Study of Higgs boson production through its decay to two photons using data collected at a centre-of-mass energy of 13 TeV with the CMS detector

Louie Dartmoor Corpe
Imperial College London

A dissertation submitted to Imperial College London
for the degree of Doctor of Philosophy

Abstract

A study of $\mathrm{H} \to \gamma\gamma$ with $12.9\,\mathrm{fb}^{-1}$ of data from p-p collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ is presented. The data were collected at the start of LHC Run 2 by the CMS experiment between April and July 2016. The result is an new standalone observation of the Higgs boson in the diphoton decay channel. The significance of the excess is observed to be 5.7σ (6.3σ expected) at the combined Run 1 best-fit $m_{\mathrm{H}} = 125.09\,\mathrm{GeV}$. The maximum significance of 6.1σ (6.3σ expected) is observed at $m_{\mathrm{H}} = 125.9\,\mathrm{GeV}$. The measured Higgs boson signal strength is $\hat{\mu} = 0.94^{+0.21}_{-0.18} = 0.94 \pm 0.16\,(\mathrm{stat.})\,^{+0.10}_{-0.07}\,(\mathrm{exp.\ syst.})\,^{+0.08}_{-0.05}\,(\mathrm{theo.\ syst.})$. The signal strength is measured separately for: fermionic and bosonic production modes, giving $\hat{\mu}_{\mathrm{VBF,VH}} = 1.52^{+0.89}_{-0.77}$ and $\hat{\mu}_{\mathrm{ggH,ttH}} = 0.81^{+0.27}_{-0.25}$; and for the ggH, ttH and VBF production modes, giving $\hat{\mu}_{\mathrm{ggH}} = 0.78^{+0.25}_{-0.23}$, $\hat{\mu}_{\mathrm{ttH}} = 1.54^{+0.9}_{-0.8}$, and $\hat{\mu}_{\mathrm{VBF}} = 1.86^{+1.5}_{-1.2}$. Measurements of the Higgs boson coupling modifiers yield $\hat{\kappa}_{\mathrm{f}} = 0.68^{+0.45}_{-0.22}$, $\hat{\kappa}_{\mathrm{V}} = 0.93^{+0.11}_{-0.10}$ for the fermionic and bosonic coupling modifiers and $\hat{\kappa}_g = 0.66^{+0.27}_{-0.19}$, $\hat{\kappa}_g = 0.66^{+0.27}_{-0.19}$ for the effective coupling modifiers. All measurements are found to be consistent with the theoretical expectation for a SM-like Higgs boson.

Declaration

The results presented in this thesis are the culmination of work which I certify to be my own, although it is based in part on studies by others. The theory described in Chapter 2 has been summarised by me, but is the work of others. The LHC and CMS designs detailed in Chapter 3 are the work of others, although I was involved in the Run 1 photon energy resolution measurements (see Section 3.2.3) and Run 2 ECAL calibration (see Section 3.2.3). Due to the complexity of the analysis, Chapters 4 and 5 are the work of various individuals in the CMS collaboration and $H \to \gamma\gamma$ group, of which I am a member. Some of the steps are therefore the work of my collaborators.

The bulk of my personal work is presented in Chapters 6 and 7, which detail the signal and background modelling, handling of systematic uncertainties, statistical interpretation of the data and production of results. Indeed, these were my personal areas of responsibility in the CMS $H \to \gamma\gamma$ group. I was responsible for these topics for two studies of $H \to \gamma\gamma$ with Run 2 data [1, 2], the latter being the basis for this thesis. These chapters also include significant developments to the parametric signal modelling techniques with were conceived and implemented by me, and which I used to produce the results, but which do not feature in [2].

Where ideas or figures from others are used, appropriate sources are referenced. Figures labelled "**CMS** *Simulation*", "**CMS** *Preliminary*" or "**CMS** *Simulation Preliminary*" are either taken from CMS publications or preliminary public documents, including those produced by me, and are referenced appropriately; or result from the analysis I have performed in this thesis, in which case they have been approved by the CMS collaboration. Figures additionally labelled "$H \to \gamma\gamma$" result from the work of members of the CMS Higgs to diphoton analysis group specifically.

Louie Dartmoor Corpe

Acknowledgements

To be written!
- family
- supervisors
- friends
- IC HEP, CMS, Hgg group
- STFC
- Emma

Contents

List of figures

List of tables

"There are more things in heaven and earth, Horatio,
Than are dreamt of in your philosophy"

— Hamlet (William Shakespeare)

Chapter 1

Introduction

The discovery of the Higgs boson in 2012 [3, 4] cemented the standard model (SM) of particle physics as one of the most successful theories of modern science in terms of its predictive power and agreement with experimental observations. The SM describes the universe in terms of matter particles and fundamental forces, which are mediated by carriers. The great success of the SM was to place all forces (except Gravity) into one framework, and to use the Higgs mechanism to explain the manifest breaking of symmetry between the Electromagnetic and Weak forces. An overview of the theory and its predictions for the properties of the Higgs boson are given in Chapter 2. In particular, this chapter details some of the reasons why the $\mathrm{H} \to \gamma\gamma$ decay is an excellent channel with which to study the Higgs boson.

The LHC, as well as the two multi-purpose detectors CMS and ATLAS, were designed with the discovery of the Higgs boson as one of their primary objectives. In the case of CMS, a key feature of the design is the electromagnetic calorimeter (ECAL), which was conceived with the study of $\mathrm{H} \to \gamma\gamma$ in mind. The LHC and the CMS detector are described in Chapter 3. In 2014, CMS produced a standalone observation of the Higgs boson decaying to photons, using $24.8\,\mathrm{fb}^{-1}$ of data collected at 7 and 8 TeV [5]. In 2013, the LHC began a two-year shutdown period, during which key upgrades to the accelerator and detectors were implemented. The hadron collider has now started up again, colliding particles at 13 TeV. The dataset analysed in this thesis contains $12.9\,\mathrm{fb}^{-1}$ collected in this regime. The rates of the Higgs boson production processes are expected to evolve with the collision energy. For this reason, it so happens that the analysed dataset has a similar statistical sensitivity to the entire 7 and 8 TeV dataset. An important first task for the CMS collaboration is therefore to confirm the existence of the Higgs boson in the 13 TeV data, and to check if the Higgs boson's properties evolve as expected.

The analysis described in this thesis follows closely the methods used in previous studies of the $H \to \gamma\gamma$ decay at CMS [1, 2, 5]. This involves the selection and reconstruction of data containing two photons which are $H \to \gamma\gamma$ candidates. This procedure is detailed in Chapter 4. The selected data are then categorised to optimise the overall sensitivity of the analysis on one hand, and to target specific Higgs boson production processes on the other hand. This scheme is described in Chapter 5.

The categorised data and simulation samples are used to produce models of how the $H \to \gamma\gamma$ signal and background are expected to be manifested in the $m_{\gamma\gamma}$ spectrum. This task, which includes modelling the sources of systematic uncertainty which enter the analysis, were the specific area of responsibility of the author. Chapter 6 describes the procedure by which these models are produced. It also contains a description of two new techniques which were developed by the author.

Finally, the results of the analysis are detailed in Chapter 7. They entail a new observation of the Higgs boson, and confirm that the measured properties of the new particle agree with the SM expectation for 13 TeV within uncertainties, which are largely dominated by the statistical component. In Chapter 8, a discussion on the conclusions which can be drawn from these results is provided, alongside a view to the future of Higgs boson physics.

Chapter 2

Theory

2.1 The Standard Model of particle physics

2.1.1 Introduction

The SM of particle physics was developed in the early 1970s. It has been immensely successful, accurately describing all processes so far encountered in high energy physics [6]. In the SM, the fundamental elements of matter, as well as the forces which govern their interaction, are represented as relativistic quantum fields, the excitations of which are manifested as particles (see Section 2.1.2). The SM places all the forces except for gravity in the same framework, and unites the electromagnetic and weak forces as the electroweak force (see Section 2.1.3). The Brout-Englert-Higgs mechanism explains the breaking of the underlying symmetry between these forces and leads to the prediction of an observable particle, the Higgs boson (see Section 2.1.4).

2.1.2 Particles and Forces

In the SM, matter is made up of spin-1/2 particles, called *fermions*. Fermions come in two types: those which interact exclusively via the electroweak force, the *leptons*, and those which can also interact via the strong nuclear force, the *quarks*. The fermions can be arranged into three *generations*. The SM fermions and their properties are displayed in Table 2.1. Each of the particles mentioned in the table has a corresponding *antiparticle*, with the same mass and opposite charge.

type	Generation I		Generation II		Generation III	
leptons	e	$m = 0.511\,\text{MeV}$	μ	$m = 105\,\text{MeV}$	τ	$m = 1.777\,\text{GeV}$
	electron	$q = -1$	muon	$q = -1$	tau	$q = -1$
	ν_e	$m \sim 0\,\text{MeV}$	ν_μ	$m \sim 0\,\text{MeV}$	ν_τ	$m \sim 0\,\text{MeV}$
	electron neutrino	$q = 0$	muon neutrino	$q = 0$	tau neutrino	$q = 0$
quarks	u	$m = 2.3\,\text{MeV}$	c	$m = 1.275\,\text{GeV}$	t	$m = 173\,\text{GeV}$
	up	$q = +\frac{2}{3}$	charm	$q = +\frac{2}{3}$	top or truth	$q = +\frac{2}{3}$
	d	$m = 4.8\,\text{MeV}$	s	$m = 95\,\text{MeV}$	b	$m = 4.18\,\text{GeV}$
	down	$q = -\frac{1}{3}$	strange	$q = -\frac{1}{3}$	bottom or beauty	$q = -\frac{1}{3}$

Table 2.1: The fundamental particles of the SM. The mass m and electric charge q are indicated for each particle [6]. The uncertainties on the masses have been omitted, although some are large.

The matter particles interact via the fundamental forces. There are four known fundamental forces in the universe: the electromagnetic force, the weak nuclear force, the strong nuclear force and the gravitational force. The gravitational force is many orders of magnitude weaker than any of the other forces, and therefore has a negligible effect on the interactions of the SM particles in high energy physics experiments. Furthermore, no adequate quantum theory of gravity currently exists, so it cannot be easily included in the SM. Consequently, the SM describes only with the strong, weak and electromagnetic forces. In the SM, the fundamental forces, are represented by the exchange of spin-1 *mediator particles*, the *vector bosons*. The forces described by the SM are listed in Table 2.2.

Force	Mediator	Mass
strong	gluons $\text{g} \in \{\text{g}_1, ..., \text{g}_8\}$	0
electromagnetic	photon γ	0
weak	W-bosons $\text{W}^\pm$	80.4 GeV
	Z-boson Z	91.2 GeV

Table 2.2: The three fundamental forces described by the SM. The mediator particle of each force is indicated along with its measured mass, where the uncertainties have been omitted [6, 7].

The γ and g are massless, in contrast to the $\text{W}^\pm$ and Z, which are massive. This difference is explained by the process of *electroweak symmetry breaking* via the Brout-Englert-Higgs mechanism [8–13] described in Section 2.1.4. This mechanism introduces an additional scalar field, which implies the existence of a massive spin-0 particle, the Higgs boson.

2.1.3 Gauge groups of the SM Lagrangian

The SM is a quantum field theory (QFT), in particular a renormalisable gauge theory. The Lagrangian $\mathcal{L}_{\text{QFT}}$ of a QFT describes the dynamics and interactions of its particles. It is constructed by considering the nature of the particles involved in the QFT and imposing the symmetries of the theory. Nöther's Theorem [14] states that for every symmetry in a Lagrangian, there is an associated conservation law. For example, if a theory respects conservation of energy or momentum, its Lagrangian must be invarariant in time or space respectively. Imposing a symmetry in a Lagrangian places requirements on how the particles in the theory are allowed to propagate and interact.

A gauge theory is a particular type of QFT where local gauge transformations are a symmetry of the Lagrangian, leading to conservation of charge. Such gauge symmetries are of principal importance in particle physics, as they lead to the introduction of gauge fields, which generate the mediator particles. For this reason, the mediators of the forces are sometimes referred to as *gauge bosons*. Typically, a gauge transformation takes the form of shifting the phase of all wavefunctions. It is reasonable to require such transformations to leave the dynamics of the theory intact, since the phase of a wavefunction is never manifest in any physical observable. Thus the Lagrangian of any realistic theory should be *gauge invariant*.

Quantum Electrodynamics

A simple example of a gauge theory is quantum electrodynamics (QED). This theory must incorporate fermions, which are described by the Dirac Lagrangian [15]:

$$\mathcal{L}_{\text{fermion}} = i\overline{\psi}\gamma^{\alpha}\partial_{\alpha}\psi - m\overline{\psi}\psi, \tag{2.1}$$

where ψ is a Dirac spinor and $\overline{\psi}$ is its adjoint, m is the mass of the fermion, γ^{α} represents the four Dirac gamma matrices and ∂_{α} is the 4-gradient.

Requiring *local gauge invariance* means that the Lagrangian should be invariant under transformations such as $\psi \rightarrow \psi' = \psi e^{i\theta(x^{\mu})}$, where $\theta(x^{\mu})$ is an arbitrary differentiable function of space-time x^{μ}. Applying the transformation to Equation 2.1 gives:

$$\mathcal{L}_{\text{fermion}} \rightarrow \mathcal{L}'_{\text{fermion}} = \mathcal{L}_{\text{fermion}} - \overline{\psi}\gamma^{\alpha}\psi(\partial_{\alpha}\theta(x^{\mu})). \tag{2.2}$$

Evidently $\mathcal{L}_{\text{fermion}}$ is not gauge invariant. This is remedied by introducing an additional field A_α. An extra term, $-g_{\text{EM}}\overline{\psi}\gamma^\alpha\psi A_\alpha$, is added to the Lagrangian to account for the interaction of the fermion with A_α, where g_{EM} is the strength of the interaction. Local gauge invariance is restored so long as A_α changes in the following way [15]:

$$A_\alpha \to A'_\alpha = A_\alpha - \frac{1}{g_{\text{EM}}}\partial_\alpha\theta(x^\mu), \tag{2.3}$$

The Lagrangian can also accommodate a term for A_α propagating freely through space. Since A_α is a 4-vector, it is described by the Proca equation for spin-1 bosons [15]:

$$\mathcal{L}_{\text{boson}} = -\frac{1}{16\pi}F^{\alpha\beta}F_{\alpha\beta} + \frac{1}{8\pi}m_{\text{boson}}A^\alpha A_\alpha, \tag{2.4}$$

where $F^{\alpha\beta} = (\partial^\alpha A^\beta - \partial^\beta A^\alpha)$ and m_{boson} is the mass of the spin-1 boson. Equation 2.4 is locally gauge invariant so long as $m_{\text{boson}} = 0$, i.e. the boson is required to be massless.

It is convenient to define the *covariant derivative*, incorporating the interaction term:

$$D_\alpha = \partial_\alpha + ig_{\text{EM}}A_\alpha. \tag{2.5}$$

The Lagrangian $\mathcal{L}_{\text{QED}}$ can then be written compactly as:

$$\mathcal{L}_{\text{QED}} = i\overline{\psi}\gamma^\alpha D_\alpha\psi - m\overline{\psi}\psi - \frac{1}{16\pi}F^{\alpha\beta}F_{\alpha\beta}. \tag{2.6}$$

The full Lagrangian for QED describes a free fermion, a free massless spin-1 boson and a completely determined term for interactions between the boson and the fermion. The boson is identified as the photon. The factor multiplying g_{EM} is interpreted as the electric charge of the fermion.

The local gauge transformation is equivalent to applying a unitary 1×1 matrix to the wavefunction. The group of all such transformations is $U(1)$. It is a general result that the number of degrees of freedom in the underlying symmetry group dictates the number of additional bosons needed to keep the theory locally gauge invariant [15].

Quantum Chromodynamics

The strategy described in Section 2.1.3 can be used to derive the Lagrangian for quantum chromodynamics (QCD), which codifies the dynamics of the strong force [15]. The situation is more complicated because the underlying group is not $U(1)$ but $SU(3)$, which has eight degrees of freedom. This leads to eight massless gluons. In QED, the interaction between the boson and the fermion is dictated by electric charge; the analogue for QCD is *colour charge*. However, colour charge cannot be represented by a single quantity. Colours charges are instead linear combinations of three quantities, designated *red, green* and *blue*. Each SM quark has three identical copies with different colour charge. Another important distinction is that $U(1)$ is abelian while $SU(3)$ is not. The consequence of this is that unlike the QED photon, which carries no electric charge, the QCD gluons do carry colour charge. Gluons can therefore feel the strong force and self-interact. This fact leads QCD to display properties such as confinement of quarks and asymptotic freedom [16, 17]. The gauge group for QCD, $SU(3)_{\mathrm{C}}$, is generally written with a subscript to indicate that it generates colour charge.

Electroweak unification

The development of a combined theory of the electromagnetic and weak forces, called electroweak theory (EWT), was a major achievement by Glashow, Weinberg and Salam [18–20]. EWT considers the gauge group $SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}}$. The $SU(2)_{\mathrm{L}}$ group has three degrees of freedom, so three gauge bosons are obtained from imposing local gauge invariance. The conserved charge is *weak isospin*. This quantity is a vector, the third component of which is labelled i_3. The subscript in $SU(2)_{\mathrm{L}}$ indicates that only left-handed particles carry nonzero weak isospin charge. The $U(1)_{\mathrm{Y}}$ group behaves as in QED, but generates *weak hypercharge* y. The electric charge q is related to weak isospin and weak hypercharge by the relation $q = y/2 + i_3$. The subscript in $U(1)_{\mathrm{Y}}$ refers to the fact that hypercharge is the generated charge.

Right- and left-handed fermions are considered separately in EWT. It is helpful to think of the right-handed fermion spinors as *singlets*, or column vectors of one spinor. For example, the right-handed electron singlet is labelled e_R. The left-handed fermions come in *doublets* (a column vector of two fermion spinors), e.g. :

$$L_{\mathrm{L}} = \begin{pmatrix} \nu_{\mathrm{e}} \\ \mathrm{e} \end{pmatrix}_{\mathrm{L}},$$

which is the left-handed lepton doublet containing the left-handed electron and electron neutrino. No right-handed neutrinos are included in this scheme. This is a deliberate feature which reflects the fact that no right-handed neutrinos have been observed or detected experimentally. However, the discovery of neutrino oscillations implies the existence of such neutrinos, so clearly the SM is only an approximate theory.

Imposing gauge invariance leads to the introduction of gauge fields: $W^1_\alpha, W^2_\alpha, W^3_\alpha$ and B_α, for $SU(2)_\mathrm{L}$ and $U(1)_\mathrm{Y}$ respectively. The physical states observed in nature are mixtures of the underlying weak isospin and weak hypercharge gauge bosons:

$$\mathrm{W}^{\pm}{}_\alpha = \sqrt{\frac{1}{2}}(W^1_\alpha \mp W^2_\alpha), \tag{2.7}$$

$$\mathrm{Z}_\alpha = \cos\theta_\mathrm{W} W^3_\alpha - \sin\theta_\mathrm{W} B_\alpha, \tag{2.8}$$

$$A_\alpha = \sin\theta_\mathrm{W} W^3_\alpha + \cos\theta_\mathrm{W} B_\alpha, \tag{2.9}$$

where θ_W is the Weinberg angle, which relates strengths of the electromagnetic (g_EM) and weak (g_W) forces via the relation $\tan\theta_\mathrm{W} = g_\mathrm{W}/g_\mathrm{EM}$. The value of $\sin^2\theta_\mathrm{W}$ has been determined experimentally to be ~ 0.23 [6].

Quarks are also accommodated in this framework. For example, for the first generation of quarks, two right-handed quark singlets u_R and d_R and one left-handed quark doublet Q_L (containing u_L and d_L) are introduced. The full SM, incorporating the electroweak and strong forces, is described by the gauge group $SU(3)_\mathrm{C} \times SU(2)_\mathrm{L} \times U(1)_\mathrm{Y}$.

An issue arises when considering masses of particles. The left- and right-handed components of the fermions transform as doublets and singlets respectively, so the usual fermion mass term is no longer gauge invariant. Furthermore, the gauge bosons should be massless to preserve the gauge symmetry. This is the case for the photon and gluons, but the $\mathrm{W}^{\pm}$ and Z need to have masses to explain weak decays. These masses were later experimentally measured to be of the order of 90 GeV [6]. A mechanism is needed to account for the masses of these particles.

2.1.4 Electroweak Symmetry Breaking and the Higgs Mechanism

The process which allows the $W^{\pm}$ and Z to acquire a mass is *electroweak symmetry breaking*. This occurs in the SM via the Brout-Englert-Higgs mechanism [8–13]. In this scheme, an additional complex scalar $SU(2)_L$ doublet ϕ is introduced. The Lagrangian for ϕ is gauge invariant but the ground state is not:

$$\begin{aligned}\mathcal{L}_\phi =& \frac{1}{2}(D^{\mathrm{EWK}}_\alpha \phi)^*(D^\alpha_{\mathrm{EWK}}\phi) \text{ (kinetic term)} \\ &+ \mu^2(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2 \text{ (potential term) ,}\end{aligned} \tag{2.10}$$

where μ and λ are constants. The covariant derivative D^{EWK}_α, defined analogously to Equation 2.5, has been used here to ensure gauge invariance. In this case, D^{EWK}_α accounts for the interactions of the electroweak gauge bosons W^1_α, W^2_α, W^3_α and B_α:

$$D^{\mathrm{EWK}}_\alpha = \partial_\alpha - ig_1\frac{Y}{2}B_\alpha + ig_2\frac{\tau^i}{2}W^i_\alpha, \tag{2.11}$$

where g_Y and g_L refer to the coupling constants of the weak hypercharge and weak isospin fields respectively, Y is the constant generator of the $U(1)_Y$ group and τ^i are the generators of the $SU(2)_L$ group. The coupling constants g_Y and g_L are related to the weak and electromagnetic coupling constants by $g_W = g_L$ and $g_{EM} = g_Y \cos\theta_W$.

The first term of the Lagrangian corresponds to the kinetic part, while the second and third correspond to a potential. The term $\frac{1}{2}\mu^2(\phi^*\phi)$ resembles a mass term, but it is not: the sign needs to be negative. If $\mu^2 < 0$, the potential has a non-zero vacuum expectation value (VEV), and the ground state is represented by a circle of minima. In order to re-write the Lagrangian in terms of physical particles, an expansion around one of the minima is required. Since they are all equivalent, an arbitrary minimum is chosen:

$$\phi_0 = \mathrm{VEV} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{2.12}$$

where $v = \sqrt{-\mu^2/\lambda}$. At this stage, any of the minima lying upon the circle of ground states could have been chosen, but a particular choice had to be made. This step breaks the manifest symmetry in the physical states while preserving it in the Lagrangian.

To obtain the physical states from this Lagrangian, a small perturbation field H around the VEV, is introduced. Any perturbation to the first component of the VEV would represent a move to an equivalent minimum. Therefore such a perturbation can be ignored, and H only acts on the second component:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{2.13}$$

which can be substituted back into Equation 2.1.4. Expanding out the interactions with the gauge bosons from the covariant derivative, and expressing them in terms of the physical bosons using Equation 2.7, the equation becomes:

$$\begin{aligned} \mathcal{L}_\phi = \frac{1}{2}(\partial_\alpha H)(\partial^\alpha H) - \mu^2 H^2 \\ + \frac{v^2}{8}(g_{\mathrm{W}}\mathrm{W}^+{}_\alpha \mathrm{W}^{+\alpha} + g_{\mathrm{W}}\mathrm{W}^-{}_\alpha \mathrm{W}^{-\alpha} + (g_{\mathrm{W}}^2 + g_{\mathrm{EM}}^2)\mathrm{Z}_\alpha \mathrm{Z}^\alpha) + ... \end{aligned}, \tag{2.14}$$

where terms which are not mass-like have been omitted. The first two terms can be identified as the Klein-Gordon equation for a massive scalar boson of mass $\sqrt{2}\mu$: this is the Higgs boson. The other terms represent the great success of the Brout-Englert-Higgs mechanism: the $\mathrm{W}^\pm$ and Z bosons have acquired masses. The absence of equivalent terms for the photon means that it does not acquire mass.

It is possible to add additional gauge invariant terms involving the scalar field. These are known as the Yukawa interactions. For example, in the case of the first generation leptons:

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= \kappa_e(\overline{L}\phi e_R + \overline{e_R}\phi^\dagger L) \\ &= \kappa_e v(\overline{e_L}e_R + \overline{e_R}e_L) + \kappa_e(\overline{e_L}He_R + \overline{e_R}He_L), \end{aligned} \tag{2.15}$$

where κ_e is a real constant. The first term gives a mass to the electron and the second represents the interaction between the electron and the Higgs boson. In this way, the

strength of the interaction is directly proportional to the mass of the particle which is considered. The neutrino does not acquire any mass via this mechanism. The scheme described above does not *prescribe* the masses of the Higgs boson or fermions: these are free parameters and must be specified from experimental measurements.

To summarise, the Brout-Englert-Higgs mechanism adds gauge-invariant terms to the SM Lagrangian which permit the $W^{\pm}$ and Z bosons and the fermions to acquire masses while leaving the photon massless. The outcome is one additional spin-0 particle, the Higgs boson, the mass of which is not directly predicted by the theory.

2.2 Higgs boson phenomenology

2.2.1 History of Higgs boson searches

Since the Higgs boson was postulated in the 1960s, there have been many efforts to try to detect it. Theoretical considerations precluded large Higgs boson masses above the order of 1 TeV [21]. A Higgs boson of mass below about 10 GeV was already excluded before the start of LEP [22]. Direct searches at LEP and the Tevatron excluded a Higgs boson mass below 114 GeV [23, 24], and precision measurements of the electroweak parameters suggested that the Higgs boson mass should be below 200 GeV [25]. The search for the Higgs boson was one of the goals prompting the construction of the LHC at CERN. Two multi-purpose detectors, ATLAS and CMS, were designed with the Higgs observation as one of their main physics goals. In 2012, the two experiments jointly announced the observation of a Higgs-like particle, ending a 50-year interval between postulation and discovery [3, 4]. The most precise measurement of the Higgs boson's mass to date is $m_{\mathrm{H}} = 125.09 \pm 0.24\,\mathrm{GeV}$ [26].

2.2.2 Higgs boson production at the LHC

According to the SM, the Higgs boson interacts with particles proportional to their masses. Four types of process can lead to the production of a Higgs boson in p-p collisions at the LHC. The Feynman diagrams for these processes can be seen in Figure 2.1. The most likely production mode for $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ is gluon-gluon fusion (ggH) via a loop of top quarks, which has a cross-section of approximately 49 pb at 13 TeV. The other production modes are VBF at 3.8 pb, vector boson associated production (VH) at 2.3 pb,

and ttH at 0.5 pb [27]. The VH production can be further split into W boson associated production (WH) (1.4 pb) and Z boson associated production (ZH) (0.6 pb), which have different final states.

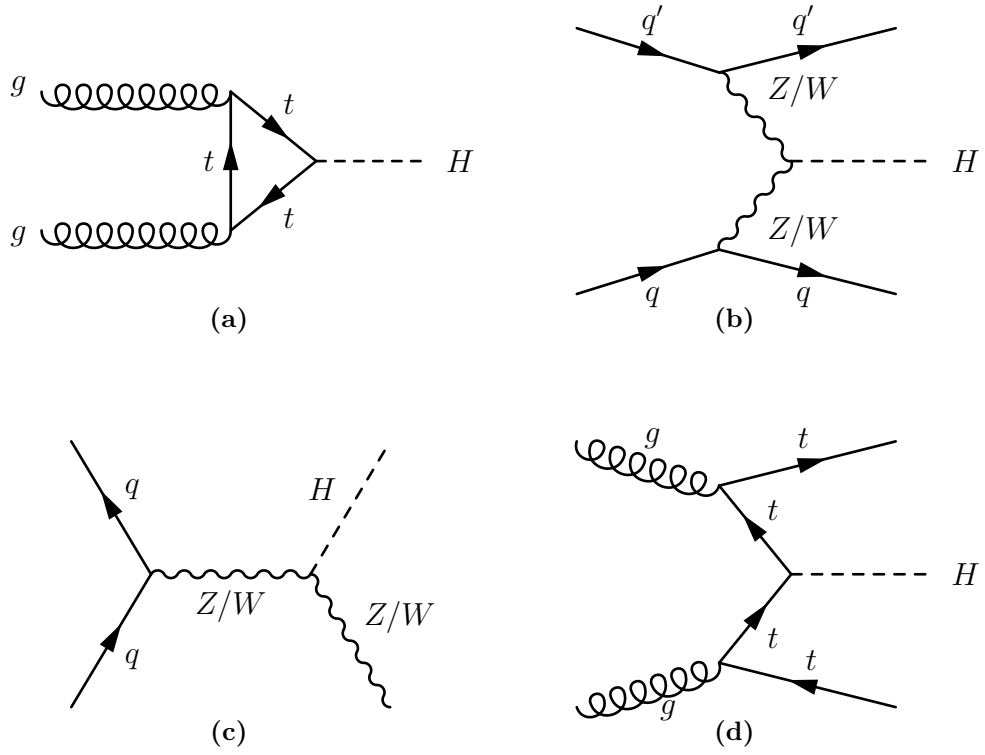


Figure 2.1: Higgs production modes at the LHC: (a) gluon-gluon fusion, via a loop of top quarks, (b) vector boson fusion, with associated quark production, (c) associated vector boson production with either the Z or W boson and (d) top quark fusion with associated top quark production.

2.2.3 Higgs boson decays

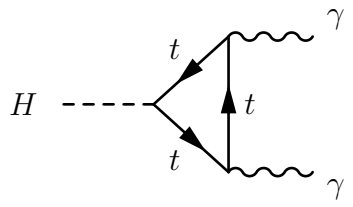
The SM Higgs boson can decay either directly to pairs of particles, or via virtual loops. In direct decays to pairs of particles, the branching ratios are proportional to the mass of the decay product for fermions and the square root of the mass of the decay product for vector boson. The most likely direct decay modes to massive particles for a SM Higgs boson with $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ are $\mathrm{H} \to \mathrm{bb}$ (58.2%); $\mathrm{H} \to \mathrm{WW}^*$ (21.4%, where W^* refers to a virtual W); $\mathrm{H} \to \tau\tau$ (6.3%); $\mathrm{H} \to \mathrm{cc}$ (2.8%); $\mathrm{H} \to \mathrm{ZZ}^*$ (2.6%, where Z^* refers to a virtual Z) [27]. The production of a pair of t quarks is strongly suppressed by kinematics. The Higgs boson's couplings to electrons, muons, up quarks, down quarks and strange

quarks are very small due to the mass of the decay products. In the SM, the neutrinos do not have mass, so no coupling to the Higgs boson is predicted. In addition, Higgs boson decays can occur via a loop of virtual massive particles to a pair of gluons (8.2%), to a pair of photons (0.23%) or to Zγ (0.02%) [27].

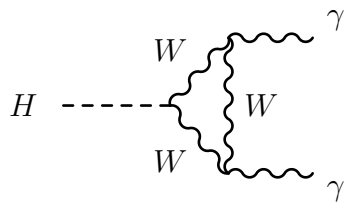
For the CMS and ATLAS detectors, two channels are particularly suited to studies of the Higgs boson despite their relatively low branching fractions. These are $\mathrm{H} \to \gamma\gamma$ and $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ (where ℓ refers to leptons, and the rate is reduced because the branching fraction of $\mathrm{Z} \to \ell\ell$ needs to be taken into account). The fact that both detectors are able to reconstruct the energy an transverse momenta of electrons, muons and photons mean that a narrow Higgs boson mass peak can be reconstructed in these two channels. The other channels which contributed, to a lesser extent, to the discovery of the Higgs boson were $\mathrm{H} \to \mathrm{bb}$, $\mathrm{H} \to \mathrm{WW}^*$ and $\mathrm{H} \to \tau\tau$. These other channels have a higher rate but difficulties in reconstructing the decay products or excessive noise from the LHC p-p collisions drastically reduce the experimental sensitivity.

2.2.4 Studying the Higgs boson using the $\mathrm{H} \to \gamma\gamma$ decay

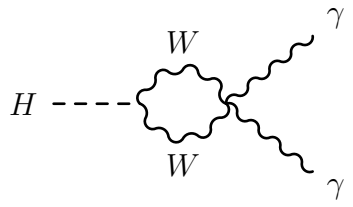
The work presented in this thesis focusses on the observation of the Higgs boson and a measurement of its properties via the $\mathrm{H} \to \gamma\gamma$ decay, for which the leading order Feynman diagrams are shown in Figure 2.2. The data collected by the CMS detector are selected for such measurements if they contain two photons which are candidates to have originated from a Higgs boson decay. The analysis of these data involves defining categories which target the individual production modes of the Higgs boson presented in Section 2.2.2. Indeed, the final states of each production mode lead to different signatures in the detector, which can be exploited for the purposes of categorisation. This approach not only increases the overall sensitivity of the analysis, but also gives a handle with which to probe the SM predictions for the rate at which the Higgs boson is produced in each mode, as well as the strength of its interaction with SM particles. Many extensions to the SM predict variations in these properties, so such measurements could put limits on new models or could yield clues to the nature of physics beyond the SM.



(a)



(b)



(c)

Figure 2.2: A Higgs boson decaying to photons via a loop of top quarks (a) or via loops of W bosons (b, c).

Chapter 3

Overview of the LHC and CMS

3.1 The Large Hadron Collider (LHC)

The LHC [28] is currently the largest and most powerful synchrotron in the world, and is installed in the tunnel which previously contained the LEP [29] collider at CERN. The tunnel is located roughly 100 m underground near Geneva, on the border between Switzerland and France. The LHC was designed to perform collisions of two types: p-p (proton-proton) collisions and, less frequently, heavy ion collisions, for example Pb-Pb (lead-lead). The former is used to search for new particles and perform SM measurements, such as the ones presented in this thesis. The latter is used, for example, for studies of quark-gluon plasma, and is not discussed in detail here.

The LHC is the last stage in a series of machines which form the CERN accelerator complex, which is illustrated in Figure 3.1. The procedure by which particles are accelerated using this chain of machines is described in [28], and is summarised below. In p-p collisions, protons are first obtained from a hydrogen gas, which is stripped of electrons. The particles are then brought up to an energy of 50 MeV using Linear Accelerator 2 (LINAC2). The particles coming out of the initial linear accelerator are transfered into the Proton Synchrotron Booster (PSB), which raises the beam energy to 1.4 GeV, before the beams enter the Proton Synchrotron (PS). At this stage, the protons energies are increased to 25 GeV. Once the beams reach the required energy, they are passed into the Super Proton Synchrotron (SPS), and boosted to 450 GeV. Finally, the beams are injected into the LHC rings, which are two concentric beampipes within the same set of bending magnets. The LHC then brings the beams up to their final energy, which, in the most recent run, was approximately 6.5 TeV, leading to a centre-of-mass collision energy of $\sqrt{s} = 13$ TeV. A further increase in the beam energy is foreseen in the

LHC programme, which would bring it up to its design value of 7 TeV per beam, and $\sqrt{s} = 14\,\text{TeV}$.

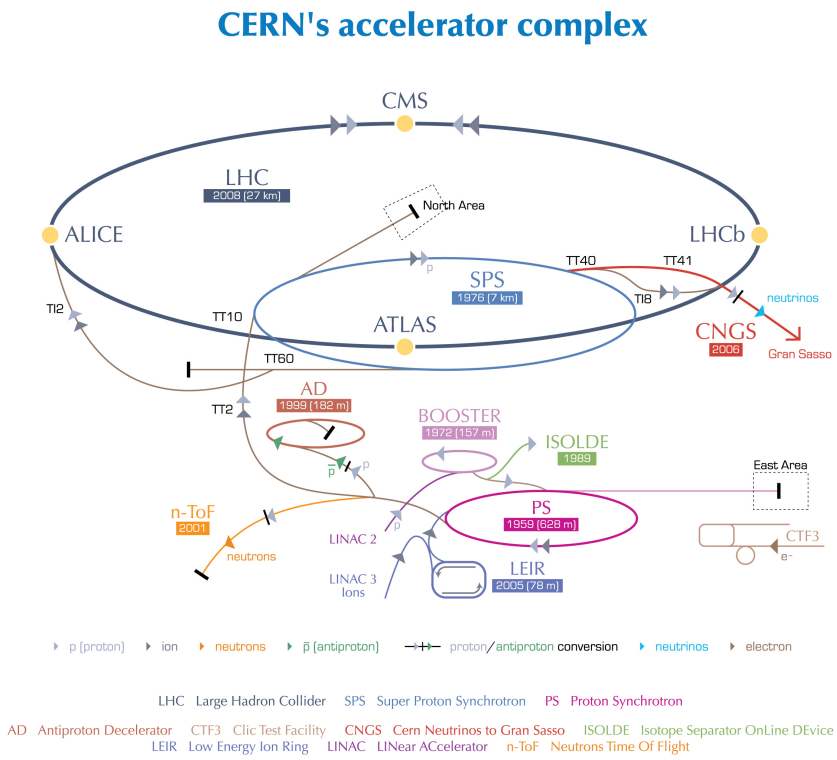


Figure 3.1: Schematic view of the CERN accelerator complex, showing the chain of machines which allow the energies of the particles to increase to 6.5 TeV: LINAC2, PS Booster, PS, SPS and finally LHC [30].

The counter-circulating beams in the LHC are bunched, with each bunch containing several billion protons. The bunch spacing was 50 ns in the initial running of the LHC, but has now been reduced to its design value of 25 ns, to allow a faster accumulation of data. The average number of times a process occurs in collision experiments (N_{process}) can be obtained from the following relation [31]:

$$N_{\text{process}} = \sigma_{\text{process}} \times \mathcal{L}_{\text{int}}, \tag{3.1}$$

where σ_{process} is the cross-section of the process and $\mathcal{L}_{\text{int}}$ is the integrated luminosity, which is the time integral of the instantaneous luminosity L. The luminosity depends only on the machine parameters, and assuming a Gaussian beam distribution, is given

by the relation [31]:

$$L = \frac{n_b N_b^2 f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{3.2}$$

where n_b is the number of bunches in each beam, N_b is the number of particles per bunch, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the normalised transverse beam emittance, β^* is the beta function at the collision point and F is a luminosity reduction factor which takes into account the fact that beams cross at a slight angle.

The LHC began operation at $\sqrt{s} = 7\,\text{TeV}$ in 2010, collecting $44\,\text{pb}^{-1}$ of data in 2010 and $6.1\,\text{fb}^{-1}$ in 2011. In 2012, the collision energy was successfully increased to $\sqrt{s} = 8\,\text{TeV}$ and $23.3\,\text{fb}^{-1}$ of data were recorded. This period corresponded to the first physics run of the LHC (Run 1). After a shutdown period for planned upgrades to the machine, the LHC began Run 2 and raised its collision energy to $\sqrt{s} = 13\,\text{TeV}$, delivering $4.3\,\text{fb}^{-1}$ in 2015 and $35.9\,\text{fb}^{-1}$ in 2016, as can be seen in Figure 3.2. The analysis of the first $12.9\,\text{fb}^{-1}$ collected in 2016 is presented in this thesis. The peak instantaneous luminosity of the LHC, achieved at the start of periods of colliding beams, is currently around $1.4 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$, significantly exceeding the design luminosity of $1 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. Run 2 is scheduled to continue until 2018 before another shutdown for upgrades is envisaged.

There are eight points along the LHC ring which are equipped with access shafts, as well as surface and underground structures. Four of these points host LHC infrastructure elements: the collimators at points 3 and 7, the RF system at point 4 and the beam dump at point 6. The remaining four points house experimental caverns where beams are focused and brought into collision. Two general purpose detectors are located at diametrically opposed sides of the ring: ATLAS [33] at point 1 and CMS [34] at point 5. Two additional specialised detectors are located at points 2 and 8, namely A Large Ion Collider Experiment (ALICE) [35] (used in the study of heavy ion collisions) and Large Hadron Collider beauty (LHCb) [36] (specialising in flavour physics) respectively.

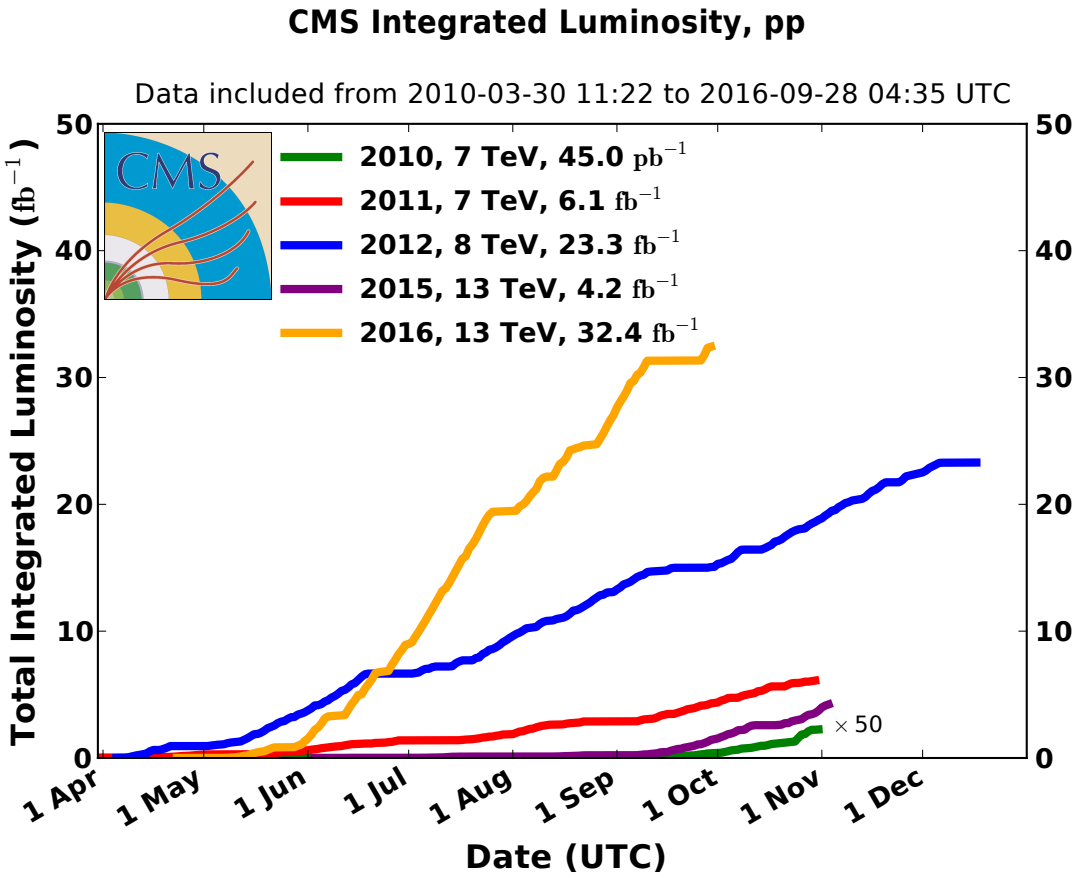


Figure 3.2: Overview of the integrated luminosity delivered by the LHC throughout its operation until the time of writing, as recorded by CMS [32].

3.2 The Compact Muon Solenoid (CMS)

3.2.1 Overview

CMS is located approximately 100 m underground at access point 5 of the LHC, near the French village of Cessy. It is over 21 m long and 14 m in diameter, weighing over 12,500 tons. It consists of a superconducting solenoid magnet 13 m long and 5.9 m in diameter, generating a 3.8 T magnetic field, which is embedded within an iron return yoke containing the muon detection system. The other sub-detectors are contained within the solenoid. The layout of the CMS detector can be seen in Figure 3.3. CMS is composed of a cylindrical barrel region closed by two endcaps. The tracker (described in Section 3.2.2), ECAL (described in Section 3.2.3), and hadron calorimeter (HCAL) (described in Section 3.2.4) are housed within the solenoid. Outside of the solenoid, four

layers of iron act as a return yoke for the magnet and house muon detector chambers (described in Section 3.2.5).

