Precision measurements of Higgs boson properties with beyond-the-Standard Model interpretations at the CMS experiment

Jonathon Mark Langford

Imperial College London
Department of Physics

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

Abstract

This thesis details the precision measurements of Higgs boson properties at the Compact Muon Solenoid experiment. The measurements use proton-proton collision data at a centre-of-mass energy of 13 TeV, collected during Run 2 of the LHC, corresponding to a total integrated luminosity of 137 fb^{-1}. Production cross sections and Higgs boson couplings are measured in the diphoton decay channel. Events with two isolated photons are selected and subsequently categorised to target different kinematic regions of the Higgs boson production phase-space. The total Higgs boson signal-strength, relative to the Standard Model prediction, is measured to be 1.12 ± 0.09. Other properties of the Higgs boson are measured. This includes a simultaneous measurement of 27 independent kinematic regions, representing the most granular measurement of Higgs boson production in a single decay channel to-date. Following this, a beyond-the-Standard Model interpretation of Higgs boson cross section measurements is provided. The interpretation is performed in an effective field theory framework, which benefits from being agnostic to the specifics of the complete beyond-the-Standard Model theory. Measurements from multiple Higgs boson decay channels are combined, enabling tighter constraints on a larger number of effective field theory parameters, and thereby reducing the parameter space for potential beyond-the-Standard Model physics. Ultimately, all results presented in this thesis are found to be consistent with the predictions of the Standard Model.

To Mum, Dad,
and my sister Hannah,

Declaration

I, the author, declare that all work presented in this thesis is my own. The studies have been conducted in collaboration with several individuals, as members of the Compact Muon Solenoid (CMS) Collaboration. All figures produced by the author, or in collaboration with the author, are labelled as "CMS" if taken directly from a CMS publication, "CMS Preliminary" if taken from a CMS preliminary public document, and figures which have not explicitly been made public by the CMS Collaboration attain the label "CMS Work in Progress", or are simply left unlabelled. The figures and studies that are taken from external sources, and are not produced directly by the author, or in collaboration with the author, are appropriately referenced throughout.

Jonathon Mark Langford

Acknowledgements

First and foremost, I would like to thank the Imperial HEP group for inviting me to perform my doctoral research at the college. The group is a collection of fascinating and vibrant minds, that have imparted on me an unwavering enthusiasm for the subject. Also, I thank the President's PhD Scholarship scheme at Imperial College for providing my funding over the course of this PhD, and allowing me to travel to CERN, in addition to other international schools and conferences.

Although there are many people at Imperial that have had a measurable impact on my PhD experience, there are a few individuals that I would like to thank in particular. Nick, thank you for always directing me towards, what I would say are, interesting analyses. All the way through the PhD, your experience and knowledge has been invaluable, and without which, a large part of this thesis would not have been possible. It has been a pleasure working together, and I am looking forward to carrying this on in the future. Gavin, you have always been there if I have needed anything; it's quite remarkable considering how much you are doing. Thank you for your incredible guidance, and always making me feel valued within the group. Ed, I could not have asked for a better mentor. I think the way you approach your work has rubbed off on me, enabling me to stay calm (and sane), and always appreciate the bigger picture. Together with Joe and Shameena, I am very proud of our work on the $H \rightarrow \gamma\gamma$ analysis, which makes up a large part of this thesis. Most importantly, thanks for keeping my drink shelf stocked. Finally, to all my friends doing a PhD at both Imperial and CERN, it has been a great pleasure going through this experience with you. Although cut short, I am grateful for the times we did have, which were always filled with interesting conversation and laughter. Good luck to you all!

Of course, thank you to the CMS Collaboration. This thesis is built upon the hard work of thousands of CMS members in designing, building, and operating a beautiful experiment at the LHC. I am honoured to be a part of it all.

Ultimately, this thesis not only represents the accumulation of my work over the past three years, but it serves as a reflection of the many amazing people that have cared for, helped and influenced me throughout my life. To my friends and family, thanks for the

laughter and all-round good times. You have helped me escape from physics when I have needed it most. Olivia, thank you for enriching my life with your beautiful soul. The emotional support you have provided has kept me going through the hardest times in the past few years. To Hannah, thank you for all you have done for me in our growing up together; I am immensely proud to call you my sister, and to little Zo, your smile and laughter provides unrivalled joy and I cannot wait to read this thesis to you next Christmas at bedtime. To Dad, thanks for teaching me (through Wigan Athletic) that the hard times make the good times so much sweeter. Despite not being able to see you over the last year, I know that your support is ever present and our celebrations, when possible, will make it all seem worthwhile. Your guidance and knowledge put me on the trajectory towards this PhD, and for that I am truly thankful. And to Mum, I cannot stress the impact your love and support has had on my life. Thanks to you and Jim for welcoming me back into your home, and providing a calm and (almost) distraction-free space in which to write this thesis. It was a time that I did not expect to have with you, but it was filled with many happy memories, and for that I am extremely grateful. I love you very much.

Contents

List of Figures

List of Tables

Preface

This thesis includes a complete description of the $H \rightarrow \gamma\gamma$ analysis documented in Ref. [1]. The analysis targets Higgs boson production cross sections and couplings, using proton-proton collision data collected by the CMS experiment during Run 2 of the LHC. The paper has recently been published by the Journal of High Energy Physics (JHEP). I was predominantly responsible for the statistical inference and the extraction of results, detailed in Chapters 5 and 6, respectively. The event reconstruction and categorisation, described in Chapter 4, was developed and optimised by other members of the $H \rightarrow \gamma\gamma$ analysis group at the CMS Collaboration. Nevertheless, the techniques have been summarised in my own words, and I am responsible for the derivation of the final analysis category yields.

Described in Chapter 7, is a beyond-the-Standard Model interpretation of Higgs boson cross section measurements at CMS, using an effective field theory approach. The interpretation was made public by the CMS Collaboration in the preliminary public document of Ref. [2], which describes the combination of Higgs boson measurements across the major Higgs boson decay channels using the partial Run 2 data set. Whilst the strategy for such an interpretation was developed by others, I was responsible for implementing this strategy at the CMS experiment, deriving the signal cross section parametrisation, and subsequently extracting the results. Furthermore, the progression of EFT measurements at CMS, described at the end of the chapter, includes only my own studies.

