Master's thesis
Physics

Searching for degenerate Higgs bosons using a profile likelihood ratio method

Jaana Heikkilä

May 20, 2015

Tutor: Dr. Giovanni Petrucciani

Censors: Prof. Paula Eerola
Dr. Giovanni Petrucciani

UNIVERSITY OF HELSINKI
Department of Physics

PL 64 (Gustaf Hällströmin katu 2)
00014 Helsingin yliopisto

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Laitos — Institution — Department
Faculty of Science	Department of Physics

Tekijä — Författare — Author

Jaana Heikkilä

Työn nimi — Arbetets titel — Title

Searching for degenerate Higgs bosons using a profile likelihood ratio method

Oppiaine — Läroämne — Subject

Physics

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's thesis	May 2015	100

Tiivistelmä — Referat — Abstract

ATLAS and CMS collaborations at the Large Hadron Collider have observed a new resonance consistent with the standard model Higgs boson. However, it has been suggested that the observed signal could also be produced by multiple nearly mass-degenerate states that couple differently to the standard model particles.

In this work, a method to discriminate between the hypothesis of a single Higgs boson and that of multiple mass-degenerate Higgs bosons was developed. Using the matrix of measured signal strengths in different production and decay modes, parametrizations for the two hypotheses were constructed as a general rank 1 matrix and the most general 5×4 matrix, respectively. The test statistic was defined as a ratio of profile likelihoods for the two hypotheses.

The method was applied to the CMS measurements. The expected test statistic distribution was estimated twice by generating pseudo-experiments according to both the standard model hypothesis and the single Higgs boson hypothesis best fitting the data. The p-value for the single Higgs boson hypothesis was defined from both expected test statistic distributions, and it was $(8.0 \pm 0.3)\%$ and $(11.0 \pm 0.3)\%$, respectively. In addition to this, a p-value was also estimated in an alternative way using a χ^2 distribution, fitted to the pseudo-experiments for the standard model Higgs boson hypothesis. The resulting p-value was 10.8%. Thus the three estimates yield similar p-values for the single Higgs boson hypothesis.

These results suggest that the CMS data is compatible with the single Higgs boson hypothesis, as in the standard model. Furthermore, the result is insensitive to choice of the single Higgs boson hypothesis used to derive it, and it does not depend on the precise shape of the test statistic distribution.

The developed method can be applied also to other arbitrarily-sized matrices, and it takes into account the uncertainties on the measurements, missing elements of data, and possible correlations. This thesis is an extensive description of the method that has also been published in EPJC (David, Heikkilä and Petrucciani), and the method has been used in the final Run 1 Higgs combination and properties article (CMS Collaboration, incl. Heikkilä).

Avainsanat — Nyckelord — Keywords

Higgs physics, Degenerate states, Profile likelihood ratio, CMS

Säilytyspaikka — Förvaringsställe — Where deposited

Kumpula Campus Library (Gustaf Hällströmin katu 2, 00014 Helsingin yliopisto)

Muita tietoja — övriga uppgifter — Additional information

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Laitos — Institution — Department
Matemaattis-luonnontieteellinen tdk	Fysiikan laitos

Tekijä — Författare — Author

Jaana Heikkilä

Työn nimi — Arbetets titel — Title

Degeneroituneiden Higgsin bosonien etsiminen osamäärätestisuureeseen perustuvaa menetelmää käyttäen

Oppiaine — Läroämne — Subject

Fysiikka

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro Gradu -tutkielma	Toukokuu 2015	100

Tiivistelmä — Referat — Abstract

Suuren hadronitörmäyttimen kaksi koetta, ATLAS ja CMS, ovat havainneet uuden hiukkasen, joka on standardimallin Higgsin bosonin kaltainen. On kuitenkin ehdotettu, että havaittu signaali voisi aiheutua useasta, lähes samanmassaisista hiukkasista, jotka kytkeytyvät eri vahvuuksilla standardimallin hiukkasiin.

Työssä kehitettiin menetelmä, jonka avulla voidaan tutkia kuinka hyvin hypoteesi yhdestä Higgsin bosonista sopii mittaustuloksiin verrattuna hypoteesiin samanmassaisista Higgsin bosoneista. Nollahypoteesi ja yleinen hypoteesi parametrisoitiin vastaavasti asteen 1 matriisina sekä yleisimpänä 5×4 matriisina. Testisuure määriteltiin osamäärätestisuureen, eli hypoteesejä vastaavien uskottavuusfunktioiden suhdeluvun, avulla.

Menetelmää sovellettiin CMS-mittauksiin. Testisuureen odotettua jakaumaa arvioitiin näennäiskokeilla, jotka luotiin kahta lähtöasetelmaa - standardimallia sekä suurimman uskottavuuden estimaatteja - käyttäen. *P*-arvo nollahypoteesin mukaisessa tilanteessa määritettiin molemmista testisuureen odotetuista jakaumista ja saadut *p*-arvot olivat vastaavasti $(8.0 \pm 0.3)\%$ sekä $(11.0 \pm 0.3)\%$. *P*-arvo määritettiin myös vaihtoehtoisella tavalla, jossa χ^2-jakauma sovitettiin siihen testisuureen odotettuun jakaumaan, joka arvioitiin standardimallin mukaisilla näennäiskokeilla. Tällöin *p*-arvo oli 10.8%. Tuloksena saadut kolme *p*-arvoa olivat näin ollen samansuuruisia.

Saadut tulokset osoittavat, että CMS-mittaustulokset ovat yhteensopivia nollahypoteesin eli yhden Higgsin bosonin hypoteesin kanssa, aivan kuten standardimallissa. Huomioitakoon, että saatu *p*-arvo ei riipu näennäiskokeiden luomiseen valitusta lähtöasetelmasta tai testisuureen odotetun jakauman tarkasta muodosta.

Työssä kehitettyä menetelmää voidaan soveltaa myös muihin mielivaltaisen kokoisiin matriiseihin, ja se huomioi mittausten virheet, puuttuvat mittaukset sekä mahdolliset korrelaatiot. Tämä työ on laajempi kuvaus menetelmästä, joka on myös julkaistu EPJC:ssä (David, Heikkilä ja Petrucciani) ja jota on käytetty CMS-kokeen lopullisessa Run 1 Higgs-kombinaatiojulkaisussa (CMS-kollaboraatio, ml. Heikkilä).

Avainsanat — Nyckelord — Keywords

Higgsin fysiikka, Degeneroituneet tilat, Osamäärätestisuure, CMS-koe

Säilytyspaikka — Förvaringsställe — Where deposited

Kumpulan kampuskirjasto (Gustaf Hällströmin katu 2, 00014 Helsingin yliopisto)

Muita tietoja — övriga uppgifter — Additional information

Contents

Acknowledgements

This thesis arose from a CERN Summer Student project that was started in June 2013. At that time, Giovanni Petrucciani and André David were my supervisors, and I will always be grateful for the chance they gave me when they chose me as their summer student.

I would like to thank my thesis supervisor Giovanni Petrucciani and Paula Eerola for their continuous support. Giovanni, I am extremely grateful for all the effort you have put into teaching me. You have always encouraged me to do things in my own way, and given me chances to grow and become a better physicist. Thank you for letting me to learn from the best - I appreciate everything you have done for me. Paula, thank you for always having time to answer my questions concerning my thesis or studying in Helsinki. From the very first moments of my arrival in Helsinki, you made me feel welcomed and showed a genuine interest in the things I was doing.

I would also like to thank my colleagues and friends at CERN and in Helsinki who have helped me and given me feedback and advice during my research career and this project. Because of people like you, I am proud to say that I am a part of CMS, Helsinki Institute of Physics and University of Helsinki.

I also wish to thank my friends and family who have supported me from the very beginning of my university studies: you have believed in me even if I have not and pushed me forward when it has been necessary. Especially “Pojat” in Oulu, you had a major role in my path of becoming a physicist and who I am this day. Last but definitely not least, I would like to thank Ilkka. You are my partner in crime, for better and for worse, and I thank you for always being there for me.

1. Introduction

The standard model of particle physics (SM) describes the elementary particles and the interactions between them. The elementary particles do not have an inner structure. The 12 fermions are divided into quarks and leptons, and the lightest ones of them form all the ordinary matter. The gauge bosons carry the strong, electromagnetic and weak forces. In addition to this, the electromagnetic and weak forces can be unified into an electroweak force at high energies. [1–3]

Gluons and photons mediate the strong and electromagnetic interactions, respectively. Since the gluons and photons are massless, they can be generated through the local gauge invariance. However, the W and Z bosons, the mediators of weak interaction, are massive and adding a mass term into the Lagrangian describing the standard model breaks the local gauge invariance. Carrying on without the local gauge invariance leads to a nonrenormalizable theory, the validity of which cannot extend energies much higher than the W and Z boson masses [4].

In the mid-sixties, it was suggested that the masses of W and Z bosons could be generated by breaking the electroweak symmetry spontaneously while introducing additional scalar fields, also known as the Higgs fields [5–10]. In the minimal implementation of this mechanism only one Higgs field is added, giving rise to a new scalar particle, the Higgs boson, in addition to the massive W and Z bosons. The mechanism, sometimes called Brout–Englert–Higgs mechanism, explains also the masses of fermions. Since the Brout–Englert–Higgs mechanism does not predict the mass of the Higgs boson m_{H}, it should be determined experimentally. Furthermore, if the Higgs field existed and the standard model particles obtained their masses by interacting with it, one evidence would be to observe the Higgs boson. Also other options are available. For example, one could examine the behaviour of the electroweak scattering at high energies. In a case where the Higgs field did not exist and the Higgs boson did not mediate the electroweak scattering, the scattering amplitude would grow arbitrarily with the energy, violating unitarity [11].

A high energy particle collider, the Large Hadron Collider (LHC), was built and harnessed to search for the Higgs boson as one of its main purposes. The Higgs boson had already been searched for, with negative results, at the Large Electron-Positron Collider (LEP), setting a lower bound on m_{H} [12]. The accelerators and experiments were designed so that a SM Higgs boson could not escape detection if it existed anywhere in the mass range from the LEP bound up the highest mass allowed in the SM.

In 2012 two experiments at LHC, CMS (Compact Muon Solenoid) and ATLAS (A Toroidal LHC Apparatus), announced the discovery of a new boson [13–15]. Even though the discovery itself was a remarkable milestone in the history of LHC, it was more important to understand if the new boson was the standard model Higgs boson. After all, there was a chance that the new boson was *a Higgs boson* predicted by a beyond standard model (BSM) theory.

Beyond standard model theories have been developed to solve the flaws of the standard model. Indeed, the standard model is imperfect: it is, for example, incapable of explaining gravity or cold dark matter, while the Brout–Englert–Higgs mechanism gives rise to the hierarchy problem. Moreover, the standard model cannot be considered as a Grand Unified Theory (GUT) where the gauge forces, strong and electroweak interactions, are unified at some (high) energy. In general, beyond standard model theories try to solve the problems of SM by introducing new particles, including more Higgs bosons [16]. Depending on the BSM theory, some of the additional Higgs bosons could be degenerate in mass.

It has been suggested that the signal observed by CMS and ATLAS could be produced by two or more nearly mass-degenerate states that couple differently to the standard model particles. If the difference between the masses of the states is small, below the experimental resolution of ~ 1 GeV [17], the two resonant peaks cannot be resolved. This is the case especially if one of the two states has only a small branching fraction in the high-resolution decay modes. However, it has been pointed out in Ref. [18] that by arranging the experimental information about the signal in a matrix and determining the rank of it, the hypothesis of multiple resonances could be tested: if there is a single resonance, the observed signal yields in different production and decay modes are not independent. Furthermore, the observed signal yields are determined by product of a set of cross sections (independent of the decay mode) and a set of branching ratios (independent of the production mode). If there is more than one resonance, the single state factorization no longer holds since the observation contains the sum of all the signals. Thus, the analysis is performed by examining a matrix of signal strengths, where the elements are written as a product of a production cross section and a decay branching ratio, normalized to the standard model expectations.

There are two challenges that prevent one from determining the rank of the signal strength matrix in conventional ways. First, there are missing matrix elements since not all combinations of the Higgs boson production and decay modes can be measured at LHC. Second, each matrix element has an associated uncertainty and the uncertainties vary greatly. Since rank is a discrete quantity, it could be computed only if the uncertainties were neglected and the data were complete. The previous attempt was able to take into account the uncertainties on the matrix elements, but could not deal with the missing measurements [18].

In this work, a method to test the hypothesis of multiple mass-degenerate Higgs bosons has been developed. The method evaluates the statistical compatibility of the measured matrix with a rank 1 matrix. Starting from the signal strength matrix, two hypotheses have been constructed on the connection between rank and the number of resonances: the single Higgs boson hypothesis has been modelled as a rank 1 matrix, whereas the hypothesis of multiple mass-degenerate states has been modelled as the most general 5×4 matrix, that can have an arbitrary value of rank. The single Higgs boson parametrization can be obtained from the most general parametrization of a 5×4 matrix by setting constraints on the parameters.

The test statistic has been defined as a difference of profile likelihood ratios for the two aforementioned hypotheses, which assures proper treating of uncertainties and/or missing measurements. The p-value for the single Higgs boson hypothesis is defined from the distribution of test statistic, determined using pseudo-experiments that were generated assuming two different hypotheses; the standard model Higgs boson hypothesis and the single Higgs boson hypothesis best fitting the data.

This thesis is divided into two parts. The first part concentrates on the Higgs physics in the standard model and in beyond standard model theories. Also the experimental set-up at CERN is described. The second part is devoted to the data analysis: the construction of the method and its application to the latest (at the time of writing) CMS measurements are presented.

Part I

Theory

2. The standard model and the Higgs mechanism

The standard model is a theory explaining what are the fundamental particles and their interactions. The elementary particles can be categorized into fermions and bosons - or in other words to matter particles and force carriers. [1–3]

2.1 Fermions in the standard model

As described above, fermions are also known as the matter particles because they form all known matter. Fermions have spin of 1/2 and they can be divided into quarks and leptons according to their properties. The theory is first defined in terms of massless fermions of definite chirality, left- or right-handed, depending on the projection of the spin along the momentum direction. The left- and right-handed fermions take part to different interactions, as will be explained later on in the following section. Only after electroweak symmetry breaking, as described in Section 2.3, left-handed and right-handed fermions can be associated in pairs to make up massive fermions.

There are two types of quarks; up- and down-type, that have electric charges of 2/3 and $-1/3$ in units of electric charge e, respectively. There are three generations of quarks, each with an up-type quark and a down-type one, and they are (u, d), (c, s) and (t, b). The generations have been categorized in terms of the quark properties and increasing masses. The first generation consists of stable particles, whereas the particles of the second and third generation are unstable. The quarks carry also a color charge, thus they participate in strong interactions, and they experience color confinement meaning they can exist only as bound states of quarks and/or anti-quarks of null total color charge.

Leptons, in contrast to quarks, do not have a color charge. However, the six leptons - electron, muon, tau and their neutrinos - also form three generations known as electronic, muonic and tauonic leptons: (e, ν_e), (μ, ν_μ), (τ, ν_τ). Electron, muon and tau carry the electric charge of $-e$, whereas their neutrinos are neutral. Neutrinos have very small mass, thus they interact weakly and as a result, they are difficult to observe directly. Each doublet has a leptonic flavour number, and the three are separately conserved except for neutrino mass effects that have been observed so far only in neutrino oscillations that were first detected in 1998 [19]. The overall leptonic number, the sum of the three, is conserved in all standard model interactions.

2.2 Interactions and their mediators in the standard model

The standard model includes the interactions that can be described by quantum field theories: electromagnetic, strong and weak forces. Since there is no quantum field theory formalism for the gravitational force, it is not described by the standard model. The standard model interactions are mediated by gauge bosons.

In quantum field theories all particles are represented as excited states of a field and the motion of the field is described by a Lagrangian density $\mathcal{L}$. The standard model is a gauge theory meaning that it must be invariant under local gauge transformations, assuring that the theory is the same in every part of universe.

Each interaction included in the standard model can be constructed starting from the Lagrangian for the free particle that experiences that interaction. The interactions are generated by requiring a local gauge invariance and introducing a covariant derivative that replaces the partial derivative.

The interaction between electrically charged particles and photons is described by quantum electrodynamics (QED). Combining quantum mechanics and special relativity, QED explains how electrically charged particles interact by exchanging photons. Since the interacting particles are fermions, let us consider the Lagrangian density for a free Dirac field

$$\mathcal{L} = \overline{\Psi}(\mathrm{i}\gamma^\mu \partial_\mu - m)\Psi, \tag{2.1}$$

where $\overline{\Psi}$ is the Dirac adjoint that creates one particle, γ^μ the Dirac matrices, m the mass of particle and Ψ the Dirac spinor that annihilates one particle. In a gauge local transformation the Dirac spinor transforms as $\Psi' \to \exp(\mathrm{i}\alpha(x))\Psi$, where $\alpha(x)$ is a real parameter that depends on space and time arbitrarily. If the local gauge transformation is performed, the Lagrangian density becomes

$$\mathcal{L}^{'} = \overline{\Psi}(\mathrm{i}\partial_\mu \gamma^\mu - m)\Psi - \overline{\Psi}\gamma^\mu \Psi \partial_\mu \alpha(x) = \mathcal{L} - \overline{\Psi}\gamma^\mu \Psi \partial_\mu \alpha(x) \neq \mathcal{L},$$

which means that the Lagrangian density is not invariant under local gauge transformations. However, by introducing the covariant derivative D_μ in terms of the partial derivative ∂_μ and a gauge field A_μ

$$D_\mu = \partial_\mu - \mathrm{i}eA_\mu$$

and requiring that the gauge field transforms in the local gauge transformation as follows

$$A_\mu^{'} \to A_\mu + \frac{1}{e}\partial_\mu \alpha(x),$$

the Lagrangian remains invariant under local gauge transformations.

Indeed, requiring local gauge invariance adds an interaction term, $\mathcal{L}_{int} = e\overline{\Psi}\gamma^\mu\Psi A_\mu$, to the Lagrangian density. As expected, the interaction term describes the electromagnetism and conservation of electric charge. It should be noted that requiring the local gauge invariance does not give rise to a mass term for the gauge field A_μ. Furthermore, an additional mass term for the gauge field, i.e. $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$, would break the gauge invariance, thus it is not possible to add the mass term to the Lagrangian density. Since the photon is known to be massless, this is not a problem unlike in cases where the gauge boson is massive.

Also gluons, the force carriers of the strong interaction, are massless. The strong interaction is described by quantum chromodynamics (QCD), which - as the name suggests - describes how particles with a color charge, i.e. quarks, interact by exchanging a gluon. The strong interaction can be generated by requiring the local gauge invariance of the Lagrangian density (2.1), but since it arises from SU(3) symmetry, the covariant derivative is formed according to the symmetry group, where the electric charge e is replaced with the charges g of SU(3)

$$D_\mu = \partial_\mu - \mathrm{i}gT^a A_\mu^a,$$

where T^a are the eight generators of the fundamental representation of SU(3). The operators are related by $[T^a, T^b] = if^{abc}T^c$, where the structure constants f^{abc} are real. Since the operators do not commute, SU(3) is a nonabelian group. As a result, gluons carry a color charge and self-interact. The 3-gluon and 4-gluon interaction terms are

$$\mathcal{L}^{(3)} = -\frac{1}{2}gf_{abc}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)A_b^\mu A_c^\nu$$

and

$$\mathcal{L}^{(4)} = -\frac{1}{4}g^2 f_{abc}f_{ade}A_b^\mu A_c^\nu A_\mu^d A_\nu^e.$$

The strong coupling and its properties change dramatically with respect to energy scale. The dependence between the strong coupling and the energy arises from the renormalization, i.e. from the redefinition of an individual interaction at a finite scale as an infinite sum of contributions including all possible interactions happening at smaller distance scales (and thus higher energy scales). At low energies, the coupling becomes strong and quarks and gluons experience the confinement, where the coupling between two quarks increases as the distance between these quarks increases. It has been shown experimentally that a single quark or gluon cannot be observed in isolation: when confined quarks are separated in high energy collisions, a new quark-antiquark pair arises from the energy of the strong coupling and hadrons - that do not carry a color charge - are created. Even though this phenomenon, called hadronization, is not fully understood theoretically, it is well established experimentally. The experimental signature of hadronization is a jet, a cone of hadrons. At high energies (or extremely low distances), the strong coupling becomes weak, allowing a quark to be treated approximately as free. This phenomenon is known as asymptotic freedom and since the coupling becomes small, it can be theoretically described by a perturbation theory.

The weak force acts on leptons and quarks. The weak interaction follows from a SU(2)×U(1) symmetry similarly to how the strong interaction is generated from the SU(3) symmetry, thus it has three gauge bosons: Z^0_μ and $W^\pm_\mu$. Accordingly, neutral and charged weak interactions exist. Neutral currents are flavour conserving, whereas charged currents change the flavour of the incoming particle. In the weak interaction, the conserved charge is the weak isospin T. Only particles belonging to a non-trivial representation of SU(2), i.e. with the weak isospin $T \neq 0$, interact with charged currents. Out of the SM fermions, only the left-handed ones do, and are all in SU(2) doublets. However, both left-handed and right-handed fermions interact with neutral currents.

The electromagnetic and weak interactions can be unified into an electroweak interaction, which is achieved by considering a symmetry group that takes into account the SU(2) and U(1) representations of the weak and electromagnetic forces, respectively. Thus, the symmetry group becomes $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. The lower index L refers to left-handed fields, whereas the lower index Y refers to weak hypercharge Y_W, defined in terms of electric charge Q and weak isospin T as $Y_\mathrm{W} = 2(Q - T)$. The gauge bosons of this group are W^i_μ $(i = 1, 2, 3)$ and B^0_μ, where W^1_μ and W^2_μ are charged gauge bosons, whereas W^3_μ and B^0_μ are neutral ones. All gauge bosons are massless as expected from the requirement of gauge invariance. Gauge invariance generates the interaction terms, also cubic and quartic self-interaction terms for the gauge bosons.

The charged gauge bosons of weak force, $W^\pm_\mu$, can be easily written in terms of W^1_μ and W^2_μ: $W^\pm_\mu = (W^1_\mu \mp \mathrm{i}W^2_\mu)/\sqrt{2}$. However, identifying W^3_μ and B^0_μ with the Z^0_μ boson and photon A_μ is more complex. Since the photon interacts with both left- and right-handed fermions and the electric charge is conserved in the electromagnetic interaction, the photon cannot be either W^3_μ or B^0_μ that conserve the hypercharge: this would require different electric charges for left- and right-handed fermions. Since W^3_μ and B^0_μ are both neutral and massless, the photon and the Z^0_μ boson can be written as their linear combination.

The electroweak theory, however, exhibits a very different feature from QCD, namely that the symmetry is broken, as the $W^\pm_\mu$ and Z^0_μ gauge bosons and the fermions are massive.

2.3 The Brout–Englert–Higgs mechanism

The masses of electroweak gauge bosons can be generated by spontaneously breaking the SU(2)×U(1) symmetry while introducing at least one new field - the Higgs field - and requiring the local gauge invariance. [5–10]

Spontaneous symmetry breaking (SSB) is a phenomenon where the Lagrangian is symmetric, but the vacuum state of the system is not. In other words, the

Lagrangian has a collection of ground states that share the symmetry, but each individual ground state is asymmetric. Thus, choosing one ground state over the other ground states breaks the symmetry. Since the asymmetry arises out of the selection of the vacuum state instead of an external reason, the symmetry breaking is called "spontaneous".

Before the electroweak spontaneous symmetry breaking (EWSSB), the Higgs field is a SU(2) doublet with complex scalars:

$$\Phi = \begin{pmatrix} \phi^{(+)} \\ \phi^{(0)} \end{pmatrix}, \tag{2.2}$$

where the upper indexes of both components refer to the electric charge Q, thus the hypercharge is $Y_W = 1$ as required from the coupling to the photon. The Lagrangian for the Higgs field is

$$\mathcal{L} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\phi) = (D_\mu\Phi)^\dagger(D^\mu\Phi) - \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2, \tag{2.3}$$

where μ and λ are real constants and D_μ is the covariant derivative in the electroweak theory. The constant λ is chosen to be positive - otherwise the potential $V(\phi)$ becomes arbitrarily negative for large values of the field, i.e. the theory is unstable. Since the potential has only one minimum when $\mu^2 > 0$, the constant μ^2 is chosen to be negative. As a result, there are infinite number of degenerate ground states. The two possible forms of the potential can be seen in Fig. 2.1. Furthermore, the potential resembles a sombrero when μ^2 is chosen to be negative, as it can be seen in Fig. 2.2.

