Higgs boson anomalous couplings using the di-photon channel with CMS experiment

Facoltà di Scienze Matematiche Fisiche e Naturali
Fisica

Federica De Riggi
ID number 1875555

Advisor
Emanuele Di Marco

Academic Year 2021/2022

Higgs boson anomalous couplings using the di-photon channel with CMS experiment
Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: deriggi.1875555@studenti.uniroma1.it

Ringraziamenti

Prima di procedere con la trattazione, vorrei dedicare qualche riga per ringraziare il gruppo di CMS Roma che con infinita disponibilità e pazienza ha corretto le mie presentazioni e mi ha dato consigli che hanno reso questa esperienza di tesi così formativa. Un ringraziamento speciale ai dottorandi della baita per il loro sostegno pratico e morale in questo percorso. Un sentito grazie al mio relatore, il Dot. Emanuele Di Marco, per tutte le ore di colloquio, di cui conservo gelosamente ancora i fogli, ed i consigli che mi ha dato. Non sarei così fiera di queste pagine senza il suo contributo

Abstract

Studies of CP violation and anomalous couplings of the Higgs boson to vector bosons are presented. The data were acquired by the CMS experiment at the LHC and correspond to an integrated luminosity of 138 fb^{-1} at a proton-proton collision energy of 13 TeV. The kinematic effects in the Higgs boson two-photons decay and its production in association with two jets (Vector Boson Fusion) is analyzed. A matrix element technique and multivariate algorithms are employed to identify the production mechanisms and to increase sensitivity to the Higgs boson tensor structure of the interactions, optimized using the full simulation of the detector. The measurement of the strength of the coupling of the Higgs boson to vector boson is performed on data through a fit which considers the simultaneous presence of SM and BSM contributions to the coupling parameters.

Contents

Introduction

The Higgs boson (H) discovered in 2012 at the CERN LHC [1; 2] has so far been found to have properties consistent with the standard model (SM) expectations. In particular, its spin-parity quantum numbers are consistent with $J^{PC} = 0^{++}$ according to measurements performed by the CMS [3] and ATLAS experiments, but leaving room for the possibility of small anomalous HVV couplings, where V stands for vector bosons (V=$W^{\pm}, Z^0$). The Higgs boson couplings, once the Higgs boson mass is defined, are precisely predicted by the Standard Model. Finding a deviation on these values would lead to evidence of the presence of new physics, such as new particles entering at higher orders of the production or decay amplitude thus altering the predictions of the Standard Model.

The purpose of this thesis is to search for new physics effects beyond the standard model (BSM) through the measurement of the Higgs boson couplings with the vector-bosons $W^{\pm}$ or Z^0. Because non-zero spin assignments of the H boson have been excluded, we focus on the analysis of couplings of a spin-0 Higgs boson. Previous studies of anomalous HVV couplings were performed by both the CMS and ATLAS experiments using either decay-only information [4], including associated production information [5], or including off-shell H boson production [6].

In this thesis I report a study of HVV couplings using information from VBF production of the Higgs boson decaying into two photons, which has a low branching fraction, but a clean signature in the CMS detector. The data used were recorded by the CMS experiment during LHC Run 2 in 2016, 2017 and 2018 proton proton collision at a center-of-mass energy of 13 TeV and correspond to an integrated luminosity of $\sim 137 fb^{-1}$. The VBF cross-section fraction times $H \rightarrow \gamma\gamma$ branching fraction predicted by the SM (and confirmed by the LHC measurements) is $\sigma_{VBF} \times BR(H \rightarrow \gamma\gamma) \sim 3.78 pb \times 0.2\%$, so using full Run 2 data we expect ~ 300 events (considering the photon detection efficiency).

The $H \rightarrow \gamma\gamma$ measurements rely on a fit of the diphoton invariant mass spectrum in data. Due to the excellent diphoton mass resolution in CMS, the Higgs boson signal appears as a narrow peak in this distribution. By using a fit function composed of a smoothly falling background distribution plus a signal peak, the background for $H \rightarrow \gamma\gamma$ processes is estimated fully on data.

To be sensitive to anomalous couplings a multivariate analysis technique and matrix element are developed using reconstructed photons and jets variables, to gain the sensitivity on J^{CP} properties of the resonance.

This thesis is organized as follows: Chapter 1 describes The Standard Model Theory, together with the Higgs boson mechanism and the main aspects of the Higgs boson observation and proprieties characterization are presented. Chapter 2 describes the experimental setup, including the Large Hadron Collider and the CMS detector. In the third chapter, the model used in this thesis, based on an effective field theory, is explained. Chapter 3 describes the procedure to select the signal with respect to the background. Moreover it gives a brief introduction to machine learning and the multivariate techniques used in this work and explains the categorization method used to be sensitive to anomalous couplings. Chapter 5 presents the modeling of the signal and the background process, the systematic uncertainties used, and the implementation of the likelihood fit. In the chapter, the expected results for the parameter of interest are also shown. In Chapter 6 the data are unblinded in the signal region and a likelihood profile study is performed to estimate the best-fit values of the parameters of interest.

Chapter 1

The Standard Model

The Standard Model (SM) is a physics theory considered one of the most important results of the last 70 years. Since the 1930s multiple studies and experiments gave insight into the fundamental structure of matter providing in agreement with the thought that everything in the universe is made from a few basic building blocks called fundamental particles, governed by four fundamental forces. Our best understanding of how these particles and three of the forces are related to each other is encapsulated in the SM which unifies three of the four fundamental forces of nature: strong, weak and electromagnetic. It has successfully explained almost all experimental results and precisely predicted a wide variety of phenomena. In Tab. 1.1 the forces, their mediators, and their relative intensities are summarized.

Interaction	Mediator	Relative intensity
Strong	gluon	1
Electromagnetic	photon	10^{-2}
Weak	$W^{\pm}$ and Z^0	10^{-7}

Table 1.1. Fundamental interactions, their mediators and relative intensity at energy scale ~ 100 GeV.

The particles involved in the SM theory are the bosons, which have an integer spin, and fermions which have a semi-integer spin. The fundamental interactions between particles are described by a function that delineates their dynamics, called the Lagrangian ($\mathcal{L}$). This function is locally invariant under suitable symmetries described by the model. The SM has $SU(2) \otimes U(1) \otimes SU(3)$ symmetry and can be divided into two sectors: the chromodynamic (QCD) sector of the symmetry $SU(3)$ and the electroweak one $SU(2) \otimes U(1)$.

$$\mathcal{L} = \mathcal{L}_{QCD} + \mathcal{L}_{e.w.} \tag{1.1}$$

Typically any symmetry is associated with a quantum number which is a number that characterizes an eigenstate of the system. SU(2) and U(1) respectively are associated to the weak isospin (I) and weak hypercharge (Y) quantum numbers. We characterize the foundamental particle according to these quantum numbers.

The fermions, which are the constituents of the matter, are divided into two groups depending on whether they are subjected to the strong force or not: the leptons do not interact with gluons while the quarks interact through the exchange of gluons.

Nature is symmetric under the group of Lorentz transformations, rotations, and translations which all together form the *Poincaré group*. The quantum number of the particles are different for different chirality which is a technical property, defined through transformation behaviour under the Poincaré group, that does not change with the reference frame. It is contrived to agree with helicity for massless particles, and is still well defined for particles with mass. The left(right)-chiral fermions of a spin 1/2 field is also denoted ad left(right)-handed. The $W^{\pm}$ can couple only to left-handed particles, while Z^0 can couple to both left-handed and right-handed particles. Left handed particles are grouped to form isospin doublets with $I = 1/2$.

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \begin{pmatrix} t \\ b \end{pmatrix}_L, \tag{1.2}$$

while right-handed particles are isospin singlet (I=0):

$$e_R, \mu_R, \tau_R, u_R, d_R, s_R, c_R, b_R, t_R \tag{1.3}$$

In Tab. 1.2 fermions with their respective quantum numbers are described.

Left-chiral fermions				Right-chiral			
	Electric charge Q	weak Isospin I_3	weak hypercharge Y_W		Electric charge Q	weak Isospin I_3	weak hypercharge Y_W
leptons							
ν_e, ν_μ, ν_τ	0	+1/2	-1	$\nu_{eR}, \nu_{\mu_R}, \nu_{\tau_R}$	0	0	0
e^-, μ^-, τ^-	-1	-1/2	-1	e_R^-, μ_R^-, τ_R^-	-1	0	-2
quarks							
u, c, t	+2/3	+1/2	+1/3	u_R, c_R, t_R	+2/3	0	+4/3
d, s, b	-1/3	-1/2	+1/3	d_R, s_R, b_R	-1/3	0	-2/3

Table 1.2. Fermions listed with the respective quantum numbers, I_3 described the isospin component along a certain size. Neutrinos with right-chirality are also listed in the table above but given the quantum numbers of these particles even if they existed they would not interact, making them impossible to detect.

The model introduces as many mediating particles as the number of generators of the model symmetry group. Therefore we introduce 3 vector bosons for SU(2) symmetry ($W_\mu^{A=1,2,3}$) and one for U(1) (B_μ). The combination of the fields that diagonalize the mass matrix are the physical states of these particles and are given by:

$$A_\mu = sin\theta_W W_\mu^3 + cos\theta_W B_\mu \tag{1.4}$$

$$Z_\mu = cos\theta_W W_\mu^3 + sin\theta_W B_\mu \tag{1.5}$$

$$W_\mu^\pm = \frac{W_\mu^1 \mp i W_\mu^2}{\sqrt{2}} \tag{1.6}$$

Where A_μ is the massless field, identifiable as the photon so as the mediator of the electromagnetic interaction. The other fields Z_μ and W_μ are relative to the particles Z and W detected for the first time at LEP (electron-positron collider) in 1983; θ_W is the Weinberg angle.

The symmetry component of the standard model $SU(3)$ introduces the color quantum number conserved in strong interactions. The particles mediating the strong interaction are 8, like the generators of SU(3), and are recognized as gluons.

Any theory, capable to make predictions, should be re-normalizable, i.e. it must have a procedure that eliminates the infinities arising in the calculation of physical objects, such as cross sections and decay rates. In order to keep electroweak theory re-normalizable, without breaking the symmetry, its Lagrangian must not contain mass terms. This means that both gauge bosons and fermions in the SM theory are massless. But measurements show that the W^+, W^-, and Z^0 bosons actually have a relatively large mass of around 80 GeV, so a machanism which introduces the mass terms without breaking the re-normalizability is needed.

1.1 The Higgs mechanism

To generate the bosons and fermions masses the spontaneous electroweak symmetry breaking, know also as *Brout-Englert-Higgs mechanism* [7], is used. According to this mechanism, a new scalar, the Higgs boson field, should exist, which is a doublet in SU(2), has a non-zero hypercharge in U(1) and is a singlet in SU(3). The gauge bosons acquire masses by interacting with the Higgs boson field and the masses of the fermions are generated through the so-called "Yukawa interaction" with the Higgs boson.

The Higgs boson field is identified as a complex scalar field with four degrees of freedom:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \begin{pmatrix} (\Phi_1 + i\Phi_2)/\sqrt{2} \\ (\Phi_3 + i\Phi_4)/\sqrt{2} \end{pmatrix} \tag{1.7}$$

where $\Phi_{i=1,2,3,4}$ are real fields while $\Phi^\pm$ are complex. The corresponding Lagrangian is:

$$\mathcal{L}_H = (D_\mu \Phi)^\dagger (D^\mu \Phi)^\dagger - V(\Phi) \tag{1.8}$$

where:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{1.9}$$

D_μ is the covariant derivative which substitutes ∂_μ to ensure invariance under the $SU(2) \otimes U(1)$ local gauge symmetry:

$$\partial_\mu \to D_\mu = \partial_\mu + \frac{i}{2} g_W \sigma_j W_j^\mu + i g_Y Y B^\mu \tag{1.10}$$

where the sum symbol over the repeated index (j=1,2,3) is omitted, g_W and g_Y are respectively the coupling constants of the fermions to the fields W^j and B, σ_j are the Pauli matrices and Y is the weak hypercharge.

The vacuum state corresponds to the minimum of $V(\Phi)$. To obtain a finite minimum, the coefficient λ must be positive, while μ^2 can be either positive or negative. If $\mu^2 > 0$ the resulting potential has a unique minimum at $\Phi = 0$.

By requiring $\mu^2 < 0$, the expectation value in vacuum is different from 0 and the symmetry is broken. In fact, in correspondence of:

$$\Phi^\dagger \Phi = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{1.11}$$

there is a set of equivalent minima connected by an SU(2) transformation, as shown in Fig. 1.1. The choice of the vacuum state breaks the symmetry of the Lagrangian in a way to avoid that a spurious mass is assigned to the photon. Therefore we need to choose a ground state ϕ_0 which conserves the electric charge symmetry group U(1). This condition will be satisfied if we choose:

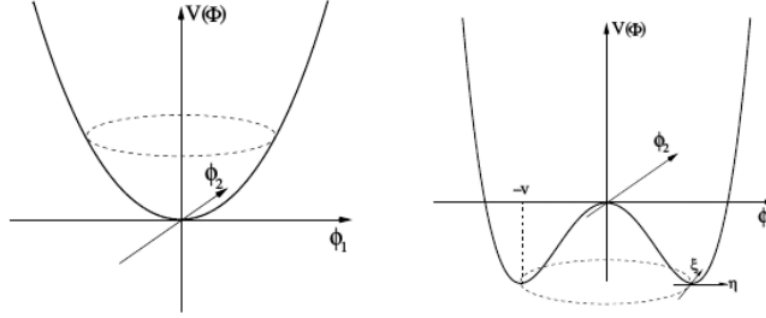$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.12}$$

Figure 1.1. Higgs boson potential before (left) and after (right) symmetry breaking.

The field Φ will then be an expansion around the fundamental state Φ_0:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.13}$$

using this definition the Eq. 1.8 can be written as:

$$
\begin{aligned}
\mathcal{L} = \left(0, \frac{v+h(x)}{\sqrt{2}}) \right)_i \left(\overleftarrow{\partial_\mu} \delta_k^i + i\frac{g_W}{2} W_\mu^A (\sigma^A)_k^i + i\frac{g_y}{2} B_\mu \delta_k^i \right) \\
\left(\overrightarrow{\partial_\mu} \delta_j^k + i\frac{g_W}{2} W_\mu^A (\sigma^A)_j^k + i\frac{g_y}{2} B_\mu \delta_j^k \right) \begin{pmatrix} 0 \\ \frac{v+h(x)}{\sqrt{2}} \end{pmatrix}_j
\end{aligned}
\tag{1.14}
$$

where the indices k, i, j refer to the spinor dimension, redefining W and Z with Eq. 1.6, 1.5 where:

$$cos\theta_W = \frac{g_W}{\sqrt{g_W^2 + g_Y^2}} \tag{1.15}$$

The masses of each boson are given by the coefficients of the quadratic terms of the fields that represent them. By making these terms explicit, we obtain:

$$m_{W^\pm} = \frac{g_W \times v}{2} \tag{1.16}$$

$$m_Z = \frac{v\sqrt{g_W^2 + g_Y^2}}{2} \tag{1.17}$$

$$m_A = 0 \tag{1.18}$$

We can now give mass to the fermions by introducing a Yukawa interaction term, coupling a left-handed fermionic doublet Q_L, a right-handed singlet (d_R)

and the Higgs doublet (Φ). The Lagrangian of interaction between a quark (or a lepton) and Higgs field will be of the form:

$$\mathcal{L} = \sum_{i,k=1}^{F} \left[\lambda_{ik}^{d} \bar{Q}_{L}^{i} d_{R}^{k} \Phi + \lambda_{ik}^{u} \bar{Q}_{L}^{i} u_{R}^{k} \Phi \right] \tag{1.19}$$

where d (u) define the field of the quark down (up); λ is the coupling and i, k are families index. By diagonalizing the mass matrices, the fermion fields gain mass:

$$m_f = \frac{v}{\sqrt{2}} \lambda^f \tag{1.20}$$

So the coupling of the Higgs boson to a given fermion is proportional to the mass.

1.2 Higgs boson production and decay at LHC

CMS [2] and ATLAS [1] experiments at LHC announced in 2012 the observation of a new particle, compatible with the Higgs boson predicted in the SM. The resonance had already been searched at LEP but the center of mass energy was insufficient for a discovery of a resonance at such high energies.

From the theory predicted by the standard model the Higgs boson was expected as a neutral particle with spin 0 and $J^P = 0^+$. Since the coupling of the Higgs boson with the fermions is proportional to the fermion mass, it is expected to decay mainly in a pair of high-mass fermions.

The SM does not predict a value for the Higgs boson mass, therefore it was necessary to probe different energies to search for the resonance. The hadronic collider, also called discovery collider, is ideal for this purpose. In an hadron collider the interactions happen at partonic level, and since each parton carries only a fraction of the proton momentum, the effective center of mass energy is that of the interacting partons, and depends on their parton density functions (i.e the momentum distribution functions of the partons within the proton).

The Higgs boson is produced via four mechanisms: *gluon gluon fusion* (ggH), *vector boson fusion* (VBF), *associated production* with an electroweak vector boson (VH) or with the top quark (ttH). The Feynman diagrams related to these process are shown in Fig. 1.2

- Gluon Gluon fusion (ggH): it is the production with the highest cross-section at the LHC. Since the Higgs boson is not involved in the strong interaction and the gluon is massless, it could not be coupled directly with the gluon, so the process $gg \rightarrow H$ takes place through a loop of fermions. As the Higgs boson coupling to fermions is proportional to the fermion masses, the dominant contribution is given by the loop of a top quark

- Vector Boson Fusion (VBF): the process is $\bar{q}q \rightarrow \bar{q}qH$ and the Higgs boson is produced through the fusion of two vector bosons ($W^\pm$ or Z^0), radiated off by the two colliding quarks. The two leading final state jets are energetic and characterized by a large rapidity gap between the two

- Production in association with an electroweak vector boson (VH), where V=$W^\pm$, Z^0: the process involves gluon or quarks that must be taken from the sea[1] of the proton and enough center-of-mass energy is requested to produce both the Higgs boson and the vector boson. For this reason the cross section of this process is small.

- Production in association with a top quark (ttH): the process could involve both gluon ($gg \rightarrow t\bar{t}H$) or quarks ($q\bar{q} \rightarrow t\bar{t}H$). The Higgs boson can be produced in association with a pair of any fermion type, but since the top is the heaviest one, a final state with two top quark is favorite.
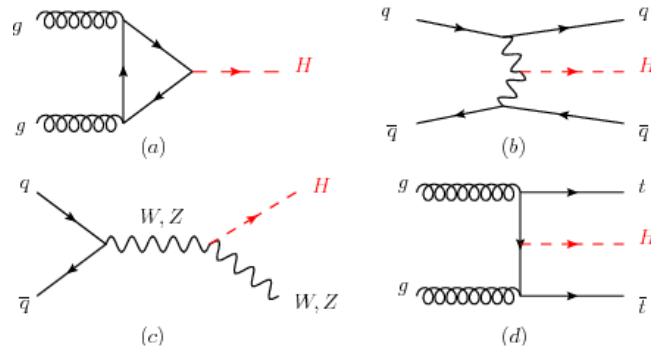

Figure 1.2. Feynman diagrams contributing to the Higgs boson production in (a) gluon fusion, (b) weak-boson fusion, (c) associated production with a gauge boson and (d) associated production with top quarks.

The SM predicts different cross sections for each Higgs boson production metho as a function of its mass m_H as shown in Fig. 1.3.

---

[1]Any hadron may contain an indefinite number of virtual "sea" quarks, antiquarks, and gluons, which do not influence its quantum numbers. This type of particle transport a small fraction of the hadron momentum.
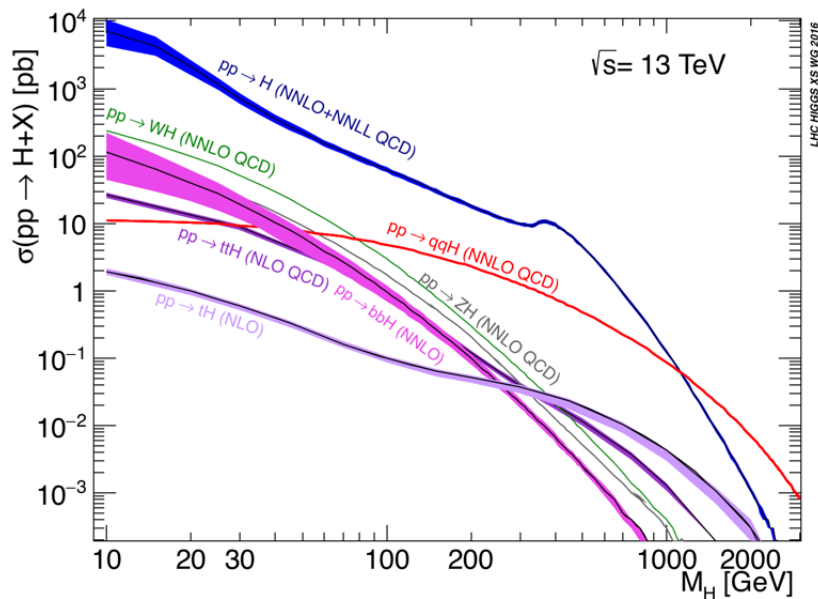
Figure 1.3. Higgs boson production cross sections at LHC at a center of mass energy $\sqrt{s}$ = 13 TeV as a function of the Higgs boson mass M_H. The blue line is the cross-section for the ggH production modes, the red for the VBF production mode, the green and the black ones are the cross-section of WH and ZH, respectively, and the violet line is the ttH cross section.

Since in the SM the width of the Higgs boson ($\Gamma_H \sim 4.1$ MeV) is such that it decays immediately after its production, the experiments can detect the products of the Higgs boson decays. Once the Higgs boson mass is fixed, the production and decay rates of the Higgs boson particle are uniquely determined [8].

The Higgs boson is a neutral particle that couples only to massive particles so it can not decay directly into two photons but the process is possible at NLO through fermion and massive vector boson loops, as shown by the Feynman diagrams in Fig. 1.4. The Branching Ratio of this decay is quite low (0.23 % for $m_H = 125$ GeV, as shown in Tab. 1.3), but the clean final state and its excellent mass resolution made this channel essential for the discovery of the Higgs boson.
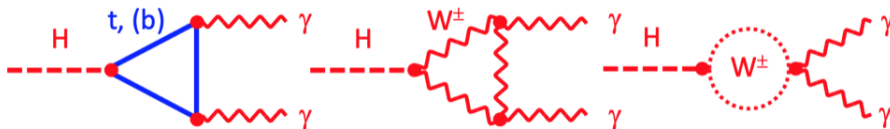

Figure 1.4. Feynman diagrams for the $H \to \gamma\gamma$ decay.

Fig. 1.6 shows the invariant mass distributions of the two photons in the channel $H\gamma\gamma$ made by CMS and ATLAS experiment for all the production

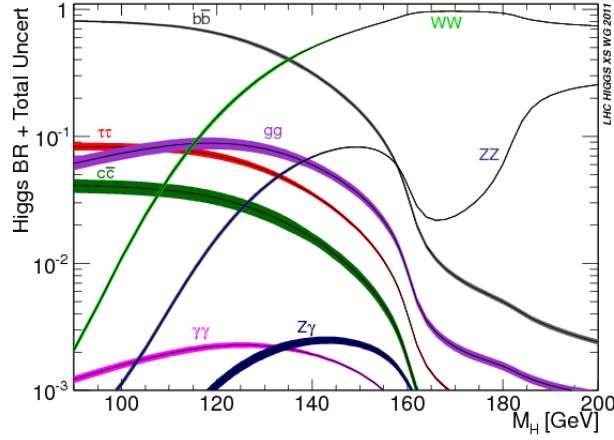modes together with all data at the time of the discovery.

Figure 1.5. Higgs boson decay Branching Ratio (i.e. fraction of particles that decay following a certain decay channel with respect to the total number of particles that decay in any channel) as a function of M_H.

Decay Channel	Branching Ratio
$H \to b\bar{b}$	5.84×10^{-1}
$H \to W^+W^-$	2.14×10^{-1}
$H \to \tau^+\tau^-$	6.27×10^{-2}
$H \to ZZ$	2.62×10^{-2}
$H \to \gamma\gamma$	2.27×10^{-3}
$H \to Z\gamma$	1.53×10^{-3}
$H \to \mu^-\mu^+$	2.18×10^{-4}
$H \to ZZ \to 4l \ (l = e^\pm, \mu^\pm)$	1.24×10^{-4}
$H \to WW \to 2l2\nu \ (l = e^\pm, \mu^\pm)$	1.24×10^{-2}

Table 1.3. Braching ratios for each possible decay channel for $m_H = 125 GeV$
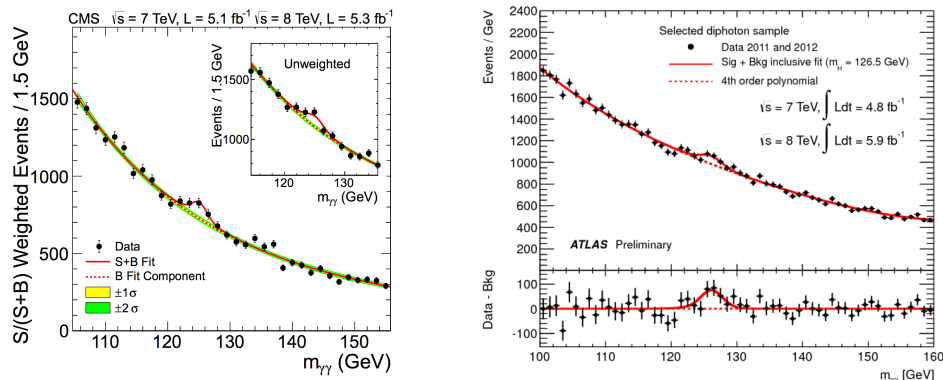
Figure 1.6. On the left the CMS diphoton invariant mass distribution with each
event weighted by the S/(S+B) value of its category (the photon were classified
into four categories using the information of the detector part in which they
are produced) where S stands for "signal" and B for "background". The lines
represent the fitted background and signal, and the colored bands represent the
±1 and ±2 standard deviation uncertainties in the background estimate. On the
right, the ATLAS distribution where the red line is the fitted distribution, and
in the lower part the residuals of the data with respect to the respective fitted
background component is shown.