Finally, Chapter 8 looks ahead to the High-Luminosity operation of the LHC machine (HL-LHC). The training of the algorithm responsible for distinguishing electrons and photons from hadronic activity is my own work. A description of this algorithm was included in the CMS Phase-2 Level-1 Trigger technical design report [3]. The chapter also details a projection study looking at the sensitivity to top-associated Higgs boson differential cross sections with the CMS Phase-2 detector. I am responsible for all elements of this analysis, including the extraction of constraints on the Higgs boson self-coupling. This study was published in a collection of HL-LHC sensitivity studies for the CERN Yellow Report in Ref. [4].

"Things are so hard to figure out
when you live from day to day
in this feverish and silly world."

Jack Kerouac, On the Road

Chapter 1

Introduction

Our current best theory for describing elementary particles and their interactions is known as the Standard Model (SM) of particle physics [5–7]. The SM successfully describes the strong force as well as the unification of the weak and electromagnetic forces. Central to this so-called electroweak theory is the Brout-Englert-Higgs (BEH) mechanism [8–13], responsible for the spontaneous symmetry breaking of the electroweak interaction and the subsequent generation of mass for other elementary particles. Additionally, the BEH mechanism predicts the existence of a scalar boson with mass situated around the electroweak scale; namely the Higgs boson. In 2012, the Higgs boson was observed experimentally [14–16] by the ATLAS and CMS Collaborations [17, 18] at the Large Hadron Collider (LHC) [19], CERN. This discovery was heralded as a great triumph of the SM, marking the completion of the particle content of the theory.

Despite all its glory, the SM is known to be incomplete. For one, the SM fails to explain the fourth fundamental force of nature: gravity. Secondly, cosmological models which successfully describe the large scale structure of the universe predict the existence of dark matter and dark energy [20], which hold no place in SM theory. The SM also fails to explain the fine-tuning of the Higgs field vacuum expectation value, known as the hierarchy problem [21, 22], and it does not explain the fact that neutrinos have mass, as required by the experimental observation of neutrino oscillations [23]. On top of this, the SM lacks a mathematical elegance by requiring a relatively large number of input parameters, whose values are not predicted by the theory and appear to be somewhat arbitrary, often provided by experiment [24]. It is these shortcomings of the theory which make necessary the existence of new physics beyond-the-Standard Model (BSM).

At the LHC, there are two complementary methods used to search for BSM physics. The direct approach aims to explicitly observe new particles in data. Alas, since the discovery of the Higgs boson in 2012 there has been no direct evidence of new particles, suggesting that any BSM physics lies beyond the energy reach of the collider, or that it has

a too-small cross section and therefore remains (so far) undetected. As a result, attention has shifted towards the second approach: indirectly probing new physics via precision measurements. Short-range interactions with BSM particles may leave a measurable imprint on the properties of SM particles. Hence, quantities well predicted in SM theory offer a unique tool for discovery, where deviations between measured and predicted values provide an indication of new physics.

Precision measurements of Higgs boson properties will provide a better understanding of electroweak symmetry breaking, help to elucidate the nature of the Yukawa interactions with fermions, and even shed light on the origins of the universe via the shape of the Higgs potential [25, 26]. Furthermore, being the only fundamental scalar in the SM, the Higgs boson lies at the heart of many proposed BSM theories, such as supersymmetry, composite models or extra dimensions [27–29]. This has led to the development of a broad and comprehensive program of work to characterise the Higgs boson and measure its couplings to other particles. Since discovery, the ATLAS and CMS Collaborations have observed all the major Higgs boson production modes, as well as the couplings of the Higgs boson to the third generation quarks and charged lepton [30–35]. Most recently, the CMS experiment reported the first direct evidence of the Higgs boson coupling to the muon [36]. Moreover, both experiments have performed differential measurements of Higgs boson properties to further scrutinise SM theory in specific regions of the Higgs boson phase space [37–41]. So far all measurements are consistent with the SM predictions.

This thesis details precision measurements of Higgs boson properties using proton-proton collision data collected by the CMS experiment during Run 2 of the LHC. Firstly, Chapter 2 provides the theoretical foundations on which the measurements reside. Chapter 3 then describes the CMS experiment at the LHC, focusing on the design choices which enable Higgs boson precision measurements. The following three chapters provide a complete description of the analysis documented in Ref. [1], which measures Higgs boson production cross sections and couplings in the diphoton decay channel ($H \rightarrow \gamma\gamma$). This channel is particularly powerful due to its clean final state topology, where the excellent energy resolution of the CMS electromagnetic calorimeter leads to a narrow peak in the diphoton invariant mass spectrum, effectively distinguishing Higgs boson production from SM background processes. Furthermore, it is one of the few channels which has reasonable sensitivity to all of the principal Higgs boson production modes.

The $H \rightarrow \gamma\gamma$ analysis is configured to perform measurements within the simplified template cross section (STXS) framework [42], which offers a coherent approach to Higgs boson production cross section measurements with increasing statistics. In the framework, the inclusive Higgs boson production phase space is divided into kinematic regions, which are split first by the production mode and subsequently by the kinematics of the event

constituents. By measuring the cross section in these bins, a more complete description of Higgs boson production is achieved. The three chapters dedicated to this analysis are structured as follows. In Chapter 4, the event reconstruction and categorisation are described. Following this, Chapter 5 explains the statistical inference techniques which are used to extract the Higgs boson cross sections and couplings. This includes the construction of the likelihood function, the modelling of both signal and background events, and the treatment of systematic uncertainties. Finally, Chapter 6 presents the results of the analysis in terms of signal-strengths, coupling-modifiers and production cross sections in the STXS framework.

