

Anomalous Higgs boson couplings to vector bosons in Vector-Boson-Fusion production with CMS experiment

Facoltà di Scienze Matematiche, Fisiche e Naturali
Dipartimento di Fisica
Corso di Laurea Magistrale in Fisica

Davide Porzio
Matricola 1560819

Relatore
Dr. Emanuele Di Marco

A.A. 2020-2021

Contents

1	Theory	4
1.1	Standard Model	4
1.2	Higgs mechanism	6
1.3	Higgs boson production and decay processes	8
1.4	Higgs boson discovery at LHC	10
1.5	CP violation	10
1.6	Anomalous Higgs boson couplings to vector bosons	11
2	LHC and CMS	15
2.1	LHC	15
2.2	CMS	18
2.2.1	Magnet	20
2.2.2	Inner tracking system	20
2.2.3	Electromagnetic calorimeter	22
2.2.4	Hadronic calorimeter	27
2.2.5	Muon detector	28
2.2.6	Trigger system	28
3	Event selection	31
3.1	Signal characteristics	31
3.2	Backgrounds	32
3.3	Measurement of the mass of the Higgs boson	33
3.4	Photon reconstruction in the ECAL	34
3.4.1	Superclustering in the ECAL	35
3.4.2	Energy corrections	36
3.4.3	Reconstruction of converted photons	37
3.5	Reconstruction and identification of jets	38
3.6	Kinematic variables	38
4	Multivariate analysis	42
4.1	MVA selection based on SM-only hypothesis	42
4.2	DNN training	44
4.3	Event selection optimization	48
4.4	DNN variables	51
5	Comparison between data and Monte Carlo	52
5.1	Diphoton loose control sample	52
5.2	$Z \rightarrow e^+e^- + 2$ jets control sample	56

6	Statistical signal extraction	62
6.1	Invariant mass fit	62
6.2	Statistical results with sPlots	66

Chapter 1

Theory

1.1 Standard Model

The Standard Model (SM) of particle physics is a theory that describes the elementary constituents of matter and three of their four interactions: the electromagnetic, the weak and the strong interactions. It was formulated in the '60 by Glashow, Salam and Weinberg [1–3]. It is a quantum field theory which is renormalizable and compatible with the special theory of relativity. Its Lagrangian has a gauge symmetry of the group $SU(3) \times SU(2) \times U(1)$.

We can divide the SM in two sectors: the Quantum Chromodynamics (QCD) sector and the electroweak sector. Therefore the Lagrangian of the SM can be written as the sum of two terms:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} \quad (1.1)$$

The QCD describes the strong interactions of quarks and gluons, which are mediated by the exchange of a color charge. The Lagrangian of the strong sector has a $SU(3)_C$ symmetry of color and is given by:

$$\mathcal{L}_{QCD} = -\frac{1}{4} \sum_i F_{\mu\nu}^i F^{i\mu\nu} + i \sum_r \bar{q}_{r\alpha} \gamma^\mu D_{\mu\beta}^\alpha q_r^\beta \quad (1.2)$$

where $F_{\mu\nu}^i$ are the tensors of the 8 gluon fields G_μ^i ($i = 1, \dots, 8$) which are given by:

$$F_{\mu\nu}^i = \partial_\mu G_\nu^i - \partial_\nu G_\mu^i - g_F f^{ijk} G_\mu^j G_\nu^k \quad (1.3)$$

where g_F is the coupling constant of the strong interactions and f^{ijk} are the structure constants of the group $SU(3)$. In Eq. (1.2), $q_{r\alpha}$ is the quark field of flavor r and color α , and the covariant derivative $D_{\mu\beta}^\alpha$ is given by:

$$D_{\mu\beta}^\alpha = \partial_\mu \delta_\beta^\alpha + \frac{i}{2} g_F \sum_i G_\mu^i \lambda_\beta^{i\alpha} \quad (1.4)$$

where $\lambda_\beta^{i\alpha}$ are the matrices of the generators of $SU(3)$.

The electroweak sector is described by a Lagrangian which is invariant under gauge transformations of the symmetry group $SU(2)_W \times U(1)_Y$. The group

$SU(2)_W$ describes the weak isospin charge I , and the $U(1)_Y$ group describes the weak hypercharge Y .

The fermions, which are leptons and quarks, are grouped in three families. Their left handed components are grouped in weak isospin doublets with $I = 1/2$, and their right handed components are grouped in weak isospin singlets with $I = 0$:

$$I = \frac{1}{2}$$

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$$

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \quad \begin{pmatrix} c \\ s \end{pmatrix}_L \quad \begin{pmatrix} t \\ b \end{pmatrix}_L \quad (1.5)$$

$$I = 0$$

$$(e)_R \quad (\mu)_R \quad (\tau)_R$$

$$(u)_R \quad (c)_R \quad (t)_R$$

$$(d)_R \quad (s)_R \quad (b)_R \quad (1.6)$$

The existence of these gauge symmetries introduces four vector bosons, three for the group $SU(2)_W$, the fields W_μ^i ($i = 1, 2, 3$), and one for the group $U(1)_Y$, the field B_μ . The physical fields are given by linear combinations of these fields:

$$A_\mu = \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \quad (1.7)$$

$$Z_\mu = \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu \quad (1.8)$$

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \quad (1.9)$$

Eq. (1.7) and (1.8) describe two neutral particles, the photon and the Z boson, and Eq. (1.9) describes two charged particles, the W^\pm bosons. θ_W is called Weinberg angle.

In this way the theory is symmetric under local gauge transformations and is described by the following Lagrangian:

$$\mathcal{L}_{EW} = \mathcal{L}_{fermions} + \mathcal{L}_{gauge\ bosons} = i \sum_f \bar{f} D_\mu \gamma^\mu f - \frac{1}{4} \sum_G F_{\mu\nu}^G F^{G\mu\nu} \quad (1.10)$$

where the sums run over all the fermionic fields f and over all the gauge bosons fields G . The covariant derivative is given by:

$$D_\mu = \partial_\mu - ig_G (\lambda^\alpha G_\alpha)_\mu \quad (1.11)$$

where g_G is the generic coupling constant of the fermion to the gauge boson field G , and λ^α are the generators of the symmetry group to which G is related.

All the particles described in this theory are massless, so at this level the theory cannot describe reality. On the other hand, explicit mass terms in the Lagrangian break the gauge symmetries on which the model is based. The Higgs mechanism solves this problem introducing the spontaneous symmetry breaking [4].

1.2 Higgs mechanism

A Lagrangian is said to be symmetric when it is invariant under a group of transformations. Degenerate eigenstates of a symmetric Lagrangian, when they undergo these kind of transformations, are transformed in linear combinations of themselves. If the Lagrangian has many degenerate fundamental states, by choosing just one of them, this will not be invariant under the gauge transformation under which the Lagrangian is symmetric. This kind of symmetry breaking is called spontaneous symmetry breaking.

The simplest way to break spontaneously the symmetry $SU(2)_W \times U(1)_Y$ is to introduce a scalar field ϕ which is an isospin doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \frac{\phi_1 + i\phi_2}{\sqrt{2}} \\ \frac{\phi_3 + i\phi_4}{\sqrt{2}} \end{pmatrix} \quad (1.12)$$

where ϕ_i ($i = 1, 2, 3, 4$) are real fields and ϕ^+ and ϕ^0 are complex fields. The simplest Lagrangian for the Higgs scalar field has the form:

$$\mathcal{L}_H = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \quad (1.13)$$

where $V(\phi)$ is given by:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (1.14)$$

and the covariant derivative D_μ is given by:

$$D_\mu = \partial_\mu + \frac{i}{2} g_W \sigma_j W_j^\mu + i g_Y Y B_\mu \quad (1.15)$$

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^μ and B_μ , σ_j are the Pauli matrices and Y is the weak hypercharge.

The potential $V(\phi)$ depends on the parameters μ^2 and λ . By requesting that $\lambda > 0$ we ensure that the energy is bounded from below and therefore that there exists a stable fundamental state. By requesting that $\mu^2 < 0$ the symmetry can be spontaneously broken. In fact, in correspondence of:

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2} \quad (1.16)$$

the potential $V(\phi)$ has a minimum, which implies that the field ϕ has a non null expectation value on the vacuum $\langle \phi_0 \rangle = v/\sqrt{2}$.

In Fig. 1.1, we show the shape of the Higgs field potential $V(\phi)$ with $\mu^2 < 0$ in the $(Re \phi, Im \phi)$ plane.

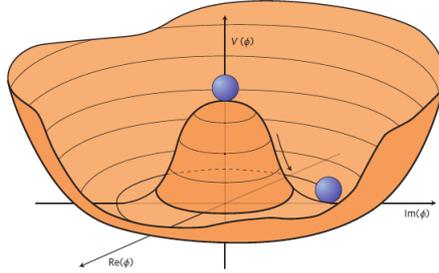


Figure 1.1: Higgs field potential $V(\phi)$ with $\mu^2 < 0$ in the $(\text{Re } \phi, \text{Im } \phi)$ plane.

We need to use perturbation theory to expand ϕ around its fundamental state. This will have to be chosen between all the states that satisfy Eq. (1.16), and each of these will break at least one symmetry of the Lagrangian. This implies the assignment of a mass to each boson connected to a broken symmetry.

The electroweak Lagrangian has a $SU(2)_W \times U(1)_Y$ symmetry. To keep the photon massless we need to choose a fundamental state that preserves the symmetry $U(1)_Q$ of electric charge. From the Gell-Mann Nishijima relation:

$$Q = I_3 + \frac{Y}{2} \quad (1.17)$$

we obtain this by choosing a fundamental state ϕ_0 which is an isospin doublet with $I = 1/2$ and which is a hypercharge singlet with $Y = 1$:

$$\phi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \quad (1.18)$$

The field ϕ will then be an expansion around the fundamental state ϕ_0 :

$$\phi(x) = \begin{pmatrix} 0 \\ \frac{v+h(x)}{\sqrt{2}} \end{pmatrix} \quad (1.19)$$

In this way the bosonic fields W^\pm and Z acquire mass, while the photon stays massless:

$$m_W = \frac{g_W v}{2} \quad (1.20)$$

$$m_Z = \frac{v \sqrt{g_W^2 + g_Y^2}}{2} \quad (1.21)$$

$$m_\gamma = 0 \quad (1.22)$$

The W and the Z bosons have been discovered at the UA1 experiment at CERN, and their mass has been precisely measured at the Large Electron Positron Collider (LEP), Tevatron and at the Large Hadron Collider (LHC). The current best values of their mass are $m_W = 80.379 \pm 0.012$ GeV, $m_Z = 91.1876 \pm 0.0021$ GeV.

In the theory we also obtain a physical particle, the Higgs boson, described by the field $h(x)$, with a mass given by:

$$m_H = \sqrt{2\lambda}v \quad (1.23)$$

The SM does not predict the value of the mass of the Higgs boson.

1.3 Higgs boson production and decay processes

At the CERN's LHC, the highest Higgs boson production cross section is achieved in the gluon gluon fusion (ggH), where the Higgs boson is produced through a loop, created by two gluons, where there are mostly top quarks, since the coupling of a fermion to the Higgs boson is proportional to the mass of the fermion, and since the top is the heaviest of the fermions [5].

The second largest production process is vector boson fusion (VBF), where two quarks emit a vector boson (a W or a Z) which fuse together in order to produce a Higgs boson.

Other production processes with lower cross section are the Higgs-strahlung, where two quarks go in a vector boson which subsequently emits a Higgs boson, and the associated production with two top quarks ($t\bar{t}H$), where two gluons emit two top quarks and a Higgs boson.

In Fig. 1.2 we show the Higgs boson production cross sections at LHC as a function of the Higgs boson mass M_H , while in Fig.1.3 we show the Higgs boson production Feynman diagrams.

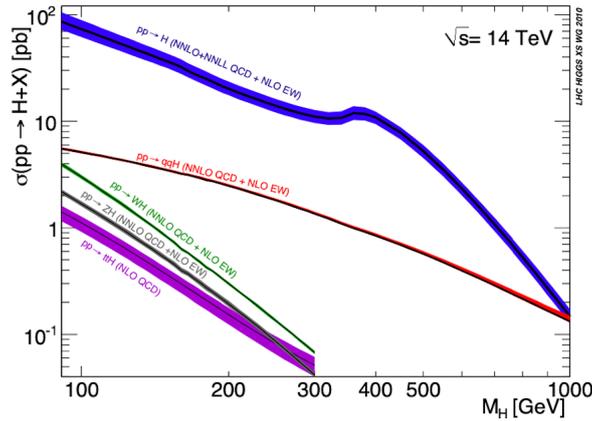


Figure 1.2: Higgs boson production cross sections at LHC at a center of mass energy $\sqrt{s} = 14$ TeV as a function of the Higgs boson mass M_H .

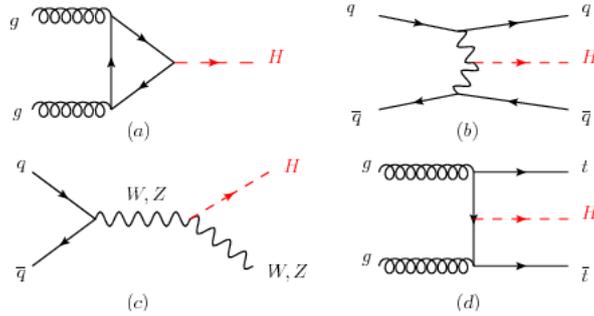


Figure 1.3: Higgs boson production Feynman diagrams: (a) ggH ; (b) VBF; (c) Higgs-strahlung, also known as associated WH and ZH productions; (d) $t\bar{t}H$.

In Fig. 1.4 we show the Higgs boson decay branching ratios as a function of M_H .

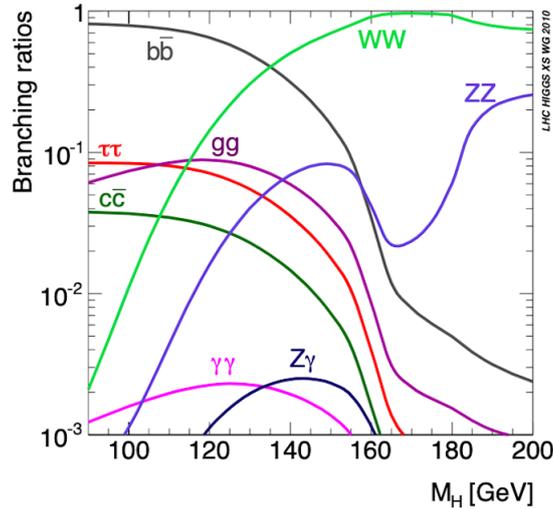


Figure 1.4: Higgs boson decay branching ratios as a function of M_H .

For $m_H \sim 125$ GeV, the highest Higgs boson branching ratio (BR) is achieved in the $b\bar{b}$ decay mode, as can be seen from Fig. 1.4, however this is the most experimentally problematic channel, because of the large QCD background.

Then we have the $H \rightarrow WW$ and $H \rightarrow ZZ$ decay modes. The channels with higher statistics are those where each W and Z decay to $2q$. However, these channels are problematic because of the presence of the QCD interactions given by the quarks that hadronize in jets. Therefore we look at the decay channels of the Z and of the W into leptons, which in a hadronic collider constitute the most experimentally accessible decay modes. A very clean decay channel of the Higgs boson is in $ZZ \rightarrow 4l$. Another important decay mode is in $WW \rightarrow 2l2\nu$, which is an experimentally challenging channel, because we do not fully reconstruct the final state, since we can only obtain the sum of the four-momenta of the two neutrinos.

The $\gamma\gamma$ decay mode is rare, in fact it has a lower branching ratio (of about $2 \cdot 10^{-3}$ for $m_H \sim 125$ GeV) respect to the WW and ZZ decay channels. However, its yield is comparable to the ones of the cleanest and most experimentally accessible channels of the WW and ZZ decays into two or four leptons. In fact, the full BRs are given by the product of the branching ratio of the Higgs boson that goes into WW and ZZ and the branching ratio of the two vector bosons that go into two or four leptons. The $\gamma\gamma$ process is also quite pure, since there are few other processes in the SM that produce only two photons in the final state. Furthermore, the region around $m_H = 125$ GeV is where the $\gamma\gamma$ branching ratio is higher, which makes this one of the most sensitive channels for a Higgs boson of about $m_H = 125$ GeV.

1.4 Higgs boson discovery at LHC

The ATLAS and CMS experiments at LHC announced in 2012 the discovery of a boson of mass around 125 GeV and excluded the Higgs boson of the SM in all the range up to 600 GeV except in a small window of mass around the observed peak. Both experiments used about 5 fb^{-1} of data collected in 2011 at $\sqrt{s} = 7$ TeV and about 20 fb^{-1} of data collected in 2012 at $\sqrt{s} = 8$ TeV.

ATLAS presented the results of the analysis of the $ZZ, WW, b\bar{b}$ and $\tau^+\tau^-$ channels at $\sqrt{s} = 7$ TeV in 2011 combined with the analysis of the $ZZ \rightarrow 4l, \gamma\gamma$ and $WW \rightarrow e\nu_e\mu\nu_\mu$ channels at $\sqrt{s} = 8$ TeV in 2012. ATLAS observed a neutral boson of mass $126.0 \pm 0.4(stat) \pm 0.4(syst)$ GeV with a significance of 5.9σ which corresponds to a probability of $1.82 \cdot 10^{-9}$ that the signal observed was due to a background fluctuation.

CMS presented the combined results between 2011 and 2012 of 5 decay channels, given by $\gamma\gamma, ZZ, WW, \tau^+\tau^-$ and $b\bar{b}$ and observed an excess of events with a significance of 5.0σ around 125 GeV which corresponds to a probability of $2.87 \cdot 10^{-7}$ that the signal observed was due to a background fluctuation. The excess is more significant combining only the most sensitive channels which are $\gamma\gamma$ and ZZ . A mass fit gives the result $125.3 \pm 0.4(stat) \pm 0.5(syst)$ GeV.

The discovery by the ATLAS and CMS Collaborations of a Higgs boson with a mass of about 125 GeV [6–8] has confirmed the predictions of the SM of particle physics [1, 2, 4, 9–11]. Since then, precise measurements have been made of the cross section, the mass, the branching ratios in various decay channels, including $b\bar{b}, \tau^+\tau^-$ and $\mu^+\mu^-$, with the Run2 at LHC, with a total integrated luminosity of 156 fb^{-1} .

1.5 CP violation

In addition to continuous gauge groups, there are a few discrete symmetries of interest, for example parity P and charge conjugation C .

The parity operation transforms a system into its mirror image, so that a particle with momentum \vec{p} is transformed into a particle with momentum $-\vec{p}$ under parity. This symmetry is preserved under electromagnetic, strong and gravity interactions, and was thus hypothesized to be conserved in every interaction. However in 1957 Wu, Ambler, Hayward, Hoppes and Hudson discovered that parity is violated in ^{60}Co β decays, and thus parity is not a fundamental

symmetry of nature.

The charge conjugation is the operation that transforms a particle into its anti-particle.

It was proposed that the combination of charge conjugation and parity, namely the CP operation, might be a fundamental symmetry of nature. However in 1964 Christenson, Cronin, Fitch and Turlay discovered evidence that CP might be violated in the kaon system. CP violation is further confirmed by many collider experiments. It is also one of the necessary conditions for the observed matter-antimatter asymmetry in the universe, as pointed out by Sakharov. However the amount of CP violation discovered in the SM cannot explain the large baryon anti-baryon asymmetry in the observable universe. Additional sources of CP violation are needed, which makes the measurement of CP properties especially interesting.

1.6 Anomalous Higgs boson couplings to vector bosons

The CMS [12–18] and ATLAS [19–24] experiments have set constraints on the spin-parity properties of the Higgs boson and anomalous HVV couplings, where V stands for W , Z , and γ electroweak gauge bosons, finding its quantum numbers to be consistent with $J^{PC} = 0^{++}$, but leaving room for small anomalous HVV couplings [25]. Anomalous stands in this case for an effect not predicted by the SM. In theories beyond the SM (BSM), Higgs boson interactions may generate several of them, which lead to new tensor structures of interactions, both CP -even and CP -odd. Possible CP violation effects in couplings to fermions, Hff , had not been experimentally probed until recently, when the first constraints were reported by CMS [26] and ATLAS [27] in $t\bar{t}H$ production using the $H \rightarrow \gamma\gamma$ channel.

The scattering amplitude describing the interaction between a spin-zero H and two spin-one gauge bosons VV , such as ZZ , $Z\gamma$, $\gamma\gamma$, or WW , includes only three independent invariant tensor structures with the coupling parameters a_i^{VV} that can have both real and imaginary parts and in general are form factors which can depend on the squared Lorentz invariant four-momenta of V_1 and V_2 , $q_{V_1}^2$ and $q_{V_2}^2$ [14]. In the following, the terms up to q_V^2 are kept in the expansion under the assumption of small contributions from anomalous couplings so that the scattering amplitude becomes:

$$A(HVV) \sim [a_1^{VV} + \frac{k_1^{VV} q_{V_1}^2 + k_2^{VV} q_{V_2}^2}{(\Lambda_1^{VV})^2}] m_{V_1}^2 \epsilon_{V_1}^* \epsilon_{V_2}^* + \quad (1.24)$$

$$+ a_2^{VV} f_{\mu\nu}^{*(1)} f^{*(2)\mu\nu} + a_3^{VV} f_{\mu\nu}^{*(1)} \tilde{f}^{*(2)\mu\nu} \quad (1.25)$$

where $f^{(i)\mu\nu} = \epsilon_{V_i}^\mu q_{V_i}^\nu - \epsilon_{V_i}^\nu q_{V_i}^\mu$ is the field strength tensor of a gauge boson with momentum q_{V_i} and polarization vector ϵ_{V_i} , $\tilde{f}_{\mu\nu}^{(i)} = (1/2)\epsilon_{\mu\nu\rho\sigma} f^{(i)\rho\sigma}$ is the dual field strength tensor defined with the aid of the Levi-Civita symbol in four dimensions $\epsilon_{\mu\nu\rho\sigma}$ with $\epsilon_{0123} = -1$ and totally antisymmetric with respect to all pairs of indices, m_{V_1} is the pole mass of the Z or W vector boson, and Λ_1 is the scale of BSM physics and is a free parameter of the model. A different coupling

in the scattering amplitude in Eq. (1.24) and (1.25) typically leads to changes of the kinematic distributions of the process.