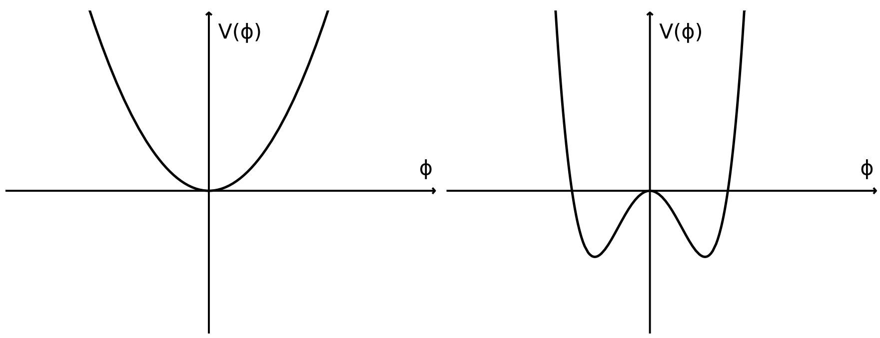

Figure 2.1: The potential $V(\phi)$ for the Higgs field when the parameter μ^2 is chosen to be positive (left) and negative (right).

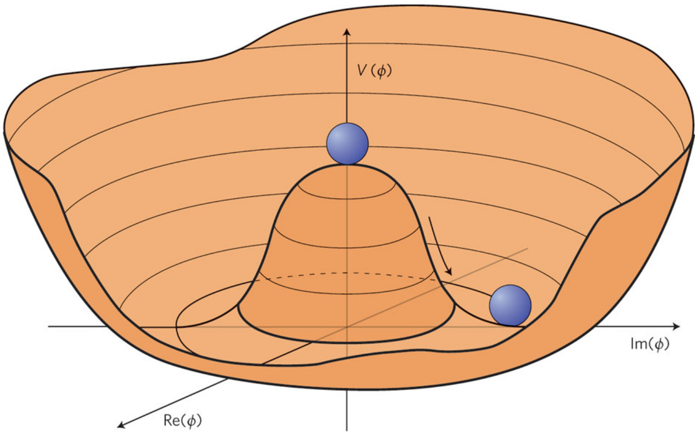

Figure 2.2: When the parameter μ^2 is negative, the potential $V(\phi)$ resembles a sombrero. The number of degenerate ground states is infinite. [20]

When the vacuum state is chosen as

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{2.4}$$

where $v = \sqrt{-\mu^2/\lambda}$ is the vacuum expectation value (VEV), the $\mathrm{SU}(2)_L\times\mathrm{U}(1)_Y$ symmetry breaks spontaneously: the system is not invariant when the original infinitesimal transformations are performed to the vacuum state. However, the system remains symmetric with respect to the combined transformation that takes into account the sum of the weak isospin and the weak hypercharge, i.e. the electric charge Q: the $\mathrm{SU}(2)_L\times\mathrm{U}(1)_Y$ symmetry has broken spontaneously to the electromagnetic group $\mathrm{U}(1)_Q$. As it follows from the Goldstone theorem, for each broken symmetry a massless state will appear.

Expanding around the chosen vacuum state, the doublet can be written in the most general form by adding the Higgs boson to the vacuum state and multiplying the state with an exponential factor that takes into account the Goldstone states:

$$\Phi = \exp\left\{\frac{\mathrm{i}}{2}\sigma^i\theta^i(x)\right\}\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+H(x)\end{pmatrix}, \tag{2.5}$$

where $\theta^i(x)$ represents the Goldstone bosons. However, the exponential term can be set to unity by using the unitary gauge where each Goldstone state is set to zero, i.e. $\theta^i(x) = 0$, allowed by the fact that the Lagrangian is invariant under the local $\mathrm{SU}(2)_L$ transformations, making the Goldstone states unphysical. Thus, the doublet becomes

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.6}$$

As it was discussed previously, any interaction can be generated by introducing the covariant derivative. In the case of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry, the covariant derivative is written as

$$D_\mu = \partial_\mu - \frac{\mathrm{i}g}{2} \vec{\sigma} \cdot \vec{W}_\mu - \frac{\mathrm{i}g^{'}}{2} B_\mu, \tag{2.7}$$

where $\vec{\sigma}$ are the Pauli spin matrices, $\vec{W}_\mu$ and B_μ the gauge bosons of SU(2)×U(1) symmetry, and g and $g^{'}$ are their couplings to the scalar field, respectively. By applying the covariant derivative to the doublet and keeping in mind the decision to use the unitary gauge, the kinetic term of the Lagrangian density (2.3) becomes

$$(D_\mu \Phi)^\dagger (D^\mu \Phi) = \frac{1}{2} \partial_\mu H \partial^\mu H + (v+H)^2 \left\{ \frac{g^2}{4} W^+ W^- + \frac{1}{8} (g^2 + g'^2) \left(\frac{g W_\mu^3 - g' B_\mu}{\sqrt{g^2 + g'^2}} \right)^2 \right\}, \tag{2.8}$$

where one can see four mass eigenstates: three for the weak interaction and one for the electromagnetic interaction. The mass eigenstates are

1. $W_\mu^\pm$: $m_\mathrm{W} = \frac{vg}{2}$
2. $Z_\mu^0 = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2+g'^2}}$: $m_\mathrm{Z} = \frac{v\sqrt{g^2+g'^2}}{2}$
3. $A_\mu = \frac{gW_\mu^3 + g'B_\mu}{\sqrt{g^2+g'^2}}$: $m_\gamma = 0$.

Since the electromagnetic group is not broken, the photon remains massless while the $\mathrm{SU}(2)_L$ gauge bosons become massive. The mass of the Higgs boson becomes $m_\mathrm{H} = \sqrt{-2\mu^2}$. The couplings of the Higgs boson to the W and Z bosons can be read from the kinetic term also.

The number of degrees of freedom should be the same before and after the Brout–Englert–Higgs mechanism. Before the spontaneous symmetry breaking, the number of degrees of freedom (d.o.f) was 12: three massless W_μ^i bosons, each having two possible transverse polarizations ($3 \times 2 = 6$ d.o.f), one massless photon B_μ (2 d.o.f) and one scalar doublet ($2 \times 2 = 4$ d.o.f). After the spontaneous symmetry breaking the number of degreed of freedom is the same as required: three massive gauge bosons $W_\mu^\pm$ and Z_μ^0, each having two transverse polarizations and one longitudinal polarization ($3 \times 3 = 9$ d.o.f), one photon A_μ (2 d.o.f) and one real scalar particle, i.e. the Higgs boson (1 d.o.f).

Also the masses of fermions can be generated from the interaction between the fermions and the Higgs field, using the gauge invariant Yukawa coupling, i.e. the gauge invariant fermion-scalar coupling:

$$\mathcal{L}_Y = -c_1(\bar{u}, \bar{d})_L \Phi \; d_R - c_2(\bar{u}, \bar{d})_L \Phi^c \; u_R - c_3(\bar{\nu_e}, \bar{e})_L \Phi \; e_R + h.c., \tag{2.9}$$

where c^i are constant 3×3 matrices in the generation space and Φ^c is the complex conjugate defined as $\Phi = \mathrm{i}\sigma^2\Phi^*$, with the hypercharge of -1. The neutrino masses have been neglected, thus a right-handed neutrino field has not been introduced and as a result, the neutrino-scalar coupling is not included in Eq. (2.9). Using the same vacuum expectation value as above, the Lagrangian density $\mathcal{L}_Y$ becomes

$$\mathcal{L}_Y = -\frac{1}{\sqrt{2}}(v + H)\left\{c_1\bar{d}d + c_2\bar{u}u + c_3\bar{e}e\right\}, \tag{2.10}$$

where the (arbitrary) fermion masses can be recognised in terms of the vacuum expectation value:

1. Mass of down quark: $m_\mathrm{d} = \frac{c_1 v}{\sqrt{2}}$

1. Mass of up quark: $m_\mathrm{u} = \frac{c_2 v}{\sqrt{2}}$

3. Mass of electron: $m_\mathrm{e} = \frac{c_3 v}{\sqrt{2}}$.

The numerical value of the vacuum expectation value has been determined experimentally, $v = 246$ GeV [21].

3. The production and decay of the standard model Higgs boson at the Large Hadron Collider

Given that the Brout–Englert–Higgs mechanism were the correct explanation for the origin of the masses of the fundamental particles, the theory could be verified for example by observing the Higgs boson predicted by it. At CERN, a high energy particle collider, the Large Hadron Collider (LHC) was built for searching the Higgs boson as one of its main purposes. In 2012, CMS and ATLAS at LHC announced the discovery of a new boson [13–15].

One of the most important aspects of the searches for the Higgs boson is to understand how the standard model Higgs boson is produced and how it decays under the experimental conditions of LHC. This chapter describes the production and decay mechanisms of the Higgs boson from the theoretical point of view - the experimental point of view on the subject, e.g. how the different processes are distinguished from each other, is discussed later on in Subsections 5.2.6 and 8.2.2. Since the phenomenology of the boson depends on its mass, the production cross sections and decay branching ratios discussed in this chapter are given for the measured mass $m_{\mathrm{H}} = 125.02\,^{+0.26}_{-0.27}\,(\mathrm{stat})\,^{+0.14}_{-0.15}\,(\mathrm{syst})$ GeV determined by CMS [17].

Following sections are based on Refs. [21–23] unless otherwise mentioned.

3.1 The production modes of the Higgs boson

The production mechanisms of the Higgs boson depend on the colliding particles and their energies. For proton-proton collisions at the LHC, the standard model Higgs boson is mainly produced in five different ways (starting from the most dominating process at LHC): gluon fusion, vector boson fusion, Higgs-strahlung, bbH and ttH associated productions. Naturally, there are many ways to produce the Higgs boson. However, since the coupling to the Higgs boson is proportional to the mass of the particle, some production mechanisms are more likely than the others.

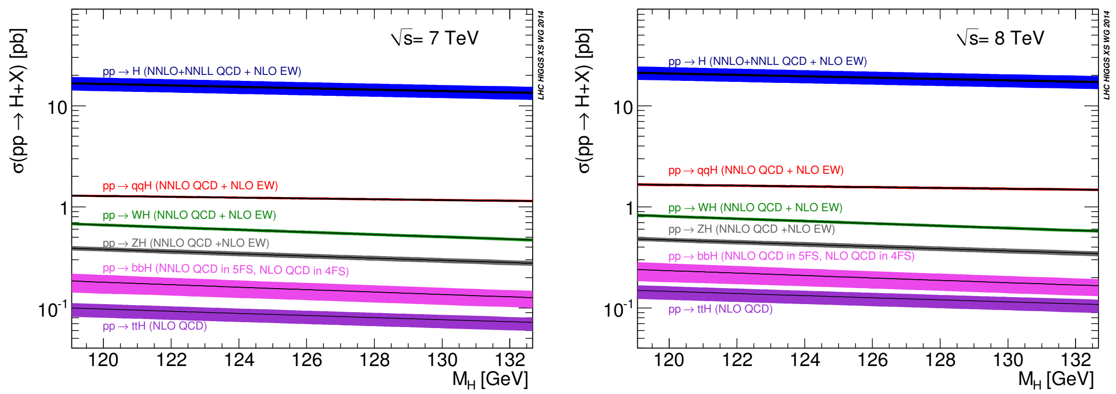

Figure 3.1: Left: The theoretical production cross sections and their uncertainties for the Higgs boson in proton-proton collisions at $\sqrt{s} = 7$ TeV. Right: The theoretical production cross sections and their uncertainties for the Higgs boson in proton-proton collisions at $\sqrt{s} = 8$ TeV. [24]

The cross section σ for a process describes the likelihood of interaction between particles and is proportional to the center-of-mass energy $\sqrt{s}$. The predicted cross sections and their uncertainties at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV can be seen in Fig. 3.1. In addition to higher orders in perturbation theory, the parton density functions (PDF) and the strong coupling α_S form together another source of uncertainty on the cross sections. For a Higgs boson mass of 125 GeV, the cross sections and their uncertainties resulting from missing higher orders at $\sqrt{s} = 7$ TeV and at $\sqrt{s} = 8$ TeV can be seen in Table 3.1. The bbH associated production has been left out from aforementioned table since it is experimentally negligible.

Process	$\sigma_7 \pm$ TH $\pm$ (PDF + α_S) (pb)	$\sigma_8 \pm$ TH $\pm$ (PDF + α_S) (pb)
Gluon fusion	15.1 $\pm$ 1.1 $\pm$ 1.1	19.3 $\pm$ 1.4 $\pm$ 1.4
Vector boson fusion	1.222 $\pm$ 0.004 $\pm$ 0.028	1.578 $\pm$ 0.003 $\pm$ 0.043
WH production	0.579 $\pm$ 0.005 $\pm$ 0.015	0.705 $\pm$ 0.007 $\pm$ 0.016
ZH production	0.335 $\pm$ 0.010 $\pm$ 0.009	0.415 $\pm$ 0.013 $\pm$ 0.010
ttH associated production	0.086 $\pm$ 0.005 $\pm$ 0.007	0.129 $\pm$ 0.009 $\pm$ 0.010

Table 3.1: The predicted numerical values for production cross sections for a Higgs boson mass of 125 GeV at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV, respectively. Here "TH" refers to missing higher orders especially in QCD. [23]

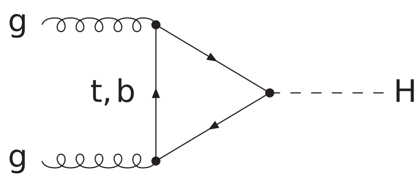

Figure 3.2: The leading order diagram for the gluon fusion production of the Higgs boson. [22]

In gluon fusion (ggH) the Higgs boson is produced via an intermediate loop of virtual quarks. The leading order (LO) Feynman diagram can be seen in Fig. 3.2. Since the Yukawa coupling to the Higgs boson is proportional to the fermion's mass, the dominant contribution to the process is from the top quark, and to a much lesser extend the bottom quark. The gluon fusion is a QCD process, thus the cross section can be calculated using perturbative QCD where higher order corrections play a major role as the strong coupling constant is not small ($\alpha_S \sim 0.1$). The gluon fusion cross section is known up to three orders in QCD (NNLO) with additional resummation of large logarithms at the same order (NNLL), whereas the electroweak corrections are known at the next-to-leading order (NLO), thus with a comparable accuracy since $\alpha_{EW} \sim 0.01 \sim \alpha_S^2$.

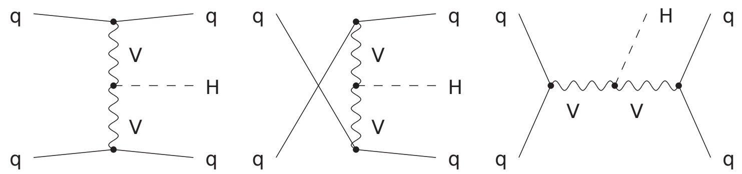

Figure 3.3: The leading order diagrams for the vector boson fusion production of the Higgs boson: t-, u- and s-channels from left to right. [22]

The vector boson fusion (VBF) has the second largest cross section at LHC. In the vector boson fusion process two scattering quarks radiate two vector bosons (W or Z bosons) that merge into a Higgs boson. The outgoing quarks are observed as forward jets, which is a distinctive trace of the vector boson fusion. The LO Feynman diagrams for t-, u- and s-channels can be seen in Fig. 3.3. Experimentally the t- and u- are more interesting than the s-channel: in the former channels the outgoing quarks become a hard jet pair, easy to distinguish from the background. Since the vector boson fusion is not a QCD process, the cross section is quite insensitive to the QCD corrections. The QCD corrections have been determined up to the next-to-next-to-leading order (NNLO) and the electroweak corrections are known at NLO.

The Higgs-strahlung (VH) is a production mechanism where a quark pair annihilate and a vector boson, which emits the Higgs boson, is created. The cross section has been calculated up to NNLO QCD corrections and NLO electroweak corrections. The LO diagrams with W and Z bosons can be seen in Fig. 3.4, where also a higher-order contribution to the ZH associated production process is shown.

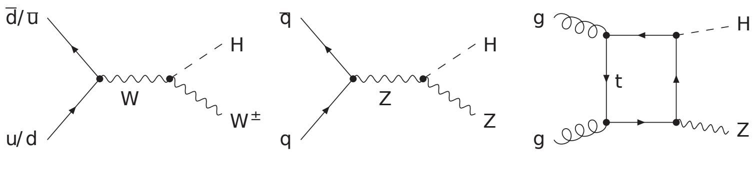

Figure 3.4: The leading order diagrams for the Higgs-strahlung production processes with W and Z bosons (two leftmost diagrams) and a higher-order contribution to the ZH associated production process (rightmost diagram). [22]

Since the ZH associated production process has an "extra" higher-order contribution, the uncertainties on the WH and ZH cross sections are determined separately. At the mass of approximately 125 GeV, the cross sections for the box-diagram at $\sqrt{s} = 7$ TeV and at $\sqrt{s} = 8$ TeV are 22.78 fb and 32.46 fb, respectively. It should be noticed from Table 3.1 that the additional higher-order contribution increases the uncertainty resulting from the QCD corrections as expected. The experimental signature of the Higgs-strahlung is the presence of a vector boson in addition to the Higgs boson.

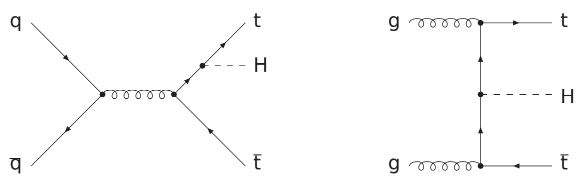

Figure 3.5: The leading order diagrams for the ttH associated production from the incoming quarks (left) and gluons (right). [22]

The process with the lowest cross section is the ttH associated production. Despite its small cross section, the ttH associated production has a striking experimental signature. In the ttH associated procution two incoming gluons or quarks form a top quark pair and the Higgs boson is radiated off them. The LO diagrams for the ttH associated production from the incoming quarks and gluons can be seen in Fig. 3.5. The cross section for the ttH associated production is known at the next-to-leading order. Even though the process has complex final state, important information about the coupling between top quark and the Higgs boson can be obtained when examining it.

3.2 The decay channels of the Higgs boson

The Higgs boson can decay to everything it couples to if the process is not prohibited by the conservation laws. The Higgs boson decays directly to fermions and gauge bosons, but decays into final states with photons or gluons are possible via intermediate loops of quarks or vector bosons. The branching ratio $\mathcal{B}(\mathrm{H} \to \mathrm{final})$ is used to express the probability for the Higgs boson to decay to a given final state.

When considering a fermion pair decay channel, the mass of the fermion plays an important role: the branching ratio is proportional to the square of the Yukawa coupling between the Higgs boson and the fermion, and the Yukawa coupling is equal to the ratio between the fermion mass and the vacuum expectation value, divided by a square root of two. If the Higgs boson were heavy ($m_{\mathrm{H}} > 350$ GeV), the dominating fermion pair decay channel would be $\mathrm{t\bar{t}}$. However, for a lighter Higgs boson, such as the observed one, the dominating decay channels are $\mathrm{b\bar{b}}$, $\tau^+\tau^-$ and $\mathrm{c\bar{c}}$, whereas other fermionic decay modes, e.g. $\mu^+\mu^-$, are rather negligible.

The Higgs boson can decay also to a gauge boson pair consisting of either on-shell (i.e. physical) or off-shell (i.e. virtual) gauge bosons. In the former case the mass of the Higgs boson must be at least twice as large as the mass of the gauge boson: $m_{\mathrm{H}} > 2m_{\mathrm{V}}$, whereas in the latter case the Higgs boson can be lighter. The gauge boson pair decays to four fermions, potentially offering a clear experimental signature. The leading order diagrams for the Higgs boson decay into a fermion and a gauge boson pairs can be seen in Fig. 3.6.

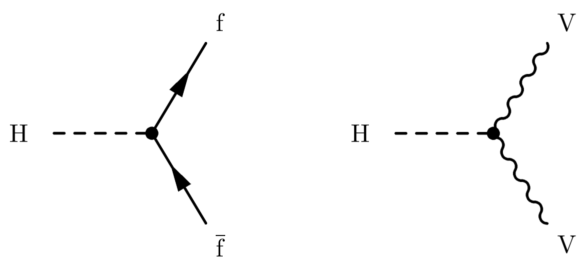

Figure 3.6: The leading order diagrams for the Higgs boson decay into a fermion and a gauge boson pairs.

The Higgs boson decay process into a photon pair, i.e. diphoton, has a small branching ratio, but it is one of the experimentally important channels: the momentum and energy of diphoton can be measured precisely, leading to high mass resolution. The dominating contributions in the fermion and gauge boson loops come from top quark and W boson loops, respectively. The leading order diagrams for the loop-induced H $\rightarrow \gamma\gamma$ decay can be seen in Fig. 3.7.

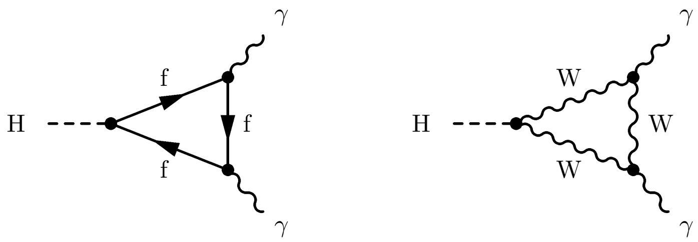

Figure 3.7: The leading order diagrams for the Higgs boson loop-induced decays into diphotons.

Other loop-induced decay modes of the Higgs boson are a gluon pair and Zγ. The decay into a gluon pair is the same interaction as the gluon fusion production mode, but it happens in the reverse order. The decay channel into two gluons has larger branching fraction than the diphoton channel, but is experimentally not interesting, since background gluon pairs are produced at an overwhelming rate by QCD interactions. The Zγ decay is almost irrelevant experimentally: the branching ratio for the decay Z $\rightarrow \ell\bar{\ell}$ is small. Thus, the decay channel of Zγ is difficult to recognise and measure.

The predicted branching ratios for all aforementioned final states and their uncertainties can be seen in Fig. 3.8. The numerical values of the branching ratios for the experimentally most important final states can be seen in Table 3.2.

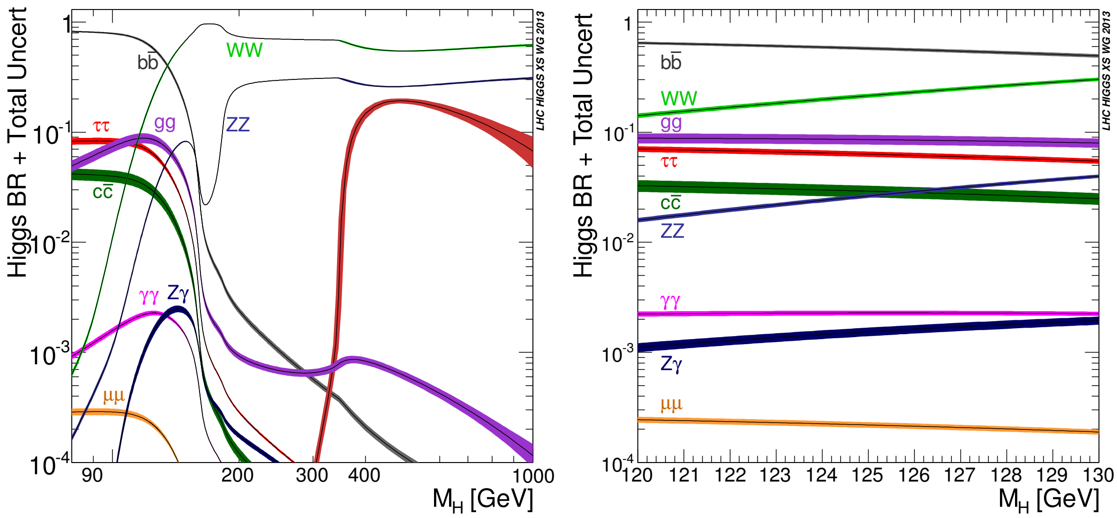

Figure 3.8: Left: Branching ratios and their theoretical uncertainties for the Higgs boson, over the full mass range. Right: Branching ratios and their theoretical uncertainties for the Higgs boson, over the relevant mass range. [24]

Final state	$\mathcal{B} \pm \sigma_{\mathcal{B}}$
$b\bar{b}$	0.577 ± 0.019
WW	0.215 ± 0.009
$\tau^+\tau^-$	0.0632 ± 0.0036
ZZ	0.0264 ± 0.0011
$\gamma\gamma$	0.0228 ± 0.0011

Table 3.2: The predicted numerical values of branching ratios for the experimentally most important final states for a Higgs boson mass of 125 GeV. The branching ratios for the final states with two vector bosons WW or ZZ exclude the decay of the boson pair. [23]

4. Extending the Higgs sector

The standard model Higgs boson is the minimal implementation of the spontaneous electroweak symmetry breaking. However, the imperfections of the standard model encourage to examine more complicated theories that in general introduce more than one Higgs field.