Following on from this, Chapter 7 describes a BSM interpretation of STXS measurements using effective field theory (EFT) [43–47]. In the EFT, BSM particles are postulated to have masses at an energy scale far beyond the accessible energy scale at the LHC. By performing a series expansion of the SM Lagrangian, we encapsulate all the information of a BSM ultraviolet (UV) complete theory, in terms of the infrared (IR) SM fields. This provides an (almost) model independent framework on which to probe BSM physics. Firstly, cross sections in the STXS framework and the Higgs boson branching fractions are expressed in terms of the EFT parameters. This parametrisation is then applied to a combination of STXS measurements performed by the CMS experiment in different Higgs boson decay channels [2]. Combining measurements in this way provides the optimal precision with regards to the current data set, and enables a number of directions in the EFT parameter space to be probed simultaneously. The results are shown as constraints on a set of EFT parameters, thereby reducing the potential parameter space for BSM physics.

The thesis concludes with a look to the future operation of the LHC machine, known as the High-Luminosity LHC (HL-LHC). Here, protons will be collided at five times the LHC design luminosity, providing a wealth of proton-proton collision data on which to base precision measurements. In Chapter 8, a machine learning algorithm trained to discriminate photons and electrons from hadronic activity in the future CMS trigger system is described [3], showing the expected performance of the algorithm in the HL-LHC environment. Following this, the future sensitivity to Higgs boson measurements at the HL-LHC is discussed. This features a projection study investigating an indirect method for probing the Higgs boson self-interaction via (single) Higgs boson production in association with top-quarks [48]. Finally, Chapter 9 draws conclusions from the results of this thesis, and offers a perspective on the future of Higgs boson precision measurements at the LHC and beyond.

Chapter 2

Theoretical foundations

2.1 The Standard Model of particle physics

2.1.1 Introduction

The Standard Model (SM) of particle physics is a relativistic quantum field theory (QFT), that encompasses all known fundamental particles of matter and their interactions via the strong nuclear force, weak nuclear force, and the electromagnetic force. Gravity, which is not included in SM theory, is many orders of magnitude weaker than the other forces and therefore can be safely ignored when studying the interactions of SM particles in high-energy physics experiments. SM theory was established in the 1970s and has remained essentially unchanged since its formation [5–7], providing extremely successful high-energy physics predictions up to the modern day. Most notably, was the prediction of the Higgs boson [8–13], which was observed experimentally in 2012 by the ATLAS and CMS Collaborations [14–16]. Nevertheless, the theory falls short in explaining a number of physical observations, including dark matter [20], neutrino oscillations [23], and the hierarchy problem [21, 22]. To be able to efficiently scrutinise the predictions of the SM in the search for beyond-the-Standard Model (BSM) physics, it is crucial that the theory is well understood.

This chapter begins by introducing the pillars of SM theory. Section 2.1.2 summarises the particle content of the SM, including both the fundamental constituents of matter (fermions) and the interaction mediators (bosons). Following this, a pedagogical approach to the construction of the SM is provided, adopting the widely-used Lagrangian formalism throughout. Here, the essential notion of a gauge theory is introduced, in the context of quantum electrodynamics (QED), quantum chromodynamics (QCD), and the unification of the electromagnetic and weak forces into the electroweak (EW) interaction. This leads to a description of spontaneous symmetry breaking in the SM, known as the Brout-

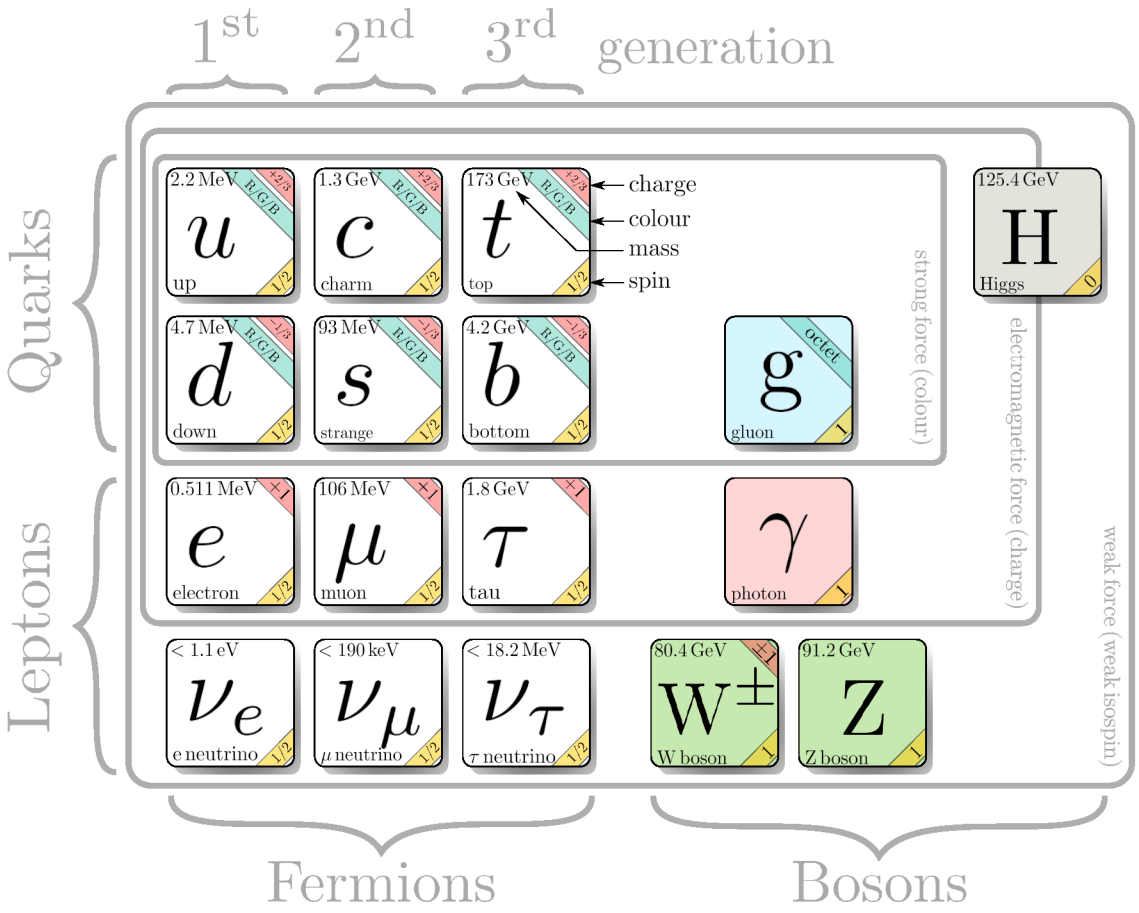


Figure 2.1: Particle content of the SM. The charge, colour, and spin quantum numbers are provided for each particle, in addition to their measured masses. All masses are taken from Ref. [24], except the Higgs boson (H) mass which is taken from Ref. [49].