1.3   The CP quantum number

In physics, the $\mathbb{C}$ and $\mathbb{P}$ operator give rise to multiplicative quantum
numbers that describe the particles behavior under the symmetry operation of
charge conjugation and parity, respectively.

The parity operation transforms a system into its mirror image, flipping the
sign of one spatial coordinate, so a particle with momentum $\vec{p}$ is transformed
into a particle with momentum $-\vec{p}$. This symmetry is not preserved under the
weak interactions, as demonstrated by the Wu experiment conducted on 1956
studying the decays of ^{60}Co.

The charge conjugation is a transformation that switches all particles in
their corresponding antiparticles, thus changing the sign of all charges but
leaving untouched the space and time variable.

The operators C and P are not conserved separately and also the CP
invariance is an approximate symmetry of Nature, as demonstrated by the
experiments on kaons on 1964 and on B-mesons at the B-factories in the 2000s.
What currently appears to be conserved is the CPT symmetry: physical
states seem to be invariant under the consecutive application of the operators
Charge-conjugation, Parity and Time-reversal.

The quantum numbers of the Higgs boson are predicted by the SM, leading
also to unique predictions for the angular and energy distributions of its decay
products. In particular the decay widths of the Higgs boson into massive gauge

bosons V = W, Z are directly proportional to the HVV couplings, which in the SM are given in terms of the fields by:

$$\mathcal{L}(HVV) = \left(\sqrt{2}G_F\right)^{1/2} M_V^2 \Phi V_\mu V^\mu \tag{1.21}$$
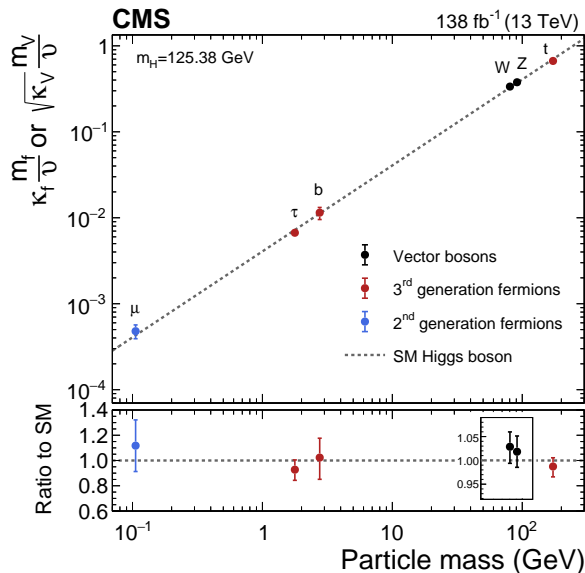
W where G_F is the Fermi coupling constant.

Figure 1.7. Higgs boson couplings to different particles. The dashed line corresponds to the SM expectation. The ordinates are different for fermions and massive vector bosons to take into account the expected SM scaling of the coupling with mass, depending on the type of particle. The result of the fit is shown as the continuous line while the inner and outer bands represent the 68 % and 95 % CL confidence regions [9].

The Eq. 1.21 is compatible with $J^{CP} = 0^{++}$ quantum numbers of the SM Higgs boson particle.

Experimentally, the charge conjugation quantum number C = +1 follows from the observation of decays into $\gamma\gamma$ for pure states in a C-invariant theory [10]. The spin/parity quantum numbers can be investigated by analyzing the helicity amplitudes of production and decay processes. They do not only provide necessary and sufficient conditions to assign the J^{PC} quantum numbers of real states but also test possible CP-violating Higgs bosons.

The spin-parity properties of the Higgs boson have already been analyzed with the data of Run 1, corresponding to an integrated luminosity of up to $5.1 fb^{-1}$ at a center-of-mass energy of 7 TeV and up to $19.7 fb^{-1}$ at 8 TeV. The decay channels $H \to 4l$ and $H \to \gamma\gamma$ were investigated discarding the hypotheses of a pure pseudoscalar Higgs at 99.95 % C.L. and the hypothesis of a spin-2 Higgs particle at 99.87% C.L. [3]. An example distribution of the test statistic and observed value in the case of the SM Higgs boson and the

spin-two hypothesis is shown in Fig. 1.8

The possibility of a spin 1 particle is discarded as a consequence of the Landau-Yang [11] theorem for which no on-shell spin-1 particle can decay into two on-shell massless photons.
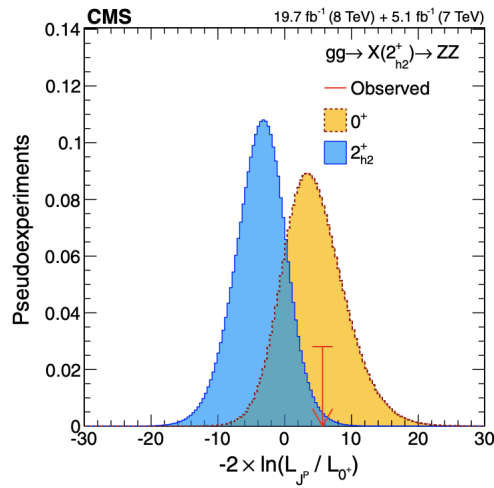

Figure 1.8. Distribution of the test statistic $q = -2ln(L_{j^P}/L_0^+)$ for the hypothesis of spin-2 Higgs (2_m^+) particle tested against the SM Higgs boson hypothesis (0^P). The expectation for the SM Higgs boson is represented by the yellow histogram. The measurement correspond to the red arrow.

Chapter 2

LHC and CMS

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator in which two beams of proton collide with a center of mass energy up to 14 TeV. The LHC consists of a 27-kilometre ring of superconducting magnets with accelerating structures to boost the energy of the particles along the way.

The collected data is used for analysis that ranges from precision measurements of standard model parameters to searches for new physics beyond the SM. Among the great successes of the accelerator is the discovery of the Higgs boson: after a 40-year search, on 4th July 2012 the CMS and the ATLAS collaboration announced the observation of a new particle, consistent with the SM Higgs boson, at a mass around 125 GeV.

In LHC protons undergo under process of collimation and acceleration before the collisions to reach the maximum energies and performances possible from the detection of the collision products. The protons are extracted from a hydrogen bottle and accelerated in small groups, "bunches", until they reach 50 MeV. They are then injected in the Proton Synchrotron Booster (PSB), a 4 rings accelerator, in which they reach the energy of 1.4 GeV. The PSB provides the protons with enough energy to be injected in the Proton Synchrotron (PS), an accelerating ring of 628 m, in which the protons are brought to 25 GeV. The bunches are then sent into the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV; finally, the beams are injected into the main ring in opposite directions and in separate beam pipes. When they reach their maximum energy (it is design to have 7 TeV per beam), they are brought to collision in four points, where LHC four major experiments are located: ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (LHC beauty).
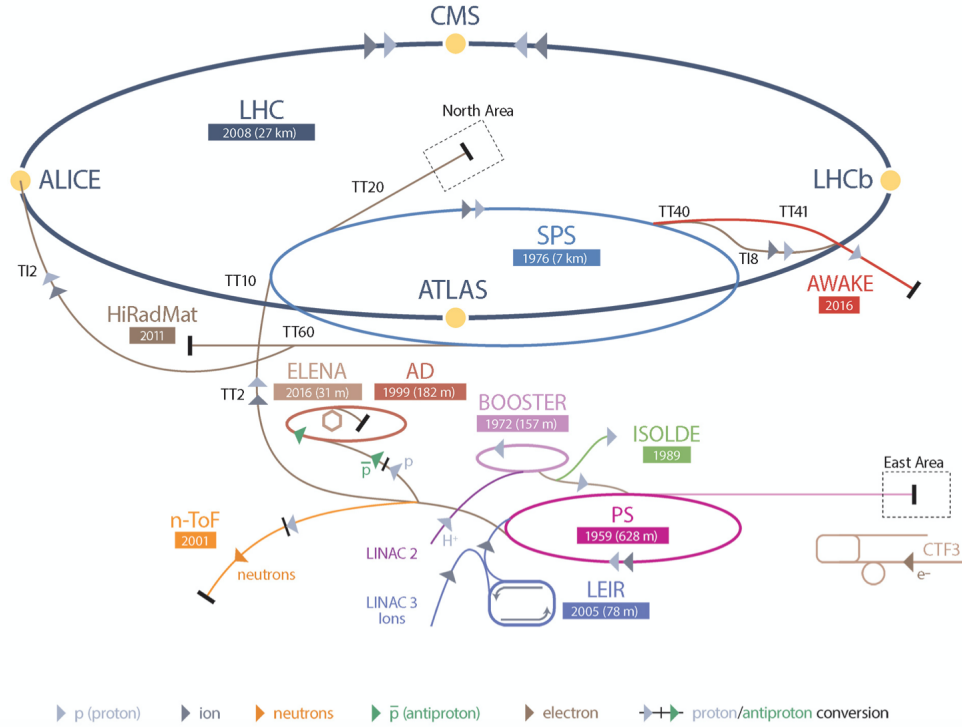
Figure 2.1. Structure of CERN accelerating complex. Linac2, Booster, SPS and LHC represent the acceleration chain protons undergo before collision; interaction points are shown along the LHC ring.

During the years of LHC data taking both collision energies and instantaneous luminosity have been increased. The instantaneous luminosity is defined as:

$$L = \frac{n_b N^2 f_{rev}}{4\pi \beta^* \epsilon_n} \gamma R \qquad (2.1)$$

where γ is the relativistic gamma factor, n_b is the number of bunches colliding at the interaction point (IP), N is the number of protons per bunch, f_{rev} is the bunches revolution frequency in the ring, β^* is the beam focal length and ϵ_n is the beam transverse normalized emittance and R is a luminosity geometrical reduction factor.

Ad seen in Eq. 2.1 the luminosity depends only on accelerator parameter and defined the ability of an accelerator to produce a certain number of interactions.

Let be σ_i the cross-section of the i process happening at LHC, the number of expected event (N_i) is:

$$N_i = \sigma_i \int dt L(t) \qquad (2.2)$$

Where $\int L(t)dt$ is the integrated luminosity.

The price to pay for brighter, more focused beams is a higher number of interactions per bunch crossing (pile up) , which will increase from an average of about 32 (2018, 13 TeV data) to 140 or 200 units per HL-LHC. To handle the pile-up the detectors must have rapid response times in order to distinguish the different particles coming from different bunch. In addition, the detectors must be resistant to radiation damage due to the high flow of particles produced as a result of collisions.
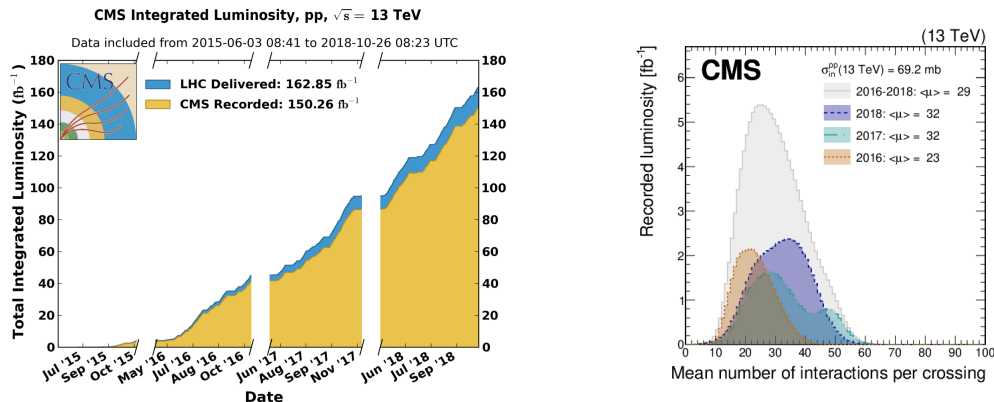
Figure 2.2. On the left Cumulative measured luminosity from 2015-06 to 2018-10. On the right distribution of the mean number of inelastic interactions per crossing (pileup) in data for pp collisions in 2016 (dotted orange line), 2017 (dotted dashed light blue line), 2018 (dashed navy blue line), and integrated over 2016-2018 (solid grey line) [12].

Since the particles contained in the two colliding beams have electric charge of the same sign, there must be two separate cavities and magnetic fields for the beams circulating in opposite directions. LHC has 1232 superconducting magnets made of Ni-Ti, which are 14.2 m long and are cooled to 1.9 K with liquid He, in order to reach a magnetic field of 8.3 T. Tab. 2.1 shows the main LHC characteristics.

Circumference	27 km
Number of magnet dipoles	1232
Dipolar magnetic field	8.3T
Magnet temperature	1.9K
Beam energy	7 TeV
$\sqrt{s}$	14 TeV
L	$2 \times 10^{34} cm^{-2} s^{-1}$
Number of protons per bunch	1.05×10^{11}
Bunch length σ_z	75 mm
Bunch radius $\sigma_x = \sigma_y$	$16 \mu m$
Number of bunches	2808
Space - time between bunches	$7.48 m^{-25} ns$

Table 2.1. LHC design parameters.

Until now, there have been two completed period of collisions and data taking, called "Run", at three different center of mass energies and instantaneous peak luminosity. Between each run a shutdowns period is always planned for maintenance and upgrades. The first run (Run1) started in 2011 with a centre of mass energy of 7 TeV and, after an upgrade, of 8 TeV. The second Run (Run 2), started in 2015, reached a centre of mass energies of 13 TeV. The third Run (Run3) scheduled for March 2022 was delayed due to the pandemic situation. The first collision of this run was on July 5 2022 and is still ongoing.

The collider is expected will continue its activity at least until 2038.

2.1 CMS

The Compact Muon Solenoid (CMS) is a general purpose detector designed to optimize the discovery potential of the LHC collider. CMS is composed by a cylindrical barrel and two endcaps; its total length is 21.6 m, the diameter is 15 m and the weight is about 12500 tons. Figures 2.3 show the CMS detector which is made of a large superconducting solenoid producing a magnetic field of 4T and containing a full silicon tracker for momentum measurements and charged particles tracks reconstruction, a precise crystal electromagnetic calorimeter (ECAL) and a hadron calorimeter (HCAL). Muon chambers are embedded in the iron return yoke of the magnet, thus forming a compact muonic system.

The absolute CMS reference frame is defined with respect to the LHC ring. The x axis points towards the center of the ring, the y axis points up and

the z axis is defined as for a right handed coordinate system. The detector is divided into 3 parts: a cylindrical part in the middle called barrel region, and two disks at the two sides of the cylinder called endcap regions. Since CMS has a cylindrical symmetry, a cylindrical coordinate system (r,η,ϕ) is adopted, where r is the radial distance from the origin, ϕ is the azimuthal angle which is measured from the x axis in the xy plane, η is the pseudorapidity which is defined as:

$$\eta = -ln(tan\frac{\theta}{2})$$ (2.3)

Where θ is the angle which runs from from the z axis (at 0°) to the y axis. In Fig. 2.4 we show the values of the pseudorapidity η corresponding to some polar angles.

The angular distance in η-ϕ coordinates is defined as:

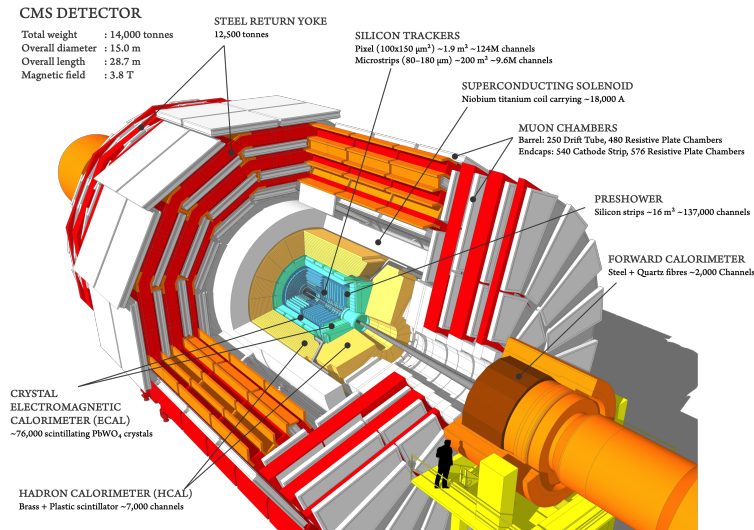$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$$ (2.4)
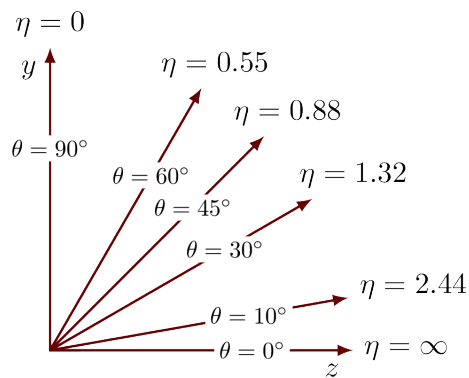

Figure 2.3. Section of CMS.

Figure 2.4. Values of the pseudorapidity η corresponding to some polar angles θ.

2.1.1 Tracker

The internal tracker [13] is the closest detector to the beams and has the purpose of detecting charged particles through their energy losses by ionization and allowing, thanks to the bending effect of the magnetic field provided by the solenoid, the measurement of their impulse.

The CMS Silicon tracker is composed of different substructures. Closest to the Interaction Point (IP) is a Silicon Pixel detector, with about 66 million $100{\times}150\ \mu m^2$ pixels arranged at distance of 4 to 11 cm from the beam line on a cylindrical barrel and end-caps structure with total length of 92 cm, this detector will not be further described in this paper. The Silicon Strip detectors are divided in the inner barrel part (TIB), the inner disks (TID), the outer barrel (TOB) and outer end-caps (TEC). The layout of the Tracker substructures is sketched in the figure below 2.5.

The material thickness in the tracker crossed by electrons and photons before reaching the ECAL depends on η . It goes from about 0.35 radiation length (X_0) at central ($\eta = 0$), increasing towards the ECAL barrel-endcap transition, and then falling back. The simulation of the material budget in front of ECAL is of great importance for the estimation of the photon conversion (photons interacting with the tracker material producing an electron-positron pair) for photon analyses as lost signal.
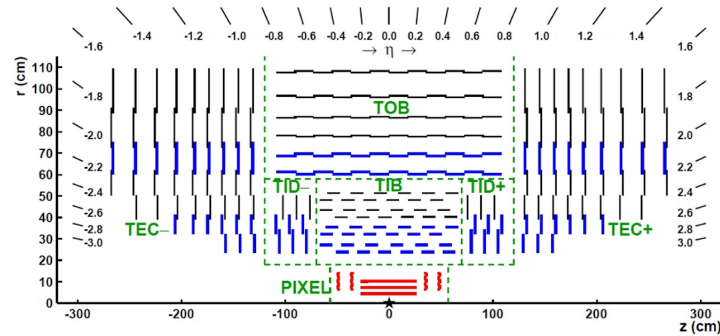
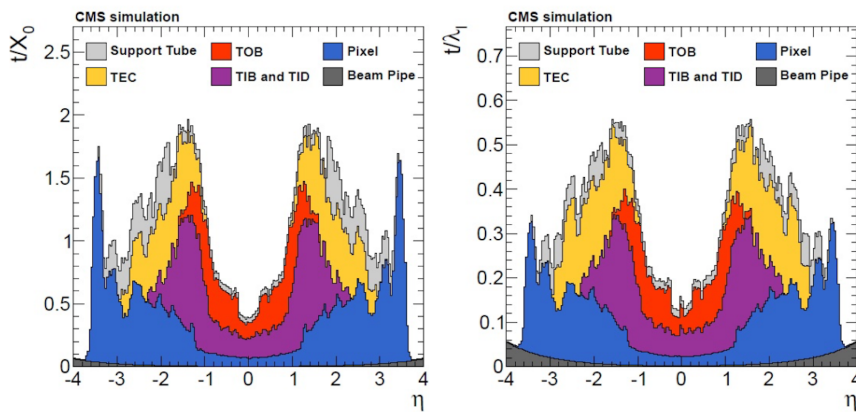**Figure 2.5.** Section of a half of the tracking system of CMS, with some values of the pseudorapidity.

Figure 2.6. Total thickness t traveled by a particle produced in the nominal center of interaction in units of radiation length X_0 (left) and nuclear interaction length λ_I (right) as a function of η.

2.1.2 The Electromagnetic Calorimeter (ECAL)

The Electromagnetic Calorimeter (ECAL) [14] of the CMS experiment at the LHC is a hermetic, fine grained, homogeneous calorimeter, containing 75,848 lead tungstate scintillating crystals. It has been designed to have an excellent energy resolution for electrons and photons, and has been optimized for research of the channel $H \rightarrow \gamma\gamma$.

The ECAL design requirements were:

- Excellent energy and position/angle resolution up to $|\eta| < 2.5$, to match the tracker coverage.

- Hermeticity, compactness and high granularity

- Fast response ($\sim 25ns$) and particle id, energy and isolation measurement at trigger level

- Large dynamic range (5 GeV to 5 TeV) and excellent linearity

- Radiation tolerance

ECAL consists of 61200 crystals of stolzite, i.e. lead tungstate ($PbWO_4$), in the barrel region, and is closed by 7324 crystals in each endcap. Particles such as electrons and photons interact in ECAL producing an electromagnetic shower and depositing all their energy in the calorimeter. The longitudinal and transverse length of the shower depend on the radiation length (X_0) and the Molière (R_M) radius, respectively, both of which are material-specific parameters. The main features of $PbWO_4$ scintillating crystals are high density ($\rho = 8.28g/cm^3$), extremely short radiation length and small Moliére radius ($X_0 = 0.85cm$, $R_M = 2.19cm$), allowing the realization of a homogeneous compact calorimeter with high granularity. It produces fast signals, 80% of the light is emitted in 25 ns. The light emission peak is at $\sim 420nm$ (i.e. blue light). The crystals are transparent to their entire scintillation emission spectrum.

The problematic aspects of $PbWO_4$ are: the reduced light yield (LY), only 100 photons per MeV for a 23 cm long crystal, that requires the use of a photodetector readout system with internal gain and a strong light yield dependence on temperature ($\Delta(LY)/\Delta T = -2\%/°C$ at $\sim 18°C$) which imposes a requirement on temperature stability of $\pm 0.05°C$.

As CMS, ECAL has a cylindrical structure. The barrel part covers the region $\eta < 1.47$ and is made by 61200 trapezoidal crystals of about 23 cm (25.8 X_0) in length and approximately 1 R_M in lateral size, measuring $22 \times 22\ mm^2$ in the frontal extremity, while $26 \times 26mm^2$ in the opposite extremity. The Crystals are arranged in a quasi-projective geometry, and use avalanche photodiodes (APD) as the photosensitive device.

The endcaps consist of two detectors, a preshower (ES) device followed by a the $PbWO_4$ scintillator calorimeter.The scintillation light is detected by vacuum phototriodes (VPTs). The preshower is used for particle identification in the endcap regions of CMS. Each ES is made of two orthogonal layers of silicon sensors, interspersed with lead layers that serve to generate electromagnetic showers. The principal aim of the preshower is to distinguish photons from pions (π^0) that decays in two close photons at high energy, and thus difficult to be discriminated. Each endcap calorimiter is made by 7324 rectangular and quasi-projective crystals of $\sim 1.3R_M$ in lateral size and about $24.7X_0$ in depth.

Notice that at value of η larger than 2.5, the radiation level is so high than precision measurements are almost impossible.

ECAL energy resolution varies with the particle energy through the expression:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{2.5}$$

where S is the stochastic term, N is the noise term and C is the constant term. The stochastic term depends on the fluctuations in the number of scintillating photons detected and in the number of processes through which the particles lose their energy in the crystals. The noise term comes from the electronic noise and from the pileup. The constant term has different causes: losses due to failures in the longitudinal containment, non uniformity in the light collection, intercalibration between crystals and geometrical imperfections.

The parameters that appear in Eq. (2.7) have been measured with test beams, and are estimated as:

$$S = 2.8\% GeV^{-1} \tag{2.6}$$

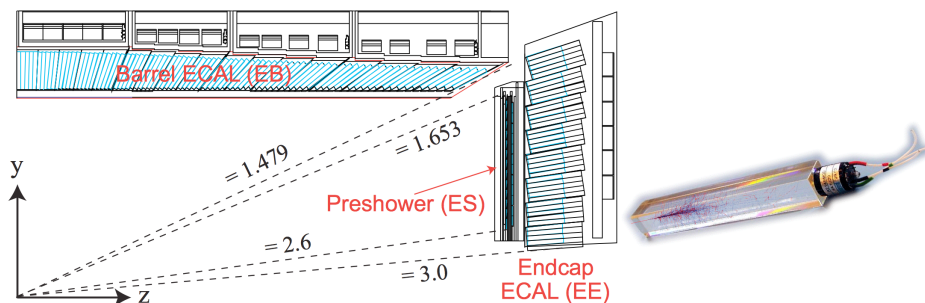$$N = 124 GeV \tag{2.7}$$

$$C = 0.3\% \tag{2.8}$$

Figure 2.7. section of a quarter of the ECAL, with some values of the pseudorapidity.