In Eq. (1.24) and (1.25), VV stands for $ZZ, Z\gamma, \gamma\gamma$, and WW . The tree-level SM-like contribution corresponds to $a_1^{ZZ}, a_1^{WW} \neq 0$, while there is no tree-level coupling to massless gauge bosons, that is $a_1^{VV} = 0$ for $Z\gamma$, and $\gamma\gamma$. Small values of the other couplings can be generated through loop effects in the SM. The other terms can be ascribed to anomalous couplings which are related to the vertices $HZZ, HWW, HZ\gamma$, and $H\gamma\gamma$.

The couplings to which we are interested are the anomalous couplings of the photons $a_2^{\gamma\gamma}$ and $a_3^{\gamma\gamma}$. In the SM, $a_2^{\gamma\gamma}$ and $a_3^{\gamma\gamma}$ are very small. Therefore measuring values significantly different from zero (of the order of the 10%) of these anomalous couplings would constitute an indication of new physics. This would mean that in the virtual loops in addition to SM particles there are also new particles that have phases different from 0 that can generate CP violation. The purpose of this thesis is to constrain a_2 and a_3 in the VBF production $VV \rightarrow H$ and using the $\gamma\gamma$ final state. The anomalous couplings can be inferred by the cross sections that we can experimentally measure.

Considerations of symmetry and gauge invariance require $k_1^{ZZ} = k_2^{ZZ} = -e^{i\phi_{\Lambda_1}^{ZZ}}, k_1^{\gamma\gamma} = k_2^{\gamma\gamma} = 0, k_1^{Z\gamma} = 0$, and $k_2^{Z\gamma} = -e^{i\phi_{\Lambda_1}^{Z\gamma}}$. $\phi_{\Lambda_1}^{VV}$ is the phase of the anomalous coupling with Λ_1^{VV} , which is either 0 or π for real couplings. In the following, the ZZ labels for the ZZ interactions will be omitted, and therefore the couplings a_1, a_2, a_3 , and Λ_1 are not labeled explicitly with a ZZ superscript.

The parity-conserving interaction of a pseudoscalar (CP -odd state) corresponds to the a_3^{VV} terms, while the other terms describe the parity-conserving interaction of a scalar (CP -even state). The a_3^{VV} terms appear in the SM only at a three-loop level and receive a small contribution. The a_2^{VV} and Λ_1^{VV} terms appear in loop-induced processes and give small contributions $O(10^{-3} - 10^{-2})$. The dominant contributions to the SM expectation of the $H \rightarrow Z\gamma$ and $\gamma\gamma$ decays are $a_2^{Z\gamma}$ and $a_2^{\gamma\gamma}$, which are predicted to be $a_2^{Z\gamma} \sim -0.007$ and $a_2^{\gamma\gamma} \sim 0.004$. Anomalous couplings may be enhanced with BSM contributions and generally acquire a non-trivial dependence on Lorentz invariant quantities and become complex. When the particles in the loops responsible for these couplings are heavy in comparison to the Higgs boson mass parameters, the couplings are real.

Under the assumption that the couplings are constant and real, the scattering amplitude formulation is equivalent to an effective Lagrangian for the $HZZ, HWW, HZ\gamma$, and $H\gamma\gamma$ interactions:

$$L(HVV) \sim a_1 \frac{m_Z^2}{2} H Z^\mu Z_\mu - \frac{k_1}{(\Lambda_1)^2} m_Z^2 H Z_\mu \square Z^\mu - \quad (1.26)$$

$$- \frac{1}{2} a_2 H Z^{\mu\nu} Z_{\mu\nu} - \frac{1}{2} a_3 H Z^{\mu\nu} \tilde{Z}_{\mu\nu} + a_1^{WW} m_W^2 H W^{+\mu} W_\mu^- \quad (1.27)$$

$$- \frac{1}{(\Lambda_1^{WW})^2} m_W^2 H (k_1^{WW} W_\mu^- \square W^{+\mu} + k_2^{WW} W_\mu^+ \square W^{-\mu}) - \quad (1.28)$$

$$- a_2^{WW} H W^{+\mu\nu} W_{\mu\nu}^- - a_3^{WW} H W^{+\mu\nu} \tilde{W}_{\mu\nu}^- + \quad (1.29)$$

$$+ \frac{k_2^{Z\gamma}}{(\Lambda_1^{Z\gamma})^2} m_Z^2 H Z_\mu \partial_\nu F^{\mu\nu} - a_2^{Z\gamma} H F^{\mu\nu} Z_{\mu\nu} - a_3^{Z\gamma} H F^{\mu\nu} \tilde{Z}_{\mu\nu} \quad (1.30)$$

$$- \frac{1}{2} a_2^{\gamma\gamma} H F^{\mu\nu} F_{\mu\nu} - \frac{1}{2} a_3^{\gamma\gamma} H F^{\mu\nu} \tilde{F}_{\mu\nu} \quad (1.31)$$

where H is the Higgs 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 $\tilde{V}_{\mu\nu} = (1/2)\epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}$.

The scenarios are parametrized in terms of the effective fractional cross sections f_{a_i} and their phases ϕ_{a_i} with respect to the two dominant tree-level couplings a_1 and a_1^{WW} in the $H \rightarrow VV \rightarrow 4l$ and $H \rightarrow WW \rightarrow l\nu_l l\nu_l$ processes, respectively.

The effective fractional ZZ cross sections f_{a_i} and phases ϕ_{a_i} are defined as follows:

$$f_{\Lambda_1} = \frac{\frac{\tilde{\sigma}_{\Lambda_1}}{(\Lambda_1)^4}}{|a_1|^2 \sigma_1 + |a_2|^2 \sigma_2 + |a_3|^2 \sigma_3 + \frac{\tilde{\sigma}_{\Lambda_1}}{(\Lambda_1)^4} + \dots}, \quad \phi_{\Lambda_1} \quad (1.32)$$

$$f_{a_2} = \frac{|a_2|^2 \sigma_2}{|a_1|^2 \sigma_1 + |a_2|^2 \sigma_2 + |a_3|^2 \sigma_3 + \frac{\tilde{\sigma}_{\Lambda_1}}{(\Lambda_1)^4} + \dots}, \quad \phi_{a_2} = \arg \frac{a_2}{a_1} \quad (1.33)$$

$$f_{a_3} = \frac{|a_3|^2 \sigma_3}{|a_1|^2 \sigma_1 + |a_2|^2 \sigma_2 + |a_3|^2 \sigma_3 + \frac{\tilde{\sigma}_{\Lambda_1}}{(\Lambda_1)^4} + \dots}, \quad \phi_{a_3} = \arg \frac{a_3}{a_1} \quad (1.34)$$

where σ_i is the cross section of the process corresponding to $a_i = 1, a_{j \neq i} = 0$, while $\tilde{\sigma}_{\Lambda_1}$ is the effective cross section of the process corresponding to $\Lambda_1 = 1$ TeV, given in units of $\text{fb} \cdot \text{TeV}^4$. The effective fractional WW and $\gamma\gamma$ cross sections are defined in complete analogy with the definitions for ZZ as shown in Eq. (1.32), (1.33) and (1.34). In Eq. (1.32), (1.33) and (1.34) the f_{a_i} parameters are bounded between 0 and 1.

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:

$$\frac{|a_i|}{|a_1|} = \sqrt{\frac{f_{a_i}}{f_{a_1}}} \sqrt{\frac{\sigma_1}{\sigma_i}} \quad (1.35)$$

$$\Lambda_1 = \frac{1}{\sqrt{|a_1|}} \left(\frac{f_{a_1}}{f_{\Lambda_1}} \right)^{\frac{1}{4}} \left(\frac{\tilde{\sigma}_{\Lambda_1}}{\sigma_1} \right)^{\frac{1}{4}} \quad (1.36)$$

where the fraction $f_{a_1} = 1 - f_{\Lambda_1} - f_{a_2} - f_{a_3} - \dots$ corresponds to the effective SM tree-level contribution, which is expected to dominate.

The signal Monte Carlo samples used in this thesis are generated with fixed values of f_{a_i} : one can cover the whole phase space with linear combinations of these samples.

Chapter 2

LHC and CMS

2.1 LHC

The LHC is an accelerator where beams of protons collide at a center of mass energy \sqrt{s} up to 14 TeV. It is the largest and the most energetic accelerator ever built up to now. It is built in a circular tunnel of about 27 km about 100 m underground, at the border between France and Switzerland, near Geneva. The same ring in the past hosted the LEP.

The LHC injection complex is indicated in Fig. 2.1. After the production of the protons, the Linear Accelerator 2 (LINAC2) accelerates them to 50 MeV, then they are brought in the first ring, the Proton Synchrotron (PS) Booster, at 1.4 GeV, then they are inserted in the Proton Synchrotron at 26 GeV and in the Super Proton Synchrotron at 450 GeV. Finally they are injected in the main ring where they can reach a center of mass energy up to 14 TeV (7 TeV per beam), as of the design value that will be reached in the future. Until now, there have been two Runs, at three different center of mass energies:

- Run1, in 2011, at 7 TeV;
- Run1, in 2012, at 8 TeV;
- Run2, from 2015 to 2018, at 13 TeV.

In March 2022 Run3 will begin, at a center of mass energy of 13.5 TeV.

The LHC injection complex

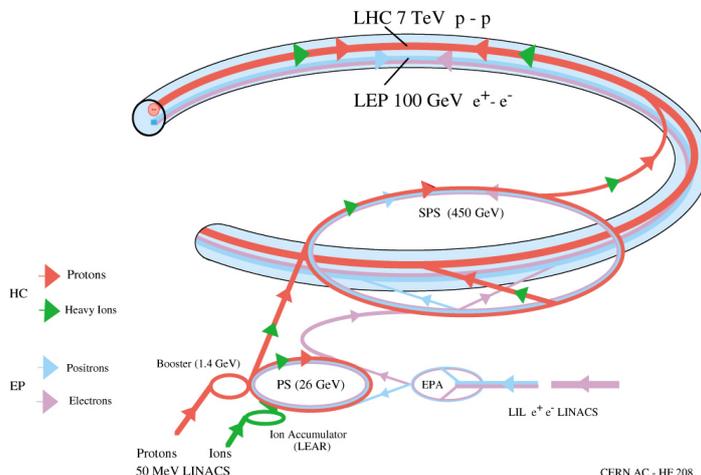


Figure 2.1: LHC injection complex.

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	8.3 T
Magnet temperature	1.9 K
Beam energy	7 TeV
\sqrt{s}	14 TeV
\mathcal{L}	$2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
Number of protons per bunch	$1.05 \cdot 10^{11}$
Bunch length σ_z	75 mm
Bunch radius $\sigma_x = \sigma_y$	16 μm
Number of bunches	2808
Space - time between bunches	7.48 m - 25 ns

Table 2.1: Main design LHC characteristics.

Unlike electronic accelerators, such as LEP, a hadronic accelerator allows to produce reactions at different energies. Since the interactions happen at partonic level, the effective center of mass energy is that of the interacting partons, and depends on their probability density functions. Therefore the reactions happen at lower energies in the center of mass respect to the interacting protons. However, 14 TeV are sufficient to reach the TeV energy scale.

At these energies, the contribution of the sea quarks becomes important, therefore the collision schemes proton-proton and proton-antiproton are basi-

cally equivalent.

In an electronic accelerator, the interacting energy is precisely known, and the cross sections have low values, therefore precision measurements can be made even with a small number of events. Instead, in a hadronic accelerator, the cross sections are much higher, therefore it is necessary to collect a large number of events in order to limit the background contribution. For this reason, an important characteristic of LHC is the high value of the luminosity.

The main purpose of LHC was to study the electroweak spontaneous symmetry breaking, looking for a Higgs boson compatible with the SM. Other aims are to find evidences of physics BSM, like Supersymmetries or the extra-dimensions, and to try to explain the nature of the dark matter and of the quark-gluon plasm. Moreover LHC tests the SM with precision measurements and studies the violation of the CP symmetry.

Four detectors study the LHC interactions, which are situated in the four crossing points of the beams. CMS (Compact Muon Solenoid) and ATLAS (A Toroidal LHC ApparatuS) are two general purpose experiments, which had as aim the discovery of the Higgs boson and which research physics BSM. LHCb, an experiment dedicated to the b quark physics, studies the CP symmetry violation and rare phenomenons in the physics of beauty flavored hadrons. ALICE (A Large Ion Collider Experiment) studies collisions between heavy ions Pb-Pb at a center of mass energy of 2.76 TeV.

The luminosity of a circular accelerator is:

$$\mathcal{L} = \frac{N_1 N_2 f k}{4\pi\sigma_x\sigma_y} \quad (2.1)$$

where N_1 and N_2 are the number of protons in the two colliding bunches, f is the revolution frequency of the protons in the ring, k is the number of bunches circulating in the ring, σ_x and σ_y are the mean dispersions of the bunches in the x and y directions perpendicular to the orbit.

The number of events is given by:

$$N = \mathcal{L}_{int}\sigma_{pp}A\epsilon \quad (2.2)$$

where \mathcal{L}_{int} is the integrated luminosity, σ_{pp} is the cross section of the pp collision process, A and ϵ are respectively the geometrical acceptance and the efficiency of the detector.

The total cross section σ_{pp}^{tot} is about 100 mb. Therefore there are about 20 proton-proton interactions per bunch crossing and 10^9 interactions per second, which results in a large multiplicity of particles. Together with the crossing rate of 25 ns, this puts very stringent requirements for the LHC detectors:

- a fine granularity, to resolve the large number of particles;
- a high response velocity, to minimize the pileup, which is the overlap of the events;
- a rapid and efficient event selection and acquisition;
- a good radiation resistance;
- a high precision, in order to distinguish the single interaction vertices.

The general purpose detectors such as ATLAS and CMS must have further requirements to fully understand the events:

- hermeticity, which means they must cover the largest possible solid angle to measure accurately the missing transverse energy;
- precise reconstruction of high energy leptons and photons;
- precise measurement of the momentum of the charged particles with a tracking system;
- good reconstruction of the energy of the hadronic particles.

In Fig. 2.2 we show the plot of the total integrated luminosity collected in Run1 and in Run2.

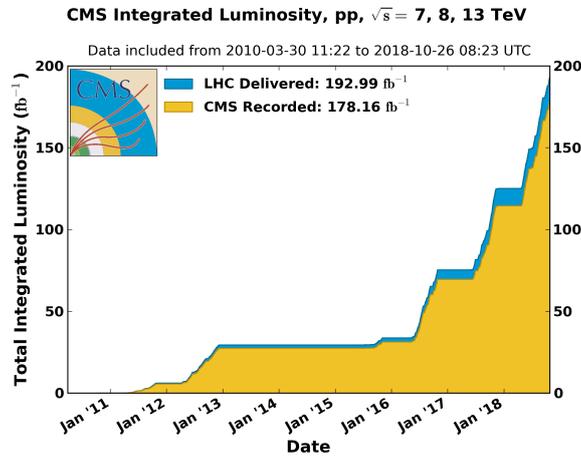


Figure 2.2: Total integrated luminosity collected in Run1 and in Run2 by LHC and CMS.

For the purpose of this thesis, we used the data collected by CMS in 2017, which correspond to an integrated luminosity of 41.5 fb^{-1} .

2.2 CMS

The Compact Muon Solenoid (CMS) experiment was designed to discover the Higgs boson, to search for evidences of new physics at LHC and to make precision measurements of already known processes. To obtain these results at LHC, the Collaboration had the following aims:

- an optimal system for the identification and the measurements of the muons. This brought to the choice of a solenoidal superconducting magnet able to produce a magnetic field of 4 T, where the muon detectors are in the iron return yoke, which allows to have a compact muonic system, precise in the tracks measurement and able to distinguish the sign of the charge for muons of energy up to 1 TeV;

- the best possible electromagnetic calorimeter for the measurements of electrons and photons, which is compatible with the dimensions of the magnet;
- a central tracking system efficient and precise in the measurement of the momentum and in the reconstruction of the vertices and the tracks of the charged particles;
- an adequate hadronic calorimeter, highly hermetic in order to permit a good measurement of the transverse missing energy due to neutrinos or other non detectable particles with trackers and calorimeters (such as new physics particles).

The structure of the CMS detector is shown in Fig. 2.3. It has a cylindrical form, with a radius of 7.5 m, a length of 12 m, and it weighs about 12500 tons. It is divided in a central region, the barrel, closed on both sides by two identical endcaps. Inside it there is the superconducting magnet, which is 13 m long and has a radius of 3 m. The magnet hosts inside of it the inner tracking system and the calorimeters. Therefore the subdetectors are arranged in layers around the axis of the cylinder.

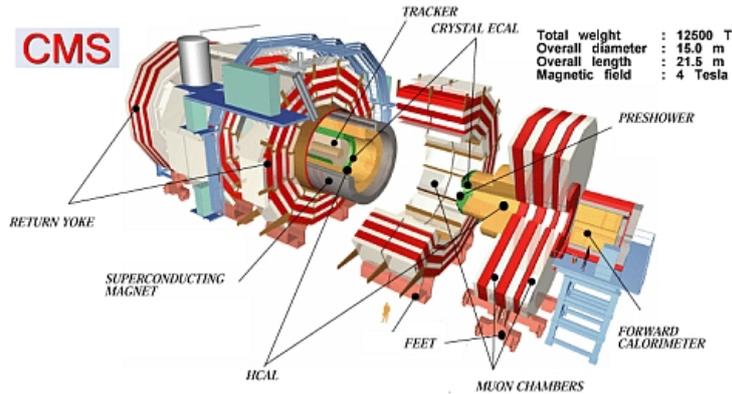


Figure 2.3: The CMS detector.

The coordinate system adopted by CMS has the origin in the nominal interaction point of the beams and adopts the following convention:

- the x axis points radially towards the center of LHC;
- the y axis points upwards;
- the z axis coincides with the symmetry axis of CMS and points towards the direction of the beams of protons.

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:

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

where θ is the polar angle, which is measured from the z axis. For high energies, η it is a good approximation of the rapidity of a particle, defined as:

$$y = \frac{1}{2} \ln \frac{E + p_L}{E - p_L} \quad (2.4)$$

where p_L is the longitudinal momentum, parallel to the beams. In Fig. 2.4 we show the values of the pseudorapidity η corresponding to some polar angles θ .

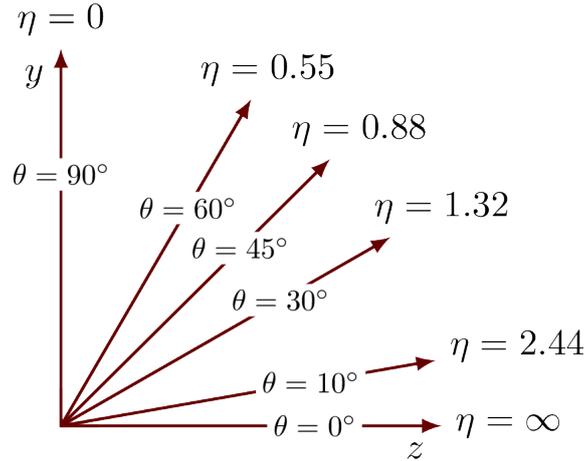


Figure 2.4: Values of the pseudorapidity η corresponding to some polar angles θ .

In the following we will denote with p_T and E_T respectively the momentum and the energy in the transverse plane xy to the beams. The transverse energy is defined as $E_T = E \sin \theta$, where E is the energy deposited in a calorimetric cell and θ is the polar angle of the position of the cell.

The angular distance in $\eta - \phi$ coordinates is defined as:

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} \quad (2.5)$$

2.2.1 Magnet

In order to measure precisely high energy muons it is necessary to have a magnetic field able to curve them enough. In order to do this in a relatively compact space a magnet was designed, which is made of a superconducting solenoid of NiTi of 13 m of length and of 5.9 m of diameter. This magnet generates a magnetic field of 3.8 T at the center, crossed by a current of 18 kA. The total magnetic energy stored is 2.4 GJ. The return yoke of the magnet is made of iron and also constitutes a part of the mechanical structure of the detector.

2.2.2 Inner tracking system

The inner tracking system reconstructs the tracks of the charged particles and measures its momentum. It is composed by detectors made in silicon.

Its main purpose is to reconstruct isolated electrons and muons with high p_T with efficiency greater than 95%, and tracks with high p_T within jets with efficiency greater than 90% in the pseudorapidity range $|\eta| < 2.4$. At the same time there are constraints on the quantity of material to be used, in order not to worsen too much the resolution on the momentum. This happens in a highly radioactive environment with high particle multiplicity, as can be seen in Tab. 2.2, where we report the fluxes of the expected particles for an integrated luminosity of 500 fb^{-1} [28].