Englert-Higgs (BEH) mechanism, which is vital for explaining how particles acquire mass.

2.1.2 Particle content

In the SM, the fundamental particles of matter and interaction mediators are described as relativistic fields; it is the quantised excitation of these fields which manifest as the physical particles that we observe in nature. The complete particle content of the SM is shown in Figure 2.1.

Ignoring the Higgs boson (H) for now, the SM particles can be divided into two groups: the spin-$\frac{1}{2}$ fermions which comprise the matter content of the SM, and the spin-1 gauge bosons which mediate the fundamental interactions. The fermions themselves come in two types: those which interact via the strong force are referred to as *quarks*, whilst those that do not are referred to as *leptons*. The quarks (q) are defined in three generations (at different mass scales) and are split into up-type and down-type quarks with charges

of $+\frac{2}{3}$ and $-\frac{1}{3}$, respectively (in units of the elementary charge, e). Quarks also possess a colour charge, which comes in one of three possible states: R, G and B. It is through this colour charge that the quarks interact with the massless gluon (g) colour octet via the strong interaction.

Likewise, the lepton sector is defined in three mass generations, and is split into the charged leptons (ℓ) and the neutral leptons (neutrinos, ν). Along with the quarks, the charged leptons interact with the massless photon field, γ, to define the electromagnetic interaction. All fermions[1], including the neutrinos, interact via the weak interaction. This occurs via the exchange of the massive vector bosons: $W^{\pm}$ bosons (charged current interactions) and the Z boson (neutral current interactions). Each fermion in the SM has a corresponding antiparticle, which has the same mass but opposite charge and parity to the respective particle[2].

The final piece in the puzzle is the spin-0 Higgs boson. Electroweak symmetry breaking via the BEH mechanism is essential in the SM as the means in which the $W^{\pm}$ and Z bosons acquire their mass. Additionally, the Yukawa interactions between the quarks/charged leptons and the Higgs field explain the masses of the fermions. The salient feature of the BEH mechanism is the existence of an additional scalar boson: the Higgs boson, the measurements of which form the basis of this thesis.

2.1.3 Constructing the Lagrangian

The SM can be neatly expressed as a Lagrangian density, $\mathcal{L}_{\mathrm{SM}}$, in terms of the particle fields. We can then infer the dynamics and interactions of the fields by applying the principle of least action to $\mathcal{L}_{\mathrm{SM}}$ (via the Euler-Lagrange equations).

At the bedrock of SM theory lies the idea of *symmetry*. Nöther's theorem [50] states that if a particular Lagrangian density is invariant under some transformation, then this implies the existence of an associated conservation law. Such an invariant transformation is referred to as a symmetry of the Lagrangian. By inverting the theorem, we find that for each conserved quantity observed in nature, there must be an associated symmetry of the corresponding Lagrangian. For example, a theory which respects the conservation of energy (momentum) is defined by a Lagrangian which is invariant under a temporal (spatial) translation.

The SM is based on a particular type of QFT, known as a *gauge theory*, where the Lagrangian is invariant under *local gauge transformations*. Such transformations shift the phase of the fundamental fields, where (in contrast to global transformations) the size

[1] There is a slight complexity here concerning the helicity states of the weakly-interacting particles, but this will be addressed in the coming sections.

[2] In fact, it is not yet known if the neutrino behaves as a Dirac fermion or a Majorana fermion, where the latter describes a fermion that is its own antiparticle. This subtle detail is not relevant for this thesis.

of the shift can be different for different points in spacetime. It is worth stressing that the fundamental fields of a QFT are not observed in nature. Instead, in high-energy physics experiments we measure quantities derived from the Lagrangian of the theory, such as cross sections, σ, and decay rates, Γ, which are combined under the blanket term of *observables*. Crucially, these observables are independent of the phase of the underlying fields, and therefore $\mathcal{L}_{\text{SM}}$ must be invariant under local gauge transformations. An important consequence of requiring this symmetry is the introduction of *gauge boson* fields which act as the mediators of the fundamental interactions. Additionally, Nöther's theorem tells us that a local gauge symmetry leads to the conservation of a physical charge.

In summary, imposing the invariance of $\mathcal{L}_{\text{SM}}$ under a local gauge transformation introduces gauge boson fields with certain properties, which interact in certain ways with the fundamental matter fields. The form of these interactions are governed by the *symmetry group* of the gauge transformation. The following sections describe the symmetries of $\mathcal{L}_{\text{SM}}$ that match what we observe in nature, and how they give rise to the electromagnetic, strong and weak interactions.