2.1.3 Hadronic calorimeter (HCAL)

The purpose of the hadron calorimeter (HCAL) [14] is to measure the energy and direction of the jets. CMS uses a a sampling calorimeter, in which the material that produces the particle shower is distinct from the material that measures the deposited energy. Very dense material can be used to produce a shower that evolves quickly in a limited space, even if the material is unsuitable for measuring the energy deposited by the shower. A disadvantage is that some of the energy is deposited in the wrong material and is not measured. HCAL is formed by alternating layers of copper absorber and fluorescent scintillating material and it is highly hermetic. The energy resolution in HCAL depends on the objects showering and it is function of the deposited energy. For hadronic

jet showers.

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

with S = 125% stochastic term and C= 5% constant term .

HCAL has a thickness between 7 and 10 interaction length (λ), which in general would not be sufficient to contain the full hadronic shower. Therefore other scintillators are set outside the solenoid, which bring the number of interaction lengths in each direction to 10.

In the barrel (hadronic barrel, HB) a calorimeter unit cell has the same surface area as a 5×5 square of ECAL cells. HB covers the zone up to $\eta = 1.3$, together with the layers of the calorimeter outside the solenoid (hadronic outside, HO), while the two endcaps (HE) reveal particles with $1.3 < \eta < 3.0$.

2.1.4   Muon Detector

Detecting muons is one of the most important tasks for the Compact Muon Solenoid. Muons are the only charged particles able to pass across the calorimeters without being absorbed. Therefore, chambers to detect muons are placed outside the magnetic coil of CMS, embedded in the return yoke, to fully exploit the 1.8 T return flux; there muons can be the only particles likely to register a signal. It is formed by four layers of drift cambers interlined with the magnet return yoke in the barrel, while is composed by Cathode Strips Chambers (CSC) and Resistive Plate Chambers (RPC) in the endcaps. A measurement of the transverse moment of muons is possible thanks to the magnetic field provided by the return yoke ($\sim 1.5T$) so that the muon trajectories are bent, and muon momentum can be measured. The use of both CSCs and RPCs in the endcaps provides for fast and precise position measurements and a redundant trigger for rapid data selection.

2.1.5   Trigger

At the nominal luminosity of LHC the event rate is equal to 109 Hz, to many data to store and analyze. it is therefore evident that the trigger [15] plays an enormous role in background rejection and process identification in CMS environment.

The trigger is composed by a two levels system: Level 1, which is an hardware trigger with a fixed latency of 3.4 μs, has to decide within this latency which data are worth taking. The time is not enough to read the information of all the detectors, therefore L1 uses only information of the calorimeters and of the muon chambers. The second level, High Level Trigger (HLT), is an array of computers running high-level physics algorithms. It is divided into three sublevels: the first access only the data of calorimeters and

muon system, the second adds the pixels, and the third reads the information of the entire event.

The L1 trigger thresholds are adjusted to provide an output o 100 kHz, while the HLT delivers a 400 Hz rate to the storage stations.

2.2   ECAL calibration with π^0 using early Run 3 data

After three years and a half of shutdown, on July 5 2022 at 4.47 p.m., there was the first proton-proton collision of the Run 3 at a $\sqrt{s} = 13.6$ TeV, slightly larger than the Run 2 one. The CMS experiment recorded these first collisions with all its systems turned on and highly operative.

The Run 3 is expected to deliver an integrated luminosity of $\sim 150 fb^{-1}$ from 2022 to 2024 with a centre of mass energy of 13.6 TeV. The expected instantaneous luminosity, up to $7.5 \times 10^{34} cm^{-2} s^{-1}$ means a higher instantaneous pile up to handle. Such a challenging pile-up warranted the recent CMS experiment upgrade of key components to resolve the multiple interactions and efficiently analyze the data. For example, the CMS pixel tracking detector, after undergoing an important upgrade between 2016 and 2017, has been further upgraded in its innermost layer in order to handle the high rates of Run3; the hadronic calorimeter (HCAL) has been improved with enhanced readout granularity along the shower depth; a new muon detectors, the Gas Electron Multiplier (GEM) detectors, have been installed. Moreover also the trigger and the data acquisition system's performances have been increased.

The excellent performance in the reconstruction and identification of high energy photons and electrons play a key role in the observation of the Higgs boson and the study of its properties. Maintaining and improving the excellent ECAL performance during Run 3 in the harsher environment is crucial for all the physics analyses that include photons or electrons in their final state. Achieving this goal requires a continuous effort in the operation, monitoring and calibration, and simulation of the calorimeter. Given the importance of the ECAL performance in the $H \rightarrow \gamma\gamma$ decay sensitivity, it is crucial to mantain optimal photon energy calibrations. I had the opportunity to analyze the eary Run 3 data from the π^0 calibration data stream. The data is used relate to about 10 pb^{-1} of the LHC fill 7963 taken on 11 July 2022. Fig. 2.8 show the integrated and istantaneous luminosity of the considered fill.
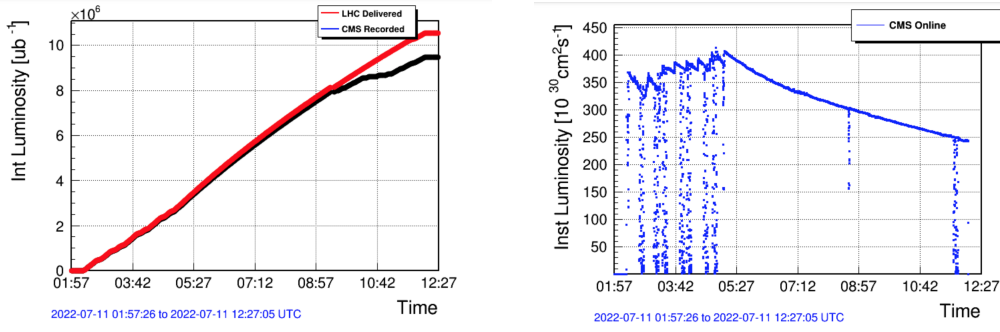
Figure 2.8. On the left. integrated luminosity through the 7963 fill. On the right instantaneous luminosity through the 7963 fill.

Photon pairs produced in neutral pion decays are selected and the invariant mass is calculated:

$$m_{\gamma\gamma} = \sqrt{2E_1E_2(1 - cos\alpha)} \tag{2.10}$$

where E1 and E2 are the photon energy and α is the aperture angle between the two. The photon invariant mass distribution is plotted and fitted with two Gaussian functions, to model the resonant signal, and a second order polynomial to model the large combinatorial background of two ECAL clusters. Let m be the mean values of the Gaussians (i=1,2) and σ_i the respective standard deviations, the peak position $(m_{\pi^0}^{FIT})$ and the resolution $\sigma_{\pi^0}^{FIT}$ are calculated as:

$$m_{\pi^0}^{FIT} = \frac{\frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \tag{2.11}$$

$$\sigma_{\pi^0}^{FIT} = \sqrt{\frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}} \tag{2.12}$$

The results are shown separately for the Endcap region (EE) and for the Barrel region in Tab. 2.2 and Tab. 2.3.

	normalization	$(401.7 \pm 1.8) \times 10^4$
Gaussian 1	m_1	$(135.3496 \pm 0.0086) MeV$
	σ_1	$(11.142 \pm 0.016) MeV$
	normalization	$(1968 \pm 18) \times 10^3$
Gaussian 2	m_2	$(137.379 \pm 0.013) MeV$
	σ_2	$(18.991 \pm 0.058) MeV$
polynomial	a	$(-319.7 \pm 2.3) \times 10^5 \ GeV^{-2}$
degree 2	b	$(851.1 \pm 6.5) \times 10^4 \ GeV^{-1}$
$ax^2 + bx + c$	c	$(1105.1 \pm 3.5) \times 10^4$
$m_{\pi^0}^{FIT}$		135.8(7) MeV
$\sigma_{\pi^0}^{FIT}$		9.6 MeV

Table 2.2. fitted parameter values for the barrel. $m_{\pi^0}^{FIT}$ and $\sigma_{\pi^0}^{FIT}$ are, respectively, the weighted mean and sigma of the two Gaussian

	normalization	$(203.4 \pm 3.5) \times 10^3$
Gaussian 1	m_1	$(123.105 \pm 0.073) MeV$
	σ_1	$(20.22 \pm 0.14) MeV$
	normalization	$(271.6 \pm 3.6) \times 10^3$
Gaussian 2	m_2	$(119.276 \pm 0.30) MeV$
	σ_2	$(10.838 \pm 0.63) MeV$
polynomial	a	$(-221.25 \pm 7.3) \times 10^5 \ GeV^{-2}$
degree 2	b	$(1028.7 \pm 2.0) \times 10^4 \ GeV^{-1}$
$ax^2 + bx + c$	c	$(-581.2 \pm 1.1) \times 10^3$
$m_{\pi^0}^{FIT}$		120.1(3) MeV
$\sigma_{\pi^0}^{FIT}$		9.5 MeV

Table 2.3. fitted parameter values for the endcaps. $m_{\pi^0}^{FIT}$ and $\sigma_{\pi^0}^{FIT}$ are, respectively, the weighted mean and sigma of the two Gaussian
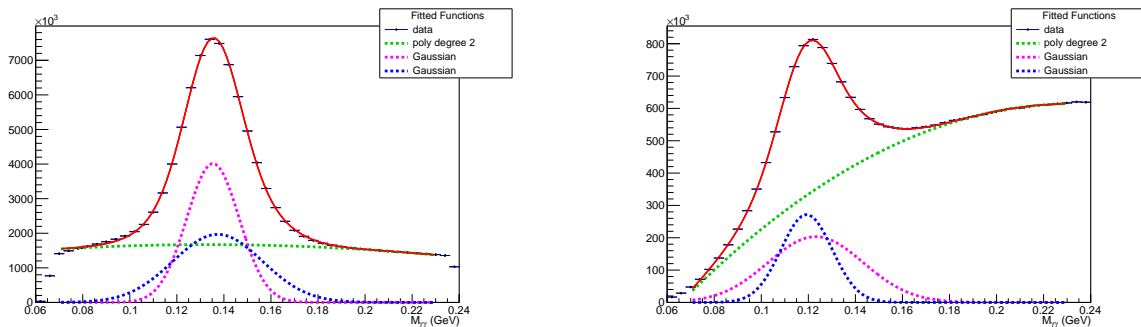
Figure 2.9. Invariant mass distributions in the EB (left) and in the EE (right). Data (black points) are fitted with the sum (solid red line) of a signal and background components (dashed green lines) modeled with two Gaussian functions and a second-order polynomial, respectively. Note the different sizes of the selected samples due to the lower signal purity and selection efficiency in EE. The difference in the shape of the background component is mainly induced by the larger p_T thresholds applied in EE. The effect is also correlated with the presence of the preshower, where a large fraction of the incoming π^0 is deposited: this reflects into a lesser energy deposited in EE.

The positions of the peaks are not compatible with the nominal mass of the π^0 meson: $m_{\pi^0} = 134.9768(5)MeV$, therefore the energy scale has to be corrected, in particular the scale of the Endcap will have to be adjusted more significantly. This is due to the fact ECAL crystals undergo a change of transparency due to radiation damage during periods of LHC operation and the loss of transparency is larger at higher η due to heavier radiation damage. The obtained π^0 peak resolution for the Barrel region is $\sigma^{EB}/E = 7.1\%$ and $\sigma^{EE}/E = 7.9\%$ for the Endcap. The results are compatible with what was expected considering the fact that the photons deriving from the pions decay have low energy ($p_T \sim 1-2$ GeV), while for photons deriving from the decay $H \to \gamma\gamma$ in the barrel we expect $\sigma_E \sim 1\%$.

2.3   High Luminosity LHC

The High Luminosity (HL-LHC) [14] program coincides with Phase II and is set to start in 2026, after the Long Shutdown 3. The accelerator will provide to CMS an additional integrated luminosity of about $2500fb^{-1}$ in over 10 years of operation. This will substantially enlarge the mass reach in the search for new particles and will also greatly extend the potential to study the properties of the Higgs boson

The increase in data acquisition is a breach towards new physics, but represents also a major challenge for the detectors. The main aim of the CMS upgrade is to preserve the experiment's current performance, in terms of background rejection and particle identification, in the more challenging HL-LHC frame. The proposed operating scenario is to level the instantaneous

luminosity at $5 \times 10^{34} cm^2 s^1$ from a potential peak value of $2 \times 10^{35} cm^2 s^1$ at the beginning of fills, and to deliver $250 fb^{-1}$ per year. Under these conditions the event PU will rise substantially to become a major challenge for the experiments, and the performance degradation due to integrated radiation dose will need to be addressed.

A complete new tracker will substitute the current one, which will suffer of significant radiation damage, with a factor 4 increase in granularity in all layers to cope with higher pile up and multiplying tracks. Endcap calorimeters will be also totally renewed: the calorimeter replacements will have higher segmentation, providing for detailed three dimensional images of showers. Precise shower timing information, with $O(20ps)$ resolution, will be obtained by new front-end electronics installed in ECAL barrel section. The muon system will be boosted by auxiliary chambers to provide information to reject background.

To further improve the CMS upgrade: the insertion of a new detector dedicated to precision timing measurements of charged particles, the Mip Timing Detector will be set. It will bring CMS the new capability of precisely measuring the timing of charged particles, thus helping in pile up reduction in HL-LHC events.

The study of the Higgs boson will continue to be central to the program. It will include precise measurements of the Higgs boson couplings, probing of its tensor structure, and the search for rare SM and BSM decays
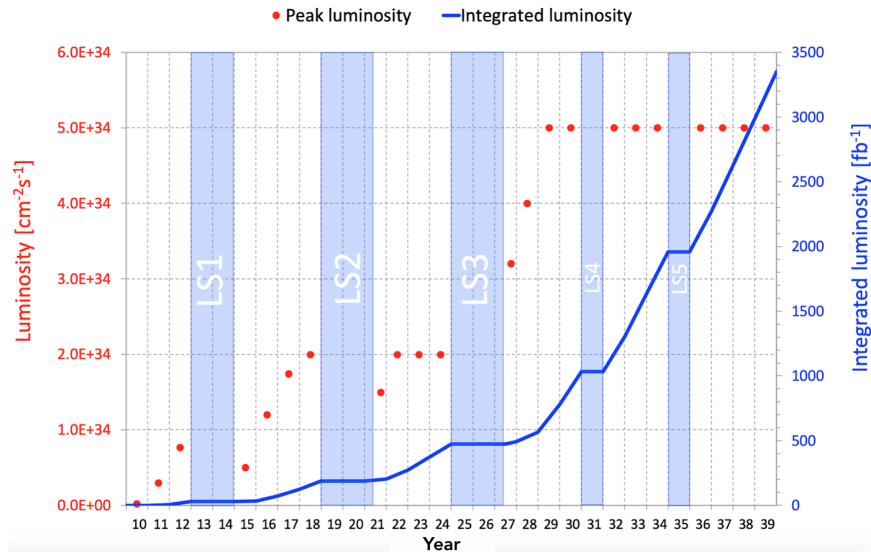

Figure 2.10. Instantaneous and integrated luminosity prospective for LHC operations until 2039.
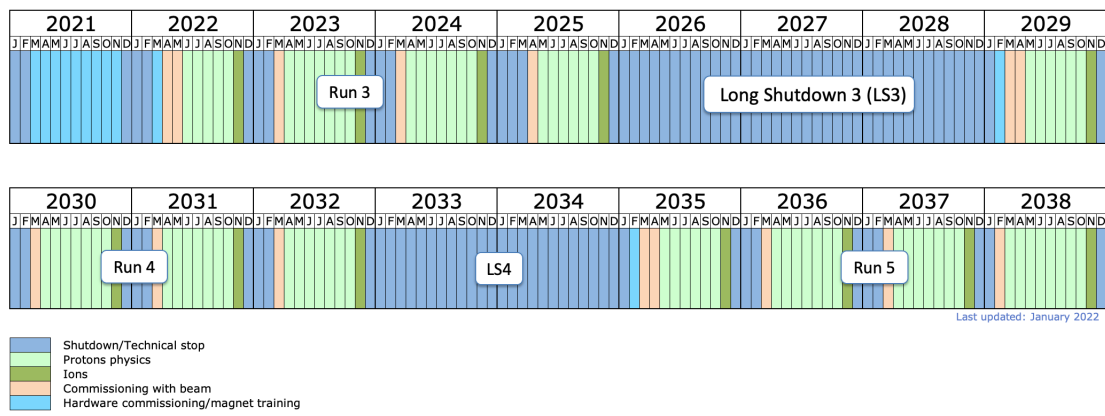
Figure 2.11. A timeline for LHC operations until 2038. [3]

Chapter 3

EFT for Higgs boson characterization

Despite CMS and ATLAS have put limits on Higgs particle quantum numbers the possibility of anomalous contributions in HVV couplings is still open.

To search for new physics effects on the Higgs boson couplings is convenient to use the *EFT-framework* approach, where *EFT* stands for *Effective Field theory.* One can define a new structure for the interaction lagrangian of the Higgs boson with vector bosons and fermions, by introducing all possible Standard Model and beyond the Standard Model gauge and Lorentz invariant interaction terms to which we try to be sensitive.

Anomalous interactions of a spin-0 H boson with two spin-1 gauge bosons VV, such as WW, ZZ, $Z\gamma$, $\gamma\gamma$, and gg, are parametrized by a scattering amplitude that includes three tensor structures with expansion of coefficients up to (q^2/Λ^2)

$$A(HVV) \sim \left[a_1^{VV} + \frac{k_1^{VV}q_1^2 + k_2^{VV}q_2^2}{(\Lambda_1^{VV})^2}\right] m_{V1}^2 \epsilon_{V1}^* \epsilon_{V2}^* + a_2^{VV} f_{\mu\nu}^{*(1)} f^{*(2)\mu\nu} + a_3^{VV} f_{\mu\nu}^{*(1)} \tilde{f}^{*(2)\mu\nu}$$

(3.1)

where q_i, ϵ_{Vi}, and m_{V1} are the 4-momentum, polarization vector, and the mass of the vector boson, indexed by i = 1, 2. $f^{(i)\mu\nu} = \epsilon_{Vi}^\mu q_i^\nu - \epsilon_{Vi}^\nu q_i^\mu$ is the gauge boson's field strength tensor and $\tilde{f}_{\mu\nu}^i = (1/2)\epsilon_{\mu\nu\rho\sigma} f^{(i)\rho\sigma}$ is the dual field strength tensor defined using the Levi-Civita symbol in four dimension ($\epsilon_{\mu\nu\rho\sigma}$). a_i^{VV} are the coupling coefficients multiply the three tensor structures, and $k_i^{VV}/(\Lambda_1^{VV})^2$, which multiply the next term in the q^2 expansion. Λ_1 is the scale of beyond the SM (BSM) physics.

In Eq. 3.1, the only nonzero SM contributions at tree level are a_1^{WW} and a_1^{ZZ}. All other ZZ and WW couplings are considered anomalous contributions, which are either due to BSM physics or small contributions arising in the SM due to

loop effects and are not accessible with the current experimental precision. The parity-conserving interaction of a pseudoscalar (CP-odd state) corresponds to the a_3 terms, while the other terms describe the parity-conserving interaction of a scalar (CP-even state). Due to the fact that kinematics of the H boson production in WW fusion and in ZZ fusion are very similar we will set $a_i^{WW} = a_i^{ZZ}$ and $k_1^{WW}/\Lambda_1^{WW} = k_1^{ZZ}/\Lambda_1^{ZZ}$.

Among the anomalous contributions, considerations of symmetry and gauge invariance require $k_1^{ZZ} = k_2^{ZZ} = -exp(i\phi_{\Lambda_1}^{ZZ})$, $k_1^{\gamma\gamma} = k_2^{\gamma\gamma} = 0$, $k_1^{gg} = k_2^{gg} = 0$, $k_1^{Z\gamma} = 0$ and $k_2^{i\phi_{\Lambda_1}^{Z\gamma}}$, where $\phi_{\Lambda_1}^{Z\gamma}$ is the phase of the corresponding coupling.

There are three other anomalous couplings targeted in this analysis : $\Lambda_1^{ZZ} = \Lambda_1^{WW} = \Lambda_1$, $a_3^{ZZ} = a_3^{WW} = a_3$ and $a_2^{ZZ} = a_2^{WW} = a_2$. The purpose of this thesis is to constrain the three sets of couplings (a_2, a_3 and k/Λ_1), under the assumption that the couplings are constant and real, studying the VBF production in the channel in which $H \to \gamma\gamma$.

It is convenient to measure the effective cross section ratios f_{ai} rather than the anomalous couplings a_i themselves, as most uncertainties cancel in the ratio. The effective cross section f_{ai} and phase ϕ_{ai} are defined as follows:

$$f_{a3} = \frac{|a_3|^2\sigma_3}{|a_1|^2\sigma_1 + |a_2|^2\sigma_2 + |a_3|^2\sigma_3 + \sigma_{\Lambda_1}|k|^2/(\Lambda_1)^4}, \phi_{a3} = arg\left(\frac{a_3}{a_1}\right) \qquad (3.2)$$

$$f_{a2} = \frac{|a_2|^2\sigma_2}{|a_1|^2\sigma_1 + |a_2|^2\sigma_2 + |a_3|^2\sigma_3 + \sigma_{\Lambda_1}|k|^2/(\Lambda_1)^4}, \phi_{a3} = arg\left(\frac{a_2}{a_1}\right) \qquad (3.3)$$

$$f_{\Lambda_1} = \frac{\sigma_{\Lambda_1}|k|^2/(\Lambda_1)^4}{|a_1|^2\sigma_1 + |a_2|^2\sigma_2 + |a_3|^2\sigma_3 + \sigma_{\Lambda_1}|k|^2/(\Lambda_1)^4}, \phi_{\Lambda1} \qquad (3.4)$$

Given the measured values of the effective fractions, it is possible to extract the ratios of the coupling constants $|a_i|/|a_1|$ and the scale of BSM physics Λ_1 as follows:

$$\frac{|a_i|}{|a_1|} = \sqrt{\frac{f_{ai}}{f_{a1}}}\sqrt{\frac{\sigma_1}{\sigma_i}} \qquad (3.5)$$

$$\frac{\Lambda_1}{\sqrt{|k|}} = \frac{1}{\sqrt{|a_1|}}\left(\frac{f_{a1}}{f_{\Lambda_1}}\right)^{1/4}\left(\frac{\tilde{\sigma_{\Lambda_1}}}{\sigma_1}\right)^{1/4} \qquad (3.6)$$

Where the fraction $f_{a1} = 1 - f_{\Lambda_1} - f_{a2} - f_{a3}$ is the SM tree-level contribution.

3.1 Lagrangian for Spin-0 particle

The couplings in Eq. 3.1 are real and constant ($\phi_{ai} = 0$ or π), as we assume in this analysis, when the particles in the loops which give rise to them are heavy in comparison to the Higgs boson mass parameters. In this scenario,

the problem could be described in an effective field theory (EFT) which is a type of approximation that describe physical phenomena without knowing the full theory at infinitely high energies. Once fixed the energy scale E which we are interested in and the level of precision of the computations, all we need to know is the EFT up to energies a few orders of magnitude bigger than E. How many orders of magnitude depend on the desired level of precision. [16]

The scattering amplitude formulation (Eq. 3.1) is equivalent to an effective Lagrangian for the $HZZ, HWW, HZ\gamma$, and $H\gamma\gamma$ interactions where no new physics below Λ_1, which in this thesis project is fixed to 1 TeV is assumed. In the Lagrangian all the particles of the Standard Model, except for the Higgs itself, are considered. The interaction lagrangian of the spin 0 state with the vector bosons W, Z, γ and gluon g can be written as follows:

$$L(HVV) \sim a_1 \frac{m_Z^2}{2} H Z^\mu Z_\mu - \frac{k_1}{(\Lambda_1)^2} m_Z^2 H Z_\mu \Box Z^\mu - \frac{1}{2} a_2 H Z^{\mu\nu} \tilde{Z}_{\mu\nu} + \frac{1}{2} a_2 H Z^{\mu\nu} \tilde{Z}_{\mu\nu} +$$

$$+ a_1^{WW} m_W^2 H W^{+\mu} W_\mu^- - \frac{1}{(\Lambda_1^{WW})^2} m_W^2 H (k_1^{WW} W_\mu^- \Box W^{+\mu} + k_2^{WW} W_\mu^+ \Box W^{-\mu}) +$$

$$- a_2^{WW} m_W^2 H W^{+\mu\nu} W_{\mu\nu}^- - a_3^{WW} H W^{+\mu\nu} \tilde{W}_{\mu\nu}^- +$$

$$+ \frac{k_2^{Z\gamma}}{(\Lambda_1^{Z\gamma})^2} m_Z^2 H Z_\mu \partial_\nu F^{\mu\nu} - a_2^{Z\gamma} H F^{\mu\nu} Z_{\mu\nu} - a_3^{Z\gamma} H F^{\mu\nu} \tilde{Z}_{\mu\nu} - \frac{1}{2} a_2^{\gamma\gamma} H F^{\mu\nu} \tilde{F}_{\mu\nu}$$

$$(3.7)$$

where H is the real Higgs field, Z_μ is the Z field, Z_μ is the Z field, W_μ is the W field, F_μ is the γ field, $V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$ is the bosonic field strength, the dual field strengths are defined as $\tilde{V}_{\mu\nu} = 1/2\epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}$, and $\Box$ is the D'Alembert operator.

3.2 Current constraints on HVV

Similar studies in which HVV coupling has been analyzed in different channel have been published by CMS. In Ref. [17] the analysis of the VBF coupling is studied from events in which the Higgs boson decays into $\tau^+\tau^-$, using 35.9 fb^{-1} of the data collected in 2016. The $H \to \tau^+\tau^-$ channel has advantages over other H boson decay channels because of the relatively high significance of the signal events in the VBF channel, moreover the results were combined using the $H \to 4l$ decay [6] from Run 1 (from 2011 and 2012) and Run 2 (from 2015, 2016, and 2017) with data corresponding to integrated luminosities of 5.1, 19.7, and 80.2 fb^{-1} at center-of-mass energies 7, 8, and 13 TeV, respectively.