Radius (cm)	Radiation dose (kGy)	Flux of charged particles ($\text{cm}^{-2}\text{s}^{-1}$)
4	840	10^8
22	70	$6 \cdot 10^6$
115	1.8	$3 \cdot 10^5$

Table 2.2: Radiation dose and flux of charged particle expected for different radiuses of the barrel of the CMS tracking system, for an integrated luminosity of 500 fb^{-1} .

Referring to Tab. 2.2, we can identify three regions exposed to different radiation doses:

- near the interaction vertex ($r < 20 \text{ cm}$), where the flux of particles is higher, there are silicon pixel detectors, with dimensions of about $100 \times 150 \mu\text{m}^2$;
- in the intermediate region ($20 < r < 55 \text{ cm}$), the flux is low enough to allow the use of silicon microstrips, with cells with minimal dimensions of $10 \text{ cm} \times 80 \mu\text{m}$;
- in the outer region ($r > 55 \text{ cm}$), the flux is low enough to allow the use of larger silicon strips, with cells with maximal dimensions of $25 \text{ cm} \times 180 \mu\text{m}$.

In Fig. 2.5, we show a section of a quarter of the inner tracking system of CMS.

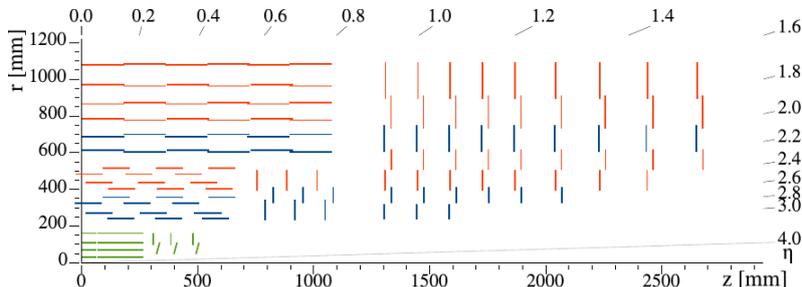


Figure 2.5: Section of a quarter of the inner tracking system of CMS, with some values of the pseudorapidity.

A section of the CMS inner tracking system is shown in Fig. 2.5. Near the interaction vertex there are 3 layers of silicon pixels, respectively at radial

distances of 4.4, 7.3 and 10.2 cm. In the barrel there are the silicon microstrips, for r between 20 and 110 cm. The barrel microstrip region is separated in two regions, one internal and one external, the internal one being shorter than the external one. The internal region is made of four layers and covers the range $20 < r < 55$ cm. The external region is made of six layers and it reaches the radius $r = 110$ cm. The forward region has 2 pixel layers and 9 layers of microstrips in each endcap. There are three additional layers, named inner disks, in the space between the border of the internal region and the endcap, on both sides of the barrel.

The inner tracking system has 66 million silicon pixels and 9.6 million silicon strips. It covers up to $|\eta| < 2.4$. The quantity of material in unity of radiation lengths is a function of η and it varies from a minimum of about 0.5 radiation lengths (at the center of the barrel) to a maximum of 1.8 radiation lengths (in the transition between barrel and endcap). Therefore it is frequent that a photon produces an electron-positron pair or that an electron emits a photon for bremsstrahlung inside the tracker. This must be held into account for a correct reconstruction of the energy of photons and electrons.

The silicon pixels yield a measurement precision of $10 \mu\text{m}$ for the coordinates in the transverse xy plane, and of $20 \mu\text{m}$ for the z coordinate. The microstrips yield a resolution which depends on the width of the cell, which is better than $55 \mu\text{m}$ in the transverse plane.

2.2.3 Electromagnetic calorimeter

The electromagnetic calorimeter of CMS is a hermetic and homogeneous calorimeter, made by 61200 lead tungstate (PbWO_4 , PWO for short) crystals in the barrel region and closed by 7324 crystals in each of the two endcaps. Its aim is to measure with high precision the energy of the electromagnetic particles, other than yield a reconstruction of the energy of the hadronic jets, in combination with the hadronic calorimeter.

The electromagnetic calorimeter played a fundamental role in the research of the Higgs boson, particularly in its decay modes $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4e^\pm/2e^\pm 2\mu^\pm$. In the mass ranges where these channels were studied for its discovery, the intrinsic width of the Higgs boson is small and therefore the experimental resolution is dominant on the uncertainty on the measurement of its mass.

The decay channel in two photons was used as a benchmark in the design of the electromagnetic calorimeter of CMS. The Collaboration chose to use an electromagnetic calorimeter with an excellent energy resolution and with a fine granularity in order to maximize the resolution in the measurement of the invariant mass of the two photons. For this reason it was chosen to use a homogeneous calorimeter, with great precision, made of scintillating crystals.

PbWO₄ crystals

The choice of the PWO was motivated by the compactness, the response rapidity and the radiation resistance of this material.

The main characteristics of PWO, compared with those of other scintillating crystals, are shown in Tab. 2.3.

	NaI (Tl)	BGO	CeF ₃	BaF ₂	PWO
X_0 (cm)	2.57	1.12	1.68	2.05	0.89
ρ (g cm ⁻²)	3.67	7.13	6.16	4.89	8.18
n	1.85	2.20	1.68	1.56	2.29
LY (a.u.)	1	0.15	0.10	0.05 ^f	0.01
				0.20 ^s	
τ (ns)	250	300	10 ^f	0.7 ^f	5 ^f
			30 ^s	620 ^s	15 ^s
λ (nm)	410	480	310 ^f	220 ^f	440 ^f
			340 ^s	310 ^s	480 ^s

Table 2.3: Main characteristics of the PWO crystals and comparison with other scintillating crystals. The apices f and s correspond respectively to the components of principal (fast) and secondary (slow) emission.

The radiation length X_0 represents the longitudinal distance over which an electron traversing a material loses in average $1/e$ of its energy through scattering processes. The electromagnetic calorimeter must ensure the complete containment of the electromagnetic shower until energies of about 1 TeV. For these energies the 98% of the longitudinal development of the shower is contained in $25 X_0$.

The Molière radius R_M is used in the description of the transversal development of the shower. It is defined as:

$$R_M = \frac{21.2 \text{ MeV } X_0}{E_C(\text{MeV})} \quad (2.6)$$

where E_C is the critical energy and it represents the energy at which the mean energy loss for ionization equals the mean energy loss for bremsstrahlung. In average, 90% of the energy of an electromagnetic shower is deposited in a cylinder of radius equal to R_M constructed around the axis of the shower.

The main scintillation characteristics of a crystal used for calorimeters are the light yield, the quantity of scintillation light emitted, and the emission time τ . The light yield LY is usually measured as the number of emitted photons per MeV of energy deposited inside the crystal. In Tab. 2.3 is reported the light yield normalized to that of sodium iodide doped with thallium NaI (Tl). For completeness, in the Tab. are reported the density ρ , the refraction index n and the wavelength of the light emission λ .

The PWO was chosen for the following reasons:

- the short radiation length of 0.89 cm allows the construction of an extremely compact calorimeter, in fact crystals of 23 cm of length correspond to $25.8 X_0$, which ensures an optimal longitudinal containment of the showers up to large energies;
- the small Molière radius allows an effective lateral containment of the showers and at the same time a good granularity;
- the high response rapidity allows to collect large part of the scintillation light between two subsequent crossings (80% of the light is emitted within 25 ns), minimizing the effects of the pileup.

- Its good radiation resistance makes it usable for decades of activity at LHC at high luminosity.

The PWO is insensitive to the neutron radiation, but the photon irradiation induces the formation of color centers in the crystal which can degrade the transparency, and therefore the light yield. A correlation has been found between the radiation resistance and the optical transmission of the crystals. Moreover it has been observed that a Nb doping optimizes the transmission spectrum and at the same time maximizes the radiation resistance of the crystal.

However the PWO has some inconvenients. A large amount of the energy stored in the crystals is dissipated in thermal emissions by the crystal lattice. Therefore the light yield is very low (about 30 photons/MeV), therefore it is necessary to use photodetectors with an amplification system. CMS uses avalanche photodiodes (APD) in the barrel and vacuum phototriodes (VPT) in the end-caps. The choice of these amplifiers was motivated by their good quantum efficiency for wavelengths in the region of the emission of the PWO, by their insensitivity to the magnetic field and by their compactness.

Another problem connected to the thermal dissipation of the absorbed energy is the dependence of the light yield on the temperature of the crystal. The PWO has a percentage variation of $-2\%/^{\circ}\text{C}$ at 18°C . To ensure a constant and reliable reaction of the detector it is necessary to keep its temperature constant, which motivated the construction of a hydraulic cooling system able to guarantee an operative temperature of $(18.00 \pm 0.05)^{\circ}\text{C}$.

Photodetectors

In the barrel the photodetectors are APDs with inverted structure (with the silicon of type n behind the $p-n$ junction) developed by the Hamamatsu company specifically for the CMS calorimeter.

The scintillation photons produced in the crystal enter the photodiode from the p^{++} layer and they are absorbed in the successive p^{+} layer, creating an electron-hole pair. The created electron is accelerated by the strong electric field produced by the $p-n$ junction, freeing other electrons by ionization. The freed electrons migrate towards the n^{++} layer where they are collected.

The motivations that led to the choice of the Hamamatsu APDs are the following:

- high internal gain, adjustable up to 200 (the working value is 50);
- good quantum efficiency;
- low capacity and low reverse saturation current;
- sufficient radiation resistance in order to work in the central region of the electromagnetic calorimeter (ECAL);
- compactness;
- insensitivity to the magnetic field;
- reduced response to charged particle thanks to the thin width of the multiplication region ($\sim 5 \mu\text{m}$).

The APDs have an active area of $5 \times 5 \text{ mm}^2$. To enhance the light collection two APDs have been stuck on the bottom face of each crystal of the barrel.

The radiation resistance of the APDs has been studied, particularly in the case of the neutron irradiation. No remarkable variation in the gain has been observed up to neutron fluxes of about $10^{13} \text{ neutrons/cm}^2$, while a linear increase of the dark current has been observed up to values of some μA for the maximum doses expected.

In the endcaps the photodetectors have to operate in a highly radioactive environment and in presence of a strong axial magnetic field. In fact the quantity of radiation is remarkably higher than that of the barrel: the flux of neutrons can reach values of about $10^{15} \text{ neutrons/cm}^2$ in the regions closest to the beams. For this reason the APDs can not be used.

It was decided to use VPTs. A phototriode is made by a photocathode made of radiation resistant glass, an anode grid distant 4-5 mm from the photocathode, and a dynode distant 2-3 mm from the anode. Typically the photocathode is grounded, the anode is put at a voltage of about 1000 V and the dynode is put at a voltage of about 800 V.

The scintillating photons coming from the crystal hit the photocathode, where they are converted in photoelectrons by photoelectric effect. The photoelectrons are accelerated by the strong electric field towards the anode. A large amount of them crosses the grid of the anode and hits the dynode, where many secondary electrons are produced (the secondary emission factor can reach 20). The secondary electrons created are accelerated towards the anode, where a large fraction of them is collected.

Geometry

The electromagnetic calorimeter has a cylindrical structure. The central part of the cylinder, named barrel, covers a pseudorapidity region of $|\eta| < 1.479$; the barrel is closed at both its sides by two lateral identical structures, named endcaps, that reach $|\eta| = 3$. In front of almost all the fiducial region of the endcaps (for $1.653 < |\eta| < 2.6$) there is a preshower, that studies the transversal development of the shower. Fig. 2.6 shows a perspective of the structure of the ECAL, and a section of a quarter of the ECAL.

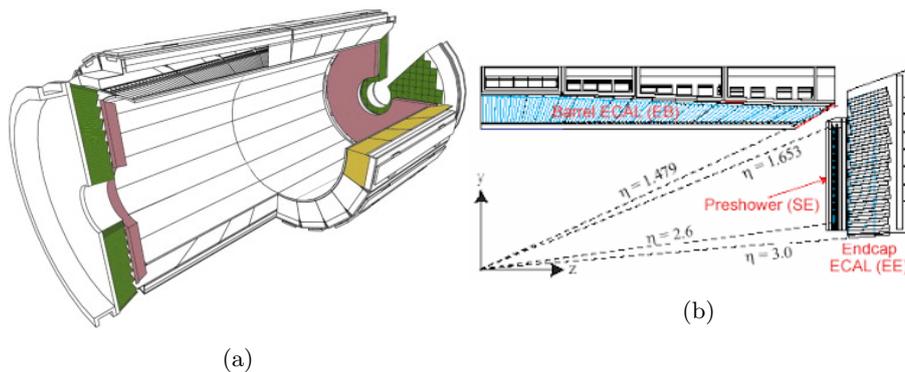


Figure 2.6: (a) Perspective of the structure of the ECAL; (b) section of a quarter of the ECAL, with some values of the pseudorapidity.

The barrel has an internal radius of 129 cm. It is made of 61200 crystals, correspondent to a granularity of 360 crystals in ϕ and of 2×85 crystals in η . The crystals have a truncated pyramid shape and are built in a quasi projective geometry so that their axes form an angle of 3° with the line joining them to the nominal interaction vertex. A single crystal corresponds approximately to a step of 0.0174×0.0174 in the $\eta - \phi$ plane (where ϕ is measured in radians), corresponding to 22×22 mm² in the front side and to 26×26 mm² in the bottom side. They are 23 cm long, which corresponds to $25.8 X_0$.

A crystal and the two APDs stucked in its bottom face constitute a subunit. The subunits are assembled in honeycomb structures made of thin foils of glass fiber, named submodules, made by five pairs of subunits. The submodules are assembled to make modules, and four modules form a supermodule. The barrel is made by 36 identical supermodules, where each of them covers half of its length.

The endcaps are placed at a longitudinal distance of 3144 mm from the interaction point, holding into account the displacement of about 2.6 cm towards the center caused by the magnetic field. They are made of identical crystals, where the front face has dimensions 28.62×28.62 mm², length equal to 220 mm (correspondent to $24.7 X_0$), and the bottom face has dimensions 30×30 mm². The crystals are grouped in mechanical units of 5×5 crystals, named supercrystals, made with a honeycomb structure in carbon fiber.

For $|\eta| > 2.5$, the radiation level and the high particle multiplicity forbid precision measurements. The crystals nearer to the beams are only used to measure the transverse energy of the event and to reconstruct the jets together with the hadronic calorimeter.

In the range $1.653 < |\eta| < 2.6$ a preshower is used, which has a circular crown shape of inner radius equal to 45.7 cm and outer radius of 1.23 m. It is a sampling calorimeter with two layers, which uses a lead sampler and a detector with silicon strips. The lead layers (of 2 and 1 X_0) cause the development of electromagnetic showers by the incoming electrons and photons, and the detectors made of silicons strips placed between the lead layers measure the deposited energy and the lateral profile of the showers.

The main aim of the preshower detector is the distinction between photons and neutral pions. In fact the π^0 decay rapidly in pair of photons which at high energies can be very close to each other and therefore they are hard to separate. Moreover the preshower is useful in the separation of electrons and particles at the minimum of ionization, and its granularity which is superior respect to that of the endcap crystals allows to determinate with greater precision the position of electrons and photons.

Energy resolution

The energy resolution of the electromagnetic calorimeter can be parametrized as:

$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \quad (2.7)$$

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:

$$S = 2.8\% \text{ GeV}^{\frac{1}{2}} \quad (2.8)$$

$$N = 124 \text{ MeV} \quad (2.9)$$

$$C = 0.3\% \quad (2.10)$$

for electrons that collide in the center of a crystal.

2.2.4 Hadronic calorimeter

The hadronic calorimeter of CMS measures the energies and the directions of the particles within the hadronic jets and estimates, together with the electromagnetic calorimeter, the missing energy of the events. For these reasons the two fundamental requirements are a good hermeticity and a good transverse granularity. It is also important to have a good energy resolution and a sufficient longitudinal containment of the hadronic showers.

The hadronic calorimeter is made by a central calorimeter ($|\eta| < 3$), and by two calorimeters for high pseudorapidities ($3 < |\eta| < 5$). It was chosen to use a sampling calorimeter which uses layers of copper as absorbers and plastic scintillators as active material. It has a tile structure, where the tiles are parallel to the beam axis. It is divided in a central cylindrical structure ($|\eta| < 1.3$) and in two endcaps ($1.3 < |\eta| < 3$), for a total of 2593 trigger towers without longitudinal segmentation. It has a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, which corresponds to the granularity of the trigger towers of the ECAL.

The central calorimeter has a depth of about 7 interaction lengths λ_I and has an energy resolution given by:

$$\frac{\sigma}{E} = \frac{100\%}{\sqrt{E(\text{GeV})}} \oplus 8\% \quad (2.11)$$

A depth of 7 λ_I is not enough to have a complete longitudinal containment of the hadronic showers. Therefore a further layer was inserted behind the solenoid, which gives 3 more λ_I and improves of 10% the energy resolution for pions of 300 GeV.

The calorimeters for high values of the pseudorapidity, placed in an environment with high doses of radiation and with high multiplicity, are sampling calorimeters made in iron and in quartz fibers. The fibers come in two different lengths: the longer start from the frontal face of the calorimeter, the shorter start 22 cm from the longer ones. In this way the electromagnetic component of the shower, which is deposited in the initial part of the calorimeter, can be obtained by subtraction. The calorimeters are made by a total of 1728 trigger towers and have a granularity of $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$.

2.2.5 Muon detector

The muon detector identifies and measures the muons, the only charged particles that are able to cross the calorimeters without being absorbed in them. The presence of muons in the final state is a characteristic of many physical processes. The most stringent requirements for the performance of the muon detector come from the decay channel of the Higgs boson $H \rightarrow ZZ^* \rightarrow 4\mu$.

The muon detector is placed outside the magnet and covers the pseudorapidity region $|\eta| < 2.4$. It is divided in a barrel and in two endcaps. Both in the barrel and in the endcaps, the system is made by four measurement stations, interspaced with the iron return yoke of the magnet. Fig. 2.7 shows a scheme of the muon system of CMS.

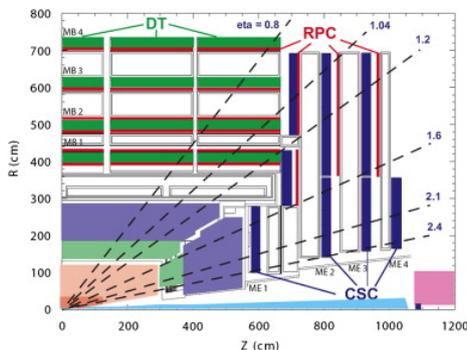


Figure 2.7: Scheme of the muon detector of CMS.

The barrel region is made of planes of drift tubes. Each station is made of a chamber made with 12 planes of tubes, for a total of 195000 tubes. In the endcaps there are cathodic strip chambers (CSC) in order to have precision measurements even in presence of a strong magnetic field and of a high particle multiplicity. The CSCs are multiwire proportional chambers where the cathodic plane is segmented in strips. They are organized in modules made of six layers.

Moreover, both in the barrel and in the endcaps, there are resistive plate chambers (RPC) which play the role of trigger. The RPCs are gas chambers with parallel planes that combine a discrete spatial resolution with an excellent time resolution (3 ns), comparable with that of a scintillator. They constitute a fast trigger system, able to identify candidate muons with high efficiency. They are organized in six stations in the barrel and in four stations in each endcap, for a total of 612 chambers.

The reconstruction efficiency for the muon tracks improves of the 90% for muons up to 100 GeV, and the assignment of the charge is correct with a confidence of 99%.

2.2.6 Trigger system

At the nominal luminosity of LHC the event rate is equal to 10^9 Hz. Since the typical dimension of an event is about 1 MB, it is not possible to store all the events in real time. On the other hand it is not very useful to do so, since a large part of the events are of QCD only, and are not interesting for the Higgs

boson and new physics researches. Fig. 2.8 shows the pp cross sections as a function of the energy in the center of mass: at the LHC energies we see that the electroweak processes (with W and Z) have a cross section about six orders of magnitude smaller than the total one, and the production of a SM Higgs boson of 120 GeV of mass has a cross section about ten orders of magnitude smaller than the total one.

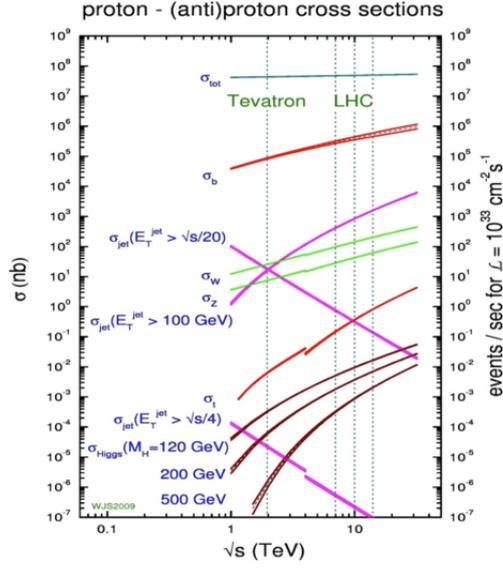


Figure 2.8: Proton-proton cross section as a function of the energy in the center of mass.

Therefore a trigger system is necessary to lower the event rate that must be stored, in order to work with an event rate that can be handled by the storage system (about 100 Hz). This is achieved with a two-level trigger system: a level-1 trigger (L1) and a high-level trigger (HLT). The system L1 is made by a series of hardware processors built for this task, while the HLT is a software system made by a farm of about a thousand of common processors.