The LHC experiments use a right-handed coordinate system whereby the x-axis points towards the centre of the LHC ring, the y-axis points upwards, and the z-axis points in the direction of the counter-clockwise beam. A more convenient coordinate system can be defined for physics analyses using the variables (η, ϕ, z). In this convention, $\eta = -\ln[\tan(\theta/2)]$ is the pseudorapidity (where θ is the polar angle relative to the beam axis) and ϕ is the angle relative to the x-axis in the (x, y) plane. The direction perpendicular to the z-axis is referred to as *transverse*, while the direction pointing along it is referred to as *longitudinal*. The transverse components of energy and momentum are denoted by E_{T} and p_{T} respectively.

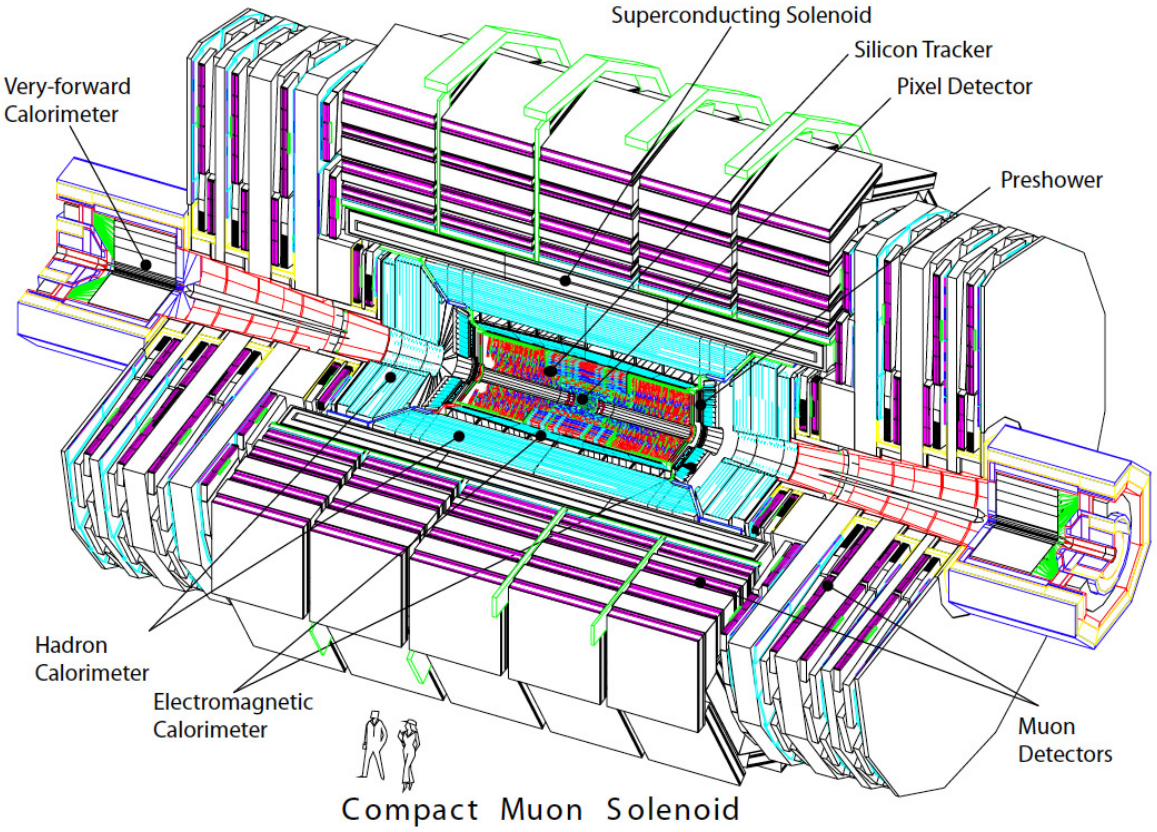


Figure 3.3: A cutaway diagram of the CMS detector, showing the main components and subdetectors, which are described in Section 3.2 [37].

3.2.2 Tracker

The closest subdetector to the beam crossing point is the tracker [38], the layout of which can be seen in Figure 3.4. This subdetector, which fits within a cylindrical volume 5.8 m long and 2.5 m in diameter, is used to measure the momenta of charged particles whose tracks are deflected in the 3.8 T magnetic field. The short interval between collisions (25 ns) requires the tracker to have a fast response time. Furthermore, the large p-p cross-section necessitates it to be resistant to radiation. The two requirements are satisfied by silicon-based detectors. Two silicon detector types are used in the CMS tracker: the pixel layers and the silicon strip layers.

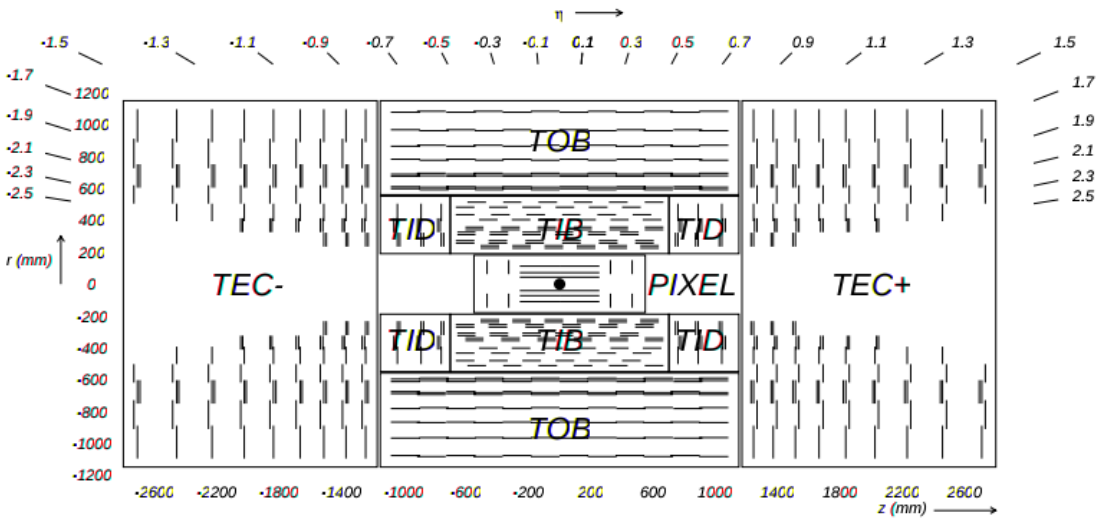


Figure 3.4: A diagram showing the layout of the CMS tracker components: the pixel tracker (labelled PIXEL) is the nearest to the interaction region marked by the black dot. The various sections of the strip tracker (TIB, TID, TOB, TEC+ and TEC−) are arranged around the pixel tracker [37].

The pixel layers are made up of 66 million 100 μm $\times$ 150 μm silicon pixels. As can be seen in Figure 3.4, these are arranged into three concentric cylinders in the barrel section (of radii between 4.4 cm and 10.2 cm) and two planes on each endcap. The spatial resolution of this part of the tracker is around 10 μm in the transverse direction. The pixel layers also have excellent longitudinal resolution (20-40 μm), which is important for vertex reconstruction. [39]

The silicon strip layers surround the pixel layers, and are composed of several sections. Four cylindrical layers form the tracker inner barrel (TIB), while the tracker inner disks (TID) are composed of three planes. Surrounding this, the tracker outer barrel (TOB)

provides a further 6 cylindrical layers, while 9 planes form the tracker endcaps (TEC). The silicon strip layers extend to 110 cm in radius. This section of the tracker uses 9.3 million strips, with each strip being 10-20 cm long and 80-183 μm wide. The transverse spatial resolution of the silicon strip layers is between 13-38 μm in the inner section and 18-47 μm in the outer section. [39]

Charged particles follow helical trajectories in the CMS magnetic field, and deposit charge as they pass through the silicon sensors. The resulting recorded signals are referred to as hits. Using multiple hits in the pixel and strip detectors, the helical trajectory can be reconstructed. The transverse momentum p_{T} of charged particles can then be extracted from the curvature of the helix. The p_{T} resolution is of the order of 2-3% in the $|\eta| < 1.6$ region and up to 11% for the outer section. Using the extrapolation from the fitted track and the longitudinal resolution of the pixel detector, tracks are grouped into common points of origin, at the primary vertex (PV) and secondary vertices.

3.2.3 Electromagnetic Calorimeter

The ECAL [37, 40] is the subdetector whose performance is the most critical to the $\mathrm{H} \to \gamma\gamma$ analysis, and its layout, operation and calibration will therefore be described in some detail.

ECAL overview

The ECAL is made up of an array of 61,200 lead tungstate ($PbWO_4$) crystals in the barrel section and 14,648 crystals in the endcaps, arranged one crystal deep. The choice of material was made because $PbWO_4$ has a short radiation length (the mean distance over which an electron loses all but $1/e$ of its energy to bremsstrahlung) of 0.89 cm. The short radiation length is important in the ECAL design because it allows electromagnetic showers to be contained within a relatively small depth of material.

The ECAL crystal front faces are 22 mm $\times$ 22 mm squares, corresponding to approximately $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$, roughly matching the Molière radius of $PbWO_4$. The individual crystal depth is approximately 26 radiation lengths, to ensure that the electromagnetic showers are fully contained within the ECAL. The array of crystals extends to $|\eta| = 3$, but precision measurements are only made up to $|\eta| = 2.5$. There are also transition regions between the ECAL barrel and endcaps around $|\eta| = 1.5$. The arrangement of the crystals in the ECAL can be seen in Figure 3.5.

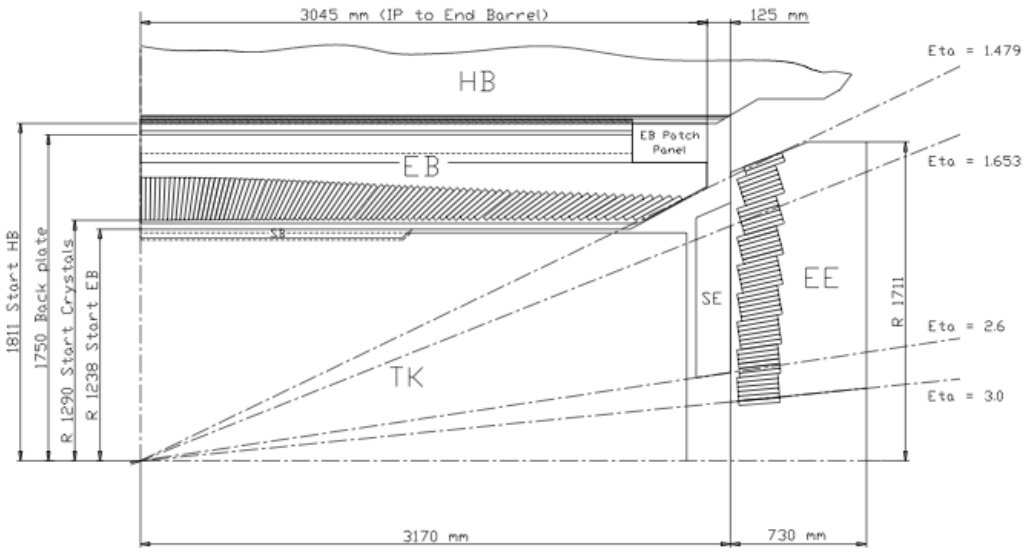


Figure 3.5: Schematic cross-section of one quadrant of the ECAL, showing the arrangement of crystals in the ECAL barrel and endcaps. The shower detector (SE, referred to as ES in the text) is also visible, as is the hadron calorimeter barrel (HB) and the tracker (TK) [40].

The ECAL consists of the ECAL barrel (EB) and the ECAL endcaps (EE). The EB provides coverage in the region $|\eta| < 1.479$, while the EE provides coverage for $1.556 < |\eta| < 2.5$. The crystals in the EB are grouped into 36 supermodules, each covering an angle of 20^o in ϕ. The crystals are arranged so that they do not point directly at the mean position of the primary interaction vertex, but are instead positioned with a 3^o offset in both θ and ϕ, to help improve the hermeticity of the detector. The EE is composed of two "D"-shaped sections, built up of *supercrystals* (units of 25 standard crystals). The EE has notably worse resolution that the EB, and this is because the calibration of the crystals is more challenging. One factor contributing to this is that the crystal transparency is affected by the high radiation doses in the EE.

An additional detector, the preshower (ES), is mounted in front of each endcap, covering the region $1.54 < |\eta| < 2.61$. The main purpose of which is to distinguish between π^0 and γ particles, and also adds three radiation lengths to the depth of the ECAL endcaps. The ES is composed of two planes of lead, of 2 and 1 radiation lengths respectively, with high granularity silicon detector strips after each.

In the barrel region, avalanche photon-diodes (APDs) operating with a gain of 50, are attached to the back of the crystals, where the scintillation light is collected. In the endcaps, vacuum photon-triodes (VPTs) are used instead of APDs as the photo-detectors.

In both cases, the photo-detectors register $\sim$4000 photoelectrons per GeV. The photo-detectors are read out by 12-bit analogue to digital converters (ADC). Ten consecutive samples are read out and stored for each crystal, and this information is used to determine the amplitude of the pulse, and therefore the amount of energy deposited in the crystal.

The resolution of the ECAL crystals is modelled with the following equation:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{3.3}$$

where S represents the stochastic term, N represents the noise term and C represents a constant term [37]. The design values of these parameters are approximately $S = 2.8\%$ GeV$^{\frac{1}{2}}$, $N = 0.12$ GeV and $C = 0.3\%$.

As can be seen in Figure 3.6, for individual $H \to \gamma\gamma$ photons in Run 1, an energy resolution of about 1% was achieved for unconverted photons in the barrel, and about 2.5% in the endcaps. For converted photons, an energy resolution of about 1.3% was observed up to $|\eta| = 1$, rising to about 2.5% at $|\eta| = 1.4$, and 3-4% in the endcaps [41].

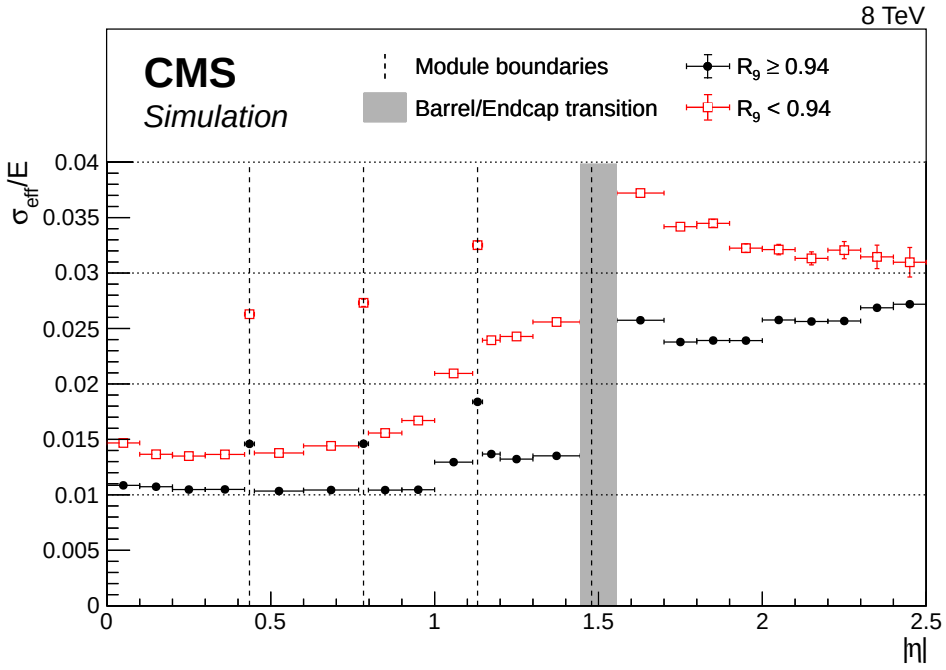


Figure 3.6: The relative energy resolution of individual simulated $H \to \gamma\gamma$ photons in Run 1 as a function of $|\eta|$, shown separately for converted photons (black circles) and unconverted photons (open red squares). The vertical lines represent the boundaries between the ECAL modules in the barrel, while the grey band indicates the transition region between the EB and the EE, where photons are not reconstructed [41].

Energy measurement

Typically, energy deposits will not be contained in a single crystal. When an electron or a photon hits the ECAL, the electromagnetic shower will spread out into adjoining crystals. In addition, particles can undergo pair conversion or emit bremsstrahlung before impacting the detector, resulting in additional associated energy deposits. Clustering algorithms are used to recover the deposits from the main impact crystal, adjacent crystals where the energy from the main shower is spread out, and the additional associated crystals, and group them into a so-called supercluster (SC). The energy of the SC (E_{SC}) can roughly be expressed as:

$$E_{\mathrm{SC}} = F_{\mathrm{SC}} \cdot G \cdot \Sigma^{i=0}_{N_{\mathrm{crystals}}}(C_i \cdot S_i(t) \cdot A_i), \tag{3.4}$$

where F_{SC} is a correction to the SC energy sum representing second-order effects, G is an ADC-to-GeV conversion factor which represents the global energy scale, C_i is a factor applied to crystal i to equalise the response (also known as an intercalibration constant), $S_i(t)$ is a time-dependent factor to correct for loss of transparency of the crystals, and A_i is the amplitude of the pulse recorded in that crystal for the bunch crossing in question. In regions covered by the ES, the signals from this subdetector are also used [42].

Calibration

The calibration of the ECAL involves using various techniques and physics objects to tune the values of $S_i(t)$, C_i and G in Equation 3.4.

The first step is to make a time-dependent correction for the transparency in the crystals. Indeed, the response of ECAL crystals varies because of radiation induced transparency loss and recovery through spontaneous annealing. Continuous monitoring and correction of the response of the crystals is required. This is achieved using a laser which periodically (every 40 minutes) injects photons of wavelength 440 nm into each crystal via a network of optical fibres. The change of the response as measured and corrected for using this mechanism is tracked in Figure 3.7. The effect of the corrections on the measured mass of the π^0 in its decay to photons can be seen in Figure 3.8.

The next step is the intercalibration, which aims to equalise the response of all crystals. This is achieved by determining a set of intercalibration constants C_i, one for each crystal.

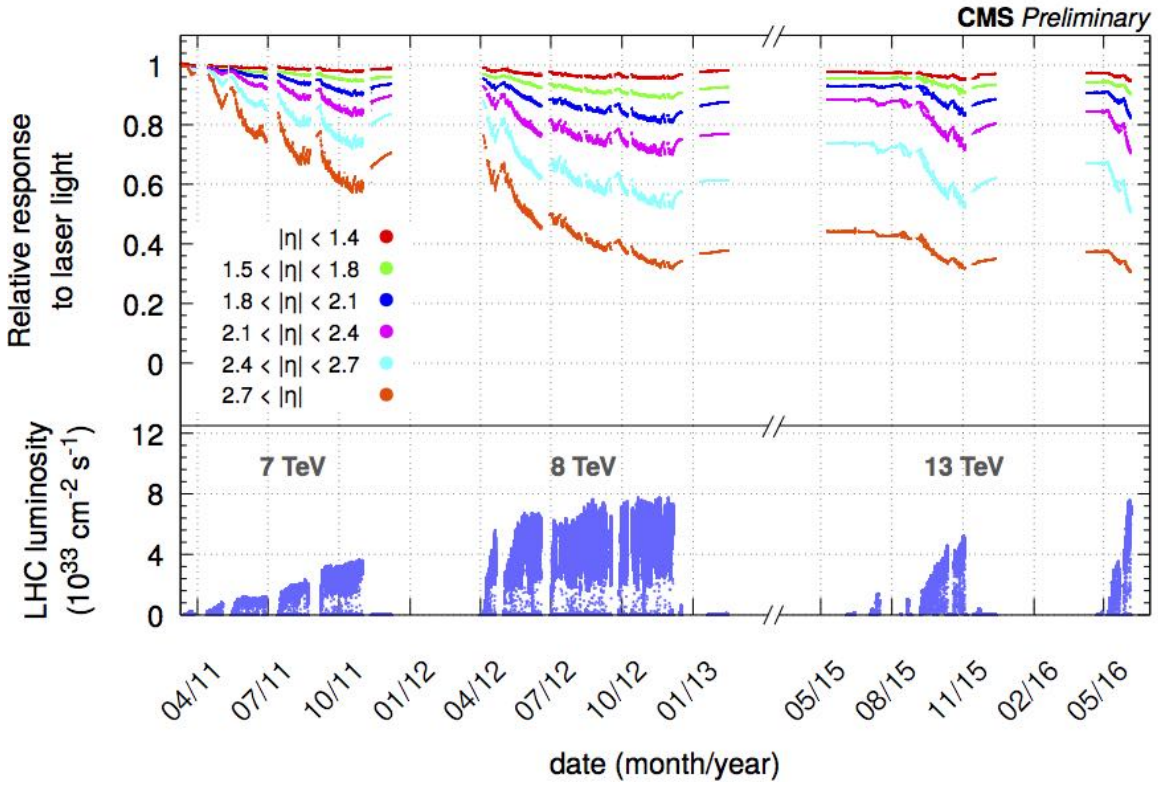


Figure 3.7: The response of the CMS ECAL lead tungstate crystals is shown as a function of time, and for different pseudorapidity ranges. The crystal response decreases as data are collected, due to transparency loss caused by exposure to radiation, and recovers during spontaneous annealing at times when no beams are present [43].

Several methods are used in this procedure. The first is to use the fact that the CMS ECAL is cylindrically symmetric. It is therefore expected that during a given period of time, the total energy measured by each crystal with the same value of η (η-ring) should be the same. Exploiting the ϕ-symmetry of the detector, one can therefore produce an intercalibration constant for each crystal in an η-ring by dividing the amount of energy that was measured by that crystal in an interval of time by the average amount measured by all crystals in that η-ring. Another method exploits the fact that the invariant mass of π^0 or η particles (as measured in their decay to photons) should be measured the same regardless of the crystal location. Therefore, one can generate intercalibration constants by dividing the measured values of the invariant masses in each crystal by the average value measured by all crystals. Intercalibration constants produced using different methods are combined to give a final set of per-crystal corrections. By construction, the average value of the intercalibration constants is unity: these corrections leave the overall scale unchanged.

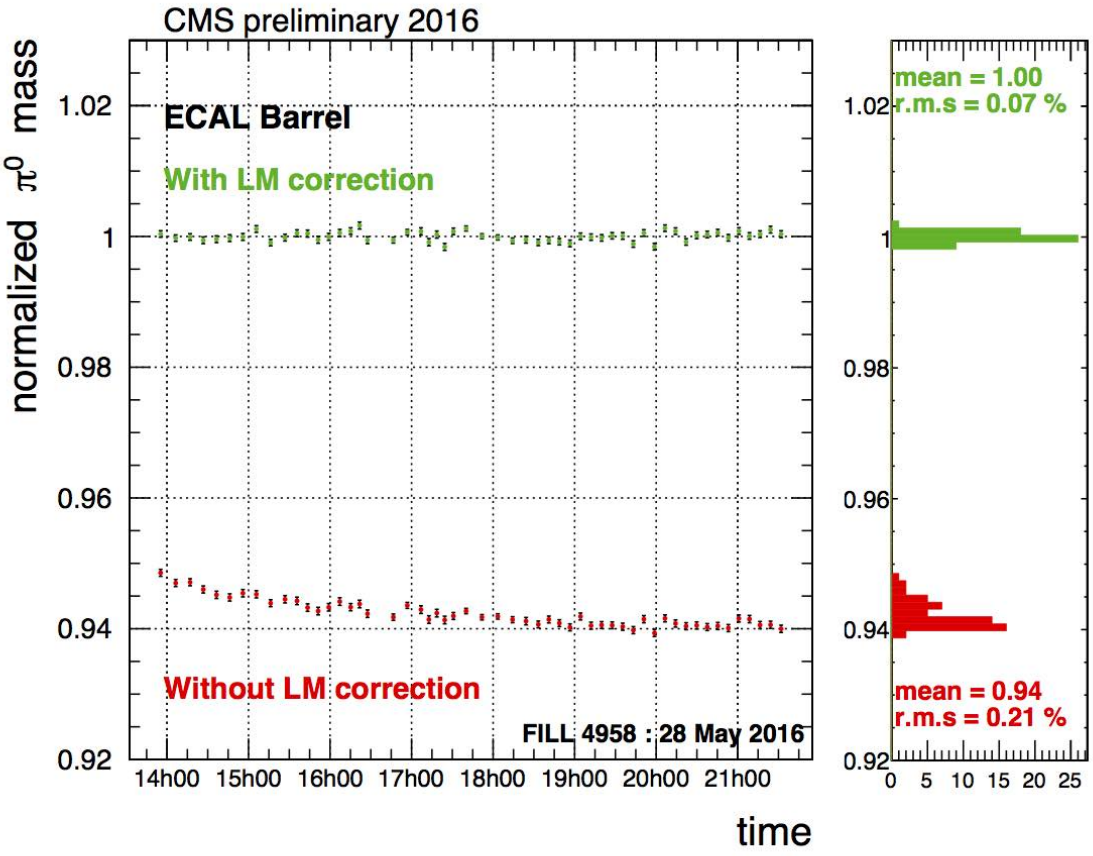


Figure 3.8: The normalised value of the invariant mass of the π^0 particle in its decay to photons, as measured by the CMS ECAL barrel, with and without the Laser Monitoring (LM) corrections for crystal transparency loss, showing the degradation of the response of the lead tungstate crystals, even over a period of hours [43].

The final step in the calibration procedure is to set the global scale G, which is also the ADC-to-GeV conversion factor. This is set by comparing the measured value of the mass of the Z boson in its decay to electrons to the nominal value. Since the mass of the Z is well known and simulated, one can set the value G such that the peak of the Z invariant mass distribution coincides with the simulated value, in different bins of η and SC type, in order to complete the calibration.

3.2.4 Hadronic Calorimeter

The HCAL is used to identify hadrons, and measure their positions and energies. In particular, it is needed to measure the energy of neutral hadrons which do not leave

any hits in the tracker or any deposits in the ECAL. Such particles need to be taken into account to accurately estimate the energies and directions of jets of particles, and to measure the magnitude and direction of any missing energy, which would indicate particles which did not interact by the CMS detector (e.g. neutrinos, or undiscovered weakly interacting particles).

The CMS HCAL is a sampling calorimeter, consisting of active material between absorber plates. The layout of the detector can be seed in Figure 3.9. The active material is a plastic scintillator read out by wavelength-shifting plastic fibres. The absorber plates are made of brass (or steel in the forward section). Brass is chosen as it is an non-magnetic material, and thus will not be affected by the strong magnetic field within the solenoid. The main body of the HCAL is composed of the hadron calorimeter barrel (HB) with coverage up to $|\eta| < 1.3$, and the hadron calorimeter endcaps (HE) with coverage up to $|\eta| < 3$, with $\Delta\phi \times \Delta\eta$ granularities between 0.087×0.087 and 0.17×0.17. In order to accurately measure missing energy, the HCAL must be as hermetic as possible. For this reason, an additional calorimeter is appended, the forward hadron calorimeter (HF), which gives coverage up to $|\eta| < 5$. This uses active quartz fibres within a steel absorber matrix. In order to fully contain hadronic showers, an additional component, the outer hadron calorimeter (HO), uses the solenoid as an absorber and is placed directly around it in the barrel region. [44]

The minimum depth of the HB is 5.8 radiation lengths, rising to 11.8 when the HO is taken into account. In the endcaps, the depth is at least 10 radiation lengths [44]. The resolution of the HCAL system was measured in test beams of single pions [45] and found to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{94.3\%}{\sqrt{E}}\right)^2 + \left(\frac{8.4\%}{E}\right)^2. \tag{3.5}$$

3.2.5 Muon detectors

The solenoid is surrounded by the outermost subdetector, the muon detector, which is built into and around the steel return yoke for the magnetic field. The CMS muon detection system consists of both endcap and barrel sections and is comprised of three types of detector: the drift tubes (DTs) in the barrel, the cathode strip chambers (CSCs) in the endcaps, and the resistive plate chambers (RPCs) in both barrel and endcaps. The muon chambers are all installed between the layers of the steel return yoke. The layout of the muon detectors can be seen in Figure 3.10. All the muon detectors are

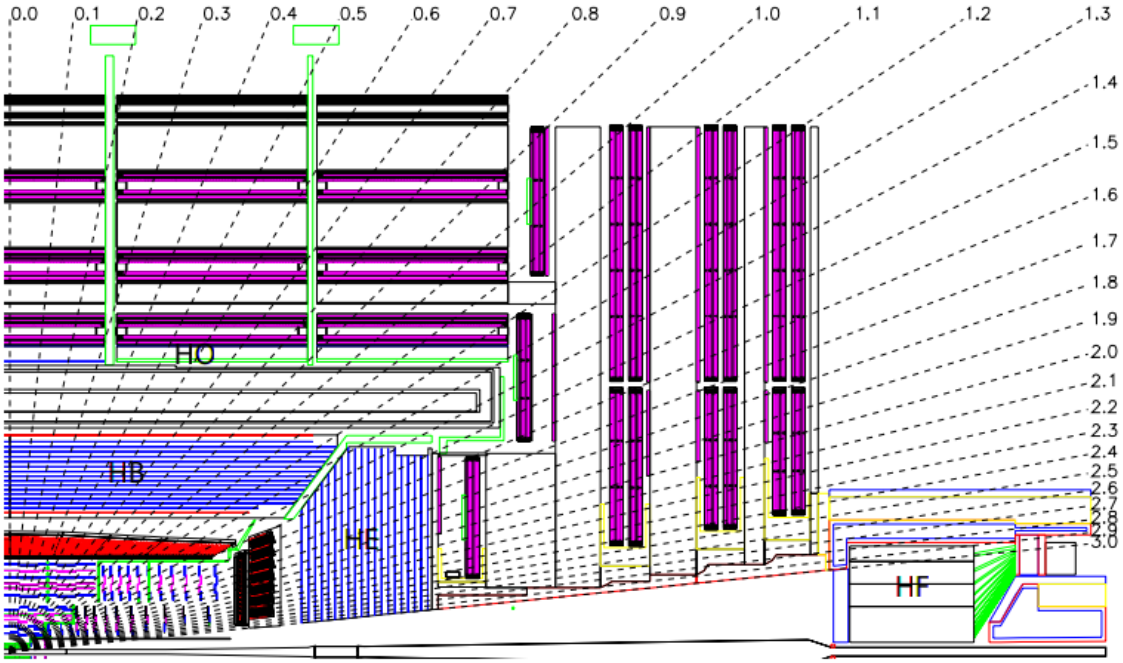


Figure 3.9: Schematic cross-section of one quadrant of the HCAL, showing the arrangement of the various components of the subdetector: The HB and HE surrounding the ECAL, with the HF at high η and the HO just outside of the solenoid [34].

gaseous detectors with a similar operational principle: as charged particles travel through a chamber, the gas contained within becomes ionised and the resulting electrons drift towards the detector's anode, which gives out an electric signal.

In the barrel section, the muon system is composed of four concentric layers of DTs, which are wire chambers filled with a mixture of gaseous Ar and CO_2. Each DT station consists of twelve layers of wire chambers, with some having their wire oriented parallel to the beam axis and other perpendicular to it. This means that each DT is able to provide a position measurement in both the transverse and longitudinal planes, with 100 μm resolution in each. The DTs have coverage up to $|\eta| < 1.3$.

In the endcaps, the field is less uniform and the neutron fluences become much larger. The muon rate is also much higher than in the barrel. Therefore, a detection system with a faster response and more resistance to radiation is needed. The CSCs, which are multi-wire chambers comprised of 6 anode wire planes interleaved among 7 cathode panels, satisfy this requirement. They are filled with a mixture of Ar, CO_2 and CF_4 gasses, and each station is also able to provide a measurement in both the longitudinal

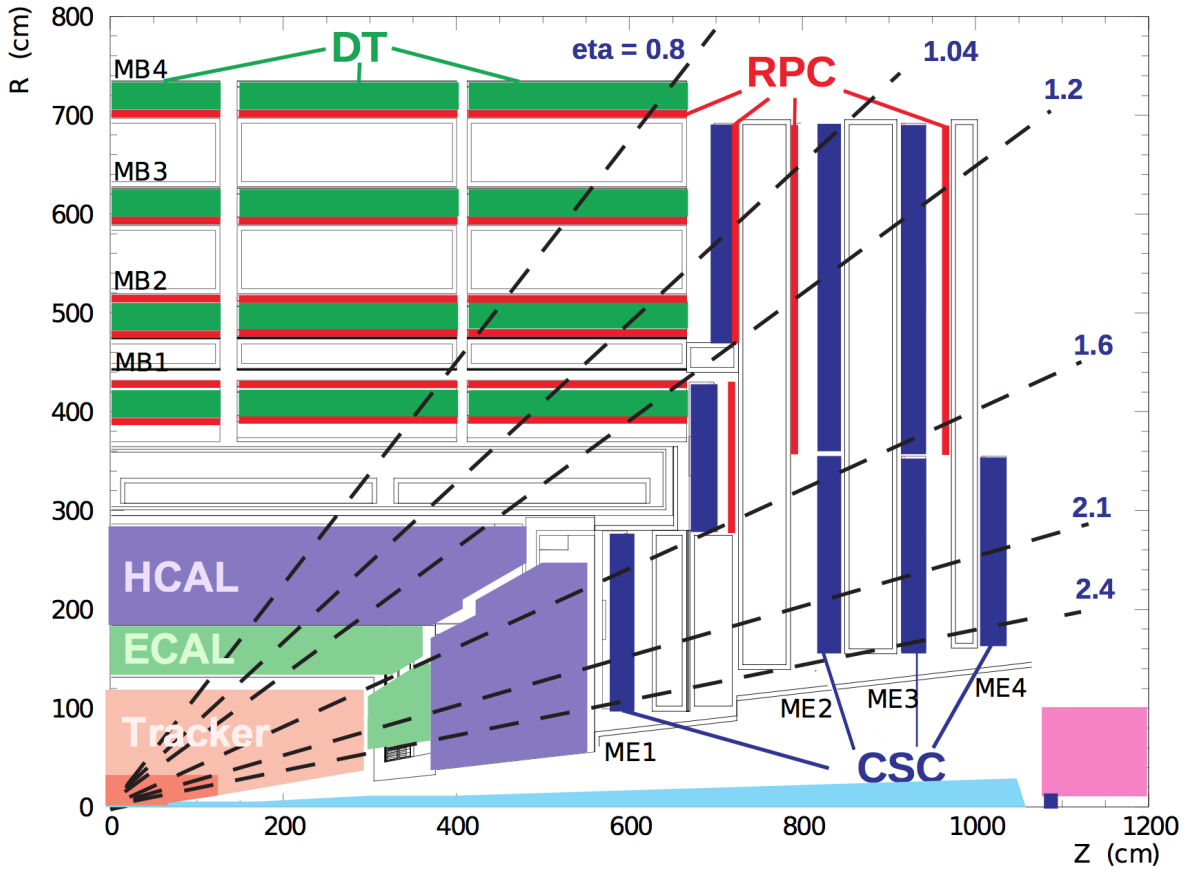


Figure 3.10: Schematic cross-section of one quadrant of the CMS muon detector, showing the arrangement of the various components of the sub-detector: The DTs in the barrel, and CSCs in the endcaps and the RPCs in both [46].

and transverse planes, with a spatial resolution around 85 μm. There are four layers of CSCs in each endcap, with coverage of $0.9 < |\eta| < 2.4$.

The final detector in the muon system is the array of RPCs. These are double-gap chambers, and have a fast response with good timing resolution but worse position resolution than the DTs or CSCs. The RPCs are used as a trigger, and can also be used to resolve ambiguities in tracking when there are multiple hits in a chamber. The RPCs are installed in both the barrel and endcap sections, up to $|\eta| < 1.6$ [34, 47].

The transverse energy resolution for muons with p_{T} below 100 GeV is 1-6% depending on their position within the detector [46].

3.2.6 Trigger and data processing

During operation, bunch crossings occur at a rate of up to 40 MHz. Each instance where the data from a bunch crossing are saved by the detector is known as an event. The storage space taken up by an event is of the order of 1 MB. If CMS were to keep all the information from each crossing, it would therefore need to save 40 TB of data per second, which is (at the time of writing) entirely impossible to store. Furthermore, the CMS sub-detector electronics are designed to read out information at a rate of at most 100 kHz, so it is not possible to read out all the collision data either. Therefore, a vast reduction in the number of events selected to be saved is required. In reality, the vast majority of collisions are of no physics interest: many will be low-energy interactions of protons instead of head-on collisions, or will come from well-understood SM processes. The strategy to reduce the number of selected events is to filter out the commonplace events and save only those which are of physics interest. This is achieved through the CMS triggering system. The trigger is used to reduce the number of saved events by a factor of order 10^5. This is achieved through two trigger levels: the level-1 trigger (L1T) and the high-level trigger (HLT).

The L1T consists of programmable, custom-designed electronics. The L1T must reduce the number of output events by a factor of at least 400, since the maximum design bandwidth of the subdetector electronics readout is 100 kHz. As collisions occur in the CMS detector, each event is stored in a buffer. A very limited time, 3.2 μs, is allocated to decide whether or not to save an event at the L1T. This must include the time taken to transmit the data from the sub-detectors to the L1T and return the decision to accept or not. The 3.2 μs latency means that the buffer must be able to hold at least 128 bunch crossings. Due to bandwidth limitations and the short latency, decisions at the L1T are made based on very coarse data from the different detector subsystems individually. There no time to transmit or exploit their full granularity and resolution, or to use detailed correlations between different sub-detectors. The short amount of processing time available also prohibits the use of information from the tracker. Based on the coarse information available, the L1T runs a series of algorithms designed to identify events where processes of interest occur. In the case of a L1T accept, the full event data are transfered off the detector and passed to the HLT [34].

The HLT consists of a farm of about 1000 commercially-available processors, which run basic reconstruction software and use the output to make a decision on which events to keep. The HLT provides a further reduction the amount of data stored, decreasing it

by a factor of about 100, down to an output of around 400 Hz. This is achieved using simplified versions of the full CMS reconstruction software, with the full granularity of the information from the CMS sub-detectors available, including from the tracker [34].

Events passing the HLT are then saved to disk and reconstructed using the full CMS software for use in physics analyses.

Chapter 4

Event reconstruction and selection

4.1 Introduction

This chapter deals with the way in which the simulation and the data collected at the CMS experiment are reconstructed for use in the $H \to \gamma\gamma$ analysis. The processing is performed for each event using the `CMSSW` package; the final selection of events specific to the $H \to \gamma\gamma$ analysis is performed using a flexible software framework called `FLASHgg`.

The CMS global event description [48, 49], referred to as the particle-flow algorithm (PF), combines information from all CMS subdetectors to reconstruct and identify individual particles. The inputs to this algorithm are the tracks reconstructed in the tracker and muon system, and the clusters of energy reconstructed in the ECAL and HCAL. The outputs of the algorithm are objects corresponding to stable particles (photons, electrons, muons, charged hadrons or neutral hadrons). These so-called PF *candidates* can then be used to reconstruct jets and identify missing energy in the event. The resolution with which momenta of particles can be measured is typically improved using PF since information from several subdetectors is available. In this scheme, ECAL SCs which are not on the extrapolated trajectory of tracks from either the muon system or the tracker are identified as photon candidates. If a track in the tracker is associated to one or more ECAL SCs, then it can be identified as an electron candidate. If a track in the tracker is consistent with a track or multiple hits in the muon system, then it is identified as a muon candidate. Tracks in the tracker which are not associated with any track in the muon system or any deposit in the ECAL are interpreted as charged hadron candidates. Finally, deposits in the HCAL which are not associated with any tracks can be identified as neutral hadron candidates.

The same reconstruction algorithms are applied both to the data collected at the CMS experiment and to simulated samples. The data samples and production of simulated samples of Monte Carlo (MC) events are discussed in Section 4.2.

It has already been noted in Section 2.2.3 that $H \rightarrow \gamma\gamma$ is one of the most sensitive decays with which to study the Higgs boson in the LHC environment. This is despite it having a small branching fraction (around 0.2%) and an irreducible background of SM processes which produce two photons. The channel benefits from a fully reconstructible final state, which is manifested as a resonant Higgs boson peak on top of a continuous diphoton invariant mass spectrum. The invariant mass of the diphoton system ($m_{\gamma\gamma}$) is given by:

$$m_{\gamma\gamma} = \sqrt{2E_{\gamma_1}E_{\gamma_2}(1-\cos\alpha)}, \tag{4.1}$$

where $E_{\gamma_1}, E_{\gamma_2}$ represent the energies of the two photons and α represents the opening angle between them. The opening angle depends on the directions of the photons. Since the CMS ECAL does not provide a directional measurement, the calculation of the opening angle relies on the spatial locations of the photons and of the Higgs boson decay vertex. The CMS ECAL can measure the spatial location of the photons with a sufficient resolution that its contribution to the uncertainty on the photon direction is negligible. However, the fact that the photon is a neutral particle and the presence of multiple interaction vertices in the event can lead to the misassignment of the vertex associated with the Higgs boson decay. This produces a large enough uncertainty in the photon direction to significantly worsen the uncertainty on the mass. Therefore, Equation 4.1 indicates that to study $H \rightarrow \gamma\gamma$, the most important steps are measuring the positions and energies of the photons, and locating the Higgs boson PV. These steps are discussed in Section 4.3 and Section 4.4 respectively.

The main Higgs bosons production processes or *modes* are described in Section 2.2.2. At leading order, in the dominant production mode (ggH), the final state consists only of the two Higgs decay photons. However, for other production modes, the Higgs boson can be produced in association with other particles. These additional particles can be reconstructed, providing information on the mode by which the Higgs boson is produced. The methods used to reconstruct such particles are described in Section 4.5.

4.2 Samples

4.2.1 Simulation samples

Samples of simulated signal events are produced for each of the main Higgs boson production modes for a range of values of m_{H} between 120 and 130 GeV. The cross-section and branching fractions used for the simulations under each m_{H} hypothesis are those recommended by the LHC Higgs Cross Section Working Group (LHCHXSWG) [27]. Signal samples are used to: prepare classifiers, for instance using the boosted decision tree (BDT) method (see Section 4.3.4); validate reconstruction and selection algorithms; and produce signal models (see Section 6.1). Signal simulations are produced at parton-level using the generator `MADGRAPH5_aMC@NLO` [50], which makes use of perturbative QCD at next-to-leading order (NLO). The parton-level samples are then interfaced with `PYTHIA 8` [51], using the tune `CUETP8M1` [52], which models the subsequent showering and hadronisation of partons.

Samples of simulated events are also generated for each of the main backgrounds of the $\mathrm{H} \to \gamma\gamma$ decay. Background samples are used to: train BDTs; validate reconstruction and selection algorithms; and optimise the categorisation scheme. The irreducible background is composed of SM processes which yield two genuine photons from a p-p interaction in the final state, and are modelled using the `Sherpa` [53] generator. The reducible background represent events where some jets are incorrectly identified as isolated photons. The largest contributors to the reducible background are γ + jet and QCD multijet events, which are modelled using the `PYTHIA 8` generator, where a filter designed to enhance the fraction of events with a large component of electromagnetic energy is applied. Drell-Yann (DY), $\mathrm{W}\gamma$ and $\mathrm{Z}\gamma$ samples are also used for validation purposes, and these are simulated using `MADGRAPH5_aMC@NLO`.