Quantum electrodynamics

The QFT of electromagnetism is known as quantum electrodynamics (QED) [51]. It describes charged fermions, photons, and their subsequent interactions. Rather profoundly, all of QED can be derived by requiring the invariance of the QED Lagrangian, $\mathcal{L}_{\text{QED}} \subset \mathcal{L}_{\text{SM}}$ under local U(1) transformations:

$$\text{U}(1) : \mathcal{L}_{\text{QED}} \mapsto \mathcal{L}'_{\text{QED}} = \mathcal{L}_{\text{QED}}. \tag{2.1}$$

The equations of motion for spin-$\frac{1}{2}$ fermion fields, ψ, are described by the Dirac equation [52]. Therefore, $\mathcal{L}_{\text{QED}}$ contains a Dirac term of the form,

$$\mathcal{L}_{\text{QED}} \supset \mathcal{L}_{\text{Dirac}} = i\bar{\psi}\not{\partial}\psi - m_\psi\bar{\psi}\psi, \tag{2.2}$$

where we have adopted the slash notation, $\not{\partial} = \gamma^\mu\partial_\mu$, such that $i\partial_\mu$ is the position space representation of the 4-momentum, p_μ. The quantity γ^μ represents the Dirac gamma matrices that obey the anti-commutation relation $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{1}_4$, where $\eta^{\mu\nu}$ is the Minkowski-space metric: $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$. The fermion field, ψ, is expressed as a four-component Dirac spinor, where the components are interpreted as the spin-up and spin-down eigenstates of the positive energy (particle) and negative energy (antiparticle) solutions of the Dirac equation. The corresponding adjoint field, $\bar{\psi}$, is defined as $\bar{\psi} = \psi^\dagger\gamma^0$, where $\psi^\dagger$ is the Hermitian conjugate of ψ. Finally, the mass of the fermion, m_ψ, is

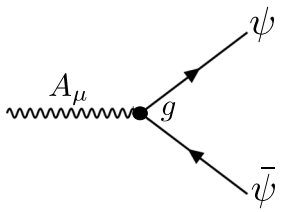


Figure 2.2: The electromagnetic interaction between the photon field, A_μ, and the charged fermion fields, ψ.

described by the mass term, $m_\psi \bar{\psi}\psi$.

Under a local U(1) transformation, the fermion field transforms as,

$$\mathrm{U}(1) : \psi \mapsto \psi' = e^{ig\theta(x)}\psi, \tag{2.3}$$

where g is a real number, and θ is a real-valued function that depends on the spacetime co-ordinate, x. Applying the transformation to $\mathcal{L}_{\mathrm{Dirac}}$,

$$\mathrm{U}(1) : \mathcal{L}_{\mathrm{Dirac}} \mapsto \mathcal{L}'_{\mathrm{Dirac}} = \mathcal{L}_{\mathrm{Dirac}} - \bar{\psi} g \gamma^\mu (\partial_\mu \theta(x))\psi, \tag{2.4}$$

introduces an additional term due to the non-vanishing derivative of $\theta(x)$.[3] To reconcile this with the U(1) gauge invariance shown in equation 2.1, it is necessary to add another term to $\mathcal{L}_{\mathrm{QED}}$,

$$\mathcal{L}_{\mathrm{int}} = -g\bar{\psi}\gamma^\mu \psi A_\mu, \tag{2.5}$$

where A_μ is a (vector) gauge boson field that transforms under local U(1) transformations as,

$$\mathrm{U}(1) : A_\mu \mapsto A'_\mu = A_\mu - \partial_\mu \theta(x). \tag{2.6}$$

This directly cancels the extra term in equation 2.4, establishing the gauge invariance of $\mathcal{L}_{\mathrm{QED}}$. What we have found is beautiful in its simplicity. By demanding the local U(1) gauge symmetry of the Dirac Lagrangian, we require the existence of the gauge boson field, A_μ, which is precisely the *photon field*. The additional term in the Lagrangian, $\mathcal{L}_{\mathrm{int}}$, describes the electromagnetic interaction between charged fermions and photons, as shown by the Feynman diagram in Figure 2.2, where the quantity $g = |e|\mathcal{Q}$ is the coupling strength of this interaction, with e being the elementary unit charge and $\mathcal{Q}$ the relative charge of fermion, ψ. The conserved quantity in the electromagnetic interaction is the electric charge.

[3] This is not present in global U(1) transformations where $\partial_\mu \theta = 0$.

The QED Lagrangian is supplemented with the gauge-invariant kinetic term,

$$\mathcal{L}_{\text{QED}} \supset \mathcal{L}_{\text{gauge}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{2.7}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor for the photon field, A_μ. This describes the propagation of the photon field in free space, such that by applying the Euler-Lagrange equations to $\mathcal{L}_{\text{gauge}}$ we recover Maxwell's equations of electromagnetism [51]. Hence, the total QED Lagrangian can be written as,

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\psi}\not{D}\psi - m_\psi\bar{\psi}\psi, \tag{2.8}$$

where we have introduced the *covariant derivative* formalism, $\not{D} = \gamma^\mu D_\mu = \gamma^\mu(\partial_\mu + igA_\mu)$, which absorbs the gauge boson field, A_μ, into the definition. Importantly, there is no mass term ($\propto m_A^2 A^\mu A_\mu$) present in the Lagrangian for the gauge boson, as terms of this type are not invariant under local gauge transformations. As a consequence, the photon field is massless and the range of the interaction is effectively infinite.

Extending to SU(N)

The strong and weak interactions are obtained by extending the concept of gauge invariance to SU(N) symmetry groups [53]. In the case of QED, gauge invariance was achieved by defining a covariant derivative containing the corresponding gauge boson field. This simple reformulation extends to all symmetry groups: gauge invariance is realised by introducing the covariant derivative,

$$D_\mu = \partial_\mu + igA_\mu^a \mathrm{T}^a \tag{2.9}$$

where A_μ^a are the gauge boson fields, T^a are the generators of the group algebra, and g is again the coupling strength of the gauge interaction. In the case of the U(1) symmetry group, there is a single generator ($\mathrm{T} = \mathbb{1}$) which results in a single gauge boson: the photon. For SU(N), the generators T^a form a basis in the group space of dimension $N^2 - 1$ (i.e. $a = 1, ..., N^2 - 1$). Symmetries of this type give rise to $N^2 - 1$ gauge bosons.