Trigger L1

The first level trigger reduces the rate from 1 GHz to about 50-100 kHz. The data are stored in queues waiting for the trigger decision, which must be taken in $3.2 \mu\text{s}$. If L1 accepts the bunch crossing the data are sent to the HLT: the time is not enough to read the informations of all the detectors, therefore L1 uses only the informations of the calorimeters and of the muon chambers. The L1 trigger is divided in:

- calorimetric trigger, based of tower triggers, which are 5×5 matrices of the ECAL crystals which correspond to the cell of the hadronic calorimeter (HCAL). The tower triggers are grouped in 4×4 squares. The calorimetric trigger identifies the candidates photons, electrons, jets and τ ;
- muonic trigger, which uses the detectors of the muon system.

The informations of the two systems are sent to the global trigger which takes the decision considering both the single objects reconstructed by the detectors and their combination.

High level trigger

The high level trigger brings the final event rate to about 100 Hz. It is a software trigger made by a set of algorithms names paths, designed to select different topologies of events. It reconstructs the parts of the event near to the objects already reconstructed by the L1 trigger. Avoiding reconstructing the entire event the decision time is diminished. Moreover the non interesting events are eliminated as soon as possible in order to free the processor from non useful jobs which would slow down the jobs of the successive events. The high level trigger is divided in three sublevels: the first accesses only the data of calorimeters and muon system, the second adds the pixels, the third reads the informations of the entire event.

Chapter 3

Event selection

3.1 Signal characteristics

We focus our attention on the Higgs boson produced through VBF and its decay in two photons. This means that we look for a signal with a final state made of two photons, coming from the Higgs boson decay, and of two jets, coming from the quarks that hadronize after the VBF scattering.

We inherited the standard photon preselection used for the $H \rightarrow \gamma\gamma$ analysis, and which will be discussed in the following. For this thesis we worked on ROOT files, where ROOT is an open-source data analysis framework developed at CERN that was used to analyze both Monte Carlo (MC) and data events. The JHUGen [29] MC program was used to simulate all anomalous couplings in the H boson production. The SM production of the Higgs boson through VBF was simulated using POWHEG [30] at next-to-leading order in QCD and AMCATNLO at next-to-leading order in QCD [31].

The $H \rightarrow \gamma\gamma$ decay channel was studied in Run1 and Run2 at LHC by CMS, which measured its differential cross sections. The results are reported in [32].

Being the Higgs boson neutral, it cannot be directly coupled to the photon. However, the coupling of the Higgs boson to a particle is proportional to the mass of the particle, thus the Higgs boson can decay in two photons through loops of heavy charged particles, like W bosons and top quarks. In Fig. 3.1 we show the Feynman diagrams for the Higgs boson decay in two photons.

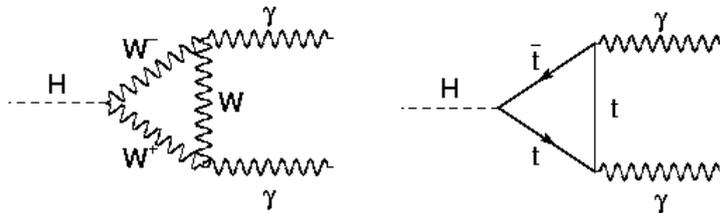


Figure 3.1: Feynman diagrams for the Higgs boson decay in two photons: (left) loop of W bosons; (right) loop of t quarks.

Since its mass was measured to be about 125 GeV, the Higgs boson will decay in two photons, each with energy $E \geq m_H/2 \sim 60$ GeV, where the lower limit of about 60 GeV is reached when the Higgs boson is produced at rest. Thus we look for events with two photons with high transverse energy E_T .

We request that the two photons are identified as isolated in the ECAL, where they form a tight cluster shape in η , and also that they are far from jets. The protons, however, are not elementary particles: the processes that lead to the production of the Higgs boson imply the presence of additional activity in the detector, which originates from their breaking. In some cases, this activity can have a typical and easily distinguishable conformation. For example, in the VBF, there are two jets produced at large pseudorapidities in the final state. When the Higgs boson is produced in association to other particles, such as W and Z in the Higgs-strahlung, we can study the decay products of these particles, which can enhance the selection efficiency.

The most striking characteristic of the signal in VBF production is the presence of a resonant Breit-Wigner peak in the invariant mass distribution of the pairs of reconstructed photons and of two forward, very energetic and with large pseudorapidity gap jets.

3.2 Backgrounds

Any event with two deposits of energy in the ECAL could in principle be considered as background for the $\gamma\gamma$ channel. Processes with two real photons in the final state with high transverse energy, which do not come from a Higgs boson and thus emulate the signal, are irreducible background, while events where at least one jet in the final state is interpreted as a photon are reducible background.

In proton-proton collisions, two photons can be produced through the annihilation of a quark-antiquark pair or through fusion of two gluons. The production of two photons through quark-antiquark annihilation is usually named Born, while the fermion loop in the Feynman diagram of the gluon fusion attributes to the process the name box.

In Fig. 3.2 we report the Feynman diagram for the Born and the box processes.



Figure 3.2: Production of two photons in proton-proton collisions: (a) Born process; (b) box process.

The topology of these events is identical to what we expect from the signal. In fact, both for the signal and for the background we have two isolated photons with high transverse energy. However, in the two background cases the diphoton production is non resonant. To discriminate between these and $H \rightarrow \gamma\gamma$ events, we can study many discriminating kinematic variables of the events, which for

example involve the diphoton invariant mass, the transverse energy and the angular distributions of the photons and the jets in the final state.

Events where at least one jet in the final state is interpreted as a photon can be of two kinds: events from multijet production ($pp \rightarrow \text{jets}$) and events with production of a photon and at least one jet ($pp \rightarrow \gamma + \text{jets}$).

In Fig. 3.3 we show the Feynman diagrams of the production of $\gamma + 1$ jet and of 2 jets in proton-proton collisions.

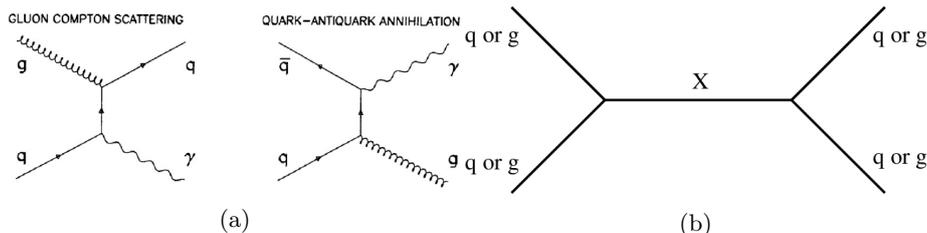


Figure 3.3: (a) Feynman diagram of the production of $\gamma + 1$ jet in proton-proton collisions: (left) gluon Compton scattering; (right) quark-antiquark annihilation; (b) Feynman diagram of the production of 2 jets in proton-proton collisions.

To constitute a source of background, the jets must deposit a large amount of energy in the ECAL, simulating the electromagnetic shower of a signal photon. Therefore they must contain a particle which can produce such a deposit, which can be a photon, an electron, or a hadron that decays electromagnetically (such as $\pi^0, \eta, \eta', \rho, \omega$). In any case, the candidate photon is part of a jet, and thus is not isolated. Typically in these kind of events, we should find some activity nearby the candidate photon, which can be tracks reconstructed in the inner tracker, other deposits in the ECAL or in the HCAL.

There is also the background relative to the process $pp \rightarrow \gamma\gamma + \text{jets}$, where the two photons can be non resonant or may have low transverse energy.

Another kind of reducible background are the Drell Yan events, where a quark-antiquark pair annihilates in a photon or in a Z boson, which then decays in a positron-electron pair ($q\bar{q} \rightarrow \gamma/Z^* \rightarrow e^+e^-$). This kind of process forms two deposits of energy in the ECAL and thus can constitute a source of background if the tracks of the electrons are not correctly reconstructed.

3.3 Measurement of the mass of the Higgs boson

Once we have defined the selection criteria, we will have events with pairs of candidate photons. The photons coming from the Higgs boson decay will have an invariant mass distribution peaked around 125 GeV, with a width given by the experimental resolution on the invariant mass, since the intrinsic width of the Higgs boson is negligible. Instead the candidate background photons will have a decreasing invariant mass distribution.

Being the photon massless, the invariant mass of a diphoton pair is given by:

$$m_{\gamma\gamma} = \sqrt{2E_1E_2(1 - \cos \alpha)} \quad (3.1)$$

where E_1, E_2 are the energies of the two photons and α is the angle between them. If, as we will see in the following, the uncertainty in the measurement of the direction of the photons is of the order of the $^\circ$, the angular resolution is negligible with respect to the energy resolution.

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 order not to attribute such a peak to a statistical fluctuation of the background, the background contribution b under the peak was calculated by performing a fit to the observed distribution in its sidebands. The contribution of the signal s was obtained by subtraction.

Assuming a Poisson distribution of the statistical fluctuations in the number of events, the significance σ of the signal is defined as:

$$\sigma = \frac{s}{\sqrt{b}} \quad (3.2)$$

When this number is greater than 5 the probability that the relative peak was due to a casual fluctuation of the background is less than $5.7 \cdot 10^{-5}$. For a significance $\sigma > 5$ a discovery is announced. For a significance $\sigma > 3$, the probability of a casual fluctuation of the background is of the order of 10^{-3} , and an evidence is announced.

From Eq. (3.1), we see that the resolution on the invariant mass depends on the resolution of the energy two photons and on the resolution of their relative angle.

The bunches of protons accelerated at LHC have a negligible section in the plane transverse to the beams, but have a standard deviation of about 7.5 cm in the direction z , which results in a longitudinal dispersion of about 5.3 cm in the distribution of the interaction vertices. Assuming that the photon is produced in the nominal interaction vertex (the origin of the CMS coordinates) can thus correspond to an error of the order of the cm, which, for a photon that collides in the center of the barrel (at $\eta \sim 0$), corresponds to an error of the order of the $^\circ$ in the measurement of its direction.

In order to have a good resolution in the angle, and thus in the invariant mass, it is important to measure the position of the interaction vertex, which is the Higgs boson decay vertex. Since the photons are not charged particles, they do not leave tracks, however we can study the tracks of the other particles produced in the proton-proton collisions to localize the Higgs boson decay vertex.

3.4 Photon reconstruction in the ECAL

A photon releases all its energy in the crystals of the ECAL. Therefore, a candidate photon is an energy deposit in the ECAL, which can extend on several crystals, thus an algorithm is needed to reconstruct all the deposited energy.

The transversal dimension of the ECAL PWO crystals was chosen to be equal to the Molière radius of the PWO, so a photon colliding in the center of a crystal's frontal face will deposit 90% of its energy in it. To fully reconstruct the energy of the photon it is necessary to group together several adjacent crystals and sum up the energy contained in them. The set of these crystals is called cluster, while the grouping procedure is called clustering algorithm.

The energy deposit of an electromagnetic shower produced in the ECAL is made of a local energy maximum embedded in a region of lower energy deposits (bump). The reconstruction process therefore must find the crystals (seeds) where there is a larger energy deposit than the deposit in the other adjacent crystals, and it must include the nearby crystals, trying to recover the maximum fraction of the energy of the shower. However the clustering process must not include the deposits relative to other particles or the noise of the detector.

To reduce the influence of the fluctuations of the noise, a procedure is followed, named zero suppression, which consists in hiding the contribution of the crystals that register an energy within three standard deviations from the mean value of the noise. The hiding implies that the value of the energy of these crystals is put equal to zero.

In the case of a photon or an electron/positron colliding in the ECAL, the bump should reproduce the cluster shape of the shower. However, at the border of it, where the energy deposits are comparable with the noise, there is the risk that part of the energy that belongs to the shower gets ignored during the zero suppression, or there is the risk that a large fluctuation of the noise simulates the presence of a second bump.

To avoid these problems, we can choose to open matrices of fixed dimensions around the seed. In fact, this kind of procedure tends to have less problems of this type respect to a bump research procedure. About the 94% of the energy of the photon is deposited in a 3×3 matrix around the seed, while the 97% in a 5×5 matrix.

While traversing the tracker material in front of the ECAL, the photons can convert in positron-electron pairs, and the electrons can irradiate photons by bremsstrahlung. The deposits are thus spread in regions of the ECAL which can be more extended respect to the characteristic regions of a single shower. This brings to elaborate more complex and flexible clustering algorithms.

3.4.1 Superclustering in the ECAL

Energy deposits in several ECAL channels are clustered assuming that each local maximum above 1 GeV corresponds to a single particle incident on the detector [33]. An ECAL energy deposit can be shared between overlapping clusters, and a Gaussian shower profile is used to determine the fraction of the energy deposit to be assigned to each of the clusters. Since electrons and photons have a significant probability of showering when traversing the CMS tracker, by the time the particle reaches the ECAL, the original object can consist of several electrons and/or photons produced from bremsstrahlung and/or pair production. The multiple ECAL clusters need to be combined into a single supercluster (SC) that contains the energy of the original electron/photon. This step is known as superclustering, and the combining process uses two algorithms:

- the “mustache” algorithm starts from a cluster above a given threshold, called seed cluster. Additional clusters are added if they fall into a zone, whose shape is similar to a mustache in the transverse plane. The name mustache is used because the distribution of $\Delta\eta = \eta_{seed-cluster} - \eta_{cluster}$ vs $\Delta\phi = \phi_{seed-cluster} - \phi_{cluster}$ is slightly bended because of the solenoidal structure of the CMS magnetic field, which tends to spread this radiated energy along ϕ , rather than along η . The size of the mustache region de-

depends on E_T , since the tracks of particles with larger transverse momenta get less bent by the magnetic field. The mustache SCs are used to seed electrons and photons;

- the “refined” algorithm uses tracking information to extrapolate bremsstrahlung tangents and conversion tracks to decide whether a cluster should belong to a SC. It uses mustache SCs as starting points, but is also capable of creating its own SCs. The refined SCs are used to determine all ECAL-based quantities of electrons and photons.

3.4.2 Energy corrections

For a correct energy reconstruction of a photon or an electron in the ECAL it is necessary to take into account the effects depending on the energy scale and on the pseudorapidity. The energy scale corrections depend on residual non linearities in the calorimeter, due to the non complete containment of the shower and to the effects of the light collection in the crystals. The dependence on the pseudorapidity comes from the structure of the detector, which has a quantity of crossed material which varies in η (it is maximum near the junctions between barrel and endcap). The quantity of crossed material influences the energy fraction lost by the electrons by bremsstrahlung and the probability of e^+e^- conversion for the photons, therefore it results in a downgrade in the energy reconstruction.

The energy deposited by electrons and photons in the ECAL and collected by the superclustering algorithm is subject to losses for several reasons. Electromagnetic shower energy in the ECAL can be lost through lateral and longitudinal shower leakage, or in intermodule gaps or dead crystals. There are also pileup interactions that produce photons/electrons that end up in the same supercluster. The shower energy can also be smaller than the initial electron energy because of the energy lost in the tracker.

These losses result in systematic variations of the energy measured in the ECAL. Without any corrections, this would lead to a degradation of the energy resolution for reconstructed electrons and photons. To improve the resolution, a multivariate technique is used to correct the energy estimation for these effects.

A set of regression fits based on boosted decision trees (BDT) is applied to correct the energy of e/γ . The minimum E_T for electrons (photons) considered for the BDT training is 1 (5) GeV at the simulation level. Each of these energy regressions is built as follows. The regression target y is the ratio between the true energy of an e/γ and its reconstructed energy, thus the regression prediction for the target is the correction factor to be applied to the measured energy to obtain the best estimate of the true energy. The regression input variables, represented by the vector \vec{x} , includes the object and event parameters most strongly correlated with the target. The regression is implemented as a gradient-BDT, and a log-likelihood function is employed:

$$\mathcal{L} = - \sum_{MC \ e/\gamma \ objects} \ln p(y|\vec{x}) \quad (3.3)$$

where $p(y|\vec{x})$ is the estimated probability for an object to have the observed value y , given the input variables \vec{x} , and the sum runs over all objects in a simulated sample in which the true values of the object energies are known.

The probability density function used in this regression algorithm is a double-sided crystal ball function that has a Gaussian core with power law tails on both sides.

After applying the previous corrections, small differences remain between data and simulation in both the electron and photon energy scales and resolutions.

An additional spreading needs to be applied to the photon and electron energy resolutions in simulation to match that observed in data. The electron and photon energy scales are corrected by varying the scale in the data to match that observed in simulated events. The magnitude of the final correction is up to 1.5% with a total uncertainty estimated to be smaller than 0.1 (0.3)% in the barrel (endcap). In Fig. 3.4 we show the ratio of the true to the reconstructed electron energy.

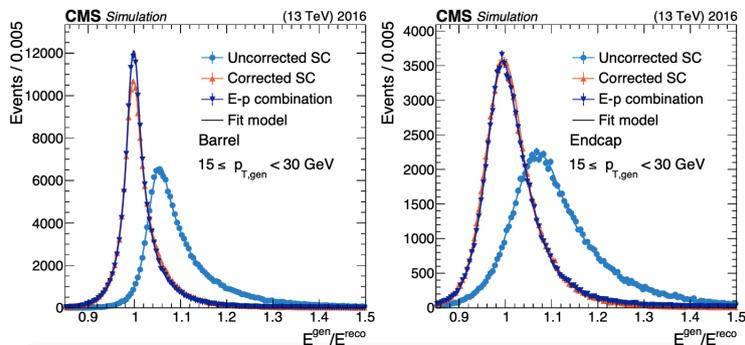


Figure 3.4: Ratio of the true to the reconstructed electron energy in the p_T range 15–30 GeV with and without regression corrections, with a DSCB function fit overlaid, in 2016 MC samples for barrel (left) and endcap (right) electrons. Vertical bars on the markers represent the statistical uncertainties of the MC samples.

3.4.3 Reconstruction of converted photons

Between 48% and 50% of the photons produced from the Higgs boson decay converts in a positron-electron pair, while traversing the material of the internal tracker.

The positron and the electron produced in the conversion are deviated in opposite directions from the magnetic field and thus can be absorbed in the ECAL in positions very distant from each other, such that they are not embedded in the same supercluster. Even if the two cluster are close together, the form of the energy deposit of a converted photon is different from that of a non converted photon.

Respect to an unconverted photon, a photon that converts in an electron-positron pair in the tracker has an additional information, which is given by the tracks left by the charged particles in the tracker, that can be studied in order to find the original direction of the photon. This procedure improves the angular resolution and therefore the invariant mass resolution of the photons, as we can see from Eq. (3.1). Even in the case of unconverted photons, the

study of the tracks of the charged particles produced in the primary vertex in association with the photons help in finding the direction of the photons.

3.5 Reconstruction and identification of jets

Hadronic jets are clustered from their reconstructed particles using the infrared and collinear safe anti- k_T algorithm [34, 35] with a distance parameter of 0.4. Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5% to 10% of the true momentum over the whole p_T spectrum and detector acceptance. Additional proton-proton interactions within the same or nearby bunch crossings (pileup) can yield additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, charged particles identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions. Jet energy corrections are derived from simulation to bring the measured response of jets to that of particle level jets on average.

In the VBF case, the final state jets will be characterized by high values of their invariant mass m_{jj} , of their transverse momentum $(p_T)_j$ and of the difference of their pseudorapidities $(\Delta\eta)_{jj}$.

3.6 Kinematic variables

Several MC simulations have been done for different hypotheses of signal, such as VBF for SM, CP -even, CP -odd, and for backgrounds, like $\gamma\gamma + \text{jets}$ and $\gamma + \text{jets}$. Different hypotheses can yield different distributions.

While typically ggH is considered as signal, in our case we will consider it as background, since we are only interested in the VVH couplings, which enter the VBF production mode even if they do not enter in the decay process in the $H \rightarrow \gamma\gamma$ channel (which occurs via quark loops). ggH constitutes in this case an irreducible background, since it contributes to the resonant invariant mass peak of the pairs of photons.

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. This is not true for the background processes, thus the distributions of m_{jj} , $(p_T)_j$ and $(\Delta\eta)_{jj}$ are expected to be different from the VBF case. Moreover, as said before, the invariant mass distribution of the pair of photons for the Higgs boson signal will be resonant, with a peak around 125 GeV, while the background will have a non resonant decreasing exponential invariant mass distribution.

In the ggH case, the final state jets are the ones of the initial state radiation. Therefore they have low p_T , and their spectrum is characterized by a decreasing exponential distribution.

Differences between the SM and the anomalous VBF samples can be found in the angular distributions of photons and jets. The anomalous samples include different CP hypotheses which have a different spin-parity with respect to a pure SM scalar Higgs boson, and also contributions of BSM particles with a mass scale of Λ_1 . In this category we are grouping the contributions of particles that

enter in a virtual way in the loops, that can have a mass even larger than that of the Higgs boson and that are short lived. In fact, in the loops the energy is not conserved, and they can receive contributions from virtual particles related to physics BSM, which can have a mass scale of order Λ_1 . We expect that the angular variables that can be reconstructed, such as $\Delta\eta$ and $\Delta\phi$ between pairs of photons and jets, are different between the SM and the anomalous samples, since for the conservation of angular momentum different spin-parity hypotheses will yield non-flat angular distributions that will depend from the type of interaction considered. Therefore the angular combinations that can be built from the photons and the jets in the final state are useful in discriminating between different hypotheses.