The simulated samples take into account the effects of the multiple interactions, other than the hard scattering interaction, taking place in each bunch crossing. These additional interactions are collectively referred to as pileup. The effect of pileup in previous and subsequent bunch crossings is also modelled. The samples are reweighted such that their pileup distributions match the data before they are used in the analysis. For all simulated samples, the detailed response of the CMS detector is modelled using `GEANT 4` [54].

4.2.2 Data samples and trigger

The data sample analysed in this thesis was recorded using the CMS detector in between March and July 2016, during p-p LHC collisions at $\sqrt{s} = 13\,\text{TeV}$. It corresponds to an integrated luminosity of $12.9\,\text{fb}^{-1}$. As described in Section 3.2.6, events recorded for analysis at CMS must pass the requirements of the two CMS triggering systems: the L1T and the HLT.

The L1T requires either a deposit in the ECAL with $p_\text{T} > 25\,\text{GeV}$ or two deposits with $p_\text{T} > 15\,\text{GeV}$ and $p_\text{T} > 10\,\text{GeV}$ respectively. Events passing the L1T are processed at the HLT, where a basic clustering is applied to the candidate photon deposits. The requirements for an event to be saved to the double-photon sample by the HLT are as follows:

- the event contains two candidate photons, with $m_{\gamma\gamma} > 90\,\text{GeV}$;
- the candidate photon with most energy satisfies $E_\text{T} > 30\,\text{GeV}$;
- the candidate photon with second-most energy satisfies $E_\text{T} > 18\,\text{GeV}$;
- both candidate photons pass a basic calorimeter-based identification using shower shape and isolation requirements.

Events passing the L1T and HLT requirements are saved for further processing. The efficiency of the trigger for signal events passing the preselection requirements described in Section 4.3.3 is studied using the tag-and-probe method [55]. This is a common way to determine the efficiency of a selection S in data, and is used at several stages in this analysis. The technique exploits resonances decaying to pairs of particles, in this case $\text{Z} \to \text{e}^+\text{e}^-$. The events used for the tag-and-probe method are selected such that the invariant mass of the decay products are near the mass peak of the resonant particle, thus ensuring high-purity sample. A strict identification requirement is imposed on one of the decay products, referred to as the *tag*, and while a very loose identification requirement is imposed for the other decay product, referred to as the *probe*. The requirements applied to the probe should be loose enough that they do not affect the efficiency of S. The efficiency of S is then the fraction of probes which satisfy S. The L1T efficiency is found to be above 97.5% in the EB and above 92% in the EE. The efficiency of the HLT is found to be above 97% in the EB and 96% in the EE.

4.3 Photon reconstruction

4.3.1 Clustering of ECAL deposits

About 95% of the energy of a single electromagnetic shower in the ECAL is contained in a 5×5 array of crystals. However, particles can have several showers associated with them. Photons travelling towards the ECAL may undergo pair conversion ($\gamma \rightarrow e^+e^-$), resulting in two nearby or overlapping showers usually separated in ϕ. Electrons or positrons are deflected by the magnetic field, and radiate photons from bremsstrahlung, resulting in multiple showers spread out in the ϕ-direction.

A *clustering* algorithm [41, 56] groups the ECAL energy deposits into SCs. Each SC is designed to associate all the individual deposits resulting from a photon or electron originating at a primary interaction vertex. The algorithm begins by identifying *seed* crystals as those with the largest amount of energy in the local area, above a predefined minimum noise threshold. The threshold represents approximately two standard deviations of the electronic noise in the ECAL. Next, clusters are obtained by iteratively grouping crystals which have a common side with a crystal already in the cluster if their energy is above another predefined threshold. The location of a cluster is defined as the logarithmic energy-weighted average position of the individual crystals which compose it. During the iterative clustering process, if a crystal could belong to different clusters, its energy is shared between them according to the distance between the crystal and each cluster, assuming a Gaussian shower profile. Finally, additional clusters can be merged with the original cluster to form a SC if they are aligned in the η-direction and lie within an extended window in the ϕ-diretcion. This final step is designed to recover deposits resulting from bremsstrahlung. The clusters and SCs are fed into the PF system, and are used to build electron and photon candidates.

4.3.2 Common variables used for photon and electron studies

A number of variables are used to characterise the development of the electromagnetic shower within the $PbWO_4$ crystals. These variables are used in various stages of the reconstruction and selection, and for convenience some of the most important ones are defined here.

Shower shape variables:

- R_9, the energy in the 3×3 array of crystals around a seed crystal divided by the energy of the SC. The R_9 variable gives information about whether a photon converted (low values of R_9) or not (high values of R_9);
- S_4, the energy of the most energetic 2×2 array of crystals containing the seed crystal, divided by the energy of the SC;
- n_{clusters}, the number of clusters in a SC;
- σ_η, the standard deviation of the logarithmic energy-weighted η-position of individual crystals within the SC (or the 5×5 array centred around the seed in some cases);
- σ_ϕ, the standard deviation of the logarithmic energy-weighted ϕ-position of individual crystals within the SC (or the 5×5 array centred around the seed in some cases);
- σ_{RR}, the width of the shower in the radial direction (defined in the endcaps only);
- $cov_{\eta\phi}$, the covariance of the single-crystal values of η and ϕ for the 5×5 array centred around the crystal with the most energy;
- E_{seed}/E_{SC}, $E_{\text{seed}}/E_{3\times3}$, $E_{\text{seed}}/E_{5\times5}$, ratios of the energy of the seed crystal and energy of: the SC, the 3×3 array centred around the seed, and the 5×5 array centred around the seed respectively;
- H/E, the ratio of the energy deposited in the HCAL in a cone of radius $R = 0.15$ directly behind a SC and the energy of that SC as recorded in the ECAL.

Isolation variables:

- $Iso^{\text{PF}\gamma}_{R=0.3}$, PF photon isolation i.e. the sum of the transverse energy of all PF photons inside a cone of radius $R = 0.3$ around the candidate photon, where the energies have been corrected according to the median energy density in the event (ρ);
- $Iso^{\text{PF ch. had.}}_{R=0.3}(V)$, PF charged hadron isolation i.e. the sum of the transverse energy of all PF charged hadrons associated with a particular vertex V inside a cone of radius $R = 0.3$ around the candidate photon. This quantity is typically calculated with respect to the selected primary vertex, and also for the vertex which has the largest isolation sum (referred to as the *worst vertex*). The latter helps to reject photons candidates which are misidentified jets originating from pileup;
- $Iso^{\text{tracker}}_{0.04<R<0.30}$ tracker hollow cone isolation, i.e. the sum of the p_{T} of all tracks inside a hollow cone of radius $0.04 < R < 0.30$ around the photon candidate.

4.3.3 Photon preselection

The photon candidates considered in the $H \to \gamma\gamma$ analysis are required to satisfy certain requirements on their kinematics, shower shapes and isolation. All photon candidates are required to pass an *electron veto*, which requires that there be no track with a hit in the inner layer of the pixel detector (not matched to a reconstructed photon conversion vertex) pointing to the SC [41]. Photons candidates are then grouped into *diphotons* by determining all possible pairs of photons in the event. For each diphoton, the photon with the largest p_T (the *leading* photon) must satisfy $p_T > 30\,\text{GeV}$ while the photon with the second-largest p_T (the *subleading* photon) must satisfy $p_T > 20\,\text{GeV}$. Additional requirements are made on a per-photon basis on the following variables: σ_η, H/E, $Iso^{\text{PF}\gamma}_{R=0.3}$, $Iso^{\text{PF ch. had.}}_{R=0.3}$, $Iso^{\text{tracker}}_{0.04<R<0.30}$ and R_9.

Since no triggers are defined in the simulation, the preselection is designed to be more stringent than the triggering requirement described in Section 4.2.2. This ensures that trigger effieciency after the preselection is close to 1.

The preselection efficiency, for all requirements aside from the electron veto, is measured in data and simulation using $Z \to e^+e^-$ events, using a tag-and-probe technique for photons in different regions of η and R_9. The results can be seen in Table 4.1. The efficiency of the electron veto is measured separately using a sample of $Z \to \mu\mu\gamma$ events, and found to be between 96% and 100%.

	DATA			Simulation		Ratio	
	Eff.	Stat. Unc.	Syst. Unc.	Eff.	Stat. Unc.	Eff.	Unc.
ECAL Barrel; R_9 >0.85	0.9451	0.0006	0.0192	0.9374	0.0007	1.0080	0.0192
ECAL Barrel; R_9 <0.85	0.8255	0.0012	0.0119	0.8258	0.0009	0.9960	0.0120
ECAL Endcap; R_9 >0.90	0.9099	0.0008	0.0212	0.9127	0.0010	0.9969	0.0212
ECAL Endcap; R_9 <0.90	0.4993	0.0018	0.0249	0.5024	0.0016	0.9938	0.0250

Table 4.1: Photon preselection efficiency (using all requirements aside from the electron veto) measured using $Z \to e^+e^-$ events in data and simulation with a tag-and-probe technique.

4.3.4 Boosted decision trees

The reconstruction of photons and other particles uses multi-variate analysis (MVA) techniques at several stages in the form of BDTs [57]. A BDT is obtained using the

decision tree (DT) method, where a technique known as *boosting* is applied. Problems where a BDT is of use always involve a list of items with N_{inputs} *features* or *input variables*, labelled here $\vec{x} = (x_1, ..., x_{N_{\text{inputs}}})$, and a property y, the *target variable* to be determined. The objective of a BDT is to produce a function $F(\vec{x})$ which is an estimate of the true value of y for a given set of input variable values [58].

The most common application for BDTs is event classification, for instance to determine whether a photon candidate was produced during a hard scatter or not: the target variable takes discrete values (background or signal). In other cases, BDTs are used for regression problems, where the target variable is continuous rather than discrete. For example, the energy correction for SCs in the ECAL (F_{SC} in Equation 3.4) is obtained using a regression BDT, as is described in Section 4.3.5. Unless otherwise specified, the BDTs used in this thesis are produced using the `TMVA` framework [59] as part of the `ROOT` [60] software package.

In general, a BDT is a linear combination of DTs [57]. A DT is obtained using a *training dataset* consisting of a list of items $(\vec{x}_m, y_m, w_m)$ for $m = 1, ..., N_{\text{items}}$, where $\vec{x}_m$ is a set of input variables values, y_m is the true value of the target variable and items can be weighted with weight w_m. In the simplest case, the value of y_m is binary: signal or background. A numerical value, 1 and -1 say, can be assigned to these two options respectively. The following description uses this binary output example, but can be generalised for y_m to take any number of discrete values for classification DTs, or continuous values in the case of regression DTs. To construct a DT, the training dataset is first split into two subsamples by applying a selection, which will be referred to as a *cut*, on one or more of the input variables. The *purity* $p(s)$ of a subsample s is the proportion of signal items, given by:

$$p(s) = \frac{\sum_{m=1}^{N^s_{\text{items}}} w_m \cdot \text{Bool}(y_m = 1)}{\sum_{m=1}^{N^s_{\text{items}}} w_m}, \tag{4.2}$$

where N^s_{items} is the number of items in the subsample and $\text{Bool}(X)$ is equal to 1 (0) if X it true (false). The cuts on the input variables are chosen to maximise the separation of signal and background in the resulting subsamples. This is achieved by minimizing a separation criterion. A common separation criterion is the *Gini index* $2p(s)(1 - p(s))$, which has a maximum at 0.5 for subsamples with an equal amount of signal and background items and gives 0 in cases where all items are of the same type (signal or background). Many other

separation criteria exist, for instance *cross-entropy, misclassification error, statistical significance* and *average squared error*. Each subsample can then be further split by a new set of cuts on the input variables. This procedure is repeated iteratively for each subsample until either the number of iterations reaches some predefined threshold known as the *tree depth*, or if the subsample satisfies some predetermined requirement on the value of the separation criterion. Each subsample obtained after the final set of cuts has been applied is known as a *leaf*. The output score of the items in a given leaf is then 1 if $p(s) > 0.5$ and -1 if $p(s) \leq 0.5$.

The procedure known as boosting [58] helps to improve the performance of a DT, for example by reducing the risk of overfitting the DT to statistical fluctuations in the training sample. Many boosting algorithms exist, but in all cases several individual DTs are produced, each trained on subsets or modified versions of the training dataset. The final BDT is a linear combination of the individual DTs. Supposing that there are N_{DT} individual DTs, labelled as $f_l(\vec{x}, \vec{\alpha}_l)$ where $l = 1, .., N_{\mathrm{DT}}$ and $\vec{\alpha}_l$ is the set of cuts in the corresponding DT, the full BDT, F, is written as:

$$F(\vec{x}, \vec{\beta}, \vec{\alpha}) = \sum_{l=1}^{N_{\mathrm{DT}}} \beta_l f_l(\vec{x}, \vec{\alpha_l}), \tag{4.3}$$

where $\vec{\beta} = (\beta_1, ..., \beta_{N_{\mathrm{DT}}})$ is the set of coefficients applied to each DT in the BDT. The values of of $\vec{\beta}$ are determined by the boosting algorithm [57, 59]. A consequence of Equation 4.3 is that the output of the BDT is no longer a discrete value of ± 1, but instead a semicontinuous variable between -1 and 1.

4.3.5 Photon energy reconstruction BDT

The energy of a SC is given by Equation 3.4, where F_{SC} is a correction to the SC energy which takes into account second order effects such as how well a SC is contained within a crystal. This correction is obtained using a per-SC regression BDT, referred to here as $BDT_{\gamma\mathrm{E}}$, assuming that the SC corresponded to a photon deposit.

The target variable for the $BDT_{\gamma\mathrm{E}}$ is the ratio of the true photon energy E_{true} and the raw energy of the SC E_{SC}. The $BDT_{\gamma\mathrm{E}}$ is trained separately for SCs in the barrel and endcaps, using a simulated sample of double-photon events reweighted to flatten the p_{T} and η distributions of the individual photons. The set of input variables is listed below:

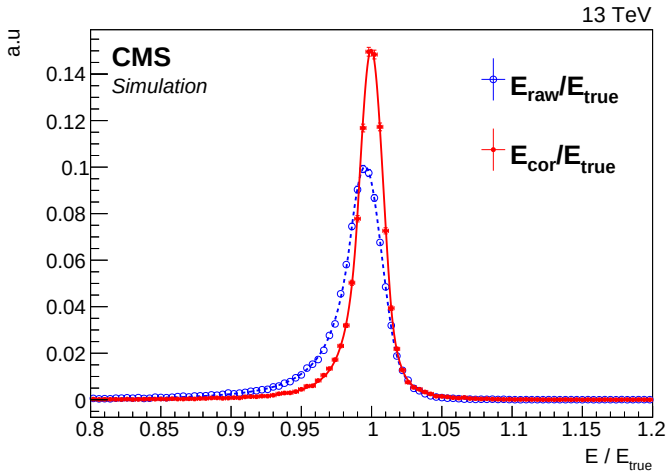
- Shower shape variables, such as those defined in Section 4.3.2, which provide information about whether a photon converted, the extent to which it began showering before hitting the ECAL, and the extent to which the shower was contained within the ECAL;
- Position variables, i.e. the position of the seed crystal in terms of crystal indices $i\eta$ and $i\phi$, the distance between the positions of the SC and the seed crystal and the distance between the seed crystal and the boundaries between ECAL modules. These variables provide information about energy lost through gaps in between the detector modules and between crystals;
- Noise variables, i.e. variables related to pileup such as the number of reconstructed vertices in the event and the median energy density in the ECAL as a function of position.

The training is done using a *semiparametric likelihood* technique. In this technique, the distribution of the target variable from the training sample is fitted to a functional form, in this case a double Crystal Ball function (DCB) [61]. The DCB is chosen as it is typically a good fit for processes modelling detector resolutions. It consists of a Gaussian core and power-law tails on each side, with 6 independent parameters: the width σ_{DCB} and mean μ_{DCB} of the core, the left tail cutoff and power parameters $\alpha^L_{DCB}, n^L_{DCB}$ and the right tail cutoff and power parameters $\alpha^R_{DCB}, n^R_{DCB}$. The fitting is performed by minimizing twice the negative log-likelihood (2NLL):

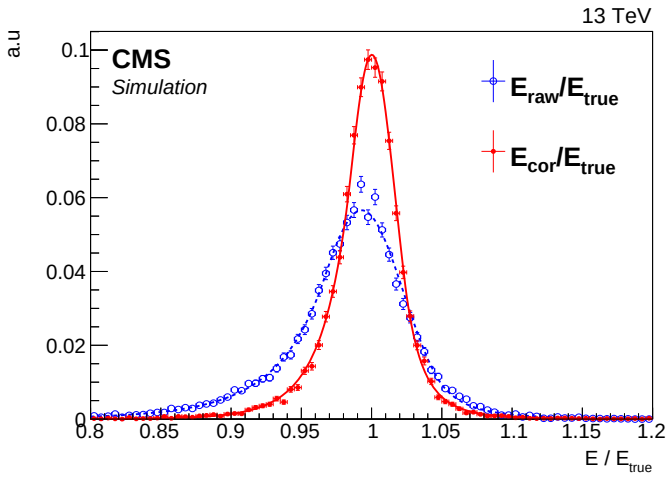
$$-2\ln\mathcal{L} = -2\sum_{\text{photons}} \ln P_{DCB}(\frac{E_{true}}{E_{SC}}|\mu_{DCB}, \sigma_{DCB}, \alpha^L_{DCB}, n^L_{DCB}, \alpha^R_{DCB}, n^R_{DCB}), \qquad (4.4)$$

where P_{DCB} is the DCB probability distribution and all the DCB parameters are functions of $\vec{x}$, the set of input variables for $BDT_{\gamma\mathrm{E}}$. The nonparametric dependence of the DCB parameters on $\vec{x}$ is obtained using separate BDTs, where a new tree is produced for each iteration of the 2NLL fit. The most probable value of E_{true} is then given by $\mu_{DCB}(\vec{x}) \times E_{SC}$ and the relative energy resolution for each photon is approximated by $\sigma_{DCB}(\vec{x})/\mu_{DCB}(\vec{x})$.

The performance of the $BDT_{\gamma\mathrm{E}}$ is illustrated using a simulated sample of $\mathrm{H} \to \gamma\gamma$ photons where $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ in Figure 4.1.



(a) EB



(b) EE

Figure 4.1: The ratio of the SC energy, shown before the $BDT_{\gamma \mathrm{E}}$ correction is applied (E_{raw}) and after the correction is applied (E_{cor}), and true energy (E_{true}) of simulated photons in a sample of $\mathrm{H} \to \gamma\gamma$ decays where $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, separately for the EB (a) and the EE (b). All distributions have been fitted to DCB functions.

4.3.6 Photon identification BDT

To reduce the large background originating from the decay of neutral hadrons (predominantly neutral pions) a per-photon BDT referred to as $BDT_{\gamma\mathrm{ID}}$ is applied to photon candidates which pass the preselection described in Section 4.3.3. The $BDT_{\gamma\mathrm{ID}}$ is trained on a $\gamma+\mathrm{jet}$ sample where the photon candidates which are geometrically matched to a generator-level photon from a p-p interaction are defined as signal. Photon candidates which have no generator-level photon match, and are therefore likely to have resulted from a misidentified neutral hadron or jet, are defined as the background. In both cases, photon candidates are required to pass the event preselection. To reduce the dependence of the $BDT_{\gamma\mathrm{ID}}$ on the kinematics of the photon, the signal photons are reweighted such that their p_{T} and η distributions match those of the background photons. The input variables for the $BDT_{\gamma\mathrm{ID}}$ are σ_η, $cov_{\eta\phi}$, S_4, R_9, σ_η, σ_ϕ, σ_{RR}, $Iso^{\mathrm{PF}\gamma}_{R=0.3}$, $Iso^{\mathrm{PF\ ch.\ had.}}_{R=0.3}$(selected vertex), $Iso^{\mathrm{PF\ ch.\ had.}}_{R=0.3}$(wrong vertex), ρ, η_{SC} and E_{SC}.

A loose requirement on the output of the $BDT_{\gamma\mathrm{ID}}$ is applied to all photons considered in the analysis, such that 99% of the signal photon candidates are kept while a large fraction of background photon candidates are removed. The $BDT_{\gamma\mathrm{ID}}$ output score of each photon is then used as a measure of the "quality" of each photon, and used as an input for the classification BDT described in Section 5.2. The $BDT_{\gamma\mathrm{ID}}$ is validated by comparing the output score for data and simulation for diphoton events (Figure 4.2) and for $\mathrm{Z}\to\mathrm{e}^+\mathrm{e}^-$ events (Figure 4.3) where the electron veto requirement is inverted. A systematic uncertainty of approximately 3% on the value of the $BDT_{\gamma\mathrm{ID}}$ output score is introduced to account for the differences between data and simulation.

4.4 Vertex reconstruction

4.4.1 Vertex identification BDT

As is discussed in Section 4.1, the determination of the location of Higgs decay is an important step in the reconstruction and selection of $\mathrm{H}\to\gamma\gamma$ events, as it impacts the calculation of the invariant mass of diphoton system. For events where the distance Δz between the true vertex and the selected vertex in the z-direction is less than 1 cm, then the impact of the opening angle uncertainty on the mass resolution is negligible. Conversely, for events where $\Delta z > 1\,\mathrm{cm}$ the $m_{\gamma\gamma}$ distribution is wider, corresponding to a degradation of the mass resolution.

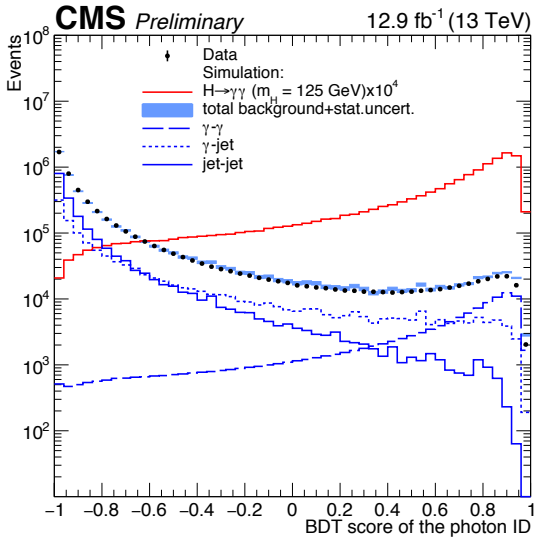


Figure 4.2: The $BDT_{\gamma\mathrm{ID}}$ output score for the lower-scoring photon in each diphoton pair in the range $100 < m_{\gamma\gamma} < 180\,\mathrm{GeV}$ for data and simulation. The simulation is composed of signal ($\mathrm{H} \to \gamma\gamma$ photons with $m_{\mathrm{H}} = 125\,\mathrm{GeV}$) and background, which has been split into γ-γ, γ-jet and jet-jet components. The sum of the background components has been scaled to the number of events in data.

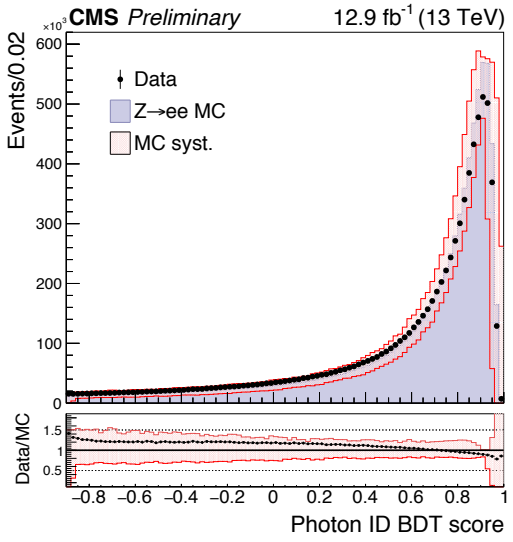


Figure 4.3: The $BDT_{\gamma\mathrm{ID}}$ output score for $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ events in data and simulation (labelled MC), where the electrons are reconstructed as photons with the electron veto requirement inverted. The red shaded region corresponds to a systematic uncertainty of approximately 3% on the value of the output score, to cover discrepancies between data and simulation.

Since the CMS ECAL is composed of a single layer of crystals, it cannot be used to point towards the vertex location. None the less, it is possible to exploit tracks recoiling from the diphoton system and the tracks of any electrons resulting from pair conversion to help determine the location of the vertex.

The first step is to produce a list of candidate vertex locations by considering all the tracks recorded in the tracker and grouping them into common points of origin by determining their closest point of approach to the beamline. Next, a per-vertex BDT is used to determine which of the candidate vertices is most likely to be the point of origin of the Higgs boson decay. This BDT is referred to as $BDT_{\text{VTX ID}}$. The set of input variables is listed below, where N^{vtx}_{tracks} is the number of charged PF candidates associated with a given vertex, $\vec{p_{\text{T}}}^{\,i}$ is the transverse momentum of the i^{th} candidate and $\vec{p_{\text{T}}}^{\,\gamma\gamma}$ is the transverse momentum of the diphoton system :

- the sum of squared transverse momenta of all tracks, $\sum_{i=0}^{N^{vtx}_{tracks}} |\,\vec{p_{\text{T}}}^{\,i}|^2$;
- the recoil of the tracks relative to the diphoton system, $\sum_{i=0}^{N^{vtx}_{tracks}} (-\,\vec{p_{\text{T}}}^{\,i} \cdot \frac{\vec{p_{\text{T}}}^{\,\gamma\gamma}}{|\vec{p_{\text{T}}}^{\,\gamma\gamma}|})$;
- the transverse momentum asymmetry, $\frac{(|\sum_{i=0}^{N^{vtx}_{tracks}} \vec{p_{\text{T}}}^{\,i}| - |\vec{p_{\text{T}}}^{\,\gamma\gamma}|)}{(|\sum_{i=0}^{N^{vtx}_{tracks}} \vec{p_{\text{T}}}^{\,i}| + |\vec{p_{\text{T}}}^{\,\gamma\gamma}|)}$.

Two additional variables are also considered if one of the two photons has converted into an e^+e^- pair, where additional information is available to help identify the vertex:

- the number of converted photon candidates in the event;
- the pull $|z_{vertex} - z_{conv}|/\sigma_{z_{conv}}$, where z_{vertex} and z_{conv} are the z-components of the positions of the reconstructed vertex under consideration and the position of the vertex extrapolated from the conversion tracks respectively, and $\sigma_{z_{conv}}$ is the uncertainty on the extrapolated vertex position.

The $BDT_{\text{VTX ID}}$ is trained using simulated Higgs boson events where the contribution from each production mode is weighted by the respective SM cross-section. Reconstructed vertices matched to a generator-level Higgs boson decay vertex are defined as the signal. All other reconstructed vertices not associated with a Higgs boson decay ares treated as background. The training samples are reweighted to account for the fact that the width of the *beamspot* (the distribution of reconstructed vertices as a function of longitudinal position) in data and simulation is not the same. This width is modelled as 5.1 cm in simulation but is measured to be 3.6 cm in the data samples used in this analysis. The reweighting is performed as a function of Δz. After the reweighting, the width of the

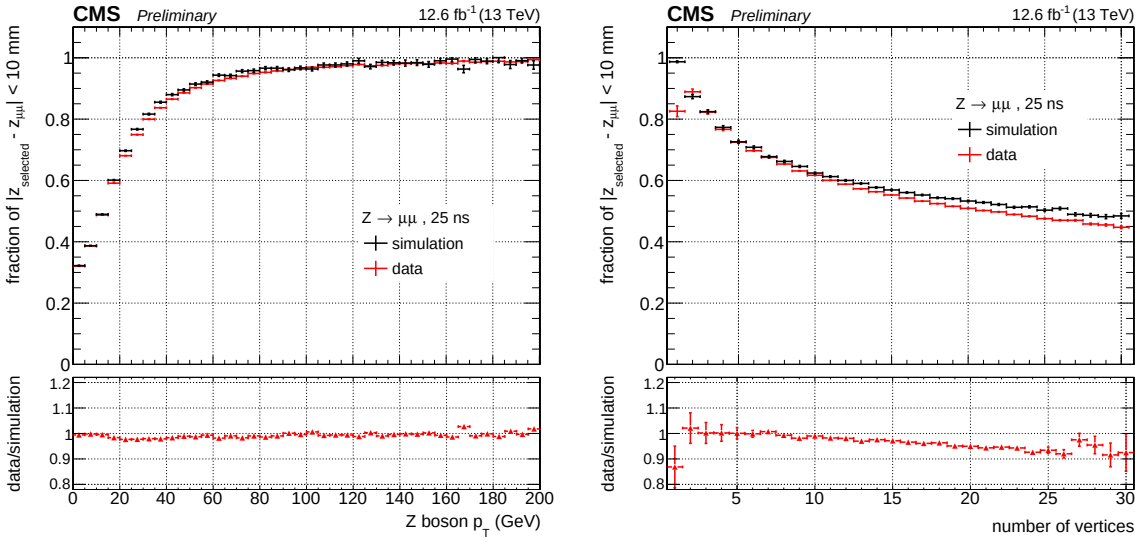


Figure 4.4: The efficiency of selecting a vertex within 1 cm of the true vertex in $Z \to \mu^-\mu^+$ events, as a function of the p_T of the Z-boson(left) and as a function of the number of vertices (right) in the event.

distribution of Δz in simulation matches the width of the Δz distribution in data, which is the beamspot width multiplied by $\sqrt{2}$.

The $BDT_{\text{VTX ID}}$ is validated for unconverted photons using $Z \to \mu^-\mu^+$ events in data and simulation. After determining the vertex of the decay from the muon tracks, the events are re-reconstructed removing the muon tracks to mimic the $H \to \gamma\gamma$ system. For converted photons, the $BDT_{\text{VTX ID}}$ is validated using a similar technique with $\gamma + \text{jet}$ samples, where the vertex is obtained from the tracks associated with the jet, and the events are re-reconstructed removing the tracks associated with the jet to imitate a diphoton system. The vertex-finding efficiencies as a function of p_T and as a function of the number of vertices in $Z \to \mu^-\mu^+$ and $\gamma + \text{jet}$ events can be seen in Figure 4.4 and Figure 4.5.

The efficiency of the $BDT_{\text{VTX ID}}$ to select the right vertex within 1 cm of the true one is estimated in simulated signal events ($m_H = 125\,\text{GeV}$), shown as a function of the number of vertices in the event and as a function of p_T in Figure 4.6. The average efficiency is of the order of 80%.

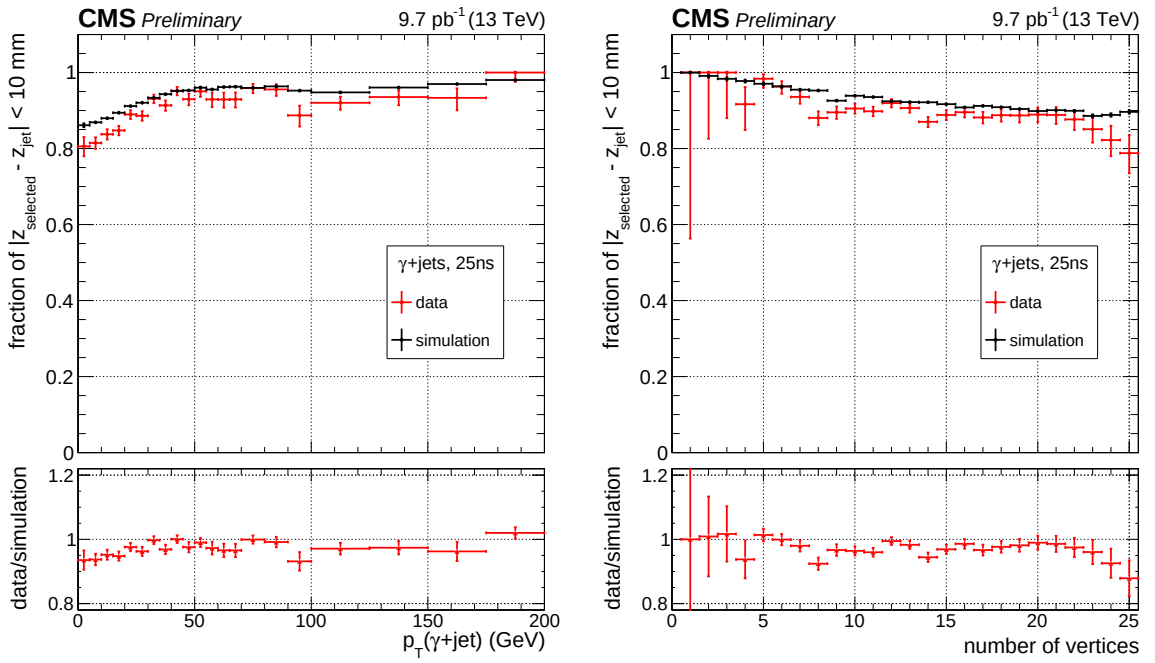


Figure 4.5: The efficiency of selecting a vertex within 1 cm of the true vertex in $\gamma+$jet events, as a function of the p_{T} of the $\gamma+$jet system (left) and as a function of the number of vertices (right) in the event.

4.4.2 Correct vertex probability BDT

An additional BDT, labelled $BDT_{\mathrm{VTX\ PROB}}$, is used to estimate of the per-event probability that the correct vertex was chosen by the $BDT_{\mathrm{VTX\ PROB}}$. The $BDT_{\mathrm{VTX\ PROB}}$ is trained on simulated $\mathrm{H} \to \gamma\gamma$ events and uses the following input variables:

- the number of reconstructed vertices in the event;
- the p_{T} of the diphoton system;
- the output scores of the three vertices ranked highest by the $BDT_{\mathrm{VTX\ ID}}$;
- the distance in the z-direction between the first- and second-highest ranked vertices;
- the distance in the z-direction between the first- and third-highest ranked vertices;
- the number of converted photons in the diphoton.

The correct vertex probability is parametrized by a 4^{th}-order polynomial as a function of $BDT_{\mathrm{VTX\ ID}}$ output score. This is done separately for converted and unconverted photons. The estimated correct vertex identification probability is shown on the same plot as the

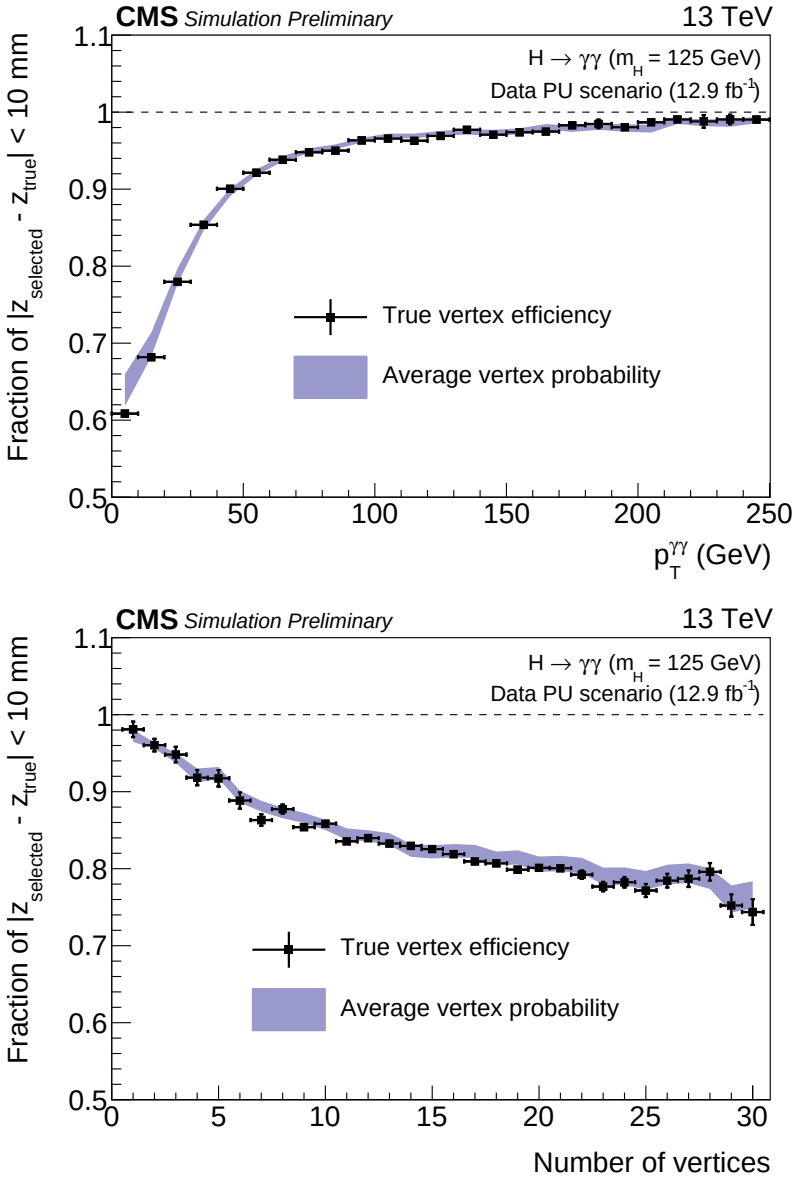


Figure 4.6: The efficiency to select a vertex within 1 cm of the true vertex in simulated $H \to \gamma\gamma$ events as a function p_T and the number of vertices in the event. The estimated probability that the vertex is chosen within 1 cm is superimposed. The uncertainty on the vertex-finding probability is determined using $Z \to \mu^-\mu^+$ events. The simulation is reweighted such that the distribution of the number of vertices and the width of the interaction region matched in data and simulation.

vertex efficiency measured in simulation in Figure 4.6. The $BDT_{\text{VTX PROB}}$ is validated using $\mathrm{Z} \to \mu^{-}\mu^{+}$ and $\gamma +$ jet events.

4.5 Reconstruction of other particles

4.5.1 Electrons

Electrons are used for the categorisation of $\mathrm{H} \to \gamma\gamma$ events where the Higgs boson events are produced by the ZH or WH mechanism. Candidate PF electrons are reconstructed either starting from ECAL deposits which are matched to tracks (called *ECAL-driven* electrons), or starting from tracks which are matched to ECAL deposit (called *tracker-driven* electrons). Typically, energetic and isolated electron candidates will be reconstructed as ECAL-driven, while low-energy ($p_{\mathrm{T}} \lesssim 10\,\mathrm{GeV}$) electrons will be reconstructed as tracker-driven. Electrons from both seeding algorithms are eventually grouped together to form the set of PF electron candidates. The electrons used in this thesis originate from $\mathrm{W}^{\pm}$ or Z decays, and hence are mostly ECAL-driven.

The ECAL-driven electrons are obtained via a procedure analogous to that described for photons in Section 4.3, but with the additional step of associating a track based on geometrical requirements. Candidate tracks are obtained from tracker hits within some window in z and ϕ around the SC position. The tracks are fitted with a special algorithm which accounts for changes in direction caused by the emission of bremsstrahlung. The SC is associated to the track whose extrapolated position in the ECAL is nearest to the energy-weighted position of the SC, but requiring that the distance in the η-direction (ϕ-direction) be no more than 0.02 (0.15). The energy of electrons is obtained from the SC energy, where the final energy correction F_{SC} is obtained using a BDT method analogous to that described in Section 4.3.5, but specially trained for electron candidates.

4.5.2 Muons

Muons are used for the selection of Higgs bosons which are produced by the ZH or WH mechanism. Muons are constructed by geometrically matching tracks reconstructed independently in the tracking system and in the muon chambers. Muon candidates must have some hits in both subdetectors to qualify as PF muons: this helps to avoids cases

where cosmic rays or muons produced in jets are misreconstructed as muons from the hard scattering interaction [46].

4.5.3 Jets

Gluons or quarks exiting the p-p interaction hadronize, forming collimated *jets* of charged and neutral hadrons. In the $H \to \gamma\gamma$ analysis, jets are used to identify events where the Higgs boson is produced by the VBF process. Jets are reconstructed using the anti-k_t algorithm [62] from PF candidates, using a cone of radius $R = 0.4$. Jets originating from pileup can sometimes overlap with jets which originate from particles produced in the hard scattering interaction. To mitigate this effect, PF charged hadron subtraction (PFCHS) is used. In this scheme, the PF charged hadron candidates associated to a vertex other than the vertex selected by the procedure described in Section 4.4 are ignored during the jet reconstruction. Since the tracker acceptance is $|\eta| < 2.5$, no PF charged hadron candidates are available outside this range, so PFCHS has no effect. For the jets reconstructed outside this region, but still in acceptance, a different pileup mitigation technique is used using a selection on the width of the jet. The width of the jet is described by the variable $\sigma_{\mathrm{RMS}} = \sum_{\mathrm{PF\ candidates}} p_{\mathrm{T}}^2 \Delta R^2 / \sum_{\mathrm{PF\ candidates}} p_{\mathrm{T}}^2$, where ΔR is the distance between the PF candidate and the jet axis from the cone. Jets must have $\sigma_{\mathrm{RMS}} < 0.03$ to pass the pileup mitigation requirement. Finally, all jets are required to be within $|\eta| < 4.7$.

Parametric corrections to the energy of the jets are made to account for the additional energy of PF neutral hadrons from pileup which are included in jets and the nonuniformity of the detector response.

4.5.4 Missing energy

Certain particles,such as neutrinos, do not leave any deposits in the detector, and therefore carry away a certain amount of energy which cannot be reconstructed. This results in an imbalance in the sum of transverse momentum. The amount of Missing Transverse Energy (MET) is calculated by considering the magnitude and direction of p_{T} required to balance all the jets and PF objects in an event. Reconstructed MET is used to identify decays from $\mathrm{W}^{\pm}$ bosons, for example when identifying Higgs boson decays originating from the WH production mode.

Chapter 5

Event categorisation

5.1 Introduction

The basic experimental method for the observation of the $H \to \gamma\gamma$ decay is to search for a resonance above the diphoton continuum background. The sensitivity of this method is enhanced by the categorisation of events according to their expected signal-to-background ratio. Furthermore the use of additional particles in the event allows the measurement of the cross-section of each production mode individually, thus probing the strength of the Higgs boson's interaction with different types of particle.

The most common production mode, ggH, produces a Higgs boson in isolation. This leaves only the photons resulting from $H \to \gamma\gamma$ decay in the final state. The other production modes (VBF, VH and ttH) produce the Higgs boson accompanied by additional particles. The VBF mode has two quarks in the final state, which hadronize to form jets. The VH mode produces a Higgs boson in association with a W or Z boson, which then decays to charged leptons, neutrinos or quarks, leading to reconstructed leptons, MET or jets. Finally, the ttH mode produces the Higgs boson in association with two top quarks, which decay to bottom quarks and either hadrons or leptons. These reconstructed additional particles can be used to categorise events.

The BDT used for categorisation of diphoton events is described in Section 5.2. The categorisation of VBF and ttH events is then discussed in Sections 5.3 and 5.4. This analysis does not have any categories which specifically target VH events. Finally, the inclusive categories and the categorisation hierarchy are respectively discussed in Section s 5.5 and 5.6.

5.2 Diphoton BDT

For events reconstructed and selected as described in Chapter 4, a BDT referred to as $BDT_{\gamma\gamma}$ is used rank diphotons by their expected signal-to-background ratio. The $BDT_{\gamma\gamma}$ is required to assess diphotons independently of their invariant mass, otherwise the m_{H} value of the training sample would introduce a bias in the output score. The input variables for the $BDT_{\gamma\gamma}$ are therefore chosen to be uncorrelated with the invariant mass of the diphoton system:

- the transverse momentum of each photon divided by $m_{\gamma\gamma}$;
- the η-position of each photon;
- the $BDT_{\gamma\mathrm{ID}}$ output score of each photon;
- $\cos(\Delta\phi)$, the cosine of the angle between the photons in the ϕ-direction;
- $\sigma_{\gamma\gamma}^{\mathrm{RV}}/\,m_{\gamma\gamma}$, the per-event estimated mass resolution of the diphoton, assuming that the correct vertex was identified;
- $\sigma_{\gamma\gamma}^{\mathrm{WV}}/\,m_{\gamma\gamma}$, the per-event estimated mass resolution of the diphoton, assuming that the vertex was not correctly identified;
- p_{rv}, the per-event estimate of the probability that the correct vertex was chosen, calculated using the $BDT_{\mathrm{VTX\ PROB}}$ described in Section 4.4.2.