In Ref. [18] the data acquired by CMS experiment at the LHC and correspond to an integrated luminosity of 137 fb^{-1} for study the kinematics effects in the Higgs bosons' four-lepton final decay $H \to 4l$ (in particular $H \to 4\mu, 4e$ and $2e2\mu$), which is an interesting channel for our analysis cause

we are able to detect the coupling of interest both in production and in a decay sector. The $H \to 4l$ is a channel with a high reconstruction efficiency and in a few calculations we will notice that the channel considered in our analysis and the one from $H \to 4l$ have similar statistics (~ 300 events with $138 fb^{-1}$). In this study the measurements of $(f_{a2}, f_{a3}, f_{\Lambda_1})$ is made both living all the $f = (f_{a2}, f_{a3}, f_{\Lambda_1})$ as floats parameters and leaving only one of the three effective cross-section factor as a free parameter (fixing the other fractions to zero). Due to the fact that production and decay test different ranges of q_i^2, will be interesting to study the consistency of the constraints from the VBF and VH production and from $H \to 4l$ decay.

Results related to the f_{a3} parameter from previous searches are shown in Table 3.1. Notice that in the results a different definition of the effective cross section fraction is used:

$$f_{ai} = \frac{|a_i|^2 a_{ii}^{(2e2\mu)}}{\sum_j |a_j^{VV}|^2 a_{jj}^{(2e2\mu)}} sign\left(\frac{a_i^{VV}}{a_1}\right) \qquad (3.8)$$

where $a_{ii}^{(2e2\mu)}$ are parameters from the $H \to ZZ/Z\gamma^*/\gamma^*\gamma^* \to 2e2\mu$.

Channel	Approach	Observed (10^{-3})	Expected (10^{-3})
$H \to 4l$	$f_{a2} = f_{\Lambda_1} = 0$	$0.04^{+0.40}_{-0.03}[-0.55, 1.68]$	$0.00^{+0.81}_{-0.81}[-4.12, 4.12]$
$H \to 4l$	float f_{a2}, f_{Λ_1}	$0.05^{+0.56}_{-0.05}[-0.72, 2.18]$	$0.00^{+1.2}_{-1.2}[-5.7, 5.7]$
combination of $H \to \tau^+\tau^-$			
and $H \to 4l$	$f_{a2} = f_{\Lambda_1} = 0$	$0.00^{+0.27}_{-0.27}[-92, 14]$	$0.00^{+0.23}_{-0.23}[-1.2, 1.2]$

Table 3.1. Allowed 68% CL (central values with uncertainties) and 95% CL (in square brackets) intervals on f_{ai} parameters

Chapter 4

Selection of VBF-produced $H \to \gamma\gamma$

In Fig.1.3 and 1.4 the Feynman diagrams for the production and decay, respectively, for the channel of interest are shown. The final state will be characterized by two isolated photons and two jets, from the hadronization of the quarks. These jets are energetic, with high di-jets invariant mass distribution (M_{jj}) and with a large rapidity gap between the two. A sketch of the kinematics of the process is seen in Fig. 4.1, as the figure shows the two jets are characterized by a high η value and high $|\Delta\eta|$ between the two.

To study possible BSM contributions in the interaction vertex, simulated data samples that include anomalous coupling is needed. For this purpose, we use MC samples generated using an extension of the JHUGen generator, which computes the matrix element at LO, that enables simulations with the most general anomalous couplings [19]. To better model the signal we also use MC samples with Standard Model couplings of the H boson, generated using aMC@NLO and PYTHIA which include NLO QCD corrections. For this analysis, the full Run-2 data sample of *pp* collisions recorded by CMS during the years 2016,2017 and 2018 were studied, corresponding to $L_{int} = 139 fb^{-1}$.
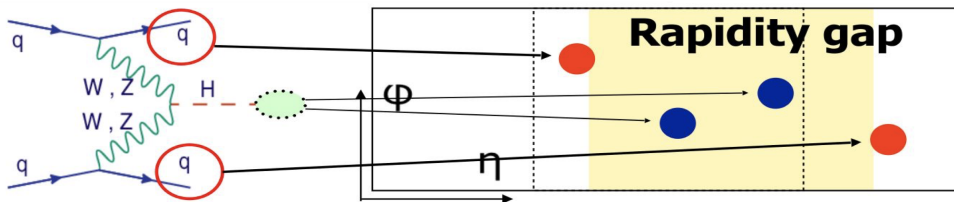

Figure 4.1. Topology of the final state of the channel of interest. The two jets (light dots) are very energetic and forward, with large $|\Delta\eta|$ between them. The two photons (dark dots) typically are more central.

4.1 Photon Selection

Di-photon triggers are used to select two well-identified isolated photons. Photons inside ECAL produce an electromagnetic shower which is a "spray" of photons and electrons. A photon is reconstructed with a clustering algorithm, which groups the shower constituents according to their kinematic properties. Because of the presence of upstream material and the magnetic field, the shower from an electron or a photon spreads into more than one crystal. The lateral size of the crystal used in ECAL is, approximately, 1 R_M, so the electromagnetic shower is almost completely contained by a matrix of 3x3 crystals. The electromagnetic showers are narrower than the hadronic ones and this information is exploited by the trigger to reduce the rate from jets faking photons. The trigger selection requires a loose identification using cluster shower shapes, a loose isolation requirement and a selection on the hadronic over the electromagnetic energy of the photon candidates.

On the events selected by the HLT a further offline pre-selection is applied, which uses more precise variables than those of the HLT. This step takes advantages of more precise reconstructed information of the event using a larger computing of time.

All the selections are applied both on data and simulated events to keep the same phase space. First, the photon with the largest transverse momentum p_T (leading) is required to have $p_T > 35$ GeV and the second largest (sub-leading) to have $p_T > 25$ GeV. Both have to satisfy an electron veto which removes the photon candidate if its supercluster is geometrically matched to an electron track with no missing hits in the innermost tracker layers.

Moreover the event selection requires $p_T^{\gamma 1}/m_{\gamma\gamma} > 1/3$ and $p_T^{\gamma 2}/m_{\gamma\gamma} > 1/4$. The use of p_T thresholds scaled by $m_{\gamma\gamma}$ prevents the distortion of the low end of the $m_{\gamma\gamma}$ spectrum.
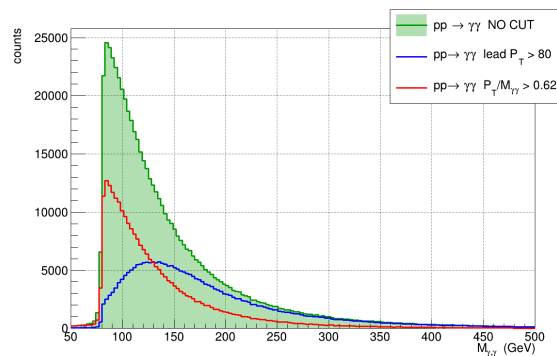

Figure 4.2. The plot shows the distortion of the low end of $m_{\gamma\gamma}$ spectrum if a cut on P_T is applied. In green the distribution of $m_{\gamma\gamma}$ if any cut is applied, in blue and red the distribution if, respectively, a cut on P_T and $P_T/M_{\gamma\gamma}$ is applied. The choice of the cut is made in such a way that the integral of the two distributions is the same

One of the most efficient ways to reject electron and photon backgrounds is the use of isolation energy sums. The isolation variables are obtained by summing the transverse momenta of charged hadrons, photons, and neutral hadrons inside an isolation cone of $\Delta R = 0.3$ with respect to the photon direction. The thresholds applied on the isolation quantities are made relative to the particle trasverse energy. Since one of the most effective isolation sum is the one computed with charge particles, photons that travel in regions within the acceptance of the tracker and to exclude the barrel-endcap transition region, both photons are required to satisfy $|\eta_{\gamma 1,2}| < 1.44$ or $1.57 < |\eta_{\gamma 1,2}| < 2.5$

Moreover the following selection criteria are applied:

	H/E	$\sigma_{i\eta i\eta}$	R_9
ECAL barrel; $R_9 > 0.80$	< 0.08	-	> 0.5
ECAL barrel; $R_9 \leq 0.80$	< 0.08	< 0.015	> 0.5
ECAL endcaps; $R_9 > 0.90$	< 0.08	-	> 0.8
ECAL bencapsarrel; $R_9 \leq 0.90$	< 0.08	< 0.03	> 0.8

Table 4.1. Preselection requirements, already applied in the MC generator.

In Tab 4.1 the following observables are used:

- Hadronic over electromagnetic energy ratio (H/E): The H/E ratio is defined as the ratio between the energy deposited in the HCAL in a cone of radius $\Delta R = 0.15$ around the supercluster direction and the energy of the photon or electron candidate.

- $\sigma_{i\eta i\eta}$: the energy-weighted standard deviation of single crystal η (in crystal index) within a 5×5 array of crystals centered on the crystal with maximum energy. This variable is particularly useful for distinguishing between photons and hadrons. Since the transverse hadronic shower development is sizably larger than the one of an electromagnetic shower, this variable is useful for discriminating hadrons from photons. The mathematical expression is:

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_i^{5x5} w_i(\eta_i - \bar{\eta}_{5x5})^2)}{\sum_i^{5x5} w_i}} \qquad (4.1)$$

Where, η_i is the pseudorapidity of the i-th crystal, $\bar{\eta}_{5x5}$ denotes the pseudorapidity mean position, and w_i is a weight factor accounting for ECAL noise of that channel. The size of the crystal in η in the barrel is 0.0175 and in the endcaps it varies from 0.0175 to 0.05. $\sigma_{i\eta i\eta}$ depends on the distance between two crystals in η thus, as shown in Fig. 4.3, its

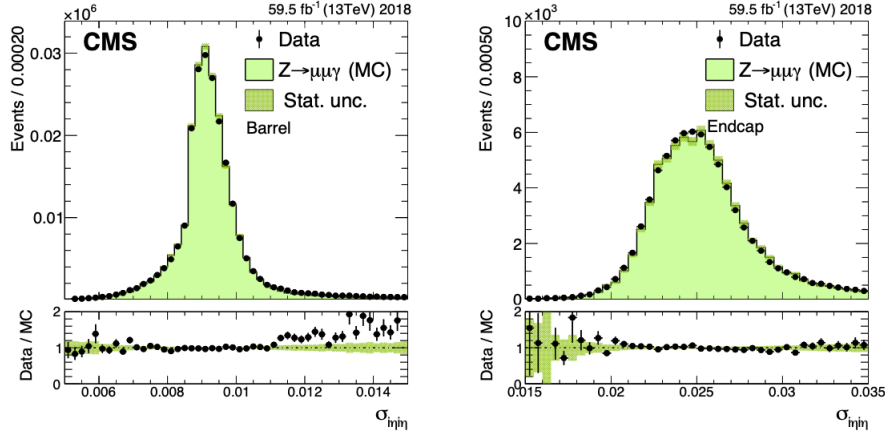distribution significantly change when evaluated in the barrel or in the endcaps.

Figure 4.3. Distribution of $\sigma_{i\eta i\eta}$ for photons from $Z \to \mu\mu\gamma$ events in the barrel (left) and in the endcaps (right). Photons are selected from 2018 data and simulation. $Z \to \mu\mu\gamma$ events, where the photons are produced by final-state radiation (FSR) provide a rather pure (99%) source of prompt photons. The study is from [20].

- R_9 : the energy sum of the 3×3 crystal array centered around the most energetic crystal in the supercluser divided by the energy of the supercluster. As the other cluster shapes variable is a useful variable to distinguish between photons and hadrons. Moreover it is employed to reject photons converting into electron pairs upstream of ECAL for which the energy resolution is worst and R_9 is smaller since a fraction of the showers' component is lost inside the tracker.

There are many other variables, besides those described above, useful for discriminating the photons from the hadronic background, since they are all more or less correlated: a Boosted Decision Tree (BDT) is used to summarize them. It is trained with shower shape observables, isolation variables based on the sums of the p_T around the candidate and the energy median density per unit area in the event (which makes the BDT independent of pileup). This BDT is then used to distinguish between prompt and background non-prompt photons. The algorithm outputs a continuous variable, which peaks around 0 for hadron and around 1 for photons, is defined as MVA^{ID}. Is required to be $MVA^{ID} > 0.5$ for both photons. Fig. 4.4 shows the distribution of the BDT output variable for events of resonant and non-resonant processes. This variable is particularly efficient to select "true" photons.
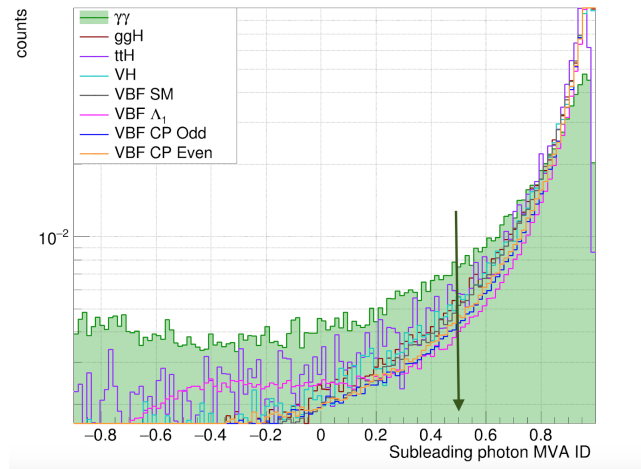
Figure 4.4. Distribution of the BDT output variable, on the subleading photon, for events of resonant and non-resonant processes normalized to total events. The arrow indicates the cut applied.

4.2 Jet Selection

We search for a final state with two energetic jets, with a large rapidity gap between the two. In the hadronic environment at the LHC, many quarks and gluons are produced, which, due to QCD confinement, create a collimated spray of hadrons, which appear as a cluster of energy deposited in a localised area of the detector, called a jet.

Hadronic jets are clustered from their reconstructed particles using the infrared and collinear safe anti-kT [21] algorithm with $\Delta R = 0.4$. Jet momentum is determined as the vectorial sum of all particle momenta in the jet.

The presence of multiple collisions in the same bunch-crossing (pileup) represents a challenge to the reconstruction and calibration procedure. The main techniques used at CMS to handle the pileup's contribution are the Charged Hadron Subtraction (CHS) algorithm that uses information from the tracker to remove the charged particles that are associated with a pileup vertex [22].

4.3 Search strategy

Once the selection criteria are applied, we have events with pairs of candidate photons and two forward jets. Since the irreducible background (mainly $pp \to \gamma\gamma$), produces a continuous spectrum in $M_{\gamma\gamma}$, the search strategy is that of a bump hunt, that is the presence of a narrow peak on a smooth background. The experimentally measured Higgs boson mass, with the Run 2 data from CMS experiment, is $M_H = 125.10 \pm 0.14$ GeV [23]. In the SM, the Higgs boson

width is very precisely predicted once the Higgs boson mass is known. For a
mass of 125 GeV, the Higgs boson has a very narrow width of 4.1 MeV. The
experimental mass resolution at CMS in $H \to \gamma\gamma$ is much larger than the
expected width of the SM Higgs boson. Therefore the width of the peak in the
invariant mass spectrum is dominated by the experimental resolution. The cur-
rent measurements of σ_{eff}, computed as the smallest interval containing 68.3%
of the photons invariant mass distribution, is 1.94 GeV, which is compatible
with the ECAL resolution.

The position measurement of photons im-
pacting on the calorimeter is used in deter-
mining their direction with respect to the
collision vertex which is located and, in case
of multiple vertices, identified with analysis
dependent algorithms exploiting track infor-
mation.

Since the Higgs boson decay of interest is
possible at NLO, through loops, this channel
is available for detection in association with
low p_T jets. From the tracks of these jets,
which come from the same production point
as the Higgs boson, the vertex is detected.

Figure 4.5. schematics rep-
resentation of how a wrong-
assigned vertex impact on the
photon direction.

From the formula 2.10 it is evident that the accuracy of the measurement
of the opening angle between the two decay photons from the Higgs boson
contributes to its reconstructed invariant mass resolution and an error on
this variable impacts directly on the invariant mass spectrum recostruction.
Therefore it is important to determine the primary vertex associated with the
Higgs boson. It results that the diphoton mass resolution is dominated by
the ECAL energy contribution only if the interaction point is known with a
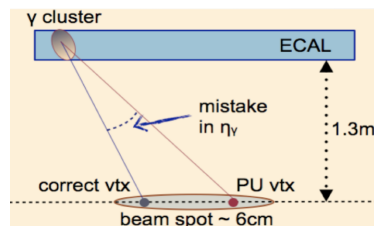precision of at least about 1 cm.

4.4 Signal and backgrounds

To increase the sensitivity of the analysis, selections are applied to the
kinematic variables in order to increase the signal to background ratio. In our
analysis signal events are the ones in which the Higgs boson is produced in
VBF (Fig. 1.2) with a coupling not foreseen by the standard model.

The dominant background to $H \to \gamma\gamma$ consists of the resonant background
coming from production channels of the Higgs boson different from the one
studied with the Feynman diagrams shown in Fig.4.6, and the non-resonant
backgrounds which appears as a smooth decreasing distribution in the $m_{\gamma\gamma}$'s
spectrum. Those events came from $pp \to \gamma + \gamma$ and $pp \to \gamma + jets$ (Fig. 4.8)
where the jets are misidentified as isolated photons. Unfortunately, a reliable
Monte Carlo sample is not available for the $pp \to \gamma + jets$ process so the
contribution of this process hat to be estimate directly from data. For that

reason diphoton pairs with a wide invariant mass range with respect the signal peak width, $100 < m_{\gamma\gamma} < 180 GeV$ are selected, i.e. keeping enough sidebands around the expected mass peak.

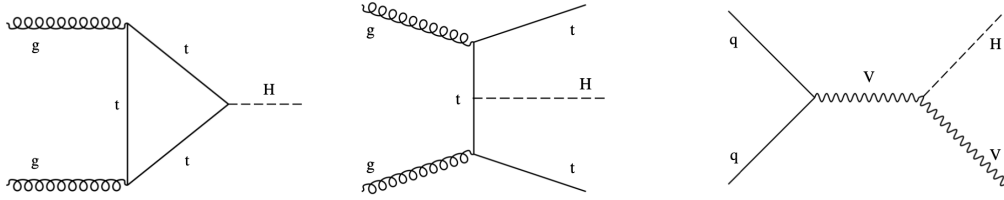**Figure 4.6.** Feynman diagram for Higgs productin modes different from VBF. From left to right: ggH ($\sigma_{ggH} \sim 48.6pb$) , ttH ($\sigma_{ttH} \sim 2pb$) and VH ($\sigma_{ttH} \sim 0.5pb$)
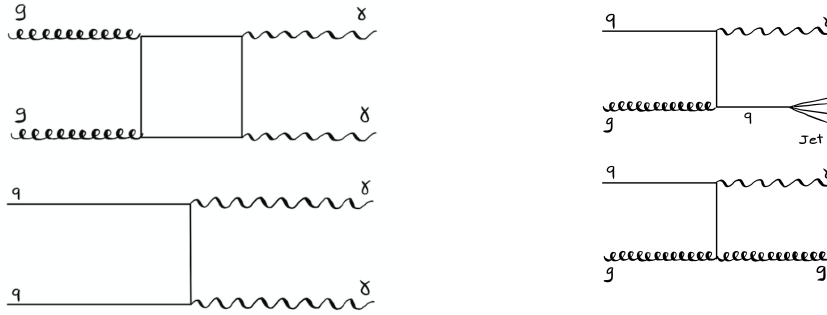

Figure 4.7. On the left the production of two photons in proton-proton collisions in a Born process (top) and in a box process (botton). On the right the Feynman diagram for $pp \rightarrow \gamma jets$.

In order to select VBF-like events futher consideration are made. The final state jets of the VBF production come from the hadronization of the partons originating from the high center of mass energy proton proton collisions, that scatter with high longitudinal momentum and that continue almost in the original direction of the beam.

In the ggH as well as ttH,VH case, the final state jets are the ones from the initial state radiation of the incoming partons. Therefore they have low p_T , and their spectrum is a decreasing exponential distribution. Consequently a constraint on the M_{jj} variable will be very effective in background rejection.

The following cuts are applied to improve the signal to background ratio when selecting VBF-like events:

- $(p_T)_{j1} > 40\,\text{GeV}, (p_T)_{j2} > 30\,\text{GeV}$;

- $|\Delta\eta_{jj}| > 0.0, |\eta_{j1}| < 4.7, |\eta_{j2}| < 4.7$;

- $min(\Delta R_{jet-\gamma}) < 0.4$;

- $M_{jj} > 250\,\text{GeV}$;

The request that $|\eta_{j1,j2}| < 4.7$ is necessary to avoid region with larger instrumental noise in the forward hadronic calorimeters. With the cut $min(\Delta R_{jet-\gamma}) < 0.4$ we select jets far from the photons to avoid photons coming from jets.
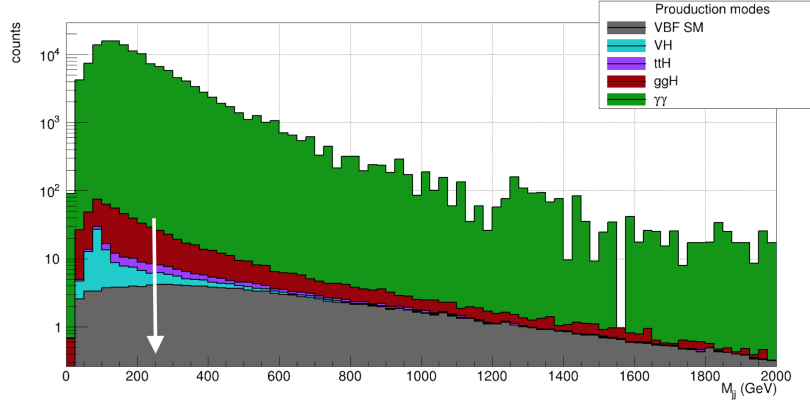

Figure 4.8. Stack distibution of M_{jj} variable. In the plot all the other selections described in the text are applied but the one on the M_{jj} variable which is represent by the white arrow. The data are normalize at the number of events expected for $137 fb^{-1}$

In Figure 4.9 the leading photon $p_T/m_{\gamma\gamma}$ data's distribution is represented and compared to the one of simulated data. It can be noticed that we have a shift between the two distributions due to NNN..LO terms and due to the absence of MC sample $\gamma + jets$ that should populate the region $0.2 < \frac{p_T}{m_{\gamma\gamma}} < 0.5$ where the excess of data is evident. Table 4.2 list the expected efficiency of each selection, with respect to the previous one, for each MC sample considered.

The kinematic distributions of particles produced in the H boson decay or in association with the H boson production are sensitive to the quantum numbers and anomalous couplings of the Higgs bosons. In particular, differences could be found in angular variables like $\Delta\eta$, $\Delta\phi$ between photons and jets due to different spin-parity hypothesis between SM and anomalous samples. Moreover, anomalous coupling would result in higher transverse momentum of the H boson as Fig. 4.10 shows.
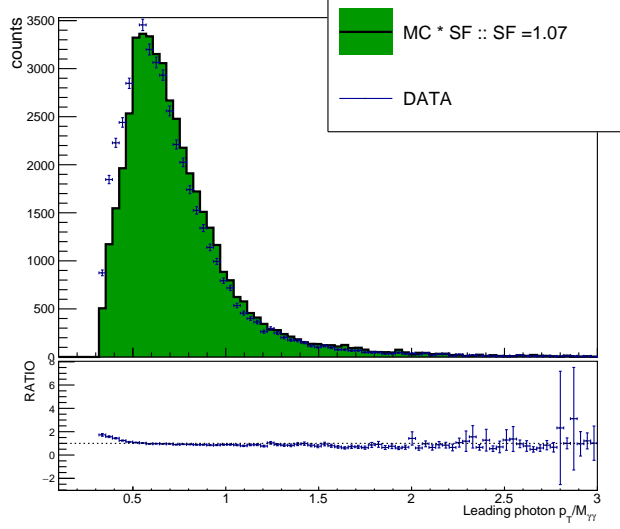
Figure 4.9. Leading photon p_T/M's specrum for data and MC. The distribution of the MC has been renormalized in order to have the same entries as the distribution of the data .On the lower box the ratio between the two.

	$\gamma\gamma$	ggH	ttH	VH	VBF SM	$VBF\Lambda_1$	VBF CP Odd	VBF CP Even
No cut	709313	1500	73	172	325	623	607	584
η_γ	-1%	-1%	-1%	-1%	-1%	-1%	-1%	-1%
$m_{\gamma\gamma}$	-1%	-1%	-1%	-1%	-1%	-1%	-1%	-1%
$\gamma p_t/M$	-24 %	-10 %	-11 %	-10 %	-11 %	-5 %	-7 %	-8%
$|\Delta\eta|_{dijets}$	-24 %	-10 %	-11 %	-10 %	-11 %	-5 %	-7 %	-8 %
$p_t jet$	-53 %	-57 %	-12 %	-36 %	-32 %	-14 %	-11 %	-12 %
M_{jj}	-85 %	-86 %	-62 %	-88 %	-46 %	-22 %	-25 %	-25%
$IDMVA_\gamma$	-93 %	-90 %	-78 %	-92 %	-60 %	-45 %	-45 %	-46%
All Cuts	46221	154	16	13	129	342	335	318

Table 4.2. In the first line the number of MC entries for each process with at least two jets. The following lines represent the fraction of events rejected by each cut with respect to the previous one. In the last line the number of expected events in 139 fb^{-1} after applying all the cuts.