For a generic covariant derivative, D_μ, the corresponding field strength tensor is,

$$\begin{aligned}
\mathbf{F}_{\mu\nu} &= \frac{i}{g}[D_\mu, D_\nu] = \partial_\mu A_\nu^a \mathrm{T}^a - \partial_\nu A_\mu^a \mathrm{T}^a + ig[A_\mu^b \mathrm{T}^b, A_\nu^c \mathrm{T}^c] \\
&= \mathrm{T}^a\Big(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g f^{abc} A_\mu^b A_\nu^c\Big) \\
&= \mathrm{T}^a F_{\mu\nu}^a,
\end{aligned} \tag{2.10}$$

where f^{abc} are the structure constants of the symmetry group defined by $[\mathrm{T}^a, \mathrm{T}^b] = if^{abc}\mathrm{T}^c$.

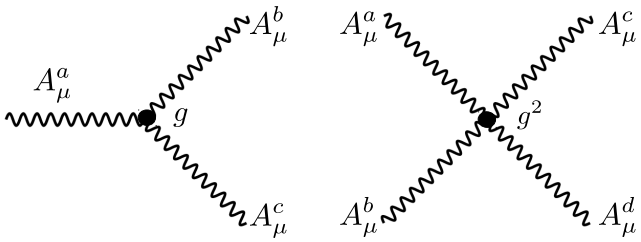


Figure 2.3: The triple (left) and quartic (right) gauge boson self-interactions.

Abelian groups are defined by the property that the generators of the group algebra commute: $f^{abc} = 0 \forall a, b, c$.[4] For non-Abelian gauge theories, such as those based on an SU(N) symmetry, the generators in general do not commute. This has dramatic consequences for the physics that the theory describes, as can be seen from the gauge term in the Lagrangian,

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F^{a,\mu\nu} F^a_{\mu\nu}. \tag{2.11}$$

Unlike in the case of QED, $\mathcal{L}_{\text{gauge}}$ now contains additional terms representing the self-interactions of the gauge bosons: a term $\propto g f^{abc} (\partial^\mu A^{a,\nu}) A^b_\mu A^c_\nu$ which corresponds to the triple gauge boson interactions, and a term $\propto g^2 f^{abc} f^{ade} A^{b,\mu} A^{c,\nu} A^d_\mu A^e_\nu$ which corresponds to the quartic gauge boson interactions. These self-interactions are shown as Feynman diagrams in Figure 2.3.

Quantum chromodynamics

The QFT of the strong interaction, known as quantum chromodynamics (QCD), is established by requiring $\mathcal{L}_{\text{QCD}} \subset \mathcal{L}_{\text{SM}}$ to be symmetric under local SU(3)$_{\text{C}}$ transformations:

$$\text{SU(3)}_{\text{C}} : \mathcal{L}_{\text{QCD}} \longrightarrow \mathcal{L}'_{\text{QCD}} = \mathcal{L}_{\text{QCD}}, \tag{2.12}$$

where the subscript $_{\text{C}}$ indicates that the transformation only applies to fields with colour charge. The gauge invariance is achieved by replacing ∂_μ with the covariant derivative,

$$D_\mu = \partial_\mu + i g_s G^a_\mu \frac{\lambda^a}{2}. \tag{2.13}$$

Here, $\frac{\lambda^a}{2}$ are the $3^2 - 1 = 8$ generators of the SU(3)$_{\text{C}}$ group algebra, known as the Gell-Mann matrices, and G^a_μ, correspond to the eight gluon fields which mediate the strong interaction. The strength of this interaction is encoded by the coupling strength, g_s.

By Nöther's theorem, we can relate this SU(3)$_{\text{C}}$ symmetry to the conservation of

[4] The U(1) group of QED is trivially Abelian, since it is defined by a single generator, T = $\mathbb{1}$.

colour charge. Quark fields possess a colour charge and therefore interact with the gluon fields. Including the Dirac terms for the quark fields, we arrive at an expression for the QCD Lagrangian,

$$\mathcal{L}_{\mathrm{QCD}} = \sum_j i\bar{\psi}_j(\not{D} - m_j)\psi_j - \frac{1}{4}G^{a,\mu\nu}G^a_{\mu\nu}. \tag{2.14}$$

where ψ_j are the quark spinors of flavour $j = \{u, d, c, s, t, b\}$. As described in the previous section, the non-commutative properties of the Gell-Mann matrices lead to the gluon self-interactions [54]. These self-interactions have major consequences for the physics of the strong interaction. Firstly, the coupling strength, g_s, decreases as a function of the interaction energy scale[5]. At very high energies, such as those probed in the hard scatterings at the LHC, the strong force is sufficiently weak to describe the quarks and gluons as independent particles using perturbation theory. This effect is referred to as *asymptotic freedom* [55]. On the other hand, at lower energies quarks and gluons are never observed as isolated colour charged particles, but are instead *confined* to colourless bound states, known as hadrons. This gives rise to the plethora of meson ($q\bar{q}$) and baryon/anti-baryon ($qqq/\bar{q}\bar{q}\bar{q}$) states that we observe in nature, and explains why outgoing quarks and gluons in high-energy collisions form collimated sprays of hadrons, known as jets [56].

Electroweak unification

One of the great successes of SM theory is the unification of the electromagnetic and weak interactions [5–7]. This was achieved by requiring the electroweak Lagrangian, $\mathcal{L}_{\mathrm{EW}} \subset \mathcal{L}_{\mathrm{SM}}$ to be symmetric under local $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ transformations,

$$\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y} : \mathcal{L}_{\mathrm{EW}} \longrightarrow \mathcal{L}'_{\mathrm{EW}} = \mathcal{L}_{\mathrm{EW}}, \tag{2.15}$$

which requires the introduction of the covariant derivative,

$$D_\mu = \partial_\mu + igW^i_\mu \frac{\sigma^i}{2} + ig'B_\mu \mathrm{Y}. \tag{2.16}$$