The per-event estimated mass resolutions are calculated from the individual photon energy resolution estimates, labelled $\sigma_{\gamma_1}^{E}/E_{\gamma_1}$ and $\sigma_{\gamma_2}^{E}/E_{\gamma_2}$, which are given by the semiparametric regression $BDT_{\gamma\mathrm{E}}$ described in Section 4.3.5. If the vertex was correctly identified, the dominant contributions to the uncertainty on the mass resolution are the energy resolutions of each photon. Assuming Gaussian resolution functions, $\sigma_{\gamma\gamma}^{\mathrm{RV}}/\,m_{\gamma\gamma}$ can be obtained by simply adding the individual relative photon energy resolutions in quadrature:

$$\sigma_{\gamma\gamma}^{\mathrm{RV}}/\,m_{\gamma\gamma} = \frac{1}{2}\sqrt{(\sigma_{\gamma_1}^{E}/E_{\gamma_1})^2 + (\sigma_{\gamma_2}^{E}/E_{\gamma_2})^2}. \tag{5.1}$$

However, if the vertex was incorrectly identified, the uncertainty on the opening angle contributes significantly to the mass resolution. The effect is modelled by including an additional term which represents the uncertainty on the mass due to the uncertainty on the vertex position, labelled $\sigma_{\gamma\gamma}^{V}$. The distance between the true vertex and the selected vertex in the z-direction follows a Gaussian distribution which has a width equal to the width in z of the beamspot multiplied by $\sqrt{2}$. Given the spatial positions of the photons,

$\sigma^{V}_{\gamma\gamma}$ can therefore be calculated explicitly, and included in the sum in quadrature:

$$\sigma^{\mathrm{WV}}_{\gamma\gamma} / \, m_{\gamma\gamma} = \sqrt{(\, \sigma^{\mathrm{RV}}_{\gamma\gamma} / \, m_{\gamma\gamma})^2 + (\sigma^{V}_{\gamma\gamma} / \, m_{\gamma\gamma})^2}. \tag{5.2}$$

The $BDT_{\gamma\gamma}$ is trained on simulated samples of signal and background processes. For the signal samples, events from different productions modes (all with $m_{\mathrm{H}} = 125\,\mathrm{GeV}$) are mixed according to their SM cross-sections. The signal events used for training are also re-weighted by a factor w^{sig} given by:

$$w^{\mathrm{sig}} = \frac{p_{rv}}{\sigma^{\mathrm{RV}}_{\gamma\gamma} / \, m_{\gamma\gamma}} + \frac{1 - p_{rv}}{\sigma^{\mathrm{WV}}_{\gamma\gamma} / \, m_{\gamma\gamma}}, \tag{5.3}$$

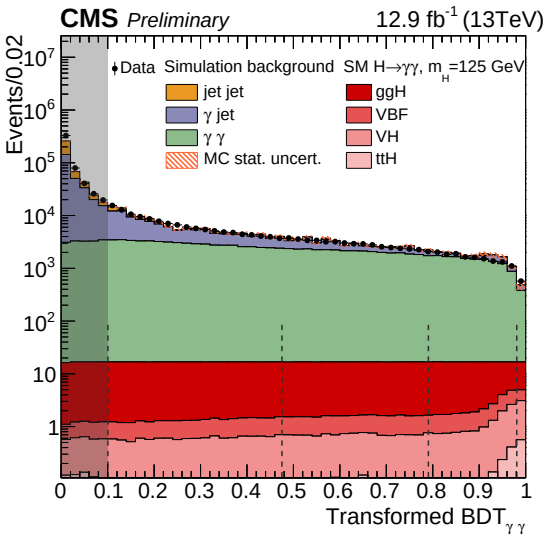
which codifies the fact that the signal-to-background ratio is inversely proportional to the mass resolution. This step ensures that the $BDT_{\gamma\gamma}$ gives a high score to events with good mass resolution. The background for the training is composed of simulated diphotons originating from the irreducible and reducible SM background processes.

Figure 5.1a shows the transformed $BDT_{\gamma\gamma}$ output score for signal and background events in the range $100 < m_{\gamma\gamma} < 180\,\mathrm{GeV}$. The transformation is applied to the $BDT_{\gamma\gamma}$ output score to give a flat distribution for signal events. The transformed $BDT_{\gamma\gamma}$ output score is validated using $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ events in data and simulation, as can be seen in Figure 5.1b.

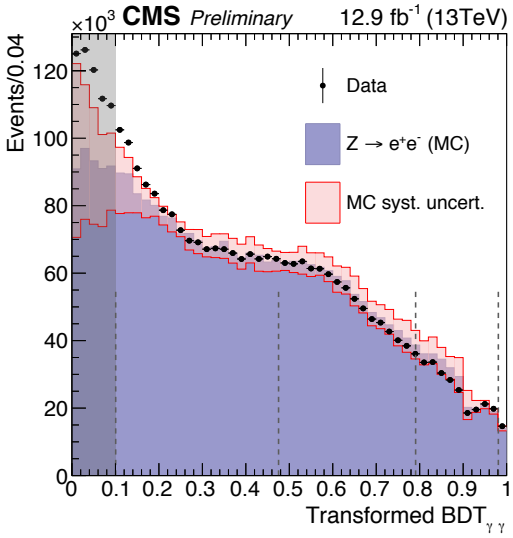
5.3 VBF-tagged categories

The VBF production mode has a cross-section approximately ten times smaller than that of the ggH mode. The additional high-p_{T} jets in the event allow the identification of VBF-like events with a high signal-to-background ratio. For this reason, although VBF events occur much less frequently than ggH events, defining VBF-tagged categories significantly improves the overall sensitivity of the analysis.

Candidate VBF events are selected by requiring that they contain two jets reconstructed as described in Section 4.5.3 after PFCHS with respect to the selected diphoton vertex. Furthermore, the jets must pass requirements aimed at removing jets that originate from other p-p interactions in the event, and must be separated from the leading and subleading photons by $\Delta R > 0.4$. The leading (subleading) jet is required to satisfy $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ ($p_{\mathrm{T}} > 20\,\mathrm{GeV}$). For events passing these requirements, an additional



(a)



(b)

Figure 5.1: (a) The transformed $BDT_{\gamma\gamma}$ score for simulated signal and background events in the range $100 < m_{\gamma\gamma} < 180\,\text{GeV}$. The transformation flattens the signal distribution. (b) The transformed $BDT_{\gamma\gamma}$ score for $\text{Z} \to \text{e}^+\text{e}^-$ events in data and simulation, where the electrons are reconstructed as photons. The pink shading represents the systematic uncertainty associated with the $BDT_{\gamma\text{ID}}$ and the $BDT_{\gamma\text{E}}$. For both (a) and (b), the vertical dashed lines represent the boundaries of the Untagged categories described in Section 5.5, while the grey shading represents the area for which diphotons are rejected.

selection on the invariant mass of the *dijet* composed of the leading and subleading jets, m_{jj}, is imposed: $m_{\mathrm{jj}} > 250\,\mathrm{GeV}$.

For events passing the dijet preselection described above, the VBF-tagged categorisation proceeds as follows. First, a BDT referred to as BDT_{jj} is trained to give a high score to events where the dijets are VBF-like. In particular, this is trained to reject ggH events, and so cannot directly incorporate information about the diphoton quality from the $BDT_{\gamma\gamma}$. A further BDT referred to as the $BDT_{\mathrm{jj},\gamma\gamma}$, which treats ggH events as neither signal nor background, is used to include the expected signal-to-background ratio of the diphoton. The $BDT_{\mathrm{jj},\gamma\gamma}$ has the $BDT_{\gamma\gamma}$ and the BDT_{jj} output scores as its inputs variables. It is thus able to combine the VBF-like dijet identification power of one BDT with with mass resolution information from the other. A selection on the output score of the $BDT_{\mathrm{jj},\gamma\gamma}$ is then used to categorise the candidate VBF-tagged events.

The BDT_{jj} is trained on simulated samples of diphoton events where the signal is defined as VBF ($\mathrm{H} \to \gamma\gamma$) events. The background consists of samples of SM events with a diphoton and a dijet in the final state, in addition to a simulated sample of ggH events where dijets are formed from pileup and initial or final state radiation. The input variables for this BDT are listed below:

- the invariant-mass-scaled transverse momentum ($p_{\mathrm{T}}/\,m_{\gamma\gamma}$) for the leading and subleading photons in the diphoton candidate;
- the transverse momenta of the leading and subleading jets in the dijet;
- m_{jj};
- $|\eta_{\mathrm{j}_1} - \eta_{\mathrm{j}_2}|$, the separation of the jets in the dijet in the η-direction;
- $\eta^* = |\eta_{\gamma\gamma} - (\eta_{j_1} + \eta_{j_2})/2|$, the *Zeppenfeld* variable [63];
- $|\phi_{\gamma\gamma} - \phi_{\mathrm{jj}}|$, the separation of the dijet and the diphoton in the ϕ-direction.

The distributions of the BDT_{jj} output scores for data and simulated background samples (and some simulated signal samples) are shown in Figure 5.2a.

The $BDT_{\mathrm{jj},\gamma\gamma}$ is trained on simulated events where the signal is a sample of VBF $\mathrm{H} \to \gamma\gamma$ events, while the background is composed of the SM diphoton background samples, as for the BDT_{jj} training. In this case, the ggH events are used neither as signal nor as background. The inputs to the BDT are the following:

- the output score of the $BDT_{\gamma\gamma}$;

- the output score of the BDT_{jj};
- $p_{\mathrm{T}}^{\gamma\gamma} / m_{\gamma\gamma}$, the invariant-mass-scaled momentum of the diphoton system, which is included since it has a significant correlation to both the other inputs.

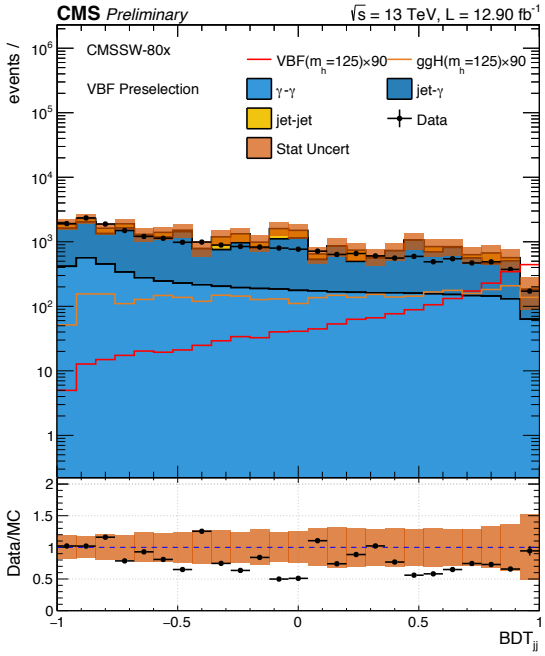
The distributions of the BDT_{jj} output scores for data and simulated background samples are shown in Figure 5.2b.

Two VBF-tagged categories (labelled 0 and 1) are defined by selections on the $BDT_{\mathrm{jj},\gamma\gamma}$ output score. The location of the two boundaries is optimised first by maximising the value of the signal-to-background ratio in the VBF-tagged 0 category, and then repeating the procedure after fixing the first boundary to maximise the signal-to-background ratio in the VBF-tagged 1 category. Of the simulated signal events in the VBF-tagged 0 category, approximately 72% are VBF events and 27% are ggH events. The corresponding values for the VBF-tagged 1 category are 55% for VBF events and 43% for ggH events. Events for which the $BDT_{\mathrm{jj},\gamma\gamma}$ output score is below the lowest boundary fail the VBF-tagged categorisation, but may be still included in other analysis categories. Repeating the optimisation procedure for three VBF-tagged categories did not lead to an improvement in the expected sensitivity of the analysis.

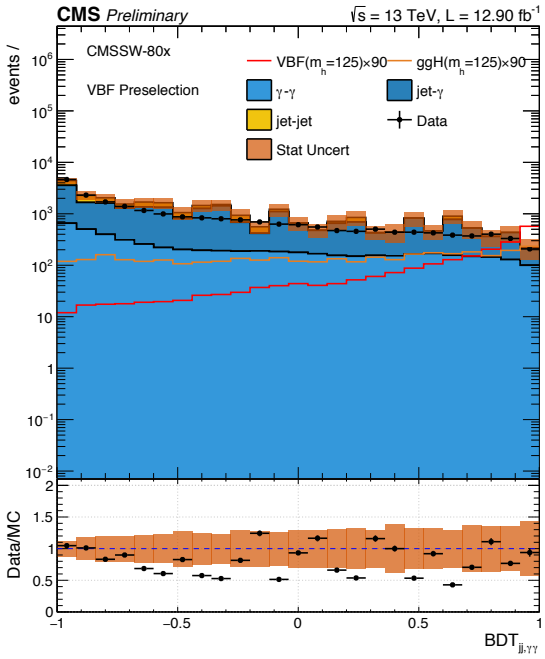
5.4 ttH-tagged categories

Events where a Higgs boson is produced in association with a pair of top quarks will contain a pair of b quarks and W bosons from their decay. The W bosons will then decay either hadronically or leptonically. The cross-section of ttH production is low, so the benefit to the analysis in terms of final significance of an observation is small. However, the categorisation of ttH-like events is important because it allows an estimate of the strength of the interaction of the Higgs boson with top quarks. Various extensions to the SM predict enhanced values of the strength of the ttH interaction, and such models can be tested through the experimental measurement of the cross-section of ttH events decaying to $\mathrm{H} \to \gamma\gamma$.

Two exclusive ttH-tagged categories are defined in this analysis. On the one hand, the leptonic ttH-tagged category aims to select ttH events where at least one of the W bosons decayed leptonically. On the other hand, the hadronic ttH-tagged category targets events where both W bosons decayed to quarks. In addition to the usual preselection applied to candidate events, the requirement on the leading photon p_{T} is increased to $p_{\mathrm{T}} > m_{\gamma\gamma}/2$,



(a) BDT_{jj} output comparing data and simulated signal and background.



(b) $BDT_{\mathrm{jj},\gamma\gamma}$ output comparing data and simulated signal and background.

Figure 5.2: The output scores of the BDT_{jj} (a) and $BDT_{\mathrm{jj},\gamma\gamma}$ (b) split by simulated production mode and comparing data and simulation. The dijet preselection has been applied in both cases.

because the p_{T} spectrum of Higgs decay particles is shifted towards higher values due to recoil against the tt system.

The leptons which are used in the selection must pass the same identification requirements as those uised for the leptonic VH-tagged categories. In order to be included in the leptonic ttH-tagged category, events must satisfy the following conditions:

- the diphoton must satisfy a loose selection on the $BDT_{\gamma\gamma}$ output which has approximately 70% signal selection efficiency and 15% background selection efficiency;
- the event must contain at least one selected lepton with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$;
- the event must contain at least 2 jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, $|\eta| < 2.4$ and separated by at least a distance $\Delta R = 0.4$ from a photon or lepton candidate;
- at least one of the jets should be tagged as a b jet using the CVSv2 algorithm medium requirement, as described in [64].

For events to be included in the hadronic ttH-tagged category, the following selections are made:

- the diphoton must satisfy a loose selection on the $BDT_{\gamma\gamma}$ output score which has approximately 95% signal selection efficiency and 45% background selection efficiency;
- there must be no leptons in the event which meet the requirements for the leptonic ttH-tagged category;
- there must be at least five jets in the event satisfying $p_{\mathrm{T}} > 25\,\mathrm{GeV}$;
- at least one of the jets should be tagged as a b jet using the CVSv2 algorithm medium requirement, as described in [64].

Events which fail the selections for the ttH-taggeds may still be selected for other categories.

5.5 Inclusive categories

The remaining events are split into inclusive categories using the $BDT_{\gamma\gamma}$ output score. The number of inclusive categories and the locations of the boundaries between them is optimised using simulated samples which are independent from those used to train the $BDT_{\gamma\gamma}$.

For a given number of inclusive categories N_{cat}, the boundaries are initially spaced evenly throughout the $BDT_{\gamma\gamma}$ output score distribution. Events falling below the lowest boundary are discarded. Simplified models are used to parametrise the signal and background $m_{\gamma\gamma}$ distributions in the remaining categories. For the signal, a sum of two Gaussians is used, representing the detector resolution and the uncertainty due to incorrect vertex assignment respectively. The background $m_{\gamma\gamma}$ spectrum is parametrised using an exponential. The expected significance is then obtained by producing and fitting an Asimov dataset [65] in each category. The estimation of the expected significance is iteratively repeated, allowing the boundaries to float. The final set of boundaries is chosen such that the expected significance is maximised for a given N_{cat}.

The procedure can be repeated for different values of N_{cat}. In this analysis, $N_{\text{cat}} = 4$ was chosen, as moving to $N_{\text{cat}} = 5$ produced a negligible improvement in the expected significance. The boundaries are represented by the vertical dashed lines in Figures 5.1a and 5.1b. The resulting categories (labelled Untagged) are numbered from 0 (highest signal-to-background ratio) to 3 (lowest signal-to-background ratio).

5.6 Categorisation hierarchy

Each event is tested to see if it can be included in the categories described previously, according to a hierarchy. If it satisfies the requirements of the first category, the event is assigned and the next event is tested. If not, the event is tested for next category in the hierarchy, and so on, until no further categories remain and the event is discarded. Each event is assigned to only one category in the hierarchy, which is ordered as follows: leptonic ttH-tagged, hadronic ttH-tagged, hadronic VH-tagged, VBF-tagged 0, VBF-tagged 1, Untagged 0, Untagged 1, Untagged 2, Untagged 3.

Chapter 6

Signal and background modelling

The statistical interpretation of the data requires models of the expected $m_{\gamma\gamma}$ distribution for each of the signal processes and the background in each category. These models must account for the sources of systematic uncertainty which affect the analysis. For the signal processes, the $m_{\gamma\gamma}$ distribution takes the form of a resonant peak around m_{H}, the width of which is entirely dominated by detector resolution. The construction of the signal model from simulation is described in Section 6.1. By contrast, the $m_{\gamma\gamma}$ distribution for background events is a smoothly-falling non-resonant continuum. The derivation of the background model from data is described in Section 6.2. Finally, the handling of the systematic uncertainties in the modelling is described in Section 6.3.

6.1 Signal modelling

6.1.1 Parametrisation of the signal $m_{\gamma\gamma}$ distributions

The shape of the $m_{\gamma\gamma}$ distribution in simulated $\mathrm{H} \to \gamma\gamma$ events is parametrised separately for each event category and for each production process. Since the vertex choice affects the shape of the $m_{\gamma\gamma}$ distribution (see Section 4.4), the modelling is performed separately for cases where the right vertex (RV) or wrong vertex (WV) was selected (within 1 cm in the z-direction).

The signal $m_{\gamma\gamma}$ distributions are all parametrised using a DCB function [61] summed with an additional Gaussian function sharing the same mean. We refer to this functional form hereafter as a DCB+1G function. The DCB shape consists of a Gaussian function core and two tails which are described by power laws. The Gaussian core models the

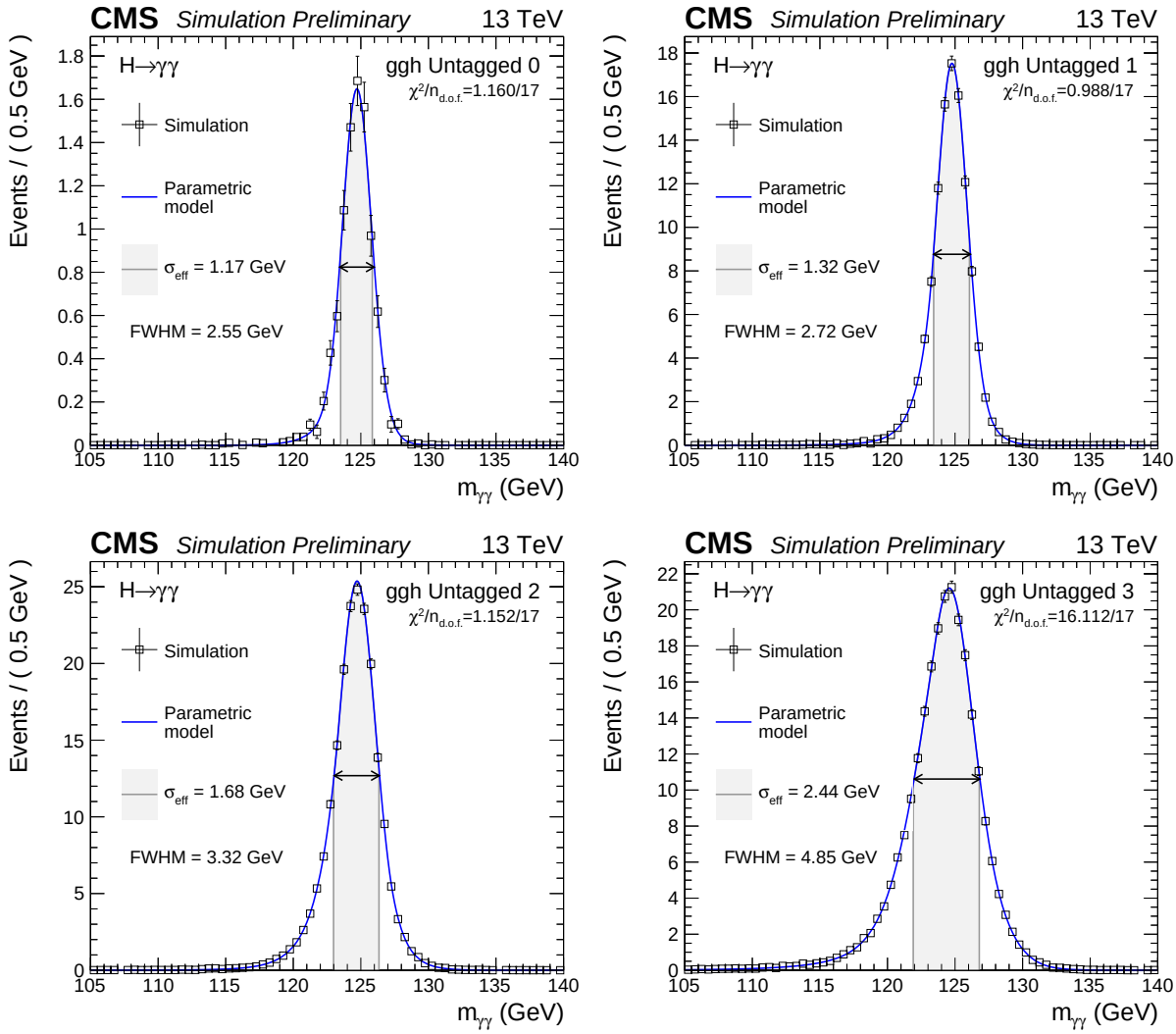
width of the distribution, while the power laws cover the extended non-Gaussian tails caused by systematic under- or over-estimation of the value of $m_{\gamma\gamma}$. The explicit form of a DCB function is :

$$f(x) = N \cdot \begin{cases} \exp(-\frac{(\alpha^L_{DCB})^2}{2}) \cdot (1 - \frac{\alpha^L_{DCB}}{n_1} \cdot (\alpha^L_{DCB} + \frac{x-\mu_{DCB}}{\sigma_{DCB}}))^{-n^L_{DCB}}, \\ \qquad\qquad \text{when } \frac{x-\mu_{DCB}}{\sigma_{DCB}} \leq -\alpha^L_{DCB} \text{ (Low-tail power law);} \\ \exp(-\frac{(x-\mu_{DCB})^2}{2\sigma^2_{DCB}}), \\ \qquad\qquad \text{when } -\alpha^L_{DCB} < \frac{x-\mu_{DCB}}{\sigma_{DCB}} < \alpha^R_{DCB} \text{ (Gaussian core);} \\ \exp(-\frac{(\alpha^R_{DCB})^2}{2}) \cdot (1 - \frac{\alpha^R_{DCB}}{n^R_{DCB}} \cdot (\alpha^R_{DCB} - \frac{x-\mu_{DCB}}{\sigma_{DCB}}))^{-n^R_{DCB}}, \\ \qquad\qquad \text{when } \frac{x-\mu_{DCB}}{\sigma_{DCB}} \geq \alpha^R_{DCB} \text{ (High-tail power law).} \end{cases} \tag{6.1}$$

In the definition above, σ_{DCB} and μ_{DCB} are the width and mean of the Gaussian core, α^L_{DCB} and α^R_{DCB} are the cross-over points to the power laws on the left and right sides, and n^L_{DCB} and n^R_{DCB} are the orders of the two power laws. The DCB+1G function therefore has a total of eight parameters when the width and amplitude of the additional Gaussian (with the same mean μ_{DCB}) are taken into account. This is in contrast with the functional form used for signal modelling in previous studies of $H \to \gamma\gamma$ by CMS [1, 2, 5], which was a sum of up to five Gaussian functions, or a total of up to fourteen parameters.

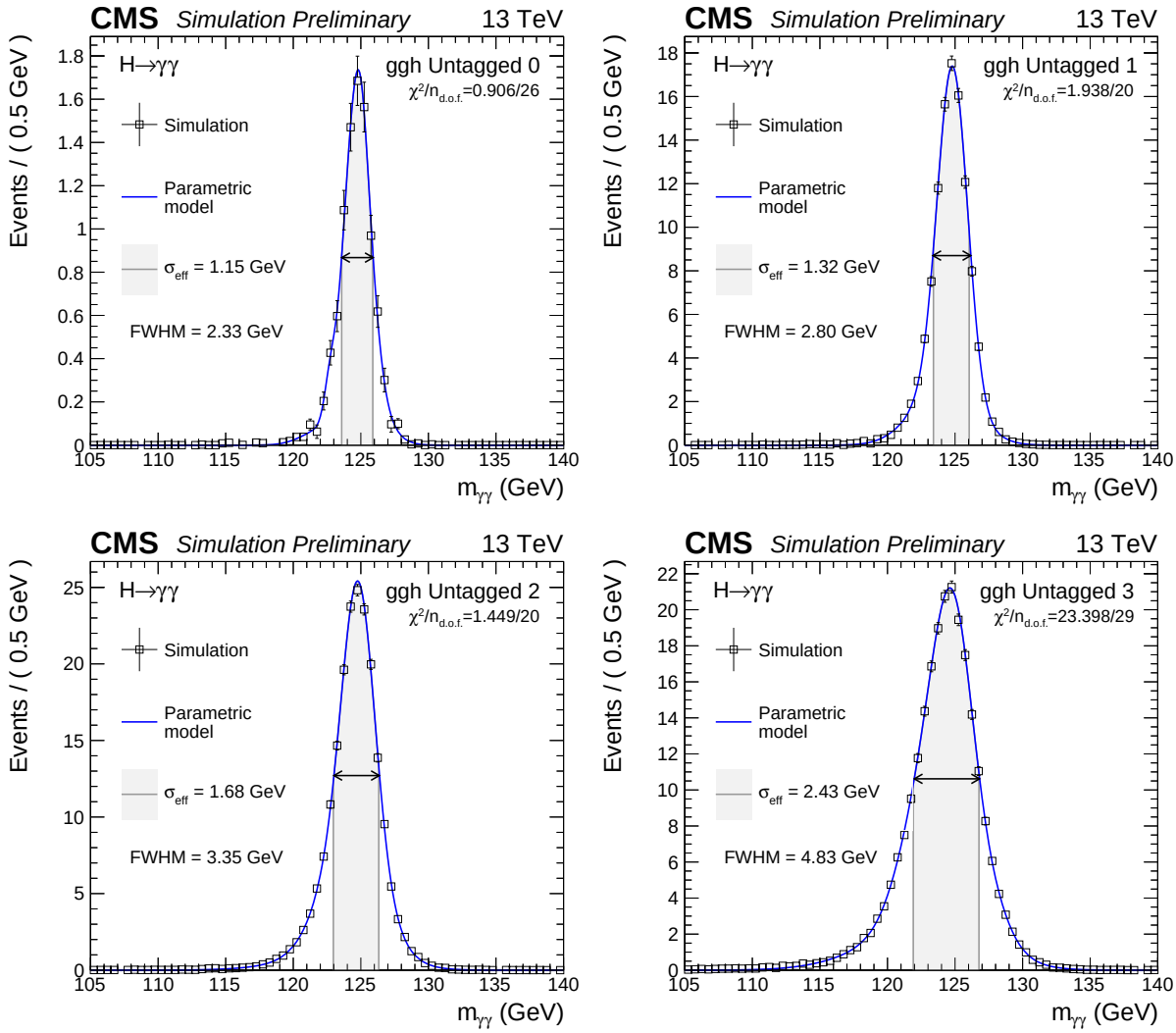
The values of the parameters of the functional form are determined by fitting the models to simulated $m_{\gamma\gamma}$ distributions. The choice of the DCB+1G function is motivated by the fact that it is simpler to handle and often gives closer agreement than the sum of Gaussian functions with fewer degrees of freedom. The DCB shape is also more susceptible to a simple physical interpretation than an arbitrary sum of Gaussian functions.

The parametrisations of the $m_{\gamma\gamma}$ distributions for the ggH process (with $m_H = 125\,\text{GeV}$) in the inclusive categories are shown in Figure 6.1 for the DCB+1G functional form and in Figure 6.2 for the sum of Gaussians. In both cases the RV and WV components have been summed according to their relative event count. The DCB+1G generally gives equivalent or better agreement considering the lower number of degrees of freedom. This is also true for the other processes and categories.



(a) Functional form: DCB+1G

Figure 6.1: Examples of the shape of the simulated $m_{\gamma\gamma}$ distribution (m_{H} =125 GeV) for the ggH process for the inclusive categories when parametrised with the DCB+1G functional form, where the RV and WV contributions have been summed according to their relative event count. The parametrisation has 17 degrees of freedom: eight for the DCB+1G shape for each of the vertex scenarios, and one additional one representing the mixing fraction. The plots show the agreement between the simulation and the parametrisation expressed as a χ^2, alongside the total number of degrees of freedom in the parametrisation. This figure is to be compared with the sum of Gaussians parametrisation shown in Figure 6.2.



(a) Functional form: Sum of Gaussians

Figure 6.2: Examples of the shape of the simulated $m_{\gamma\gamma}$ distribution (m_{H} =125 GeV) for the ggH process for the inclusive categories when parametrised with the sum of Gaussians functional form, where the RV and WV contributions have been summed according to their relative event count. The models contain between three and five Gaussians in each vertex scenario, leading to between 17 and 29 degrees of freedom after summing the RV and WV contributions with a mixing fraction. The plots show the agreement between the simulation and the parametrisation expressed as a χ^2, alongside the total number of degrees of freedom in the parametrisation. This figure is to be compared with the DCB+1G parametrisation shown in Figure 6.1.

6.1.2 Dependence of model on m_{H}

Since the mass of the Higgs boson is not exactly known, the parametrisations from simulated signal samples assuming different m_{H} values are combined to form a single parametric model for each process and each category. There are two methods to perform this combination.

The first option is to follow the fitting procedure described above for each m_{H} case separately. The individual parameters of the functional form can then be linearly interpolated from one m_{H} case to the next to produce the parametric model. This is the approach taken in previous $\mathrm{H} \to \gamma\gamma$ analyses at CMS [1, 2, 5].

The second option, which is used in this analysis, instead performs a simultaneous fit of all the different m_{H} samples, where the individual parameters of the functional form are themselves polynomials of m_{H}. The floating parameters in the fit are then the coefficients of these polynomials. The advantage of this method, which is referred to as simultaneous signal fitting (SSF), is that it guarantees a sensible parametric model. By contrast, the linear interpolation method can lead to discontinuous or unphysical models. This is because the individual m_{H} hypotheses are parametrised separately and often must be adjusted by hand to produce a consistent model. Furthermore, the SSF method reduces the total number of parameters used to determine the full parametric model.

The SSF method is applied separately for each process, category and RV or WV case, using seven different m_{H} values: 120, 123, 124, 125, 126, 127 and 130 GeV. The description of the parameters of the DCB+1G function uses polynomials of order 1. Polynomials of order 0 and 2 were also tested. No substantial improvement in the agreement was observed when using an order greater than 1: the floating parameters were not sufficiently constrained by the simulated $m_{\gamma\gamma}$ distributions to bring any meaningful improvement.

The signal models for the RV and WV contributions are then combined. The fraction of events where the selected vertex was within 1 cm in the z-direction from the true vertex is evaluated for each m_{H} sample, and then parametrised as a first order polynomial to get a smooth dependence on m_{H} for the RV/WV mixing fraction.

6.1.3 Normalisation of signal models

The signal models are normalised so that their integrated contents correspond to the number of events predicted by the SM after detector acceptance (A) and selection

efficiency (ϵ) are taken into account. The expected number of events in each sample (ignoring A and ϵ) is first computed as the product of the relevant process cross-section, the $H \to \gamma\gamma$ branching fraction and the integrated luminosity of the data sample. The values of the cross-sections and branching fraction as a function of m_H are taken from [27]. Values of $\epsilon \times A$, for each production process and each event category, are obtained from the analysis of the simulated samples, by comparing the final event content to the expected number of events. The dependence on m_H of these values is parametrised by polynomial fits. Figure 6.3 shows the overall $\epsilon \times A$ for all categories combined as a function of m_H.

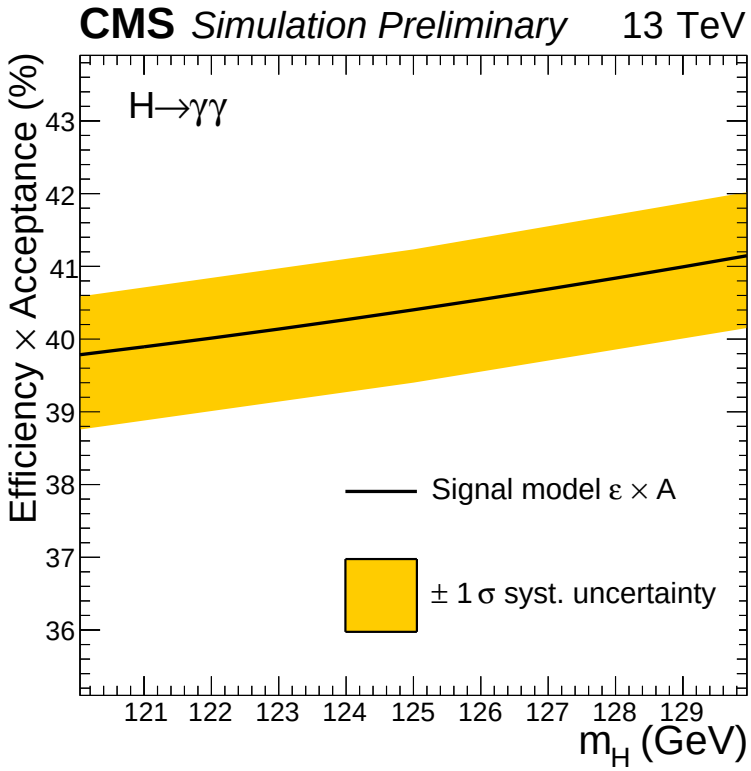


Figure 6.3: The $\epsilon \times A$ of all categories combined shown as a function of m_H. The orange band shows the effect of the systematic uncertainties associated with trigger efficiency, photon identification and selection, photon energy scale and resolution, and vertex identification.

The final normalisation of the signal models for an arbitrary value of m_H can then be obtained from the parametrised $\epsilon \times A$ multiplied by the relevant cross-section, branching fraction and integrated luminosity. As an example, the dependence of the full normalised parametric signal model on m_H for the ggH process in each of the inclusive analysis categories is shown in Figures 6.4. The equivalent figures for all processes in all categories are available in Appendix A in Figures A.6, A.7, A.8, A.9, and A.10.

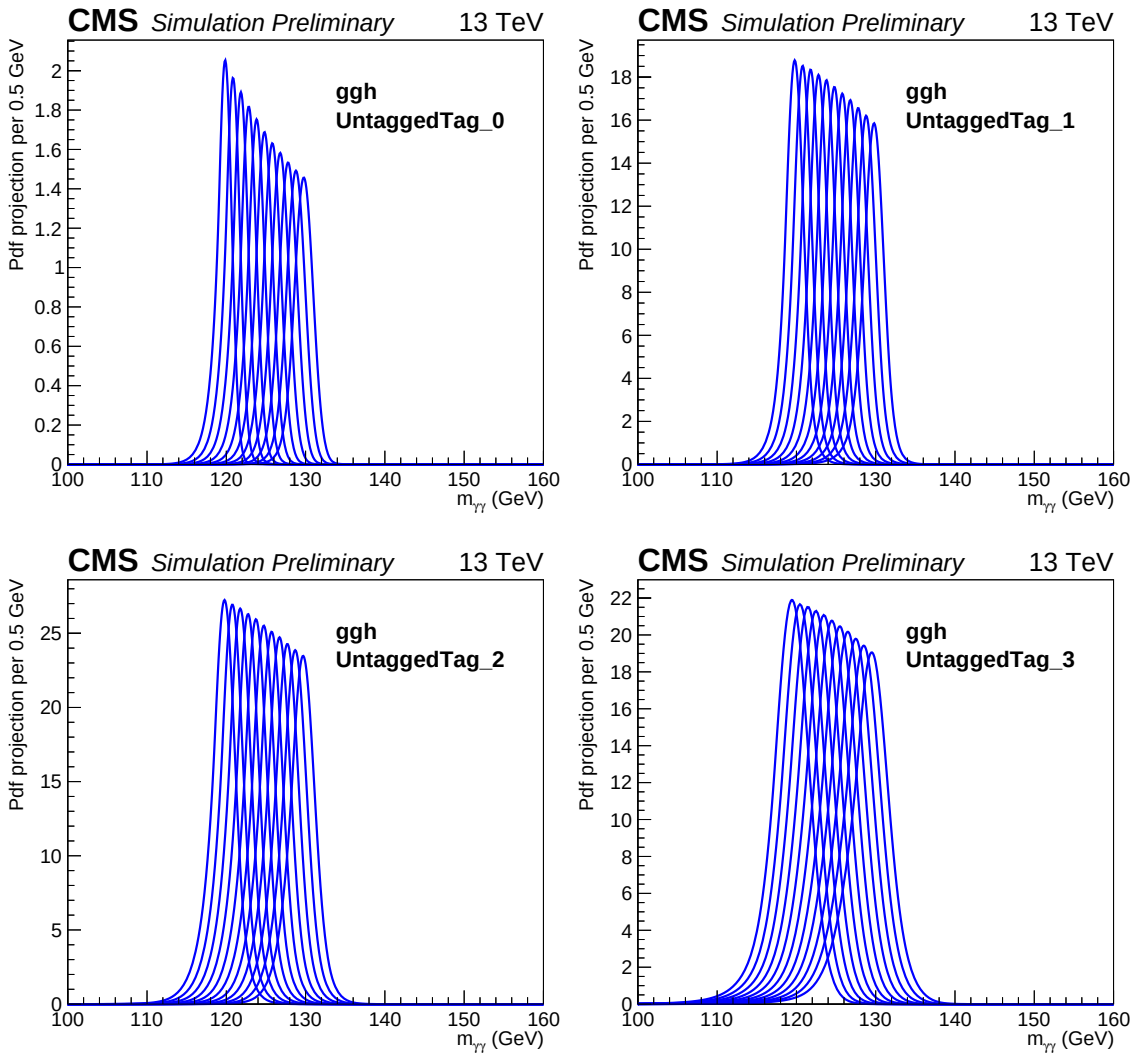


Figure 6.4: The m_{H}-dependence of the signal models for the ggH process for each of the Untagged categories is shown. Each curve shows the signal model for a given value of m_{H}. The contributions from the RV and WV components of each model were interpolated between the samples for different m_{H} using the SSF method, and summed together according to their relative event content.

The normalised signal models for each process are summed together to obtain the expected signal $m_{\gamma\gamma}$ distribution for all categories combined (as shown in Figures 6.5), or for individual analysis categories (as shown in Figures 6.6 and 6.7).

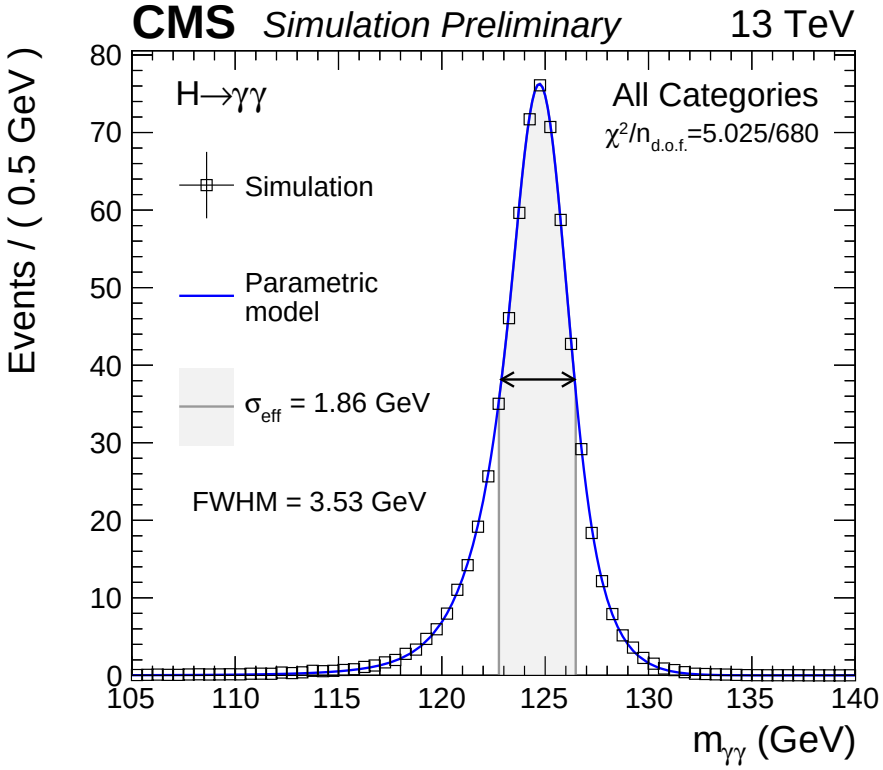


Figure 6.5: The signal model for all analysis categories combined for $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained by summing the contributions from each production process according to the $\epsilon \times A$. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown.