The standard model is not capable describing some experimental observations: dark matter or neutrino oscillations. Furthermore, the gauge forces of the standard model - strong and electroweak interactions - cannot be unified at some (high) energy. In addition to this, the Brout-Engler-Higgs mechanism gives rise to the hierarchy problem. [16]

Many theories have been introduced to solve the problems of the standard model. For example, supersymmetry (SUSY) [25] solves the hierarchy problem with new particles called superpartners or sparticles that have the same quantum numbers but different spin as the SM particles. When considering theories reaching beyond the standard model and including additional Higgs fields, a parameter defined in terms of the W and Z boson masses and the gauge couplings,

$$\rho_{\mathrm{EW}} = \frac{M_{\mathrm{W}}^2}{M_{\mathrm{Z}}^2 \cos^2\theta_{\mathrm{W}}}, \tag{4.1}$$

and its value - experimentally very close to unity - must be taken into account [21].

The parameter ρ_{EW} can also be interpreted as an evidence of the scalar structure of a theory, and written - at tree level - in terms of all scalar multiplets ϕ_i and their weak isospins T_i, weak hypercharges Y_i and vacuum expectation values v_i

$$\rho_{\mathrm{EW}} = \frac{\sum_{i=1}^{n}[T_i(T_i+1) - \frac{1}{4}Y_i^2]\ v_i}{\sum_{i=1}^{n}\frac{1}{2}Y_i^2\ v_i}, \tag{4.2}$$

leading to a conclusion that the simplest extension of the standard model can be achieved with additional SU(2) scalar doublets and singlets that implement $T(T+1) = \frac{3}{4}Y^2$ with hypercharges $Y = \pm 1$ and $Y = 0$, respectively. Also more complex extensions are possible, including an additional SU(2) triplet with hypercharge $Y = 2$, but they can lead to $\rho_{\mathrm{EW}} \neq 1$. [26]

Depending on the extension of the standard model, a few of the resulting physical Higgs bosons could be mass-degenerate. In the simplest extensions the mass-degeneracy can be obtained by imposing constraints on the model's parameters, but with more complex extensions it can rise "naturally". Regardless of the way the phenomenon is caused, the mass-degenerate Higgs bosons could produce an observation of a SM-like signal with the mass of $\sim$ 125 GeV. Inspired by this, the next section discusses a well-motivated model, the two-Higgs-doublet model (2HDM) [27]. After a short introduction to the model, the discussion concentrates on the phenomenology of mass-degenerate states.

4.1 Two-Higgs-Doublet Model (2HDM)

This section follows Ref. [28] loosely. The two-Higgs-doublet model is well-motivated by many aspects. One motivation comes from supersymmetry, where the scalars and their complex conjugates belong to chiral and opposite chiral multiplets, respectively. An additional Higgs doublet is required because multiplets of different chirality cannot couple together, i.e. a single doublet cannot generate masses of the up- and down-type quarks simultaneously. Another motivation is axion models, where the second Higgs doublet is necessary when the CP-violating in the QCD Lagrangian is rotated away using a global U(1) symmetry. The two-Higgs-doublet can be motivated also by a sufficient baryon asymmetry of the Universe, a phenomenon that can be generated due to the flexible scalar mass spectrum of the 2HDM and the additional sources of CP violation.

As the name suggest, the two-Higgs-doublet model contains two SU(2) doublets Φ_1 and Φ_2 with hypercharges $Y_1 = Y_2 = 1$ and thus there are eight degrees of freedom. In the spontaneous symmetry breaking the vacuum expectation values are chosen as $v_1/\sqrt{2}$ and $v_2/\sqrt{2}$, respectively, and the VEVs satisfy $v_1^2 + v_2^2 = v_{\rm SM}^2 \approx (246 \text{ GeV})^2$. After the spontaneous symmetry breaking, five physical Higgs fields remain: two CP-even scalars (h, H), one CP-odd scalar (A) and two charged Higgs bosons ($\mathrm{H}^{\pm}$). The most important parameter of two-Higgs-doublet model is the ratio of the VEVs

$$\tan\beta = \frac{v_2}{v_1}. \tag{4.3}$$

Another parameter, the mixing angle α, is defined for the CP-even scalars h and H. The parameters α and β determine the interactions between the Higgs field and the vector bosons and the fermions.

There are different types of 2HDMs and they all, in general, give rise to tree level flavour-changing neutral currents (FCNC), which is a serious problem since experimental data does not show such phenomenon. However, in some types of 2HDMs, the FCNC can be assumed not to exist due to a symmetry. The models leading to natural flavour conservation are called type I, type II, lepton-specific and flipped 2HDMs. The type II model is the most studied, mainly because the structure of the Minimal Supersymmetric Standard Model (MSSM) is a special case of the type II 2HDM where the mass of the lightest Higgs boson has a upper bound, the scalar self-couplings and the mixing parameter α are not arbitrary. The type I model is the second most examined and it includes the inert doublet models where one of the VEVs is set to zero, leading to a stable doublet that can produce dark matter candidates. The lepton-specific and flipped 2HDMs are less studied than types I and II. However, since they differ from types I and II only in the lepton and neutrino sector, many conclusions on types I and II apply to the lepton-specific and flipped models as well. Thus, the main focus in this section is on the type I and II models.

In type I model all quarks and leptons couple to the second Higgs doublet Φ_2, whereas in type II model the up-type quarks coupe to Φ_2 but the down-type quarks and leptons couple to the first Higgs doublet Φ_1. The coupling constants of the fermions and the vector bosons with respect to the standard model couplings can be seen in Table 4.1. The couplings of the vector bosons are the same in types I and II, whereas the couplings of the fermions are not. Since the couplings do not depend only on the masses, the production cross sections and the decay branching ratios will depend also on the parameters α and β.

	Type I			Type II		
	h	H	A	h	H	A
up-type quarks	$\cos\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\cot\beta$	$\cos\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\cot\beta$
down-type quarks and leptons	$\cos\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$-\cot\beta$	$-\sin\alpha/\cos\beta$	$\cos\beta/\sin\beta$	$\tan\beta$
vector bosons (W or Z)	$\sin(\beta-\alpha)$	$\cos(\alpha-\beta)$	-	$\sin(\beta-\alpha)$	$\cos(\alpha-\beta)$	-

Table 4.1: The couplings of the fermions and vector bosons to the CP-even (h and H) and CP-odd (A) Higgs bosons in the type I and II 2HDMs, normalized to the SM couplings. [28]

4.1.1 Mass-degenerate states in 2HDM

The possibility to have (nearly) mass-degenerate states at 125 GeV have been studied extensively. Immediately after the discovery of new boson, one of the possible explanations for the apparent excess in the $\gamma\gamma$ channel was the presence of degenerate Higgs bosons, justified with the fact that the CP-odd scalar A does not couple to vector bosons at tree level, but can decay to diphoton, thus enhancing that signal in that final state but not in final states with two vector bosons, i.e. WW or ZZ.

The work done in Refs. [29–31] discussed especially cases where $m_{\mathrm{h}} \simeq m_{\mathrm{A}}$, $m_{\mathrm{H}} \simeq m_{\mathrm{A}}$ and $m_{\mathrm{h}} \simeq m_{\mathrm{H}} \simeq m_{\mathrm{A}}$ and investigated some possible predictions for ratios of signal strengths for the standard model, constructed using a product of a cross section and a branching ratio for certain production and decay channels.

It was shown in Refs. [30, 31] that the enhanced $\gamma\gamma$ signal could be produced by mass-degenerate Higgs bosons especially when $\tan\beta \sim 1$ and $m_{\mathrm{h}} \simeq m_{\mathrm{A}}$. Even though the original motivation of the study is less relevant now as results on the full Run 1 data no longer call for a large enhancement in the $\gamma\gamma$ channel, the study remains interesting. An example benchmark point, taken from Ref. [31], for 2HDM with mass-degenerate Higgs bosons contributing to the 125 GeV peak is presented below.

Type I 2HDM with nearly mass-degenerate CP-even and CP-odd Higgs bosons, $m_h = 125$ *GeV and* $m_A = 125.1$ *GeV, while* $\tan\beta = 1.2$ *and* $\sin\alpha = -0.6$.
When the mass of the CP-odd Higgs boson is set to 125.1 GeV, the other Higgs bosons have the masses $m_{\mathrm{H}} = 625$ GeV and $m_{\mathrm{H}^\pm} = 612$ GeV. The signal strengths, defined as ratios between the observed and the expected signal yield, will be

$$\mu_{i,j} = \frac{(\sigma_i \cdot \mathcal{B}_j)_{\mathrm{h}} + (\sigma_i \cdot \mathcal{B}_j)_{\mathrm{A}}}{(\sigma_i \cdot \mathcal{B}_j)_{\mathrm{SM}}}. \tag{4.4}$$

As it will be explained in Chapter 6, the presence of overlapping resonances can be tested by arranging the signal strengths in a matrix and by examining it. Using the numerical values given in Ref. [31], the signal strength matrix is

	$\mathrm{H} \to \gamma\gamma$	$\mathrm{H} \to \mathrm{ZZ}$
ggH	1.31	1.02
VBF	0.83	0.94

(4.5)

When considering only two production modes and two decay modes, the presence of overlapping resonances can be tested from the double ratio of signal strengths [32], in this case

$$\rho = \frac{\mu_{\mathrm{ggH},\gamma\gamma}\ \mu_{\mathrm{VBF,ZZ}}}{\mu_{\mathrm{ggH,ZZ}}\ \mu_{\mathrm{VBF},\gamma\gamma}}. \tag{4.6}$$

It is evident that if the signal is only from one resonance mode, e.g. CP-even Higgs boson h, the double ratio ρ is equal to unity since deviations from the SM in production cross section or in decay branching ratios cancel in ρ:

$$\rho = \frac{(\sigma_{\mathrm{ggH}} \cdot \mathcal{B}_{\gamma\gamma})_{\mathrm{h}}}{(\sigma_{\mathrm{ggH}} \cdot \mathcal{B}_{\gamma\gamma})_{\mathrm{SM}}} \frac{(\sigma_{\mathrm{VBF}} \cdot \mathcal{B}_{\mathrm{ZZ}})_{\mathrm{h}}}{(\sigma_{\mathrm{VBF}} \cdot \mathcal{B}_{\mathrm{ZZ}})_{\mathrm{SM}}} \cdot \frac{(\sigma_{\mathrm{ggH}} \cdot \mathcal{B}_{\mathrm{ZZ}})_{\mathrm{SM}}}{(\sigma_{\mathrm{ggH}} \cdot \mathcal{B}_{\mathrm{ZZ}})_{\mathrm{h}}} \frac{(\sigma_{\mathrm{VBF}} \cdot \mathcal{B}_{\gamma\gamma})_{\mathrm{SM}}}{(\sigma_{\mathrm{VBF}} \cdot \mathcal{B}_{\gamma\gamma})_{\mathrm{h}}} = 1. \tag{4.7}$$

However, by plugging the values of the signal strengths given in (4.5) into Eq. (4.6), we see that for this 2HDM benchmark point $\rho = 1.45$, significantly different from unity. Thus, if this scenario were to be realised in nature and we were able to measure ρ with sufficient precision, we would be able to tell that the signal cannot be explained by a single resonance.

5. The experimental set-up: the Large Hadron Collider and the Compact Muon Solenoid

5.1 The Large Hadron Collider

The Large Hadron Collider (LHC), the world's largest particle accelerator, is located at CERN on the border between France and Switzerland near Geneva. LHC has a circumference of 27 kilometers and has been designed to collide proton beams with $\sqrt{s} = 14$ TeV and an instantaneous luminosity of 10^{34} cm^{-2}s^{-1} [33]. In 2012, both proton beams consisted of 1374 bunches each containing order 10^{11} protons, the bunch spacing was set to 50 ns and the center-of-mass energy was $\sqrt{s} = 8$ TeV [34].

Two beams are collided in four different collision points where seven LHC experiments record the output of collisions. The experiments are known as ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment), LHC-b (LHC-beauty), TOTEM (TOTal, Elastic and diffractive cross-section Measurement), LHC-f (LHC-forward) and MoEDAL (Monopole and Exotics Detector At the LHC).

ATLAS and CMS are general-purpose detectors, designed to investigate many phenomenon of physics such as the Higgs boson and possible extensions to the standard model, whereas the other experiments are concentrating to more specific subjects. ALICE investigates the quark-gluon plasma using heavy-ion collisions and TOTEM the total cross section but also elastic scattering and diffractive processes. LHCf measures hadronic cross sections in the kinematic regime useful for cosmic ray models, LHCb examines the interactions of b-hadrons and MoEDAL is searching for the magnetic monopole and other highly ionizing stable massive particles. The data analysed in this thesis was collected using CMS.

5.2 The Compact Muon Solenoid

In order to target multiple physics phenomena appearing in high energies, a complicated experiment is needed: this sets a long list of requirements including for example gigantic size and strong magnetic fields. CMS is called a huge onion for a reason: it is made of concentric layers of particle detectors and it is the second largest experiment at LHC with a weight of approximately 14000 tonnes. CMS is approximately 22 meters long and has a diameter of 15 meters. [35] The sectional view of CMS can be seen in Fig. 5.1, whereas a cross section of CMS can be seen in Fig. 5.2. Based on Ref. [35], following Subsections 5.2.1, 5.2.2 and 5.2.3 discuss how different kind of particles are identified and their momentum and energy are measured.

5.2.1 The tracker: for measuring the momentum

The CMS tracker measures the momentum of charged particles by measuring the curvature of their trajectory in the magnetic field. The tracker is made of silicon semiconductors: when a charged particle passes the semiconductor, a small electric signal is produced by ionization - neutral particles do not leave a trace. Combining the signals from different silicon strips and pixels, the path can be reconstructed.

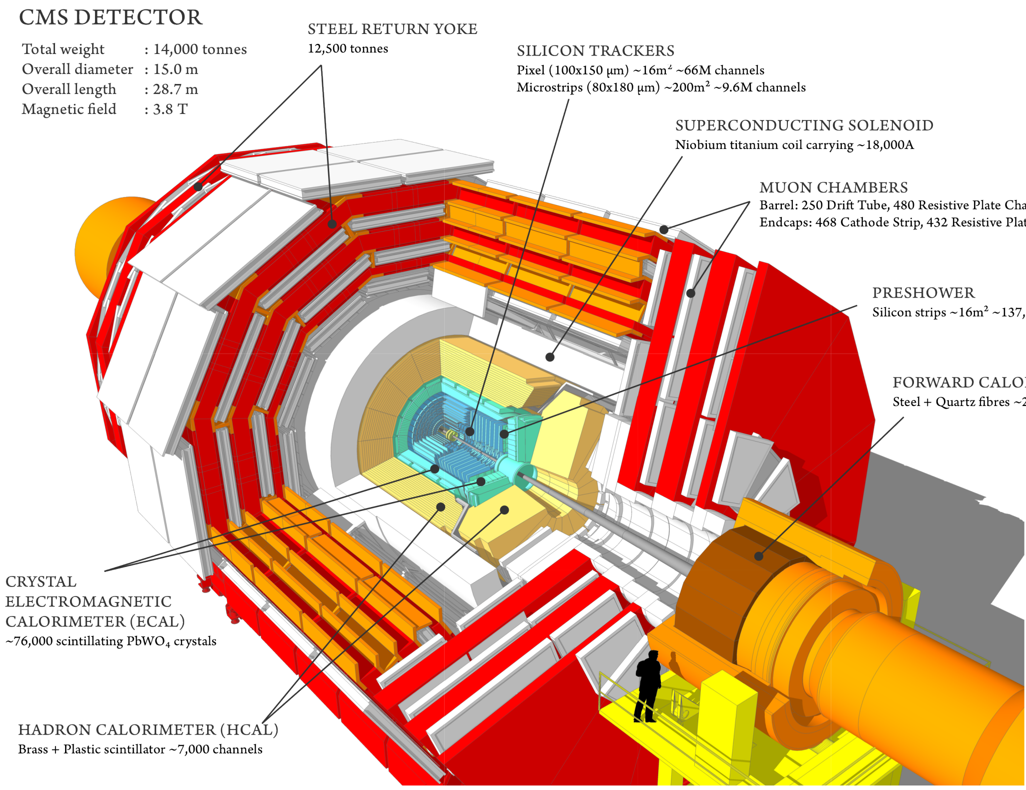

Figure 5.1: A sectional view of CMS detector, showing its multiple subdetectors in the onionlike form and revealing some technical facts of each layer. [36]

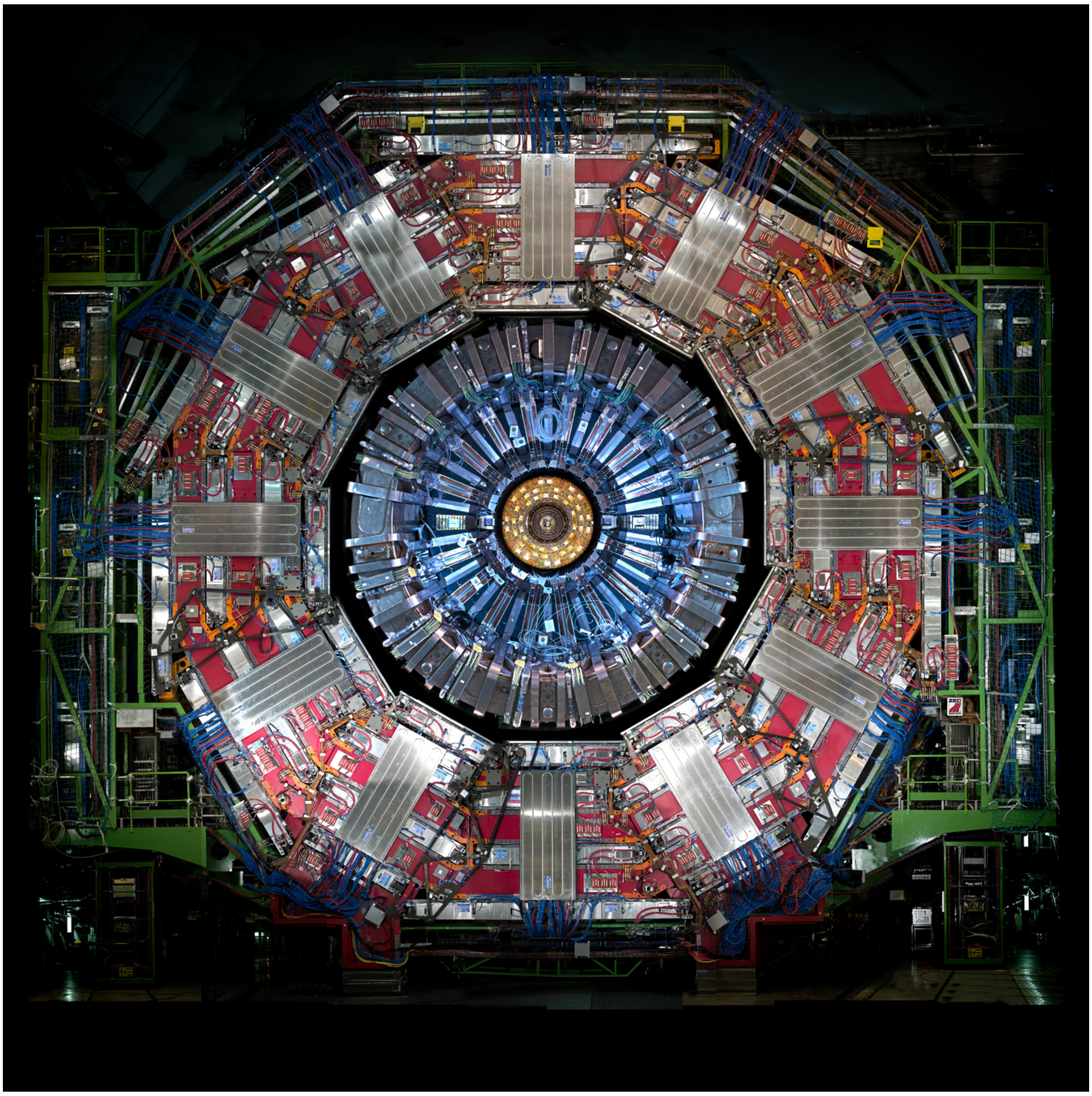

Figure 5.2: A cross sectional view of the constructed CMS detector. [37]

The strong magnetic field of 3.8 T curves the path in a direction depending on the charge. The more the path is curved, the smaller is the momentum of the particle. Using the information about the path and the magnetic field, we can determine not only the momentum but also the charge. In Fig. 5.3 one can see how different particles curve in the magnetic field (which is oriented perpendicularly to the plane of this drawing): negatively charged particles curve to the left, whereas positively charged particles curve to the right.

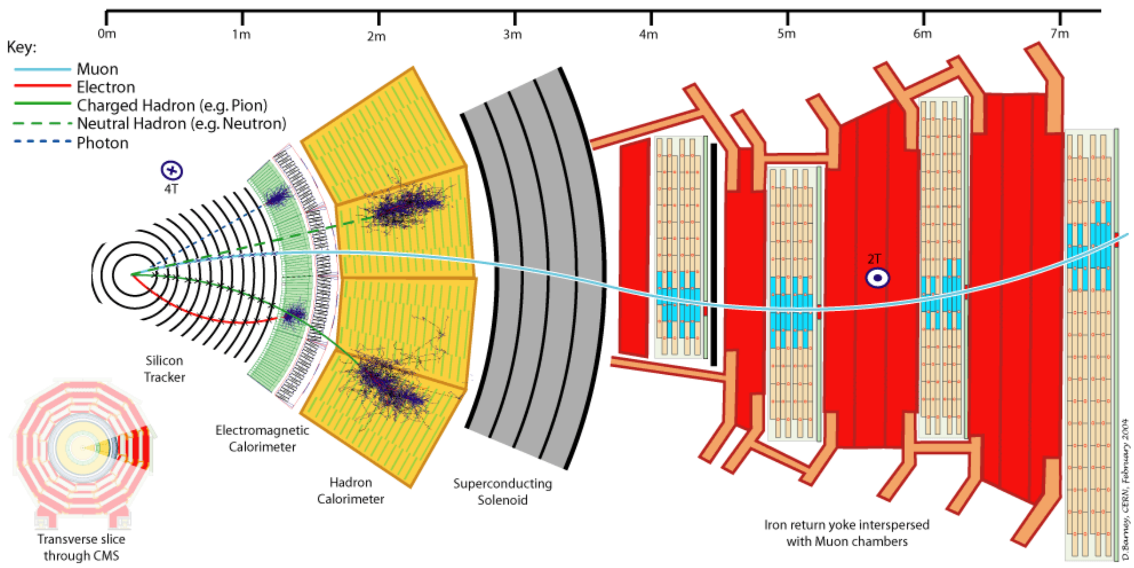

Figure 5.3: Charged particles curve in the magnetic field and leave traces to the CMS tracker, unlike neutral particles. Image by David Barney et al., CERN, 2004.

5.2.2 The calorimeters: for measuring the energy

The calorimeters measure the energy of the particles by making them loose all their energy in the material the calorimeters are composed of. Different particles interact differently with different materials, thus a certain amount of the material is needed to stop specific particles. The calorimeters of CMS have been constructed bearing this fact in mind, making sure that for example electrons, hadrons and muons are identified from each other at the most basic level.

In following subsections, the calorimeters are handled separately starting from the one closest to the interaction point and then moving to the outer layers. First we discuss the Electromagnetic calorimeter and then the Hadronic calorimeter - the Muon detector will be handled in the next section.