4.5 Event categorization

To increase the sensitivity to the SM and BSM signals, I choose to divide the data set into categories with different signal-over-background ratio. Then the categories will be statistically recombined. Given that the kinematic distributions of the particles produced in association with the Higgs boson or from its decay are sensitive to Higgs boson quantum numbers, they could be used to distinguish signal from background processes. Multiple kinematic
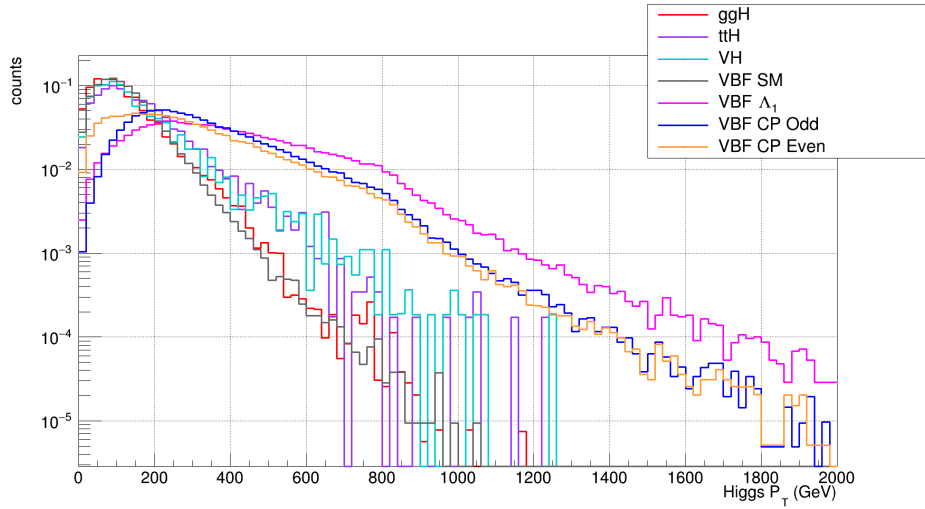
Figure 4.10. Trasverse Higgs momentum distribution, normalized to 1, of different production modes

variables can be constructed in this final state, with a different level of cross-correlation among them, so an analyis based on selection on all of them such that includes all of them would be very impractical. Limiting the study using only the more informative variables could drastically worsen the results as it would cause a loss of information. Hence, I choosed to combine them into new variables that are optimal for the problem of interest. Two possible strategies that I studied are machine learning methods and the Matrix Element Approach. Both will be used in the subsequent analysis.

4.5.1 Machine Learning Approach: Regression

Machine learning algorithms could be trained on Monte Carlo data sets to learn how to assign a probability that an event belongs to a certain class from input variables. The class label is the discrete attribute whose value you want to predict based on the values of other attributes. The number and definition of the classes can be chosen in order to maximize the discriminative power of the output variables. In particular, in this study, I consider 3 classes. The "multiclass" approach (i.e. more than 2 classes) allows to optimize the discrimination not only of the signal with respect only one background, but of each specific production process of the Higgs boson. The output of the algorithm is a real number between 0 and 1 which can be interpreted as the probability that an event belongs to a certain class.

I choose to define the classes as follows:

- BKG : not VBF Higgs boson: ggH, $pp \rightarrow \gamma\gamma$,

- SM VBF : VBF with a SM coupling type of event

- BSM VBF : VBF with a BSM coupling type of event

For each Monte Carlo event where the true nature of each event, is known, half of the set will be used to train the algorithm and the remaining part to test it, using a statistically independent set of events. The events used in the training and test sets are weighted with a weight proportional to the respective cross-section.

For the training I used the Toolkit for Multivariate Analysis (TMVA) [24] which easily handles ROOT trees thanks to object-oriented implementations in C++/ROOT.

Different types of Regression algorithms exist and, with several MVA methods, whose performance depend on the problem they must be tested on the problem of interest. In the following analysis, I tested two algorithms: a boosted decision tree (BDT) and a Deep Neural Network (DNN).

As a benchmark to define the the best algorithm, the Area Under the Curve (AUC) is used, i.e. the area subtended by the Receiver Operating Characteristic Curve (ROC). The ROC curve plots the signal efficiency vs. the background rejection at different classification thresholds. In other words, the signal efficiency is the fraction of the desired events which are correctly selected (true positive rate TPR) while the background rejection is the fraction of background events that are excluded from it (true negative rate TNR):

$$TPR = \frac{TP}{TP + FN} \tag{4.2}$$

$$TNR = \frac{TF}{TF + FP} \tag{4.3}$$

A random classifier on this plane would be represented by a diagonal while the area under the ROC curve of an optimal classifier, being able to perfectly distinguish signal on background, is equal to one.

In order to choose among the two possible MVA approaches available, BDT and DNN, the representation of the ROC curve the "One vs Rest" method is chosen. As the name suggests, it is a method used to evaluate multiclass models by comparing each class against all the others at the same time. In this scenario we take one class and consider it as our "signal" class, while all the others are considered as the "background" class. By doing this, we reduce the multiclass classification output into a binary classification one, making it possible possible to compute the ROC curve as explained above. In Fig. 4.11, for example, on the horizontal axis the efficiency referred to one class is shown, while on the vertical axis the rejection of other classes is shown. The use of this method implies a number of ROC curves equal to the number of classes used in the algorithm: in the considered case three. Fig. 4.11 shows the ROCs for the BDT and DNN approaches.

It is evident that the DNN (even if not optimized) has better performances than the BDT. Therefore, we choose to optimize the DNN to improve the
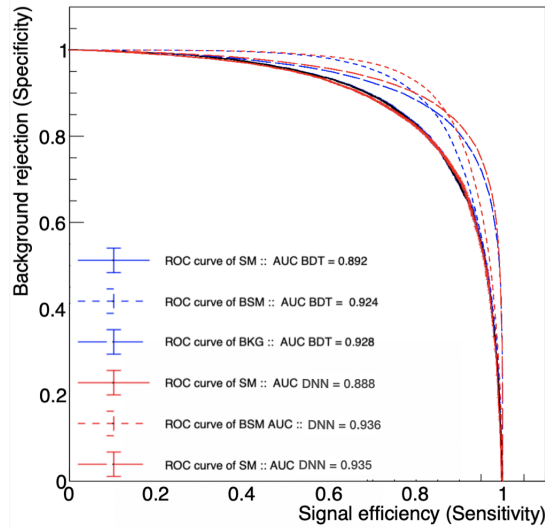
algorithm's performance.

Figure 4.11. ROC curve computed with the "One vs Rest" method for each class. The blue curve are computed using a BDT while the red one using a DNN.

4.5.2 DNN training

A neural network is a function $f : R_n \rightarrow R_m$ from the space of the features to the space of the labels. It is structured in layers and consists of an input layer of size n, an output layer of size m and several hidden layers of various sizes. Each layer element is called neuron (or node) and is the "place" where computation happens. Each neuron is connected with all previous layer neurons and with all those of the next layer. These connections produce a certain response at a given set of input signals and are mathematically described by weights, which characterize the strength of a node's influence on another, and bias which helps to shift the output values by adding constants. A function, called *activation function*, is then applied to each neuron and provides a nonlinear response to the whole algorithm. Among the most used activation functions there's the tangens hyperbolicus (Tanh) and the Rectified Linear Unit function (Relu) thet can be represented as follows:

$$f_{Relu}(x) = max(0, x) \tag{4.4}$$

The final goal is to teach the model the ability to correctly predict the expected values on a set of instances not present in the training set (test set). This is done by minimizing the *loss function*, which quantifies the difference between the expected result and the result produced by the machine learning model.

The loss function is periodically evaluated so it's possible to derive the gradients which are used to update the weights and biases of each node and

gradually adjust them. The *learning rate* is a hyperparameter (parameter whose value is used to control the learning process) that controls the extent of this adjustment [25]. A graphical representation of a deep neural network is show in Fig. 4.12.
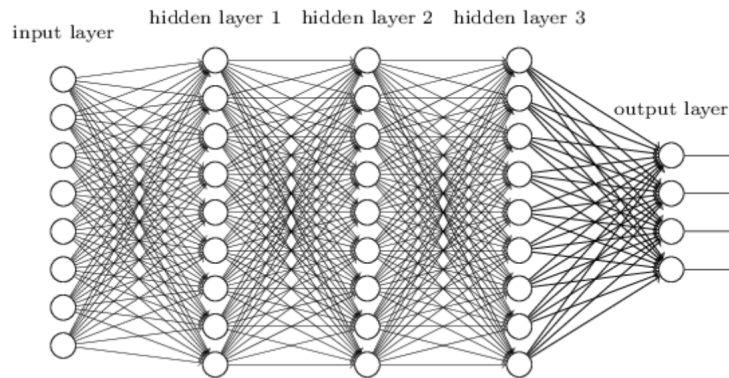

Figure 4.12. Graphical representation of a deep neural network which has three hidden layers apart from the input and output layers. The circles represent the neurons and the lines the connections between them.

The DNN was trained to discriminate between the 3 classes (SM VBF, BSM VBF and BKG), the two highest AUC mean values are chosen as a benchmark to optimize it.

Different layouts were tested (number of layers and nodes, and different activation functions) for different learning rates. In particular, the following hyperparameters were tested:

- Learning Rate : 10^{-2}, 10^{-3}, 10^{-4}

- Activation Function : Relu, Tanh

- Number of Layer: 1,2,3

- Number of Nodes for each Layer: 10,50,100

Figure 4.13 shows the optimal configuration (the one that maximize the average of the two major AUCs). In conclusion, the following hyperparameter are set for the training:

- Learning Rate : 10^{-3}

- Activation Function : Relu

- Number of Layer: 1

- Number of Nodes for each Layer: 100

Note that neural networks are stochastic before being trained. The stochastic process refers to a variable process in which the result implies randomness and presents some uncertainty. This means that, repeating the training with the same parameters (e.g. random initialization, data shuffling, hyperparameters) it is possible to find differences in the output.

To verify the extent of the phenomenon, I chosed a set of hyperparameters and I trained the DNN 10 times with the chosen setup. Figure 4.13 shows that the average of the two major AUCs is different at each iteration and the previous fluctuations (due to a different hyperparameters' choice) are compatible with the fluctuation due to the use of a non-deterministic algorithm. It is concluded that the results are stable with regard to the layout's variations, therefore the simpler structure of the model, which gives good results, is chosen.
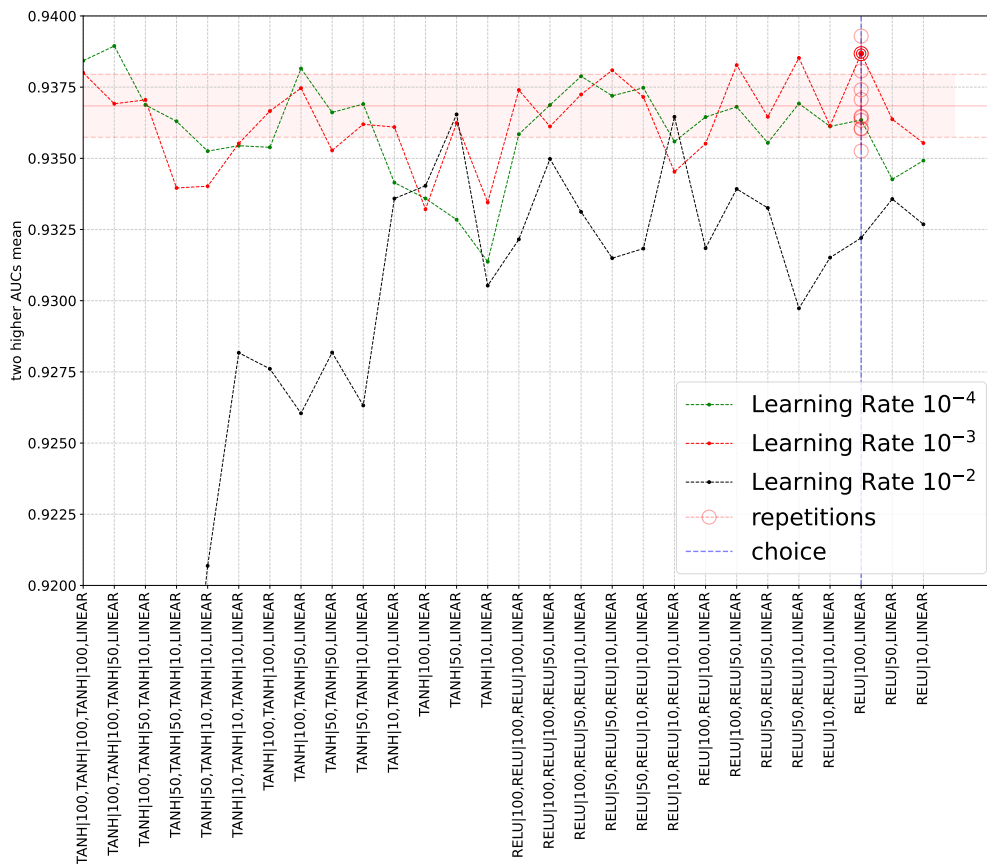
Figure 4.13. Two highest AUCs mean value for different layout. The empty dots represent the results of a fixed layout setting and the light red line/bar is the mean/standard deviation of these results.

In principle, with more variables, we have more information about the event and the DNN should gain greater separating power. However, reducing the number of variables used when training could benefit the model by removing information redundancy, reducing training time, and avoiding overfitting. Over-

fitting occurs when the model starts to describe the statistical fluctuations in the training sets and loses generality, resulting in a seemingly high performance that does not hold if data sets are changed [26].

This motivates us to rank the variables i.e., order the features features based on their importance for outcome prediction, and select the most important ones of them in the training. The ranking of the variables, is calculated as the sum of the weights-squared of the connections between the variable's neuron in the input layer and the first hidden layer. This is shown in Fig. 4.15

By removing the variables linearly correlated more than 75% and with a ranking lower than 10^{-4}, only 14 variables remain. I performed the training with the N best-ranked variables, with N=[1,...14]. The results, shown in Fig. 4.15, highlight that the gain deriving from the use of more than 8 variables is negligible and that the trend reaches a plateau. Taking into account possible fluctuations in the results, we choose to use 10 input variables in the final DNN training.
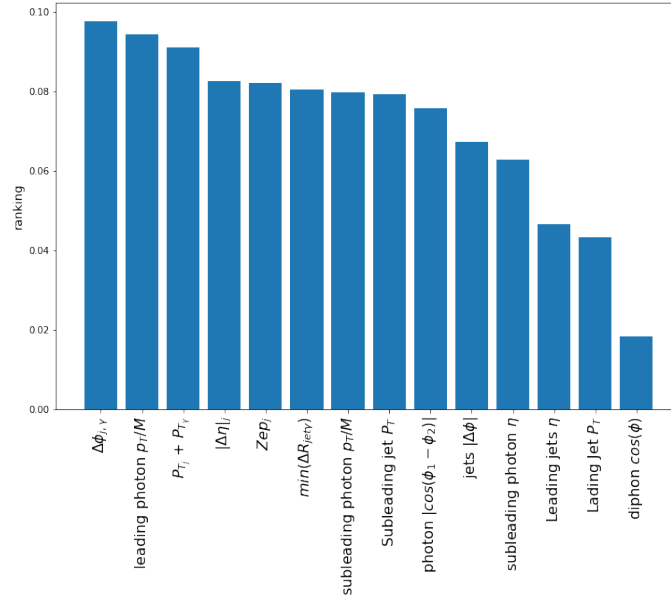

Figure 4.14. Ranking score for the best 14 variables used in the study. The Zeppenfeld variable was first proposed in ref. [27], it is defined for VBF as: $\eta^* = \eta_{j3} - <\eta_{j1j2}>$, where j_3 is a minijet emitted from one of the tagged jets.
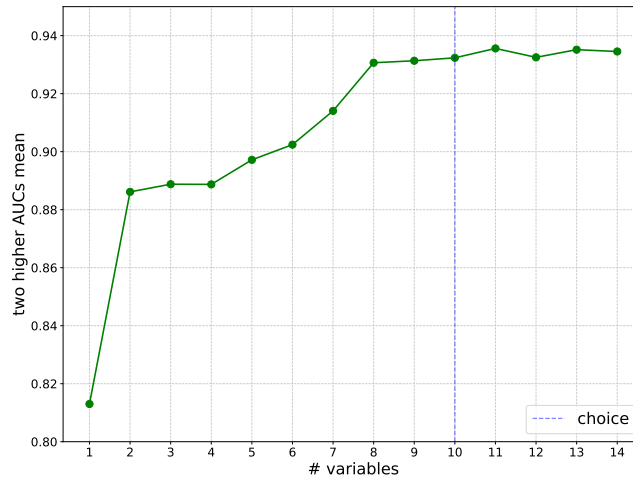
Figure 4.15. Model performances using a different number of variables. Note that the plateau of the plot is formed when using more than eight training variables.

As a result for each input event, the algorithm returns the probability of each event belonging to a class, thus producing three output parameters. The sum of the probabilities assigned to each event is normalized to one, meaning only two of the three output variables are independent. Fig 4.16, 4.17, 4.18 show histograms normalized to their areas of the three output variables for different MC samples. The plots also shows the distributions relating for the VH and ttH samples, not used in the training phase. These samples bring only a small contribution to the resonant background so not having used them in the training phase will not drastically affect the analysis.
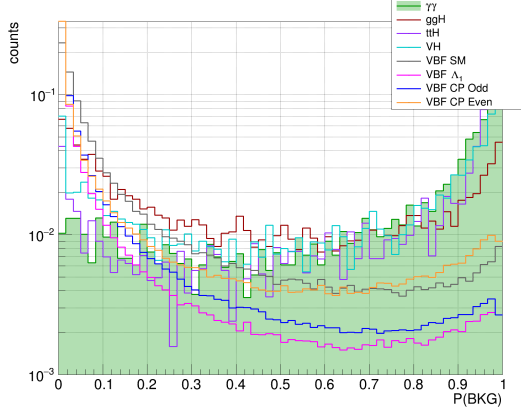
Figure 4.16. probability distribution belonging to the BKG class for different MC samples
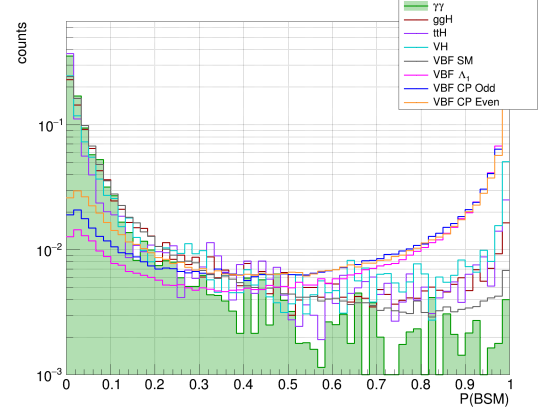

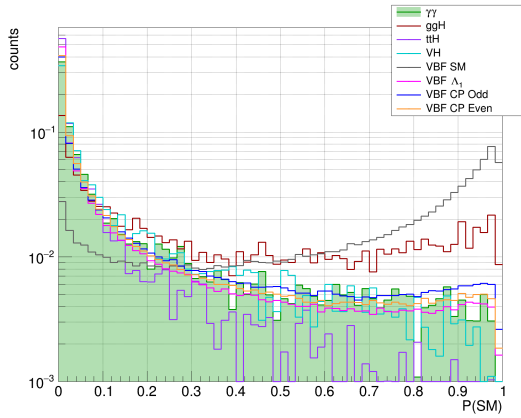**Figure 4.17.** probability distribution belonging to the BSM class for different MC samples

Figure 4.18. probability distribution belonging to the SM class for different MC samples

AUC	Value
AUC of P(BKG)	0.929
AUC of P(VBF BSM)	0.935
AUC of P(VBF SM)	0.878

Table 4.3. AUC values for the optimal DNN training.

4.5.3 MELA variables

The MELA (Matrix Element Likelihood Approach) is a calculation of probability distributions based on the first-principle matrix element or Lorentz-invariant scattering amplitude [19; 28].
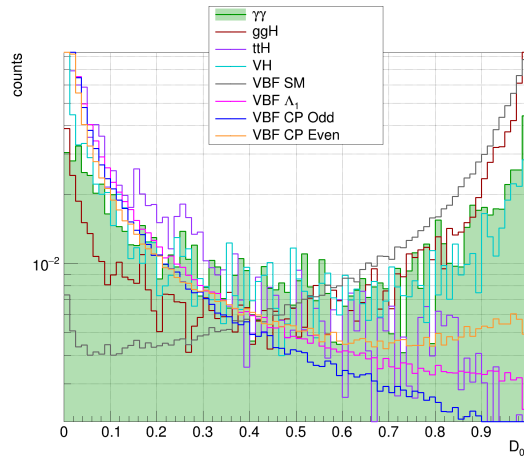
Figure 4.19. D_0^- distribution for MC samples for different H boson production and couplings and non-resonant background ($\gamma\gamma : pp \rightarrow \gamma\gamma$).

Given a set of observables $\vec{\Omega}$ and the matrix element it is possible to compute the probability P that a $\vec{\Omega}$ matches a certain hypothesis. In this way, all relevant kinematic information is summarized in a minimal set of discriminants D computed from ratios of probabilities P, as

$$D_{alt} = \frac{P_{sig}(\vec{\Omega})}{P_{sig}(\vec{\Omega}) + P_{alt}(\vec{\Omega})} \qquad (4.5)$$

where "sig" stands for the SM signal; "alt" denotes an alternative hypothesis, which could be background, an alternative H boson production mechanism, or, as in our case, an alternative H boson coupling model [29].

In the following analysis, we will use the variable D_0^-. This variable is built to be a discriminator between VBF Higgs boson with SM couplings and VBF Higgs boson with CP-odd couplings but works well as a discriminator for all BSM hypotheses as shown in Fig 4.19.

4.5.4 Optimizing the categorization

The two best performing output variables (P(BSM) and P(BKG)) deriving from the DNN training and the discriminatory D_0^- are exploited to categorize the data in the 3-dimensional space. Each of these variables has a different discriminating power, hence the choice of variables and the size of the space. The expected number of events for each process belonging to each category will be different as well as the signal-to-background ratio. The categorization using these variables has two targets: increase the sensitivity to the signal, by dividing events in classes with different S/B, and to provide sensitivity to H coupling differences. The categorization, in fact, uses the kinematic information contained in the variables constructed with multivariate analysis, providing the sensitivity to the BSM processes. Without categorization, the

signal-to-background ratio would be much lower while by separating the events into categories with different purities, I can extract the signal independently with a fit in $m_{\gamma\gamma}$ (for this reason the diphoton mass was not used to "build" the categorization variables so is independent from them), and then recombine them. With this method the high signal-to-background ratio is preserved without cutting on the variables (which would inevitably lead to a loss of a signal events.

For this purpose, a boundary is chosen on each dimension that divides the sample into two bins (so the sample is divided into 2^3 categories). I have chosen not to use more than two bins in each dimension given that the number of VBF events expected is O(100), so a different choice would most likely lead to empty categories. To choose the best boundaries I scan different threshold and select the set that maximizes the statistical significance (σ). A graphical representation of the method is shown in the figure 4.20. The significance is calculated from the number of entries in each category (i = 1..8) for each process:

$$\sigma = \sqrt{\sum_{i=1}^{8} \left(\frac{s_i}{\sqrt{s_i + b_i}} \right)^2} \qquad (4.6)$$

where:

$$s_i = f_{a3} \times s_i^{a3} \qquad (4.7)$$

$$b_i = b_i^{ggH} + b_i^{ttH} + b_i^{VH} + b_i^{\gamma\gamma} + b_i^{\gamma+jets} + (1 - f_{a3}) \times b_i^{SMVBF} \qquad (4.8)$$

Here f_{a3} is the effective cross section fraction of VBF events in which Higgs boson decays into two photons with coupling a_3. In the optimization of the categories, the value of f_{a3} is fixed at 0.07 which is the upper limit, at 95% confidence level, deriving previous analysis described in [29]. The term s_i is the number of signal events expected in 139 fb^{-1} in the case of VBF Higgs boson signal, with a relative fraction of f_{a3} anomalous coupling and (1-f_{a3}) SM coupling, while b_i indicates the number of background events in the i-th category. The number of background/signal events is computed in the *signal region* i.e. phase space defined by the preselection cuts and in a mass region around the expected Higgs boson peak (120 < $m_{\gamma\gamma}$ < 130 GeV). Beyond the 5σ, where $\sigma \sim 1$ GeV is the approximate $m_{\gamma\gamma}$ resolution, (~ 5 GeV) from the mass of the Higgs boson I expect only background events, which will not affect the final results deriving from the fit, but are kept in the pre-selection because they are useful for the modeling of the non-resonant background. Therefore it is chosen, only for this optimization, to ignore the events in the $m_{\gamma\gamma}$ sidebands.
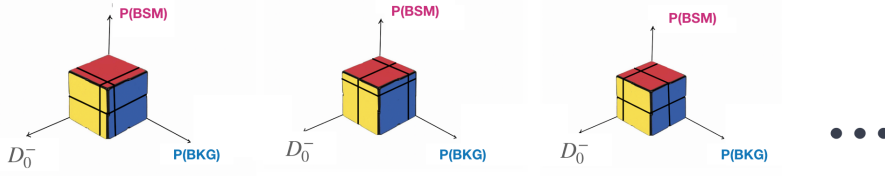
Figure 4.20. Graphic representation of the method used. Significance is computed for different configurations (different threshold) to select the one that maximizes it.