6.2 Background modelling

6.2.1 The discrete profiling method

The background model is derived from data by fitting a function to the invariant mass spectrum in each category. However, the underlying functional form of a given background distribution is unknown. In some cases, several different families of function could in principle be chosen for the parametrisation, all giving acceptable agreement with the data. Even within a given family of functions, it is not always clear which to choose. The

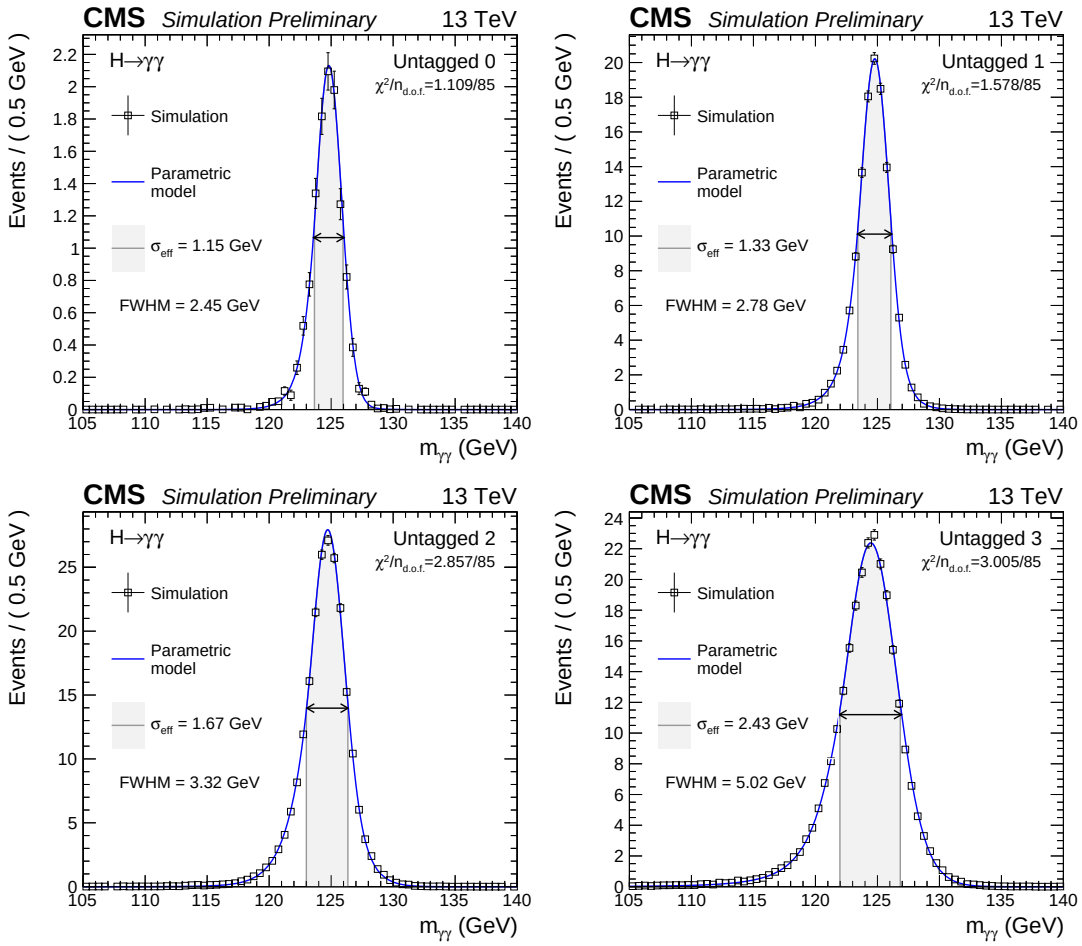


Figure 6.6: The signal models for the Untagged analysis categories for $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained by summing the contributions from each production process according to their $\epsilon \times A$. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown.

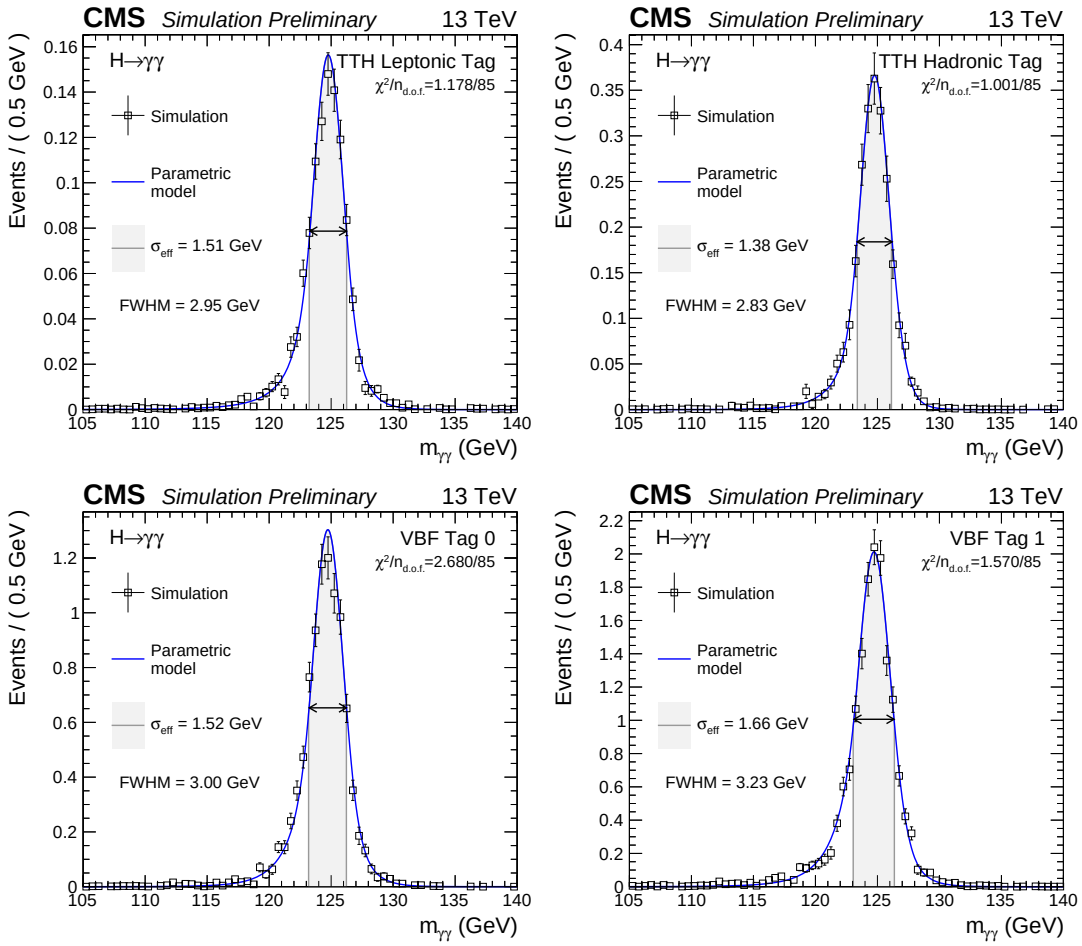


Figure 6.7: The signal models for the VBF-tagged and ttH-tagged analysis categories for $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained by summing the contributions from each production process according to their $\epsilon \times A$. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown.

uncertainty associated with making a particular choice must be accounted for. These issues are addressed using the discrete profiling method [66], which treats the choice of functional form as a discrete nuisance parameter in the final 2NLL fit to the data. This method provides a natural way to include the uncertainty in the choice, and is described below.

When making a measurement of a parameter of interest using a 2NLL minimisation, nuisance parameters representing systematic uncertainties are profiled: they are allowed to float during the minimisation, but their final value is not of interest. The additional freedom produces a wider 2NLL curve, representing the additional uncertainty attributed to the floating nuisances. The same result can be obtained in a different way: if the value of one of the nuisance parameters is instead fixed at the best-fit value, the width of the resulting 2NLL curve will be narrower but still with its minimum at the same place as the full profiled 2NLL curve. This width represents the uncertainty of the measurement without the effect of the nuisance parameter in question. If the procedure is repeated for different fixed values of the nuisance parameter, different 2NLL curves, not necessarily at the minimum, will be produced. In the limit that many different values of the fixed nuisance parameter are sampled, the minimum envelope of all the fixed-nuisance 2NLL curves will converge to the full profiled 2NLL curve. The uncertainty can then be obtained from the envelope as for a usual 2NLL curve. This procedure is illustrated in Figure 6.8, and can also be applied for nuisance parameters for which only some discrete, particular values are possible. When parametrising the background distribution, the chosen functional form can be treated as a such a discrete nuisance parameter. The method described above can then be used to account for the uncertainty associated with the parametrisation of the background. The fact that different functional forms can have different numbers of degrees of freedom is taken into account by adding a penalty term to the 2NLL proportional to the number of parameters in the function.

6.2.2 Application in the $H \to \gamma\gamma$ analysis

In theory, a complete set of all analytic functions should be considered to obtain an exact result using the discrete profiling method. In practice, it is only necessary to include the subset of all analytic function which give a good description of the data. In this analysis, where the background distribution is smoothly falling, the following four families of functions are considered:

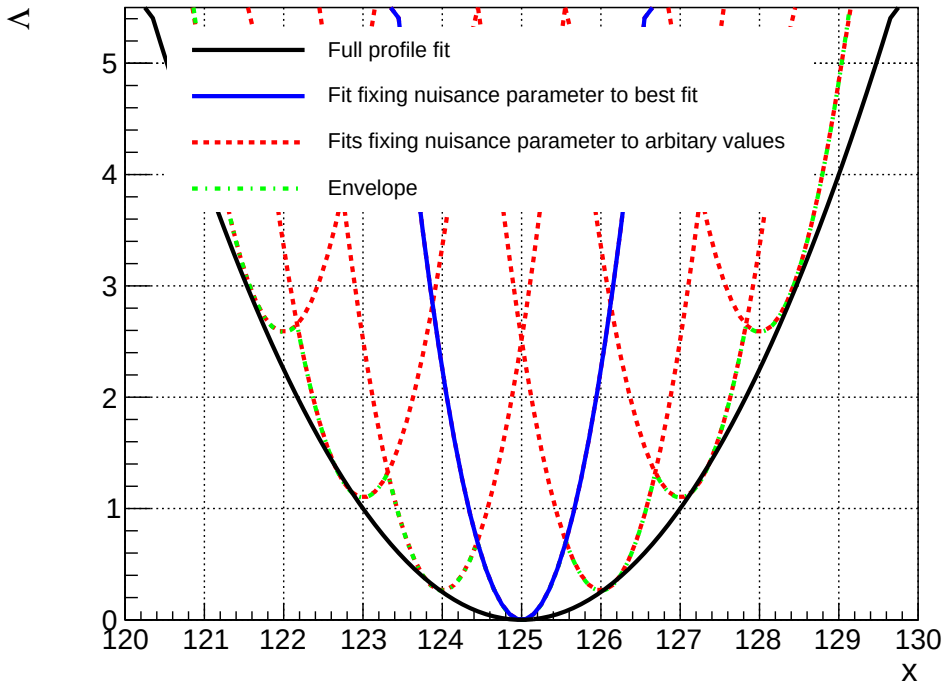


Figure 6.8: An illustration of the construction of the envelope to estimate the effect of a nuisance parameter. The NLL (denoted as Λ) curve obtained when performing a likelihood scan of parameter of interest x if the nuisance parameter is profiled is shown in black. The NLL curve obtained by fixing the nuisance to the best-fit value is shown in blue. The NLL curves for various fixed values of the nuisance other than the best-fit are shown in red. The minimum envelope of these curves, shown in green, approximates the original NLL curve obtained by profiling the nuisance parameter [66].

- sums of exponentials,

$$f_N(x) = \sum_{i=1}^{N} p_{2i-} e^{p_{2i} x};$$

- sums of polynomials (in the Bernstein basis),

$$f_N(x) = \sum_{i=0}^{N} p_i b_{(i,N)}, \text{ where } b_{(i,N)} := \binom{N}{i} x^i (1-x)^{N-i};$$

- Laurent series,

$$f_N(x) = \sum_{i=1}^{N} p_i x^{-4+\sum_{j=1}^{i} (-1)^j (j-1)};$$

- sums of power-law functions:

$$f_N(x) = \sum_{i=1}^{N} p_{2i-1} x^{-p_{2i}};$$

where for all k, the p_k are a set of parameters, and N represents the order of a particular function in the family. In order to keep the computing time required for the fitting to a tolerable level, only a subset of functions from each family are considered. The maximum order of the candidate function considered from each family is obtained separately for each analysis category using the following procedure. Starting with the lowest-order function in the family, the parameters of the candidate function are varied to minimize the 2NLL with respect to the $m_{\gamma\gamma}$ distribution. A penalty of 1 times the number of parameters in the functional form is added to the value of the 2NLL to account for differences in the number of degrees of freedom. The same procedure is applied to the function of next-highest order in the family. In the limit of large sample size, the difference in the minimum 2NLL between functions of successive orders N and $N+1$, $2\Delta NLL_{N+1} = 2(NLL_{N+1} - NLL_N)$, is distributed as a χ^2 with M degrees of freedom where, M is the difference in the number of free parameters in the order-$(N+1)$ and order-N functions. A p-value is then calculated as:

$$p\text{-value} = \text{prob}(2\Delta NLL > 2\Delta NLL_{N+1} | \chi^2(M)).$$

If the p-value is less than a predetermined threshold, chosen as 0.05, the higher-order function is supported by the data, otherwise the higher-order function is assumed unnecessarily flexible given the data. In the former case, the next-highest order function is then considered, and so on. Otherwise, the procedure terminates having found the highest-order suitable function. An additional constraint is applied to remove low order functions which do not fit the data well. These would not contribute to the minimum envelope anyway, and removing them reduces the amount of time needed to perform the final minimisation described in Chapter 7. The remaining functions from each of the four families are added to the final set of candidate functions to be used in the discrete profiling. As an example, the chosen functions are shown for the Untagged categories are shown in Figure 6.9. The equivalent figures for the other analysis categories are available in Appendixapp:modelling in Figure A.11. For each category, the candidate functions give acceptable agreement with the data, but can lead to large variations in the predicted number of events in the region of interest between 120 and 130 GeV.

A series of tests demonstrated that this method provides good coverage of the uncertainty associated with the choice of the function and provides an unbiased estimate of the signal strength. These tests are described in detail in [66].

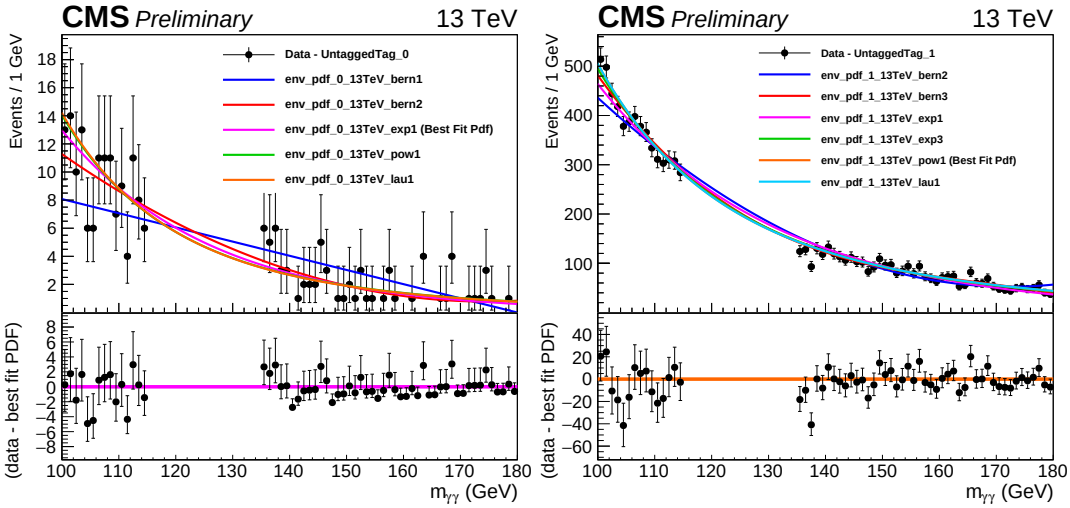
6.3 Systematic uncertainties

The systematic uncertainty on the choice of the background fitting function is handled directly by the discrete profiling method described in Section 6.2. The uncertainties which affect the signal model are more numerous, and are implemented according to their effect.

Systematic uncertainties which affect the shape of the $m_{\gamma\gamma}$ distribution are built directly into the signal models described in Section 6.1. Most systematics of this type impact individual photon energies, and therefore the invariant mass. The effect of the systematic variation is propagated to the mean, σ_{eff} and normalisation of the signal $m_{\gamma\gamma}$ distribution for each category and signal process. Corresponding nuisance parameters are then inserted which can modify the normalisation, mean and width of the DCB+1G parametrisations. An exception is the nuisance corresponding to the vertex efficiency, which instead modifies the relative mixing fraction of the RV and WV components of the model. The nuisance parameters are then Gaussian-constrained and allowed to be profiled in the final 2NLL minimisation when making measurements of the parameters of interest. Nuisances of this type are referred to hereafter as *shape nuisances*.

Uncertainties on selection efficiencies do not affect the shape of the $m_{\gamma\gamma}$ distribution, but do change the final event count (or *yield*) of the signal model for each category. Depending on the source of the systematic uncertainty, individual categories can be scaled differently, although all categories will simultaneously either increase or decrease. The corresponding nuisance parameters are profiled in the 2NLL minimisation with a log-normal [67] constraint, which cannot give rise to negative yields. Systematics of this type are referred to as *yield nuisances*, which are either symmetric or asymmetric.

Finally, some uncertainties affect the categorisation of events. In this case, the systematic variations can cause events to move from one event category to another, or altogether out of the acceptance of the analysis. These systematic uncertainties are implemented as nuisance parameters which change the relative yield of event categories, and are referred to as *category migration nuisances*. They are implemented analogously to yield nuisances, except that the yield of certain categories will increase while the yield of others



(a) Untagged 0 **(b)** Untagged 1

(c) Untagged 2 **(d)** Untagged 3

Figure 6.9: The set of candidate functions chosen to parametrise the background using the discrete profiling method in the Untagged categories. For each category, all candidate functions give acceptable agreement with the data, but can lead to large variations in the predicted number of events in the region of interest between 120 and 130 GeV. The resulting uncertainty in the choice of parametrisation is handled by the discrete profiling method.

must decrease in turn. Furthermore, separate nuisances are implemented to account for different types of migration: for instance, between the individual VBF-tagged categories on the one hand, and then between all VBF-tagged and all Untagged categories on the other.

6.3.1 Theory uncertainties

Parton Distribution Functions

The uncertainty on signal process cross-sections due to uncertainties in the parton distribution functions (PDF) produces both an overall change in signal yield and category migrations.

One symmetric yield nuisance for each signal process is included, accounting for both the uncertainties in the PDFs and the uncertainty on the strong force coupling constant (α_s). Each varies the category yields according to the corresponding uncertainty in the process cross-section, as provided by the LHCHXSWG recommendation [27]. Specifically, the ggH, VBF, WH, ZH, and ttH process cross-sections are scaled by 3.2%, 2.1%, 1.9%, 1.6% and 3.6% respectively.

The relative yield change is modelled using a series of category migration nuisances. The size of the variations is determined according to the PDF4LHC prescription [68], by re-weighting individual events according to the NNPDF30 PDF set [69]. The effect of the variations is normalised by their effect on the overall yield in each category, such that they represent only migrations. The procedure results in 60 uncorrelated symmetric category migration nuisances. The migrations are of the order of 0.5% in most categories.

QCD scale

The uncertainty on the scale of the QCD interaction is parametrised in terms of the renormalisation (μ_R) and factorisation (μ_F) scales. The asymmetric yield nuisances corresponding to variations of these parameters are taken direction from the LHCHXSWG recommendations for cross-sections [27]. The size of the effect for the ggH, VBF, WH, ZH and ttH processes is $^{+4.6\%}_{-6.7\%}$, $^{+0.4\%}_{-0.3\%}$, $^{+0.5\%}_{-0.7\%}$, $^{+3.8\%}_{-3.0\%}$ and $^{+5.8\%}_{-9.2\%}$ respectively.

Three additional asymmetric category migration nuisances are also included. The size of their effect is estimated at generator level, by varying the values of μ_R and μ_F by factors

of 2 (upward variation) or 0.5 (downward variation). The three category migration nuisances correspond to varying μ_R up or down while keeping μ_F constant ; varying μ_F up or down while keeping μ_R constant ; and varying both μ_R and μ_F up or down uniformly. In each case the migrations are found to be of the order of 1-4% of the category yield.

Strong force coupling constant

The effect of the uncertainty on the value of α_s is modelled as described in the section for the uncertainties on the PDFs. The yield nuisances for the uncertainties on the PDFs also incorporate the uncertainty on α_s. Three additional asymmetric category migration nuisances are included, the effects of which are determined according to the PDF4LHC prescription [68]. The migrations are of the order of 0.5% in most categories.

$H \to \gamma\gamma$ branching ratio

The uncertainty on the SM $H \to \gamma\gamma$ branching fraction is taken directly from [27]. It is implemented as a yield nuisance affecting all analysis categories, and the variation is of 2.08% on the category yield.

Gluon fusion contamination of VBF-tagged and ttH-tagged categories

The theoretical prediction for the jet multiplicity in gluon fusion events is unreliable for high multiplicities. This leads to the introduction of an uncertainty on the contamination of ggH events in other analysis categories which use jets in their selections.

For VBF-tagged categories, the uncertainty is estimated using the Stewart-Tackmann procedure [70]. This results in a set of category migration nuisances: a migration between VBF-tagged categories (at most 39% on category yield) and a migration between Untagged and VBF-tagged categories (at most 10% on category yield);

For the ttH-tagged categories, the equivalent effect is modelled with yield nuisances instead of categorisation migration nuisances. This is because there is no migration between the two ttH-tagged categories, and the effect of migrating events into the Untagged categories would be negligible. Three yield nuisances are considered, which represent:

- parton shower modelling uncertainty, which is estimated by comparing the jet multiplicity in data and simulation for $t\bar{t} \rightarrow$ jets events, where the tops quarks both decay to leptons, leading to variations of the order of 45% on the yield of the ttH-tagged categories;
- "gluon splitting" modelling, which is estimated from the ratio of cross-sections of ttbb and tt + 2 jet events in 13 TeV data. The size of this variation is of the order of 18% on the yield of the ttH-tagged categories;
- an additional nuisance included to account for the small size of the simulated samples used in these studies. The size of this variation is of the order of 10% on the yield of the ttH-tagged categories.

Underlying event and parton shower modelling

The *underlying event* refers to all the soft interactions between partons which occur in addition to the hard scatter in a p-p collision. In simulation, these are modelled as multiple independent scatterings between the other partons. Tuning the parameters of the modelling of the underlying event (such as the momentum cutoff for such additional interactions, and the matter density of the proton) can lead to a modified jet production cross-section.

Parton shower modelling refers to the emission of QCD radiation in the form of gluons from partons during p-p collisions. These gluons can themselves emit further QCD radiation or convert to quark-antiquark pairs, until hadronisation occurs. The modelling of these showers depends on certain parameters, such as the hadronisation scale, which is not predicted perturbatively.

Both of these uncertainties can affect the shape, number or energy of jets, and so the VBF categorisation. The sizes of the variations are obtained from dedicated simulated samples where the parameters relating to the underlying event and parton shower modelling have been tuned differently. The uncertainties are treated together as a set of category migrations nuisance between: VBF-tagged 0 and VBF-tagged 1 categories (of the order of 8%); and all VBF-tagged categories and all Untagged categories (of the order of 9%).

6.3.2 Photon uncertainties

Photon preselection

The efficiency of the photon preselection is quantified using the tag-and-probe method described in Section 4.3.3, which also provides systematic uncertainties for different photon classes. The systematic uncertainties are propagated to a yield nuisance, the effect of which is around 4% on the category yields.

Photon identification

An uncertainty on the photon identification efficiency is introduced to account for the difference in distributions of the $BDT_{\gamma\mathrm{ID}}$ output score observed between data and simulation in $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ events (see Figure 4.3). The size of this uncertainty is approximately 3% on the value of the output score. This is treated as a yield nuisance, where the uncertainty on the output score is translated to an uncertainty on the yield of the order of 3%.

Photon energy scale and resolution

After the calibration of the ECAL described in Section 3.2.3, there are still some discrepancies in the photon energy scale and resolution between simulation and data. Since electrons and photons are both reconstructed as SCs, these discrepancies can be studied using $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ events where the electrons are reconstructed as photons. For SCs in eight R_9 and $|\eta|$ classes, the invariant mass distributions in data and simulation are both fitted with a Breit-Wigner (BW) function convoluted with a Crystal Ball function (CB) function. The BW function models the natural shape of the Z-peak, while the CB function describes the ECAL resolution and losses due to unrecovered energy from bremsstrahlung. The natural width and pole mass of the Z boson are fixed to their accepted values [6] in this parametrisation.

The corrections to the photon energy scale are given by the relative differences between the best-fit means of the CB in data and simulation, divided by the Z boson pole mass, in each bin. The corrections to the photon energy resolution are applied by adding additional smearing terms to the width of the CB in quadrature. The additional scale and resolution corrections described above each have uncertainties, which are related to choices made for the $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ event selection and classification, as well as the difference

between the final electron and photon energy regression BDTs. The uncertainties on the energy scale and resolution are quantified for each photon class, and propagated to the $m_{\gamma\gamma}$ distribution in each analysis category. This results in shape nuisances for the photon energy scale in four classes (high and low R_9, each for EB and EE), and eight shape nuisances for the photon energy smearing (parametrised as constant and stochastic contributions).

The size of the systematic uncertainties are of the order of 0.15% to 0.50% depending on the photon class. The effect on the mean of the $m_{\gamma\gamma}$ distribution is at most 0.25%, while the effect on the σ_{eff} is at most 20%, depending on the analysis category.

Per-photon energy resolution

The uncertainty of the per-photon energy resolution is conservatively evaluated by scaling the output of the $BDT_{\gamma\text{E}}$ described in Section 4.3.5 by $\pm 5\%$. This uncertainty is propagated throughout the analysis and modelled as a yield nuisance. The size of the variation is typically of the order of 2% depending on the analysis category.

Non-linearity of detector response

The uncertainty associated with the fact that the ECAL response is not linear is estimated by comparing boosted $\text{Z} \to \text{e}^+\text{e}^-$ decays in data and simulation. Individual photon energies are affected by up to 0.2%. The effect is propagated as a shape nuisance, which varies the mean of the $m_{\gamma\gamma}$ distribution by 0.1% in each category.

Shower shape corrections

The uncertainty deriving from the imperfect modelling of shower shape variables is estimated using simulated samples with and without the corrections. This effect is of order 0.06% on the photon energy scale, and is implemented as four shape nuisances for photons in different η and R_9 classes, which vary the $m_{\gamma\gamma}$ distribution mean by at most 0.20% and σ_{eff} by at most 1.80%.

Non-uniformity of the light collection

The uncertainty on the response of the ECAL crystals depending on their position in η is modelled separately as shape nuisances for photons in the barrel and in the endcaps. The size of the uncertainty is 0.07% on the photon energies. The effect is propagated to the $m_{\gamma\gamma}$ distribution of each category and applied as separate shape nuisances in the EE and EB, which vary the mean by up to 0.2% and the σ_{eff} by up to 5%.

Modelling of detector response in `GEANT 4`

Imperfect modelling of the differences between electromagnetic showers for electrons and photons in the detector simulation software `GEANT 4` may have a small impact on the photon energy scale. The size of the variation is determined with a dedicated simulated sample where the parameters of shower modelling are modified. The size of the effect is found to be consistent with zero, but an upper bound of approximately 0.05% is applied on the uncertainty on the mean of the $m_{\gamma\gamma}$ distribution in each category.

Modelling of the material budget

The imperfect modelling of the amount of material between the vertex and the ECAL affects the simulation of the photon and electron showers. The uncertainty related to this effect is estimated with dedicated samples where the amount of simulated material is uniformly varied by $\pm\,5\%$. It is treated as two separate shape nuisances, for EB and EE photons separately, which affect the mean of the $m_{\gamma\gamma}$ distribution by at most 0.1% (0.03%) and the σ_{eff} by at most 3.5% (7.6%) for the EB (EE) photons.

Electron veto

The electron veto element of the photon preselection is validated separately using $\text{Z} \to \mu\mu\gamma$ events. A small uncertainty is introduced to account for the discrepancy between data and simulation. This source of uncertainty is treated as a yield nuisance, the size of which is of the order of 0.5%.

6.3.3 Per-event uncertainties

Integrated luminosity

The uncertainty on the value of the integrated luminosity of the data sample is modelled as a scale nuisance, the size of which is 6.2% on the yield of all signal processes.

Jet energy scale and resolution

The uncertainties on the jet energy scale are described by category migration nuisances: between VBF-tagged 0 and VBF-tagged 1 (typically of order 3% on the category yield); between all VBF and Untagged categories (of order 10%) ; and between ttH-tagged and Untagged categories (of order 12%).

The nuisances for the jet energy resolution are treated analogously, with migrations of order 0.5% between VBF-tagged 0 and VBF-tagged 1 categories; of order 1.5% between VBF and Untagged categories ; and of order 4% between ttH-tagged and Untagged categories.

Diphoton selection

The $BDT_{\gamma\gamma}$ output score is used to make a selections on the diphotons which enter the analysis. A small uncertainty is introduced to account for the residual discrepancy between data and simulation in $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ events after the uncertainties from the per-photon energy resolutions estimate and the $BDT_{\gamma\mathrm{ID}}$ scores are taken into account. This source of uncertainty is treated as a yield nuisance, the size of which is of 0.2% in all categories.

Vertex-finding efficiency

The uncertainty on the vertex-finding efficiency is determined by comparing the fraction of correctly identified vertices in $\mathrm{Z} \to \mu^{-}\mu^{+}$ between data and simulation. The uncertainty is modelled as a shape nuisance which alters the RV/WV mixing fraction of each signal model. The variation is of the order of 1.5% on the RV fraction.

Trigger efficiency

The uncertainty on the trigger efficiency is estimated using a tag-and-probe method as described in Section 4.2.2. The uncertainty is applied as a yield nuisance parameter of size 0.1%.

Lepton reconstruction and b-tagging efficiencies

The uncertainty on the reconstruction of electrons and muons is determined by considering the ratio of leptons reconstructed in data and simulation. The result is implemented as a yield uncertainty for the ttH-tagged categories of the order of 1% for electrons and 5% for muons. Similarly, the uncertainty on the tagging of b jets is evaluated by varying the ratio between the measured b-tagging efficiency in data and simulation within their uncertainty. This is propagated to a yield nuisance, which has an effect of the order of 2% for ttH-tagged categories.

Rejection of jets from pileup

The uncertainty on pileup jet rejection using the selection of the σ_{RMS} (as described in Section 4.5.3) is described by a category migration nuisance between all VBF and Untagged categories of order 4%.

A further uncertainty, relating to the number of pileup jets in simulation which are reconstructed as genuine jets from the hard scatter, is described by category migration nuisances: between VBF-tagged 0 and VBF-tagged 1 (of order 1% on the category yield); and between all VBF and Untagged categories (of order 0.5%).

6.4 Signal and background modelling summary

The results of the signal and background modelling for a signal with $m_{\text{H}} = 125\,\text{GeV}$ are shown in Table 6.1. The expected number of signal events for each category is broken down by the contribution from each production mode. The σ_{eff} (half the width of the smallest window containing 68.3% of the distribution) and σ_{HM} (the width of the distribution at half of the maximum value, divided by 2.35) are also shown. The expected number of background events predicted by the best-fit function in a $\pm 1\,\sigma_{\text{eff}}$ window around 125 GeV is shown for each analysis category.

The information from Table 6.1 is presented in a visual format in Figure 6.10. This clearly shows the categories preferentially selecting events from the targeted processes. Also shown is the signal-to-background ratio ($S/(S+B)$) in each analysis category, illustrating that the Untagged and VBF-tagged categories are ordered by $S/(S+B)$.

The effect of each of the sources of systematic uncertainty listed above is summarised in Table 6.2. The effect is quoted as the contribution to the expected relative uncertainty on the measurement of the signal strength (either total or by production mode).

Event Categories	SM 125GeV Higgs boson expected signal								Bkg Events
	Total	ggH (%)	VBF (%)	WH (%)	ZH (%)	ttH (%)	σ_{eff}	σ_{HM}	in $\pm\sigma_{eff}$
Untagged Tag 0	11.92	79.10	7.60	7.11	3.59	2.60	1.15	1.04	11.49
Untagged Tag 1	128.78	85.98	7.38	3.70	2.12	0.82	1.33	1.18	530.11
Untagged Tag 2	220.12	91.11	5.01	2.18	1.23	0.47	1.67	1.41	2244.59
Untagged Tag 3	258.50	92.35	4.23	1.89	1.06	0.47	2.43	2.14	9060.32
VBF Tag 0	9.35	29.47	69.97	0.29	0.07	0.20	1.52	1.28	9.42
VBF Tag 1	15.55	44.91	53.50	0.86	0.38	0.35	1.66	1.37	73.76
TTH Hadronic Tag	2.42	16.78	1.28	2.52	2.39	77.02	1.38	1.20	3.09
TTH Leptonic Tag	1.12	1.09	0.08	2.43	1.06	95.34	1.51	1.26	1.27
Total	647.77	87.93	7.29	2.40	1.35	1.03	1.86	1.50	10265.82

Table 6.1: The expected number of signal and background events per category. The σ_{eff} of the signal model is also provided as an estimate of the $m_{\gamma\gamma}$ resolution in that category. The expected number of background events in a $\pm 1\,\sigma_{\text{eff}}$ window around 125 GeV is also quoted.

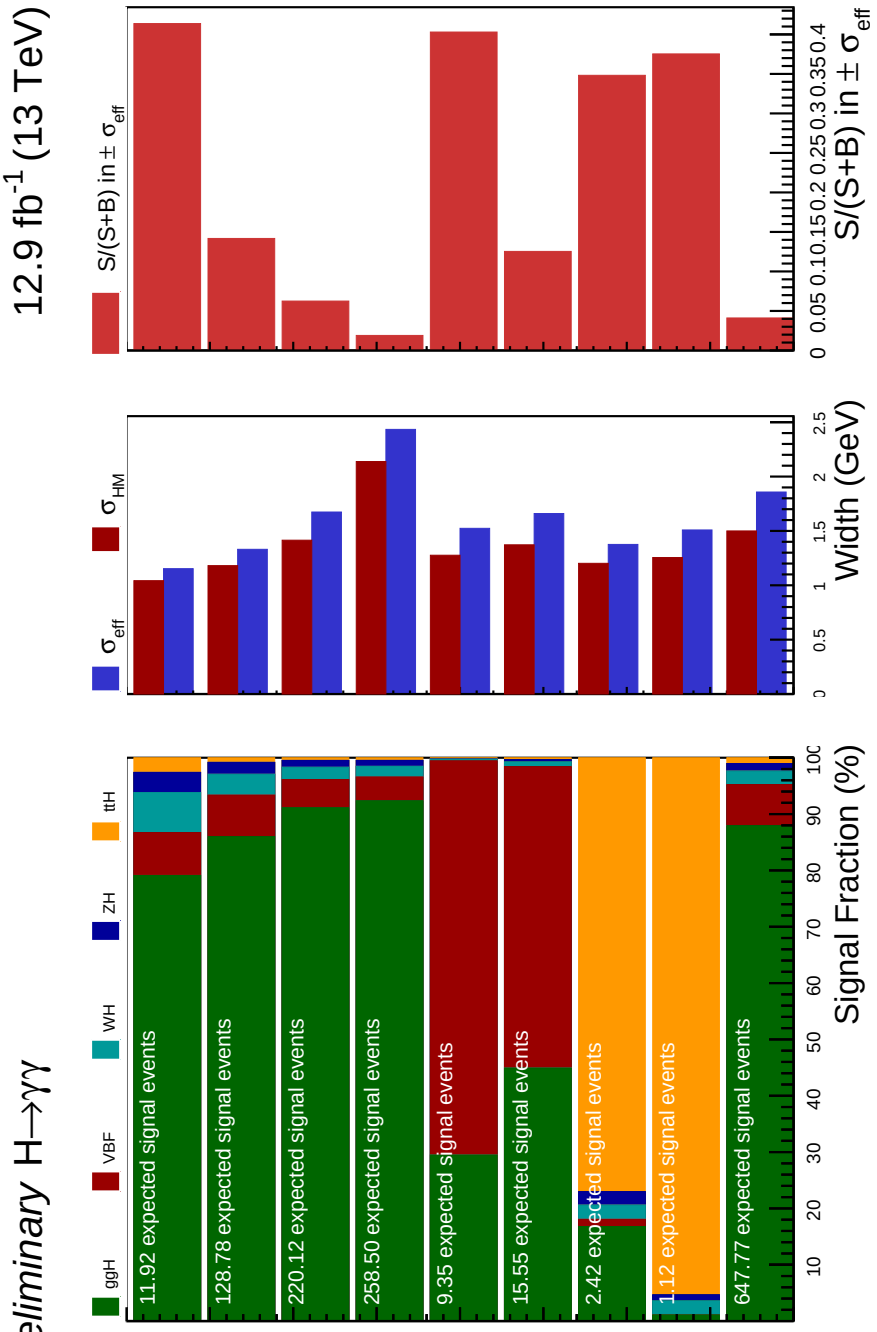


Figure 6.10: The signal composition of the analysis categories in terms of the Higgs boson production modes. The σ_{eff} and σ_{HM} are also shown, as is the expected signal to background ratio in a $\pm\,\sigma_{\mathrm{eff}}$ window around $m_{\mathrm{H}} = 125\,\mathrm{GeV}$.

Expected relative uncertainty for SM Higgs boson ($m_{\mathrm{H}} = 125GeV$)				
Systematic	μ	μ_{ggH}	μ_{qqH}	μ_{ttH}
Integrated luminosity	6.32 %	6.55 %	9.74 %	9.18 %
QCD scale yield	5.26 %	6.89 %	1.07 %	12.96 %
Photon preselection	4.13 %	4.16 %	6.49 %	5.10 %
Photon identification	4.03 %	4.49 %	1.32 %	5.40 %
Photon energy scale and smearing	2.79 %	2.85 %	10.64 %	8.11 %
QCD scale migrations	2.72 %	2.23 %	14.71 %	1.03 %
PDF and alphaS yield	2.53 %	3.30 %	3.58 %	4.99 %
Branching ratio	2.19 %	2.25 %	3.78 %	2.33 %
Per photon energy resolution estimate	1.64 %	0.55 %	3.53 %	2.12 %
AlphaS migrations	1.56 %	2.36 %	3.57 %	0.58 %
PDF migrations	1.24 %	1.56 %	2.90 %	1.44 %
Jet energy scale and resolution	1.17 %	2.36 %	21.62 %	7.02 %
Underlying event and parton shower	1.16 %	2.19 %	17.53 %	1.37 %
Vertex finding efficiency	0.60 %	0.33 %	3.93 %	0.99 %
Modelling of material budget	0.55 %	1.66 %	3.05 %	2.24 %
ggF contamination in VBF categories	0.51 %	1.52 %	22.58 %	1.12 %
Electron veto	0.45 %	0.43 %	0.18 %	0.30 %
Trigger efficiency	0.37 %	1.27 %	0.60 %	0.89 %
Unmatched pileup	0.31 %	1.54 %	0.70 %	1.55 %
Lepton reconstruction and btag efficiencies	0.30 %	1.55 %	1.40 %	1.55 %
Diphoton selection	0.27 %	0.14 %	2.10 %	1.63 %
ggF contamination in ttH categories	0.27 %	1.46 %	2.74 %	7.34 %
Modelling of detector response in GEANT4	0.24 %	0.06 %	0.23 %	0.94 %
Nonlinearity of detector response	0.19 %	0.05 %	2.41 %	1.26 %
Shower shape corrections	0.17 %	1.56 %	2.68 %	2.30 %
Nonuniformity of light collection	0.16 %	1.26 %	3.10 %	2.19 %

Table 6.2: The contribution to the expected relative uncertainty on the measurement of the signal strength for a SM Higgs boson. The effect is quoted for the overall signal strength and also for the individual signal strengths of each production mode.

Chapter 7

Statistical analysis and results

The signal and background models described in Chapter 6 are used to perform the statistical interpretation of $12.9\,\text{fb}^{-1}$ of Run 2 data collected at $\sqrt{s} = 13\,\text{TeV}$ by CMS in 2016. The author was responsible for the signal and background modelling, systematics handling and statistical interpretation in the official CMS preliminary result [2] produced with the same dataset. The results presented in this section differ from those presented in [2] insofar as they use improvements to the signal modelling techniques developed by the author. Namely the DCB+1G functional form was used instead of a sum of Gaussians, and interpolation using SSF was used instead of linear interpolation. The results from this thesis and [2] are almost identical, as expected, and are compared at various points throughout this chapter.

Due to the scaling of the SM Higgs boson and background cross-sections moving from 7 or 8 TeV to 13 TeV, the dataset used in this thesis has roughly the same statistical sensitivity as the full Run 1 dataset, which had an integrated luminosity of $24.8\,\text{fb}^{-1}$. The final analysis of the Run 1 dataset [5] led to an independent observation of the Higgs boson in the diphoton decay channel. The objective of the work presented in this thesis is therefore to make a confirmation of the observation (or "rediscovery") of the Higgs boson in the diphoton decay channel, and measurements of some of its properties.

The statistical analysis proceeds in several stages. The first step is to determine the best signal-plus-background fit of the models to the data. The floating parameters of the signal and background models are varied simultaneously in each analysis category to obtain the closest overall agreement with the observed invariant mass distributions. This procedure is described in Section 7.1, and leads to a Higgs boson signal being measured in the data. The next step is to quantify the significance of this signal and formally reject the hypothesis that there is no Higgs boson. The procedure for hypothesis testing

and calculation of the significance is detailed in Section 7.2. Having established the existence of a SM-like Higgs boson which decays to photons, the final step is to make measurements of some of its properties, in particular those which relate to the rate at which it is produced or interacts with other SM particles. These measurements are described in Section 7.3 and Section 7.4.

As with all Higgs boson analyses performed within the CMS collaboration, the statistical interpretation uses the frequentist approach. The corresponding statistical tools and techniques are briefly explained as they are needed throughout this chapter.

7.1 Best fit of models to the data

The observed $m_{\gamma\gamma}$ distributions obtained from the data (labelled $m_{\gamma\gamma}^{obs}$ hereafter) are parametrised with models consisting of signal and background components. The tool which is used determine the model parameters and to assess the agreement with the data in a category C is called the *likelihood function*,

$$\mathcal{L}_C(\mu,\, m_{\rm H}; \mathbf{n}|\, m_{\gamma\gamma}^{obs,C}) = \mu \cdot f_S^C(\, m_{\gamma\gamma}^{obs,C}|\, m_{\rm H}; \mathbf{n}_S) + f_B^C(\, m_{\gamma\gamma}^{obs,C}|\mathbf{n}_B). \tag{7.1}$$

In the definition above:

- $\mu = \sigma^H/\sigma_{SM}^H$ is a parameter of interest (POI) called the signal strength, where in this context σ^H is the observed Higgs boson cross section, and σ_{SM}^H is the SM Higgs boson cross-section. The mass of the Higgs boson $m_{\rm H}$ is a second POI. The likelihood function can in fact be generalised for an arbitrary number of POIs (as in Sections 7.3 and 7.4);
- $\mathbf{n}$ is a set of floating nuisance parameters, composed of those which affect the signal model and those which affect the background model, labelled $\mathbf{n}_S$ and $\mathbf{n}_B$ respectively;
- $m_{\gamma\gamma}^{obs,C}$ is the invariant mass distribution for a particular category C;
- f_S^C and f_B^C represent the probability distribution functions of the signal and background components of the model in category C, normalised to the expected number of signal and background events in that category respectively. Some of the nuisances parameters can modify the normalisation of f_S^C and f_B^C.