As said before, a discriminating variable that can be considered is the transverse momentum of the photons and the jets. In this case, the new physics anomalous samples will have a harder spectrum respect to the SM sample, especially the one with the large new physics scale Λ_1 .

All these differences in the distributions are useful in discriminating between different hypotheses of Higgs boson signal and between signal and background (i.e. between VBF and ggH, $\gamma\gamma + \text{jets}$, $\gamma + \text{jets}$).

A set of simple selections on kinematic variables is applied to reject most of the backgrounds and enhance the VBF phase space: in the following we will call this set of cuts *VBF preselection*. These cuts are chosen in order to maximize the discrimination between signal and background, and are given by the following conditions:

- $|\eta_{\gamma_1}| < 2.5$, $|\eta_{\gamma_2}| < 2.5$, $|\eta_{\gamma_1}| < 1.44$ or $|\eta_{\gamma_1}| > 1.57$, $|\eta_{\gamma_2}| < 1.44$ or $|\eta_{\gamma_2}| > 1.57$;
- $100 \text{ GeV} < m_{\gamma\gamma} < 180 \text{ GeV}$;
- $(p_T)_{\gamma_1}/m_{\gamma\gamma} > 0.333$, $(p_T)_{\gamma_2}/m_{\gamma\gamma} > 0.25$;
- $|\Delta\eta|_{jj} > 0.0$, $|\eta_{j_1}| < 4.7$, $|\eta_{j_2}| < 4.7$, $\min(\Delta R_{jet \ \gamma}) > 0.4$;
- $(p_T)_{j_1} > 40 \text{ GeV}$, $(p_T)_{j_2} > 30 \text{ GeV}$;
- $m_{jj} > 250 \text{ GeV}$;
- $MVA_{\gamma_1}^{ID} > 0.5$, $MVA_{\gamma_2}^{ID} > 0.5$.

The cuts in the acceptance of the photons $|\eta_\gamma| < 2.5$ consider the photons that travel in regions within the acceptance of the tracker, even if the ECAL covers up to $|\eta| = 3$. In fact, in the regions beyond the borders of the tracker, the photons have a worse energy resolution, and there is a greater contribution from jets, because one cannot require that the photons are isolated in the tracker.

The photons coming from the decay of the Higgs boson have a Jacobian peak around 60 GeV in their p_T distribution. A cut on $(p_T)_\gamma$ would alterate the invariant mass distribution of the photons. Instead, we choose to apply a cut to $(p_T)_\gamma/m_{\gamma\gamma}$, which does not modify the invariant mass distribution, and which yields a smoothly decreasing exponential distribution in $(p_T)_\gamma$ for the background.

As jet acceptance we require $|\eta_j| < 4.7$. Beyond that the contribution of noise and pileup energy in the hadronic calorimeters is too high. With the

$\min(\Delta R_{jet \gamma}) > 0.4$ cut, we require that the jets are far from the photons to avoid photons coming from jets.

A BDT was trained using as inputs some photon identification variables that exploit the cluster shape of the photons, whose output is called MVA_{γ}^{ID} . A cut in $MVA_{\gamma}^{ID} > 0.5$ is used to reject the fake photons, which can be charged or neutral hadrons (eg. kaons or pions), which typically yield different cluster shapes. This cut is particularly efficient in suppressing the $pp \rightarrow 2 \text{ jets}, \gamma + 1 \text{ jet}$ and $\gamma + 2 \text{ jets}$ events. Regarding the $\gamma + 1 \text{ jet}$ sample, by cutting on the MVA_{γ}^{ID} of the two reconstructed photons, one of them will pass with high efficiency, the other with low efficiency. Since we do not know which one of them is the fake photon, we put the same requirement on both reconstructed photons.

In Fig. 3.5, 3.6 and 3.7, we show the histograms normalized to their area of some discriminating variables for some MC samples for the events that pass the *VBF preselection* in Eq. 3.6.

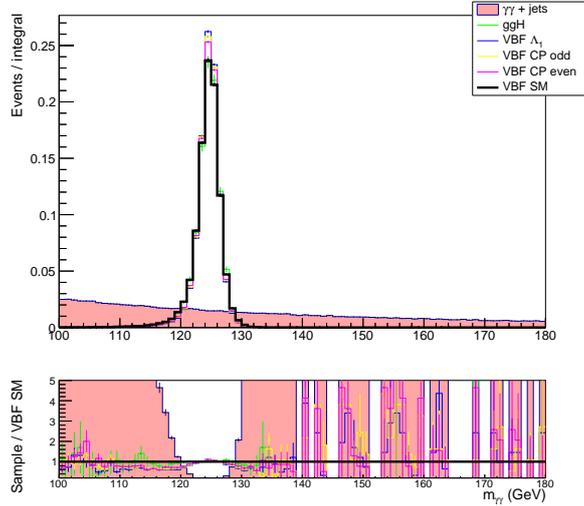


Figure 3.5: MC simulation of $m_{\gamma\gamma}$ for different hypotheses.

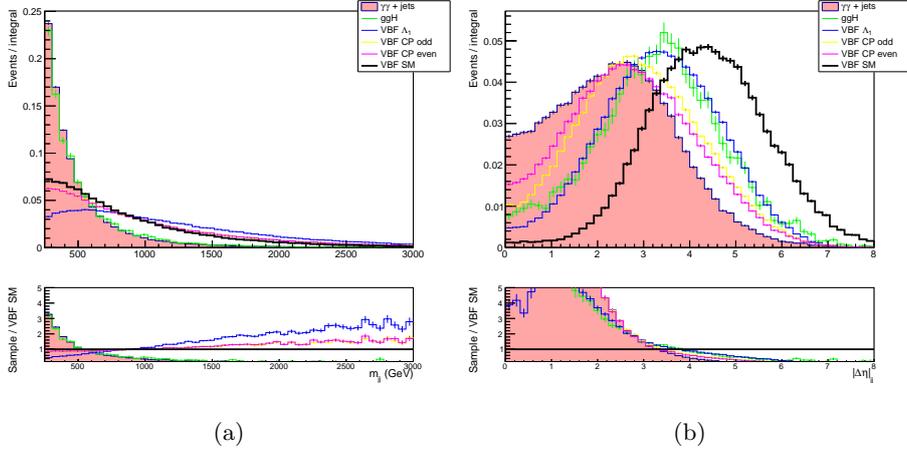


Figure 3.6: (a) MC simulation of m_{jj} for different hypotheses; (b) MC simulation of $|\Delta\eta|_{jj}$ for different hypotheses.

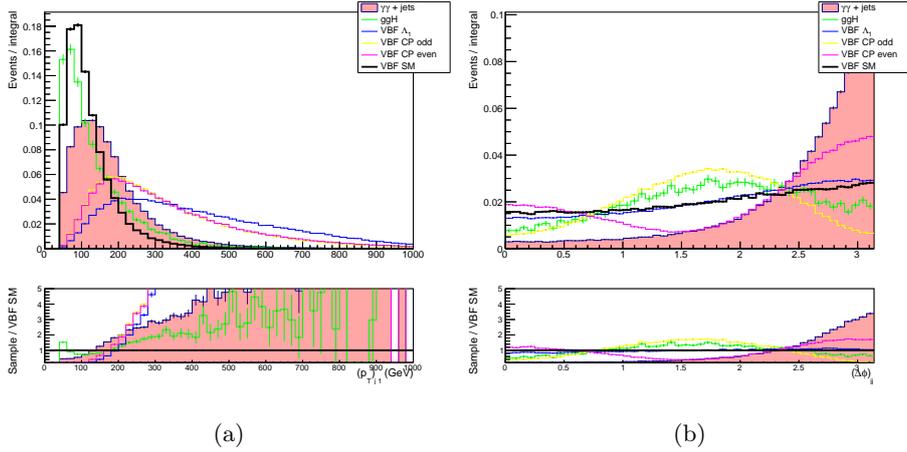


Figure 3.7: (a) MC simulation of $(p_T)_{j_1}$ for different hypotheses; (b) MC simulation of $(\Delta\phi)_{jj}$ for different hypotheses.

As we can see from Fig. 3.5, the invariant mass of the pairs of photons for the different signals does not depend on the coupling. The difference between the SM and BSM samples does not lie in the shape of $m_{\gamma\gamma}$, but in the normalization, which comes from the differences in the cross section and in the preselection efficiency for the different samples. In Sec. 6.2 we will use this variable to extract the signal, by statistically subtracting the non-resonant background with a fit: at that point we will remain with the sum of the VBF and the ggH contributions.

In Fig. 3.6 we show m_{jj} and $|\Delta\eta|_{jj}$, the typical variables used in the VBF context.

As we can see from Fig. 3.7, the anomalous VBF samples show a harder $(p_T)_j$ spectrum respect to the SM VBF and ggH.

Chapter 4

Multivariate analysis

4.1 MVA selection based on SM-only hypothesis

A multivariate analysis (MVA) is a statistical procedure for data analysis which involves the simultaneous study of many variables and which can produce a single output as a result. The output for each hypothesis class can be interpreted as the probability that the event under test is a signal or a background event. In particular, a BDT was trained for the measurement of the cross section of the SM Higgs boson in the $\gamma\gamma$ decay channel, and it discriminates between 3 classes, namely SM VBF, SM ggH and non resonating background [32].

In the study done for this thesis, contrary to what was done in the measurement of the cross section, we consider also some anomalous VBF signals. In Fig. 4.1 and 4.2, we show the histograms normalized to their area of the three MVA outputs for some MC samples, including SM and BSM cases.

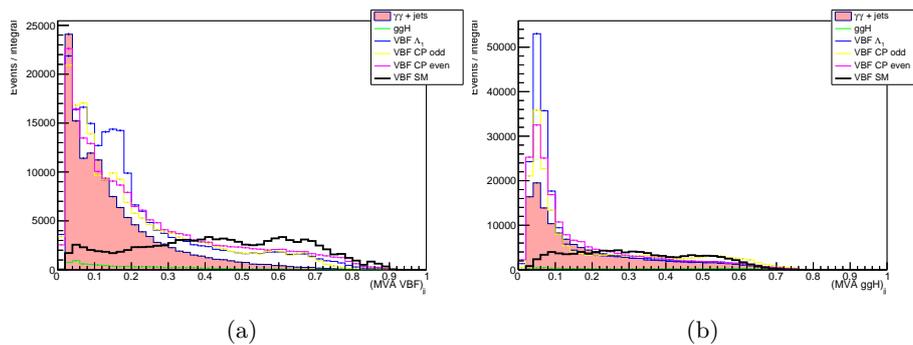


Figure 4.1: (a) $(MVA VBF)_{jj}$ for different MC samples; (b) $(MVA ggH)_{jj}$ for different MC samples.

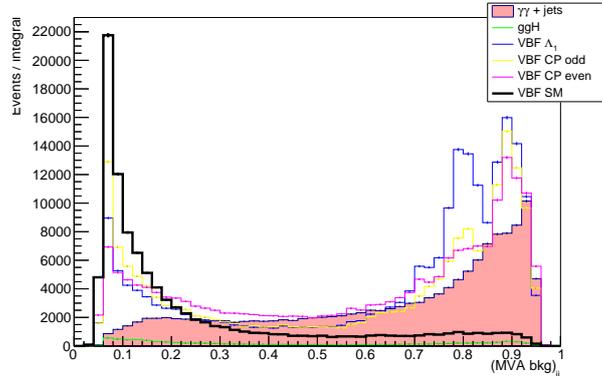


Figure 4.2: $(MVA\ bkg)_{jj}$ for different MC samples.

In [32] a selection on the three outputs of the BDT was used to discriminate the SM signal from the background: in the paper, some cuts in the MVA outputs were chosen, given by the following conditions:

- $(MVA\ VBF)_{jj} > 0.379$;
- $(MVA\ ggH)_{jj} < 0.565$;
- $(MVA)_{\gamma\gamma} > 0.800$.

Applying the *VBF preselection* in Eq. 3.6 and the MVA cuts in Eq. 4.1, we obtain the Tab. 4.1 of the selection efficiencies for the single cut with respect to the previous one and the Tab. 4.2 of the expected events in the 2017 data sample with $41.5\ \text{fb}^{-1}$.

	$\gamma\gamma + \text{jets}$	$\gamma + \text{jets}$	ggH	VBF Λ_1	VBF CP odd	VBF CP even	VBF SM
$ \eta_\gamma $	0.99	0.99	0.99	0.99	0.99	0.99	0.99
$m_{\gamma\gamma}$	0.50	1.00	1.00	0.98	0.98	0.98	0.99
$(p_T/M)_\gamma$	0.85	0.83	0.93	0.95	0.94	0.93	0.90
$ \Delta\eta _{jj}$	0.28	0.09	0.20	0.84	0.91	0.91	0.56
$(p_T)_j$	0.57	0.25	0.46	0.90	0.95	0.95	0.77
m_{jj}	0.36	0.26	0.33	0.91	0.84	0.85	0.79
MVA_γ^{ID}	0.65	0.10	0.72	0.70	0.74	0.73	0.73
$(MVA\ VBF)_{jj}$	0.10	0.12	0.17	0.16	0.18	0.22	0.59
$(MVA\ ggH)_{jj}$	0.10	0.12	0.17	0.16	0.18	0.22	0.59
$(MVA)_{\gamma\gamma}$	0.09	0.05	0.17	0.16	0.18	0.21	0.57

Table 4.1: Selection efficiencies for the single cut with respect to the previous one for different hypotheses, with the *VBF preselection* and the MVA cuts.

	$\gamma\gamma + \text{jets}$	$\gamma + \text{jets}$	ggH	VBF Λ_1	VBF CP odd	VBF CP even	VBF SM
$ \eta_\gamma $	608055	474253	2079	228	184	198	167
$m_{\gamma\gamma}$	302929	472854	2070	223	180	194	165
$(p_T/M)_\gamma$	256841	394648	1927	212	169	180	149
$ \Delta\eta _{jj}$	72142	35391	378	179	154	164	83
$(p_T)_j$	41207	8963	173	162	147	156	64
m_{jj}	15024	2327	57	147	123	132	50
MVA_γ^{1D}	9802	241	41	103	91	96	37
$(MVA\ VBF)_{jj}$	965	30	7	17	16	21	22
$(MVA\ ggH)_{jj}$	965	30	7	17	16	21	22
$(MVA)_{\gamma\gamma}$	848	13	7	17	16	21	21

Table 4.2: Number of expected events in 41.5 fb^{-1} for different hypotheses, with the *VBF preselection* and the MVA cuts.

The number of events is given by:

$$N = \mathcal{L}_{int}\sigma\epsilon \quad (4.1)$$

where \mathcal{L}_{int} is the integrated luminosity, which for the 2017 samples studied in this thesis is equal to 41.5 fb^{-1} , σ is the cross section and ϵ is the efficiency.

From Tab. 4.2, we can see that few signal events pass the selection given by the cut in the $(MVA\ VBF)_{jj}$ variable, especially for the anomalous signals, so in the process we are rejecting most of the signal, identifying it as background. It is therefore necessary to find a new set of discriminators in order to select more VBF events as signal and in order to not discard them as background.

To this end, a deep neural network (DNN) training was performed. This produced a new set of variables, analogous to the MVA ones, which were used to select efficiently the VBF events as signal and to treat the other contributions as background. In addition to the SM samples, this time the DNN considered was trained also on the anomalous signals samples, which show different kinematics respect to the SM.

4.2 DNN training

Deep learning is a class of machine learning algorithms, which are computer algorithms that can improve automatically through experience and using data [36]. These algorithms build a model based on sample data, known as training data, in order to make predictions without being explicitly programmed to do so.

A DNN is an artificial neural network (ANN), that consists of a collection of simulated neurons. Each neuron is a node, connected to other nodes via links. Each link has a weight, which determines the strength of a node's influence on another.

ANNs are composed of artificial neurons conceptually derived from biological neurons. Each artificial neuron has inputs and produces a single output which can be sent to many other neurons. The inputs can be the feature values of a sample of external data, or can be the outputs of other neurons. The outputs

of the final output neurons of the neural net accomplish the particular task in question.

The output of the neuron is given by the weighted sum of all the inputs, weighted by the weights of the connections from the inputs to the neuron. A bias term is added to the sum. The sum is called activation, and is passed to a non-linear activation function to produce the output. Therefore the activation function defines how the weighted sum of the input is transformed into an output from the nodes in a layer of the network. The initial inputs are external data, while the ultimate outputs accomplish the particular task in question.

The network consists of connections. Each connection provides the output of a neuron as an input to another neuron, and is assigned a weight that represents its relative importance. A given neuron can have multiple input and output connections.

The neurons are typically organized into multiple layers, especially in deep learning. A DNN is an ANN with multiple layers between the input and output layers. Neurons of one layer connect only to neurons of the immediately preceding and immediately following layers. The layer that receives the external data is the input layer. The layer that produces the result is the output layer. In between there can be zero or more hidden layers.

A hyperparameter is a constant parameter whose value is set before the learning process begins. Examples of hyperparameters include the number of layers, the batch size, the number of epochs, the number of nodes.

Learning is the adaptation of the network to better handle a task by considering sample observations. It involves adjusting the weights of the network to improve the accuracy of the result, by minimizing the observed errors. Learning is complete when examining additional observations does not usefully reduce the error rate. It is possible to define a loss function that is evaluated periodically during learning. As long as its output continues to decline, learning continues. The outputs are numbers, so when the error is low, the difference between the output and the correct answer is small. Learning attempts to reduce the total of the differences across the observations.

For our DNN a classification algorithm was considered.

Many discriminating variables were used as input for our DNN training, such as m_{jj} , the transverse momentum and the angular distributions already discussed in Sec. 3.6. We do not use $m_{\gamma\gamma}$ as input of the DNN training: in Sec. 6.1 we will use this variable in the fit procedure performed to extract the signal as a peak over a decreasing non-resonating background.

The DNN was trained to discriminate between 3 classes of events:

- ggH, the SM contribution only;
- $\gamma\gamma + \text{jets}$, which constitutes the main source of background, since $\gamma + \text{jets}$ has low statistics;
- VBF production.

The training was done with two VBF configurations: in the first we considered the SM only; in the second we considered the SM and the anomalous signals, which include CP -even, CP -odd, Λ_1 , $Z\gamma$ samples and their fractions (for example 50% CP -even) and their combinations (for example 50% Λ_1 and 50% $Z\gamma$).

As a result, for each of the three classes the DNN gave an output between 0 and 1, which is analogous to a MVA output and can be viewed as the probability that a given contribution is correctly interpreted as belonging to a given class.

In Fig. 4.3 and 4.4, we show the histogram normalized to their areas of the three DNN variables for different MC samples and for the three classes VBF, ggH and background.

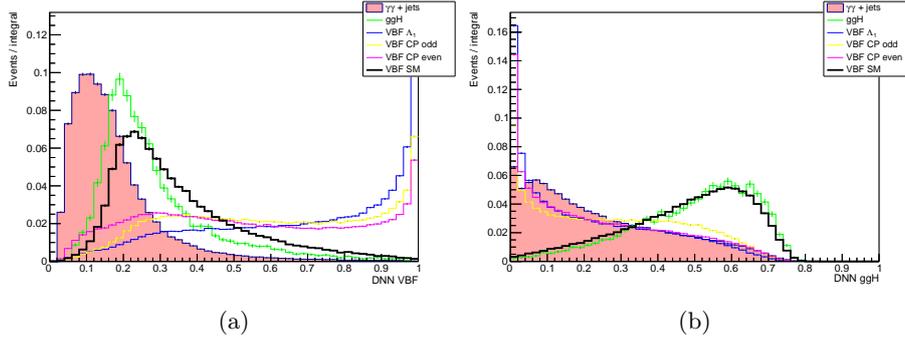


Figure 4.3: (a) $DNN VBF$ for different MC samples; (b) $DNN ggH$ for different MC samples.

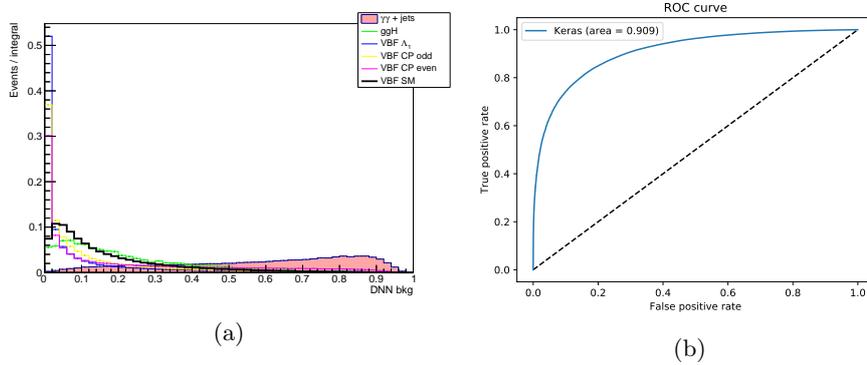


Figure 4.4: (a) $DNN bkg$ for different MC samples; (b) ROC curve, signal efficiency vs background efficiency.

The ROC curve in Fig. 4.4, where ROC stands for receiver operating characteristic, is a plot that shows the performance of a classification model at all classification thresholds. Thus each point of the ROC curve will indicate a classification threshold. The threshold is a three-dimensional selection on the three output variables. This curve plots two parameters, the true positive rate TPR and the false positive rate FPR , defined respectively as follows:

$$\begin{aligned}
 TPR &= \frac{TP}{TP + FN} \\
 FPR &= \frac{FP}{FP + TN}
 \end{aligned}
 \tag{4.2}$$

where TP, TN, FP, FN stand respectively for true positives, true negatives, false positives and false negatives.