In the formula 4.8 the background $\gamma + jets$ also contributes, however, there is no MC for this sample, because of the difficulty to model this process well with simulation. So, the ratio between data (N_{data}) and MC (N_{MC}) in the right and left sidebands, called *control region* (CR). is used to estimate the contribution of the total non-resonant background in the signal region (SR):

$$b_i^{non-res}(SR) = \frac{N_{MC_i}^{non-res}(CR)}{N_{data_i}(CR)} \times N_{MC_i}^{non-res}(SR) \qquad (4.9)$$

The maximum significance (assuming $f_{a3} = 0.7$) found through the scan of the 3D space of the thresholds in the 3 variables space is $\sigma = 1.74$ and the cuts that maximize significance are the following:

- DNN P(BSM) = 0.97

- DNN P(BKG) = 0.05

- $D_0^- = 0.6$

To show in one dimension the optimization procedure, in Fig. 4.21 each curve is calculated using the best value on the other two variables and performing a scan over the third one. As the plot shows, the significance is particularly sensitive to cut variations along the variable P(BSM) which is why a denser scan was performed along this dimension.
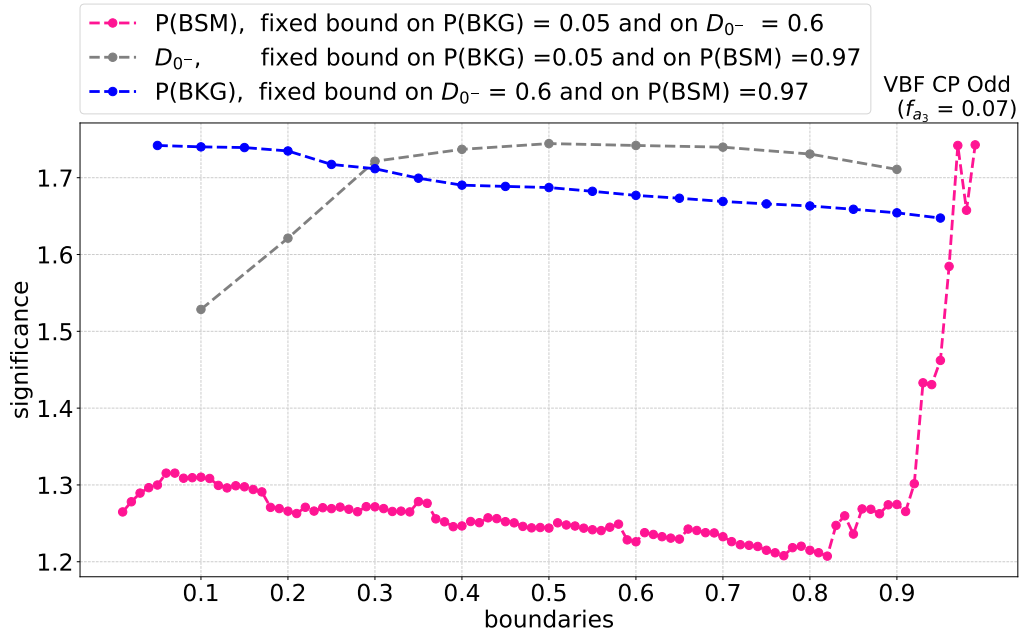
Figure 4.21. BSM CP odd's significance variation with regard to different categorization choices.

The process described above was also carried out for the other BSM couplings taken into consideration in the analysis (Fig.4.22, Fig.4.23). The trends in significance as the choice of cuts varies are similar in all the plots which indicates that the optimization does not depend on the coupling scenarios.

The table 4.4 indicates what is the contribution in each category for different types of events for the final choice of cuts. At the end of the category optimization, 3 of the 8 initial categories have zero or very small number of expected events and are not considered for the rest of the analysis.

It can be noted that even though different couplings have different cross sections (σ_{ai}), what is used in the analysis are the differences between VBF Higgs boson with SM couplings distributions and VBF Higgs boson with anomalous couplings distributions. In this way, the dependence of the analysis on the generator σ_{ai} estimation decrease.
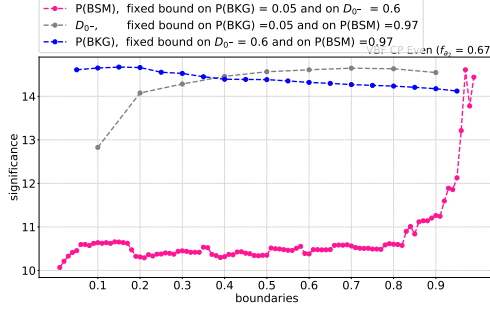
Figure 4.22. BSM CP even significance variation to different cuts choice. The effective cross section value is fixed to the upper limits found in Ref. [29]
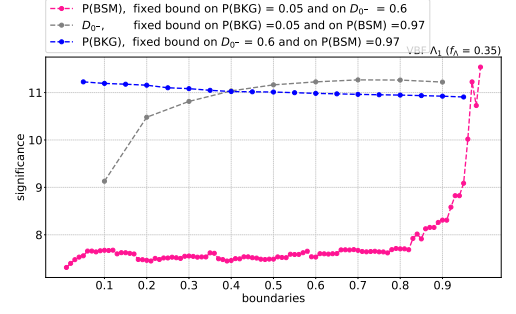

Figure 4.23. BSM Λ_1 significance variation to different cuts choice. The effective cross section value is fixed to the upper limits found in Ref. [29]

	data CR	BKG CR	$\gamma\gamma$	ggH	VH	ttH	VBF SM	VBF CP Odd	VBF CP Even	VBF Λ_1
$P(BKG) < 0.05$, $D_0^- < 0.6$, $P(BSM) < 0.97$	66	64	81	7	1	1	15	98	64	82
$P(BKG) < 0.05$, $D_0^- < 0.6$, $P(BSM) > 0.97$	27	23	19	3	1	1	2	126	104	159
$P(BKG) < 0.05$, $D_0^- > 0.6$, $P(BSM) < 0.97$	162	144	124	13	0	0	45	11	19	19
$P(BKG) < 0.05$, $D_0^- > 0.6$, $P(BSM) > 0.97$	0	0	0	0	0	0	0	0	0	1
$P(BKG) > 0.05$, $D_0^- < 0.6$, $P(BSM) < 0.97$	3367	3531	3450	47	7	13	17	84	101	58
$P(BKG) > 0.05$, $D_0^- < 0.6$, $P(BSM) > 0.97$	0	0	0	0	0	0	0	0	0	0
$P(BKG) > 0.05$, $D_0^- > 0.6$, $P(BSM) < 0.97$	3174	2709	2720	69	4	2	53	15	30	17
$P(BKG) > 0.05$, $D_0^- > 0.6$, $P(BSM) > 0.97$	0	0	0	0	0	0	0	0	0	0

Table 4.4. Contribution in each category for different process for 139 fb^{-1} in the signal region. In the table f_i is fixed to 1.

Chapter 5

Maximum Likelihood fit

The discovery of the Higgs boson in the decay channel in two photons yielded a peak in the invariant mass spectrum of the selected pair of photons. In particle physics a discovery is defined as an observation that is incompatible with a positive statistical fluctuation with respect to the expected background alone. Given that most of the processes that are being searched have low statistics, a Poisson distribution of the statistical fluctuations in the number of events is typically assumed.

Constructing a histogram of M bins with N events the probability distribution of signal plus background (s+b) and only background (b) are defined as:

$$P(N|b+s) = \prod_{i=1}^{M} \frac{e^{-(b_i+s_i)}(b_i+s_i)^{n_i}}{n_i!} \tag{5.1}$$

$$P(N|b) = \prod_{i=1}^{M} \frac{e^{-(b_i)}(b_i)^{n_i}}{n_i!} \tag{5.2}$$

where s_i (b_i) is the number of expected signal (background) events in the i-th bin and n_i is the number of events observed in the same bin.

Conventionally a discovery (exclusion) is announced if $P(b) \leq 2.86 \times 10^{-7}$ ($P(s+b) \leq 5 \times 10^{-2}$).

A powerful test between two hypotheses is the *Likelihood approach*. The Likelihood function is the product of the probability density function (PDF), calculated from the MC or using data from the side-bands, times the poisson probability for observing N events in each bin, assuming a negligible correlation between the bin contents.

If f^s and f^b are the PDFs of the signal and the background, respectively, for the event j_c and $c = 1..C$ is the index which refers to different channel or (in our case) category then the Likelihood of the signal plus background (L_s) and of the background only (L_b) are defined as:

$$L_s = \prod_{c=1}^{C} P_{poisson} (N_c | b_c + s_c) \times \prod_{j_c=1}^{N_C} f_c^{b+s}(m_{\gamma\gamma j_c}) =$$

$$= \prod_{c=1}^{C} \frac{e^{-(s_c+b_c)}(s_c + b_c)^{N_c}}{N_c!} \times \prod_{j_c=1}^{N_C} \left[\frac{s_c f_c^s(m_{\gamma\gamma j_c}) + b_c f_c^b(m_{\gamma\gamma j_c})}{s_c + b_c} \right]$$

$$(5.3)$$

$$L_b = \prod_{c=1}^{C} P_{poisson} (N_c | b_c) \times \prod_{j_c=1}^{N_C} f_c^{b}(m_{\gamma\gamma j_c}) =$$

$$= \prod_{c=1}^{C} \frac{e^{-b_c} b_c^{N_c}}{N_c!} \times \prod_{j_c=1}^{N_C} \left[f_c^b(m_{\gamma\gamma j_c}) \right]$$

$$(5.4)$$

The likelihood is expected to be larger when the correct pdf is used. Nevertheless, for convenience, it is preferred to minimize $-2 \times ln(L)$ than maximize (L).

To extract the results, simultaneous binned fits are performed on the $m_{\gamma\gamma}$ distributions of all analysis categories, in the mass range $100 < m_{\gamma\gamma} < 180$ GeV. Each category has a different peak resolution and a different background's shape, so different probability density functions (PDF) will be used to model both background and signal in each category. Once the PDF of the fit components have been modeled, the minimization is for few number of parameter of interest.

5.1 Fit for anomalous couplings contributions

To measure the presence of the Higgs VBF signal, the product $\sigma_{VBF} \times BR(H \to \gamma\gamma) \times \mu$ is defined as POI. Where μ is defined as the ratio of the measured to SM expected. Then, to consider possible contributions not foreseen by the standard model, for each anomalous coupling the effective cross section fractions (f_{ai}) is added to the model, which I leave as floating parameters as well as μ. Therefore, the case is studied in which only one of the three fractions (components of $\vec{f}_{ai}$) is different from 0, while the other two components will be fixed at 0 during the fit. For example, $f_{ai} = 0$ indicates a pure SM Higgs boson, $f_{ai} = 1$ gives a pure BSM particle and $f_{ai} = 0.5$ means that the two couplings contribute equally to the $H_{VBF} \to \gamma\gamma$ process. In particular, f_{a3} is the fractional pseudoscalar cross section in the $H_{VBF} \to \gamma\gamma$ channel. A value of f_{a3} different than 0 or 1 would indicate a CP violation, with possible scalar and pseudoscalar states. It is convenient to measure $\vec{f}_{ai}$ rather than the anomalous couplings themselves, infact most systematic uncertainties cancel in the ratio. Moreover, the effective fractions are physically bounded by 0 and 1.

Let $S_{ALT_{ai}}$ be the signal model in the invariant mass deriving from an VBF alternative complings' hypothesis , and S_{SM} the ones deriving from a VBF SM hypothesis. The VBF signal model will be defined as $S_{VBF}(f_{ai}) =$

$\mu * [f_{ai} * S_{ALT_{ai}} + (1 - f_{ai}) * S_{SM}]$ in which μ is the overall signal strength and will be a floating parameter in the model.

All measurements in each category are combined to perform a global fit using the Maximum Likelihood method to estimate the effective cross-section ratios. The probability density (P) is normalized to the total event yield in each category k as a sum over all resonant processes j and non-resonant background processes according to:

$$P_k(m_{\gamma\gamma}) = \mu \sum_j P_{jk}^{res}(m_{\gamma\gamma}, \vec{\theta}_{jk}, \vec{f}) + P_k^{non-res}(m_{\gamma\gamma}) \qquad (5.5)$$

Due to the low statistic of the expected VBF signal no interference between diagrams with different coupling is studied.

5.2   Signal and background modeling

The $m_{\gamma\gamma}$ spectrum in each event category is fitted with parameterized signal and background shapes to estimate the number of signal events.

To model the Higgs' peak, the diphoton invariant mass distribution for a nominal Higgs boson mass m_H is constructed from the simulation of each Higgs' production mode and category, to take into account possible differences in line shape in different category or production process. For example, depending on the chosen category the peak resolution varies, in particular, the ones dominated by high photons transverse momentum events will have higher photon energies and better $\sigma(E)/E$ from which better $\sigma(M)/M$

Given that there is no reliable MC sample to model, the non-resonant background a fully data-driven method in the control region is chosen. Using this strategy the falling background spectrum is determined directly from the observed data in the sideband of the diphoton invariant mass.

5.2.1   Signal and resonant background

The resolution on the photon opening angle has a negligible contribution to the mass resolution, compared to the ECAL energy resolution when the interaction point is known to be within about 1 cm. For this reason the two scenarios of a correct or wrong vertex are considered separately when constructing the signal model.

To keep track of the year-dependent resolution effects and the year-dependent systematic uncertainties, the signal model will also be construct separately for each year.

For each signal production process ($N^{proc} = 5$), year ($N^{year} = 3$), category ($N^{cat} = 5$), and vertex scenario ($N^{vrx} = 2$), $m_{\gamma\gamma}$'s distributions are fitted using a sum of at most five Gaussian functions, then, the sum of the gaussians that minimize the χ^2 is chosen as the nominal PDF. Therefore the number of signal

models is $N^{proc} \times N^{year} \times N^{cat} \times N^{vrx} = 5 \times 3 \times 5 \times 2 = 150$. In Figure 5.1 shows an example of how the nominal PDF is chosen in a category where most of the BSM VBF events are expected, using a SM Higgs boson VBF MC sample for the year 2017 and the Right Vertex Scenario.

Some of the parameters are dependent on m_H, as the position of the peak or the yield, which is known with an uncertainty ($m_H = (125.25 \pm 0.17)$ GeV [23]). To take into account the variation of these parameters as a function of m_H within its experimental uncertainy associated with the Higgs boson mass, the trends of the parameters as a function of m_H are interpolated with polynomials using Montecarlo samples generated in the hypothesis in which $m_H = 120, 125, 130$ GeV.

MC samples, related to $m_H \neq 125 GeV$, are generated using JHUGen and are available for all Standard Model (SM) processes. For the processes with alternative coupling hypotheses, the information obtained from the VBF SM samples will be used.

The the cross section and braching ratio values are set according to LHC measurements and theoretical calculations [23]. On the other hand the efficiency and the expected fraction of Right Vertex events are found from the simulation. To extract the expected number of events in an integrateld luminosity $L = 137 fb^{-1}$ the following formula is used:

$$N_{j,i} = (\sigma BR)_j \times (\epsilon A)_{j,i} \times L \tag{5.6}$$

where $N_{i,j}$ is the number of signal events of process j in category i. The $(\epsilon A)_{i,j}$ defines the fraction of signal process j, falling in the i-th category. Fig. 5.2 shows the expected distributions and fitted lineshapes for the 3 m_H points for which we have the MC simulation and the interpolated ones. The figure also shows the interpolated trend of N_{ij}, σ, BR, ϵA and RV fraction.

The function modelling the signal in each category is obtained by summing all the Gaussians found. Fig. 5.3 shows the sum of the contributions of all processes (for both vertex scenarios) by category, separated by year. Depending on the chosen category the peak resolution varies, in particular, due to being dominated by high transverse momentum events, BSM event-dominated categories have an improved resolution.
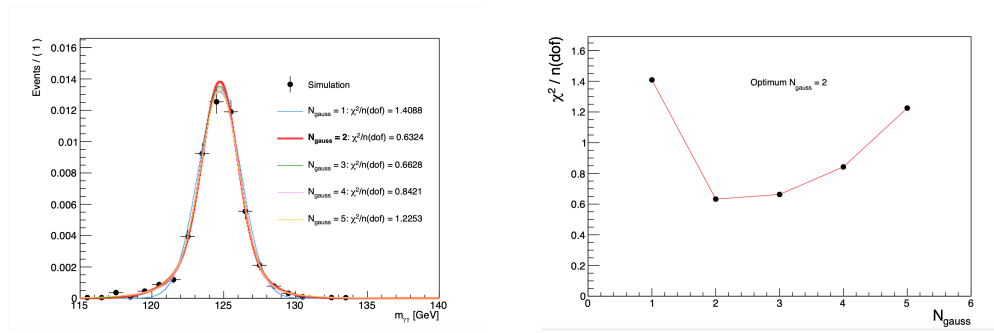
Figure 5.1. Example of a fit in category where most of the BSM VBF events are expected using a SM Higgs boson VBF MC sample for the year 2017 and Right Vertex scenario. The nominal probability density function (pdf) is the distribution that minimizes the chi-square that is the sum of two Gaussians shown in the right plot, while the best PDF is shown with the thick line in the left plot.
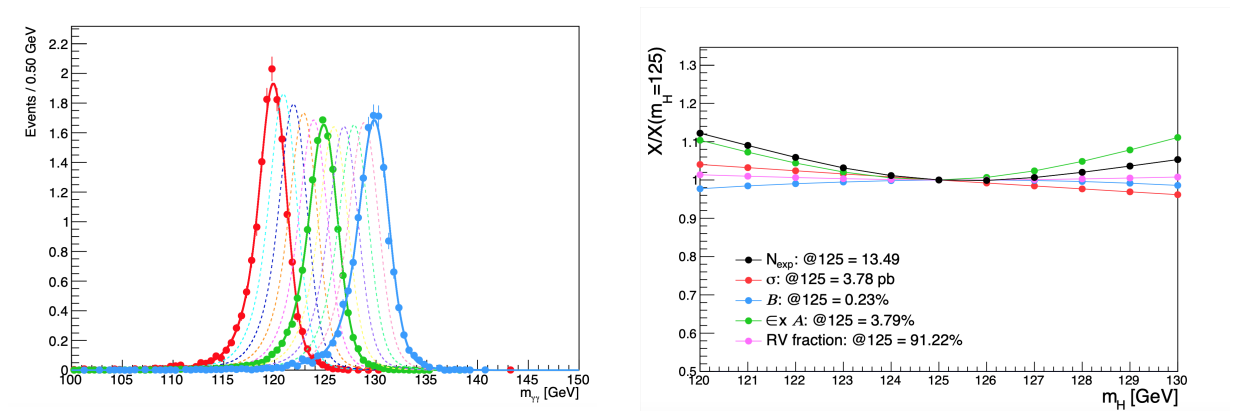

Figure 5.2. On the left the $m_{\gamma\gamma}$ distribution with Higgs mass hypothesis different than 125 GeV. On the right the trends of the parameters as a function of M_H.
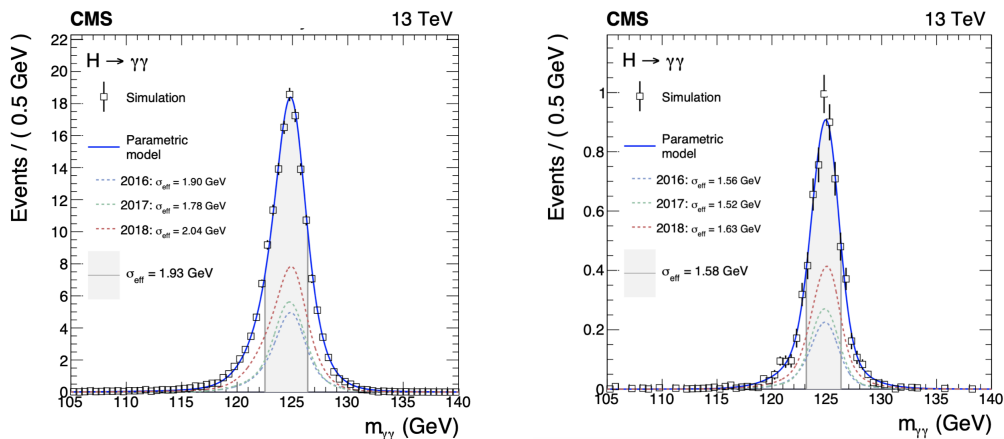
Figure 5.3. Sum of the contribution of all processes separated by year in different category. On the Right a category dominated by SM VBF type of events on the left a category dominated by BSM VBF events.

5.2.2 Non-resonant background

Data in side-bands around the expected signal peak is used to model the $m_{\gamma\gamma}$ PDF of the background and a set of candidate function families is considered, including exponential functions, Bernstein polynomials, Laurent series and power law functions.

The function with the lower chi-squared is chosen as the nominal PDF.

When studying a small signal sitting on a large background, relatively small changes to the background shape can have significant effects on the final fit result. As a consequence, there is an uncertainty about which function to use to parameterize the background.

A common approach to assess this systematic uncertainty is to fit various different plausible functions and determine the spread of the values of the parameters of interest when using these functions. However, these methods tend to have some degree of arbitrariness that's why the *Discrete profiling method* [30] is used .
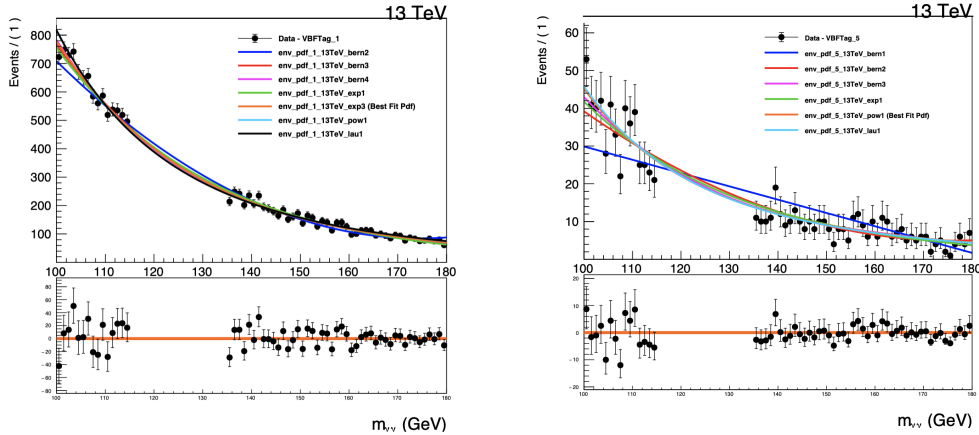
Figure 5.4. Background modeling in different categories, on the left a category with high background statistic. On the right is a category in which the choice of the functional form of the background significantly affects the result.

5.3 Discrete profiling method

The discrete profiling method uses the choice of background function as a systematic uncertainty which is handled as a nuisance parameter. A nuissance parameter is a variable that we are not interested in, however it must be accounted for to estimate the variables of interest called Parameter of Interest (POI), for example the number of signal events. When fitting for the POI, the Likelihood function is minimised with respect to it and all the other nuisance parameters.

In the so call *discrete profiling method*, each backgroud function is labelled by an integer which will be treated as a discrete nuisance parameter. If the discrete nuisance is left floating, it will be profiled by looping through the possible index values and finding during the fit the pdf which gives the best fit. In other words the envelope of all possibilities is taken, and the systematic error is converted into a statistical error in an automatic way and with less arbitrariness.

To test this approach, I generated a pseudo-dataset (toy) to perform a likelihood (L) scan as a function of the VBF signal strength modifier μ_{VBF}, defined as the ratio of the observed Higgs VBF production with regards to the SM expectation. To make the effect more noticeable the scan is made using only one category, therefore the fit result should not be regarded as a physical result. By leaving the background functional form as a floating parameter, at each point in the likelihood scan, the PDF will be set as the one which gives the lowest value of the function $-2 \times Log(L)$.

Fig. 5.5 shows, as expected, that the curve obtained by allowing the integer PDF index to float picks out the best function for each value of μ_{VBF}. Notice that the use of the "bern1" function to model the background returns poor

results and so the likelihood related to this function is higher than the plot range. This function plays no role in the envolope and so this will be ignored by the method. Each curve has a different minimum, this is a direct example of the systematic uncertainty associated with the choice of function.
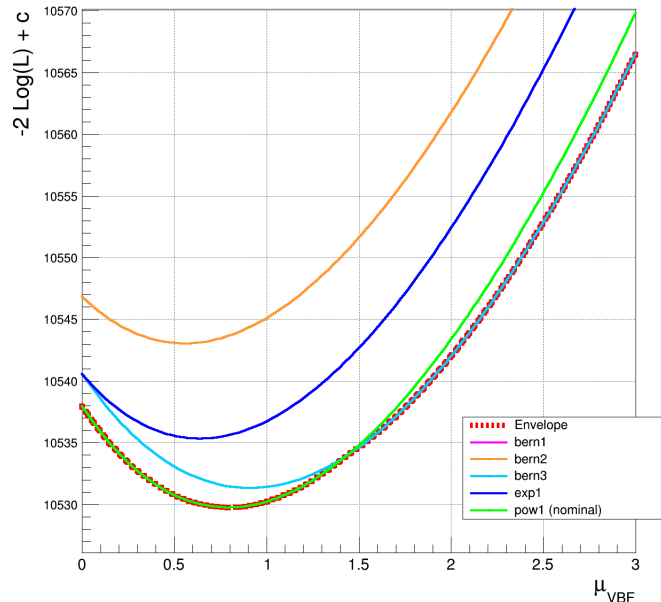

Figure 5.5. Absolute value of (twice) the negative log-likelihood scan as a function of the VBF signal strength. The curve is obtained by allowing the pdf index to float (labelled "Envelope"), picks out the best function for each value of the signal strength. The other curves are computed fixing a background-shape model into the fit. The c value is a correction to take into account the different number of parameters between different function (as defined in [30])

5.4 Systematic uncertainties

We have shown in Sec. 5.2.2 a way to treat the systematic uncertainty related to the choice of the bakground PDF. In addition to that, both theoretical systematics and experimental uncertainties have to also be introduced in the model. Uncertainties that modify the shape of the invariant mass peak (i.e. uncertainties related to the calibration that shifts the position of the peak) have to be considered when modeling the signal and resonant background or resolution effects that smear the width of the resonance. These uncertainties are reflected as an uncertainty on the shape of each Gaussian in each category and account for possible migrations of signal and background events between categories, in fact most of the selection criteria are based on cut that involve E_1 and E_2 so changing the energy scale or the resolution could affect the yield in each category.