The construction of the signal component of the model in each category, f_S^C, was described in Section 6.1, and is comprised of the individual models for each Higgs boson production process, each normalised to their respective expected number of events:

$$\begin{aligned} f_S^C(\, m_{\gamma\gamma}^{obs,C} |\, m_{\mathrm{H}}; \mathbf{n}_S) = f_{S,\mathrm{ggH}}^C(\, m_{\gamma\gamma}^{obs,C} |\, m_{\mathrm{H}}; \mathbf{n}_S) + f_{S,\mathrm{VBF}}^C(\, m_{\gamma\gamma}^{obs,C} |\, m_{\mathrm{H}}; \mathbf{n}_S) \\ + f_{S,\mathrm{VH}}^C(\, m_{\gamma\gamma}^{obs,C} |\, m_{\mathrm{H}}; \mathbf{n}_S) + f_{S,\mathrm{ttH}}^C(\, m_{\gamma\gamma}^{obs,C} |\, m_{\mathrm{H}}; \mathbf{n}_S). \end{aligned} \tag{7.2}$$

When performing the signal-plus-background fit to the data, the values of the individual parameters of the signal model functional form are fixed. Specifically, since the SSF method was used, this refers to the coefficients of the seventeen polynomial functions which describe the dependence on m_{H} of the DCB+1G parameters for both the RV and WV scenarios and their mixing fraction. The parameters which are allowed to vary are the nuisance parameters $\mathbf{n}_S$ introduced into the signal modelling to account for systematic uncertainties (see Section 6.3), as well as the POIs μ and m_{H}. The signal strength μ uniformly scales all the signal models for each process and for each category. This means that the contribution to the overall signal model from each process remains in proportion to what is predicted by the SM, but the normalisation of the overall signal model can be varied.

The handling of the background component of the model f_B^C, for a given category, was described in Section 6.2. The nuisance parameters $\mathbf{n}_B$ affecting the background model are composed of:

- the discrete nuisance parameter in each category which corresponds to the choice of background function, as prescribed by the discrete profiling method (see Section 6.2.1);
- the individual parameters of all the candidate functions, in all categories, which are allowed to float in the fit.

The overall likelihood function $\mathcal{L}$ for the simultaneous fit of all categories at once is obtained by taking the product of the likelihood functions $\mathcal{L}_C$ in each analysis category, taking care to correlate the corresponding nuisance parameters in each one:

$$\mathcal{L}(\mu,\, m_{\mathrm{H}}; \mathbf{n} |\, m_{\gamma\gamma}^{obs}) = \prod_C \mathcal{L}_C(\mu,\, m_{\mathrm{H}}; \mathbf{n} |\, m_{\gamma\gamma}^{obs,C}). \tag{7.3}$$

The fit is obtained by minimising twice the negative log-likelihood (2NLL). The best-fit values of the POIs and nuisance parameters are denoted as $\hat{\mu}$, $\hat{m}_{\mathrm{H}}$ and $\hat{\mathbf{n}}$, such that:

$$\{(\mu,\, m_{\mathrm{H}}, \mathbf{n}) : -2\ln\mathcal{L}(\hat{\mu},\, \hat{m}_{\mathrm{H}}; \hat{\mathbf{n}}|\, m_{\gamma\gamma}^{obs}) = \min_{\mu,\, m_{\mathrm{H}}, \mathbf{n}}(-2\ln\mathcal{L}(\mu,\, m_{\mathrm{H}}; \mathbf{n}|\, m_{\gamma\gamma}^{obs}))\}, \tag{7.4}$$

where $\min_{\mu,\, m_{\mathrm{H}}, \mathbf{n}}$ represents the global minimum evaluated over all allowed values of μ, m_{H} and $\mathbf{n}$, taking into account their constraints. The parameters in this analysis are constrainted in one of three possible ways: for the systematic uncertainties described by shape nuisances, a Gaussian constraint is applied ; for the systematic uncertainties described by yield or migration nuisances, a log-normal constrain is applied ; for all other nuisances, the constraint is flat.

In general, the 2NLL cannot be minimized analytically. Instead, this task is handled numerically with the `Minuit2` minimizer [71], as part of the `RooFit` software package [72]. The resulting parametrisations of the observed data are shown for each analysis category separately in Figures 7.1 and 7.2. The combined parametrisation and data resulting from a direct sum of each category is shown in Figure 7.3a, while the sum weighted by the $S/(S+B)$ in $\pm\,\sigma_{\mathrm{eff}}$ around the best-fit m_{H} in each category is shown in Figure 7.3b.

The best-fit values of the POIs are found to be $\hat{\mu} = 0.94$ and $\hat{m}_{\mathrm{H}} = 125.9\,\mathrm{GeV}$ (the best-fit values in [2] were 0.95 and 126.0 respectively). The best-fit parametrisation qualitatively indicates the presence of an excess of data in the same place in the $m_{\gamma\gamma}$ distributions in all analysis categories. When the categories are summed, particularly when weighted by their $S/(S+B)$, the presence of a narrow resonance above the falling background spectrum is visible by eye. The best-fit of the signal-plus-background model to the observed data therefore suggests a SM-like Higgs boson signal in the data, although a full statistical assessment needs to be undertaken to quantify the size of the excess, taking into account the sources of systematic uncertainty which enter the analysis.

7.2 Significance of observation

Given the best-fit value of the signal strength $\hat{\mu} = 0.94$ determined in Section 7.1, a frequentist approach is used to determine the degree of certainty with which the null hypothesis (that there is no Higgs boson) can be rejected in favour of an alternative hypothesis (that a SM-like Higgs boson exists). The hypotheses can be formulated in

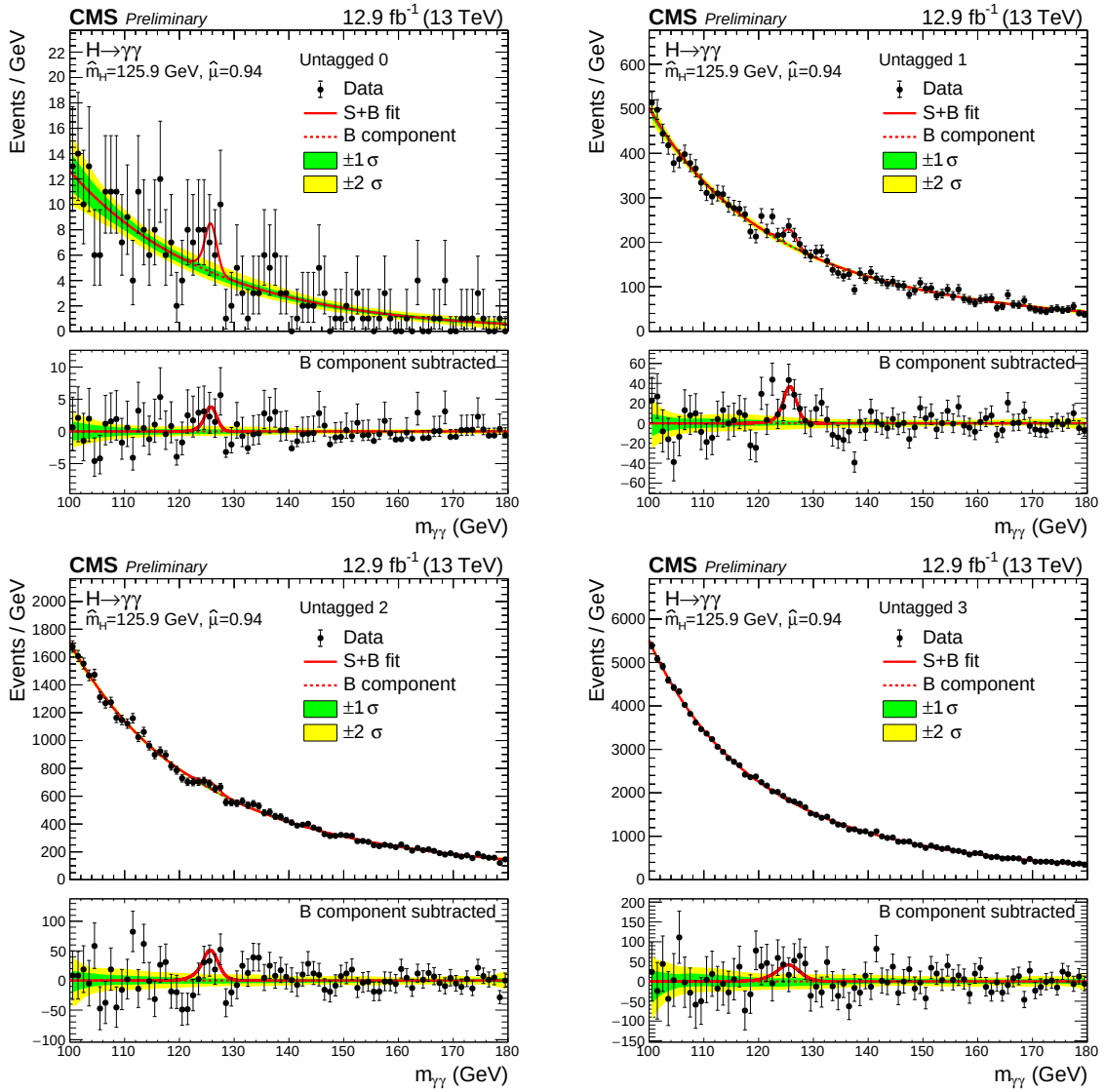


Figure 7.1: The signal-plus-background fit (solid red line) of the observed $m_{\gamma\gamma}$ distribution in data (black points) for the Untagged analysis categories. The background-only fit is shown as a dashed red line, while the green and yellow bands denote the 1σ and 2σ uncertainties on the background shape respectively.

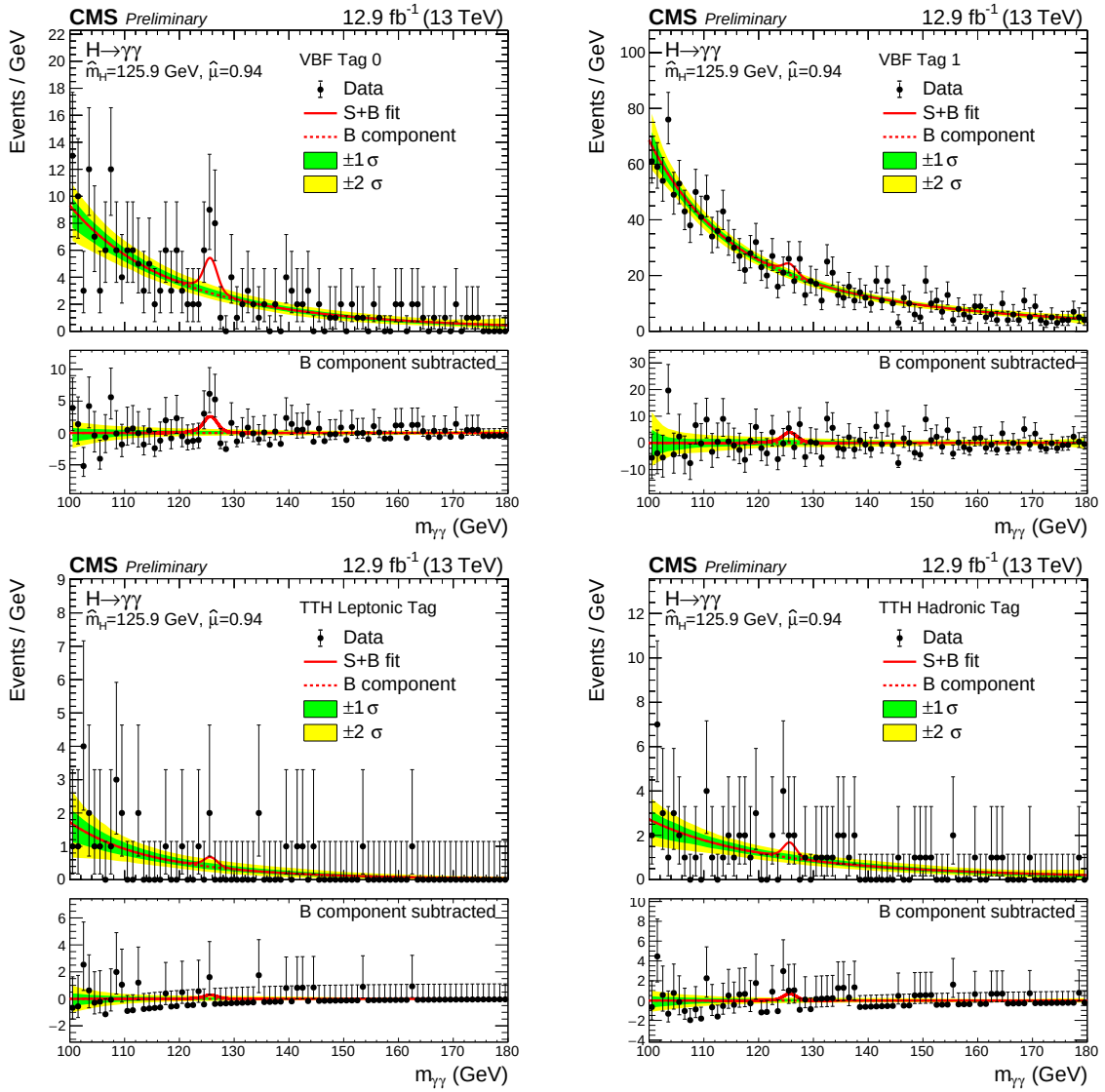
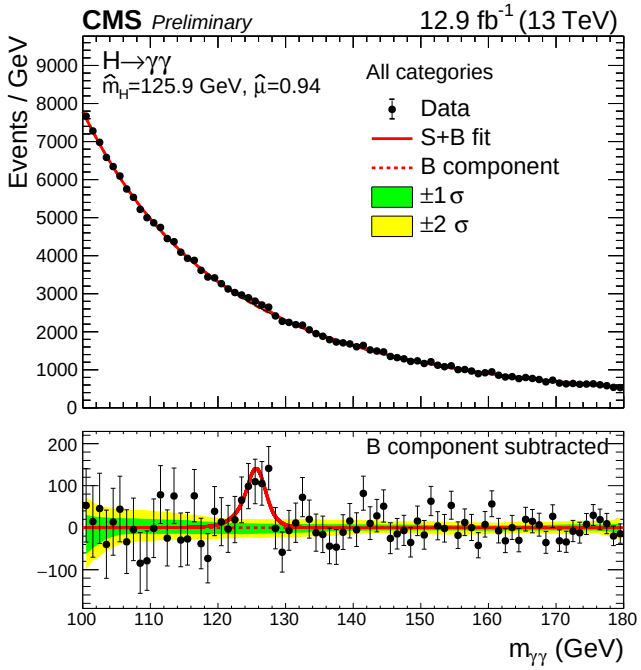
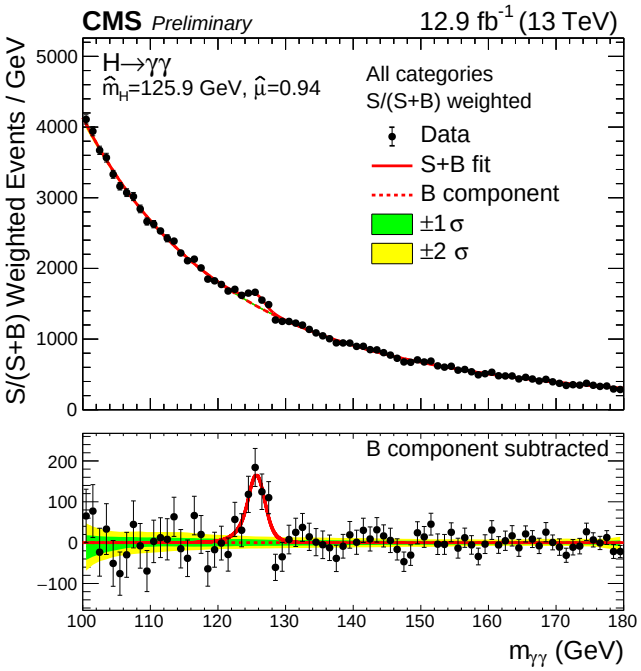


Figure 7.2: The signal-plus-background fit (solid red line) of the observed $m_{\gamma\gamma}$ distribution in data (black points) for the VBF-tagged and ttH-tagged analysis categories. The background-only fit is shown as a dashed red line, while the green and yellow bands denote the 1σ and 2σ uncertainties on the background shape respectively.



(a) Direct sum



(b) S/(S+B) weighted sum

Figure 7.3: The signal-plus-background fit (solid red line) of the observed $m_{\gamma\gamma}$ distribution in data (black points) for all categories combined, either using a direct sum (a) or a sum weighted by the $S/(S+B)$ in $\pm 1\,\sigma_{\text{eff}}$ around the best-fit value of m_{H} (b). The background-only fit is shown as a dashed red line, while the green and yellow bands denote the 1σ and 2σ uncertainties on the background shape respectively.

terms of the signal strength: the null hypothesis H_0 corresponds to the case where $\mu = 0$, while the alternative hypothesis H_μ corresponds to $\mu > 0$.

A statistical test is constructed by specifying a critical region w of the data space, such that for a given set of observed data $m_{\gamma\gamma}^{obs}$:

$$P(\, m_{\gamma\gamma}^{obs} \in w | H_0) \leq \alpha, \tag{7.5}$$

where $P(\, m_{\gamma\gamma}^{obs} \in w | H_0)$ is the probability (assuming that H_0 is correct), of observing the data inside the critical region w, and α is a small predetermined threshold [73].

The statistical power β of the test is the probability of accepting H_0 when it is false and the alternative H_μ is true. This is given by:

$$P(\, m_{\gamma\gamma}^{obs} \in w | H_\mu) = 1 - \beta. \tag{7.6}$$

The critical region is be chosen such that the power β of the test is maximised for a given α, to ensure that if $m_{\gamma\gamma}^{obs} \in w$, then H_0 has a low probability of being true while H_μ has a high probability of being true.

A common choice, which is found to maximise β [73], is to define the critical region in terms of a test statistic q_μ, corresponding to the difference between the best-fit 2NLL and the 2NLL (abbreviated as 2ΔNLL) evaluated for a particular μ. The test statistic is defined explicitly as:

$$q_\mu = \begin{cases} -2 \ln \mathcal{L}(\mu,\, m_{\mathrm{H}}; \hat{\mathbf{n}}^\mu |\, m_{\gamma\gamma}^{obs}) - 2 \ln \mathcal{L}(\hat{\mu},\, m_{\mathrm{H}}; \hat{\mathbf{n}} |\, m_{\gamma\gamma}^{obs}) & \text{when } \hat{\mu} \geq 0, \\ 0 & \text{when } \hat{\mu} < 0, \end{cases} \tag{7.7}$$

where $\hat{\mathbf{n}}^\mu$ denotes the best fit $\mathbf{n}$ for a fixed value of μ. When trying to exclude hypothesis H_0, the test statistic q_0 in particular is used to define a critical region. In the limit of a large sample of data, the probability distribution function of the test statistic (f_q), is Gaussian. The fact that $q_0 = 0$ for $\hat{\mu} < 0$ reflects the fact that only excesses in the data are regarded as significant. This simplifies the definition of the critical region, since increasingly large values of q_0 indicate increasing incompatibility with H_0, and therefore only the right-hand tail of f_q is considered when assessing probabilities. Assuming H_0, the probability of obtaining a value of q_0^{obs} (corresponding to observed data $m_{\gamma\gamma}^{obs}$) or higher is given by the integral of f_q from q_0^{obs} to infinity. This probability is commonly

referred to as the p-value. We can therefore define the critical region as:

$$w = \{ m_{\gamma\gamma}^{obs} : \int_{q_0^{obs}}^{+\infty} f_q(q_0) dq_0 \leq \alpha \}, \tag{7.8}$$

In particle physics experiments, the threshold α to reject the null hypothesis is typically 2.87×10^{-7}. If expressed as the number of standard deviations that a Gaussian-distributed variable would fluctuate to give the same p-value, then this threshold is 5σ.

The test statistic q_μ in Equation 7.7 is implicitly defined for a particular assumption on the value of m_{H}. This ensures that only excesses compatible with the Higgs boson signal distribution for that particular value of m_{H} are regarded as significant. In other words, only localised excesses in the $m_{\gamma\gamma}$ spectrum will lead to a small p-value. Thus, in this context we refer to local p-values, which represent the probability that a statistical fluctuation in the observed background distribution gave rise to a localised excess consistent with the signal model at the assumed m_{H}. The definition of H_μ thus also depends on the assumed value of m_{H}. In particular, H_μ is the hypothesis that there exists a Higgs boson with mass m_{H} and signal strength μ. If the observed data fall in the critical region for a given value of m_{H}, then the null hypothesis (that there is no Higgs boson, regardless of its mass) is rejected in favour of H_μ for that particular m_{H}. This is not, however, the same as saying that the Higgs boson has that particular value of m_{H}. The correct statement is that the alternative hypothesis, assuming m_{H}, is more likely than H_0 given the data, and that H_μ assuming a different m_{H} could be yet more likely.

The local p-value is therefore evaluated separately, given the observed data, for different assumptions about the value of m_{H} in the range 120-130 GeV in 0.1 GeV steps. The result is shown in Figure 7.4. The black solid line represents the local p-value scan for the observed data. The dashed lines represent the expected local p-values for a SM Higgs boson. These are obtained by generating an Asimov dataset [65] from the best-fit background-only model and a signal of strength $\mu = 1$, and then performing a signal-plus-background fit and following the same procedure as for observed data. For the blue dashed line, the signal was injected at $m_{\mathrm{H}} = 125.09$ GeV (the best fit value from the previous combined Run 1 measurement by CMS and ATLAS [26]), while for the red dashed line the signal was injected at the corresponding m_{H} for each step.

The local observed significance at the Run 1 best fit ($m_{\mathrm{H}} = 125.09$ GeV) is 5.7σ, where 6.3σ was expected for the SM Higgs boson. The maximum local observed significance is found at 125.9 GeV, corresponding to 6.1σ where 6.3σ was expected (these results are

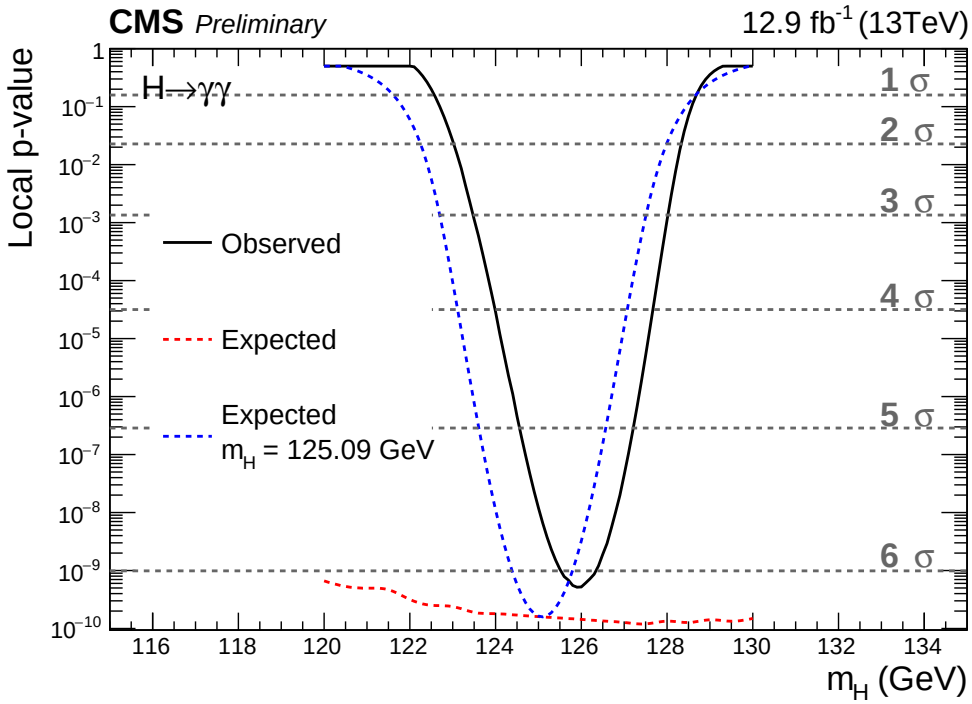


Figure 7.4: The local p-value for the observation as a function of the Higgs boson mass (black), shown with the expected local p-values for a SM Higgs boson, across the range 120-130 GeV. The expected local p-values are obtained using Asimov datasets. The blue dashed line shows the expected local p-value when the mass of the injected signal is $m_{\mathrm{H}} = 125.09\,\mathrm{GeV}$, while the red line shows the maximum significance for any injected signal in the range of 120 to 130 GeV.

consistent with [2] to within one unit of the smallest quoted decimal place). Since the observed data fall in the critical region where at least one m_{H} assumption yields a local significance above 5σ (i.e p-value is less than 2.87×10^{-7}), the null hypothesis that there is no Higgs boson is rejected in favour of the alternative hypothesis that there exists a Higgs boson. Therefore, the data correspond to an observation of the Higgs boson decaying to photons.

The maximum significance of the observation in this analysis occurs near $m_{\mathrm{H}} = 126.0\,\mathrm{GeV}$, which is somewhat different from the combined best-fit m_{H} measured in Run 1. However, the results which are presented here do not comprise a measurement of the Higgs boson mass, as the data were not reprocessed with the final set of ECAL calibrations and tuning of the $BDT_{\gamma\mathrm{E}}$ which are required for a precision measurement.

7.3 Measurements of the signal strength

7.3.1 Global signal strength

One of the advantages of using $2\Delta\text{NLL}$ as a test statistic is that to a very good approximation, the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty on the measured value of a POI can be obtained by finding the values of the POI for which $2\Delta\text{NLL} = 1$ and $2\Delta\text{NLL} = 4$ respectively [73]. This fact is used to produce a measurement of the global signal strength μ.

Two modifications are made to the definition of the test statistic in Equation 7.7. First, the m_{H} parameter is profiled in the minimisation at each step, which means that it is allowed to float with a flat constraint. Second, the requirement that the test-statistic is nonzero only for positive values of $\hat{\mu}$ is relaxed, since this was a enforced to simplify the calculation of p-values. The new definition of the test statistic is therefore given by:

$$q_{\mu} = -2\ln\mathcal{L}(\mu,\, \hat{m}_{\text{H}}{}^{\mu}; \hat{\mathbf{n}}^{\mu}|\, m_{\gamma\gamma}^{obs}) - 2\ln\mathcal{L}(\hat{\mu},\, \hat{m}_{\text{H}}; \hat{\mathbf{n}}|\, m_{\gamma\gamma}^{obs}), \tag{7.9}$$

where $\hat{\mathbf{m}}_{\text{H}}{}^{\mu}$ denotes the best-fit m_{H} for a fixed value of μ.

The measurement is made by evaluating the test statistic for fixed values of μ in small steps in the range of 0.5 to 1.5. The result of the so-called $2\Delta\text{NLL}$ *scan* of μ is shown in Figure 7.5. By definition, the best-fit point $\hat{\mu}$ has a $2\Delta\text{NLL}$ value of 0. This gives the central value for the measurement. The upper and lower uncertainties are obtained by finding the intercepts of the curve with $2\Delta\text{NLL} = 1$. The contribution to the total uncertainty on the signal strength arising from the statistical, experimental systematic and theory systematic components are assessed by repeating the process, but freezing the corresponding nuisance parameters to their post-fit values. Their effect is then calculated by taking the difference in quadrature with respect to the total uncertainty. The measured value of the signal strength is:

$$\hat{\mu} = 0.94^{+0.21}_{-0.18} = 0.94 \pm 0.16\ (\text{stat.})\ ^{+0.10}_{-0.07}\ (\text{exp. syst.})\ ^{+0.08}_{-0.05}\ (\text{theo. syst.}),$$

which is consistent with the result quoted in [2] to within one decimal place.

This measurement indicates that the observed global signal strength is compatible with the SM expectation within one standard deviation. The observed particle therefore appears to behave very closely to the predictions of the SM in its overall production

rate. None, the less, various extensions to the SM predict variations of the order of a few percent, and therefore the current uncertainties, which are of the order of 20%, cannot rule out contributions from physics beyond the SM. Further accumulation of data during the LHC programme will help to bring down this uncertainty, which is currently dominated by the statistical component. The full 2016 dataset contains approximately three times as much data, which would already be enough to bring the statistical contribution to the uncertainty to the level of the theory and experimental systematic contributions.

A similar likelihood scan can be repeated for specific values of m_{H} in the 120-130 GeV range, using the $2\Delta\mathrm{NLL}$ definition from Equation 7.7 but removing the requirement that the test-statistic is nonzero only for positive values of $\hat{\mu}$. The result is shown in Figure 7.6, where the best-fit signal strength is plotted as a function of the fixed value of m_{H} in 0.1 GeV steps. The green bands represent the $\pm 1\sigma$ uncertainty obtained by finding the crossing with $2\Delta\mathrm{NLL} = 1$ for each step. This figure illustrates that no excesses other than the one at the best-fit exist in the region of interest.

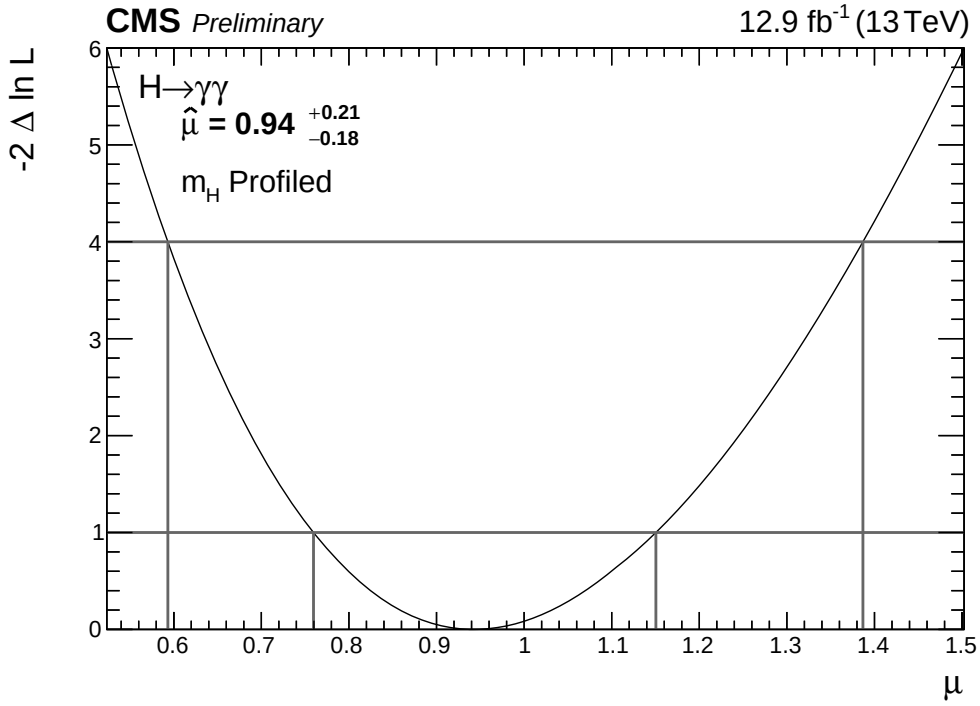


Figure 7.5: The $2\Delta\mathrm{NLL}$ scan of the overall signal strength for a Higgs boson decaying to two photons. The mass of the Higgs boson is profiled in the fit. The 1σ and 2σ uncertainties correspond to the crossings with $2\Delta\mathrm{NLL} = 1$ and $2\Delta\mathrm{NLL} = 4$.

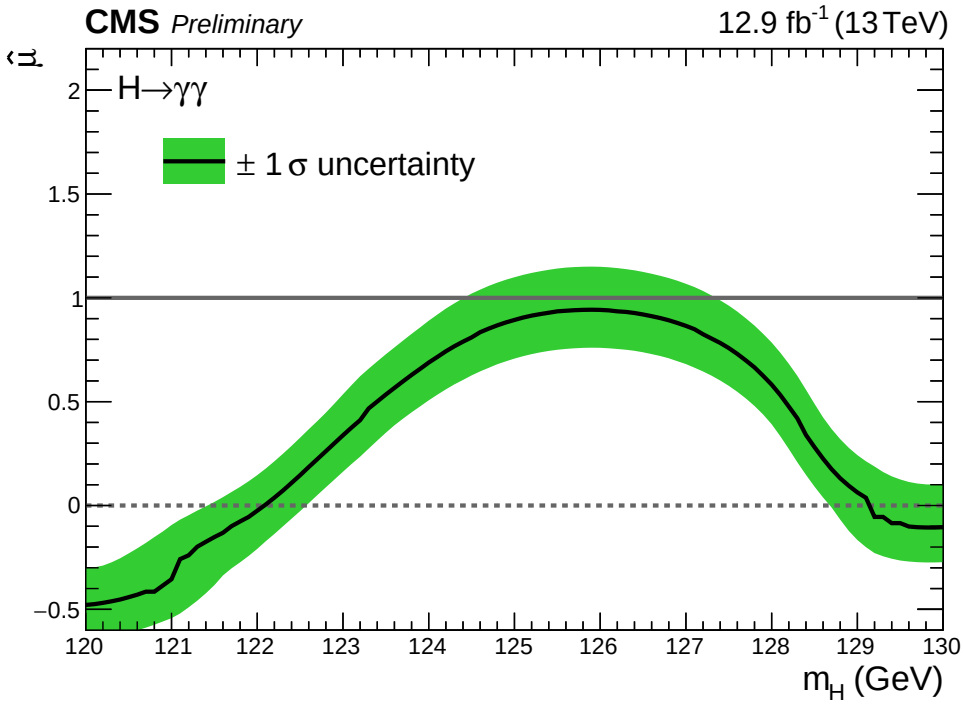


Figure 7.6: The best-fit signal strength for fixed values of m_{H} in the 120-130 GeV range, where the m_{H} parameter is fixed in the fitting procedure. The green bands show the $\pm 1\sigma$ uncertainty obtained by finding the crossing with $2\Delta\mathrm{NLL} = 1$.

7.3.2 Fermionic and bosonic components of the signal strength

When making the measurement of the global signal strength as in Section 7.3.1, a single POI which uniformly scales all production processes in all categories is defined. However, this measurement makes the assumption that the contribution of each production process to the total Higgs boson cross-section is in proportion to the SM prediction. In order to test this assumption, the single POI representing to global signal strength can be split up into components. For example, the contributions from the production modes where the Higgs boson is produced from fermions (ggH and ttH) and vector bosons (VBF and VH) are separated, to test if they individually agree with the SM expectation.

A modified likelihood function is required, which is amended from Equation 7.12 to include two POIs, $\mu_{\text{ggH,ttH}}$ and $\mu_{\text{VBF,VH}}$ in the place of μ:

$$\begin{aligned}\mathcal{L}(\,\mu_{\text{ggH,ttH}},\,\mu_{\text{VBF,VH}},\,m_{\text{H}};\mathbf{n}|\,m_{\gamma\gamma}^{obs}) = \prod_{C}\Big[f_B^C(\,m_{\gamma\gamma}^{obs,C}|\mathbf{n}_B) \\ +\ \mu_{\text{ggH,ttH}}\cdot(f_{S,\text{ggH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\text{H}};\mathbf{n}_S) + f_{S,\text{ttH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\text{H}};\mathbf{n}_S)) \\ +\ \mu_{\text{VBF,VH}}\cdot(f_{S,\text{VBF}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\text{H}};\mathbf{n}_S) + f_{S,\text{VH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\text{H}};\mathbf{n}_S))\Big].\end{aligned} \tag{7.10}$$

In this new likelihood function, the $\mu_{\text{ggH,ttH}}$ parameter scales the yield of the signal models for ggH and ttH in all categories uniformly, but does not affect the yields of the models for VBF or VH, and vice versa for the $\mu_{\text{VBF,VH}}$ parameter.

The measurement is made by producing a two-dimensional $2\Delta\text{NLL}$ scan. The test statistic $q_{(}\,\mu_{\text{ggH,ttH}},\,\mu_{\text{VBF,VH}})$ is defined as in Equation 7.9 (i.e. with m_{H} profiled), but using the amended definition of the likelihood from Equation 7.10. The result of the two-dimensional scan can be seen in Figure 7.7. The black cross shows the location of the best-fit point, with the red diamond indicating the SM expectation. In two-dimensional $2\Delta\text{NLL}$ scans, the 1σ and 2σ contours are the intersections with $2\Delta\text{NLL} = 2.30$ and $2\Delta\text{NLL} = 6.18$ respectively [73], and these are shown as solid and dashed lines in the figure.

The plot shows that the observed best-fit point is consistent with the SM within 1σ. The elliptical shape of the contours reflects the fact that the Untagged categories, which are by far the most sensitive due to their high event content and $S/(S+B)$, are populated chiefly by ggH events. This results in a strong constraint on the $\mu_{\text{ggH,ttH}}$ parameter, and a somewhat looser one on the $\mu_{\text{VBF,VH}}$ parameter. The uncertainty ellipses are also slightly inclined, which can be understood by the fact that the VBF-targeting categories contain a non-negligible amount of ggH events, and vice-versa. This leads to slight correlation: an increase in $\mu_{\text{ggH,ttH}}$ typically must come at the expense of a decrease in $\mu_{\text{VBF,VH}}$ to ensure a good fit in all analysis categories.

To correctly extract the uncertainties on the $\mu_{\text{ggH,ttH}}$ and $\mu_{\text{VBF,VH}}$ parameters individually, a $2\Delta\text{NLL}$ scan of each is performed while profiling the other. This is achieved by modifying the test statistic to treat the profiled POI analogously to m_{H} in Equation 7.9.

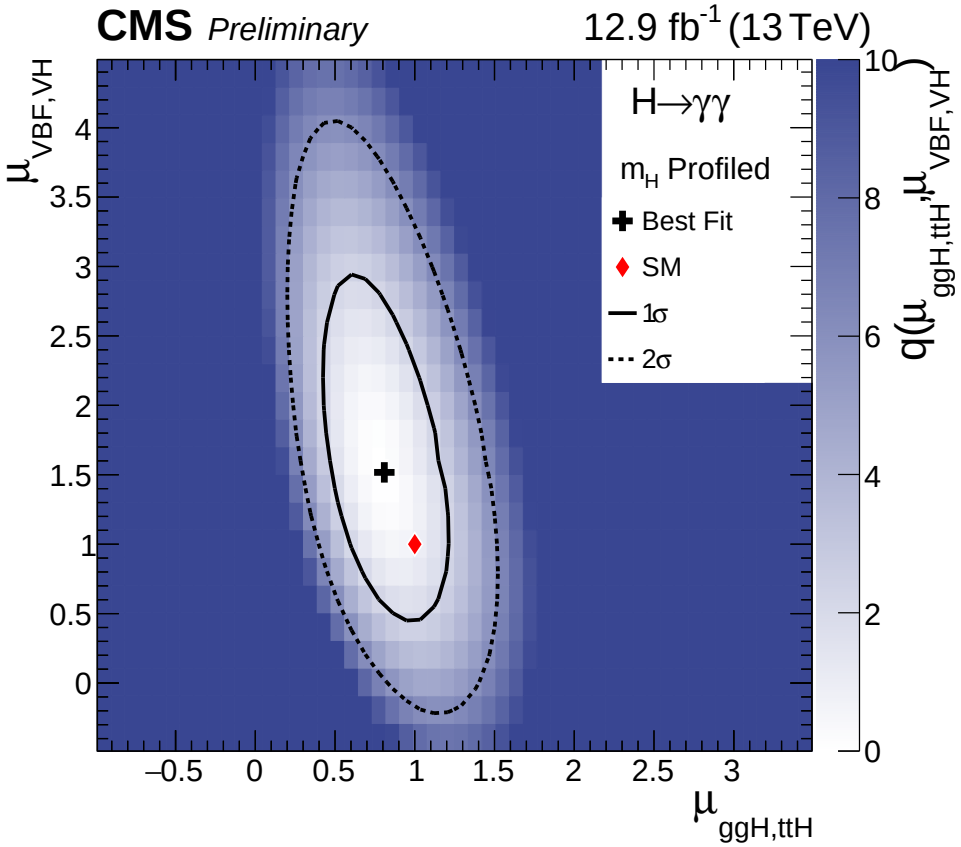


Figure 7.7: The result of a two-dimensional 2ΔNLL scan of the $\mu_{\mathrm{ggH,ttH}}$ and $\mu_{\mathrm{VBF,VH}}$ components of the signal strength. The red diamond indicates the SM expectation, while the black cross shows the location of the best-fit point. The measurement is consistent with the SM within the uncertainty contours, which are shown in solid and dashed lines for the 1σ and 2σ uncertainties respectively. The value of m_{H} was profiled in the scan.

The resulting scans are available in Appendix B in Figure B.1, which give rise to the following measurements:

$$\hat{\mu}_{\mathrm{VBF,VH}} = 1.52^{+0.89}_{-0.77},$$
$$\hat{\mu}_{\mathrm{ggH,ttH}} = 0.81^{+0.27}_{-0.25}.$$

The $\mu_{\mathrm{ggH,ttH}}$ measurement is consistent with [2] to within one unit of the smallest quoted decimal place and $\mu_{\mathrm{VBF,VH}}$ differs by less than 5% which is small compared to the size of the uncertainties. The fermionic and bosonic components of the signal strength are therefore each found to be individually compatible with the SM expectations.

7.3.3 Per-process signal strengths

Using a similar procedure to that described in Section 7.3.2, measurements of the signal strengths of the individual Higgs boson production modes can be made. In this case, the likelihood function and test statistic are modified to contain one POI for each production process (μ_{ggH}, μ_{VBF}, μ_{VH}, and μ_{ttH}). Each per-process signal strength independently scales the yield for the signal models for the corresponding process in all categories, leaving the signal models for the other processes unchanged. The likelihood function for is therefore defined as:

$$\begin{aligned}\mathcal{L}(\,\mu_{ggH},\,\mu_{ttH},\,\mu_{VBF},\,\mu_{VH},\,m_{\mathrm{H}};\mathbf{n}|\,m_{\gamma\gamma}^{obs}) = \prod_{C}\Big[& f_B^C(\,m_{\gamma\gamma}^{obs,C}|\mathbf{n}_B) \\ & + \mu_{ggH}\cdot f_{S,\mathrm{ggH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_S) \\ & + \mu_{ttH}\cdot f_{S,\mathrm{ttH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_S) \\ & + \mu_{VBF}\cdot f_{S,\mathrm{VBF}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_S) \\ & + \mu_{VH}\cdot f_{S,\mathrm{VH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_S)\Big],\end{aligned} \tag{7.11}$$

The technique described above exploits the categorisation scheme described in Chapter 5: in particular, the fact that the relative contribution from each process to the overall signal model differs from category to category. This means that varying each per-process signal strength has a different effect on the overall likelihood function. However, since no VH-tagged categories are included in this analysis, it is not possible to resolve the effect of varying μ_{VH} from other POIs, in particular μ_{ggH}, since most VH events are included in the Untagged categories along with the ggH events. To break the degeneracy, when making the measurements of the other POIs, the parameter μ_{VH} is fixed to a value of 1.

The measurements of the μ_{ggH}, μ_{VBF}, and μ_{ttH} are performed by producing a 2ΔNLL scan of each parameter, while profiling the others, where the 2ΔNLL definition has been suitably modified to accommodate the new likelihood function defined in Equation 7.11. The m_{H} parameter is also profiled. The best-fit values and their uncertainties are shown on Figure 7.8. The measurement of the global signal strength obtained in Section 7.5 is shown as the vertical black line with green bands showing the 1σ uncertainties. The SM expectation is shown as the vertical dashed red line. The per-process signal strength measurements are all compatible with the SM expectation within 1σ. The measurements also agree with those presented in [2] within one unit of the smallest significant figure for μ_{ggH}, and within 5% for μ_{VBF} and μ_{ttH}, which is small compared to the 1σ uncertainties.

This result is of particular interest because certain extension to the SM predict modified values of the per-process signal strengths. In particular, if a heavy top-like particle exists (as predicted by many theories to resolve the hierarchy problem), then an anomalous value of μ_{ttH} could be observed. However, the variations in the value of μ_{ttH} predicted by such models are typically smaller than 10%. There is evidently plenty of room to accommodate such variations in the current measurement. As more data are collected over the course of the LHC programme, this type of measurement will become increasingly important since it could reveal clues to the nature of physics beyond the SM, or put strong constraints on proposed extensions. This result also shows that the bulk of the sensitivity of the analysis resides in the Untagged categories, which are largely composed of ggH events.