In order for the ROC curve to be a good discriminant, it has to stay as far as possible from the diagonal.

The area under the ROC curve is a measure of the quality of the model's predictions across all possible classification thresholds, and it can be interpreted as the probability that the classification model ranks a random positive example more highly than a random negative example. It ranges from 0 to 1: a model whose predictions are 100% wrong has an area equal to 0, while a model whose predictions are 100% correct has an area equal to 1.

The software library Tensorflow was used to train the DNN [37]. The process maximized the accuracy and at the same time minimized the categorical crossentropy. The accuracy and the categorical crossentropy are metrics, functions used to judge the performance of a model. The accuracy calculates how often the prediction equals the input data, and is the fraction of predictions our model got right. It can be defined as follows:

$$accuracy = \frac{\text{number of correct predictions}}{\text{total number of predictions}} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.3)$$

The categorical crossentropy is a loss function, and it computes the crossentropy metric between the input data and the predictions. It is a measure of the difference between the two distributions.

For the training of our DNN, we used 19 nodes as the input layer, equal to 19 discriminating variables, then 10 nodes in the hidden layer, and then 3 nodes as the output layer, equal to the 3 output classes. In Fig. 4.5 we show the training convergence plot.

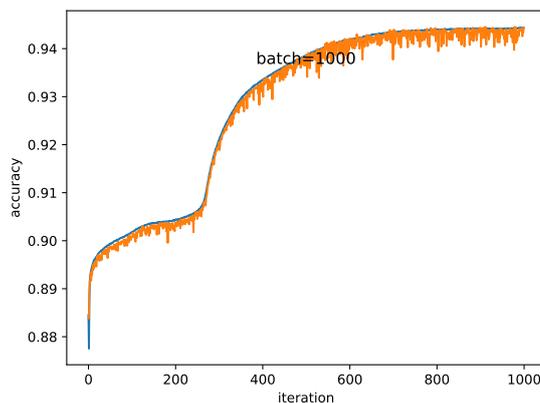


Figure 4.5: Training convergence plot.

In the training convergence plot, the batch size, which is equal to 1000, represents the number of training samples used in one iteration, while the number of epochs or iterations, which is also equal to 1000, represents the number of times the learning algorithm runs through the entire training dataset. As we

can see, both the test accuracy and the train accuracy roughly converge to a constant for large number of iterations. We then decided to stop the training there.

4.3 Event selection optimization

After the training, we obtained three DNN variables relative to the classes VBF, ggH and background, which were used to optimize the event selection. This was done in the following way. For each sample of signal and of background, we filled a three-dimensional histogram with the cut on the VBF variable in the x axis, with the cut on the ggH variable in the y axis and with the cut on the background variable in the z axis. We filled the histogram with the statistical significance computed with the number of events that pass the *VBF preselection* in Eq. 3.6 expected in 41.5 fb^{-1} , adding a selection $120 < m_{\gamma\gamma} < 130 \text{ GeV}$, i.e. that $m_{\gamma\gamma}$ is within 5σ with respect to the expected peak. We add this cut on $m_{\gamma\gamma}$ since, as will be shown in Sec. 6.1, we will extract the signal with a fit to the invariant mass of the pairs of photons: in this case, the statistical significance is determined by the amounts of signal and background present in a region where most of the signal is found.

For a fixed triplet of cuts $(x_{min}, y_{max}, z_{max})$, we defined the number of events in a histogram of a particular sample as the following integral:

$$N^{sample} = \int_{x_{min}}^1 dx \int_0^{y_{max}} dy \int_0^{z_{max}} dz N^{sample}(x, y, z) \quad (4.4)$$

where along x we integrated from x_{min} to 1, since the selection on the VBF probability is $prob(VBF) > x$, while along y and along z we integrated from 0 to y_{max} and from 0 to z_{max} respectively, since for the ggH and bkg outputs we require an upper limit, $prob(ggH) < y$ and $prob(bkg) < z$.

Then we defined the statistical significance in the following way:

$$statistical\ significance = \frac{s}{\sqrt{s+b}} \quad (4.5)$$

where s and b are the number of events, calculated with Eq. (4.4), for the samples “signal” and “background”, as described below:

$$s = s_1 + s_2 \quad (4.6)$$

$$s_1 = f \cdot N^{VBF\ SM} \quad (4.7)$$

$$s_2 = (1 - f) \cdot N^{VBF\ CP\ odd} \quad (4.8)$$

$$b = b_1 + b_2 + b_3 \quad (4.9)$$

where b_1 , b_2 and b_3 are the number of ggH, $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ events, respectively, normalized to 41.5 fb^{-1} , and the value of the fraction $f = 90\%$ was chosen since we expect a small contribution from BSM couplings, and we used a fraction of 10% of CP odd signal, as an example.

Then we looped over the triplet of cuts (x, y, z) , each from 0 to 1, in steps of 1% probability of each category, and we found the maximum of the statistical significance given in Eq. (4.5).

In correspondence of the maximum of the significance, we found the optimal values of the cuts for the three DNN variables:

- $DNN\ VBF > 0.29$;
- $DNN\ ggH < 0.71$;
- $DNN\ background < 0.09$.

To show graphically the procedure of the event selection optimization, a two-dimensional scan of the statistical significance was performed. This was done keeping fixed one of the cuts on the DNN outputs at a time, namely fixing one of the cuts of the triplet (x, y, z) at a time in correspondence of the value that maximized the statistical significance, and performing a two-dimensional scan on the other two cuts.

For example, fixing the cut on $DNN\ VBF$ means that in the integral in Eq. (4.4) we integrate in dx from the x that maximizes the statistical significance, X , to 1, and we leave the rest unchanged, so that the number of events becomes:

$$x - \text{maxed } N^{\text{sample}} = \int_X^1 dx \int_0^{y_{\text{max}}} dy \int_0^{z_{\text{max}}} dz N^{\text{sample}}(x, y, z) \quad (4.10)$$

At this point, we define a two-dimensional histogram, filled with the value of the statistical significance which corresponds to the $DNN\ ggH$ and $DNN\ bkg$ cuts. The statistical significance is calculated from Eq. (4.5), using the values of s and b in Eq. (4.6), (4.7), (4.8) and (4.9), calculated using this time the x -maxed number of events in Eq. (4.10). Thus we obtain the two-dimensional histogram in Fig. 4.6.

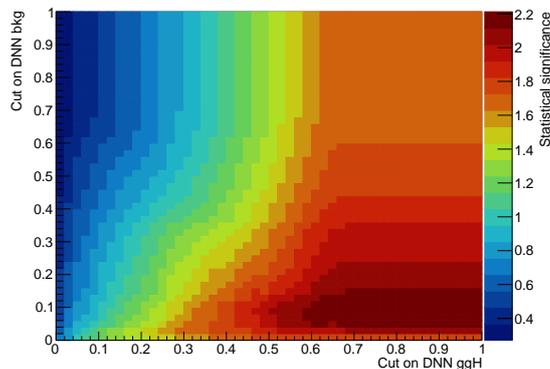


Figure 4.6: Two-dimensional histogram of the statistical significance, fixing the cut on $DNN\ VBF$.

We proceed in a similar way for the cuts on $DNN\ ggH$ and $DNN\ bkg$. Thus we obtain the two-dimensional histograms in Fig. 4.7.

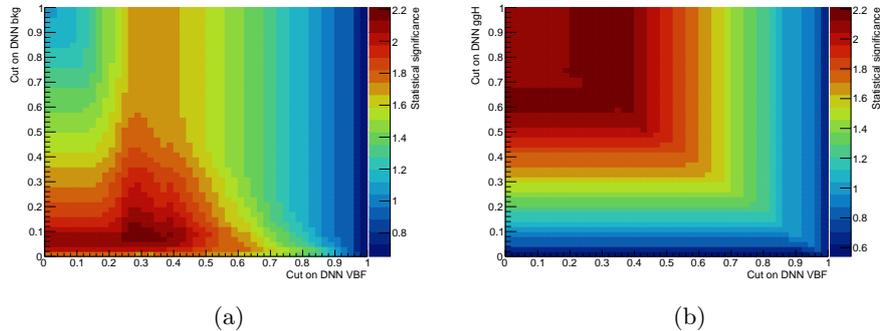


Figure 4.7: (a) Two-dimensional histogram of the statistical significance, fixing the cut on $DNN\ ggH$; (b) two-dimensional histogram of the statistical significance, fixing the cut on $DNN\ bkg$.

We then constructed the Tab. of the selection efficiencies 4.3, and the Tab. of the number of the expected events 4.4, using the $VBF\ preselection$ of Eq. 3.6 and the DNN cuts of Eq. 4.3.

	VBF Λ_1	VBF CP odd	VBF CP even	VBF SM	ggH	$\gamma + jets$	$\gamma\gamma + jets$
$ \eta_\gamma $	0.99	0.99	0.99	0.99	0.99	0.99	0.99
$m_{\gamma\gamma}$	0.98	0.98	0.98	0.99	1.00	1.00	0.50
$(p_T/M)_\gamma$	0.95	0.94	0.93	0.90	0.93	0.83	0.85
$ \Delta\eta _{jj}$	0.84	0.91	0.91	0.56	0.20	0.09	0.28
$(p_T)_j$	0.90	0.95	0.95	0.77	0.46	0.25	0.57
m_{jj}	0.91	0.84	0.85	0.79	0.33	0.26	0.36
MVA_γ^{1D}	0.70	0.74	0.73	0.73	0.72	0.10	0.65
$DNN\ VBF$	0.90	0.86	0.77	0.53	0.34	0.04	0.11
$DNN\ ggH$	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$DNN\ bkg$	0.80	0.73	0.62	0.61	0.56	0.00	0.17

Table 4.3: Selection efficiencies of the single cut respect to the previous one for different hypotheses, with the $VBF\ preselection$ and the DNN cuts.

	VBF Λ_1	VBF CP odd	VBF CP even	VBF SM	ggH	$\gamma + \text{jets}$	$\gamma\gamma + \text{jets}$
$ \eta_\gamma $	228	184	198	167	2079	474253	608055
$m_{\gamma\gamma}$	223	180	194	165	2070	472854	302929
$(p_T/M)_\gamma$	212	169	180	149	1927	394648	256841
$ \Delta\eta _{jj}$	179	154	164	83	378	35391	72142
$(p_T)_j$	162	147	156	64	173	8963	41207
m_{jj}	147	123	132	50	57	2327	15024
MVA_γ^{1D}	103	91	96	37	41	241	9802
<i>DNN VBF</i>	93	78	74	20	14	10	1099
<i>DNN ggH</i>	93	78	74	20	14	10	1099
<i>DNN bkg</i>	75	57	46	12	8	0	191

Table 4.4: Number of events expected in 41.5 fb^{-1} for different hypotheses, with the *VBF preselection* and the DNN cuts.

From Tab. 4.4, we can see that more signal events pass the whole selection respect to Tab. 4.2 which was made using the selection optimized for the SM analysis, especially for the anomalous signals. We therefore conclude that the DNN was trained correctly and that the event selection was optimized in order to select correctly the VBF events as signal and the other samples as background.

4.4 DNN variables

For the next steps of this thesis, we do not use the cuts in the DNN variables in Eq. 4.3 in our event selection, since even if they are more efficient than the SM MVAs on the anomalous coupling samples, they still reduce of about 40% the number of SM VBF events, obtaining only 12 events at the end of the selection for the case of an integrated luminosity of 41.5 fb^{-1} , as we can see from Tab. 4.3 and 4.4. Instead, we will use the shape of the DNN variables to perform a statistical check using the sPlots technique, as we will see in Sec. 6.2.

Chapter 5

Comparison between data and Monte Carlo

5.1 Diphoton loose control sample

We can compare the data sample collected by CMS in 2017 and the different MC samples at our disposal in the case of the diphoton events. In order to do this, we choose a loose preselection, which differs from the *VBF preselection* in Eq. 3.6 by the following cuts:

- $MVA_{\gamma\gamma}^{ID} > -0.2$;
- $m_{jj} > 100$ GeV;
- $(p_T)_{j_1} > 30$ GeV.

This preselection also lacks the acceptance cuts of the photons in $|\eta_\gamma|$ that were present for the *VBF preselection*. As we can see, this preselection is looser than the *VBF preselection*. In fact, for the comparison between data and MC we are less interested in having the highest purity for the Higgs boson signal, but we are interested in having the highest statistics of diphoton events in a similar phase space of the signal, in order to have enough precision to compare the kinematic and photon identification (ID) variables.

The number of expected events that pass the preselection in Eq. 5.1 is given in Tab. 5.1.

	VBF CP odd	VBF CP even	VBF SM	ggH	γ + jets	$\gamma\gamma$ + jets	Data
Loose preselection	118	125	51	114	1739	27448	47549

Table 5.1: Number of expected events in 41.5 fb^{-1} for different MC samples and for the data, for the loose preselection.

We do not apply any correction to the shape of the variables or to the efficiencies of photons or jets.

As we can see from Tab. 5.1, the γ + jets and the $\gamma\gamma$ + jets events are respectively one and two orders of magnitude larger than the VBF events, so we

will neglect its contribution in this control region in what follows. We notice that the number of data events is roughly two times the $\gamma\gamma + \text{jets}$ events. Therefore there is a different normalization between the data and the MC samples of $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$. We can adjust it by performing a two-component fit procedure to one of the discriminating variables at our disposal for the $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ samples, in order to find a scale factor that we will apply to the MC samples, so that the number of events of data and MC will become comparable, and we can then compare the shapes of the kinematic variables between the data and the simulation.

To this end, we choose $|\Delta\eta|_{jj}$, which is the most discriminating variable between the $\gamma + \text{jets}$ and the $\gamma\gamma + \text{jets}$ samples. This is best seen from the plot of the histograms normalized to their area for the two samples in Fig. 5.1.

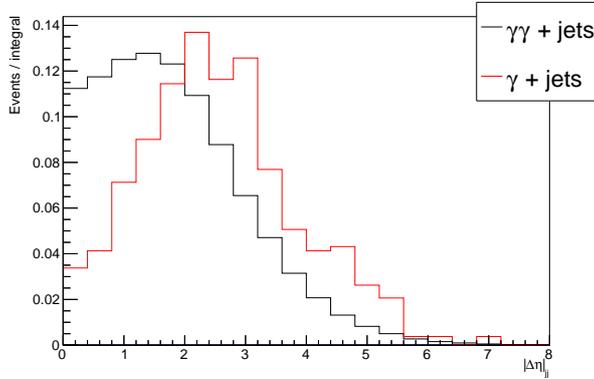


Figure 5.1: Normalized histograms of $|\Delta\eta|_{jj}$ for the $\gamma\gamma + \text{jets}$ and $\gamma + \text{jets}$ samples.

The fit procedure is articulated in the following steps:

- 1) fit the distribution of $|\Delta\eta|_{jj}$ for the $\gamma + \text{jets}$ sample, using the Probability Density Function (PDF):

$$f_{\gamma j} = C_1 G_1 + (1 - C_1) G_2 \quad (5.1)$$

- 2) fit the distribution of $|\Delta\eta|_{jj}$ for the $\gamma\gamma + \text{jets}$ sample, using the PDF:

$$f_{\gamma\gamma j} = C_2 G_3 + (1 - C_2) G_4 \quad (5.2)$$

- 3) fit the distribution of $|\Delta\eta|_{jj}$ for the data sample, using the PDF:

$$f = N f_{\gamma j} + (1 - N) f_{\gamma\gamma j} \quad (5.3)$$

The fits of $|\Delta\eta|_{jj}$ for the $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ samples are used to find the functional form respectively of $f_{\gamma j}$ and $f_{\gamma\gamma j}$, which are continuous functions that can be used to fit the data sample.

For steps 1) and 2), we use both for the PDF $f_{\gamma j}$ in Eq. (5.1) and $f_{\gamma\gamma j}$ in Eq. (5.2) the sum of two Gaussians G_i to fit the shape of the respective

MC distribution, with a normalization parameter C_i that takes into account the fraction of one Gaussian respect to the other in the sum.

In Fig. 5.2 we show the plots of the fits.

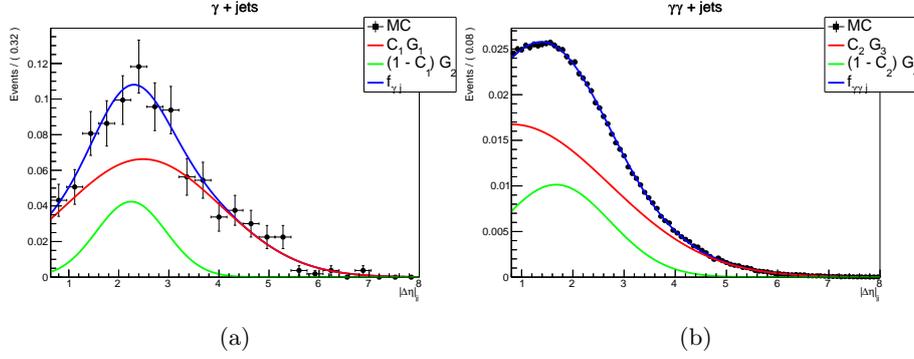


Figure 5.2: (a) Fit to the $|\Delta\eta|_{jj}$ distribution for the $\gamma + \text{jets}$ sample; (b) fit to the $|\Delta\eta|_{jj}$ distribution for the $\gamma\gamma + \text{jets}$ sample.

Each of the two fits yields the means and the standard deviations of the two Gaussians and the normalization parameters C_i . These parameters are left floating and the fit procedure adjusts them in order to better fit the shape of the MC distributions.

For step 3) we use the total PDF f in Eq. (5.3), which fits the distribution of the data, obtaining the plot shown in Fig. 5.3.

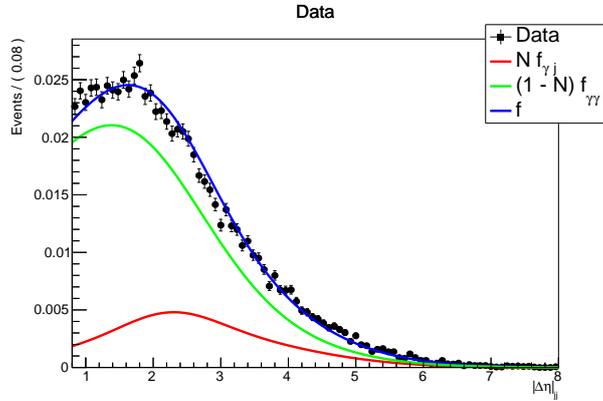


Figure 5.3: Fit to the $|\Delta\eta|_{jj}$ distribution for the data sample.

The fit yields the value of the normalization parameter N , which is left floating and which the fit procedure adjusts in order to better fit the shape of the data distribution.

Now we can compare the data and the MC samples, by scaling the $\gamma + \text{jets}$ and the $\gamma\gamma + \text{jets}$ samples respectively by the following factors:

$$scale(\gamma + jets) = \frac{Integral(data) \cdot N}{Integral(\gamma + jets)} \quad (5.4)$$

$$scale(\gamma\gamma + jets) = \frac{Integral(data) \cdot (1 - N)}{Integral(\gamma\gamma + jets)} \quad (5.5)$$

Using these scales, we can now plot in Fig. 5.4, 5.5 and 5.6 the distributions of some variables for the stack of $\gamma + jets + \gamma\gamma + jets$ and of the data.

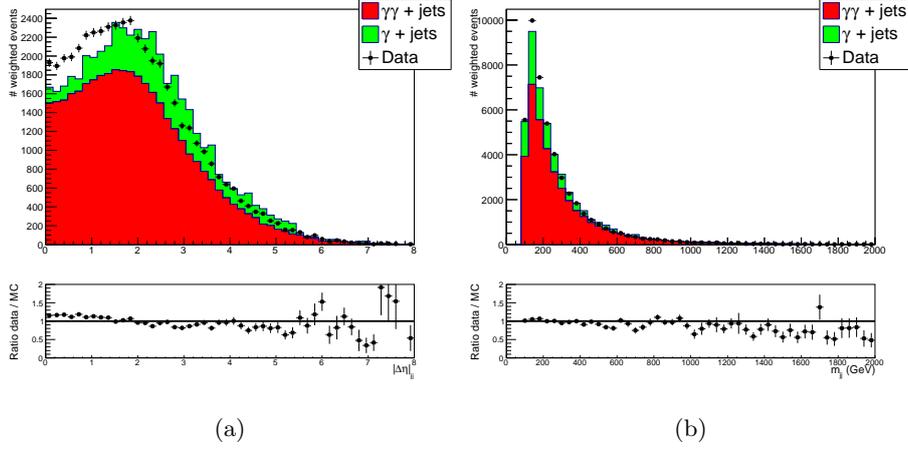


Figure 5.4: (a) Scaled histograms for $|\Delta\eta|_{jj}$ for the data and for the stack of $\gamma + jets$ and $\gamma\gamma + jets$; scaled histograms for m_{jj} for the data and for the stack of $\gamma + jets$ and $\gamma\gamma + jets$.