If the shape of the $m_{\gamma\gamma}$ is unaffected and the uncertainty is related only to the event yield (i.e. integrated luminosity systematics) a log-normal distribution is used to consider it in the Likelihood. The log-normal distribution is a function of a natural logarithm that is normally distributed. For small uncertainties, the log-normal is approximately a Gaussian.

The method used in this analysis treats systematic uncertainties as a nuisance parameter so thet the Likelihood will be minimized with respect to them.
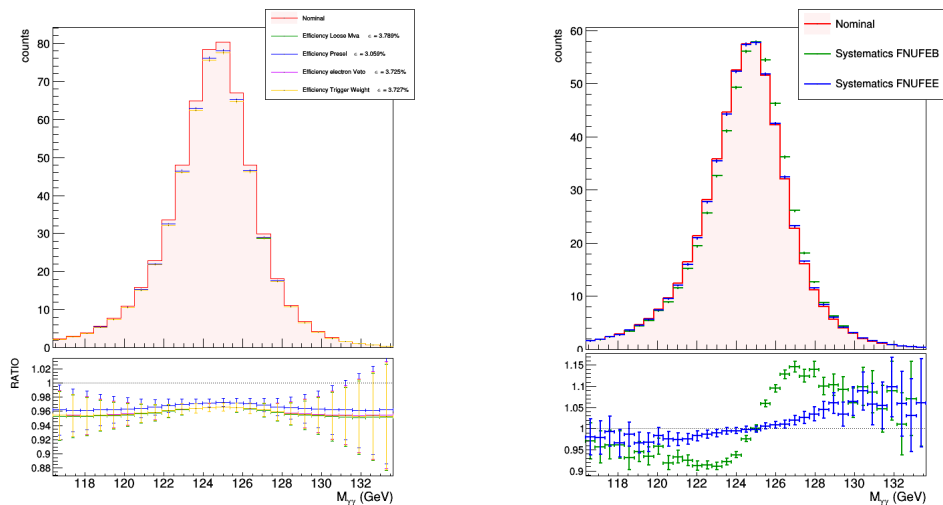
i

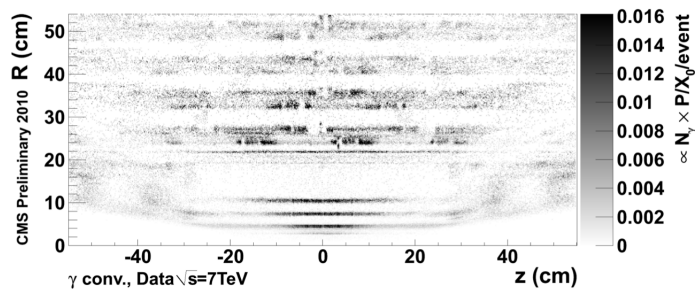**Figure 5.6.** Example of shape systematics and normalization systematics.

Figure 5.7. Maps of the z-R cross section of the material distribution for nuclear interactions. The plot highlights each part of the tracker well, just as if it were an x-ray [13].

5.4.1 Theoretical uncertainties

- *ggH associated scales and migrations* : Theoretical systematics on scale and category migrations due to the QCD scale in the gluon fusion Higgs boson production cross section. There are 8 uncertainties in total, 2 of

which correspond to scale change and 6 others to category migrations of events with associated jet productions.

5.4.2 Experimental uncertainties affecting the $m_{\gamma\gamma}$ shape

The uncertainties affecting the $m_{\gamma\gamma}$ shape have been calculated in CMS for each of the years, and I have verified the effect on the shape of the resonance adding them on the fit.

- *Photon energy scale and resolution* : The energy is corrected for the lack of the complete containment of the electromagnetic showers in the clustered crystals, the energy lost by photons that convert before reaching the calorimeter, and the effects of pileup. After applying these corrections to the photon energy, some residual differences between data an MC remain. To correct them $Z \rightarrow ee$ events in which electron shower are reconstructed as photons are used.

  The energy deposition of photons that convert before reaching the calorimeter tend to have wider transverse profiles and thus lower values of R9 than those of unconverted photons. So the systematics is separately estimated not only for different detector regions (endcap or barrel) but also for different R9 values (High R9, Low R9).

- *Non-uniformity of light collection* : The shower maximum for photons is deeper than that of electrons by approximately one radiation length. Differences in the light collection efficiency along the length of the ECAL crystals will introduce a difference in the ECAL response to electrons and photons. As such, additional systematic uncertainty is assigned to the photon energy scale. The systematics are modelled differently for the endcap and for the barrel regions due to differences in the degree of radiation and crystal size. This systematic is calculated by varying the light yield model in the crystal at the simulation level with GEANT (Fig. 5.6 right).

- *Modelling of material in front of the ECAL* : the amount of material through which objects pass before reaching the ECAL affects the behaviour of the electromagnetic showers, and may not be perfectly modelled in the simulation. Studies of photon conversions, nuclear interactions and multiple scattering are used to check the consistency of the Tracker simulation geometry with the material distribution inside the actual detector. The effect on the energy scale is at most 0.24%. Figure 5.7 shows a map of the material distribution made by studying nuclear interactions [13].

5.4.3 Experimental uncertainties of normalization

- *Integrated luminosity* : Uncertainties of 1.0%, 2.0% and 1.5% are determined by the CMS luminosity monitoring for 2016, 2017, and 2018

data-sets. The absolute calibration scan uncertainties (uncorrelated for the 3 years) is kept separately from non-linear luminometers response or from external factor such as the LHC filling scheme.

- *Photon identification BDT score* : A boosted decision tree (BDT) is used to separate prompt photons from photon candidates that arise from misidentified jet fragments, but which satisfy the preselection. This photon identification BDT is trained using simulated $\gamma + jets$ events where prompt photons are considered as signal and non-prompt photons as background. Differences between Data and MC samples of the inputs lead to a possible difference in the selection efficiency on the BDT, and therefore a systematic uncertainty.

- *Jet energy scale and smearing corrections* : The impact of jet energy scale uncertainties in event yields is evaluated by varying the jet energy corrections within their uncertainties and propagating the effect to the final result. This uncertainty leads to migration within VBF categories, within ttH categories, within VH categories, and from tagged to untagged categories.

- *Trigger efficiency* : the efficiency of the trigger selection is measured with $Z \rightarrow ee$ events.

- *Photon preselection* : the photon preselection efficiency is computed as the ratio between the efficiency measured in data and in simulation. The uncertainty associated with this value is of various kinds, for example it derives from the limited systematics of the sample $z \rightarrow ee$ used to estimate it.

5.4.4 Impact on μ_{VBF} and f_{a3}

Initially I studied the impact of systematics on μ_{VBF} in a SM hypothesis. The profile Likelihood ratio is built plotting twice the negative of the logarithm of the Likelihood ratio function (Λ):

$$\Lambda = -2ln\frac{L(\mu,\hat{\hat{\theta}})}{L(\hat{\mu},\hat{\theta})} \tag{5.7}$$

Where $\hat{\mu}$ (parameter of interest in this case) and $\hat{\theta}$ (nuisances parameters) are the best values of the parameters obtained by maximizing L; $\hat{\hat{\theta}}$ are the values of the nuisance parameters obtained by maximizing L at μ fixed.

Let Λ_{BF} be Λ's minimum the 68.3 % confidence interval of μ_{VBF} is the interval where $\Lambda < \Lambda_{BF}+1$. Similary the 95.4% confidence intervals are defined by the regions in which $\Lambda < \Lambda_{BF} + 4$.

Because of with the preliminary selection cuts that selects VBF-like events, there are not categories purely dominated by other production modes other

than VBF, so the analysis, is not sensitive to SM coupling other than the VBF one as shown in Fig. 5.8. Moreover, as expected, leaving all the strenght modifiers, of the principal production modes, as floating parameters the impact of statistical errors on the μ_{VBF} parameter is significantly greater, making the systematics' impact negligible (Fig. 5.9). We choose to focus the analysis on the VBF production so the study on the parameter μ_{VBF} is made fixing the other SM couplings at the nominal SM value (i.e. one), using this approach the significance of the VBF signal is $\sqrt{13.9}\sigma = 3.73\sigma$.
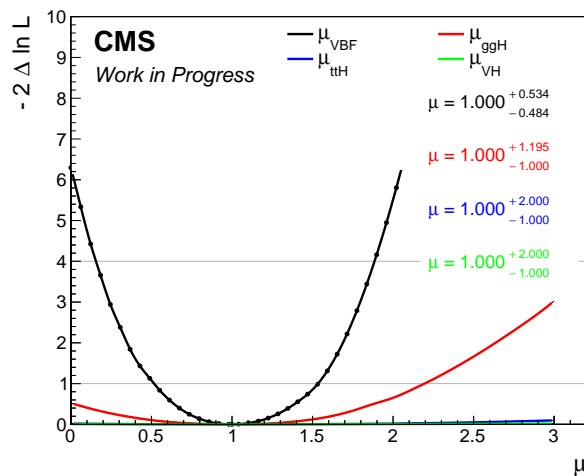
Figure 5.8. Λ profile expected where all the strength of SM couplings are floating parameters. The significance of the VBF signal in this condition is $\sqrt{6.3}\sigma = 2.5\sigma$. The Figure shows that the analysis is not sensity to strenght parameter different than the one related to the VBF coupling.
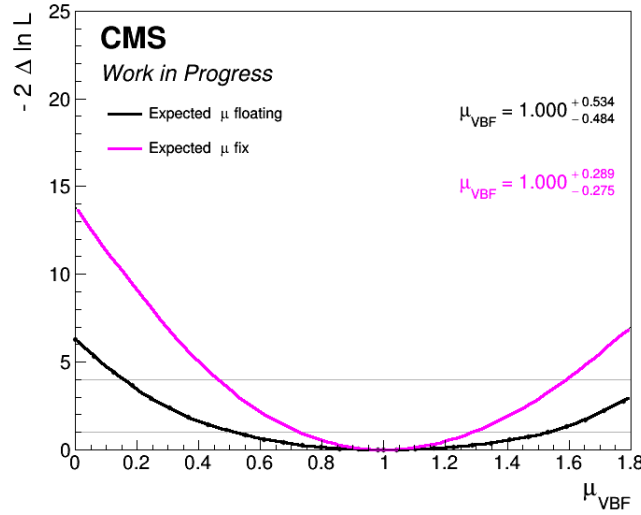
Figure 5.9. Λ profile expected as a function of μ_{VBF} where all the strength of SM couplings are floating parameters (black curve) and where the other SM couplings ($\mu_{ttH}, \mu_{VH}, \mu_{ggH}$) are fixed to the nominal value (violet curve).

Fig. 5.10 shows the Λ profile expected where the strength of the other SM couplings are fixed at the nominal value (i.e. one). To build the curve an Asimov data-set is used i.e. the one data set in which all observed quantities are set equal to their expected values. Using the Asimov datasets avoids generating several toys for averaging the results by directly estimate the parameter from an "average sample" of data.

If the nuisance parameters were fixed to their values at the best fit point (i.e. the absolute minimum in Λ) then the profile Λ curve would be narrower. This directly reflects the fact that if the nuisance parameters had no uncertainty, the error on the POI would be reduced, i.e. there would be no systematic uncertainty arising from this source and the width of this curve would only be affected by the statistical uncertainty.

In this way I evaluate the total impact of the systematics on the parameter of interest. In Fig. 5.10 is shown the Λ profile expected as a function of μ_{VBF} with and without introducing the systematics. The best fit value, considering the systematic uncertainties, is $\mu_{VBF} = 1.00^{+0.28}_{-0.27}$. In the same figure, on the right, a summary of the impact of the main sources of systematic uncertainty in the fit to signal VBF strength.
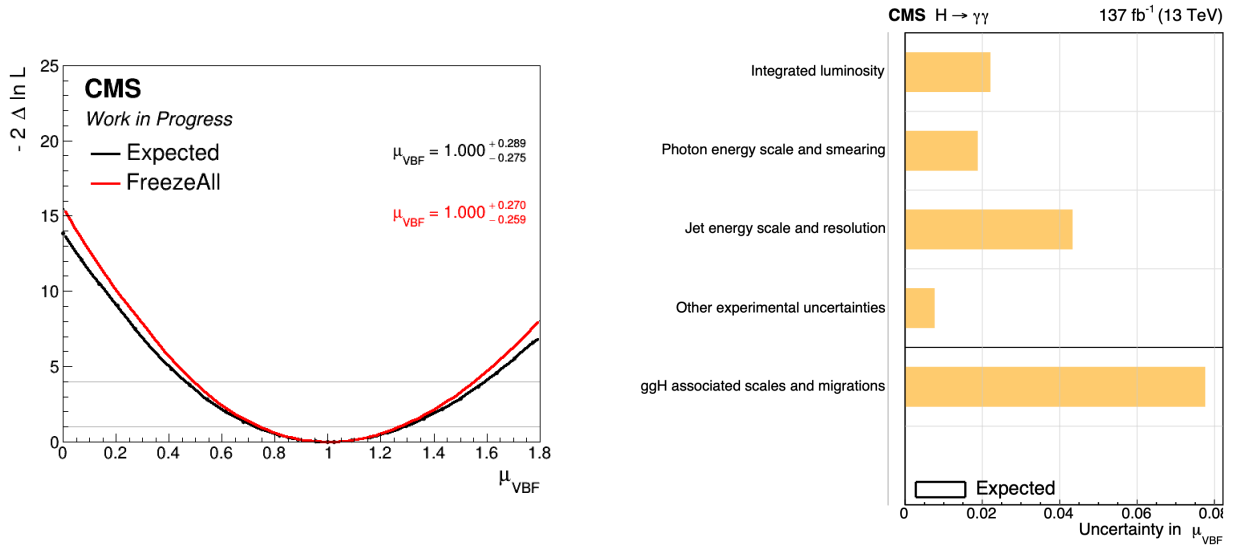
Figure 5.10. On the left the Λ profile expected where the strength of the other SM couplings are fixed at the nominal value. The curve labelled as "FreezeAll" is calculated fixing all the nuisance parameters to their values at the best fit point. On the right, a summary of the expected impact of the main sources of systematic uncertainty on μ_{VBF}.
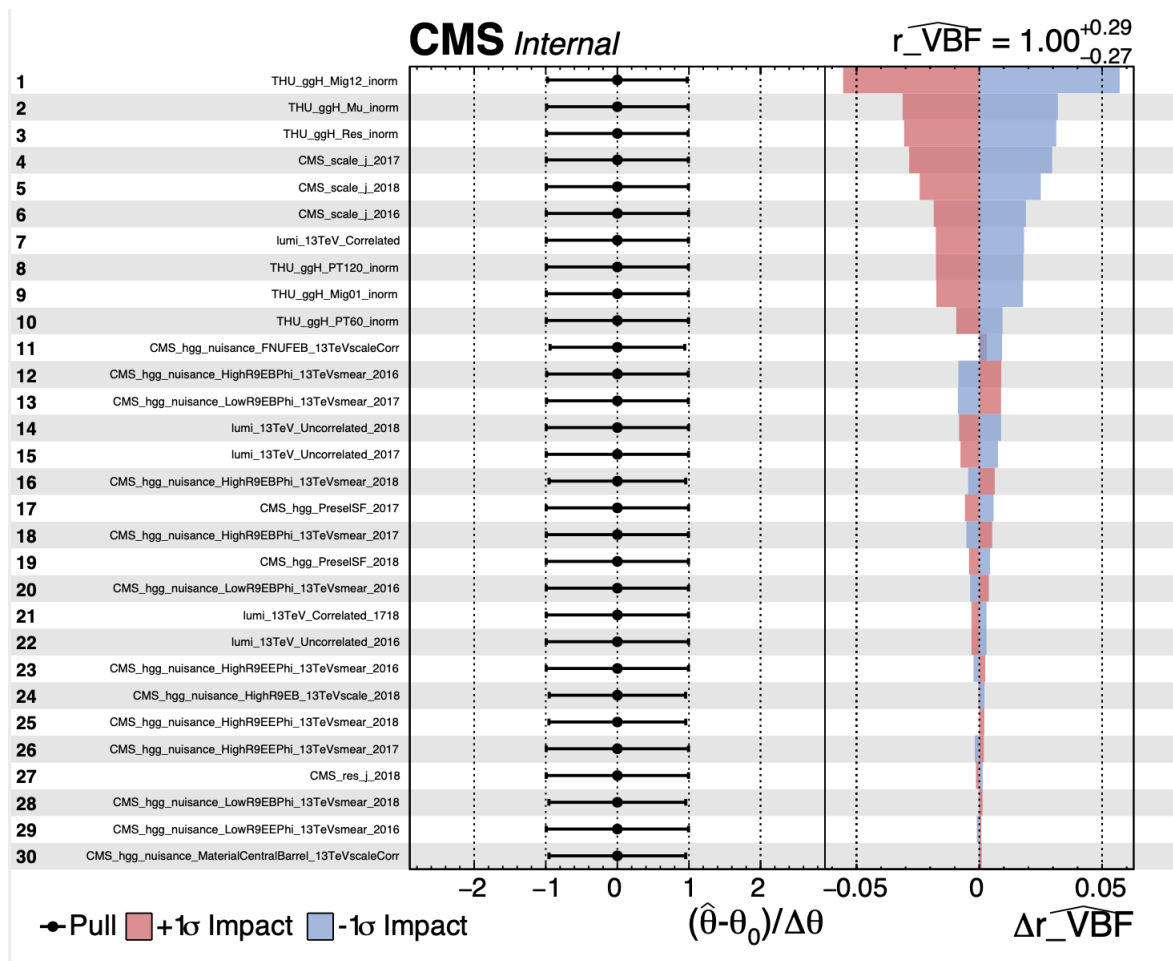
Figure 5.11. Impact of the individual systematics on μ_{VBF}, due to the fact that we are using an Asimov data-set the pull value is centered at one for each systematic.

The same working method to study the impact of systematics on μ_{VBF} is also used for the study of the f_{a_3} parameter. Most uncertainties cancel in this analysis because measurements of anomalous couplings are expressed as relative cross sections. Statistical uncertainties dominate over any systematic uncertainties in this fit.

The profile Likelihood ratio is also plotted for the other BSM couplings considered in the analysis (Fig. 5.13).
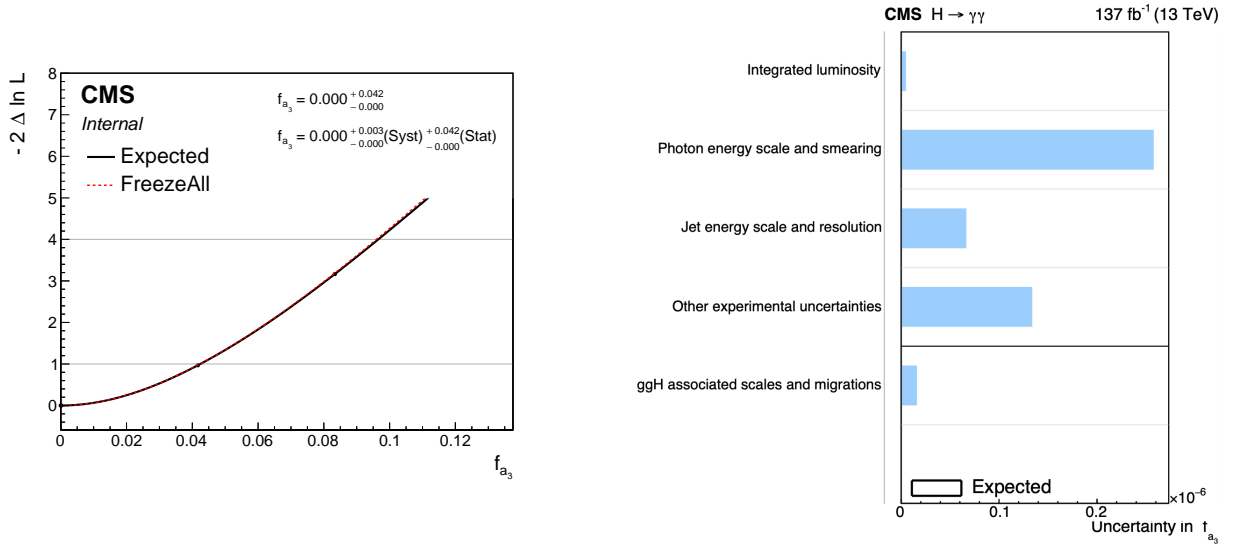
Figure 5.12. On the left, the Λ profile expected where the f_{a3}. The curve labeled as "FreezeAll" is calculated fixing all the nuisance parameters to their values at the best-fit point. On the right, a summary of the expected impact of systematics on the parameter of interest is shown.
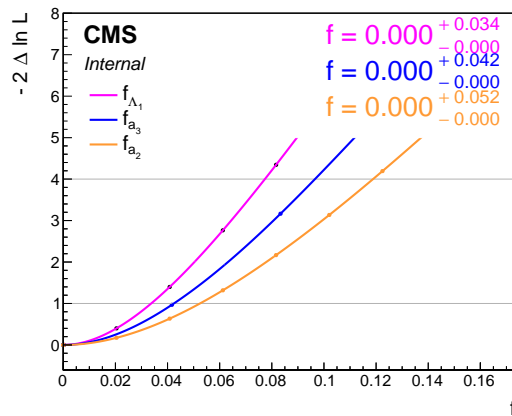

Figure 5.13. The profile Likelihood ratio expected as a function of the effective cross-section function for anomalous couplings scenario.

5.5 Fit stability

To study the stability of the fit with regards to different choices of functions to parameterize the background and to access the bias introduced by selecting a particular function it is developed a bias study using toy Monte Carlo experiments. To do that i used a statistical toolkit developed by CMS born to combine the results of the Higgs analysis in different channels. Thanks to

the sophisticated tools available today, the toolkit is largely used in the high energy sector for Likelihood estimation, shape analysis and unbinned analysis.

For the bias study, we generated 1000 pseudo-datasets (*toys*) according to nominal PDFs but using a different background function within a category to fit it, repeating the procedure for each category.

The parameter of interest (POI) fitted in this study is the VBF signal strength modifier μ_{VBF} defined as the ratio of the observed Higgs VBF production to the standard model expectation. To measure the bias we study the pull defined as the difference between the measured value of μ_{VBF} and the generated value (here we used a SM model so it's 1) relative to the uncertainty on μ_{VBF}:

$$pull = \frac{2(\mu_{VBF} - 1)}{\sigma_\mu^+ + \sigma_\mu^-} \tag{5.8}$$

The formula takes in consideration asymmetric errors.

The pull distribution can be drawn and the mean provides an estimate of the bias in units of total fit uncertainty. An unbiased fit will have a gaussian pull distribution centered at 0 an with variance equal to 1.

In the example shown in Fig. 5.14 a degree 3 Bernstein polynomial in the category most populated by SM VBF events is chosen to fit the toys when the true function (the one used to generate it) is a degree 2 Bernstein polynomial in the same category. The fitted Gaussian has a mean value of 0.03 which would indicate a bias of $3\% \times \sigma(\mu_{VBF})$ of the uncertainty on μ_{VBF}.

Table 5.14 highlights that the category with the most bias is the one with the least background statistic in the control region. This is in agreement with what was expected: a lower background statistic leads to more variations in the PDF shapes and greater uncertainty due to the function choice.
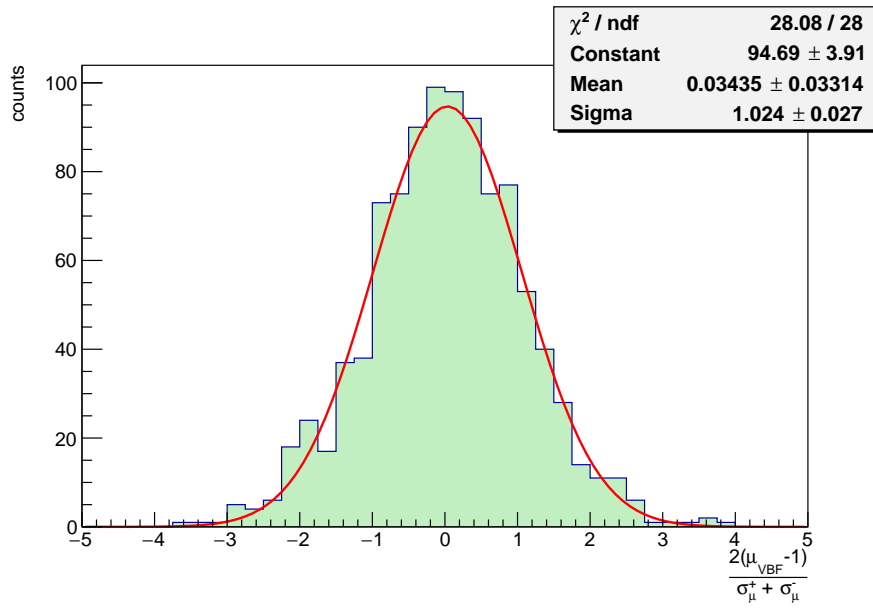
χ^2 / ndf	28.08 / 28
Constant	94.69 ± 3.91
Mean	0.03435 ± 0.03314
Sigma	1.024 ± 0.027

Figure 5.14. pull distribution when a degree 3 Bernstein polynomial in the category most populated by SM VBF events is chosen to fit the toys when the true function (the one used to generate it) is a degree 2 Bernstein polynomial.

	mean	sigma
$P(BKG) < 0.05, D_0^- < 0.6$, P(BSM) < 0.97	-0.003	0.965
$P(BKG) < 0.05, D_0^- < 0.6$, P(BSM) > 0.97	-0.029	1.012
$P(BKG) < 0.05, D_0^- > 0.6$, P(BSM) < 0.97	0.121	0.967
$P(BKG) > 0.05, D_0^- < 0.6$, P(BSM) < 0.97	0.056	1.011
$P(BKG) > 0.05, D_0^- > 0.6$, P(BSM) < 0.97	0.034	1.024

Table 5.1. Gaussian parameters of the pull distribution in different category

Chapter 6

Results

All the optimization procedure is done in a "blind" way i.e. without using the data in the signal region. The fit is performed on data unblinding the signal region and the likelihood profile is calculated for the three sets of f_{ai}, describing anomalous HVV couplings in the production sector. Fig. 6.1 shows the fit in all five categories. As expected, the signal is more prominent in the low background categories. In order to show in a unique plot the data with the fit results, preserving the higher significance of the events in the high S/B categories, the events are merged with a weight proportional to S/(S+B). The weighted data, with the fit result rescaled accordingly, is shown in Fig. 6.6. The best-fit values for the signal strength, μ (i.e. the ratio of the measured cross section and the one predicted by the Standard Model for a Higgs boson mass $m_H = 125$ GeV), under the assumption that $f_{ai} = 0$, is $\mu = 0.97^{+0.19}_{-0.17}$ i.e. consistent with the SM prediction. Tab. 6.2 shows the peak contribution yields in the various categories.

