HDMC- two-Higgs-doublet model calculator", *Computer Physics Communications* **181** (2010), no. 1, 189–205. doi:https://doi.org/10.1016/j.cpc.2009.09.011.

[121] S. Liebler et al., "Phenomenology of on-shell Higgs production in the MSSM with complex parameters", *The European Physical Journal C* **77** (May, 2017). doi:10.1140/epjc/s10052-017-4849-y.

[122] R. Brun et al., "ROOT — An object oriented data analysis framework", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **389** (1997), no. 1, 81–86. doi:https://doi.org/10.1016/S0168-9002(97)00048-X. New Computing Techniques in Physics Research V.

[123] P. David, "Readable and efficient HEP data analysis with bamboo", *EPJ Web Conf.* **251** (2021) 03052. doi:10.1051/epjconf/202125103052, arXiv:2103.01889.

[124] E. Guiraud et al., "RDataFrame enhancements for HEP analyses", *Journal of Physics: Conference Series* **2438** (feb, 2023) 012116. doi:10.1088/1742-6596/2438/1/012116.

[125] W. M. et al., "Higgs production in association with bottom quarks", *Journal of High Energy Physics* **132** (2015) 35. doi:10.1007/JHEP02(2015)132.

[126] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Computer Physics Communications* **191** (2015) 159–177. doi:https://doi.org/10.1016/j.cpc.2015.01.024.

[127] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040. doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[128] R. Harlander et al., "Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach",. arXiv:1112.3478.

[129] A. Mertens, "Search for 2HDM extensions of the scalar sector close to the alignment limit with the CMS detector", PhD thesis, UCL - SST/IRMP - Institut de recherche en mathématique et physique, *Digital Access to Library (DIAL)* 147, 2017. http://hdl.handle.net/2078.1/185406.

[130] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET",.

[131] Jiyeon Han and Arie Bodek , "Rochester corrections for muon momentum scale and resolution developed by Rochester group. ". https://www-cdf.fnal.gov/~jyhan/cms_momscl/cms_rochcor_manual.html, https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon.

[132] CMS Collaboration, "Jet algorithms performance in 13 TeV data", technical report, 2017. https://cds.cern.ch/record/2256875.

[133] "CMS Internal TWiki: Methods to apply b-tagging efficiency scale factors". https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods.

[134] CMS, "CMS Internal Twiki: L1 Prefiring. Reweighting recipe to emulate Level 1 ECAL and Muon prefiring". https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1PrefiringWeightRecipe.

[135] D. Kovalskyi et al., "Measurement of prefiring in the L1 muon triggerin Run 2". https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20AN-2021/086.

[136] CMS Collaboration, "Plans for Jet Energy Corrections at CMS",.

[137] "CMS Internal TWiki: MET Filter Recommendations for Run II ". https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2. Accessed: 2021-10-21.

[138] CMS, "The reweighting of pt(top-quark) in MC.". https://twiki.cern.ch/twiki/bin/view/Sandbox/JamesKeaveneySandbox.

[139] CMS Collaboration, "Evidence for the Higgs boson decay to a bottom quark–antiquark pair", *Physics Letters B* **780** (2018) 501–532. doi:https://doi.org/10.1016/j.physletb.2018.02.050.

[140] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys. Rev. D* **89** (Jan, 2014) 012003. doi:10.1103/PhysRevD.89.012003.

[141] CMS Collaboration, "A Deep Neural Network for Simultaneous Estimation of b Jet Energy and Resolution", *Comput. Softw. Big Sci.* **4** (2020), no. 1, 10. doi:10.1007/s41781-020-00041-z, arXiv:1912.06046.

[142] CMS Collaboration, "Measurement of the inclusive W and Z production cross sections in pp collisions at $\sqrt{s} = 7$TeV with the CMS experiment", *JHEP.* **10** (2011) 132. doi:10.1007/JHEP10(2011)132, arXiv:1107.4789.

[143] "CMS Internal TWiki: Measurement of the trigger efficiencies for a dilepton selection for a ttbar analysis with the full run 2 dataset". https://gitlab.cern.ch/tdr/notes/AN-19-140/-/raw/miniAODv2/TrigEffAN_temp.pdf.

[144] "CMS Internal TWiki: TOP Trigger (Run2) ". https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopTrigger#Dilepton_triggers.

[145] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *The European Physical Journal C* **71** (feb, 2011). doi:10.1140/epjc/s10052-011-1554-0.

[146] "CMS Internal TWiki: Summary table of samples produced for the 1 Billion campaign, with 25ns bunch-crossing". https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns.

[147] J. D. Scargle et al., "STUDIES IN ASTRONOMICAL TIME SERIES ANALYSIS. VI. BAYESIAN BLOCK REPRESENTATIONS", *The Astrophysical Journal* **764** (feb, 2013) 167. doi:10.1088/0004-637x/764/2/167.

[148] D. Sivia, "A review of: "Data Analysis: A Bayesian Tutorial" D. S. Sivia, 1996 Oxford, Clarendon Press ISBN 0-19-851889-7", *European Journal of Engineering Education* **22** (1997), no. 2, 224–224. doi:10.1080/03043799708928277, arXiv:https://doi.org/10.1080/03043799708928277.

[149] W. Cash, "Parameter estimation in astronomy through application of the likelihood ratio.", *Astrophysical Journal* **228** (March, 1979) 939–947. doi:10.1086/156922.

[150] "CMS Internal TWiki: Luminosity Physics Object Group (Lumi POG)". https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM#SummaryTable.

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

[152] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", 2011. arXiv:1103.0354.

[153] K. Jaffel, "CMS Internal CADILINES: Search for neutral heavy Higgs bosons in the context of a 2HDM via $H/A \to Z(l^+l^-) \to A/H(b\bar{b})$ with full Run 2 data set". https://cms.cern.ch/iCMS/analysisadmin/cadilines?line=HIG-22-010&tp=an&id=2595&ancode=HIG-22-010.

[154] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435. doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[155] The ATLAS collaboration, "Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the ll$b\bar{b}$ final state in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector", *Physics Letters B* **783** (2018) 392–414. doi:https://doi.org/10.1016/j.physletb.2018.07.006.

[156] D. S. Team, "Gradient Boosting". https://datascience.eu/machine-learning/gradient-boosting-what-you-need-to-know/, 2020.

[157] The HL-LHC Collaboration, "HL-LHC Project". https://project-hl-lhc-industry.web.cern.ch/content/project-schedule.

[158] A. Beniwal et al., "Global fit of 2HDM with future collider results", 2022. arXiv:2203.07883.

[159] S. Frixione et al., "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *Journal of High Energy Physics* **2007** (nov, 2007) 070–070. doi:10.1088/1126-6708/2007/11/070.

[160] R. Frederix et al., "On the reduction of negative weights in MC@NLO-type matching procedures", *JHEP* **07** (2020) 238. doi:10.1007/JHEP07(2020)238, arXiv:2002.12716.