The non-Abelian $\mathrm{SU(2)_L}$ symmetry group has $2^2 - 1 = 3$ generators, $\mathrm{T}^i = \sigma^i/2$, which are the 2×2 Pauli-spin matrices, with three associated gauge boson fields, W^i_μ. The quantity g represents the coupling strength of this interaction, whilst the corresponding conserved charge is known as weak isospin, t_3. The Abelian $\mathrm{U(1)_Y}$ symmetry has a single generator, $\mathrm{T} = \mathrm{Y} = \frac{y}{2}\mathbb{1}$, with a single gauge boson field, B_μ. This interaction has a coupling strength, g', and conserves a quantity referred to as weak hypercharge, y. We extract the physical states of the gauge fields ($\mathrm{W}^\pm$ bosons, Z boson and the photon) according to the following

[5] This property is related to the β-function of QCD; a discussion of which is left to Ref. [54].

relations,

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \end{aligned} \tag{2.17}$$

where $\theta_W = \tan^{-1}(g'/g)$ is the weak mixing angle, and is measured directly in experiment. The gauge term of the electroweak (EW) Lagrangian,

$$\mathcal{L}_{\text{EW}} \supset \mathcal{L}_{\text{gauge}} = -\frac{1}{4}W^{i,\mu\nu}W^i_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}, \tag{2.18}$$

introduces self-interactions between the W_μ^i fields, which after rotating to the physical states are observed as the $\gamma\text{W}^+\text{W}^-$ and ZW^+W^- triple gauge interactions, and the $\text{W}^+\text{W}^-\text{W}^+\text{W}^-$, $\text{W}^+\text{W}^-\text{ZZ}$ and $\text{W}^+\text{W}^-\gamma\gamma$ quartic gauge interactions.

Turning our attention to the fermions, the diagram in Figure 2.1 tells us that both quarks and leptons interact via the weak interaction. In nature, we observe the weak interaction to violate parity conservation [57], or in other words, the weak interaction behaves differently under the spatial translation: $x \mapsto -x$. A fermion field, ψ, can be projected into its left-handed and right-handed components, ψ_L and ψ_R, using the projection operators, $\hat{P}_{L/R} = \frac{1}{2}(1 \mp \gamma^5)$, where $\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3$. We can then bake the effect of parity violation into EW theory by only permitting interactions between the W_μ^i gauge fields and left-handed fermions[6] (right-handed anti-fermions) [53]. In this way, the left-handed components of the fermion fields transform as doublets under SU(2)_L transformations ($t_3 = \pm\frac{1}{2}$), whilst the right-handed components transform as singlets ($t_3 = 0$),

$$\begin{aligned} \psi_L: &\quad L_{jL} = \begin{pmatrix} \nu_{jL} \\ \ell_{jL} \end{pmatrix}, \quad Q^a_{jL} = \begin{pmatrix} u^a_{jL} \\ d^a_{jL} \end{pmatrix}, \\ \psi_R: &\quad \ell_{jR}, \quad u^a_{jR}, \quad d^a_{jR}. \end{aligned} \tag{2.19}$$

Here, ℓ and ν represent the charged leptons and neutrinos, u and d represent the up-type and down-type quarks, and the flavour indices, j, and colour indices, a, are explicitly kept in the notation for completeness. It is important to note that there is no right-handed neutrino field in the standard EW theory, as they do not interact via either the weak or electromagnetic interactions.

The fermionic EW interactions are defined by the following gauge-invariant terms in the Lagrangian,

[6] Hence the use of the subscript, $_\text{L}$, in the SU(2) symmetry group definition.

$$\mathcal{L}_{\text{EW}} \supset \bar{\psi}_L i \not{D}^L \psi_L + \bar{\psi}_R i \not{D}^R \psi_R, \tag{2.20}$$

where

$$\begin{aligned} D_\mu^L &= \partial_\mu \mathbb{1}_2 + igW_\mu^i \frac{\sigma^i}{2} + ig' B_\mu \frac{y_L}{2} \mathbb{1}_2 \\ &= \begin{pmatrix} \partial_\mu & 0 \\ 0 & \partial_\mu \end{pmatrix} + \frac{ig}{2} \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} + \frac{ig' y_L}{2} \begin{pmatrix} B_\mu & 0 \\ 0 & B_\mu \end{pmatrix}, \\ D_\mu^R &= \partial_\mu + ig' B_\mu \frac{y_R}{2}. \end{aligned} \tag{2.21}$$

Considering the diagonal elements of D_μ^L together with D_μ^R, and applying the rotation to the physical basis (equation 2.17), we find a photon-fermion interaction term ($\bar{\psi}\psi A_\mu$) of the form,

$$\begin{aligned} \mathcal{L}_{\text{EW}} \supset & -\Big[\bar{\psi}_L \gamma^\mu (g \sin\theta_W \frac{\sigma^3}{2} + g' \cos\theta_W \frac{y_L}{2} \mathbb{1}_2) \psi_L \; + \; \bar{\psi}_R \gamma^\mu g' \cos\theta_W \frac{y_R}{2} \psi_R \Big] A_\mu, \\ = & -|e| \Big[\bar{\psi}_L \gamma^\mu (\frac{\sigma^3}{2} + \frac{y_L}{2} \mathbb{1}_2) \psi_L + \bar{\psi}_R \gamma^\mu \frac{y_R}{2} \psi_R \Big] A_\mu, \\ = & -|e| \mathcal{Q} \bar{\psi} \gamma^\mu \psi A_\mu, \end{aligned} \tag{2.22}$$

where the relations, $g \sin\theta_W = |e|$ and $g' \cos\theta_W = |e|$, are inferred by relating to the equivalent interaction term of QED in equation 2.5. Moreover, the observation that the electromagnetic interaction does not depend on the handedness (chirality state) of the fermion field tell us,

$$\mathcal{Q} = t_3 + \frac{y}{2}. \tag{2.23}$$

This is ultimately the defining relation of electroweak unification. At first glance, the electromagnetic and weak interactions appear to have completely different properties. The electromagnetic interaction is parity-conserving and is mediated by the massless photon over an effectively infinite range, whilst the weak interaction is parity-violating, and is mediated by the massive weak bosons, over a very short interaction range. Nevertheless, this section has proved that the two can be unified into a common electroweak theory, where equation 2.23 tells us how the charge of the electromagnetic interaction is related to the weak isospin and hypercharge. Using this relation, the full set of electroweak quantum numbers for the SM fields are derived to be the values shown in Table 2.1. Following the same procedure as shown for the $\bar{\psi}\psi A_\mu$ interaction, the charged current ($W^\pm$) and neutral current (Z) weak interactions are derived from the Lagrangian in equation 2.20, providing the full set of fermionic EW interactions displayed in Figure 2.4.