The Electromagnetic calorimeter (ECAL)

The Electronic calorimeter is made of lead tungstate crystals. Lead tungstate is transparent material, and it scintillates when an electron or photon passes through. The total amount of light depends on the energy of a particle. The production mechanisms of the scintillation light in the cases of photons and electrons are discussed in following subsections.

As mentioned above, the energy of a particle is defined from the number of produced photons. That is, one wants to collect all the photons and make conclusions about the energy based on this value. However, the interaction between the incoming photon or electron and the calorimeter material is not simple: during the interaction, so called electromagnetic cascade is formed.

The length and width of the electromagnetic cascade depends on the energy of the incoming particle. The secondary electrons, positrons and photons, created in the first interaction, continue interacting with the material producing further secondaries until particles in the cascade have too little energy (or until they reach the end of the detector).

Since the light yield is quite low, there is a need for an amplification of the signal. This is done, for example, using Avalanche Photodiodes (APDs) that are made of silicon and placed to a strong electric field. As the name suggests, the working principle is to create an avalanche of electrons. The process starts when an electron, knocked out of an atom by the scintillation photon, is accelerated in the field. The accelerated electron kicks out other electrons that are also accelerated and the number of electrons increase exponentially. The avalanche can be detected as a very high current in a short time. The signal is then digitized and transported for readout using fibre optic cables.

In the endcaps, a preshower detector is used. The preshower detector is constructed from silicon strips and located before the lead tungstate crystals. The purpose of this subdetector is to increase the spatial precision and separate signals made by two closely-spaced photons from the decays of light neutral mesons (e.g. π^0) from the ones made by one highly energetic photon. The energy deposited at the preshower detector is added to the total energy from the lead tungstate crystals.

Electrons At higher energies, electrons lose energy mainly by bremsstrahlung, where a photon is emitted when the electron is accelerated by an atomic nucleus. The distance over which the electron loses 63% of its energy is called radiation length X_0 and it depends on the absorbing material. If the energy is smaller than few tens of MeV, ionization dominates over bremsstrahlung.

Photons As explained above, neutral particles, such as photons, do not leave traces to the tracker. Thus, the only information about photons is obtained from the ECAL. Photons interact with the calorimeter material in three different ways - the dominating process depends on the photon's energy and the atom number Z (82 for lead). If the energy is below a critical value, the energy loss happens by ionization.

High energy photons lose energy via the pair production. The pair production is a process where a high energy photon, interacting with a nucleus, creates an electron-positron pair. The process stops when the photon energy is a few MeV: the photon must have at least twice the rest energy of an electron for this interaction to happen. In such cases, where the energy of photon is enough to only exceed the limit, the nucleus is needed to absorb the momentum of the original photon -

otherwise the conservation laws would be violated. The distance of path for pair production can be expressed in terms of the radiation length; it is $9/7X_0$.

The Hadronic calorimeter (HCAL)

Unlike electrons and photons, hadrons are normally not fully contained by the ECAL. Since hadrons interact with matter by inelastic and elastic scattering, an interaction length, in analogy with a radiation length, must be defined. The interaction length depends on the material and takes into account the inelastic scattering. In most cases the interaction length is much larger than the radiation length, leading to a situation where the amount of material used for the ECAL is not enough to stop hadrons.

The HCAL is a sampling calorimeter, i.e. it is made of alternating layers of absorber and scintillator. A dense material (brass) has been used to fill the maximum space reserved for the absorber, whereas the scintillator has been made of plastic. Instead of using APDs, Hybrid Photodiodes (HPDs) measure the light signals: the light is converted into electrons by the photoelectric effect, providing a possibility to do a single photon counting. The electrons, accelerated in the electric field, hit the silicon diode target that generates an amplified signal.

The HCAL was designed to measure indirectly also neutrinos from the unbalance of the momentum flow in the transverse plane: knowing that the transverse momentum of the colliding protons is zero and measuring the total momentum of final products (with sufficient transverse momentum), it is possible to reconstruct the transverse momentum of undetected particles, such as neutrinos, from the conservation of the total momentum.

5.2.3 The Muon detector

Muons are charged particles that are hundreds of times heavier than electrons. Since muons do not interact with material as easily as electrons, the muon detector is the outermost layer of the CMS detector - even the powerful magnet mentioned earlier is inside of the muon system (outside of the HCAL). Thus, the direction of the magnetic field is opposite, ensuring clearer measurement of the momentum. In CMS, only the momentum of muons is measured.

The muon detector makes use of three technologies. The trajectories of muons are measured using drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). The drift tubes are in the barrel (around the beam line), the cathode strip chambers in the end caps (ends of the barrels) and the resistive plate chambers are located in the barrel and in the end caps. Combining the hits from the muon stations, the trajectory of a particle can be reconstructed. The measurement is precise and robust, since multiple layers of detector is used.

The muon drift tubes are gas detectors, tubes containing a wire in a gas volume. A charged particle, passing through the volume, ionises the gas. As a result, the released electrons drift in the electric field to the wire. By registering the drift time of the electrons, the original distance of the muon from the wire can be calculated.

The cathode strip chambers are made of arrays of anode wires, crossed with cathode strips and placed in a gas. Two position coordinates are obtained, thanks to the perpendicular strips and wires, when a particle passes through and ionises the gas: the electrons and positive ions move to the anodes and cathodes, respectively.

The resistive plate chambers are also gas detectors and are used to complement the DTs and CSCs. The RPCs are made of two parallel plastic plates (anode and cathode) that are separated by a gas volume. A passing particle ionises the gas, thus electrons are released and an avalanche is created. The signal is picked up by metallic strips after a time delay. A coarse but fast measurement of the muon momentum is achieved by the pattern of hit strips. The RPCs have an excellent time resolutions and a reasonable spatial resolution.

5.2.4 Reconstructing physics objects

In order to observe a certain unstable and short-lived particle, such as the Higgs boson, one must be able to identify the final state objects from the decay of it and the additional objects produced in the rest of the collision event: electrons, photons, muons, hadronically-decaying taus, jets (possibly identified as originating from b-quarks) and missing transverse energy. In CMS, this is achieved using the CMS particle-flow algorithm.

The CMS particle-flow algorithm [38] combines information from all subdetectors of CMS, aiming to reconstruct and identify all stable particles (electrons, muons, photons, charged and neutral hadrons) in the event. Using tracks, calorimeter clusters and muon tracks as building bricks ("elements"), a link algorithm connects the elements and produces "blocks" that can be used to reconstruct each particle in the event. The particle-flow algorithm then reconstructs and identifies a set of particles from each block. As a result, a list of reconstructed particles, and thus a full description of the event, is obtained.

Electrons are reconstructed from a track in the tracker and an energy deposit in the ECAL, whereas reconstructing photons requires only an energy deposit in the ECAL. Muons are reconstructed from - in addition to a track in the tracker - a track or some hits in the muon system. Since the dominant decay mode of the hadronically-decaying taus is a combination of up to two neutral pions and one or three charged hadrons, reconstructing hadronically-decaying taus requires an energy deposit in the ECAL (to distinguish photons from π^0 decay) and tracks in the tracker (to reconstruct the charged hadrons). Leptons produced from the decays of the Higgs boson or W and Z bosons can be separated from backgrounds, such as charged hadrons produced in jets or leptons produced from the decay of hadrons, by requiring them to be isolated, i.e. there are no other energetic particles, tracks or calorimetric deposits around them.

As described earlier in Section 2.2, quarks and gluons hadronize which can be observed as jets. Thus, quarks and gluons are reconstructed by clustering reconstructed particles such as charged and neutral hadrons and photons from hadron decays. When identifying jets originating from b-quarks, i.e. b-tagged jets, the main key is to look for secondary vertices, or tracks not compatible with the primary vertex, that are due to long lifetime of hadrons containing b-quarks. The missing energy is reconstructed from the unbalance of the momentum flow in the transverse plane, as explained in section 5.2.2.

In general, electrons, photons and muons can be reconstructed with very good purity and energy resolution, while jets have a comparatively poor energy resolution.

5.2.5 The trigger and the Data acquisition systems

When searching for unknown particles or physics beyond the standard model, interesting events must be separated from background events. Depending on the bunch spacing, the collision rate is 20 MHz or 40 MHz. However, every event cannot be recorded due to the disk space limitations. For example, in Run 1 data could be stored on tape at $\mathcal{O}(1000)$ Hz, implying a need for an extreme rate reduction of at least $\mathcal{O}(10^4 - 10^5)$ [39].

The rate reduction is performed in multiple steps, using a two-level trigger system that consists of the Level-1 trigger (L1) and the High-level trigger (HLT). The L1 trigger is hardware implemented and makes the decision based on the information from the calorimeters and the muon detector, thus relying on low level analysis. The L1 trigger is designed to reduce the data rate to 100 kHz at the full designed luminosity. [35]

If the event is selected for further analysis, the data is transmitted to the Data acquisition system (DAQ) that includes the HLT. The HLT does not consist of separate trigger levels, but can be still considered as a multi-level system with "virtual" levels. Event information is reconstructed incrementally, and selection criteria are applied in the process, leading to a rejection of large number of events at early stages. The HLT reduces the data rate to $\mathcal{O}(1000)$ Hz. If the event is accepted by the HLT, the information is recorded for the offline analysis. A more precise description of the trigger and the data acquisition systems can be found in Refs. [35, 40].

5.2.6 Searches for the Higgs boson at CMS

The Higgs boson is reconstructed from its decay products, i.e. the Higgs boson is searched by examining experimental signatures characteristic for different decay channels.

Since the decay channels $\mathrm{H} \to \gamma\gamma$, $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$, $\mathrm{H} \to \mu^+\mu^-$ and $\mathrm{H} \to \mathrm{b\bar{b}}$ have visible decay products, the Higgs boson's invariant mass can be reconstructed and used to identify the event as containing a Higgs boson. However, the mass resolution of $\sim 1\%$ can be achieved only with the first three modes [17].

Some decay channels have neutrinos among the visible decay products. This is inevitably the case for the channel $\mathrm{H} \to \tau^+\tau^-$, especially if one selects events with at least one leptonic tau decay ($\tau \to \mathrm{e}\bar{\nu}_\mathrm{e}\nu_\tau$) that are experimentally more accessible due to the light lepton. Nevertheless, the invariant mass of the Higgs boson can be reconstructed using also the missing energy due to available techniques, resulting the mass resolution of $\sim 15\%$ [17]. Also the decay channel $\mathrm{H} \to \mathrm{WW} \to 2\ell 2\nu$ has neutrinos in the final state, but information about the Higgs boson's invariant mass can be reconstructed from the momenta of the leptons and the missing energy.

The events are also categorized by experimental traces, called tags, that target a certain production mechanism. Events produced by the gluon fusion are required to include either 0 or 1 jets, whereas selecting the vector boson fusion requires at least two jets with a large rapidity gap. Depending on the analysis, the Higgs–strahlung can be tagged from the vector boson decays into leptons, neutrinos and jets. The ttH associated production has the same experimental traces as the VH production, but with b-tagged jets that are produced when bottom quarks hadronize.

In the end, the signal is extracted from the analysis of some kinematic distribution of the events. In final states with a precise mass reconstruction ($\gamma\gamma$, 4ℓ, $\mu^+\mu^-$), the invariant mass is used, possibly combined with other kinematic information. Other channels use approximately reconstructed invariant masses ($\tau^-\tau^+$) or more complex multivariate methods ($\bar{\mathrm{b}}\mathrm{b}$) - the latter are also used when the decay products of the Higgs boson are not often clearly identifiable, as it is the case e.g. for ttH associated production with decay channels $\mathrm{H} \rightarrow \mathrm{b}\bar{\mathrm{b}}$ or $\mathrm{H} \rightarrow \mathrm{WW}$ [17].

Part II

Data analysis

6. Determining the number of overlapping resonances from the matrix of measured signal strengths

As discussed in Chapter 4, more than one mass-degenerate state could produce an observation of a SM-like signal with the mass of $\sim$ 125 GeV. However, the two resonant peaks cannot be resolved if the mass difference of the states is below the experimental resolution of $\sim$ 1 GeV [17], and especially if one of the two states has only a small branching fraction in the high-resolution decay modes.

In Ref. [18] the problem has been approached by arranging the information about the signal - the products of cross section and branching ratio for each mode - in a matrix and by calculating the rank of it, aiming to determine a lower bound to the number of bosons in the presence of multiple Higgs bosons contributing to the signal yield in each measurement.

6.1 An example of a 2×2 matrix

Let us use a 2×2 matrix as an example. The rows of this matrix represent two production modes, gluon fusion and vector boson fusion, respectively. In reality one cannot directly detect the individual production modes, but experimental signatures (often referred to as "tags") of production modes that can be used to categorize each event containing a Higgs boson candidate. Individual experimental results on Higgs boson production and decay modes can thus also be arranged in a matrix, but with rows corresponding to production tags and not production modes. We will however overlook this complication now, deferring its discussion to Chapter 8.

The columns represent the two chosen decay modes in this example: two photons and two W bosons. As mentioned in Subsection 5.2.6, the high mass resolution can be achieved with the decay process into a photon pair, but not with the decay channel H $\to$ WW. When at least one decay channel yields a low mass resolution, the multiple Higgs bosons can more easily appear as degenerate.

Every element of the matrix is the ratio between the observed signal yield $S_{\rm obs}$ and the expected signal yield $S_{\rm exp}$. The ratio can be written in terms of the cross section and the branching ratio as

$$\left[\frac{S_{\rm obs}}{S_{\rm exp}}\right]_{i,j} = \left[\frac{N_{\rm obs} - B_{\rm exp}}{S_{\rm exp}}\right]_{i,j} = \frac{(\sigma_i \mathcal{B}_j)_{\rm data} \mathcal{L} A \epsilon}{(\sigma_i \mathcal{B}_j)_{\rm SM} \mathcal{L} A \epsilon} = \frac{(\sigma_i \cdot \mathcal{B}_j)_{\rm data}}{(\sigma_i \cdot \mathcal{B}_j)_{\rm SM}}, \tag{6.1}$$

where $N_{\rm obs}$ stands for the observed number of events, $B_{\rm exp}$ for the expected background, $\mathcal{L}$ for the integrated luminosity, A for the signal acceptance, and ϵ for the signal efficiency. The term $(\sigma_i \cdot \mathcal{B}_j)_{\rm data}$ stands for the measured cross section times the branching ratio for Higgs bosons and $(\sigma_i \cdot \mathcal{B}_j)_{\rm SM}$ for the corresponding expectation from the standard model, whereas the indices i and j identify the production tag and decay mode, respectively. Clearly, if the observation were to match exactly the expectation for the SM Higgs boson, the ratio would be equal to unity.

The 2×2 matrix can now be filled by using Eq. (6.1). Using an abbreviation

$$\mu_{i,j} = \frac{(\sigma_i \cdot \mathcal{B}_j)_{\rm data}}{(\sigma_i \cdot \mathcal{B}_j)_{\rm SM}}, \tag{6.2}$$

the matrix becomes

	H $\to \gamma\gamma$	H $\to$ WW
ggH	$\mu_{\rm ggH,\gamma\gamma}$	$\mu_{\rm ggH,WW}$
VBF	$\mu_{\rm VBF,\gamma\gamma}$	$\mu_{\rm VBF,WW}$

(6.3)

If there is only one Higgs boson, neglecting the uncertainties in the $\sigma \cdot \mathcal{B}$ measurements the determinant of the 2×2 matrix (6.3) is

$$\frac{(\sigma_{\rm gg} \cdot \mathcal{B}_{\gamma\gamma})_{\rm data}}{(\sigma_{\rm gg} \cdot \mathcal{B}_{\gamma\gamma})_{\rm SM}} \frac{(\sigma_{\rm VBF} \cdot \mathcal{B}_{\rm WW})_{\rm data}}{(\sigma_{\rm VBF} \cdot \mathcal{B}_{\rm WW})_{\rm SM}} - \frac{(\sigma_{\rm gg} \cdot \mathcal{B}_{\rm WW})_{\rm data}}{(\sigma_{\rm gg} \cdot \mathcal{B}_{\rm WW})_{\rm SM}} \frac{(\sigma_{\rm VBF} \cdot \mathcal{B}_{\gamma\gamma})_{\rm data}}{(\sigma_{\rm VBF} \cdot \mathcal{B}_{\gamma\gamma})_{\rm SM}} = 0, \tag{6.4}$$

and consequently the matrix has rank 1. However, if there are two Higgs bosons, the matrix elements are sums of two terms, i.e.

$$\begin{aligned}
\mu_{\rm ggH,\gamma\gamma} &= \mu^{(1)}_{\rm ggH,\gamma\gamma} + \mu^{(2)}_{\rm ggH,\gamma\gamma}, \\
\mu_{\rm ggH,WW} &= \mu^{(1)}_{\rm ggH,WW} + \mu^{(2)}_{\rm ggH,WW}, \\
\mu_{\rm VBF,\gamma\gamma} &= \mu^{(1)}_{\rm VBF,\gamma\gamma} + \mu^{(2)}_{\rm VBF,\gamma\gamma}, \\
\mu_{\rm VBF,WW} &= \mu^{(1)}_{\rm VBF,WW} + \mu^{(2)}_{\rm VBF,WW},
\end{aligned} \tag{6.5}$$

where the upper indexes 1 and 2 stand for the first and second Higgs boson, respectively. When replacing all four matrix elements with the corresponding sums from Eq. (6.5), there will be no cancellations and the determinant is not equal to zero in general, thus the rank is two.

6.1.1 The double ratio

In addition to calculating the determinant of a matrix, the uncertainty on the determinant must be determined to take into account the uncertainties in the signal strength measurements. However, it is not practical to evaluate the uncertainty with respect to zero. Thus, when calculating the rank of the matrix, the double ratio ρ has been examined instead of the determinant. The double ratio ρ has been defined [32] as

$$\rho = \frac{\mu_{\rm ggH,\gamma\gamma}\ \mu_{\rm VBF,WW}}{\mu_{\rm ggH,WW}\ \mu_{\rm VBF,\gamma\gamma}}. \tag{6.6}$$

From this definition, it follows that the determinant of the 2×2 matrix is zero if and only if ρ is equal to unity. Furthermore, if there is one resonance, the expected value of ρ is 1. If there are two resonances, the double ratio becomes

$$\rho = \frac{(\mu^{(1)}_{\rm ggH,\gamma\gamma} + \mu^{(2)}_{\rm ggH,\gamma\gamma})\ (\mu^{(1)}_{\rm VBF,WW} + \mu^{(2)}_{\rm VBF,WW})}{(\mu^{(1)}_{\rm ggH,WW} + \mu^{(2)}_{\rm ggH,WW})\ (\mu^{(1)}_{\rm VBF,\gamma\gamma} + \mu^{(2)}_{\rm VBF,\gamma\gamma})}, \tag{6.7}$$

and the expected value of the double ratio ρ is arbitrary, thus the rank in general is 2. For a numerical example, see Subsection 4.1.1.

The double ratio ρ is useful when considering only 2×2 matrices, but in reality there are four production modes and five analysed decay modes. Therefore a 5×4 matrix has to be defined in order to describe the whole situation.

6.2 The matrix of the production tags and decay modes

A matrix of experimental results from CMS for different production tags and decay modes [17] is given below. The rows of this matrix represent four production modes: the gluon fusion, the vector boson fusion, the Higgs–strahlung and the ttH associated production. The columns represent five decay modes[1]: $\gamma\gamma$, WW, ZZ, $\tau\tau$ and bb. Missing elements correspond to combinations of production tag and decay mode that have not been measured individually, or whose preliminary results are not included in Ref. [17] nor in this analysis.

	$H \to \gamma\gamma$	$H \to WW$	$H \to ZZ$	$H \to \tau\tau$	$H \to bb$
ggH	1.01 ± 0.28	0.77 ± 0.22	0.88 ± 0.30	0.84 ± 0.4	–
VBF	1.51 ± 0.51	0.62 ± 0.54	1.55 ± 0.81	0.95 ± 0.41	–
VH	0.57 ± 0.87	0.8 ± 1.01	–	0.87 ± 0.94	1.01 ± 0.51
ttH	2.67 ± 2.07	3.94 ± 1.57	–	-1.33 ± 4.84	0.65 ± 1.83

(6.8)

There are also other decay modes beyond the aforementioned five that are either not searched for directly, or the direct searches are not included in this analysis. They can nonetheless give small contributions to the expected signal yields in the searches used here. For the purpose of this analysis, expected yields from those extra decays are assumed to have the same $(\mathcal{B}/\mathcal{B}_{\text{SM}})_i$ ratio as the most closely related decay among the five: gg, cc, ss, uu, dd are included as bb, $\mu\mu$ as $\tau\tau$, and $Z\gamma$ as $\gamma\gamma$.

[1] Implying charge conjugation throughout; bb stands for $b\bar{b}$, $\tau\tau$ stands for $\tau^+\tau^-$, etc.

The decay to a gluon pair can be counted as a decay to bottom quark pair because the interaction between the Higgs boson and the gluon pair has to be mediated by quarks. Assimilating muons to taus is justified with the fact that both particles, muon and tau, are leptons and interact with the Higgs boson in the same way. The decay to a Z boson and a photon can be considered as a decay to diphoton because the interaction between the Higgs boson and either pair is mediated by loops of quarks and of W bosons.

Since the rank is a discrete quantity, it can be computed only if the elements of the matrix are known exactly. If there are missing measurements, uncertainties on the matrix elements or correlations, one cannot determine the rank, but if the matrix of the expected $\mu_{i,j}$ has rank N, the matrix of the observed $\mu_{i,j}$ is expected to be statistically compatible with a matrix of rank N. Thus, the aim of this analysis will be to evaluate the statistical compatibility of the observed data with a matrix of rank 1, as expected for a single Higgs boson.

7. Statistical treatment

In order to evaluate the statistical compatibility of the observed data with a matrix of rank 1, it is necessary to introduce likelihood functions to model the data.

Likelihood functions can contain two types of parameters: parameters of interest and nuisance parameters. The parameters of interest are the parameters which are examined and/or determined. The nuisance parameters are not interesting but must be taken into account in the analysis - they are profiled, which means that their values are obtained by fitting the data [41].

The likelihood is the probability to observe the data assuming that the given model is true. The likelihood function $L = L(\alpha, \beta)$ is the likelihood expressed as a function of the model's parameters α and β; the joint probability density function for n independent measurements x_i is

$$L(\alpha, \beta) = \prod_{\mathrm{i}}^{n} f_i(x_i; \alpha, \beta), \tag{7.1}$$

where α is the parameter of interest and β is the nuisance parameter [42].

Using the likelihood function $L(\alpha, \beta)$, the profile likelihood ratio has been defined as

$$\kappa(\alpha) = \frac{L(\alpha, \hat{\hat{\beta}}(\alpha))}{L(\hat{\alpha}, \hat{\beta})}, \tag{7.2}$$

where $\hat{\hat{\beta}}(\alpha)$ is the best fit value that maximises the likelihood function for a given α, whereas $\hat{\alpha}$ and $\hat{\beta}$ are the maximum likelihood estimators which maximise the likelihood function in general. The test statistic $q(\alpha)$ can be determined for the given value of the parameter α from the definition of the profile likelihood ratio [21, 41–45]:

$$q(\alpha) = -2\ln\kappa(\alpha) = -2\ln\frac{L(\alpha, \hat{\hat{\beta}}(\alpha))}{L(\hat{\alpha}, \hat{\beta})}. \quad (7.3)$$

Once the observed value of the test statistic q has been obtained, a p-value can be determined. The p-value is the probability to measure a test statistic value at least as extreme as the observed one if the null hypothesis is true. In this analysis the null hypothesis is the hypothesis of a single Higgs boson. Usually the null hypothesis is rejected if the p-value is smaller than 5%. However, to reject the well established hypothesis, for example the background-only hypothesis when searching for a new signal, one normally requires a significance of 3 or 5 standard deviations, which corresponds to a p-value smaller than 1.3×10^{-3} or 2.9×10^{-7}.

In this analysis, where results of the CMS searches are used as input, the likelihood is defined as a function of the expected signal yields (referred to a SM Higgs boson). In order to analyse them, it is necessary to provide a model for the matrix of the Higgs boson cross sections and branching ratios. In general, a model describes the production channels, the decay modes and the used parameter(s). The model defines the predicted event yields for the signal strength in each production and decay mode as a function of parameters.