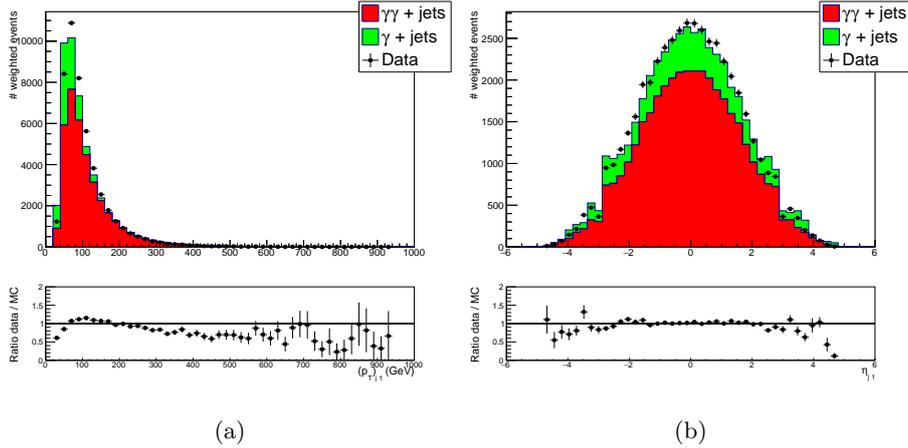


Figure 5.5: (a) Scaled histograms for $(p_T)_{j_1}$ for the data and for the stack of $\gamma + jets$ and $\gamma\gamma + jets$; (b) scaled histograms for η_{j_1} for the data and for the stack of $\gamma + jets$ and $\gamma\gamma + jets$.

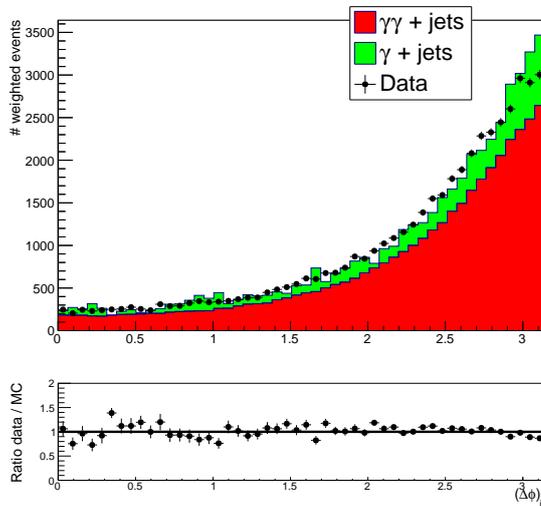


Figure 5.6: Scaled histogram for $\Delta\phi_{jj}$ for the data and for the stack of $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$.

As we can see from the previous plots, the difference in the normalization that was present in Tab. 5.1 between the background samples $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ and the data sample has been significantly reduced by scaling the histograms with the normalization factors obtained from the fit procedure, and a fair agreement was obtained in the shape of the variables interesting for our selection of the events between the data and the stacked MC samples.

After the normalizations of the components are adjusted by the fit, there are still some discrepancies in the $|\Delta\eta|_{jj}$ distribution in Fig. 5.4 between data and MC samples. This may be due to different reasons. Either the shape of $\gamma + \text{jets}$ is not entirely correct, or with the loose preselection in Eq. 5.1 used in this context there are other components which can give a relevant contribution, such as $Z + \text{jets}$, $W + \text{jets}$, $t\bar{t}$.

Both the MC for $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ are simulation at leading order in QCD, where there is no hard emission of gluons and therefore of jets. For this reason, the simulation of p_T in Fig. 5.5 is slightly shifted respect to the data.

5.2 $Z \rightarrow e^+e^- + 2 \text{ jets}$ control sample

In order to study the performances of the selections on a sample that presents characteristics similar to that of the signal, it is possible to select a control sample with large statistics, high purity and free of bias using the Tag & Probe technique, by studying the decay of the Z boson into e^+e^- .

We use the $Z \rightarrow e^+e^-$ events to build a sample with isolated electrons. Using the approximation where a reconstructed electron approximates a reconstructed photon, apart from the presence of a track, the electrons can be used in the analysis of the photons to probe the identification performances and to study the discrepancies between the efficiencies of the data and of the MC simulations. The electron/photon approximation is precise and moreover it is justified by the

fact that both are reconstructed with the same algorithms. Both are particles that are completely absorbed by the ECAL, where the behaviour of a 45 GeV electron originating from the decay of a Z boson is similar to the behaviour of a 60 GeV photon originating from the decay of a Higgs boson.

The $Z \rightarrow e^+e^-$ events are selected by the following requirements:

- we require that an electron (tag) passes a very tight selection. This selection, in order to identify the electrons, uses variables that exploit the geometrical correspondence and the energy compatibility between the ECAL energy deposits and the tracks in the tracker, and the ratio between the ECAL and the HCAL energy deposits. It also applies cuts on the isolation variables;
- the other electron (probe) is selected requiring that the combined mass of the tag-probe system is within a range around the mass of the Z boson.

In this way, the probe electrons constitute a pure and bias free sample, and they are given in input to the photon identification system in order to study its performances.

In the event selection, we require that the two electrons are well reconstructed and identified with an invariant mass near to the mass of the Z boson.

We can distinguish three cases, depending on the ECAL region where the two electrons are identified:

- both electrons are in the barrel (both $|\eta| < 1.479$);
- one electron is in the barrel ($|\eta| < 1.479$) and the other is in the endcap ($|\eta| > 1.479$);
- both electrons are in the endcaps (both $|\eta| > 1.479$).

For both electrons in the barrel, the event selection used is the following:

- $|\eta_{e_1}| < 2.5$, $|\eta_{e_2}| < 2.5$, $|\eta_{e_1}| < 1.44$ or $|\eta_{e_1}| > 1.57$, $|\eta_{e_2}| < 1.44$ or $|\eta_{e_2}| > 1.57$;
- $80 < m_{ee} < 100$ GeV;
- $(p_T)_{e_1}/m_{ee} > 0.3$, $(p_T)_{e_2}/m_{ee} > 0.3$;
- $MVA_{e_1}^{ID} > 0.5$, $MVA_{e_2}^{ID} > 0.5$;
- $|\Delta\eta|_{jj} > 0.0$, $|\eta_{j_1}| < 4.7$, $|\eta_{j_2}| < 4.7$, $\min(\Delta R_{jet\ e}) > 0.4$;
- $(p_T)_{j_1} > 40$ GeV, $(p_T)_{j_2} > 30$ GeV;
- $m_{jj} > 250$ GeV;
- $|\eta_{e_1}| < 1.479$, $|\eta_{e_2}| < 1.479$.

The cut in m_{ee} means that we select the pairs of electrons with an invariant mass centered around the peak of the Z boson. As we said for the *VBF preselection* in Eq. 3.6, the cut in $(p_T)_e/m_{ee}$ is applied in order not to modify the invariant mass distribution of the electrons for the background. Moreover, the cut in MVA_e^{ID} is applied in order to reduce the events with multijets, $t\bar{t}$ and

single t production that yield fake electrons. The previous cuts assume that the electrons are in the barrel; for electrons in the endcaps, we have $|\eta_e| > 1.479$.

In the previous event selection we also specified some cuts for the jets, whose experimental effects are the same both for the Higgs boson and for the Z boson events. Therefore, we can check the agreement between data and MC samples of both electrons and jets variables, using the cuts in Eq. 5.2. In Fig. 5.7, 5.8 and 5.9 we report some histograms normalized to their area.

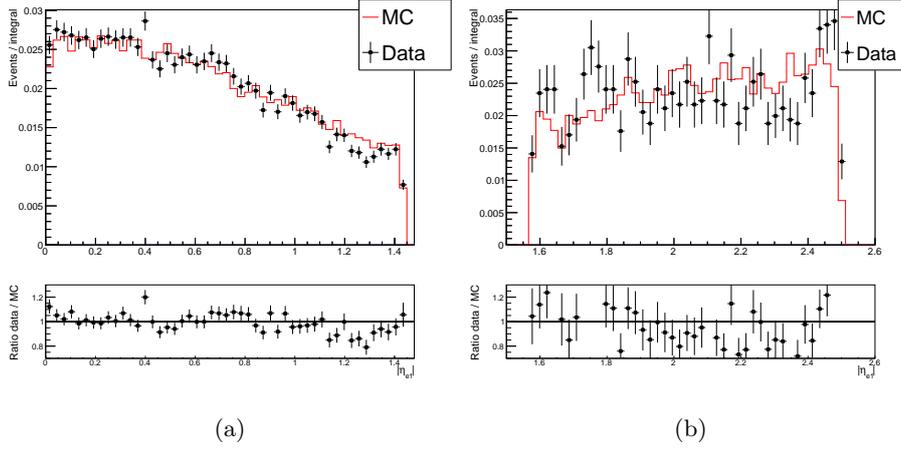


Figure 5.7: (a) Normalized histogram for $|\eta_{e1}|$ for the data and for the MC samples, for both electrons in the barrel; (b) normalized histogram for $|\eta_{e1}|$ for the data and for the MC samples, for both electrons in the endcaps.

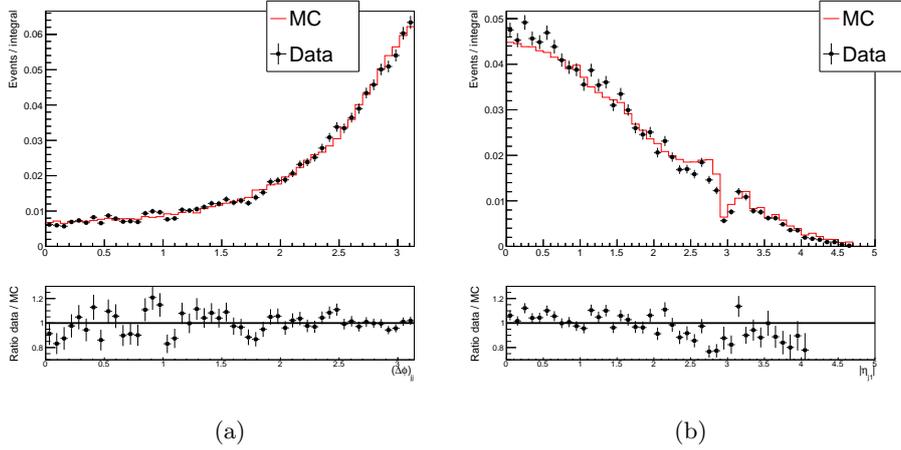


Figure 5.8: (a) Normalized histogram for $(\Delta\phi)_{jj}$ for the data and for the MC samples, for both electrons in the barrel; (b) normalized histogram for $|\eta_{j1}|$ for the data and for the MC samples, for both electrons in the barrel.

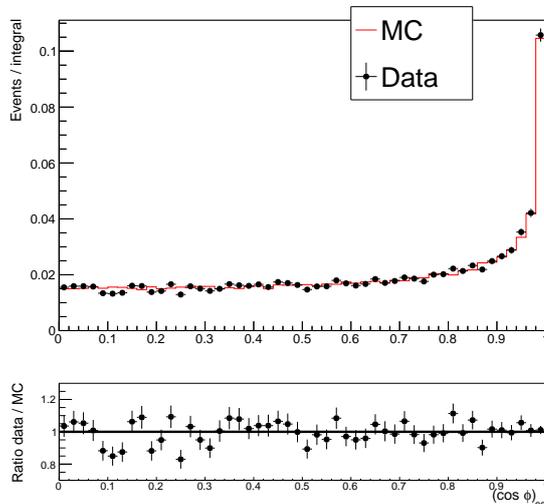


Figure 5.9: Normalized histogram for $(\cos \Delta\phi)_{ee}$ for the data and for the MC samples, for both electrons in the barrel.

We can see that there is a fair agreement between data and MC samples.

The ECAL and the HCAL both cover up to $|\eta| = 3$. The forward jets that travel between $3 < |\eta| < 3.5$ are reconstructed with the HF, a forward hadronic calorimeter made with scintillating fibers, which resists at high levels of radiation at high $|\eta|$. The fall that we can see in $|\eta_{j_1}|$ in Fig. 5.8 at $|\eta| = 3$ is an acceptance hole that corresponds to the transition region between the HCAL and the HF, where we measure less events.

We can also make a comparison between data and MC for the calorimetric variables of the electrons, using an event selection similar to the one in Eq. 5.2, apart from the fact that we do not apply cuts in the jets variables; we also put a constraint in MVA^{ID} only on the leading electron, the tight one ($MVA_{e_1}^{ID} > 0.5$), and we assume that only the subleading electron, the probe one, is in the barrel ($|\eta_{e_2}| < 1.479$). Using this new event selection, we can compare data and MC samples in the histograms in Fig. 5.10 and 5.11 normalized to their area for the probe electron.

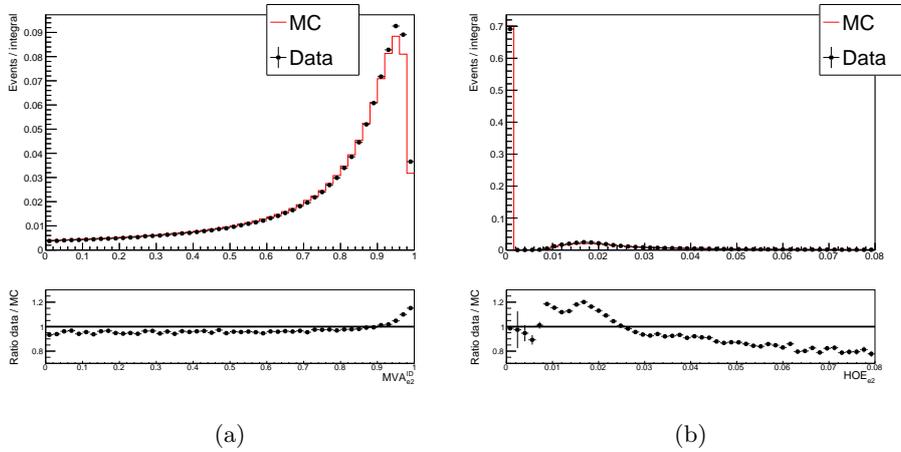


Figure 5.10: (a) Normalized histogram for $MVA_{e_2}^{ID}$ for the data and for the MC samples, for the subleading electron in the barrel; (b) normalized histogram for HOE_{e_2} for the data and for the MC samples, for the subleading electron in the barrel.

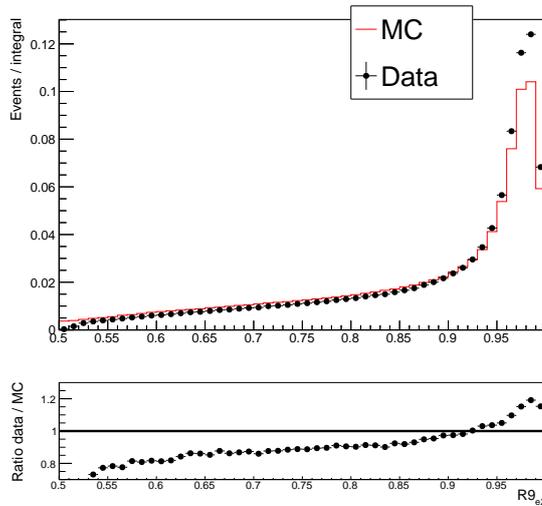


Figure 5.11: Normalized histogram for $R9_{e_2}$ for the data and for the MC samples, for the subleading electron in the barrel.

HOE stands for the ratio of the energy deposited in the HCAL and the energy of the supercluster in the ECAL. For electrons and photons, HOE is approximately 0, whereas for jets HOE would have a large distribution. Therefore HOE could be used as a discriminating variable between electromagnetic particles, such as electrons/photons, and hadronic objects, such as jets. In Fig. 5.10, we can see that most of the events of HOE_{e_2} are peaked around 0, as we expect for the probe electron. We can also note that there is a small bump

around 1.5%, which is due to the energy associated to the pileup. In this case, the electron is absorbed in the ECAL, and the hadrons coming from the pileup end up in the HCAL. Moreover there is some noise in the detector: some towers of the HCAL are over the threshold and present an energy greater than 0. This means that HOE is not exactly equal to 0. Both the MC and the data show the effects of the noise of the detector and of the pileup. However, the two components are difficult to describe in the MC. In fact, the MC is usually done at the beginning of the Run: at a later time, the radiation damage and the ageing of the photodetectors yield noise in the detector, which will give a different response respect to the MC in the data. These are the sources of discrepancies between data and MC.

$R9$ is defined as the ratio of the energy in the 3×3 matrix around the seed and the energy of the whole supercluster:

$$R9 = \frac{E_{3 \times 3}}{E_{SC}} \quad (5.6)$$

and it is used to separate the converted photons from the photons that did not interact. The side of the crystals of the ECAL was chosen to be approximately equal to the Molière radius of the PWO. Therefore the lateral development of the cluster in the shower is contained in the central seed crystal and in a crown made of a 3×3 matrix of crystals around it, as said before in Sec. 3.4. In fact the PWO was chosen for this reason, in a way that a 3×3 matrix of crystals around the central crystal contains almost all the energy of a photon/electron colliding in the seed. In Fig. 5.11 we see that there is a peak around 1, therefore for $R9$ we find the behaviour that we expect for the electrons. For the jets $R9$ is a small quantity. In fact, the distribution of the particles in a jet is quite scattered, and is usually not fully contained in a 3×3 matrix.

The data and MC samples show a similar behaviour. Therefore, we assume that the MC describes well all these variables also for photons.

It is possible to use the ratio between data and MC in the case of the $Z \rightarrow e^+e^-$ distributions to correct the MC distributions of the $H \rightarrow \gamma\gamma$ case. In fact, the calorimetric variables of the electrons originating from the Z decay have a similar behaviour to those of the photons originating from the H decay. For example, we can apply the corrections obtained from the HOE ratio between data and MC of the $Z \rightarrow e^+e^-$ case to modify the MC distribution of the same variable in the $H \rightarrow \gamma\gamma$ case. For the purpose of this thesis, this correction was not done, and we only checked that the comparison between data and MC gave good results for the case of the Z .

Chapter 6

Statistical signal extraction

6.1 Invariant mass fit

In Chap. 4, we performed the DNN training and the event selection optimization without using the invariant mass of the pairs of photons. Now we are ready to extract the signal from the peak of $m_{\gamma\gamma}$ above a smoothly falling background, by parametrizing the signal and the background with different PDFs.

We know that both VBF SM and ggH samples show a resonant peak in the $m_{\gamma\gamma}$ distribution. We can ask ourselves if the same argument is valid for the anomalous VBF samples. As we saw in Fig. 3.5, the anomalous VBF signals have the same shape of the VBF SM, therefore the anomalous couplings do not modify the invariant mass distribution of the pairs of photons.

We can also check quantitatively that the signal shape does not depend on the Higgs boson coupling by fitting two signal samples, VBF SM and VBF Λ_1 , for the invariant mass of the pairs of photons passing the *VBF preselection* in Eq. 3.6, in order to extract the resonant signal above the smoothly falling background. To fit the two MC distributions, we use in both cases the sum of three Gaussians G_i : one is needed to emulate the tail which is present in both distributions for low invariant masses; the other two are needed to emulate the resonant peak. Therefore for both distributions we use the following PDF:

$$f = C_1 G_1 + C_2 G_2 + (1 - C_1 - C_2) G_3 \quad (6.1)$$

We leave as free parameters the means and the standard deviations of the three Gaussians, and also the relative fractions of each Gaussian respect to the others in the sum C_i : the fits adjust all the parameters in order to better fit the shape of the distributions.

The histograms of $m_{\gamma\gamma}$ on which we perform the fit procedure are obtained by applying the *VBF preselection* in Eq. 3.6 with the tighter request $115 < m_{\gamma\gamma} < 135$ GeV.

For the VBF SM sample, we obtain the plot in Fig. 6.1.

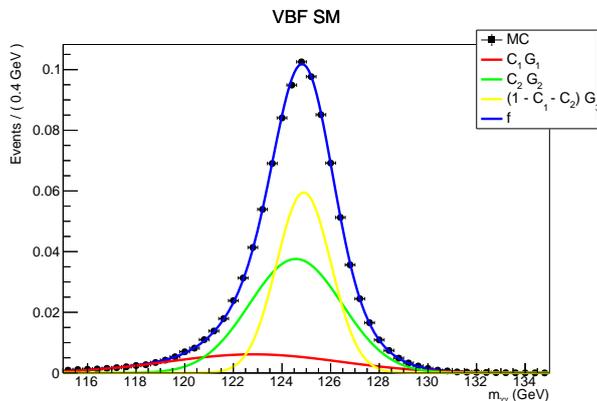


Figure 6.1: Fit to the distribution of $m_{\gamma\gamma}$ for the VBF SM sample.

For the VBF Λ_1 sample, we obtain the plot in Fig. 6.2.

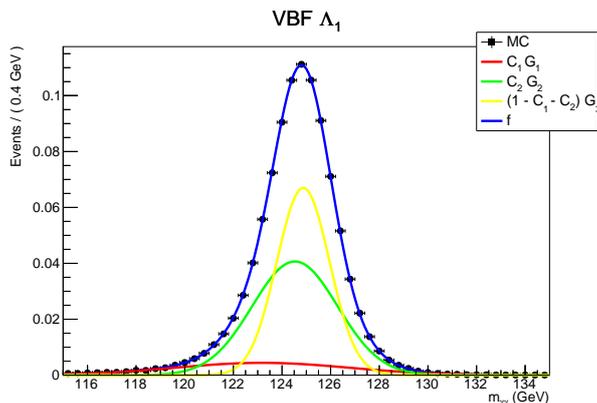


Figure 6.2: Fit to the distribution of $m_{\gamma\gamma}$ for the VBF Λ_1 sample.