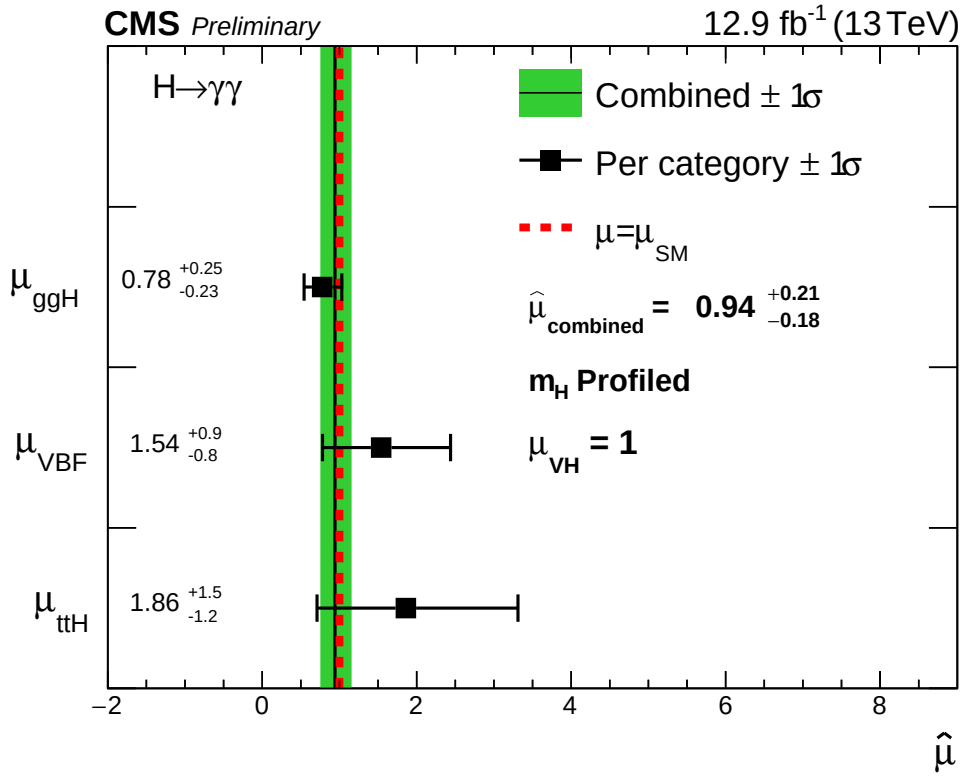


Figure 7.8: The measurements of the per-process signal strengths μ_{ggH}, μ_{VBF}, μ_{ttH}, obtained by performing 2ΔNLL scans of each one while profiling the others. In each case m_{H} is also profiled in the fit, and $\mu_{VH} = 1$ is imposed since this analysis does not include any categories specifically targeting the VH process. The vertical black line and green bands represent the measurement of the overall signal strength μ, and the SM expectation is shown in the vertical red dashed line.

7.3.4 Compatibility of result with SM in each category

Using an analogous method to the one described in Section 7.3.3, it is possible to make a measurement of the signal strength for each category separately. In this case, one POI per analysis category is defined, which scales the yield of the signal models of all processes uniformly, but independently within each category. The full likelihood function is therefore expressed as:

$$\mathcal{L}(\{\mu_C\},\, m_{\mathrm{H}}; \mathbf{n}|\, m_{\gamma\gamma}^{obs}) = \prod_C \Big[\mu_C \cdot f_S^C(\, m_{\gamma\gamma}^{obs,C}|\, m_{\mathrm{H}}; \mathbf{n}_S) + f_B^C(\, m_{\gamma\gamma}^{obs,C}|\mathbf{n}_B)\Big], \tag{7.12}$$

where $\{\mu_C\}$ represents the set of signal strengths for each category.

Although the signal strengths $\{\mu_C\}$ do not have any direct physical meaning, they can be used to check that each category gives a result consistent with the overall measurement, and that no bias is introduced by any particular category. The result of the check is shown in Figure 7.9, which determines that all the per-category signal strengths are compatible with the SM expectation and the overall result. The results for each per-category signal strength match those from [2] within a few a percent, as for previously quoted results.

Of the eight categories which are included in this analysis, all but two of them (the Untagged 2 and VBF-tagged 0 categories) fall within 1σ of the overall best-fit global signal strength, which roughly matches the expectation that a randomly distributed Gaussian variable falls within 1σ of the mean approximately 32% of the time.

This result shows once again that the Untagged categories are the ones which provide bulk of the sensitivity of the overall result. Another point of interest is that the $S/(S+B)$ of a given category does not necessarily correlate with the size of the uncertainty for that category. In particular, the category whose signal strength has the lowest uncertainty is the Untagged 2 category, while Figure 6.10 shows that the Untagged 0 category has the best $S/(S+B)$. This can be explained by the fact that the statistical component is still the dominating uncertainty in each measurement, and the lower $S/(S+B)$ of the Untagged 2 category is compensated by the larger number of events which enter it.

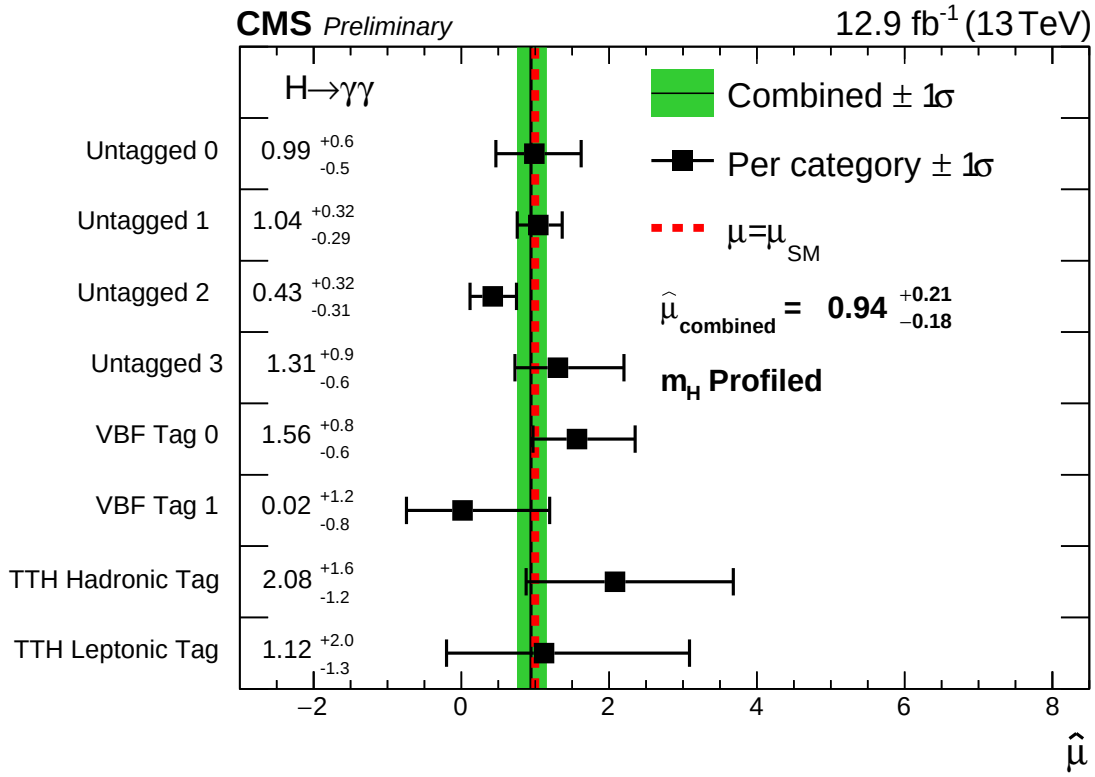


Figure 7.9: The measurements of the per-category signal strengths, obtained by performing 2ΔNLL scans of each one while profiling the others. In each case m_{H} is also profiled in the fit, The vertical black line and green bands represent the measurement of the overall signal strength μ, and the SM expectation is shown in the vertical red dashed line. The per-category signal strengths do not have a direct physical interpretation, and this result is a check that no particular category is introducing a large bias into the overall measurement.

7.4 Measurements of Higgs boson coupling modifiers

7.4.1 Motivation and theory

The measurements presented so far have all dealt with Higgs boson signal strengths. Such observables are sensitive to variations in the rate at which the Higgs boson is produced, but do not take into account the possible variations in the partial width of the subsequent decay. An alternative set of measurements addresses this shortcoming by being sensitive to variations in the coupling strength of the Higgs boson with individual particles, relative to the SM expectation.

The so-called *kappa* framework assigns a modifier to the coupling strength of the Higgs boson to a particle or group of particles X directly in the amplitude of the process. The

corresponding modifier is labelled as κ_X. A detailed description of the scheme is available in [74]. The assumptions which underly this framework are as follows:

- any calculated deviations from the SM prediction are due to only one Higgs-boson-like state with mass around 125 GeV;
- the natural width of this state is sufficiently small that it can be neglected, allowing the cross-section and branching fraction for a process $ii \to H \to ff$ to be decomposed as $(\sigma^H_{ii} \cdot \Gamma^H_{ff})/(\Gamma_H)$, where σ^H_{ii} is the cross-section for a Higgs boson to be produced from the initial state ii, Γ^H_{ff} is the partial decay width of the Higgs boson into the state ff and Γ_H is the total width of the Higgs boson.

The *coupling modifier* κ_X for a particle X interacting with the Higgs boson is applied directly as a factor to the corresponding cross-section σ^H_{XX} or partial decay width Γ^H_{XX}. For processes which only occur via loops of particles, an *effective coupling modifier* is defined as a function of the coupling modifiers for particles in which play a large role in the loop, e.g. $\kappa_\gamma = \kappa_\gamma(\kappa_b, \kappa_t)$ (for the decay $\mathrm{H} \to \gamma\gamma$) and $\kappa_g = \kappa_g(\kappa_b, \kappa_t)$ (for ggH production). Bosonic and fermionic coupling modifiers κ_f and κ_V are defined which uniformly scale the Higgs boson's interactions with all fermions and vector bosons respectively. The coupling modifiers applied to each of the main Higgs boson production processes and the $\mathrm{H} \to \gamma\gamma$ decay are shown in Table 7.1, adapted from [74].

Process	Type	Loop (interference)	Coupling modifier (in terms of κ_f and κ_V)	Effective coupling modifier
ggH	cross-section	yes (t - b)	κ^2_f	κ^2_g
VBF	cross-section	no	κ^2_V	-
VH	cross-section	no	κ^2_V	-
ttH	cross-section	no	κ^2_f	-
$\mathrm{H} \to \gamma\gamma$	partial width	yes (t - W)	$1.59 \cdot \kappa^2_\mathrm{V} + 0.07 \cdot \kappa^2_\mathrm{f} - 0.66 \cdot \kappa_\mathrm{V} \cdot \kappa_\mathrm{f}$	κ^2_γ

Table 7.1: Coupling strength modifiers attributed to each of the main Higgs boson production mechanism cross-sections and the partial width of the $\mathrm{H} \to \gamma\gamma$ decay, including QCD and EW corrections [74].

7.4.2 Bosonic and fermionic coupling modifiers

To make a measurement of fermionic and bosonic Higgs boson coupling modifiers, the likelihood function is re-written such that κ_f and κ_V are the POIs:

$$\begin{aligned}\mathcal{L}(\,\kappa_\mathrm{f},\,\kappa_\mathrm{V},\,m_\mathrm{H};\mathbf{n}|\,m_{\gamma\gamma}^{obs}) = \prod_C \Big[& f_B^C(\,m_{\gamma\gamma}^{obs,C}|\mathbf{n}_B) \\ & + \kappa_\mathrm{f}^2\cdot(1.59\cdot\kappa_\mathrm{V}^2 + 0.07\,\kappa_\mathrm{f}^2 - 0.66\,\kappa_\mathrm{V}\,\kappa_\mathrm{f})\cdot f_{S,\mathrm{ggH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_\mathrm{H};\mathbf{n}_S) \\ & + \kappa_\mathrm{f}^2\cdot(1.59\cdot\kappa_\mathrm{V}^2 + 0.07\,\kappa_\mathrm{f}^2 - 0.66\,\kappa_\mathrm{V}\,\kappa_\mathrm{f})\cdot f_{S,\mathrm{ttH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_\mathrm{H};\mathbf{n}_S) \\ & + \kappa_\mathrm{V}^2\cdot(1.59\cdot\kappa_\mathrm{V}^2 + 0.07\,\kappa_\mathrm{f}^2 - 0.66\,\kappa_\mathrm{V}\,\kappa_\mathrm{f})\cdot f_{S,\mathrm{VBF}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_\mathrm{H};\mathbf{n}_S) \\ & + \kappa_\mathrm{V}^2\cdot(1.59\cdot\kappa_\mathrm{V}^2 + 0.07\,\kappa_\mathrm{f}^2 - 0.66\,\kappa_\mathrm{V}\,\kappa_\mathrm{f})\cdot f_{S,\mathrm{VH}}^C(\,m_{\gamma\gamma}^{obs,C}|\,m_\mathrm{H};\mathbf{n}_S)\Big]. \end{aligned} \tag{7.13}$$

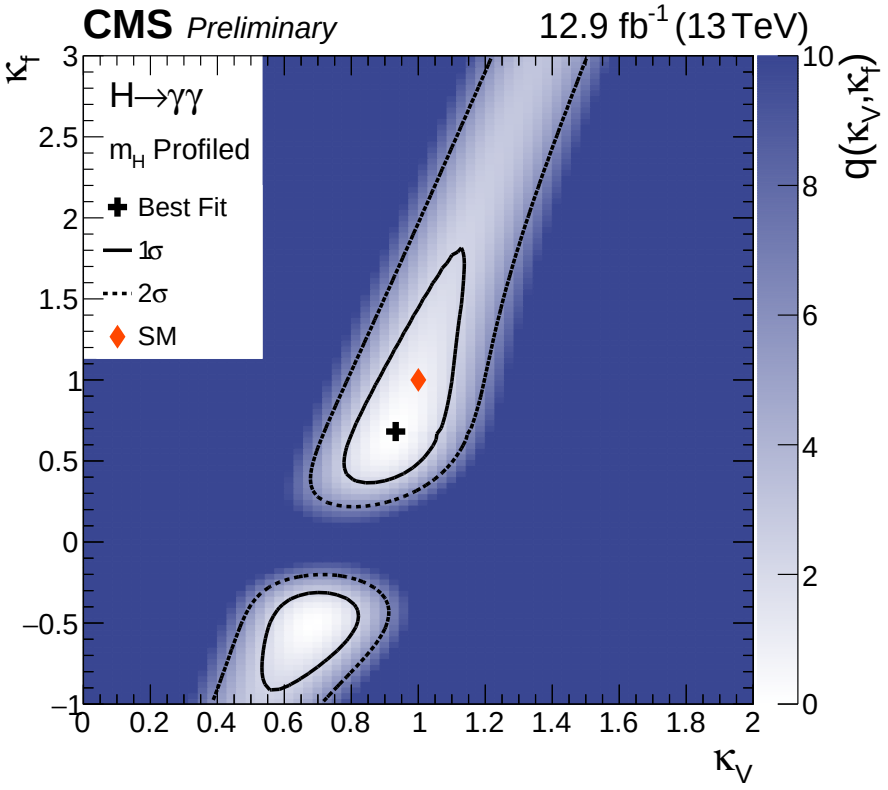
The definition of the test statistic is also modified to take the new POIs into account:

$$q(\,\kappa_\mathrm{V},\,\kappa_\mathrm{f}) = -2\ln\mathcal{L}(\,\kappa_\mathrm{V},\,\kappa_\mathrm{f},\,\hat{m}_\mathrm{H}{}^{\mu};\hat{\mathbf{n}}^{\mu}|\,m_{\gamma\gamma}^{obs}) - 2\ln\mathcal{L}(\,\hat{\kappa}_\mathrm{V},\,\hat{\kappa}_\mathrm{f},\,m_\mathrm{H};\hat{\mathbf{n}}|\,m_{\gamma\gamma}^{obs}). \tag{7.14}$$

The result of a two-dimensional 2ΔNLL scan of κ_f and κ_V is shown in Figure 7.10. The black cross indicates the best-fit while the red diamond indicates the SM expectation. The 1σ and 2σ contours are indicated by solid and dashed lines. The best-fit indicates compatibility with the SM.The location of the best-fit point agrees within a few percent with the equivalent result in the Appendix of[2]. Measurements of κ_f and κ_V can be made individually following the same method as described in Sectionsec:statandresults:rvrf. The 2ΔNLL of each POI while profiling the others can be found in Appendix B in Figure B.2, which yield the measurements:

$$\begin{aligned}\hat{\kappa}_\mathrm{V} &= 0.93^{+0.11}_{-0.10},\\ \hat{\kappa}_\mathrm{f} &= 0.68^{+0.45}_{-0.22}.\end{aligned}$$

An interesting feature of Figure 7.10 is that a second local minimum exists where κ_f takes negative values. In general, the coupling strength modifiers always occur squared in the amplitude. However, as noted in Table 7.1, the $\mathrm{H}\to\gamma\gamma$ amplitude contains destructive interference between the contributions of the t and W result in a term proportional to $(\,\kappa_\mathrm{V}\cdot\kappa_\mathrm{f})$. This means that the measurement has a small amount of sensitivity to the sign of the coupling strength modifier κ_f. In this measurement, the positive value is preferred, as expected by the SM.



(a) 2ΔNLL scan of κ_{f} versus κ_{V}

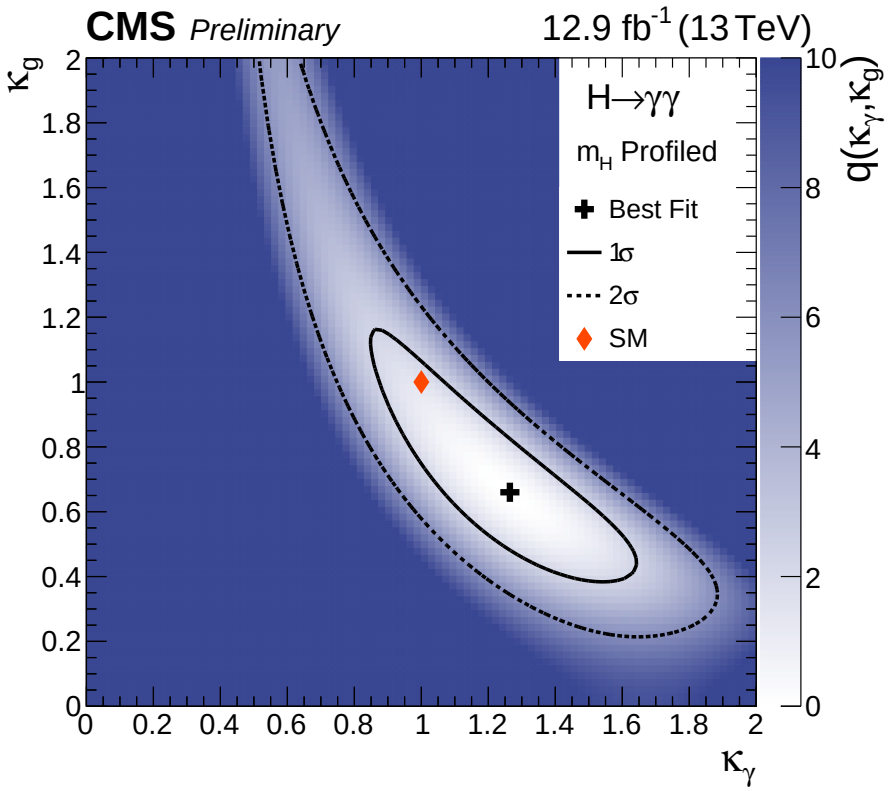
Figure 7.10: The result of a two-dimensional 2ΔNLL scan of the effective coupling strength modifiers for fermions and bosons (a). The best-fit values are denoted with black crosses and the SM expected values with red diamonds. The best-fit points agree with the SM within the 1σ and 2σ uncertainty contours denoted by the solid and dashed lines respectively.

7.4.3 Effective coupling modifiers to gluons and photons

Alternatively, the measurement can be made in terms of the Higgs boson's effective coupling to gluons and photons using a two-dimensional 2ΔNLL scan of κ_{g} and κ_{γ}. In this case, the likelihood function is modified as follows:

$$\begin{aligned}\mathcal{L}(\,\kappa_{\mathrm{f}},\,\kappa_{\mathrm{V}},\,m_{\mathrm{H}};\mathbf{n}|\,m_{\gamma\gamma}^{obs}) = \prod_{C}\Big[& f_{B}^{C}(\,m_{\gamma\gamma}^{obs,C}|\mathbf{n}_{B}) \\ & + \kappa_{\mathrm{g}}^{2}\cdot\kappa_{\gamma}^{2}\cdot f_{S,\mathrm{ggH}}^{C}(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_{S}) \\ & + \kappa_{\gamma}^{2}\cdot f_{S,\mathrm{ttH}}^{C}(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_{S}) \\ & + \kappa_{\gamma}^{2}\cdot f_{S,\mathrm{VBF}}^{C}(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_{S}) \\ & + \kappa_{\gamma}^{2}\cdot f_{S,\mathrm{VH}}^{C}(\,m_{\gamma\gamma}^{obs,C}|\,m_{\mathrm{H}};\mathbf{n}_{S})\quad\Big].\end{aligned} \tag{7.15}$$

The test statistic $q(\kappa_\gamma, \kappa_g)$ is then defined analogously to Equation 7.14. In this case the m_H parameter is profiled. The results of the 2ΔNLL scan are presented in Figure 7.11.



(a) 2ΔNLL scan of κ_γ versus κ_g

Figure 7.11: The result of a two-dimensional 2ΔNLL scan of the effective coupling strength modifiers for gluons and photons. The best-fit values are denoted with black crosses and the SM expected values with red diamonds. The best-fit points agree with the SM within the 1σ and 2σ uncertainty contours denoted by the solid and dashed lines respectively.

Measurements of κ_g and κ_γ are performed by producing a 2ΔNLL of each POI while profiling the other. The scans are available in Appendix B in Figure B.3, which yield the measurements:

$$\hat{\kappa}_g = 0.66^{+0.27}_{-0.19},$$
$$\hat{\kappa}_\gamma = 1.26^{+0.25}_{-0.27}.$$

The best-fit point agrees with the SM within the uncertainties. In this case, there is no sensitivity to negative values of the modifiers, since the $H \to \gamma\gamma$ loop which contains a term proportional to $(\kappa_V \cdot \kappa_f)$ is contained in the effective coupling κ_γ.

Chapter 8

Summary and conclusions

A study of the $\mathrm{H} \to \gamma\gamma$ decay using $12.9\,\mathrm{fb}^{-1}$ of 13 TeV data was presented, where signal-like diphoton events were sorted into orthogonal analysis categories. Signal and background models were built from the categorised data and simulation, before undergoing statistical interpretation. The analysis differed from the official CMS preliminary result for the same dataset [2] insofar as it used improvements to the parametric signal modelling.

The result of the analysis is a new observation of the Higgs boson with 13 TeV data. The significance of the excess assuming $m_{\mathrm{H}} = 125.09\,\mathrm{GeV}$ is observed to be 5.7σ (6.3σ expected). The maximum significance is observed to be 6.1σ (6.3σ expected) for $m_{\mathrm{H}} = 125.9\,\mathrm{GeV}$. Measurements of some of the properties of the observed particle were made. The best-fit global signal strength was determined to be $\hat{\mu} = 0.94^{+0.21}_{-0.18} = 0.94 \pm 0.16$ (stat.) $^{+0.10}_{-0.07}$ (exp. syst.) $^{+0.08}_{-0.05}$ (theo. syst.). The best-fit fermionic and bosonic components of the signal strength were found to be $\hat{\mu}_{\mathrm{ggH,ttH}} = 0.81^{+0.27}_{-0.25}$ and $\hat{\mu}_{\mathrm{VBF,VH}} = 1.52^{+0.89}_{-0.77}$. Furthermore, the signal strength was measured for the main Higgs boson production processes (apart from VH), giving $\hat{\mu}_{\mathrm{ggH}} = 0.78^{+0.25}_{-0.23}$, $\hat{\mu}_{\mathrm{ttH}} = 1.54^{+0.9}_{-0.8}$, $\hat{\mu}_{\mathrm{VBF}} = 1.86^{+1.5}_{-1.2}$. Finally, measurements of some of the Higgs boson coupling strength modifiers were made, yielding $\hat{\kappa}_{\mathrm{f}} = 0.68^{+0.45}_{-0.22}$ and $\hat{\kappa}_{\mathrm{V}} = 0.93^{+0.11}_{-0.10}$ for the fermionic and bosonic modifiers and $\hat{\kappa}_g = 0.66^{+0.27}_{-0.19}$ and $\hat{\kappa}_\gamma = 1.26^{+0.25}_{-0.27}$ for the effective coupling modifiers. All the measurements are found to be consistent with the SM prediction within the uncertainties, which are dominated by the statistical component. This result is a confirmation of the Higgs boson discovery in the diphoton decay channel, using in the Run 2 dataset. The new particle behaves as predicted by the SM for a Higgs boson at 13 TeV. The SM therefore continues to provide excellent agreement with data, as it has done for all measurements from collider experiments made thus far.

Despite its successes, however, the SM is not a satisfactory theory to describe our universe: it is ostensibly incomplete or approximate. It does not leave room for neutrino masses which are necessary to explain the observed neutrino flavour oscillations, e.g. in [75]. The existence dark matter cannot be explained by the SM, and yet this is evident from, for example, the behaviour of the Bullet cluster [76]. The SM also does not account for the gravitational force or the matter-antimatter asymmetry in the universe. Furthermore, the observed mass of the Higgs boson raises the fine-tuning problem: if we assume the existence of physics beyond the SM, radiative corrections to the Higgs boson mass should push it to the scale of new physics unless a fine-tuned cancellation occurs. Yet m_{H} is observed near the electroweak scale. Various extensions of the SM seek to address these concerns. In some models the Higgs boson is a composite particle [77] or part of a wider family of Higgs bosons [78]. Other theories suggest that it could be a 'portal' to new physics [79], as it could interact with unknown particles with which regular matter cannot. All such theories would have measurable effects on the Higgs boson's production rate, interactions with other particles or multiplicity.

In order to test some of these proposed models, refinements of the measurements presented above will become increasingly important. For example, the μ_{ttH} parameter is sensitive to additional top-like particles, which appear necessary to resolve the fine-tuning problem. Such additional particles are also predicted by many theories which address neutrino oscillations, dark matter, and matter-antimatter asymmetry, for example supersymmetry [80]. As presented above, the rate of $ttH(\gamma\gamma)$ is consistent with the SM, but with $\sim$150% uncertainty. Evidently, there is plenty of room for discrepancies to hide within the current measurements. Recent extrapolations of the CMS results, which the author was responsible for producing, showed that reducing these uncertainties to the $\sim$30% level could be achievable in the next five years, and down to the $\sim$15% by the end of the LHC programme [81]. Such improvements will put stringent constraints on the nature of physics beyond the SM, particularly when combined with results from other decay channels.

The field of Higgs physics therefore finds itself in a new and exciting situation. The discovery of the Higgs boson has opened a new avenue with which to study fundamental questions about our universe. This avenue will be exploited for the very first time in the coming decade, as the field pivots from the discovery era into the precision measurement era. Precision studies of the Higgs boson's properties will either set strong limits on proposed theories or reveal clues to new physics beyond the SM.

Appendix A

Additional signal and background modelling figures

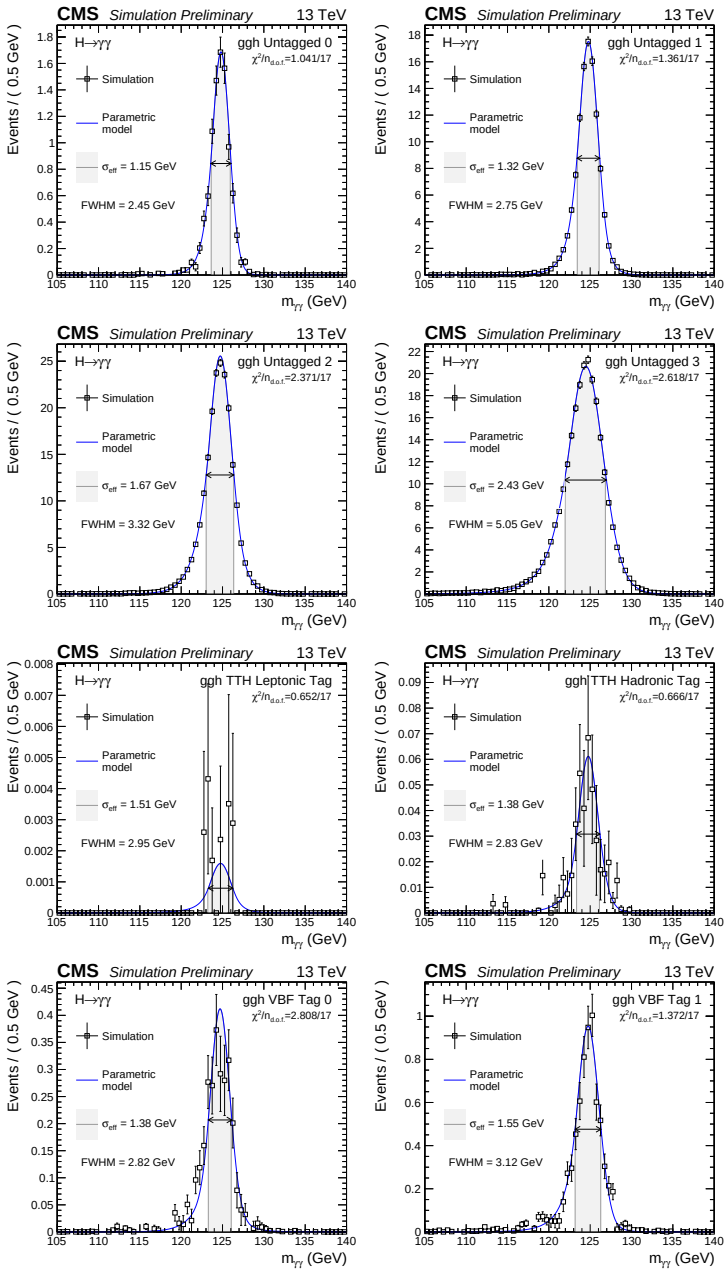


Figure A.1: The signal models for the ggH process, evaluated at $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained after application of the SSF interpolation method for the DCB+1G parametrisation of the simulated mass points. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown. Note that the fits here maybe differ slightly from those in shown Figure 6.1, which were produced by fitting the $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ samples only.

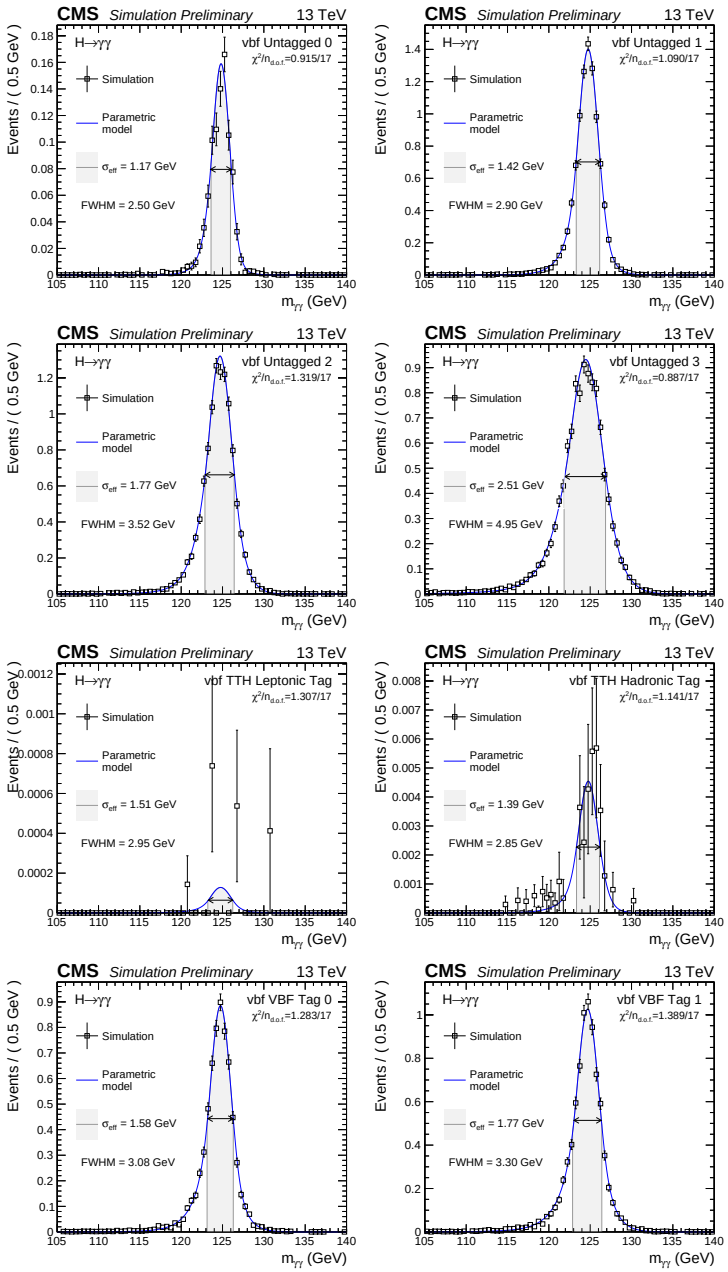


Figure A.2: The signal models for the VBF process, evaluated at $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained after application of the SSF interpolation method for the DCB+1G parametrisation of the simulated mass points. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown. Note that the fits here maybe differ slightly from those in shown Figure 6.1, which were produced by fitting the $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ samples only.

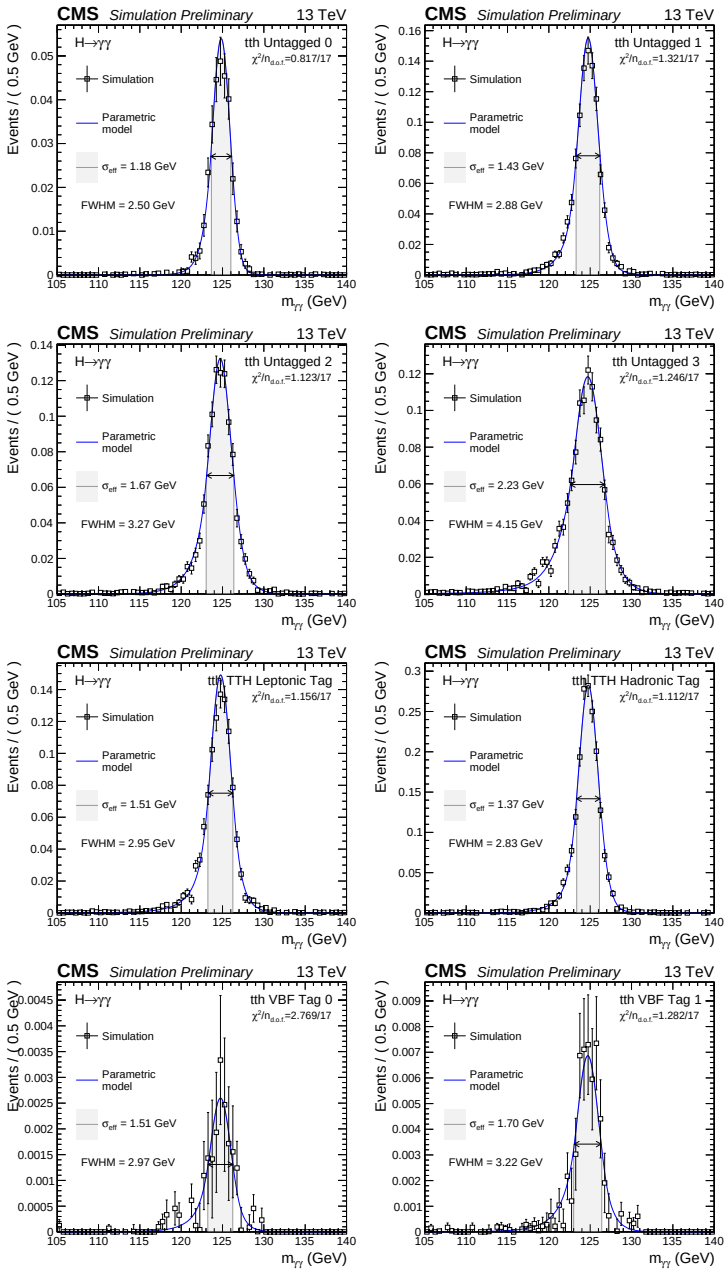


Figure A.3: The signal models for the ttH process, evaluated at $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained after application of the SSF interpolation method for the DCB+1G parametrisation of the simulated mass points. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown. Note that the fits here maybe differ slightly from those in shown Figure 6.1, which were produced by fitting the $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ samples only.

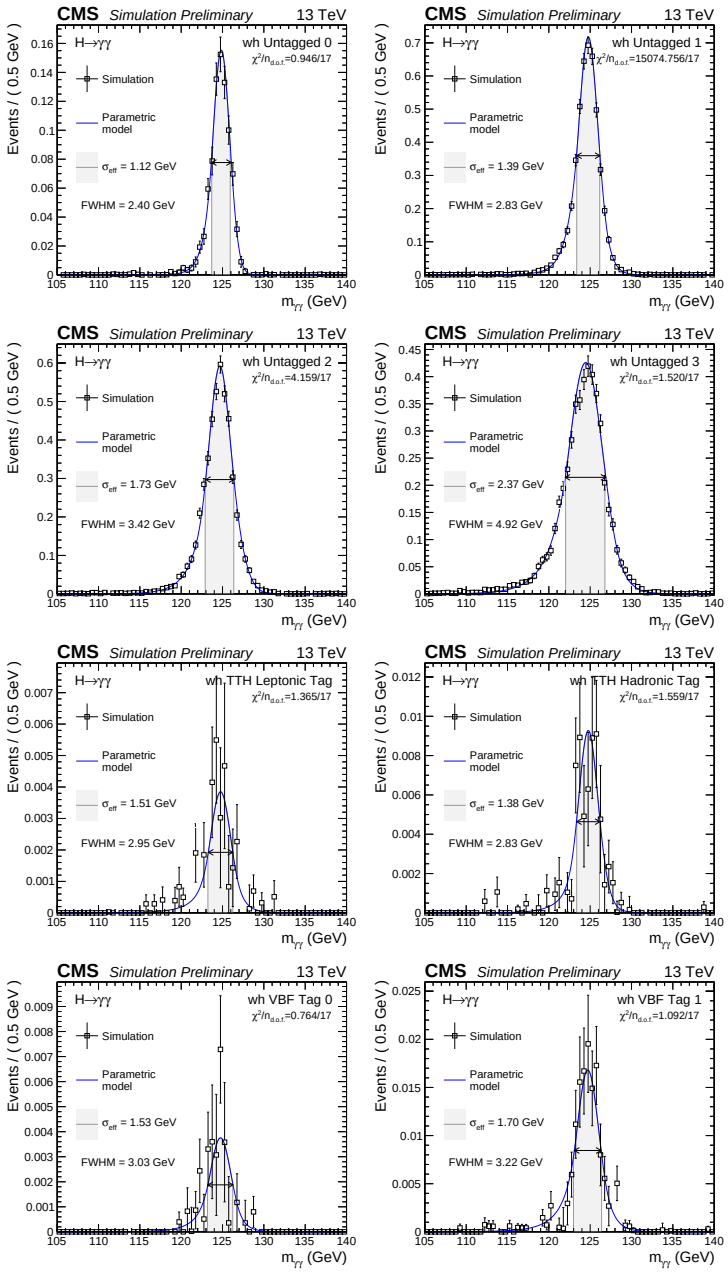


Figure A.4: The signal models for the WH process, which is later combined with ZH to model the VH process, evaluated at $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained after application of the SSF interpolation method for the DCB+1G parametrisation of the simulated mass points. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown. Note that the fits here maybe differ slightly from those in shown Figure 6.1, which were produced by fitting the $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ samples only.

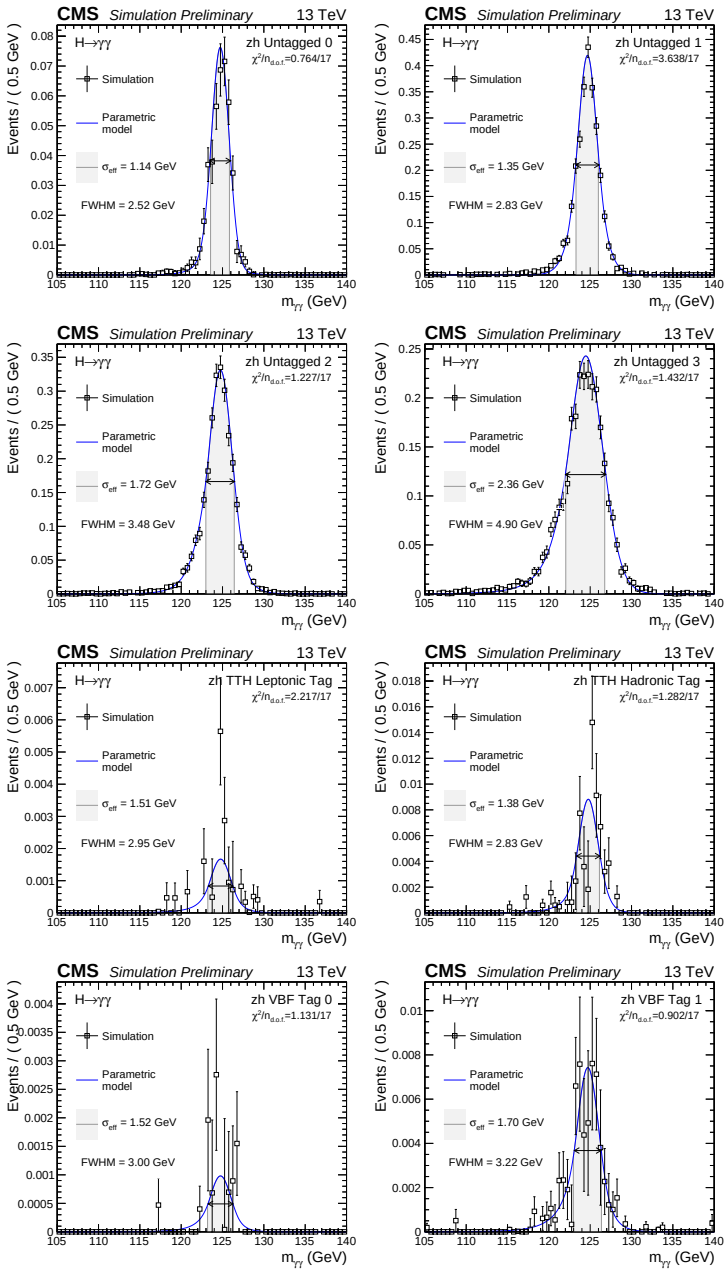


Figure A.5: The signal models for the ZH process, which is later combined with ZH to model the VH process, evaluated at $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, obtained after application of the SSF interpolation method for the DCB+1G parametrisation of the simulated mass points. The σ_{eff} (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the FWHM (the width of the distribution at half of the maximum value) are also shown. Note that the fits here maybe differ slightly from those in shown Figure 6.1, which were produced by fitting the $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ samples only.