When considering the 2×2 matrices, the model can be constructed using the definition of the double ratio ρ (6.6). Since the double ratio is the only parameter of interest, the test statistic q can be written as above in Eq. (7.3). From Wilks' theorem it follows that the distribution of the test statistic $q(\rho)$ is approximately a χ^2 distribution with one degree of freedom [46]. Thus, the test against the hypothesis of the rank 1, i.e. at least one resonance, becomes

$$q(\rho = 1) = \chi^2(1 \text{ dof}). \quad (7.4)$$

The uncertainties on the observed value of ρ, corresponding to the 68% confidence level interval for it, are derived from Eq. (7.4) imposing

$$\Delta q(\rho) = 1. \tag{7.5}$$

The test statistic $q(\rho)$ can be considered as a likelihood ratio,

$$q(\rho) = -2 \ln \frac{L(\rho)}{L(\hat{\rho})}, \tag{7.6}$$

where in the numerator there is the likelihood function of the fixed double ratio and in the denominator there is the maximised likelihood function, i.e. the likelihood function of the maximum likelihood estimator $\hat{\rho}$. In other words, when testing for compatibility with rank 1, the numerator is the likelihood to observe data assuming a rank 1 matrix ($\rho = 1$) and the denominator is the likelihood to observe data assuming the most general 2×2 matrix (arbitrary value of ρ).

The likelihood ratio defined in Eq. (7.6) can be generalized to 5×4 matrices such that in the denominator there is the model of the most general 5×4 matrix instead of the most general 2×2 matrix. In the numerator there is the model of the general rank 1 matrix, decomposed as a tensor product of two vectors (μ_i) and $(1, \lambda_{\mathrm{VBF}}, \lambda_{\mathrm{VH}}, \lambda_{\mathrm{ttH}})$, where

$$\mu_i = \frac{[\sigma_{\mathrm{ggH}} \cdot \mathcal{B}(\mathrm{H} \to i)]_{\mathrm{data}}}{[\sigma_{\mathrm{ggH}} \cdot \mathcal{B}(\mathrm{H} \to i)]_{\mathrm{SM}}} \tag{7.7}$$

is the signal strength in the gluon fusion and

$$\lambda_j = \frac{[\sigma_j / \sigma_{\mathrm{ggH}}]_{\mathrm{data}}}{[\sigma_j / \sigma_{\mathrm{ggH}}]_{\mathrm{SM}}} \tag{7.8}$$

is the relative scaling factor between the signal strengths in the gluon fusion and the selected production mode j. This choice of parametrization will be discussed further in Chapter 10. A parametrization of the most general 5×4 matrix can be obtained conveniently from the previous parametrization by making the λ_j also depend on the decay mode i, i.e. defining

$$\lambda_j^i = \frac{[(\sigma \cdot \mathcal{B})(i,j) / (\sigma \cdot \mathcal{B})(i, \mathrm{ggH})]_{\mathrm{data}}}{[(\sigma \cdot \mathcal{B})(i,j) / (\sigma \cdot \mathcal{B})(i, \mathrm{ggH})]_{\mathrm{SM}}}. \tag{7.9}$$

In this general case, the test statistic is a profiled likelihood ratio with N parameters constrained at the numerator. Under the same conditions as for the test statistic (7.6) with $N = 1$, by Wilks' theorem the test statistic should be distributed as a χ^2 with N degrees of freedom. However, in this case we do not expect Wilks' conditions to be satisfied to a sufficient degree, as it will be explained in Section 10, and thus we rely on a Monte Carlo simulation to generate the expected distribution of the test statistic. The p-value is then obtained from the test statistic distribution by counting how many of the generated test statistic values are above the observed value $q^{\rm obs}$.

8. Incremental steps of implementing the algorithm

In order to test if a matrix has the rank 1, an algorithm has been implemented in incremental steps. The algorithm constructs the model, which is needed to evaluate the statistical compatibility of the observed data with a matrix of the rank 1. The input data to analyse, referred to as “data cards”, describe the observed event yields and kinematic distributions, and the expectations for a SM Higgs boson and all the background processes. The fit is done starting from these more fundamental quantities, not from the fitted signal strengths in each channel, since this guarantees preserving the full shape of the likelihood and the full set of correlations.

To explain in a clearer way the statistical treatment of the data, the analysis will first be done using simplified inputs instead of the full inputs from the CMS searches. When constructing the simplified data cards, we replace the contents of real analysis, i.e. its signal and background yields and kinematic distributions, with a very simple analysis - a counting experiment. The counting experiment may have the observed event yield, the expectation for a SM Higgs boson, and the expected background unrelated to that of the real analysis, but it reproduces closely the same signal strength value and uncertainty of the real analysis.

In the next two subsections the incremental steps and the simplified data cards are described: starting from one production mode and one decay channel (one matrix element) and going to two production modes and one decay channel (2×1 matrix). Also, simplified data cards will be used to illustrate some important features of the searches.

8.1 One production mode

The simplest situation with the initial state of two gluons and the final state of two photons can be modelled as a counting experiment. The likelihood function is very simple and consists of a single Poisson probability

$$L = P(N|B + \mu S) = \frac{(B + \mu S)^N \exp(-B - \mu S)}{N!}, \tag{8.1}$$

where S stands for the expected signal for a SM Higgs boson, B for the expected background and N for the observed number of events.

The only parameter of interest is the signal strength μ, which is, as described in Chapter 6, the ratio between the observed and expected signal yield for the Higgs boson. The best fit value for μ has been calculated by maximising L (8.1) with respect to μ

$$\hat{\mu} = \frac{N - B}{S}, \tag{8.2}$$

and its uncertainty has been calculated from the following approximate formula

$$\Delta\hat{\mu} = \frac{\sqrt{N}}{S}. \tag{8.3}$$

The number of observed events N, the expected total signal S and the background B must be chosen in order to reproduce the observed result $\mu = 1.01\ ^{+0.29}_{-0.26}$ displayed in Eq. (6.8) and taken from Ref. [17]. Therefore the values have been set as $N = 44.3$, $S = 24.1$ and $B = 20$. Using the chosen values, the best fit value for μ, that is $\hat{\mu}$, and its uncertainty $\Delta\hat{\mu}$ become

$$\hat{\mu} = 1.01 \pm 0.27. \tag{8.4}$$

The fit has been computed maximising the likelihood from Eq. (8.1), the uncertainties have been determined from the profile likelihood ratio as explained in Eq. (7.5) and the result is

$$\hat{\mu}_{\text{fit}} = 1.01\ ^{+0.29}_{-0.26}. \tag{8.5}$$

As expected, the result is the same in both cases. In Fig. 8.1 the profile likelihood ratio $q(\mu)$ is shown as a function of μ in two cases: using the simplified data and the standard model Higgs boson expectation. In the latter case the number of observed events is defined as $N = B+S$. Since the best fit value $\hat{\mu}_{\text{fit}}$ is approximately the same as the standard model expectation $\mu_{\text{SM}} = 1$, the two likelihood scans are similar. However, this is not the case for other matrix elements for which the best fit value differs from unity.

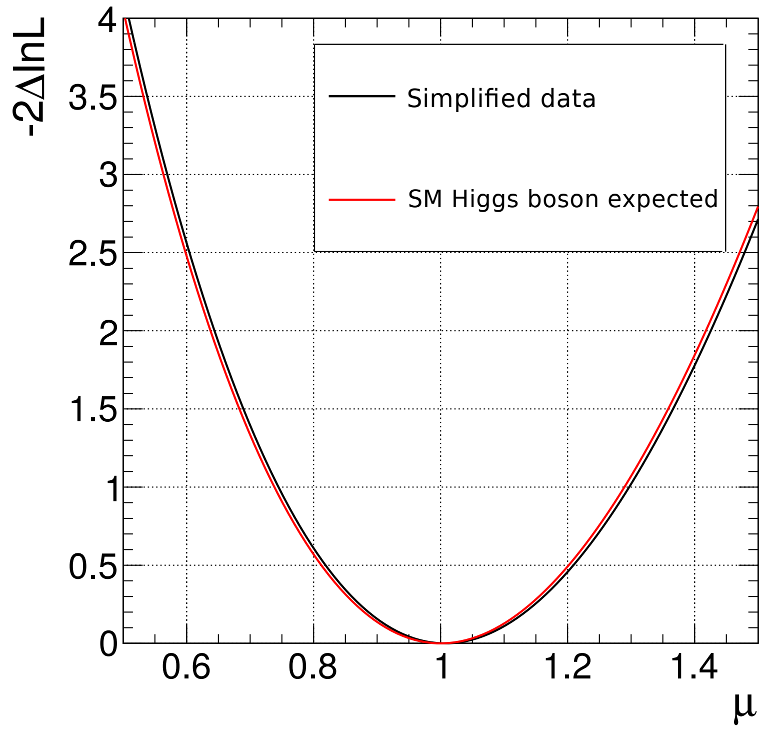

Figure 8.1: Fit for the signal strength μ (best fit value $\hat{\mu} = 1.01$) using the simplified data and the standard model Higgs boson expectation.

8.2 Two production modes

In this section we study a 2×1 matrix, i.e. two production modes and one decay channel. We will examine the gluon fusion and the vector boson fusion as the production modes, and two photons as the decay mode.

8.2.1 Pure production modes

Let us consider the ideal case, in which we have two experimental final states: one receiving signal from one Higgs boson production mode (ggH) followed by the $H \to \gamma\gamma$ decay mode; the other receiving signal from a different Higgs boson production mode (VBF) followed by the same $H \to \gamma\gamma$ decay mode. In other words, we treat the experimental final states as "pure" states, as if there were no possibility to miscategorize the signal yield according to the production processes.

As described in Section 3.1, vector boson fusion is characterized by the production of a Higgs boson in association with two hadronic jets, while in gluon fusion at lowest order the Higgs boson is produced alone. In this simplified picture, the Higgs boson production in gluon fusion can be measured in the final state with a Higgs boson and no hadronic jets, while VBF can be measured in the final state with a Higgs boson and two jets.

With two production channels, the expected yield can be parametrized in terms of an overall factor θ common to both production modes and a relative scaling factor λ between the two. The likelihood function is the product of two Poisson probabilities

$$L = P_{0j}P_{2j} = P(N_{0j}|B_{0j} + \theta S_{0j})P(N_{2j}|B_{2j} + \theta\lambda S_{2j}), \tag{8.6}$$

expressed in terms of observed number of events N_i, expected SM Higgs boson signal S_i, and expected background B_i for the final states with zero jets ($0j$) and two jets ($2j$).

Compared to the case with only one production mode, the simplified data card includes also information of the vector boson fusion: the number of observed events

$N_{2j} = 38.2$, the expected total signal $S_{2j} = 12.0$ and the background $B_{2j} = 20$. The equations for the best fit values of θ and λ have been derived by differentiating the likelihood function L (8.6) with the respect of θ and λ respectively

$$\hat{\theta} = \frac{N_{0j} - B_{0j}}{S_{0j}} \text{ and } \hat{\lambda} = \frac{S_{0j}(N_{2j} - B_{2j})}{S_{2j}(N_{0j} - B_{0j})}. \tag{8.7}$$

The best fit value of λ, i.e. $\hat{\lambda}$ has been obtained using the equation above and the result is $\hat{\lambda} = 1.5$, which can be seen also in Fig. 8.2.

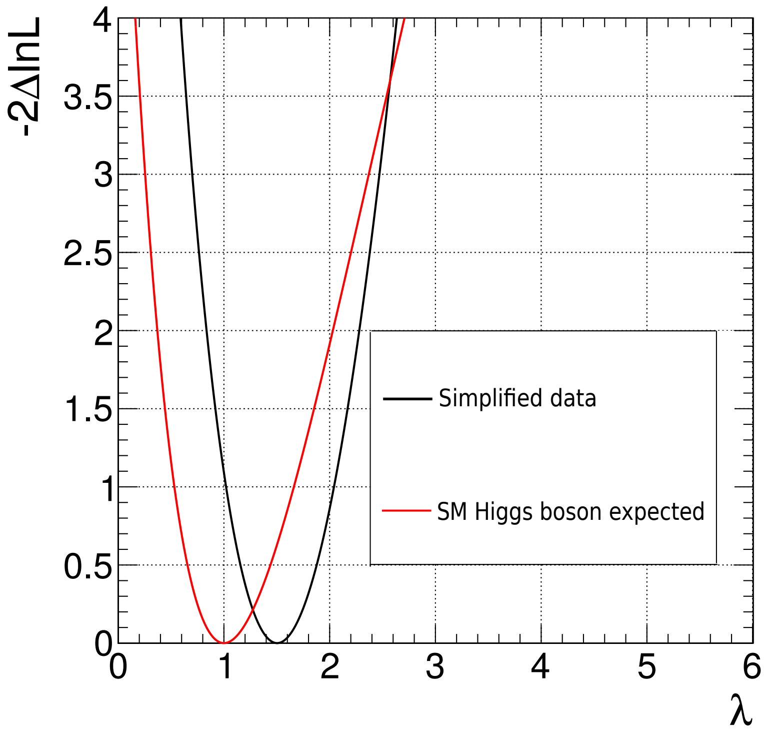

Figure 8.2: Fit for the relative scaling factor λ (best fit value $\hat{\lambda} = 1.5$) using the simplified data and the standard model Higgs boson expectation.

8.2.2 Mixing between production modes

In order to make the simplified data cards resemble more the real data cards, it is important to understand that it is impossible to categorize perfectly the signal yields according to the production modes.

The 0-jet tagged signal yield (mostly from the gluon fusion) can contain a contribution from the vector boson fusion that appears as the gluon fusion: the jets in the vector boson fusion might not be observed, e.g. because they fall outside of the acceptance of the calorimeters, have too low transverse momentum, or fail selection criteria used in the analysis to reject jets from noise or pile-up. In turn, the 2-jet tagged signal yield (mostly from the vector boson fusion) can contain a contribution from the gluon fusion, mostly because in the gluon fusion there might be extra QCD radiation, producing jets which can be misinterpreted as the ones from the quarks from the vector boson fusion.

The Higgs–strahlung and the top fusion categories can have experimental signatures that are cleaner to select: tagging VH with V decaying into $\ell\nu$, $\ell\ell$, $\nu\nu$ requires leptons and/or missing transverse energy and ttH tags require leptons, missing transverse energy or a large multiplicity of jets (including b-tagged ones). These experimental signatures are less likely to be mimicked by extra QCD radiation in a gluon fusion event. The Higgs–strahlung with V decaying hadronically into a pair of jets has instead challenges similar to that with VBF.

Different Higgs boson decay modes can sometimes also be hard to categorize. In particular, Higgs boson decays into W or tau pairs can have similar signatures when the W bosons and taus decay leptonically, especially in events that contain additional W bosons such as from WH or ttH production, complicating the event interpretation. Separation of WW and $\tau\tau$ decay modes is instead easier to achieve in gluon fusion production through the different kinematic distributions of the leptons and neutrinos in the two final states[1]. For ttH searches with electrons or muons (targeting $H \to WW$), and with hadronic taus (targeting $H \to \tau\tau$), the ratio between expected signal yields from the two Higgs boson decay modes are $WW/\tau\tau \approx 3$ and $WW/\tau\tau \approx 1/4$, respectively. Also the decay channel of two bottom quarks can contain some signal from two W bosons (contamination level approximately 10%), because the W boson can decay into jets containing the charm quark that could be misidentified as b-jets.

[1]The vector sum of the neutrino transverse momenta can be inferred from the missing transverse energy, since there are no other invisible particles in the event.

Approximate assumptions about the signal composition in the different final states were made from the simulations and the contamination level can be seen in Table 8.1. The VH and ttH are intentionally left out because, as explained above, they are considered as pure production modes.

	$\gamma\gamma_{0j}$	$\gamma\gamma_{2j}$	WW_{0j}	WW_{2j}	ZZ_{0j}	ZZ_{2j}	$\tau\tau_{0j}$	$\tau\tau_{2j}$
ggH	0.90	0.30	0.95	0.15	0.90	0.70	0.80	0.20
VBF	0.10	0.70	0.05	0.85	0.10	0.30	0.20	0.80

Table 8.1: The fraction of the signal from the gluon fusion and the vector boson fusion in each tagged final state.

In reality, the composition in the diphoton channel is more complex than in the approximation: there are several subcategories in 0-jet and 2-jet channels, which have different mixtures. In the real data cards there is also signal from the VH, which in some subcategories looks like the signal from the vector boson fusion. These are the reasons why the simplified data cards do not correspond to the real data cards perfectly.

The likelihood function is the same as the Eq. (8.6) with two additional terms which take the mixing into account

$$L = P_{0j}P_{2j} = P(N_{0j}|B_{0j} + \theta(S_{0j} + \lambda S_{VBF\to 0j}))P(N_{2j}|B_{2j} + \theta(S_{gg\to 2j} + \lambda S_{2j})), \quad (8.8)$$

where $S_{VBF\to 0j}$ stands for the part of the signal in 0-jet channel from the vector boson fusion and $S_{gg\to 2j}$ for the part of the signal in 2-jet channel from the gluon fusion.

The contamination is considered in the simplified data cards by splitting the expected signal yield for each channel in the different production modes according to Table 8.1. The fitting has been done by maximising the likelihood (8.8) and the result is $\hat{\lambda} = 1.92$. The difference between the pure and mixed cases can be seen in the Fig. 8.3: the uncertainty on the relative scaling factor λ increases when the mixing between production modes is taken into account.

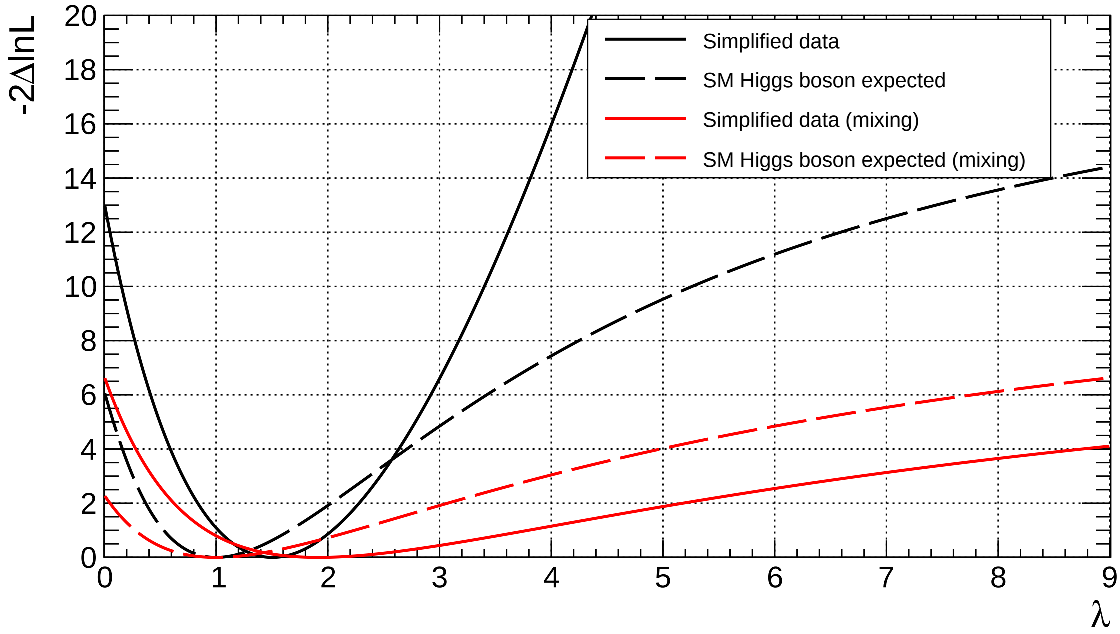

Figure 8.3: The black line represents the pure case and the red line the case in which the mixing between the production modes is taken into account - the dashed lines represent the standard model expectation.

8.2.3 Theoretical uncertainty

As a further step in complexity towards the real data cards, theoretical uncertainties are introduced. The theoretical uncertainty takes into account two sources of uncertainty on the predicted cross section. The first source is the uncertainties on the inputs used to calculate the predicted cross section: knowledge of the parton distribution functions and of other standard model parameters, e.g. the strong coupling constant and quark masses. The second source is the missing higher order contributions and other theoretical approximations in the calculation done in perturbation theory. While the second of these two is certainly theoretical, the first at least in part arises from experimental measurements of the input parameters. The theoretical uncertainty is modelled as a multiplicative correction of the expected event yield, and described with a log-normal distribution.

The systematic error assumed on the gluon fusion cross section is 10% and 30% in 0 jets and 2 jets final states, respectively. The uncertainty on the gluon fusion cross section in 2 jets final state is larger because the process is more complicated and has been computed only to a lower order in QCD, whereas the simpler process, i.e. the gluon fusion in 0 jet final state, has been computed to a next-to-next-leading order producing a smaller uncertainty. Since the vector boson fusion is an electroweak process, it is less sensitive to QCD corrections. Thus the theoretical uncertainty on the cross section is much smaller, and has been neglected in this study.

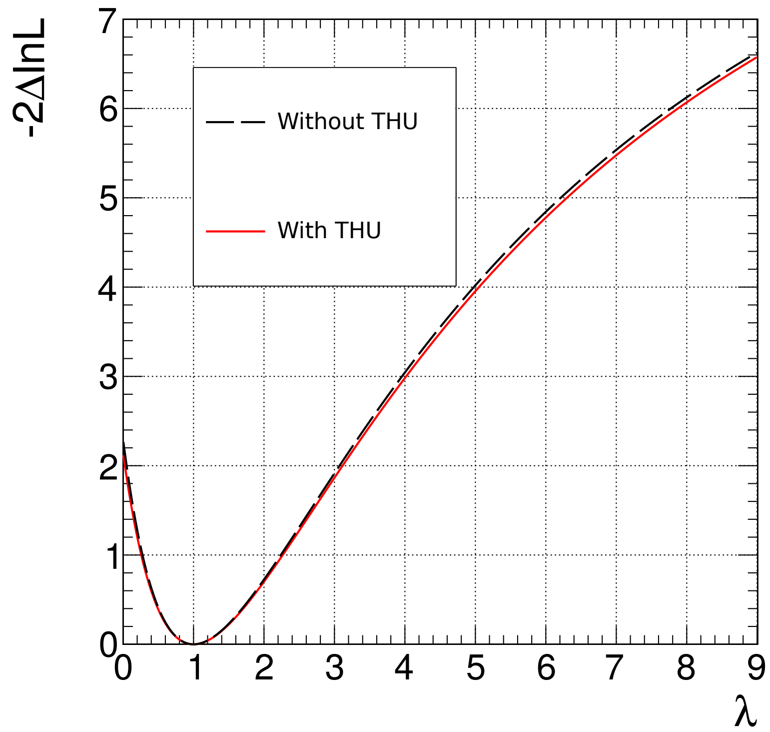

Figure 8.4: Fit for the relative scaling factor λ with the theoretical uncertainty (red line) and without the theoretical uncertainty (black dashed line) when expecting the SM Higgs boson. The abbreviation "THU" stands for the theoretical uncertainty.

The effect of the theoretical uncertainty is shown in Fig. 8.4, where one can see a difference between the simplified data card without the theoretical uncertainty and with the theoretical uncertainty: in the latter case the shape of the log-likelihood function is broader, resulting in a weaker constraint on the parameter of interest λ. The two-dimensional likelihood scan for the parameters λ and θ can be seen in Fig. 8.5.

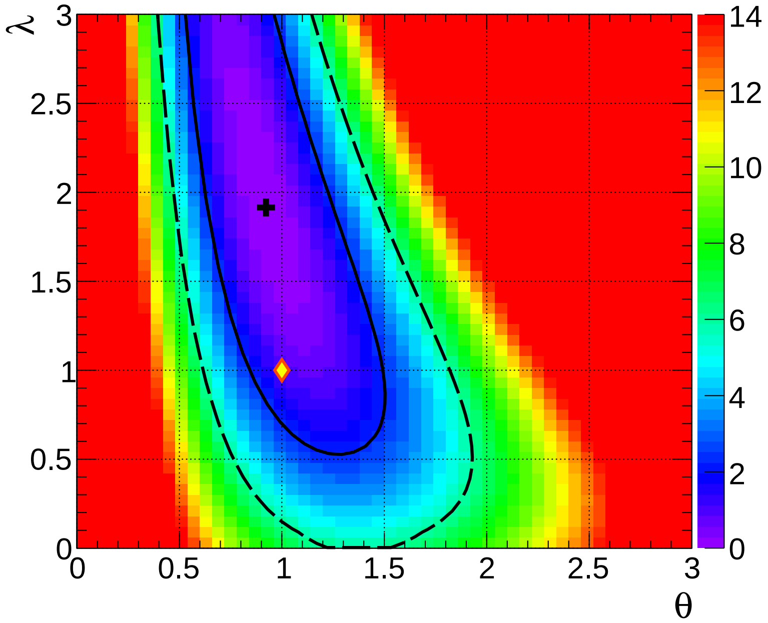

Figure 8.5: The 2D likelihood scan for the parameters θ and λ, taking into account the theoretical uncertainty on the gluon fusion cross section. The likelihood ratio q is at its minimum when $\lambda = 1.92$ and $\theta \approx 0.92$. The agreement between the pseudo-experiments and the standard model is better than 1σ.