In Tab. 6.1 we report the means and the standard deviations of the three Gaussians G_i with their errors for the VBF SM and VBF Λ_1 samples.

	VBF SM	VBF Λ_1
Mean G_1 (GeV)	122.8 ± 6.6	123.2 ± 8.0
Mean G_2 (GeV)	124.6 ± 7.2	124.5 ± 6.0
Mean G_3 (GeV)	124.9 ± 4.1	124.9 ± 3.7
Standard deviation G_1 (GeV)	3.6 ± 1.4	3.6 ± 1.5
Standard deviation G_2 (GeV)	1.9 ± 1.5	1.8 ± 1.3
Standard deviation G_3 (GeV)	1.1 ± 1.2	1.1 ± 1.1

Table 6.1: Means and standard deviations of the three Gaussians G_i with their errors for the VBF SM and VBF Λ_1 samples.

As we can see from Tab. 6.1, the means and the standard deviations of each of the three Gaussians are consistent between VBF SM and VBF Λ_1 samples

within one standard deviation from each other. We therefore conclude that we can use the same PDF for all the VBF signals (both SM and BSM) and for ggH.

To find the PDF of the inclusive background, we do not rely at all on the MC, because we have shown in Sec. 5.1 that it does not describe well the variables for the main backgrounds. Thus we fit directly the data distribution blinding the signal region between 120 and 130 GeV (which means that we do not consider the region where we expect to find the signal), in order to find the parameters of the PDF of the background, which we choose to be a second order polynomial which decreases with $m_{\gamma\gamma}$:

$$g = 1 + Am_{\gamma\gamma} + Bm_{\gamma\gamma}^2 \quad (6.2)$$

We do not care about the contributions of $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ to the total background: we assume that everything that is not resonant is smooth. The result is shown in Fig. 6.3.

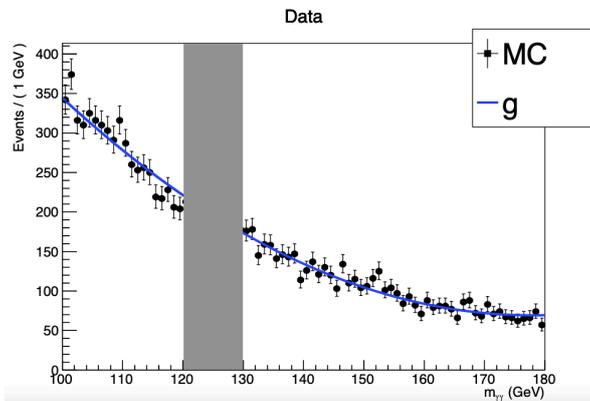


Figure 6.3: Fit to the distribution of $m_{\gamma\gamma}$ for the data sample, blinding the region between 120 and 130 GeV.

Also in this case the histogram of $m_{\gamma\gamma}$ on which we performed the fit procedure was obtained by applying the *VBF preselection* in Eq. 3.6.

We can now fit the data sample in the whole range $100 < m_{\gamma\gamma} < 180$ GeV, unblinding the signal region between 120 and 130 GeV. In order to do this, we choose the following PDF:

$$h = N_{sig}f + N_{bkg}g \quad (6.3)$$

where this time for f , which is given in Eq. (6.1), we choose to use a unique mean for the two Gaussians that model the resonant peak, and we leave this parameter floating in the fit; g is given in Eq. (6.2). We also leave as free parameters the fractions N_{sig} and N_{bkg} respectively of the signal and of the background in Eq. (6.3). All these parameters are yielded by the fit procedure, which was done with an extended likelihood with a Poissonian number of events:

$$Poisson(data|N_{sig} + N_{bkg}) = \prod_i \frac{(N_{sig,i} + N_{bkg,i})^{n_i}}{n_i!} e^{-N_{sig,i} - N_{bkg,i}} \quad (6.4)$$

which is a product of Poisson probabilities to observe n_i events in bins i . When the bin size is infinitely small, an unbinned likelihood is used for the data sample with k events:

$$k^{-1} \prod_i [N_{sig} f(m_{\gamma\gamma,i}) + N_{bkg} g(m_{\gamma\gamma,i})] e^{-N_{sig} - N_{bkg}} \quad (6.5)$$

We choose to extract the signal with such an unbinned maximum likelihood fit to use the information of the single event, which is the optimal approach with finite statistics.

The result is shown in Fig. 6.4.

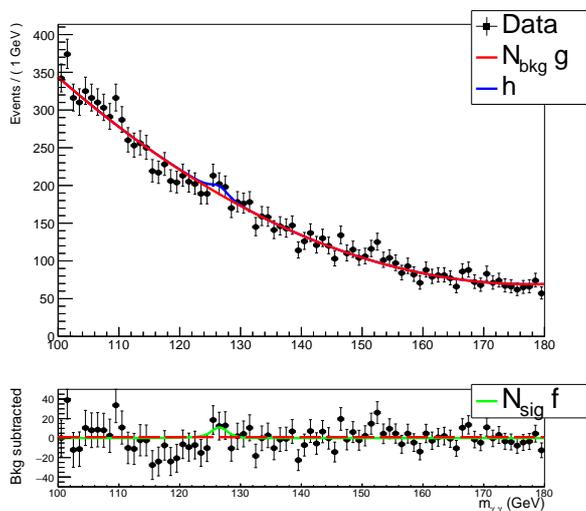


Figure 6.4: Fit to the distribution of $m_{\gamma\gamma}$ for the data sample.

We neglect the systematic uncertainties associated to the parametrization of the background as a decreasing polynomial and to the parametrization of the signal with a sum of Gaussians. In fact, regarding the signal, there are systematic uncertainties in the energy scale that can shift the position of the peak and in the energy resolution that may change the width of the peak. We did not consider these systematic uncertainties. A more detailed analysis on this topic can be found in [32].

The number of signal events N_{sig} and its error $\sigma(N_{sig})$ that we obtain from the fit are given by:

$$N_{sig} \pm \sigma(N_{sig}) = 48 \pm 37(stat) \quad (6.6)$$

We now compare the number of signal events obtained from the fit in Eq. (6.6) to the sum of the number of expected events from ggH and different VBF processes after the *VBF preselection* in Eq. 3.6. This is shown in Tab. 6.2.

	ggH	VBF Λ_1	VBF CP odd	VBF SM	Fit result
<i>VBF preselection</i>	41	103	91	37	48 ± 37

Table 6.2

Without looking at any angular distribution or at other variables, we can directly see from the expected and measured number of events that the fit yields a result which is in good agreement with the sum of ggH and VBF SM, consistent with it within one σ , while the sum of ggH and other VBF samples are more than two σ away from the result of the fit.

In addition to comparing the expected and measured number of events, we can also compare the angular distributions and other variables for the MC samples and for the data, using a particular statistical background subtraction technique.

6.2 Statistical results with sPlots

We know that the Higgs boson signal can be viewed as a resonant peak in the invariant mass distribution over a decreasing background. However, we do not know the shape of the signal for variables that differ from the invariant mass.

In order to find these shapes, we can use the sPlot technique [38]. This procedure is used to statistically subtract the background from the data: what remains is the distribution of the signal only, that uses informations contained in the covariance matrix of the fit.

In order to do this procedure, we first perform a fit to the invariant mass distribution. This fit yields the sWeights, which are quantities proportional to the probability that an event is a signal event. When we plot the distribution of a variable weighting it with the sWeights, we obtain the sPlot, which is the distribution of that variable for the signal only. Furthermore, the integral of the sPlot, by construction, equals the number of signal events.

The invariant mass fit to get the sWeights is the one described in Sec. 6.1. In Fig. 6.5, 6.6 and 6.7 we show the sPlots of some variables, compared with the MC samples with expected events in 41.5 fb^{-1} of the sum of different VBF hypotheses and ggH.

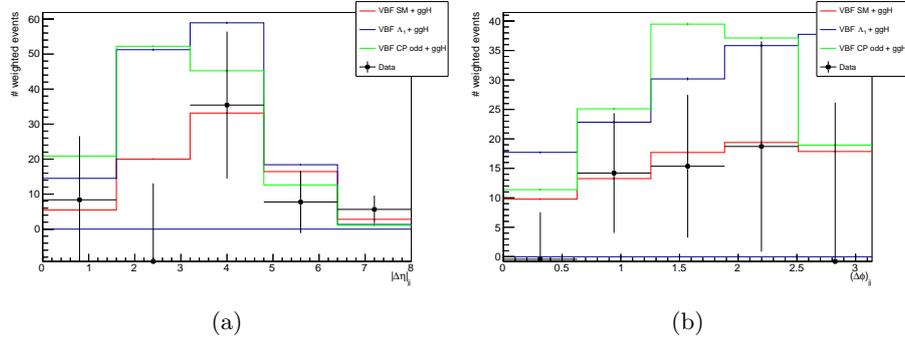


Figure 6.5: (a) sPlot and MC with expected events in 41.5 fb^{-1} of different VBF hypotheses + ggH of $|\Delta\eta|_{jj}$; (b) sPlot and MC with expected events in 41.5 fb^{-1} of different VBF hypotheses + ggH of $(\Delta\phi)_{jj}$.

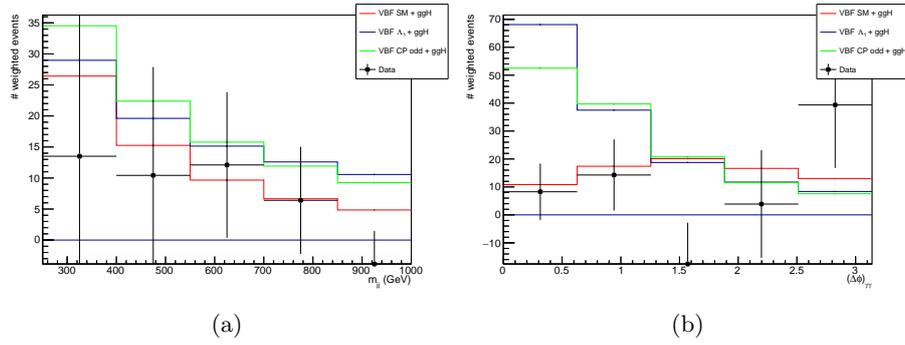


Figure 6.6: (a) sPlot and MC with expected events in 41.5 fb^{-1} of different VBF hypotheses + ggH of m_{jj} ; (b) sPlot and MC with expected events in 41.5 fb^{-1} of different VBF hypotheses + ggH of $(\Delta\phi)_{\gamma\gamma}$.

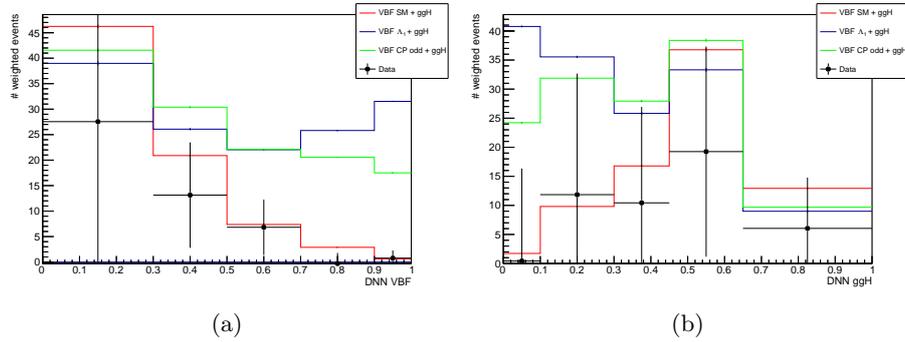


Figure 6.7: (a) sPlot and MC with expected events in 41.5 fb^{-1} of different VBF hypotheses + ggH of $DNN \text{ VBF}$; (b) sPlot and MC with expected events in 41.5 fb^{-1} of different VBF hypotheses + ggH of $DNN \text{ ggH}$.

We plot the sum of different VBF hypotheses and ggH, because the resonant peak of the Higgs boson receives contribution not only from the VBF events but also from the ggH events.

The yield of the fit is more in agreement with the SM respect to the alternative hypotheses, as can be seen from Tab. 6.2. Moreover, the distributions of the single variables in Fig. 6.5, 6.6 and 6.7 are all in better agreement with the SM respect to the anomalous hypotheses: this is particularly evident in the distribution of $DNN\ VBF$ in Fig. 6.7, which combines all the discriminating variables in an optimal way.

In a qualitative way, this excludes the hypotheses Λ_1 and CP odd. However, a more refined fit and all the data of Run2 are necessary in order to find the exclusion limits for the anomalous couplings fractions f_{a_2}, f_{a_3} , while here we have shown only the cases where all the Higgs boson VBF cross sections are due to a given anomalous coupling, i.e. $f_{a_2}, f_{a_3}, f_{\Lambda_1}$.

Conclusions

The CMS and ATLAS experiments at LHC have set constraints on the spin-parity properties of the Higgs boson and anomalous HVV couplings, where V stands for W , Z , and γ electroweak gauge bosons, finding its quantum numbers to be consistent with $J^{PC} = 0^{++}$, but leaving room for small anomalous HVV couplings. In theories BSM, Higgs boson interactions may generate several of them, which lead to new tensor structures of interactions, both CP -even and CP -odd. The couplings to which we are interested are the anomalous couplings of the photons a_2 and a_3 , which in the SM are very small. Therefore measuring values significantly different from zero (of the order of the 10%) of these anomalous couplings would constitute an indication of new physics. This would mean that in the virtual loops in addition to SM particles there are also new particles that have phases different from 0 that can generate CP violation. The purpose of this thesis is to constrain a_2 and a_3 in the VBF production $VV \rightarrow H$ and using the $\gamma\gamma$ decay channel. The anomalous couplings can be inferred by the cross sections that we can experimentally measure.

We look for a signal with a final state made of two photons, coming from the Higgs boson decay, and of two jets, coming from the quarks that hadronize after the VBF scattering. To this end, several MC samples of VBF signal and of background were studied, where the VBF signal includes both SM and anomalous samples, and the background includes gluon gluon fusion, $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ production. A comparison was made between the distributions of several discriminating variables. A deep neural network was trained in order to enhance the discrimination between signal and background samples. A comparison of different nature was made between the distributions of MC samples and of the data sample collected by CMS in 2017, which corresponds to an integrated luminosity of 41.5 fb^{-1} , by using $\gamma\gamma + \text{jets}$ and $\gamma + \text{jets}$ events and the $Z \rightarrow e^+e^-$ control sample. An invariant mass fit was performed to the distribution of the $\gamma\gamma$ invariant mass in the data sample to extract the $H \rightarrow \gamma\gamma$ signal from the large non-resonating background. From the results of the fit, the sPlots were constructed in order to statistically subtract the background from the data. Finally, the sPlots were inspected in order to establish whether the results of the fit are compatible with the SM or the anomalous hypotheses. We have found that the data are in agreement with the SM and a 100% BSM production is qualitatively excluded.

Bibliography

- [1] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.
- [2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [3] Oscar González López. Searches for a high-mass higgs boson in the zz and $w\gamma$ decay channels with the cms detector. *Nuclear and Particle Physics Proceedings*, 273-275:907–912, 2016. 37th International Conference on High Energy Physics (ICHEP).
- [4] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [5] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016.
- [6] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A.A. Abdelalim, O. Abdinov, R. Aben, B. Abi, M. Abolins, and et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, Sep 2012.
- [7] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, and et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, Sep 2012.
- [8] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, and et al. Observation of a new boson with mass near 125 gev in pp collisions at $\sqrt{s} = 7$ and 8 tev. *Journal of High Energy Physics*, 2013(6), Jun 2013.
- [9] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [10] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132–133, 1964.
- [11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

- [12] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, and et al. Study of the mass and spin-parity of the higgs boson candidate via its decays to z boson pairs. *Physical Review Letters*, 110(8), Feb 2013.
- [13] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl, and et al. Measurement of the properties of a higgs boson in the four-lepton final state. *Physical Review D*, 89(9), May 2014.
- [14] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, M. Friedl, R. Frühwirth, V. M. Ghete, and 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*, 92(1), Jul 2015.
- [15] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, and et al. Limits on the higgs boson lifetime and width from its decay to four charged leptons. *Physical Review D*, 92(7), Oct 2015.
- [16] A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogio, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, and et al. Constraints on anomalous higgs boson couplings using production and decay information in the four-lepton final state. *Physics Letters B*, 775:1–24, Dec 2017.
- [17] A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogio, E. Asilar, T. Bergauer, J. Brandstetter, M. Dragicevic, J. Erö, A. Escalante Del Valle, and et al. 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*, 99(11), Jun 2019.
- [18] A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogio, E. Asilar, T. Bergauer, J. Brandstetter, M. Dragicevic, J. Erö, A. Escalante Del Valle, and et al. Constraints on anomalous hvv couplings from the production of higgs bosons decaying to τ lepton pairs. *Physical Review D*, 100(11), Dec 2019.
- [19] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O. S. AbouZeid, and et al. Evidence for the spin-0 nature of the higgs boson using atlas data. *Physics Letters B*, 726(1-3):120–144, Oct 2013.
- [20] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, and et al. Study of the spin and parity of the higgs boson in diboson decays with the atlas detector. *The European Physical Journal C*, 75(10), Oct 2015.
- [21] G. Aad, B. Abbott, O. Abdinov, J. Abdallah, B. Abeloos, R. Aben, M. Abolins, R. Aben, M. Abolins, and et al. Test of cp invariance in vector-boson fusion production of the higgs boson using the optimal observable method in the $di\tau$ decay channel with the atlas detector. *The European Physical Journal C*, 76(12), Nov 2016.

- [22] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al. Measurement of inclusive and differential cross sections in the $h \rightarrow zz^* \rightarrow 4l$ decay channel in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector. *Journal of High Energy Physics*, 2017(10), Oct 2017.
- [23] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al. Measurement of the higgs boson coupling properties in the $h \rightarrow zz^* \rightarrow 4l$ decay channel at $\sqrt{s} = 13$ tev with the atlas detector. *Journal of High Energy Physics*, 2018(3), Mar 2018.
- [24] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S.H. Abidi, O.S. AbouZeid, N.L. Abraham, H. Abramowicz, H. Abreu, and et al. Measurements of higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ tev with the atlas detector. *Physical Review D*, 98(5), Sep 2018.
- [25] A.M. Sirunyan, A. Tumasyan, W. Adam, J.W. Andrejkovic, T. Bergauer, S. Chatterjee, M. Dragicevic, A. Escalante Del Valle, R. Frühwirth, M. Jeitler, and et al. 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*, 104(5), Sep 2021.
- [26] A.M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, T. Bergauer, M. Dragicevic, J. Erö, A. Escalante Del Valle, M. Flechl, R. Frühwirth, and et al. Measurements of $t\bar{t}h$ production and the cp structure of the yukawa interaction between the higgs boson and top quark in the diphoton decay channel. *Physical Review Letters*, 125(6), Aug 2020.
- [27] G. Aad, B. Abbott, D.C. Abbott, A. Abed Abud, K. Abeling, D.K. Abhayasinghe, S.H. Abidi, O.S. AbouZeid, N.L. Abraham, H. Abramowicz, and et al. Cp properties of higgs boson interactions with top quarks in the $t\bar{t}h$ and th processes using $h \rightarrow \gamma\gamma$ with the atlas detector. *Physical Review Letters*, 125(6), Aug 2020.
- [28] Sandor Czellar, Edward Hægström, Aatos Heikkinen, Jaakko Härkönen, Veikko Karimäki, Ritva Kinnunen, Tapio Lampen, Katri Marjaana Lassila-Perini, Sami Lehti, Tomas Linden, Panja-Riina Luukka, S Michal, Teppo Mäenpää, Jukka Nysten, M Stettler, Eija Tuominen, Jorma Tuominiemi, Lauri Wendland, and The CMS Collaboration. *CMS Physics: Technical Design Report : Vol 1 Detector performance and software*, volume LHCC-2006-001. CERN, Switzerland, cms-tdr-008-01 edition, 2006.
- [29] Andrei V. Gritsan, Jeffrey Roskes, Ulascan Sarica, Markus Schulze, Meng Xiao, and Yaofu Zhou. New features in the jhu generator framework: Constraining higgs boson properties from on-shell and off-shell production. *Physical Review D*, 102(5), Sep 2020.
- [30] H.B. Hartanto, B. Jäger, L. Reina, and D. Wackerth. Higgs boson production in association with top quarks in the powheg box. *Physical Review D*, 91(9), May 2015.

- [31] Stefano Frixione and Bryan R Webber. Matching nlo qcd computations and parton shower simulations. *Journal of High Energy Physics*, 2002(06):029–029, Jun 2002.
- [32] A. M. Sirunyan, A. Tumasyan, W. Adam, J. W. Andrejkovic, T. Bergauer, S. Chatterjee, M. Dragicevic, A. Escalante Del Valle, R. Frühwirth, and et al. Measurements of higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s} = 13$ tev. *Journal of High Energy Physics*, 2021(7), Jul 2021.
- [33] A.M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, A. Escalante Del Valle, R. Frühwirth, M. Jeitler, N. Krammer, L. Lechner, and et al. Electron and photon reconstruction and identification with the cms experiment at the cern lh. *Journal of Instrumentation*, 16(05):P05014, May 2021.
- [34] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, Apr 2008.
- [35] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. Fastjet user manual. *The European Physical Journal C*, 72(3), Mar 2012.
- [36] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, Jan 2015.
- [37] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [38] M. Pivk and F.R. Le Diberder. : A statistical tool to unfold data distributions. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 555(1-2):356–369, Dec 2005.