When doing the fit simultaneously for μ_{VBF} and one for f_{ai}, the sum of the peak events is less than expected and the best fit parameter of the VBF output signal strength is $\mu_{VBF} = 0.80^{+0.26}_{-0.25}$, which is underestimated from what is expected from the standard model. This affectes the C.L. (Confidence Level) of the effective cross-section fraction found: due to the under-fluctuation of the signal in all categories, the f_{ai} best-fit value, not imposing any physical constraints ($f_{ai} \geq 0$), results negative.

The results of the likelihood scan are shown in Fig. 6.2, 6.3 and 6.4 and in Tab. 6.1.

Parameter	Observed	Expected
f_{a3}	$-0.045^{+0.036}_{-0.031}[-0.110, 0.037]$	$0.000^{+0.042}_{-0.035}[-0.068, 0.099]$
f_{a2}	$-0.057^{+0.044}_{-0.039}[-0.13, 0.041]$	$0.000^{+0.052}_{-0.043}[-0.087, 0.119]$
f_{Λ_1}	$-0.037^{+0.028}_{-0.025}[-0.082, 0.025]$	$0.000^{+0.034}_{-0.027}[-0.054, 0.078]$

Table 6.1. Allowed 68% CL (central values with uncertainties) and 95% CL (in square brackets) intervals on f_{ai} parameters

The results shown by previous searches in the chapter 3.2 use the cross-section ratio defined in the $H \to 2e2\mu$ decay as the common convention across various measurements, knowing from a LO calculation of the cross-section that $\sigma_1^{VBF}/\sigma_3^{VBF} = 0.089$ and that $\sigma_1^{H\to 2e2\mu}/\sigma_3^{H\to 2e2\mu} = 6.53$. Rescaling the cross-section by these factors to make a comparison (this is needed because in VBF $H \to 2e2\mu$ a possible anomalous coupling with Z bosons appears both in the production side and in the $H \to ZZ$ decay), the fitted f_{a3} becomes $f_{a3} < 0.00057$ at 68% of C.L. This result is then less sensitive as the current best limit, obtained using the combination of $H \to ZZ \to 4l$ and the VBF $H \to \tau^+\tau^-$ decays.
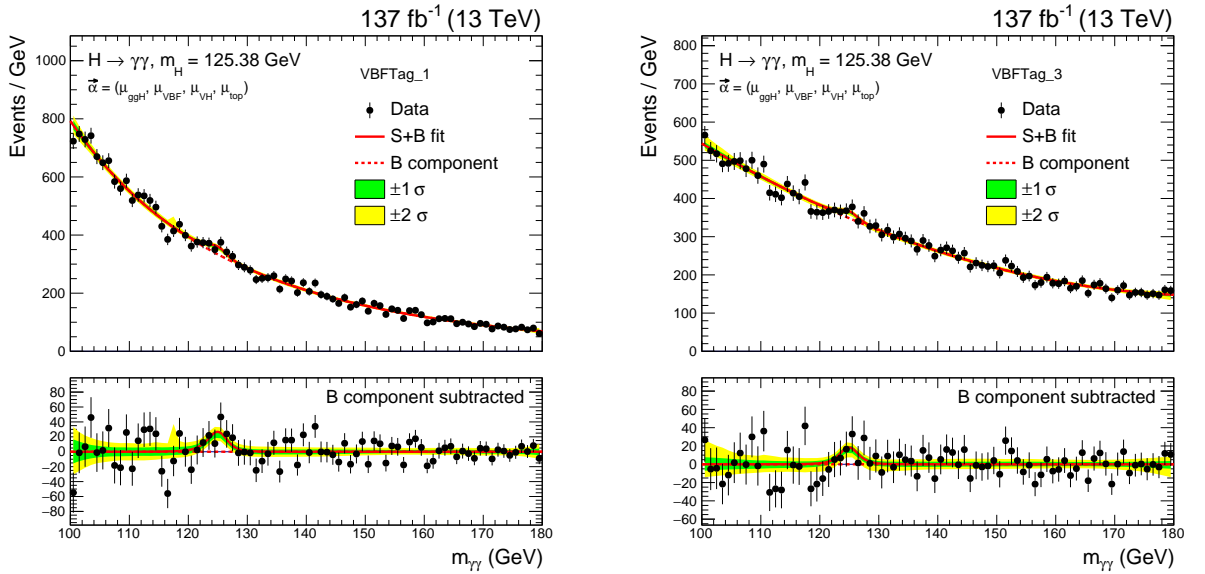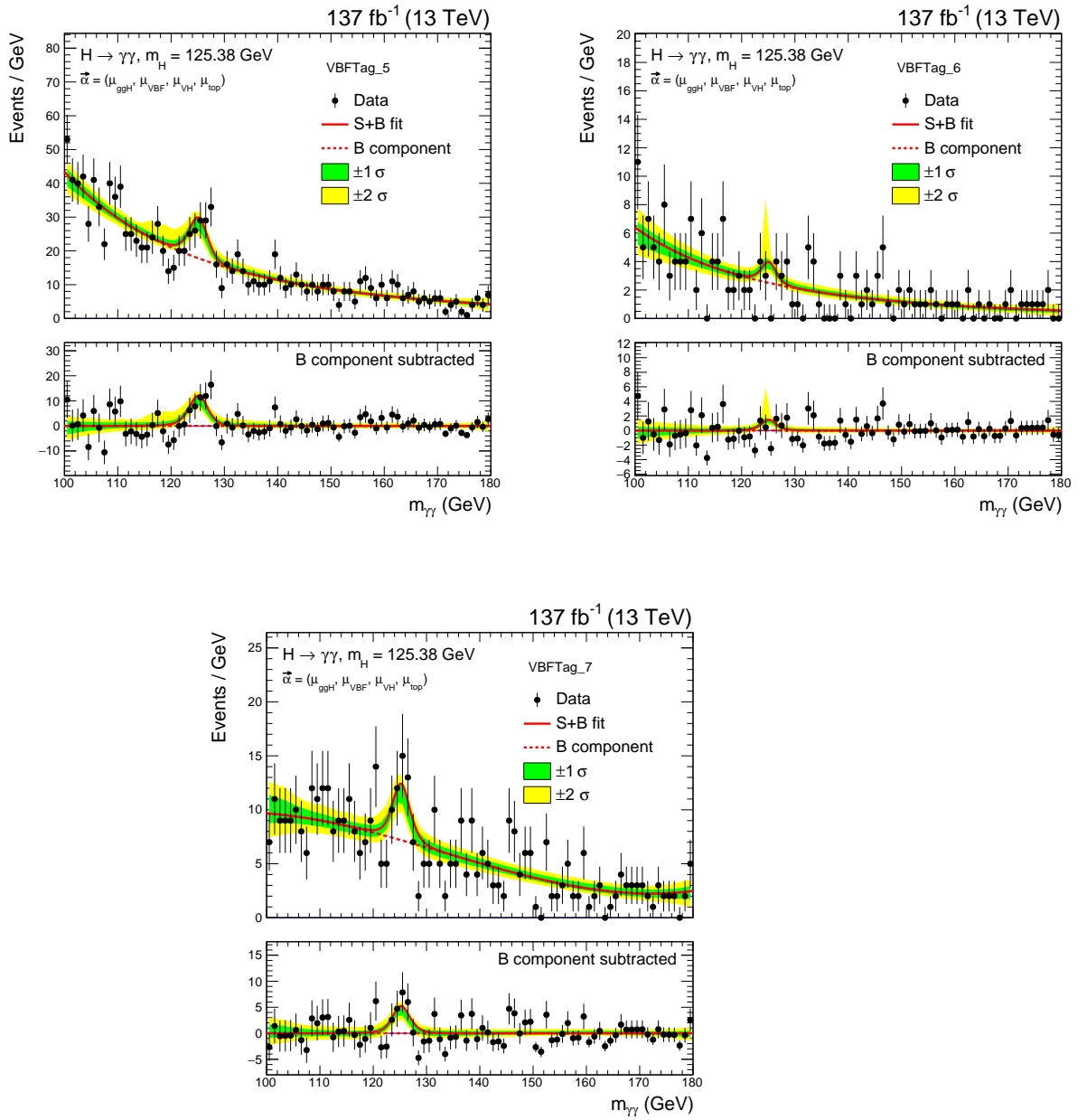
Figure 6.1. Di-photon invariant mass distribution in different category (). The lines represent the fitted background and signal, and the colored bands represent the ±1 and ±2 standard deviation uncertainties in the background estimate.
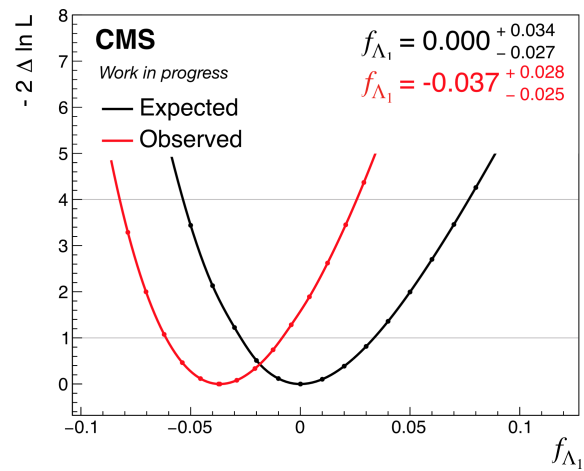
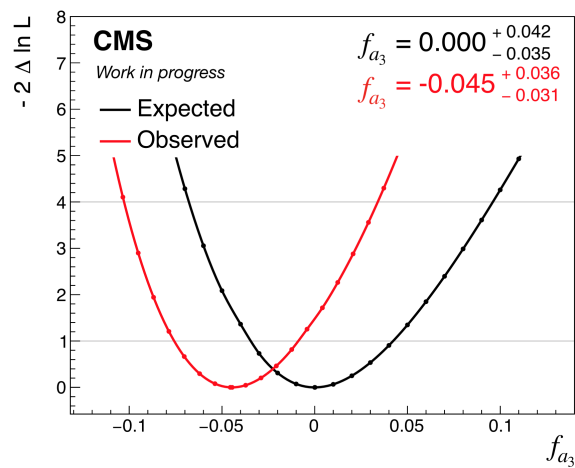**Figure 6.2.** The profile Likelihood ratio expected and observed as a function of f_{Λ_1}.

Figure 6.3. The profile Likelihood ratio expected and observed as a function of f_{a3}.

Figure 6.4. The profile Likelihood ratio expected and observed as a function of f_{a2}
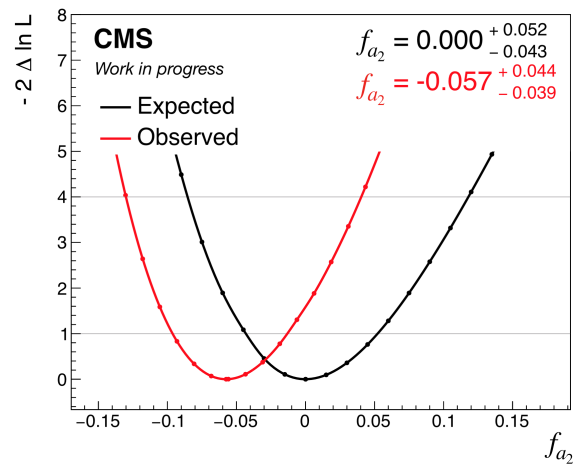
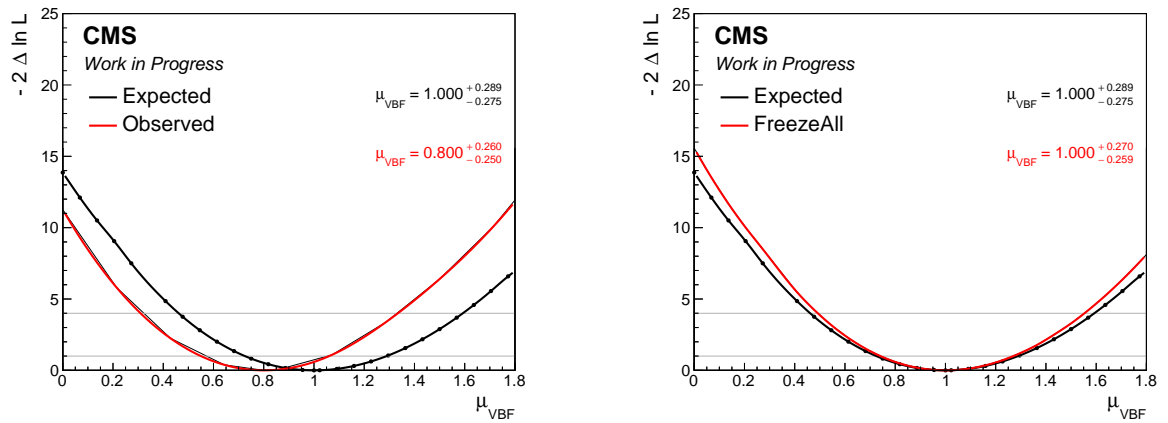

Figure 6.5. On the left the Λ profile expected and observed as a function of μ_{VBF} when the other SM strength value are fixed at the nominal value. On the right the Λ profile expected and observed as a function of μ

	S_{obs}	S_{exp}
$P(BKG) < 0.05$, $D_0^- < 0.6$, $P(BSM) < 0.97$	23	24
$P(BKG) < 0.05$, $D_0^- < 0.6$, $P(BSM) > 0.97$	6	7
$P(BKG) < 0.05$, $D_0^- > 0.6$, $P(BSM) < 0.97$	57	58
$P(BKG) > 0.05$, $D_0^- < 0.6$, $P(BSM) < 0.97$	83	84
$P(BKG) > 0.05$, $D_0^- > 0.6$, $P(BSM) < 0.97$	126	128
ToT	295	301

Table 6.2. Sum of the events belonging to the peak aspects and observed, calculated by subtracting the signal plus background spectrum to the background spectrum.
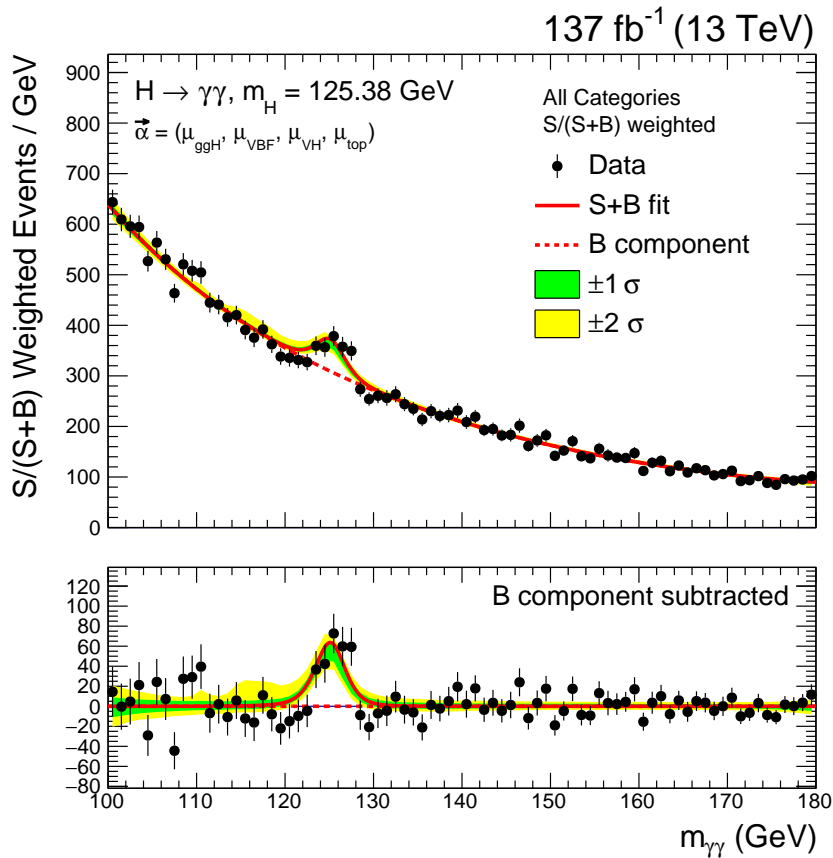

Figure 6.6. The diphoton invariant mass distribution with each event weighted by the S/(S+B) value of its category. Note that by S we mean the Higgs boson production events participating to the invariant mass peak. The lines represent the fitted background and signal, and the coloured bands represent the ±1 and ±2 standard deviation uncertainties in the background estimate.

Conclusions

In this thesis I studied possible anomalous interactions of the Higgs boson (H) with vector bosons (V), including CP-violating effects, using the VBF production and the $H \to \gamma\gamma$ decay channel. Constraints on the CP-violating parameter f_{a3} and on the CP-conserving parameters f_{a2} and f_{Λ_1} are set using 2016–2018 data from pp collisions recorded with the CMS detector during the Run 2 of the LHC. The data used correspond to an integrated luminosity of $137\,fb^{-1}$ at a center of mass energy of 13 TeV.

After selecting the signal with respect to the background, the events have been categorized exploiting machine learning techniques and matrix element approach to maximize the sensitivity of the analysis to anomalous couplings. Then the signal and background process are modeled to perform a global fit and estimate the parameters of interest using the likelihood method. The simple case is studied in which only one of the BSM amplitudes contributes to the production cross-section, simultaneously to the SM, is studied, i.e. only one f_{ai} at a time is different from 0. The constraints at 95% C.L. observed on the CP-violating parameter, f_{a3}, and on the CP-conserving parameters are found: $f_{a3} < 0.037$, $f_{a2} < 0.041$ and $f_{\Lambda_1} < 0.025$. The measured VBF signal strength is $\mu_{VBF} = 0.80^{+0.26}_{-0.25}$, consistent with the SM expectation. The f_{ai} measured constraints presented are limited by statistical precision and are consistent with the expectations for the standard model Higgs boson. The impact of the main systematics on the parameters was also analyzed and found to play a minor role. The obtained limits on the three f_{ai} effective fraction of anomalous cross section are competitive with the best ones from the combination of $H \to ZZ \to 4l$ and $H \to \tau^-\tau^+$ analyses.

A futher extension of the analysis would be to consider simultaneous BSM amplitudes and their interference effects to probe the complete BSM Lagrangian of the HVV interaction.

Bibliography

[1] ATLAS-Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc," *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 716, pp. 1–29, 9 2012.

[2] The-CMS-Collaboration, "Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc," 7 2012. [Online]. Available: http://arxiv.org/abs/1207.7235http://dx.doi.org/10.1016/j.physletb.2012.08.021

[3] V. Khachatryan, "Constraints on the spin-parity and anomalous hvv couplings of the higgs boson in proton collisions at 7 and 8 tev," *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 92, 7 2015.

[4] V. K. et al., "Constraints on the spin-parity and anomalous hvv couplings of the higgs boson in proton collisions at 7 and 8 tev," *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 92, 7 2015.

[5] The-CMS-Collaboration, "Constraints on anomalous higgs boson couplings using production and decay information in the four-lepton final state," *Physics Letters B*, vol. 775, pp. 1–24, dec 2017. [Online]. Available: https://doi.org/10.1016%2Fj.physletb.2017.10.021

[6] T. C. collaboration, "Measurements of the higgs boson width and anomalous hvv couplings from on-shell and off-shell production in the four-lepton final state," *Physical Review D*, vol. 99, 6 2019.

[7] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," 1964.

[8] L. H. C. S. W. Group, "Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector," 10 2016. [Online]. Available: http://arxiv.org/abs/1610.07922http://dx.doi.org/10.23731/CYRM-2017-002

[9] The-CMS-Collaboration, "Precise determination of the mass of the higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 tev," *European Physical Journal C*, vol. 75, 5 2015. [Online]. Available: http://arxiv.org/abs/1412.8662http://dx.doi.org/10.1140/epjc/s10052-015-3351-7

[10] M. M. Mühlleitner, "Determination of higgs spin and parity at the lhc."

[11] S. Y. Choi and J. H. Jeong, "Selection rules for the decay of a particle into two identical massless particles of any spin," *Physical Review D*, vol. 103, 2 2021. [Online]. Available: http://arxiv.org/abs/2102.11440http://dx.doi.org/10.1103/PhysRevD.103.096013

[12] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, T. Bergauer, M. Dragicevic, J. Erö, A. Valle, M. Flechl, R. Frühwirth, M. Jeitler, N. Krammer, I. Krätschmer, D. Liko, T. Madlener, I. Mikulec, N. Rad, J. Schieck, S. Xie, and L. Finco, "Pileup mitigation at cms in 13 tev data," *Journal of Instrumentation*, vol. 15, pp. P09 018–P09 018, 09 2020.

[13] The-CMS-collaboration, "Studies of tracker material." [Online]. Available: https://cds.cern.ch/record/1279138/files/TRK-10-003-pas.pdf

[14] J. Butler, D. Contardo, M. Klute, J. Mans, L. Silvestris, D. Abbaneo, A. Ball, P. Bloch, O. Buchmueller, F. Cavallari, A. Dabrowski, P. D. Barbaro, I. Fisk, M. Girone, F. Hartmann, J. Hauser, C. Jessop, D. Lange, F. Meijers, C. Seez, W. Smith, and S. Cittolin, "Technical proposal for the phase-ii upgrade of the compact muon solenoid."

[15] The-CMS-collaboration, "Cms physics: Technical design report volume 1: Detector performance and software," 2006.

[16] M. Maggiore, "A modern introduction to quantum field theory."

[17] The-CMS-collaboration, "Constraints on anomalous hvv couplings from the production of higgs bosons decaying to lepton pairs," *Physical Review D*, vol. 100, 12 2019.

[18] N. A. et al, "Constraints on anomalous higgs boson couplings in production and decay with the full run-2 h $\rightarrow$ 4 dataset."

[19] A. V. Gritsan and et al, "New features in the jhu generator framework: constraining higgs boson properties from on-shell and off-shell production," *Physical Review D*, vol. 102, 2 2020. [Online]. Available: http://arxiv.org/abs/2002.09888http://dx.doi.org/10.1103/PhysRevD.102.056022

[20] The-CMS-collaboration, "Electron and photon reconstruction and identification with the cms experiment at the cern lhc," *Journal of Instrumentation*, vol. 16, 12 2020. [Online]. Available: http://arxiv.org/abs/2012.06888http://dx.doi.org/10.1088/1748-0221/16/05/P05014

[21] M. Cacciari, G. P. Salam, and G. Soyez, "Fastjet user manual," 2011.

[22] The-CMS-collaboration, "Pileup mitigation at cms in 13 tev data," *Journal of Instrumentation*, vol. 15, 3 2020. [Online]. Available: http://arxiv.org/abs/2003.00503http://dx.doi.org/10.1088/1748-0221/15/09/P09018

[23] M. Carena, C. Grojean, M. Kado, and V. Sharma, "Status of higgs boson physics," 1905.

[24] K. Albertsson, S. Gleyzer, A. H. L., Moneta, and P. Speckmayer, "Tmva 4: Toolkit for multivariate data analysis with root," 2020.

[25] "https://programmathically.com/understanding-backpropagation-with-gradient-descent."

[26] S. An, "Optimising variable selection for machine learning analysis in atlas tth search," 2017.

[27] D. Rainwater, R. Szalapski, and D. Zeppenfeld, "Probing color-singlet exchange in $z + 2$-jet events at the lhc," *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 54, pp. 6680–6689, 5 1996. [Online]. Available: http://arxiv.org/abs/hep-ph/9605444http://dx.doi.org/10.1103/PhysRevD.54.6680

[28] S. Sumowidagdo, "A pedagogical derivation of the matrix element method in particle physics data analysis," *Journal of Physics: Conference Series*, vol. 988, p. 012003, 3 2018. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/988/1/012003https://iopscience.iop.org/article/10.1088/1742-6596/988/1/012003/meta

[29] The-CMS-collaboration, "Constraints on anomalous higgs boson couplings to vector bosons and fermions in its production and decay using the four-lepton final state," *Physical Review D*, vol. 104, 4 2021. [Online]. Available: http://arxiv.org/abs/2104.12152http://dx.doi.org/10.1103/PhysRevD.104.052004

[30] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, "Handling uncertainties in background shapes: the discrete profiling method," *Journal of Instrumentation*, vol. 10, pp. P04 015–P04 015, 8 2014. [Online]. Available: http://arxiv.org/abs/1408.6865http://dx.doi.org/10.1088/1748-0221/10/04/P04015