9. Examining 2×2 matrices using the double ratio approach

Calculating the rank of the 5×4 signal strength matrix is impossible because of the missing elements. However, if the 5×4 matrix has rank 1, any sub-matrix of it must have rank 1, too. Let us start from the smallest sub-matrices that can be used for the rank 1 test, i.e. the 2×2 matrices - smaller matrices than those have rank 1 by definition. We approach this case using the double ratio method described in Chapter 7.

9.1 Determining interesting 2×2 matrices

Since numerous 2×2 sub-matrices can be extracted from the 5×4 matrix, it is unfeasible to analyse each and every one of them in detail. Moreover, our interest is in applying the double ratio method to the sub-matrices for which the corresponding double ratio is measured more accurately.

A program has been written for determining the interesting 2×2 matrices. The program goes through all the possible 2×2 matrices and calculates the double ratio and the double ratio's uncertainty. The uncertainty on the double ratio ρ for the matrix (6.3) has been estimated using

$$\Delta\rho = |\rho|\sqrt{\left(\frac{\Delta\mu_{\text{ggH},\gamma\gamma}}{\mu_{\text{ggH},\gamma\gamma}}\right)^2 + \left(\frac{\Delta\mu_{\text{VBF},\gamma\gamma}}{\mu_{\text{VBF},\gamma\gamma}}\right)^2 + \left(\frac{\Delta\mu_{\text{ggH,WW}}}{\mu_{\text{ggH,WW}}}\right)^2 + \left(\frac{\Delta\mu_{\text{VBF,WW}}}{\mu_{\text{VBF,WW}}}\right)^2}. \quad (9.1)$$

Production modes	Decay modes	$\rho \pm \Delta\rho$	$\Delta\rho_{\text{exp}}$
ggH, VBF	$\gamma\gamma$, WW	0.54 ± 0.54	0.82
ggH, VBF	$\gamma\gamma$, ZZ	1.17 ± 0.89	1.0
ggH, VBF	$\gamma\gamma$, $\tau\tau$	0.75 ± 0.58	0.82
ggH, VBF	WW, $\tau\tau$	1.4 ± 1.5	0.81
ggH, VBF	ZZ, $\tau\tau$	0.64 ± 0.57	1.0
ggH, VH	WW, $\tau\tau$	1.0 ± 1.7	1.5
ggH, ttH	$\gamma\gamma$, WW	1.9 ± 1.9	2.6
ggH, ttH	WW, $\tau\tau$	-0.3 ± 1.1	5.1
VBF, VH	WW, $\tau\tau$	0.7 ± 1.4	1.5
VBF, ttH	WW, $\tau\tau$	-0.22 ± 0.84	5.1
VH, ttH	$\gamma\gamma$, $\tau\tau$	-0.3 ± 1.4	5.4
VH, ttH	$\gamma\gamma$, bb	0.14 ± 0.46	2.9
VH, ttH	WW, $\tau\tau$	-0.3 ± 1.3	5.3
VH, ttH	WW, bb	0.13 ± 0.41	2.7

Table 9.1: The 2×2 submatrices and their double ratios which have the uncertainty under 200%. The expected uncertainty ($\Delta\rho_{\text{exp}}$) has been calculated by setting all $\mu_{i,j}$ to unity, but by keeping each $\Delta\mu_{i,j}$ unchanged. In cases that include the top fusion production, the expected uncertainties are very large due to signal strength values and their uncertainties that are much larger than unity, and the small values of ρ.

All the ratios which have the uncertainty under 200% are listed in Table 9.1, where also each expected uncertainty $\Delta\rho_{\text{exp}}$ is shown. The expected uncertainty has been calculated using Eq. (9.1) by setting all signal strengths $\mu_{i,j} = 1$, as expected for the SM Higgs boson, but by keeping $\Delta\mu_{i,j}$ unchanged, assuming that the uncertainties reflect only the sensitivity of the analysis and are approximately independent of the observed $\mu_{i,j}$. Especially in cases that include the top fusion production as the other production mode, the expected uncertainties are very large compared to the observed uncertainties. This originates from signal strength values and their uncertainties that are much larger than unity, and the small observed values of ρ.

The cases $(\mathrm{ggH}, \mathrm{VBF}) \times (\gamma\gamma, \mathrm{WW})$ and $(\mathrm{ggH}, \mathrm{VBF}) \times (\mathrm{ZZ}, \tau\tau)$ have been selected as the most promising ones for the analysis because they have the smallest uncertainties. The cases $(\mathrm{VBF}, \mathrm{ttH}) \times (\mathrm{WW}, \tau\tau)$ and $(\mathrm{VH}, \mathrm{ttH}) \times (\mathrm{WW}, \mathrm{bb})$ have been selected for the analysis as good examples of mixing of the decay modes.

In order to test the data against the hypothesis of the rank 1, a parametrisation of the expected signal yield has been developed. As described in Chapter 7, the model used for this purpose has been written by using the definition of the double ratio ρ. For example, in the case $(\mathrm{ggH}, \mathrm{VBF}) \times (\gamma\gamma, \mathrm{WW})$, signal strengths in the four possible combinations of production and decay mode are parametrized in terms of: two signal strengths, $\mu_{\gamma\gamma}$ and μ_{WW}, for gluon fusion production followed by the decay in $\gamma\gamma$ and WW; one common ratio λ between signal strengths in VBF versus gluon fusion; and the double ratio ρ, to allow the matrix to have rank different from 1. The resulting expressions for the four signal strengths are

$$\begin{array}{c|cc} & \mathrm{H} \to \gamma\gamma & \mathrm{H} \to \mathrm{WW} \\ \hline \mathrm{ggH} & \mu_{\gamma\gamma} & \mu_{\mathrm{WW}} \\ \mathrm{VBF} & \lambda \cdot \mu_{\gamma\gamma} & \lambda \cdot \mu_{\mathrm{WW}} \cdot \rho \end{array} \tag{9.2}$$

For convenience in presenting the likelihood as function of ρ, however, we choose to associate it to the production mode that is less accurately measured, and to the decay mode which yields $\rho < 1$. The algorithm has been tested using the simplified and the real CMS data cards.

9.2 Simplified data cards

The simplified data cards have been written such that the mixing between the production modes, the possible mixing between the decay modes, and the theoretical uncertainties on gluon fusion production process are taken into account.

9.2.1 The likelihood scans

The likelihood scan for the double ratio ρ was done in each aforementioned case.

Case $(\text{ggH}, \text{VBF}) \times (\gamma\gamma, \text{WW})$

The value of the double ratio ρ from the naive calculation was 0.54 ± 0.54, but the obtained best fit value is $0.40\,^{+0.90}_{-0.53}$, as shown in Fig. 9.1. These two values are different because of the mixing between production modes. The bigger the mixing between production modes, the more the value in Table 9.1 and the best fit value may differ. From the value of profile likelihood at $\rho = 1$, $q(\rho = 1) = 0.58$, a p-value of 0.45 is computed using (7.4). This means there is 45% probability to measure $q(1) > 0.58$ if the matrix has rank 1.

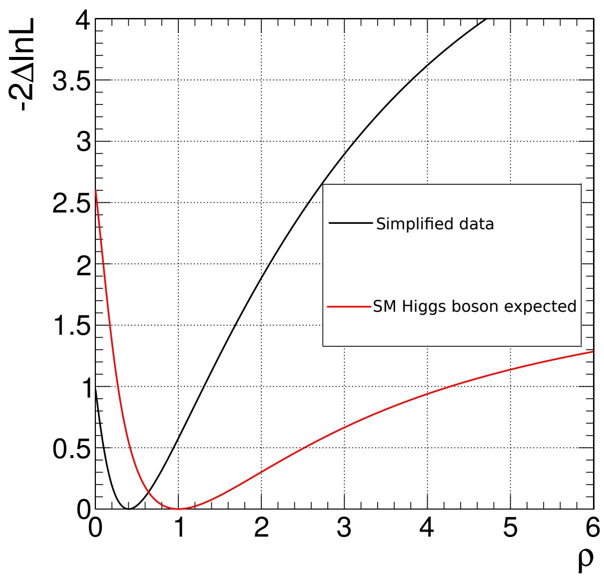

Figure 9.1: Case $(\text{ggH}, \text{VBF}) \times (\gamma\gamma, \text{WW})$. Fit for the double ratio ρ (best fit value $\hat{\rho} = 0.40\,^{+0.90}_{-0.53}$). The observed profile likelihood ratio $-2\Delta \ln \text{L}$ is 0.58 at $\rho = 1$.

Case $(\text{ggH}, \text{VBF}) \times (\text{ZZ}, \tau\tau)$

In this case the double ratio ρ is defined as

$$\rho = \frac{\mu_{\text{ggH,ZZ}}\ \mu_{\text{VBF},\tau\tau}}{\mu_{\text{ggH},\tau\tau}\ \mu_{\text{VBF,ZZ}}}. \tag{9.3}$$

The best fit value is $\hat{\rho} = 0.17$, whereas the value obtained from the naive calculation was 0.64 ± 0.57.

The observed and expected graphs can be seen in Fig. 9.2. The expected graph is extremely wide, resulting from the low purity of the 2-jet category in the decay mode of two Z bosons as shown in Table 8.1. However, the observed graph is less flat. This is because of the small best fit value $\hat{\rho}$ which produces a compressed x axis. Furthermore, since ρ is a ratio of numbers with uncertainties, the confidence intervals for ρ can extend up to positive infinity if one of the terms in the denominator is compatible with zero within that confidence level. As it can be seen from the observed graph, large values of the double ratio can still be compatible at 68.2% confidence level with the simplified data, i.e. the positive uncertainty on ρ is much larger than the 0.57 from the naive calculation. The observed profile likelihood ratio $-2\Delta \ln \text{L}$ is 0.50 at $\rho = 1$. The p-value for the $\rho = 1$ hypothesis is 0.48.

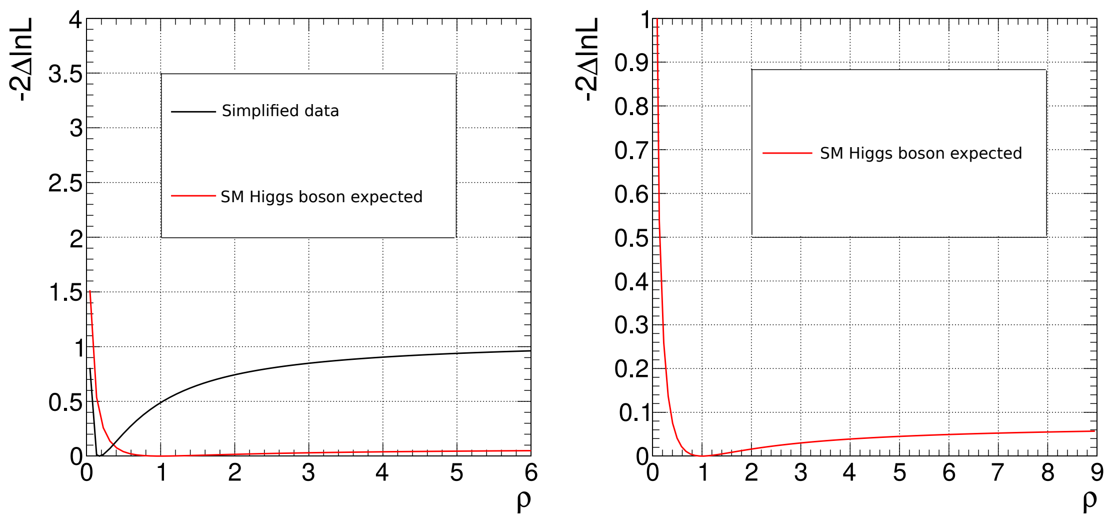

Figure 9.2: Case $(\text{ggH}, \text{VBF}) \times (\text{ZZ}, \tau\tau)$. Left: Fit for the double ratio ρ (best fit value $\hat{\rho} = 0.17$, value from the naive calculation was 0.64 ± 0.57). The observed profile likelihood ratio $-2\Delta \ln \text{L}$ is 0.50 at $\rho = 1$. Right: The scan for the double ratio ρ when the standard model Higgs boson is expected. The sensitivity is bad, as it can be seen from the flat form of the graph, resulting from the low purity of the 2-jet category in the decay mode of two Z bosons.

Case $(\mathrm{VH}, \mathrm{ttH}) \times (\mathrm{WW}, \mathrm{bb})$

In this case the double ratio ρ multiplies the yield of the ttH associated production to bottom quarks, i.e. ρ is defined as

$$\rho = \frac{\mu_{\mathrm{VH,WW}}\ \mu_{\mathrm{ttH,bb}}}{\mu_{\mathrm{VH,bb}}\ \mu_{\mathrm{ttH,WW}}}. \tag{9.4}$$

The best fit value is $\hat{\rho} = 0.06\,^{+0.81}_{-0.06}$ (value from the naive calculation was 0.13 ± 0.41). As described in Subsection 8.2.2, the Higgs–strahlung and the ttH associated production are considered as pure production modes. However, the expected graph, shown in Fig. 9.3 with the observed graph, is wide because of the mixing between the decay modes and the large uncertainties on the observed signal strengths, especially on $\mu_{\mathrm{ttH,WW}}$ and $\mu_{\mathrm{ttH,bb}}$. The observed profile likelihood ratio $-2\Delta\ln\mathrm{L}$ is 1.18 at $\rho = 1$, and the p-value for the hypothesis $\rho = 1$ is 28%.

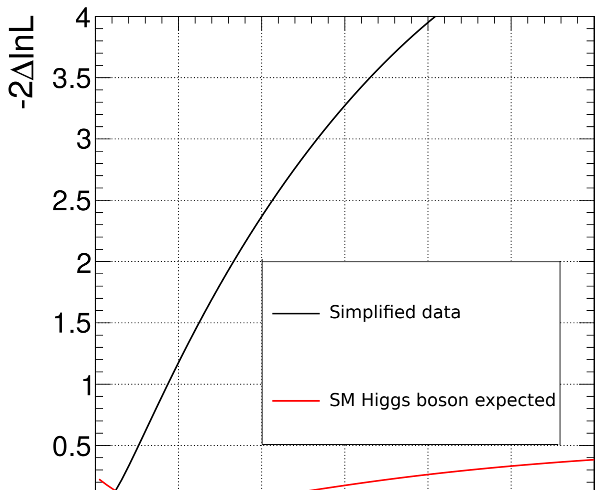

Figure 9.3: Case $(\mathrm{VH}, \mathrm{ttH}) \times (\mathrm{WW}, \mathrm{bb})$. Fit for the double ratio ρ (best fit value $\hat{\rho} = 0.06\,^{+0.81}_{-0.06}$, value from the naive calculation was 0.13 ± 0.41). The observed profile likelihood ratio $-2\Delta\ln\mathrm{L}$ is 1.18 at $\rho = 1$.

The effect of taking the mixing between the decay modes into account can be seen in Fig. 9.4: the uncertainty on the double ratio ρ increases when the mixing between the decay modes is present in the simplified data cards.

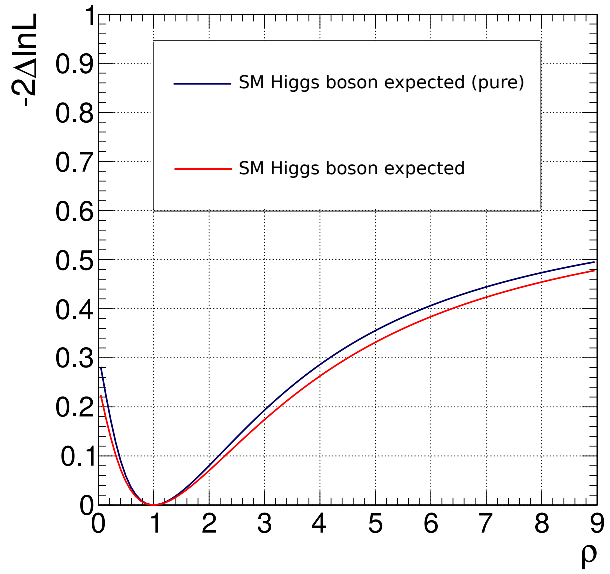

Figure 9.4: Case $(\mathrm{VH}, \mathrm{ttH}) \times (\mathrm{WW}, \mathrm{bb})$. Comparison plot of two likelihood scans when the standard model Higgs boson is expected: considering the decay modes as pure and taking the mixing between the decay modes into account, marked with dark blue and red, respectively. The uncertainty on ρ increases in the latter case.

Case $(\mathrm{VBF}, \mathrm{ttH}) \times (\mathrm{WW}, \tau\tau)$

In this case the double ratio ρ multiplies the yield of the ttH associated production to two tau leptons, i.e. ρ is defined as

$$\rho = \frac{\mu_{\mathrm{VBF,WW}}\ \mu_{\mathrm{ttH},\tau\tau}}{\mu_{\mathrm{VBF},\tau\tau}\ \mu_{\mathrm{ttH,WW}}}. \tag{9.5}$$

The result of the likelihood scans in the expected and observed cases can be seen in Fig. 9.5. The best fit value $\hat{\rho}$ is $0.0\,^{+1.1}_{-0.0}$ (value from the naive calculation was -0.22 ± 0.84); a deficit compared to the background-only hypothesis is observed in ttH with $\mathrm{H} \to \tau\tau$, leading to a negative signal strength, but in this analysis only physical particles ($\mu \geq 0$, $\lambda \geq 0$, $\rho \geq 0$) are considered. The observed profile likelihood ratio $-2\Delta \ln \mathrm{L}$ is approximately 0.92 when the double ratio ρ is at unity. The p-value against the single Higgs boson hypothesis is 0.34.

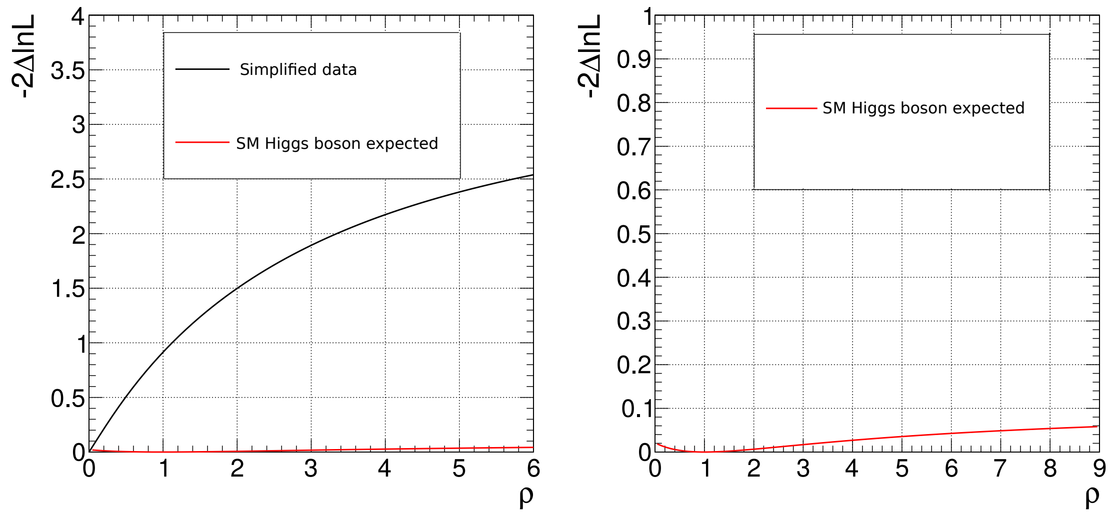

Figure 9.5: Case $(\mathrm{VBF}, \mathrm{ttH}) \times (\mathrm{WW}, \tau\tau)$. Left: Fit for the double ratio ρ (best fit value $\hat{\rho} = 0.0\,^{+1.1}_{-0.0}$, value from the naive calculation was -0.22 ± 0.84). The observed profile likelihood ratio $-2\Delta \ln \mathrm{L}$ is approximately 0.92 at $\rho = 1$. Right: The scan for the double ratio ρ when the standard model Higgs boson is expected. The sensitivity is bad, as it can be seen from the flat form of the graph, due to the poor sensitivity of the input measurements and to the mixing between the decay modes.

The observed graph is again less wide, resulting from the small best fit value that compresses the x axis. The expected uncertainty is larger than in the case $(\mathrm{ggH}, \mathrm{VBF}) \times (\mathrm{ZZ}, \tau\tau)$ (see Fig. 9.2) even though there is no mixing between the vector boson fusion and the top production: the contamination from gluon fusion in the VBF measurements, and the larger amount of mixing between the decay modes lead to wider expected graph. As it can be seen in Fig. 9.6, the expected uncertainty on the double ratio ρ increases when the mixing between the decay modes of two W bosons and τ leptons is taken into account. The effect is more notable than the one seen in Fig. 9.4 because the amount of mixing between the decay modes is larger in this case.

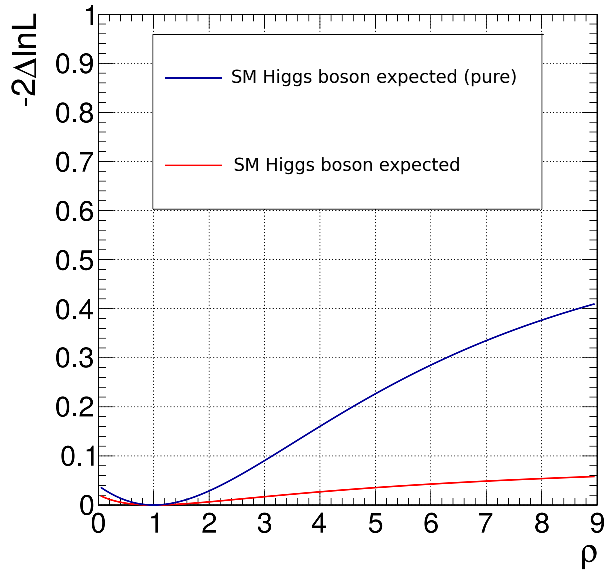

Figure 9.6: Case $(\mathrm{VBF}, \mathrm{ttH}) \times (\mathrm{WW}, \tau\tau)$. Comparison plot of two likelihood scans when the standard model Higgs boson is expected: considering the decay modes as pure and taking the mixing between the decay modes into account, marked with dark blue and red, respectively. The uncertainty on ρ increases in the latter case. The contamination from gluon fusion in the VBF measurements is present in both cases.

9.2.2 The test statistic distribution

As explained in Section 7, the test statistic is approximately distributed as a χ^2 distribution with one degree of freedom. The expected test statistic distribution for the case $(\mathrm{ggH}, \mathrm{VBF}) \times (\gamma\gamma, \mathrm{WW})$ was generated in order to see if the p-value obtained interpreting the test statistic $q(\rho = 1)$ as a χ^2 distribution with one degree of freedom is a good approximation of the correct result.

The pseudo-experiments were generated assuming the standard model hypothesis, i.e. by setting all parameters equal to unity. The value of the test statistic at $\rho = 1$ was evaluated for each pseudo-experiment, to derive the expected distribution of the test statistic under that hypothesis.

The observed value of the test statistic (7.6) is $q_\rho^{\mathrm{obs}} = 0.58$ and the fraction of pseudo-experiments for which $q_\rho \geq 0.58$ is 0.442 ± 0.007, where the uncertainty quoted is calculated from the 68.2% Clopper-Pearson interval [47]. The p-value defined from the test statistic distribution is compatible with the p-value of 0.45 obtained from the likelihood scan (Fig. 9.1).

Since it is interesting to see how well the expected distribution agrees with a χ^2 distribution with one degree of freedom, the data was fitted using a probability density function that generalizes the χ^2 distribution also for non-integer values of the number of degrees of freedom. The expected test statistic distribution and the result of the fit can be seen in Fig. 9.7, where the observed value of the test statistic is marked with the blue arrow, the fit result with the red line and the χ^2 distribution with one degree of freedom with the green line. The χ^2 fit is successful (χ^2/ndf $\approx$ 1.5) and the number of degrees of freedom is approximately one as it should be.