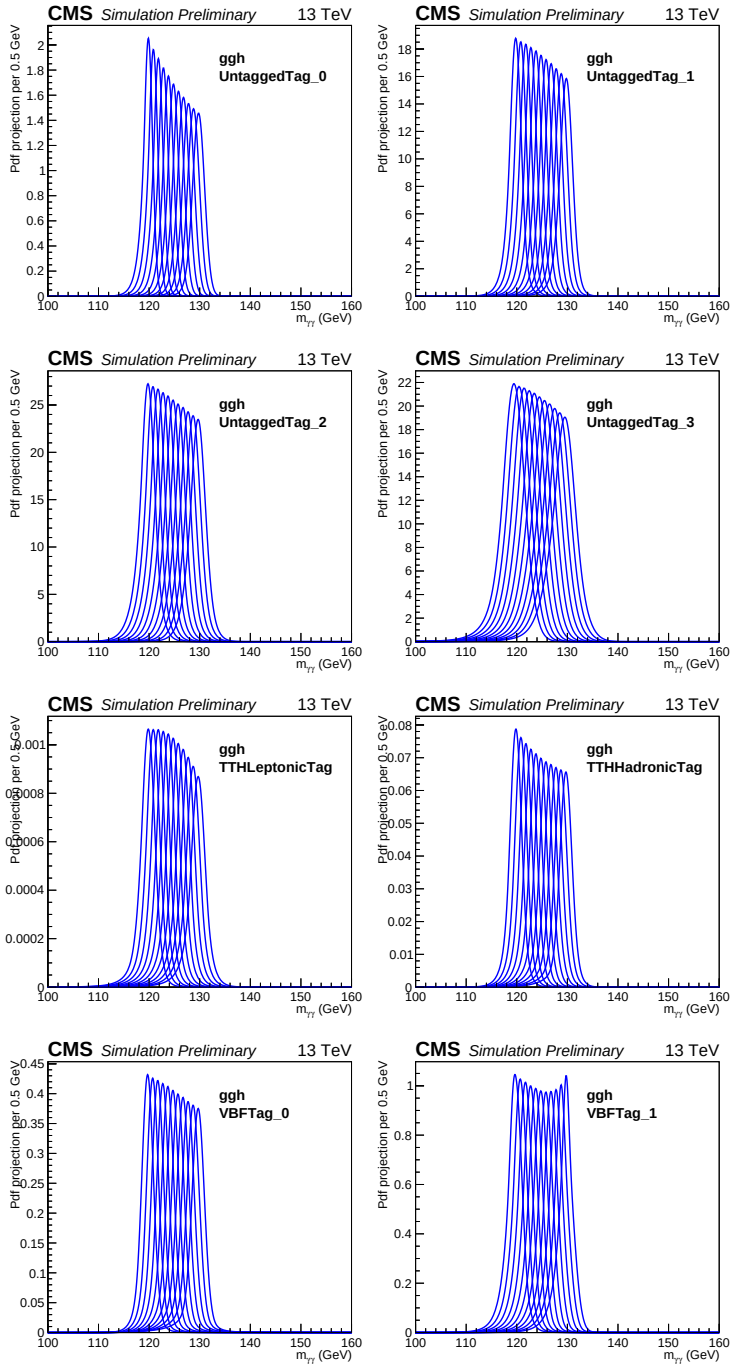


Figure A.6: The m_{H}-dependence of the signal models for the ggH process for each of the categories is shown. Each curve shows the signal model for a given value of m_{H}. The contributions from the RV and WV components of each model were interpolated between the samples for different m_{H} using the SSF method, and summed together according to their relative event content.

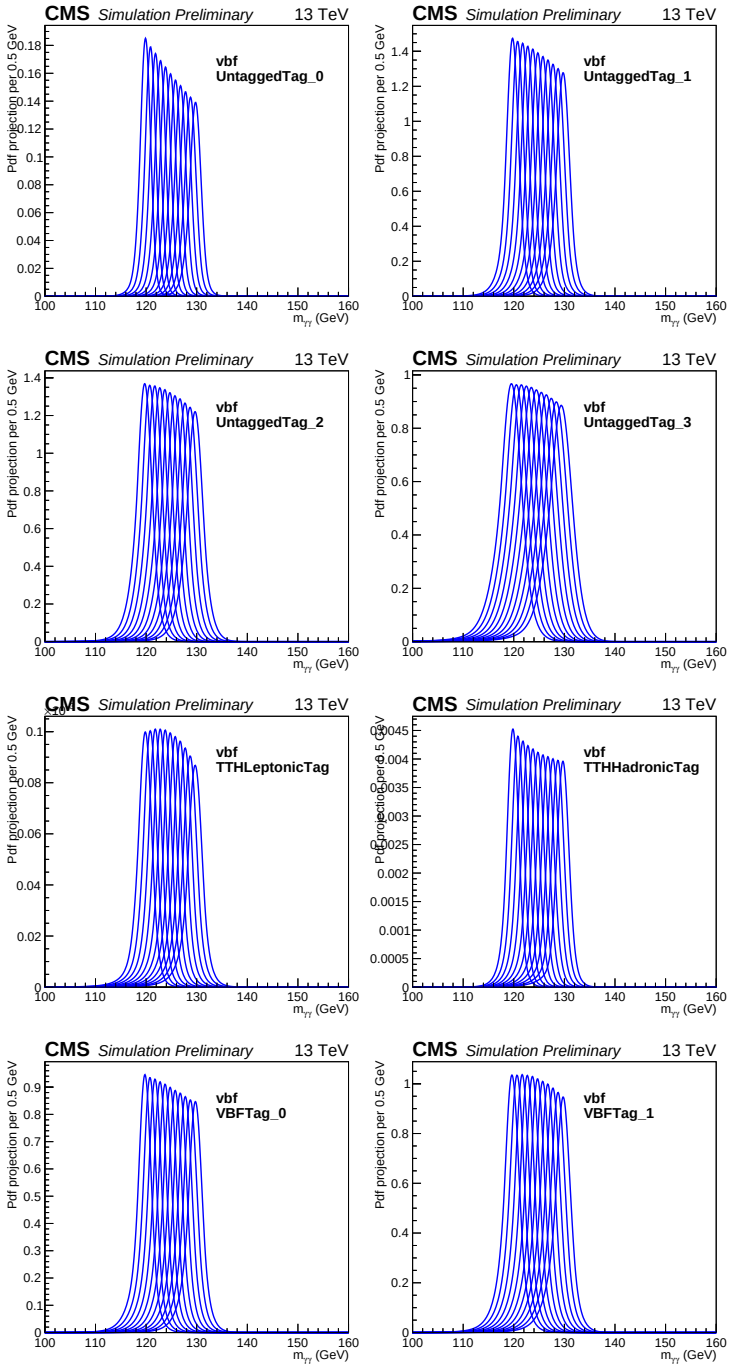


Figure A.7: The m_{H}-dependence of the signal models for the VBF process for each of the categories is shown. Each curve shows the signal model for a given value of m_{H}. The contributions from the RV and WV components of each model were interpolated between the samples for different m_{H} using the SSF method, and summed together according to their relative event content.

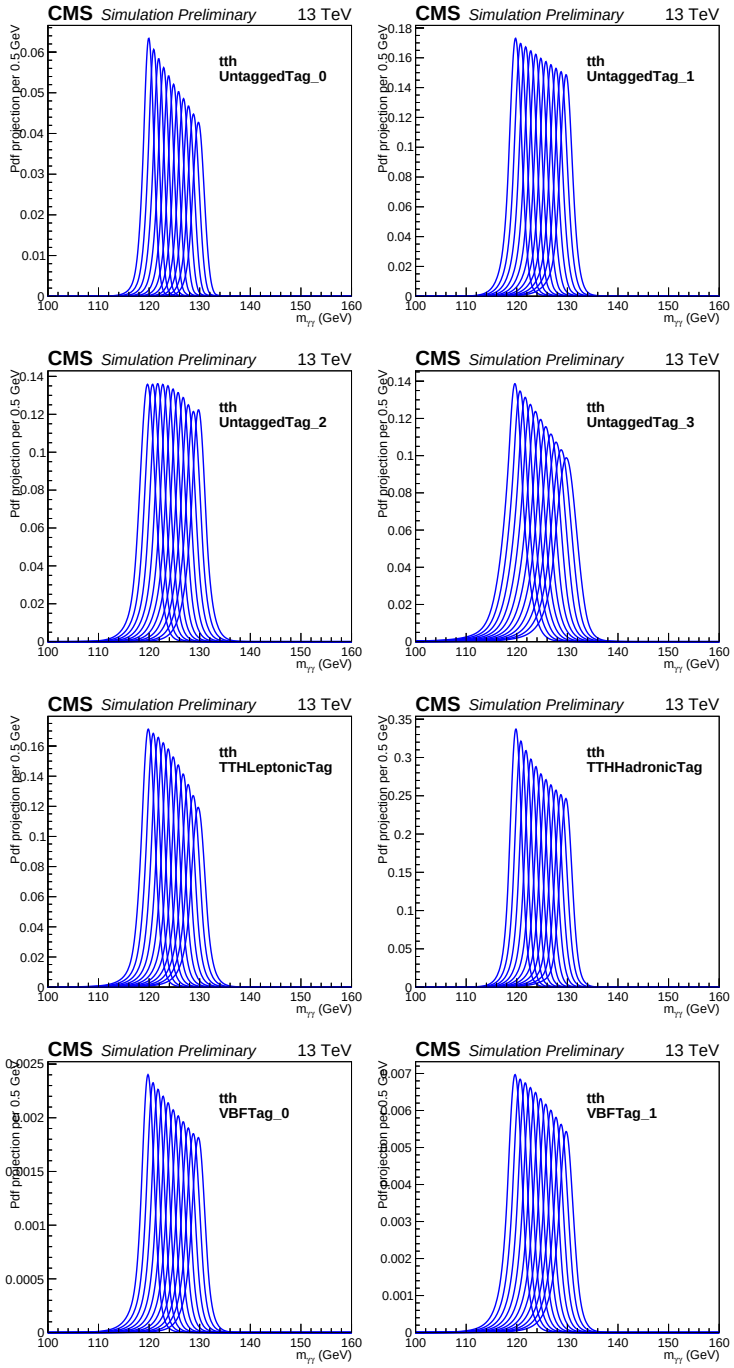


Figure A.8: The m_{H}-dependence of the signal models for the ttH process for each of the categories is shown. Each curve shows the signal model for a given value of m_{H}. The contributions from the RV and WV components of each model were interpolated between the samples for different m_{H} using the SSF method, and summed together according to their relative event content.

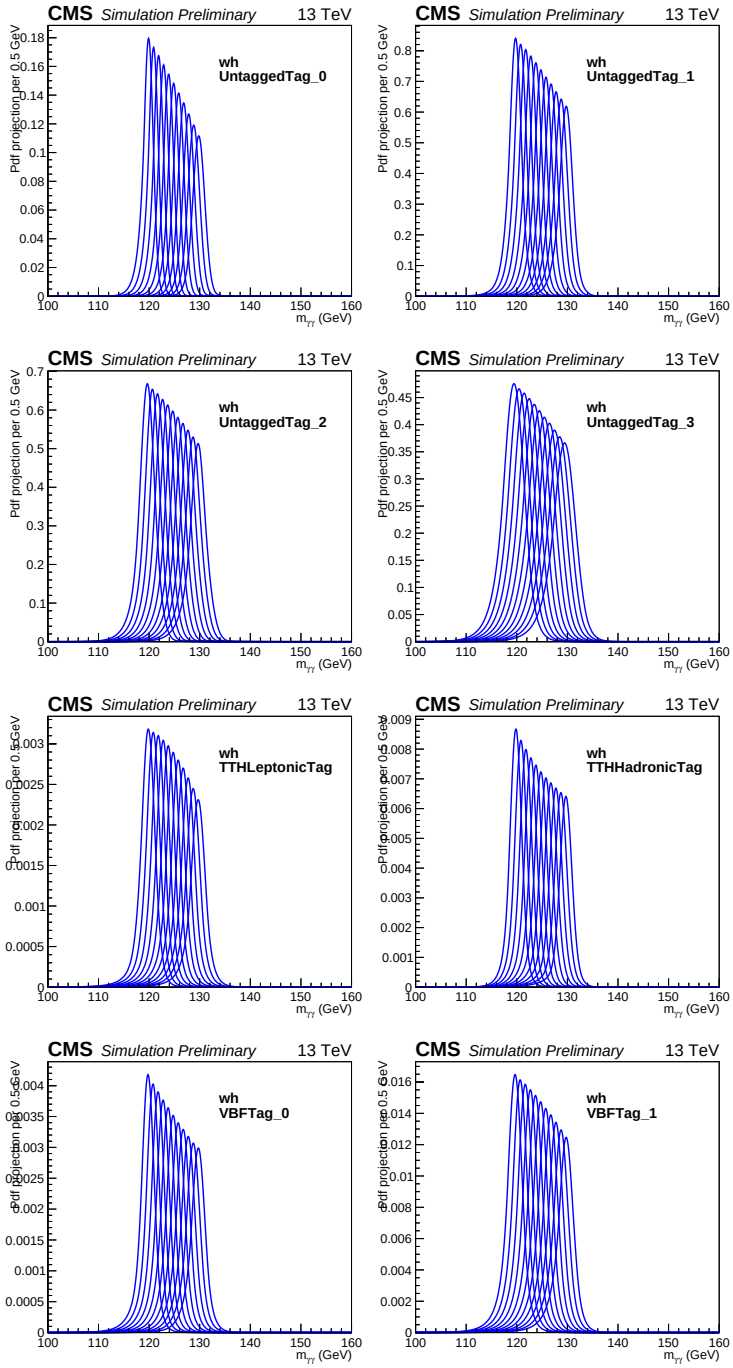


Figure A.9: The m_{H}-dependence of the signal models for the WH process for each of the categories is shown. Each curve shows the signal model for a given value of m_{H}. The contributions from the RV and WV components of each model were interpolated between the samples for different m_{H} using the SSF method, and summed together according to their relative event content.

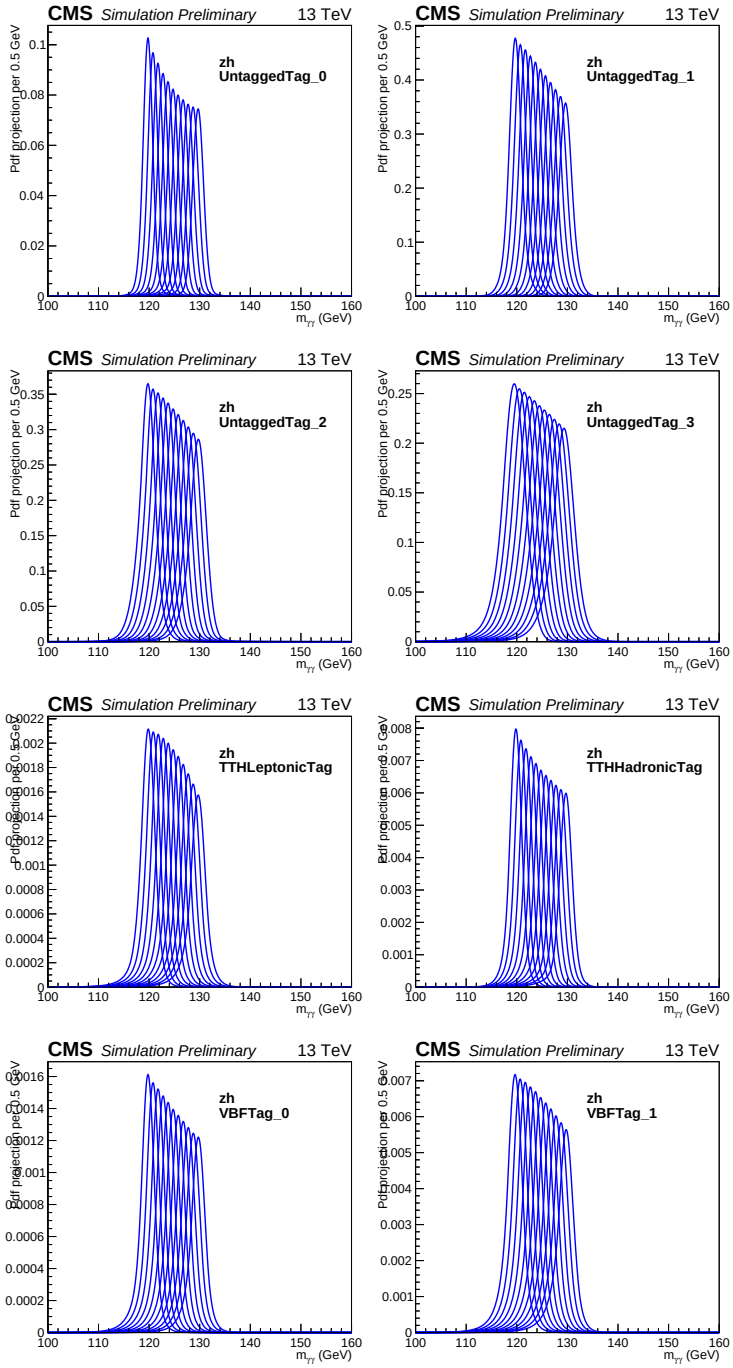
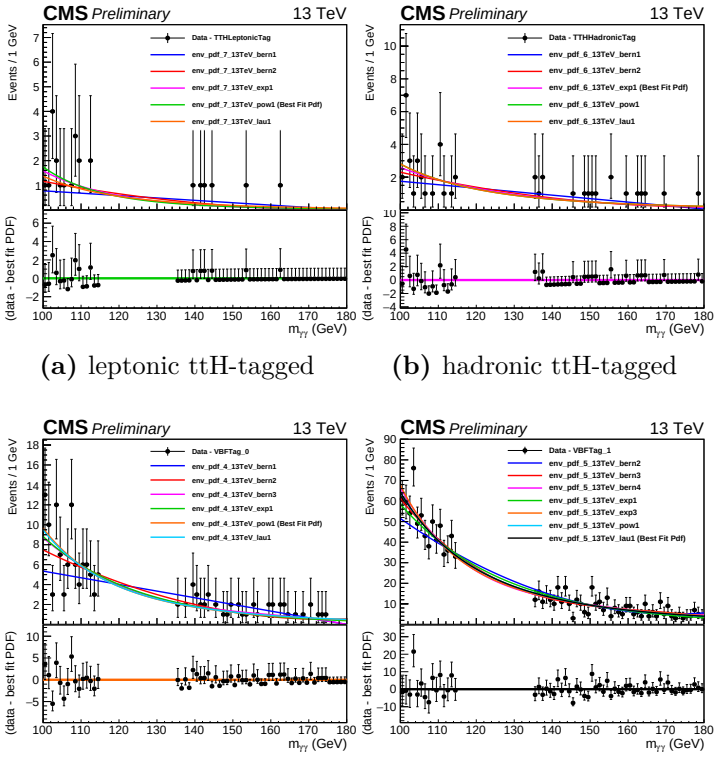


Figure A.10: The m_{H}-dependence of the signal models for the ZH process for each of the categories is shown. Each curve shows the signal model for a given value of m_{H}. The contributions from the RV and WV components of each model were interpolated between the samples for different m_{H} using the SSF method, and summed together according to their relative event content.



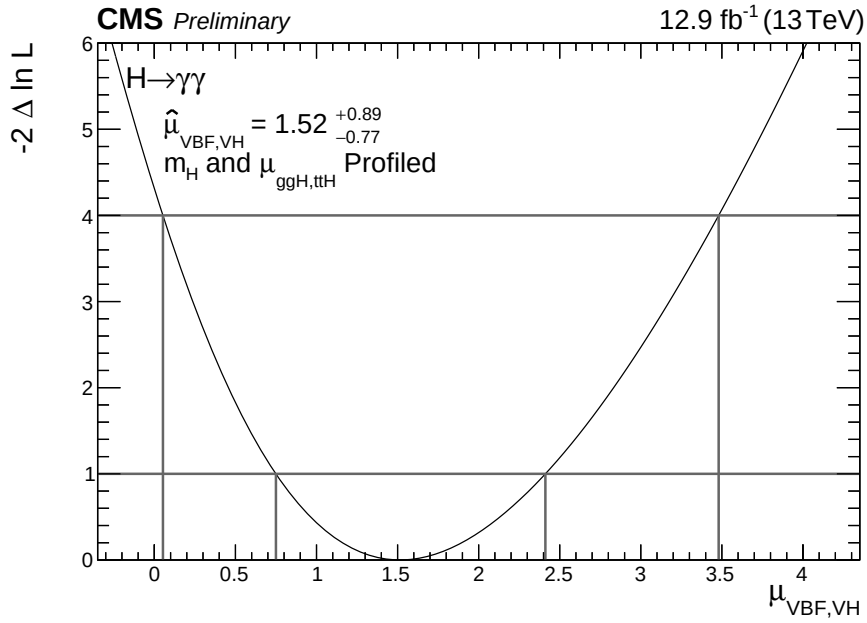
(a) leptonic ttH-tagged **(b)** hadronic ttH-tagged

(c) VBF-tagged 0 **(d)** VBF-tagged 1

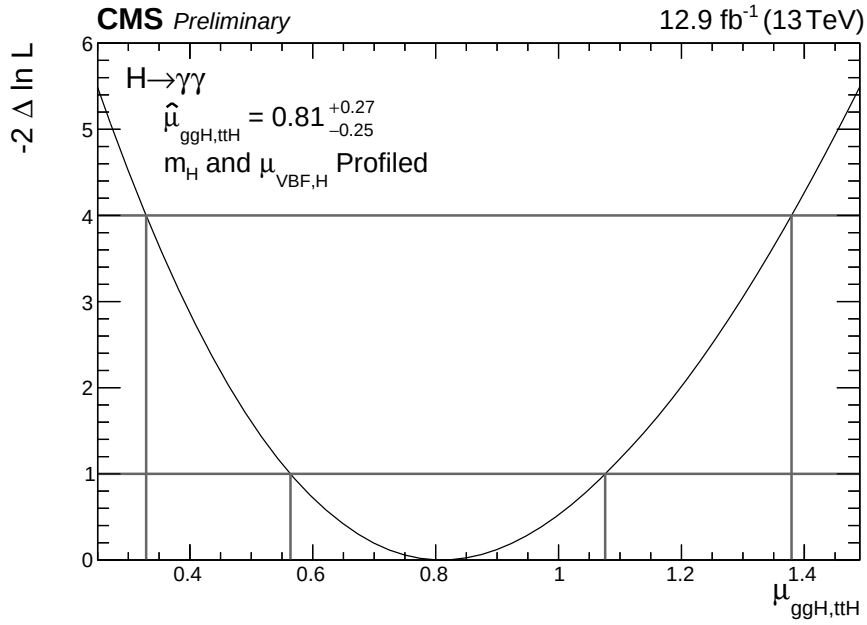
Figure A.11: The set of candidate functions chosen to parametrise the background using the discrete profiling method in the VBF and ttH categories. For each category, all candidate functions give acceptable agreement with the data, but can lead to large variations in the predicted number of events in the region of interest between 120 and 130 GeV. The resulting uncertainty in the choice of parametrisation is handled by the discrete profiling method.

Appendix B

Additional 2ΔNLL scans for measurements

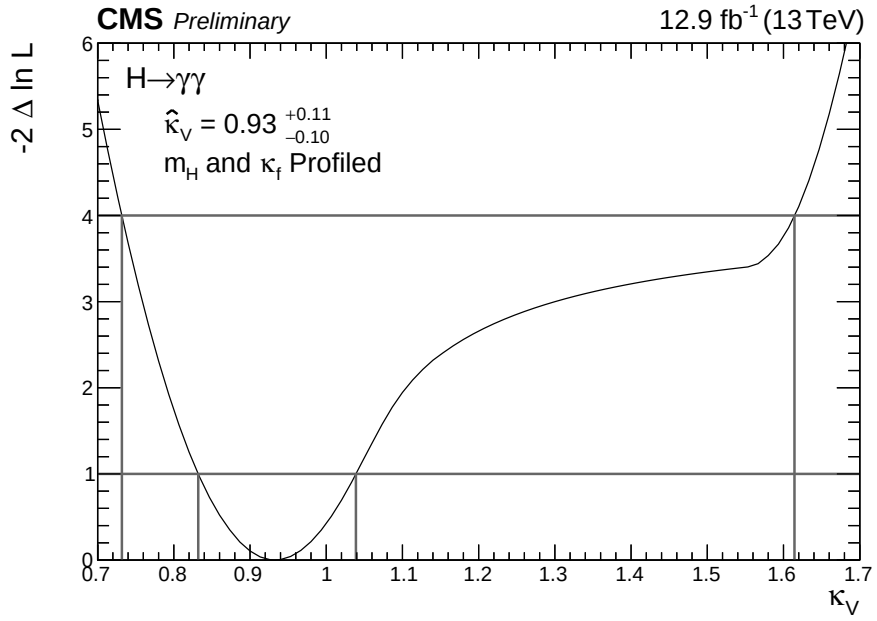


(a) 2ΔNLL scan of $\mu_{\mathrm{VBF,VH}}$, profiling $\mu_{\mathrm{ggH,ttH}}$ and m_{H}

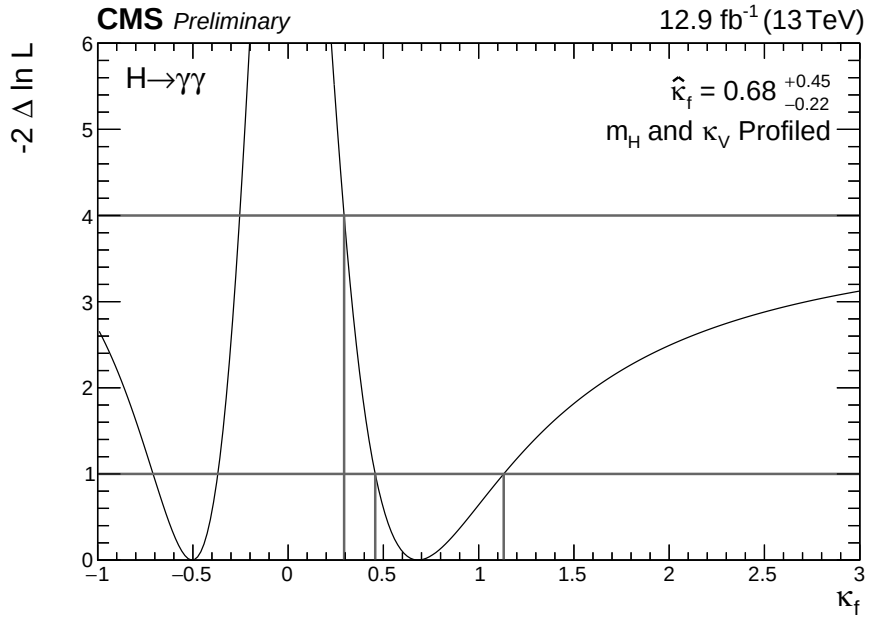


(b) 2ΔNLL scan of $\mu_{\mathrm{ggH,ttH}}$, profiling $\mu_{\mathrm{VBF,VH}}$ and m_{H}

Figure B.1: The result of performing 2ΔNLL scans of $\mu_{\mathrm{ggH,ttH}}$ while profiling $\mu_{\mathrm{VBF,VH}}$ (a) and vice versa (b). In both cases the mass of the Higgs boson is profiled in the fit.

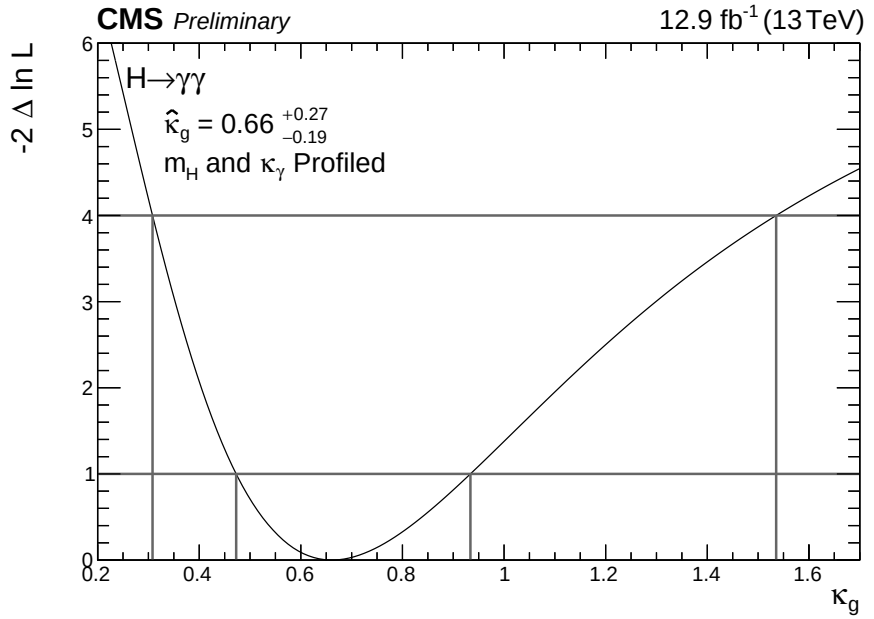


(a) 2ΔNLL scan of κ_{V}, profiling κ_{f} and m_{H}

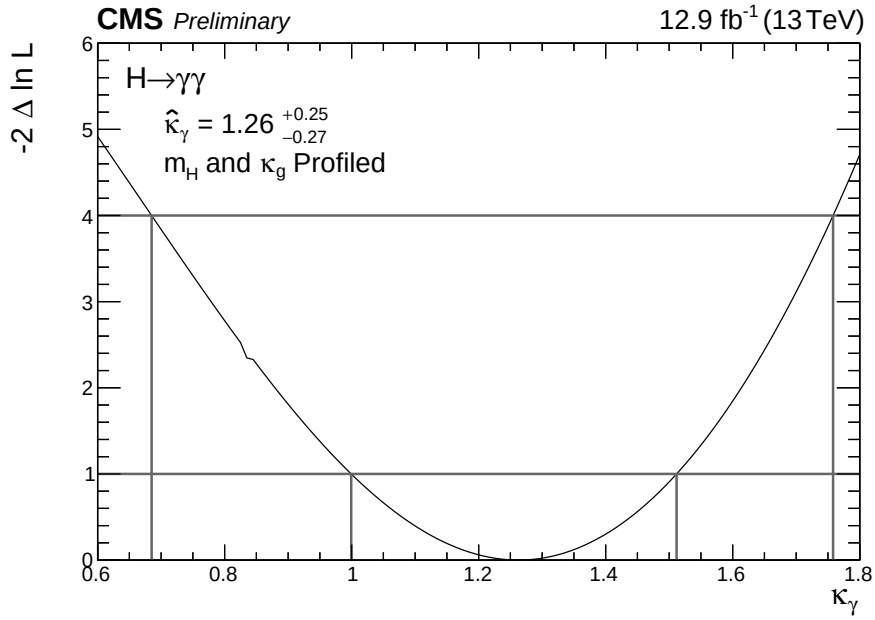


(b) 2ΔNLL scan of κ_{f}, profiling κ_{V} and m_{H}

Figure B.2: The result of performing 2ΔNLL scans of κ_{f} while profiling κ_{V} (a) and vice versa (b). In both cases the mass of the Higgs boson is profiled in the fit.



(a) 2ΔNLL scan of κ_{g}, profiling κ_{γ} and m_{H}



(b) 2ΔNLL scan of κ_{γ}, profiling κ_{g} and m_{H}

Figure B.3: The result of performing 2ΔNLL scans of κ_{g} while profiling κ_{γ} (a) and vice versa (b). In both cases the mass of the Higgs boson is profiled in the fit.

Bibliography

[1] CMS Collaboration, "First results on Higgs to $\gamma\gamma$ at 13 TeV", Technical Report CMS-PAS-HIG-15-005, CERN, Geneva, (2016).

[2] CMS Collaboration, "Updated measurements of Higgs boson production in the diphoton decay channel at $\sqrt{s} = 13$ TeV in pp collisions at CMS.", Technical Report CMS-PAS-HIG-16-020, CERN, Geneva, (2016).

[3] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[4] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012), no. 1, 1 – 29, doi:/10.1016/j.physletb.2012.08.020.

[5] CMS Collaboration, "Observation of the diphoton decay of the Higgs boson and measurement of its properties", *Eur. Phys. J.* **C74** (2014), no. 10, 3076, doi:10.1140/epjc/s10052-014-3076-z, arXiv:1407.0558.

[6] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys.* **C38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.

[7] M. Thomson, "Modern particle physics". Cambridge University Press, New York, 2013.

[8] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", *Phys. Rev. Lett.* **13** (1964) 321, doi:10.1103/PhysRevLett.13.321.

[9] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132, doi:10.1016/0031-9163(64)91136-9.

[10] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 508, doi:10.1103/PhysRevLett.13.508.

[11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles", *Phys. Rev. Lett.* **13** (1964) 585, doi:10.1103/PhysRevLett.13.585.

[12] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons", *Phys. Rev.* **145** (1966) 1156, doi:10.1103/PhysRev.145.1156.

[13] T. W. B. Kibble, "Symmetry breaking in non-Abelian gauge theories", *Phys. Rev.* **155** (1967) 1554, doi:10.1103/PhysRev.155.1554.

[14] E. Noether, "Invariant Variation Problems", *Gott. Nachr.* **1918** (1918) 235–257, doi:10.1080/00411457108231446, arXiv:physics/0503066. [Transp. Theory Statist. Phys.1,186(1971)].

[15] D. Griffiths, "Introduction to Elementary Particles". Physics textbook. Wiley, 2008.

[16] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349, doi:10.1103/PhysRevLett.30.1346.

[17] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Non-Abelian Gauge Theories", *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346, doi:10.1103/PhysRevLett.30.1343.

[18] S. L. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* **22** (1961) 579, doi:10.1016/0029-5582(61)90469-2.

[19] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[20] A. Salam, "Weak and electromagnetic interactions", in *Elementary particle physics: relativistic groups and analyticity*, N. Svartholm, ed., p. 367. Almqvist & Wiskell, Stockholm, 1968. Proceedings of the eighth Nobel symposium.

[21] U. M. Heller, M. Klomfass, H. Neuberger et al., "Numerical analysis of the Higgs mass triviality bound", *Nucl. Phys.* **B405** (1993) 555–573, doi:10.1016/0550-3213(93)90559-8, arXiv:hep-ph/9303215.

[22] S. L. Wu, "Brief history for the search and discovery of the Higgs particle - A personal perspective", *Mod. Phys. Lett.* **A29** (2014), no. 09, 1330027, doi:10.1142/S0217732313300279, arXiv:1403.4425.

[23] OPAL Collaboration, DELPHI Collaboration, LEP Working Group for Higgs boson searches, ALEPH Collaboration, L3 Collaboration, "Search for the standard model Higgs boson at LEP", *Phys. Lett.* **B565** (2003) 61–75, doi:10.1016/S0370-2693(03)00614-2, arXiv:hep-ex/0306033.

[24] TEVNPH Collaboration, CDF Collaboration, D0 Collaboration, CDF Collaboration, "Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data", Technical Report FERMILAB-CONF-12-065-E, CDF-NOTE-10806, D0-NOTE-6303, FERMILAB, Chicago, (2012).

[25] P. B. Renton, "Electroweak fits and constraints on the Higgs mass", in *Proceedings, 32nd International Conference on High Energy Physics (ICHEP 2004): Beijing, China, August 16-22, 2004*, pp. 564–567. 2004. arXiv:hep-ph/0410177.

[26] ATLAS Collaboration and CMS Collaboration, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", *Phys. Rev. Lett.* **114** (May, 2015) 191803, doi:10.1103/PhysRevLett.114.191803.

[27] B. Mellado Garcia, P. Musella, M. Grazzini et al., "CERN Report 4: Part I Standard Model Predictions", Technical Report LHCHXSWG-DRAFT-INT-2016-008, CERN, Geneva, (May, 2016).

[28] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008), no. 08, S08001, doi:10.1088/1748-0221/3/08/S08001.

[29] LEP Injector Study Group, "LEP design report, volume I: The LEP injector chain; LEP design report, volume II: The LEP Main Ring". CERN, Geneva, 1983.

[30] C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN.", Technical Report CERN-DI-0812015, CERN, Geneva, (Dec, 2008).

[31] M. Benedikt, P. Collier, V. Mertens et al., "LHC Design Report", Technical Report CERN-2004-003-V-3, Geneva, (2004).

[32] CMS Collaboration, "CMS Luminosity - Public Results". http://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[33] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

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

[35] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[36] LHCb Collaboration, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[37] CMS Collaboration, "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006.

[38] CMS Collaboration, "The CMS tracker system project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

[39] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[40] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

[41] CMS Collaboration, "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015) P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.

[42] CMS Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in *pp* Collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009, doi:10.1088/1748-0221/8/09/P09009, arXiv:1306.2016. [JINST8,9009(2013)].

[43] CMS Collaboration, "CMS ECAL first results with 2016 data". https://twiki.cern.ch/twiki/bin/view/CMSPublic/EcalDPGResultsCMSDPS2016031.

[44] CMS Collaboration, "The CMS hadron calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

[45] CMS Collaboration, "The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c", *Eur. Phys. J.* **C60** (2009) 359–373, doi:10.1140/epjc/s10052-009-0959-5,10.1140/epjc/s10052-009-1024-0. [Erratum: Eur. Phys. J. C61, (2009) 353].

[46] CMS Collaboration, "Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[47] CMS Collaboration, "The CMS muon project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

[48] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", cms physics analysis summary, (2009).

[49] CMS Collaboration, " Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", Technical Report CMS-PAS-PFT-10-001, CERN, Geneva, (2010).

[50] J. Alwall, R. Frederix, S. Frixione et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[51] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036.

[52] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J.* **C76** (2016), no. 3, 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[53] T. Gleisberg, S. Hoeche, F. Krauss et al., "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622.

[54] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[55] CMS Collaboration, "Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **10** (2011) 132, doi:10.1007/JHEP10(2011)132, arXiv:1107.4789.

[56] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[57] T. J. Hastie, R. J. Tibshirani, and J. H. Friedman, "The elements of statistical learning : data mining, inference, and prediction". Springer series in statistics. Springer, New York, 2009. Autres impressions : 2011 (corr.), 2013 (7e corr.).

[58] J. H. Friedman, "Greedy function approximation: A gradient boosting machine.", *Ann. Statist.* **29** (10, 2001) 1189–1232, doi:10.1214/aos/1013203451.

[59] A. Hocker, J. Stelzer, F. Tegenfeldt et al., "TMVA - Toolkit for Multivariate Data Analysis", *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[60] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework", in *Proceedings AIHENP'96 Workshop, Lausanne*, volume 389 (1997) 81-86. September, 1996. See also http://root.cern.ch/.

[61] J. Gaiser, "Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'". PhD thesis, SLAC, 1982.

[62] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063.

[63] D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, "Probing color singlet exchange in Z + two jet events at the CERN LHC", *Phys. Rev.* **D54** (1996) 6680–6689, doi:10.1103/PhysRevD.54.6680, arXiv:hep-ph/9605444.

[64] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[65] G. Cowan, K. Cranmer, E. Gross et al., "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.* **C71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727. [Erratum: Eur. Phys. J. C73, (2013) 2501].

[66] P. D. Dauncey, M. Kenzie, N. Wardle et al., "Handling uncertainties in background shapes: the discrete profiling method", *JINST* **10** (2015), no. 04, P04015, doi:10.1088/1748-0221/10/04/P04015, arXiv:1408.6865.

[67] "Lognormal Distributions: Theory and Applications". Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 1987.

[68] F. Demartin, S. Forte, E. Mariani et al., "The impact of PDF and alphas uncertainties on Higgs Production in gluon fusion at hadron colliders", *Phys. Rev.* **D82** (2010) 014002, doi:10.1103/PhysRevD.82.014002, arXiv:1004.0962.

[69] S. Carrazza, S. Forte, Z. Kassabov et al., "An Unbiased Hessian Representation for Monte Carlo PDFs", *Eur. Phys. J.* **C75** (2015), no. 8, 369, doi:10.1140/epjc/s10052-015-3590-7, arXiv:1505.06736.

[70] I. W. Stewart and F. J. Tackmann, "Theory Uncertainties for Higgs and Other Searches Using Jet Bins", *Phys. Rev.* **D85** (2012) 034011, doi:10.1103/PhysRevD.85.034011, arXiv:1107.2117.

[71] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput.Phys.Commun.* **10** (1975) 343–367, doi:10.1016/0010-4655(75)90039-9.

[72] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling", *eConf* **C0303241** (2003) MOLT007, arXiv:physics/0306116. [,186(2003)].

[73] G. Cowan, "Statistical Data Analysis". Oxford University Press, Oxford, 1998.

[74] ATLAS, CMS Collaboration, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.

[75] F. P. An, J. Z. Bai, A. B. Balantekin et al., "Observation of Electron-Antineutrino Disappearance at Daya Bay", *Phys. Rev. Lett.* **108** (April, 2012) 171803, doi:10.1103/PhysRevLett.108.171803, arXiv:1203.1669.

[76] D. Clowe, M. Bradac, A. H. Gonzalez et al., "A Direct Empirical Proof of the Existence of Dark Matter", *Astrophys. J. Lett.* **648** (2006), no. 2, L109.

[77] K. Agashe, R. Contino, and A. Pomarol, "The Minimal composite Higgs model", *Nucl. Phys.* **B719** (2005) 165–187, doi:10.1016/j.nuclphysb.2005.04.035, arXiv:hep-ph/0412089.

[78] N. Craig, J. Galloway, and S. Thomas, "Searching for Signs of the Second Higgs Doublet", arXiv:1305.2424.

[79] B. Patt and F. Wilczek, "Higgs-field portal into hidden sectors", arXiv:hep-ph/0605188.

[80] S. P. Martin, "A Supersymmetry primer", doi:10.1142/9789812839657_0001,10.1142/9789814307505_0001, arXiv:hep-ph/9709356. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[81] CMS Collaboration, "Updates on Projections of Physics Reach with the Upgraded CMS Detector for High Luminosity LHC", Technical Report CMS-DP-2016-064, CERN, Geneva, (Oct, 2016).

List of Acronyms

2NLL twice the negative log-likelihood

ADC analogue to digital converters

ALICE A Large Ion Collider Experiment

APDs avalanche photon-diodes

BDT boosted decision tree

CB Crystal Ball function

CSCs cathode strip chambers

DCB double Crystal Ball function

DTs drift tubes

DT decision tree

DY Drell-Yann

EB ECAL barrel

ECAL electromagnetic calorimeter

EE ECAL endcaps

ES preshower

EWT electroweak theory

ggH gluon-gluon fusion

HB hadron calorimeter barrel

HCAL hadron calorimeter

HE hadron calorimeter endcaps

HF forward hadron calorimeter

HLT high-level trigger

HO outer hadron calorimeter

L1T level-1 trigger

LHCHXSWG LHC Higgs Cross Section Working Group

LHCb Large Hadron Collider beauty

LINAC2 Linear Accelerator 2

MC Monte Carlo

MET Missing Transverse Energy

MVA multi-variate analysis

MVA multi-variate analysis

NLO next-to-leading order

PDF parton distribution functions

PFCHS PF charged hadron subtraction

PF the particle-flow algorithm

POI parameter of interest

PSB Proton Synchrotron Booster

PS Proton Synchrotron

PV primary vertex

QCD quantum chromodynamics

QED quantum electrodynamics

QFT quantum field theory

RPCs resistive plate chambers

RV right vertex

SC supercluster

SM standard model

SPS Super Proton Synchrotron

SSF simultaneous signal fitting

ttH top quark fusion and associated production

VBF vector boson fusion

VEV vacuum expectation value

VH vector boson associated production

VPTs vacuum photon-triodes

WH W boson associated production

WV wrong vertex

ZH Z boson associated production