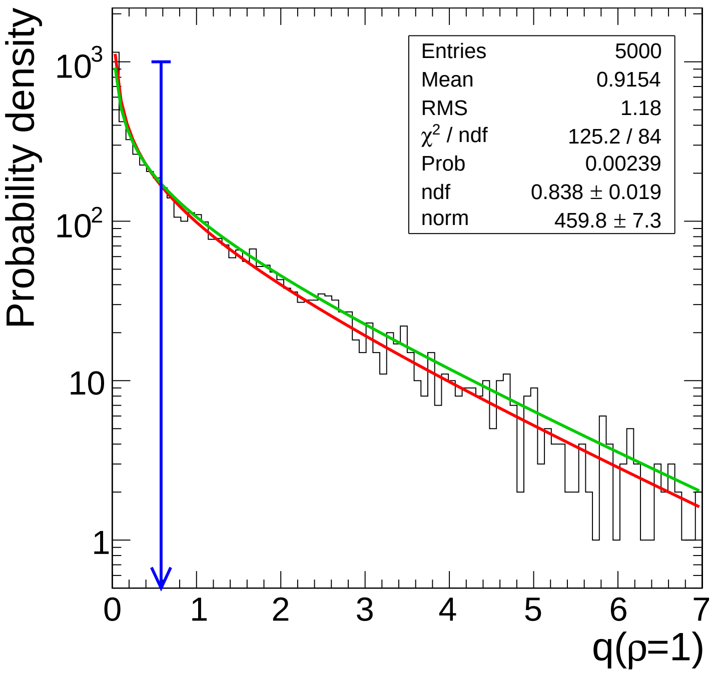

Figure 9.7: The expected test statistic distribution for the case (ggH, VBF) $\times$ ($\gamma\gamma$, WW), the result of χ^2 fit (red line) and the χ^2 distribution with one degree of freedom (green line). The observed value of the test statistic, marked with the blue arrow, is $q_\rho^{\text{obs}} = 0.58$ and the p-value against the hypothesis of a single Higgs boson is 0.442 ± 0.007.

9.3 Real data cards

After going through the case of the 2×2 matrices with simplified data cards, we move on to the real data cards. Since these real data cards include all four production modes, the model has to be modified so that it takes into account all of them; the gluon fusion, the vector boson fusion, the Higgs–strahlung and the ttH associated production:

$$\begin{array}{c|cc} & \mathrm{H}\to\gamma\gamma & \mathrm{H}\to\mathrm{WW} \\ \hline \mathrm{ggH} & \mu_{\gamma\gamma} & \mu_{\mathrm{WW}} \\ \mathrm{VBF} & \lambda\cdot\mu_{\gamma\gamma} & \lambda\cdot\mu_{\mathrm{WW}}\cdot\rho \\ \mathrm{VH} & \lambda_{\mathrm{v}}\cdot\mu_{\gamma\gamma} & \lambda_{\mathrm{v}}\cdot\mu_{\mathrm{WW}}\cdot\rho_{\mathrm{v}} \\ \mathrm{ttH} & \lambda_{\mathrm{t}}\cdot\mu_{\gamma\gamma} & \lambda_{\mathrm{t}}\cdot\mu_{\mathrm{WW}}\cdot\rho_{\mathrm{t}} \end{array} \tag{9.6}$$

The only parameter of interest is ρ: $\mu_{\gamma\gamma}$, μ_{WW}, λ, λ_{v} , λ_{t}, ρ_{v} and ρ_{t} have been profiled. By introducing the parameters λ_{t}, λ_{v}, ρ_{t} and ρ_{v}, we ignore the additional production modes VH and ttH in the test for ρ.

9.3.1 The likelihood scan

The likelihood scan using the real data cards can be seen in Fig. 9.8. The best fit value $\hat{\rho}$ is $0.68\,^{+1.36}_{-0.62}$. The p-value for the hypothesis $\rho = 1$ is 0.74.

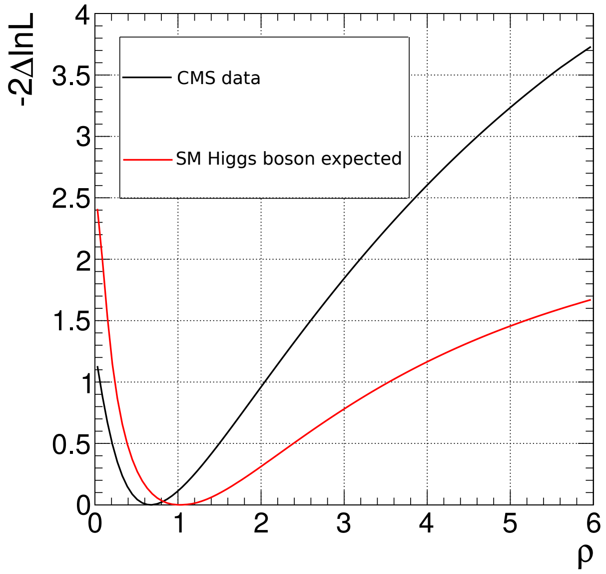

Figure 9.8: Likelihood scan for the double ratio ρ (best fit value is $\hat{\rho} = 0.68\,^{+1.36}_{-0.62}$), using the real data cards for the sub-matrix $(\mathrm{ggH}, \mathrm{VBF}) \times (\gamma\gamma, \mathrm{WW})$.

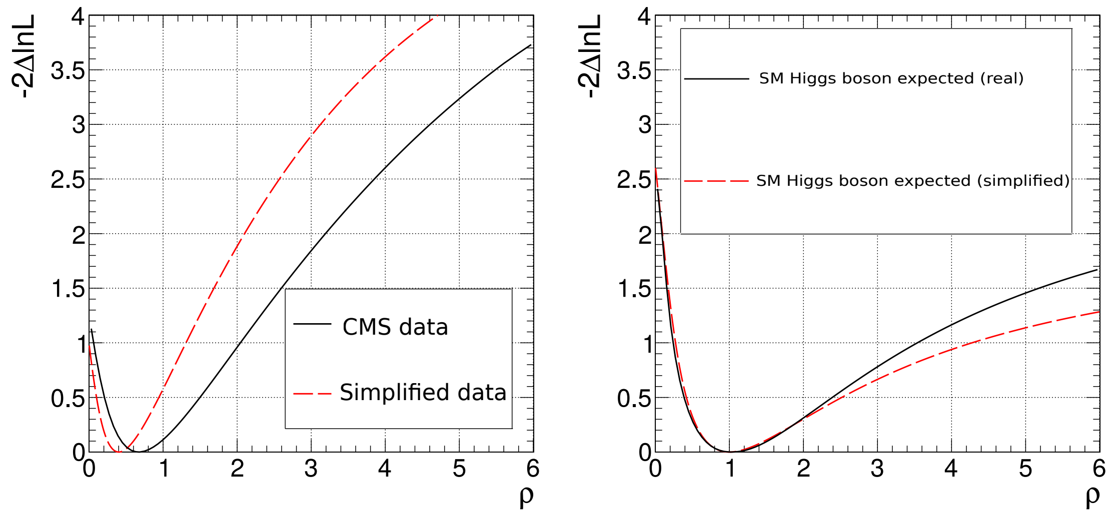

Figure 9.9: Left: The observed likelihood scans using simplified (red dashed line) and real CMS data (black solid line). Right: The expected likelihood scans for a standard model Higgs boson using simplified (red dashed line) and real CMS data (black solid line) uncertainties.

The differences between the likelihood scans for ρ, obtained using the simplified and real data, are easy to see from Fig. 9.9. The observed graph is slightly wider than the one obtained with the simplified data and the best fit values of ρ are different. This most likely results from the difference between the signal compositions in the simplified and real data cards. The approximated signal compositions, presented in Table. 8.1, do not perfectly reflect the signal compositions in the real data cards in which each production tag is a combination of many channels with different signal compositions.

The expected graph from the full data is more narrow than the one obtained with the simplified data when ρ is large. The way the simplified data cards were constructed as counting experiments is based on an assumption that the expected absolute uncertainty on the signal strength is approximately equal to the observed uncertainty. However, this might be a bad approximation for channels that in reality have a lower number of background events than the constant value used to construct the simplified data cards, i.e. 20. Lower background yields lead a less Gaussian likelihood than what could be obtained in cases where the background number of events is large. Indeed, the usage of the uncertainty on the signal strength in building the simplified data cards guarantees that the simplified and full likelihoods match only near the minimum, not at small or large values of ρ that in this case are excluded more strongly in the real data than in the simplified data.

9.3.2 The test statistic distribution

To examine if the asymptotic expression for the p-value from Eq. (7.4) is a good approximation also in the case of real data cards, the expected test statistic distribution was estimated assuming the standard model hypothesis.

Using CMS data, the observed value of the test statistic (7.6) is $q_\rho^{\text{obs}} = 0.113$ and the fraction of pseudo-experiments for which $q_\rho \geq 0.113$ is $0.709\,^{+0.006}_{-0.007}$. The p-value defined from the test statistic distribution is in consensus with the p-value of 0.74 obtained from the likelihood scan (Fig. 9.8).

The expected test statistic distribution was again fitted using a probability density function that generalizes the χ^2 distribution to non-integer values of the number of degrees of freedom. The distribution and the result of χ^2 fit can be seen in Fig. 9.10, where also a χ^2 distribution with one degree of freedom is shown.

The χ^2 fit is as successful as earlier ($\chi^2/\text{ndf} \approx 1.5$) but the number of degrees of freedom is slightly smaller than obtained from the previous fit, shown in Fig. 9.7. This is because of the small event yields in the vector boson fusion categories, leading to a less Gaussian-like likelihood: a Poisson distribution can be derived from a Gaussian distribution when the total number of events is large. As a result, the situation is not ideal for Wilks' theorem which states that the test statistic is approximatively distributed as a χ^2 distribution with one degree of freedom if the sample size approaches infinity.

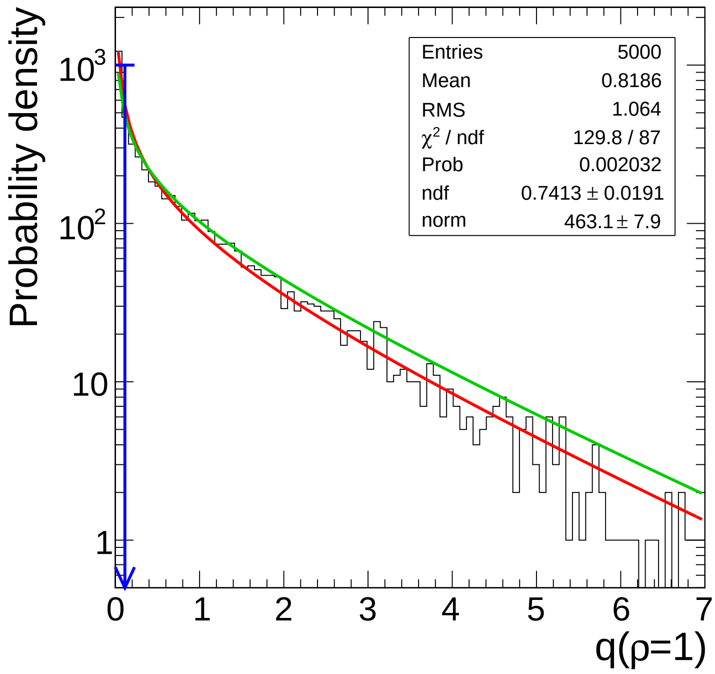

Figure 9.10: The expected test statistic distribution for the case $(\mathrm{ggH}, \mathrm{VBF}, \mathrm{VH}, \mathrm{ttH}) \times (\gamma\gamma, \mathrm{WW})$, the result of χ^2 fit (red line) and the χ^2 distribution with one degree of freedom (green line). The observed value of the test statistic is $q_\rho^{\mathrm{obs}} = 0.113$ (marked with the blue arrow) and the p-value against the hypothesis of a single Higgs boson is $0.709\,^{+0.006}_{-0.007}$.

10. Exploiting all the available information

As described in Section 7, generalizing the statistical test from 2×2 matrices to 5×4 matrices requires introducing more general parameters than the double ratio ρ. Namely, the model of the general rank 1 matrix is chosen to be decomposed as the tensor product of two vectors (μ_i) and $(1, \lambda_{\mathrm{VBF}}, \lambda_{\mathrm{VH}}, \lambda_{\mathrm{ttH}})$, whereas in the model of the most general $5{\times}4$ matrix each λ_j is redefined such that it also depends on the decay mode i.

In principle, the tensor product of the two vectors (μ_i) and $(1, \lambda_{\mathrm{VBF}}, \lambda_{\mathrm{VH}}, \lambda_{\mathrm{ttH}})$ is not the most general parametrization of a rank 1 matrix since the gluon fusion measurement (μ_i) is not allowed to be at zero if the other measurements for the same decay mode are non-zero - fitting this kind of case would lead to incorrect result.

A solution would be to use an alternative parametrization where the components of the vector $(1, \lambda_{\mathrm{VBF}}, \lambda_{\mathrm{VH}}, \lambda_{\mathrm{ttH}})$ are replaced with four non-negative expressions constrained to have a fixed sum, which can be achieved with:

$$\begin{aligned} 1 &\to 4 \cdot f/4 \\ \lambda_{\mathrm{VBF}} &\to 4 \cdot (1 - f/4) \cdot g/3 \\ \lambda_{\mathrm{VH}} &\to 4 \cdot (1 - f/4) \cdot (1 - g/3) \cdot h/2 \\ \lambda_{\mathrm{ttH}} &\to 4 \cdot (1 - f/4) \cdot (1 - g/3) \cdot (1 - h/2), \end{aligned} \tag{10.1}$$

where f is in range $[0, 4]$, g in range $[0, 3]$ and h in range $[0, 2]$. In the hypothesis of a single Higgs boson the parameters f, g and h would be set to unity. That said, in practice the observed signal strength for the gluon fusion in each decay mode is non-zero so it is not necessary to redefine the parametrization of the general rank 1 matrix.

10.1 Defining the models and the test statistic

The expected signal yields are defined as function of the parameters according to

1) μ_i, λ_{VBF}, λ_{VH} and λ_{ttH} (the general rank 1 matrix)

$$\begin{array}{c|ccccc} & \mathrm{H}\to\gamma\gamma & \mathrm{H}\to\mathrm{WW} & \mathrm{H}\to\mathrm{ZZ} & \mathrm{H}\to\tau\tau & \mathrm{H}\to\mathrm{bb} \\ \hline \mathrm{ggH} & \mu_{\gamma\gamma} & \mu_{\mathrm{WW}} & \mu_{\mathrm{ZZ}} & \mu_{\tau\tau} & \mu_{\mathrm{bb}} \\ \mathrm{VBF} & \lambda_{\mathrm{VBF}}\cdot\mu_{\gamma\gamma} & \lambda_{\mathrm{VBF}}\cdot\mu_{\mathrm{WW}} & \lambda_{\mathrm{VBF}}\cdot\mu_{\mathrm{ZZ}} & \lambda_{\mathrm{VBF}}\cdot\mu_{\tau\tau} & \lambda_{\mathrm{VBF}}\cdot\mu_{\mathrm{bb}} \\ \mathrm{VH} & \lambda_{\mathrm{VH}}\cdot\mu_{\gamma\gamma} & \lambda_{\mathrm{VH}}\cdot\mu_{\mathrm{WW}} & \lambda_{\mathrm{VH}}\cdot\mu_{\mathrm{ZZ}} & \lambda_{\mathrm{VH}}\cdot\mu_{\tau\tau} & \lambda_{\mathrm{VH}}\cdot\mu_{\mathrm{bb}} \\ \mathrm{ttH} & \lambda_{\mathrm{ttH}}\cdot\mu_{\gamma\gamma} & \lambda_{\mathrm{ttH}}\cdot\mu_{\mathrm{WW}} & \lambda_{\mathrm{ttH}}\cdot\mu_{\mathrm{ZZ}} & \lambda_{\mathrm{ttH}}\cdot\mu_{\tau\tau} & \lambda_{\mathrm{ttH}}\cdot\mu_{\mathrm{bb}} \end{array} \tag{10.2}$$

2) μ_i, λ^i_{VBF}, λ^i_{VH} and λ^i_{ttH} (the most general 5×4 matrix)

$$\begin{array}{c|ccccc} & \mathrm{H}\to\gamma\gamma & \mathrm{H}\to\mathrm{WW} & \mathrm{H}\to\mathrm{ZZ} & \mathrm{H}\to\tau\tau & \mathrm{H}\to\mathrm{bb} \\ \hline \mathrm{ggH} & \mu_{\gamma\gamma} & \mu_{\mathrm{WW}} & \mu_{\mathrm{ZZ}} & \mu_{\tau\tau} & \mu_{\mathrm{bb}} \\ \mathrm{VBF} & \lambda^{\gamma\gamma}_{\mathrm{VBF}}\cdot\mu_{\gamma\gamma} & \lambda^{\mathrm{WW}}_{\mathrm{VBF}}\cdot\mu_{\mathrm{WW}} & \lambda^{\mathrm{ZZ}}_{\mathrm{VBF}}\cdot\mu_{\mathrm{ZZ}} & \lambda^{\tau\tau}_{\mathrm{VBF}}\cdot\mu_{\tau\tau} & \lambda^{\mathrm{bb}}_{\mathrm{VBF}}\cdot\mu_{\mathrm{bb}} \\ \mathrm{VH} & \lambda^{\gamma\gamma}_{\mathrm{VH}}\cdot\mu_{\gamma\gamma} & \lambda^{\mathrm{WW}}_{\mathrm{VH}}\cdot\mu_{\mathrm{WW}} & \lambda^{\mathrm{ZZ}}_{\mathrm{VH}}\cdot\mu_{\mathrm{ZZ}} & \lambda^{\tau\tau}_{\mathrm{VH}}\cdot\mu_{\tau\tau} & \lambda^{\mathrm{bb}}_{\mathrm{VH}}\cdot\mu_{\mathrm{bb}} \\ \mathrm{ttH} & \lambda^{\gamma\gamma}_{\mathrm{ttH}}\cdot\mu_{\gamma\gamma} & \lambda^{\mathrm{WW}}_{\mathrm{ttH}}\cdot\mu_{\mathrm{WW}} & \lambda^{\mathrm{ZZ}}_{\mathrm{ttH}}\cdot\mu_{\mathrm{ZZ}} & \lambda^{\tau\tau}_{\mathrm{ttH}}\cdot\mu_{\tau\tau} & \lambda^{\mathrm{bb}}_{\mathrm{ttH}}\cdot\mu_{\mathrm{bb}} \end{array} \tag{10.3}$$

The first model is the general rank 1 matrix with eight parameters: μ_i, λ_{VBF}, λ_{VH} and λ_{ttH}, where the lower index i stands for the decay mode. The parameters of interest are λ_{VBF}, λ_{VH} and λ_{ttH} - the rest of the parameters are profiled.

The second model is the most general 5×4 matrix, which has 20 parameters: μ_i, λ^i_{VBF}, λ^i_{VH} and λ^i_{ttH}, where the index i stands for the decay mode. The 15 parameters of interest are λ^i_{VBF}, λ^i_{VH} and λ^i_{ttH}, while the signal strength parameters are profiled.

If the matrix has rank 1, the rows of these two models are the same, i.e. $\lambda_{\text{VBF}} \equiv \lambda^i_{\text{VBF}}$, $\lambda_{\text{VH}} \equiv \lambda^i_{\text{VH}}$ and $\lambda_{\text{ttH}} \equiv \lambda^i_{\text{ttH}}$. In other words, the general rank 1 parametrization can be recovered from the most general 5×4 parametrization by setting such constraints in the parameters of the latter parametrization.

The test statistic q_λ has been defined as a likelihood ratio between the likelihoods of two aforementioned models

$$q_\lambda = -2\ln \frac{L(\text{data}|\lambda^i_{\text{VBF}} = \hat{\lambda}_{\text{VBF}},\ \lambda^i_{\text{VH}} = \hat{\lambda}_{\text{VH}},\ \lambda^i_{\text{ttH}} = \hat{\lambda}_{\text{ttH}}, \hat{\mu}_i)}{L(\text{data}|\hat{\lambda}^i_{\text{VBF}}, \hat{\lambda}^i_{\text{VH}}, \hat{\lambda}^i_{\text{ttH}}, \hat{\mu}_i)}, \tag{10.4}$$

where the numerator is the likelihood to observe data assuming the general rank 1 matrix, and the denominator is the likelihood to observe data assuming the most general 5×4 matrix. There are 12 parameters constrained at the numerator of q_λ ($\#\{\lambda^i_j\} - \#\{\lambda_j\} = 15 - 3 = 12$), thus if Wilks' conditions were to be satisfied, q_λ should be distributed as a χ^2 distribution with 12 degrees of freedom.

The ratio between the production modes is not dependent of the decay mode if there is only one Higgs boson, and so the value of the profile likelihood ratio q_λ is small because both models fit the data equally well. However, if the rank is not equal to unity, the model of most general 5×4 matrix will fit the data better than the model of general rank 1 matrix, resulting a large value of the profile likelihood ratio q_λ.

Since the constraints in the test statistic (10.4) are simpler to implement if a parameter is set equal to some constant than equal to another parameter, the test statistic q_λ has been written in the form

$$\begin{aligned} q_\lambda &= q_2(\lambda^i_{\text{VBF}} = 1, \lambda^i_{\text{VH}} = 1, \lambda^i_{\text{ttH}} = 1) - q_1(\lambda_{\text{VBF}} = 1, \lambda_{\text{VH}} = 1, \lambda_{\text{ttH}} = 1) \\ &= -2\ln \frac{L(\text{data}|\lambda^i_{\text{VBF}} = 1, \lambda^i_{\text{VH}} = 1, \lambda^i_{\text{ttH}} = 1, \hat{\mu}_i)}{L(\text{data}|\hat{\lambda}^i_{\text{VBF}}, \hat{\lambda}^i_{\text{VH}}, \hat{\lambda}^i_{\text{ttH}}, \hat{\mu}_i)} \\ &\quad + 2\ln \frac{L(\text{data}|\lambda_{\text{VBF}} = 1, \lambda_{\text{VH}} = 1, \lambda_{\text{ttH}} = 1, \hat{\mu}_i)}{L(\text{data}|\hat{\lambda}_{\text{VBF}}, \hat{\lambda}_{\text{VH}}, \hat{\lambda}_{\text{ttH}}, \hat{\mu}_i)}, \end{aligned} \tag{10.5}$$

that is, as a difference of profile likelihood ratios for both models. The special case where each parameter is set to unity follows from the standard model hypothesis and is chosen only for convenience.

If a gluon fusion measurement $\mu_{\mathrm{ggH},i}$ is missing, as is the case for $\mathrm{H} \to \mathrm{bb}$, the parametrization for the most general 5×4 matrix has more free parameters than there are measurements, resulting in a technical difficulty for the numerical minimization used in calculating the profile likelihood. This can be solved by setting the missing signal strength to a constant in the denominator of the test statistic q_λ (10.4) without loss of generality: the free values of λ^i_{VBF}, λ^i_{VH} and λ^i_{ttH} reabsorb any deviation from the constant value. In the following analysis the signal strength $\mu_{\mathrm{ggH,bb}}$ was set to unity, purely for convenience and inspired by the expectation from the standard model hypothesis. However, any other non-zero value would give the same results.

Setting the missing measurement to a constant in the denominator of Eq. (10.4) must be taken into account in the right way when evaluating the test statistic q_λ as the difference between q_2 and q_1. The definition of q_2 does not change in any other way than having one constant parameter (μ_{bb}), whereas one extra parameter of interest must be added to the test statistic q_1:

$$\begin{aligned}
q_{\lambda,\mathrm{mis.}} &= q_2(\lambda^i_{\mathrm{VBF}}=1,\lambda^i_{\mathrm{VH}}=1,\lambda^i_{\mathrm{ttH}}=1) - q_1(\lambda_{\mathrm{VBF}}=1,\lambda_{\mathrm{VH}}=1,\lambda_{\mathrm{ttH}}=1,\mu_{bb}=1)\\
&= -2\ln\frac{L(\mathrm{data}|\lambda^i_{\mathrm{VBF}}=1,\lambda^i_{\mathrm{VH}}=1,\lambda^i_{\mathrm{ttH}}=1,\mu_{\mathrm{bb}}=1,\mu_i=\hat{\mu}_i)}{L(\mathrm{data}|\hat{\lambda}^i_{\mathrm{VBF}},\hat{\lambda}^i_{\mathrm{VH}},\hat{\lambda}^i_{\mathrm{ttH}},\mu_{bb}=1,\mu_i=\hat{\mu}_i)}\\
&\quad + 2\ln\frac{L(\mathrm{data}|\lambda_{\mathrm{VBF}}=1,\lambda_{\mathrm{VH}}=1,\lambda_{\mathrm{ttH}}=1,\mu_{\mathrm{bb}}=1,\mu_i=\hat{\mu}_i)}{L(\mathrm{data}|\hat{\lambda}_{\mathrm{VBF}},\hat{\lambda}_{\mathrm{VH}},\hat{\lambda}_{\mathrm{ttH}},\hat{\mu}_{bb},\hat{\mu}_i)}\\
&= -2\ln\frac{L(\mathrm{data}|\lambda^i_{\mathrm{VBF}}=\hat{\lambda}_{\mathrm{VBF}},\ \lambda^i_{\mathrm{VH}}=\hat{\lambda}_{\mathrm{VH}},\ \lambda^i_{\mathrm{ttH}}=\hat{\lambda}_{\mathrm{ttH}},\ \hat{\mu}_{\mathrm{bb}},\mu_i=\hat{\mu}_i)}{L(\mathrm{data}|\hat{\lambda}^i_{\mathrm{VBF}},\hat{\lambda}^i_{\mathrm{VH}},\hat{\lambda}^i_{\mathrm{ttH}},\mu_{\mathrm{bb}}=1,\mu_i=\hat{\mu}_i)}.
\end{aligned} \tag{10.6}$$

As a result, the test statistic $q_{\lambda,\mathrm{mis.}}$ has 11 parameters constrained at numerator $(\#\{\lambda^i_j\} - \#\{\lambda_j\} - \#\{\mu_{\mathrm{bb}}\} = 15 - 3 - 1 = 11)$.

The observed value of the profile likelihood ratio for CMS data was evaluated using Eq. (10.6) and it is $q^{\mathrm{obs}}_\lambda \simeq 12.15$. To quantify whether the observed value of the test statistic is small or large, i.e. compatible or not with a single Higgs boson hypothesis, we rely on the p-value of that hypothesis. As mentioned in Section 7, the p-value is defined as the probability to measure as or more extreme value than the observed value of the test statistic, computed from the expected test statistic distribution under the hypothesis of rank 1.

10.2 Generating the pseudo-experiments: the test statistic distribution(s)

Pseudo-experiments were used to estimate the expected distribution of the test statistic. Since the expected test statistic distribution of q_λ can depend on the specific single Higgs boson hypothesis, a general expected test statistic distribution is not well defined. To understand if the shape of the expected test statistic distribution is sensitive on the choice of the specific single Higgs boson hypothesis, pseudo-experiments were generated according to two hypotheses; the standard model Higgs boson hypothesis and the best fit to the data under the hypothesis of a single Higgs boson.

In the standard model Higgs boson hypothesis each μ and λ are set equal to unity, whereas in the the best fit to the data under the hypothesis of a single Higgs boson each μ and λ are set equal to the their best fit values. The best fit values can be seen in Table 10.1. The value of the test statistic was evaluated for each pseudo-experiment, resulting a test statistic distribution for both aforementioned hypotheses.

Parameter	SM Higgs boson	Best fit to the data
$\mu_{\gamma\gamma}$	1.0	0.855
μ_{WW}	1.0	0.797
μ_{ZZ}	1.0	0.886
$\mu_{\tau\tau}$	1.0	0.603
μ_{bb}	1.0	0.347
λ_{VBF}	1.0	1.597
λ_{VH}	1.0	1.775
λ_{ttH}	1.0	4.649

Table 10.1: The values of the parameters in the standard model Higgs boson hypothesis and in the best fit to the data under the hypothesis of a single Higgs boson.

The expected test statistic distributions can be seen in Fig. 10.1. The tails of the histograms, especially on the negative side, are resulting from the few failed pseudo-experiments. In these cases the numerical maximization of the likelihood, used to compute the numerator or denominator of the q_1 (q_2), did not converge properly. However, since the number of failed pseudo-experiments is small, it does not affect the result.

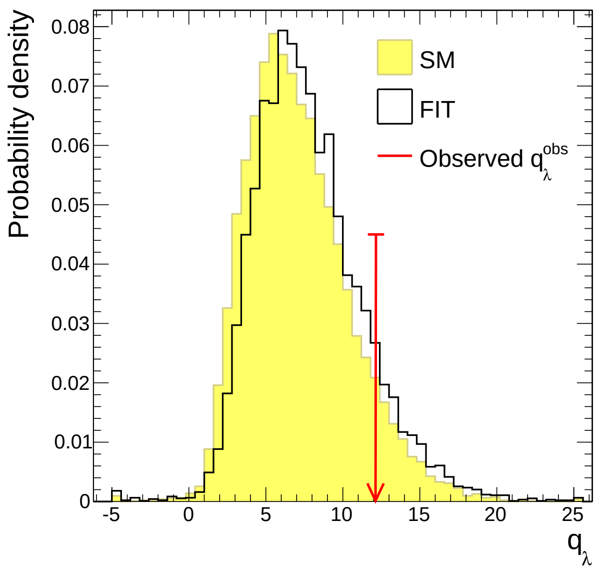

Figure 10.1: The distributions of the test statistic, using CMS data. The distributions according to the standard model Higgs boson hypothesis and the single Higgs boson hypothesis best fitting the data are displayed as filled yellow and outlined black histograms. The observed value of the test statistic q_λ^{obs} is marked with the red arrow.

As introduced in Section 7, according to Wilks' theorem, the expected distributions of a likelihood ratio test statistic with 11 constrained parameters at numerator can be approximated as a χ^2 distribution with 11 degrees of freedom in an ideal situation. Thus, one would expect to see the test statistic distributions to be peaked at $q_\lambda \approx 11$, but clearly this is not the case: both expected test statistic distributions are peaked at smaller value of q_λ. This is due to Wilks' conditions not being sufficiently satisfied: some parameters are not measured at all or they are measured with large uncertainties compared to the distance from the physical boundaries imposed on them ($\mu \geq 0$, $\lambda \geq 0$), and so the statistical fluctuations they can have around their expected value are limited. As a consequence, these parameters each contribute to less than one degree of freedom.

From the comparison of the two test statistic distributions in Fig. 10.1 one can see that the distribution generated assuming the best fit to the data under the hypothesis of a single Higgs boson is shifted to slightly higher values of q_λ than the distribution assuming the standard model Higgs boson hypothesis. This can be explained by the best fit values: the gluon fusion measurements are below the standard model expectations, whereas we observe more than what we expect from the standard model in the vector boson fusion, the Higgs–strahlung and the ttH associated production measurements. In other words, each λ is more far away from the set physical boundaries (μ, $\lambda \geq 0$) than the SM expectation. Thus they are able to fluctuate more and this increases the count of effective degrees of freedom. As a result, the distribution generated assuming the best fit to the data under the hypothesis of a single Higgs boson is shifted to higher values of q_λ.

10.2.1 Connection between the number of parameters in the model(s) and the number of degrees of freedom in the χ^2 distribution

Since in this analysis the test statistic is a likelihood ratio where the hypothesis in the numerator is a restriction of the denominator hypothesis with less parameters, it is expected to be approximately distributed as a χ^2 distribution if the hypotheses of Wilks' theorem were verified. However, as pointed out above, the test statistic appears to have smaller number of constrained parameters at numerator than expected. To examine more in detail why the Wilks' theorem is not perfectly satisfied for the test statistic $q_{\lambda,\text{mis.}}$, we will study the expected distributions of test statistics q_1 and q_2 from Eq. (10.6) separately.

According to the Wilks' theorem the test statistic q_1, defined from a likelihood ratio of fits for the general rank 1 model with 4 parameters constrained at numerator $(\lambda_j, \mu_{\rm bb})$, should correspond to a χ^2 distribution with 4 degrees of freedom, while the test statistic q_2, defined from fits with the most general 5×4 model with 15 parameters fixed at numerator (λ_j^i), should correspond to a χ^2 distribution with 15 degrees of freedom. To test the validity of this approximation, the two test statistic distributions, assuming the standard model Higgs boson hypothesis, have been fit using a χ^2 distribution with floating number of degrees of freedom.

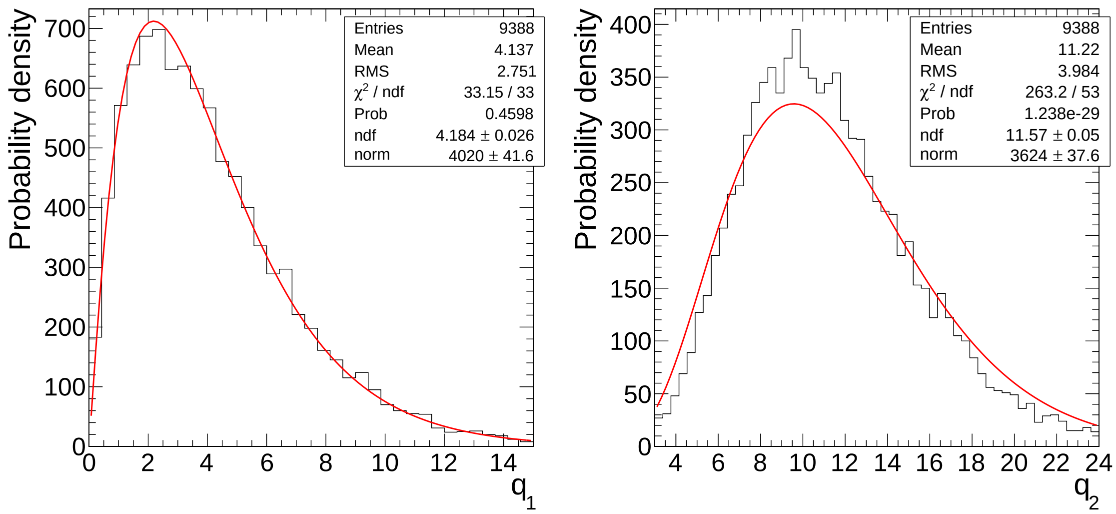

Figure 10.2: Left: The distribution of the test statistic q_1 and the χ^2 fit. Right: The distribution of the test statistic q_2 and the χ^2 fit.

The test statistic distributions and the χ^2 fits can be seen in Fig. 10.2. The fit for the first model is successful ($\chi^2/\text{ndf} \approx 1$) and the number of degrees of freedom is approximately four, as expected. However, the fit for the second model is poor ($\chi^2/\text{ndf} \approx 5$) and the number of degrees of freedom is approximately 12 instead of 15. The lack of the number of degrees of freedom and the bad agreement of the test statistic distribution with a χ^2 distribution are because of the missing or poorly measured elements of the signal strength matrix.

10.3 Computing the *p*-value for the single Higgs boson hypothesis

The *p*-value under the hypothesis of the standard model Higgs boson is obtained from the test statistic distribution, by counting how many of the generated test statistic values are above $q_\lambda^{\rm obs}$:

$$p_{\rm SM} = P(q_\lambda > q_\lambda^{\rm obs} | \mu_{\rm i} \equiv 1, \lambda_{\rm j} \equiv 1) = (8.0 \pm 0.3)\%, \tag{10.7}$$

where the uncertainty quoted is statistical, reflecting the limited number of pseudo-experiments used to determine the distribution. This means that for a standard model Higgs boson we would expect an outcome as extreme or more extreme than the observation in (8.0 ± 0.3)% of the cases. In the case of the second hypothesis, i.e. the best fit to the data under the hypothesis of one Higgs boson, the *p*-value becomes

$$p_{\rm Fit} = P(q_\lambda > q_\lambda^{\rm obs} | \hat{\mu}_{\rm i}^{\rm obs}, \hat{\lambda}_{\rm j}^{\rm obs}) = (11.0 \pm 0.3)\%. \tag{10.8}$$

Since both *p*-values are relatively large compared to the value corresponding to a significance of 3 or 5 standard deviations, 1.3×10^{-3} or 2.9×10^{-7}, the null hypothesis of a single Higgs boson cannot be rejected. That is, the CMS data is compatible with the hypothesis of a single Higgs boson, as in the standard model. From the comparison of the *p*-values obtained from the two different test statistic distributions (Fig. 10.1) one should notice that the *p*-value is approximately insensitive to choice of the single Higgs boson hypothesis used to derive it (the standard model Higgs boson versus best fit to data). This is not unexpected, as the best fit to the data under the single Higgs boson hypothesis is not far from the SM predictions.

10.3.1 An alternative way to compute the *p*-value: the χ^2 fit

As explained in Section 10.2, sometimes the evaluation of the test statistic yields incorrect values, which can result in an incorrect estimation of the tails and thus might affect on the results.

Motivated by the asymptotic behaviour of the test statistic, let us consider an alternative way to compute the p-values, using a χ^2 distribution with a number of degrees of freedom derived from a fit to the histogram of test statistic values obtained from pseudo-experiments. This alternative way is also less sensitive to the tails, since the fit is done for the overall distribution. Confidence in that the results do not depend dramatically on the precise shape of the distribution can be gained if the p-values obtained from the χ^2 fit and the test statistic distribution are similar.

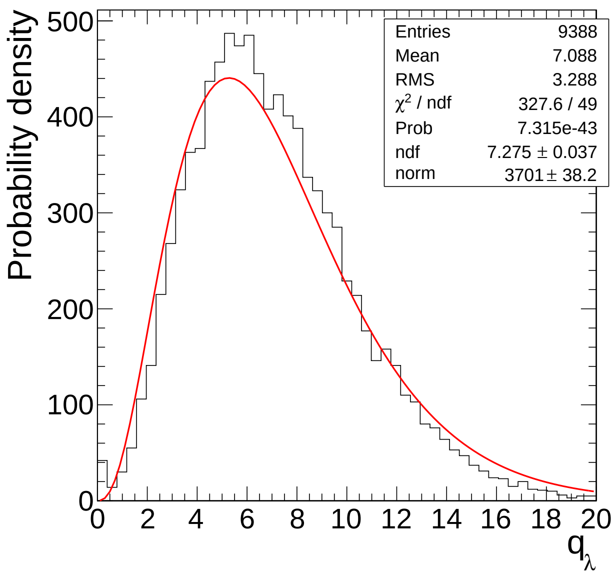

Figure 10.3: The distribution of the test statistic q_λ and the χ^2 fit with 51 bins.

The χ^2 fit was done for the test statistic distribution assuming the standard model Higgs boson hypothesis and the result can be seen in Fig. 10.3. Similarly to the case of the second test statistic q_2, the distribution is not well described by a χ^2

distribution ($\chi^2/\text{ndf} \approx 6.7$). That is, the test statistic distribution does not agree with a χ^2 distribution quantitatively even though these two distributions seems to agree qualitatively.

The probability for the fitted χ^2 distribution to yield a value of q_λ larger than the observed q_λ^{obs} is 0.108, which is close to the p-value of 0.08 obtained from the test statistic distribution. The p-value is insensitive to the binning. Thus, the result does not depend on the precise shape of the test statistic distribution.

11. Results and Conclusions

In 2012, ATLAS and CMS collaborations at the Large Hadron Collider announced the discovery of a new boson consistent with the standard model Higgs boson. However, since theories with an extended Higgs sector are well-motivated due to the imperfections of the standard model, it is not self-evident that the observed signal is from the standard model Higgs boson. One of the possible explanations for the SM-like signal could be multiple nearly mass-degenerate Higgs bosons that couple differently to the standard model particles.

In this thesis, we tested for mass-degenerate Higgs bosons in CMS data from the matrix of observed signal strengths in each production and decay mode using a profile likelihood method. The developed method can be used in the presence of uncertainties, correlations, and unmeasured elements of the matrix. A condensed description of the method has also been published in EPJC [48], and the method has been used in the final Run 1 Higgs combination and properties article [17].

The test statistic was defined as a difference of profile likelihood ratios for two hypotheses, a single Higgs boson and multiple mass-degenerate Higgs bosons hypotheses, that were constructed inspired by the connection between rank and the number of resonances represented in Ref. [18].

The expected test statistic distribution was estimated using pseudo-experiments that were generated according to the standard model Higgs boson hypothesis and the single Higgs boson hypothesis best fitting the data. The p-value for the single Higgs boson hypothesis was defined from the distributions of the test statistic, and it was $(8.0 \pm 0.3)\%$ and $(11.0 \pm 0.3)\%$, respectively. The p-value for the single Higgs boson hypothesis was also computed in an alternative way, by fitting the distribution assuming the standard model Higgs boson hypothesis with a probability density function that generalizes the χ^2 distribution also for non-integer values of the number of degrees of freedom. The resulting p-value was 10.8%.

To summarize, CMS data was found to be compatible with the hypothesis of a single Higgs boson as in the standard model. Moreover, the result is approximately insensitive to choice of the single Higgs boson hypothesis used to derive it, and it does not depend on the precise shape of the test statistic distribution.

As the individual measurements will become more accurate during Run 2 of LHC, this test will naturally become more stringent, and it can be also extended to include other production and decay modes of the Higgs boson once they will be accessible.

Bibliography

1 S. L. Glashow, "Partial-symmetries of weak interactions", *Nuclear Physics* **22** (1961), no. 4, 579 – 588, doi:http://dx.doi.org/10.1016/0029-5582(61)90469-2.

2 S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.

3 A. Salam, "Weak and Electromagnetic Interactions", *Conf.Proc.* **C680519** (1968) 367–377.

4 F. Halzen and A. D. Martin, "Quarks and leptons: an introductory course in modern particle physics". Wiley, New York, NY, 1984.

5 F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323, doi:10.1103/PhysRevLett.13.321.

6 P. Higgs, "Broken symmetries, massless particles and gauge fields", *Physics Letters* **12** (1964), no. 2, 132 – 133, doi:http://dx.doi.org/10.1016/0031-9163(64)91136-9.

7 P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, doi:10.1103/PhysRevLett.13.508.

8 G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587, doi:10.1103/PhysRevLett.13.585.

9 P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons”, *Phys. Rev.* **145** (May, 1966) 1156–1163, doi:10.1103/PhysRev.145.1156.

10 T. W. B. Kibble, “Symmetry Breaking in Non-Abelian Gauge Theories”, *Phys. Rev.* **155** (Mar, 1967) 1554–1561, doi:10.1103/PhysRev.155.1554.

11 J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, “The Higgs hunter’s guide”. Addison-Wesley, Redwood City, cop. 1990.

12 ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration and OPAL Collaboration, “Search for the Standard Model Higgs boson at LEP”, *Physics Letters B* **565** (2003), no. 0, 61 – 75, doi:http://dx.doi.org/10.1016/S0370-2693(03)00614-2.

13 ATLAS Collaboration, “Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

14 CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

15 CMS Collaboration, “Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV”, *J. High Energy Phys.* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

16 Y. Nagashima, “Beyond the standard model of elementary particle physics”. Wiley-VCH, Weinheim, 2014.

17 CMS Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J. C* **75** (2015), no. 5, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

18 Y. Grossman, Z. Surujon, and J. Zupan, "How to test for mass degenerate Higgs resonances", *Journal of High Energy Physics* **2013** (2013), no. 3, doi:10.1007/JHEP03(2013)176.

19 Y. Fukuda et al., "Evidence for Oscillation of Atmospheric Neutrinos", *Phys. Rev. Lett.* **81** (Aug, 1998) 1562–1567, doi:10.1103/PhysRevLett.81.1562.

20 L. Alvarez-Gaume and J. Ellis, "Eyes on a prize particle", *Nature Phys.* **7** (2011), no. 1, 2–3, Editorial Material.

21 Particle Data Group, K. A. Olive et al., "Review of Particle Physics", *Chin. Phys. C* **38** (2014), no. 9, 090001, doi:10.1088/1674-1137/38/9/090001.

22 LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al., "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables". CERN, Geneva, 2011.

23 LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group", Technical Report arXiv:1307.1347. CERN-2013-004, Geneva, (2013).

24 LHC Higgs Cross Section Working Group Collaboration. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGCrossSectionsFigures.

25 S. P. Martin, "A Supersymmetry primer", *Adv.Ser.Direct.High Energy Phys.* **21** (2010) 1–153, doi:10.1142/9789814307505_0001, arXiv:hep-ph/9709356, and references therein.

26 P. Langacker, "Grand unified theories and proton decay", *Physics Reports* **72** (1981), no. 4, 185 – 385, doi:http://dx.doi.org/10.1016/0370-1573(81)90059-4.

27 T. Lee, "A Theory of Spontaneous T Violation", *Phys.Rev.* **D8** (1973) 1226–1239, doi:10.1103/PhysRevD.8.1226.

28 G. Branco et al., "Theory and phenomenology of two-Higgs-doublet models", *Physics Reports* **516** (2012), no. 1–2, 1 – 102, doi:http://dx.doi.org/10.1016/j.physrep.2012.02.002, and references therein.

29 J. F. Gunion, Y. Jiang, and S. Kraml, "Could two NMSSM Higgs bosons be present near 125 GeV?", *Phys. Rev. D* **86** (Oct, 2012) 071702, doi:10.1103/PhysRevD.86.071702.

30 P. M. Ferreira, R. Santos, H. E. Haber, and J. P. Silva, "Mass-degenerate Higgs bosons at 125 GeV in the two-Higgs-doublet model", *Phys. Rev. D* **87** (Mar, 2013) 055009, doi:10.1103/PhysRevD.87.055009.

31 A. Drozd, B. Grzadkowski, J. F. Gunion, and Y. Jiang, "Two-Higgs-doublet models and enhanced rates for a 125 GeV Higgs", *Journal of High Energy Physics* **2013** (2013), no. 5, doi:10.1007/JHEP05(2013)072.

32 J. F. Gunion, Y. Jiang, and S. Kraml, "Diagnosing Degenerate Higgs Bosons at 125 GeV", *Phys. Rev. Lett.* **110** (Jan, 2013) 051801, doi:10.1103/PhysRevLett.110.051801.

33 L. R. Evans and P. Bryant, "LHC Machine", *J. Instrum.* **3** (2008) S08001. 164 p.

34 M. Lamont, "Status of the LHC", *J.Phys.Conf.Ser.* **455** (2013) 012001, doi:10.1088/1742-6596/455/1/012001.

35 CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

36 T. Sakuma and T. McCauley, "Detector and Event Visualization with SketchUp at the CMS Experiment", *J.Phys.Conf.Ser.* **513** (2014) 022032, doi:10.1088/1742-6596/513/2/022032, arXiv:1311.4942.

37 M. Hoch and M. Brice, "CMS experiment", (Aug, 2008). CMS Collection.

38 CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, (Apr, 2009).

39 D. Trocino, "The CMS High Level Trigger", *Journal of Physics: Conference Series* **513** (2014), no. 1, 012036, doi:10.1088/1742-6596/513/1/012036.

40 CMS Collaboration, S. Cittolin, A. Rácz, and P. Sphicas, "CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project". Technical Design Report CMS. CERN, Geneva, 2002.

41 G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011), no. 2, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727.

42 G. Cowan, "Statistical data analysis". Oxford Univ. Press, Oxford, 1998.

43 J. Neyman and E. S. Pearson, "On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference: Part I", *Biometrika* **20A** (1928), no. 1-2, 175–240, doi:10.1093/biomet/20A.1-2.175.

44 J. Neyman and E. S. Pearson, "On the Use and Interpretation of Certain Test

Criteria for Purposes of Statistical Inference: Part II", *Biometrika* **20A** (1928), no. 3-4, 263–294, doi:10.1093/biomet/20A.3-4.263.

45 A. Stuart, J. K. Ord, and S. Arnold, "Kendall's advanced theory of statistics. Vol. 2A, Classical inference and the linear model". Arnold, London, 6th edition, 2004.

46 S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", *The Annals of Mathematical Statistics* **9** (03, 1938) 60–62, doi:10.1214/aoms/1177732360.

47 C. J. Clopper and E. S. Pearson, "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial", *Biometrika* **26** (1934), no. 4, 404–413, doi:10.1093/biomet/26.4.404.

48 A. David, J. Heikkilä, and G. Petrucciani, "Searching for degenerate Higgs bosons", *Eur.Phys.J.C* **75** (2015), no. 2, 49, doi:10.1140/epjc/s10052-015-3279-y, arXiv:1409.6132.