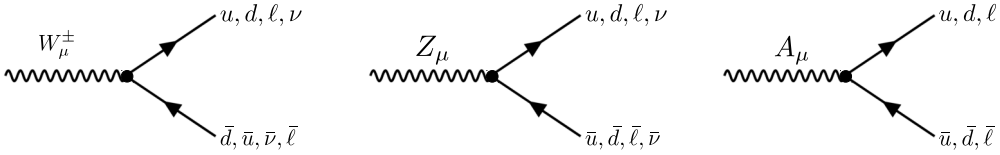


Figure 2.4: The electroweak interactions including fermions.

Table 2.1: The weak isospin (t_3), weak hypercharge (y) and electric charge ($\mathcal{Q}$) quantum numbers of the Standard Model fields. Here, the fermion flavour states are explicitly shown.

Particle field	t_3	y	$\mathcal{Q} = t_3 + \frac{y}{2}$
u^a_L, c^a_L, t^a_L	$+\frac{1}{2}$	$+\frac{1}{3}$	$+\frac{2}{3}$
d^a_L, s^a_L, b^a_L	$-\frac{1}{2}$	$+\frac{1}{3}$	$-\frac{1}{3}$
u^a_R, c^a_R, t^a_R	0	$+\frac{4}{3}$	$+\frac{2}{3}$
d^a_R, s^a_R, b^a_R	0	$-\frac{2}{3}$	$-\frac{1}{3}$
ν_{eL}, $\nu_{\mu L}$, $\nu_{\tau L}$	$+\frac{1}{2}$	-1	0
e_L, μ_L, τ_L	$-\frac{1}{2}$	-1	-1
e_R, μ_L, τ_L	0	-2	-1
$W^{\pm}_{\mu}$	± 1	0	± 1
Z_{μ}	0	0	0
A_{μ}	0	0	0
G^a_{μ}	0	0	0
H	$-\frac{1}{2}$	1	0

Spontaneous symmetry breaking

By simply applying the elegant principle of gauge invariance, we have constructed a Lagrangian density that describes all of the fundamental interactions between the fermions and the vector gauge bosons, as well as the gauge-boson self-interactions of the strong and weak forces. Nevertheless, there is a striking omission: a consistent description of the *particle masses*. As in QED, gauge-boson mass terms of the form $\propto m_{\mathrm{W}}^2 W^{i,\mu} W_\mu^i$ are forbidden in electroweak theory as they are not invariant under local $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ transformations. This contradicts the masses of the $\mathrm{W}^\pm$ and Z bosons observed in experiment. Moreover, mass terms for the fermion fields, $\propto m_\psi(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$ are also not gauge invariant due to the different transformation properties of the left-handed and right-handed fields. In SM theory, we introduce the Brout-Englert-Higgs (BEH) mechanism [8–13] to generate the particle masses by spontaneously breaking the symmetry of the electroweak interaction.

A spontaneous symmetry breaking (SSB) refers to the case where the underlying Lagrangian of the theory obeys certain symmetries, whereas the lowest energy (vacuum) state of the corresponding Hamiltonian does not [53]. The BEH mechanism explains what happens when the broken symmetry is a local gauge symmetry. In the SM, the BEH mechanism breaks the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry of the electroweak interaction into a $\mathrm{U(1)_{EM}}$ symmetry, as shown in equation 2.24:

$$\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y} \quad \overset{\mathrm{BEH}}{\mapsto} \quad \mathrm{SU(3)_C} \otimes \mathrm{U(1)_{EM}}. \tag{2.24}$$

Note, the $\mathrm{SU(3)_C}$ symmetry of the strong interaction remains unbroken and therefore the gluon fields, G_μ^a, are unaffected. The SSB is achieved via the introduction of two complex scalar fields[7], ϕ^+ and ϕ^0, which transforms as a doublet, H, under $\mathrm{SU(2)_L}$ transformations,

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{2.25}$$

The doublet, H, is referred to as the Higgs field. In total, this corresponds to adding four additional degrees of freedom into the Lagrangian via the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ invariant term for a complex scalar field,

$$\mathcal{L}_{\mathrm{SM}} \supset \mathcal{L}_{\mathrm{Higgs}} = (D^\mu H)^\dagger (D_\mu H) - \mu^2 H^\dagger H - \frac{1}{4}\lambda (H^\dagger H)^2 \tag{2.26}$$

where the covariant derivative D^μ acting on H is that shown in equation 2.16, such that the Higgs doublet possesses a weak hypercharge of $y_H = 1$. The latter two terms describe

[7] Because the BEH mechanism is required to generate the masses of the weak gauge bosons, one of the complex scalar fields must be neutral, ϕ^0, whilst the other must be charged, ϕ^+, such that $(\phi^+)^* = \phi^-$ [58].

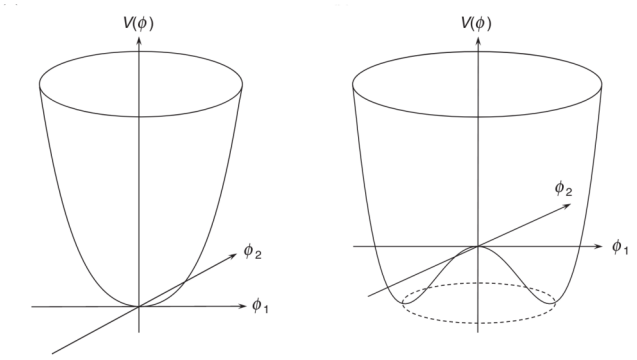


Figure 2.5: The potential, $V(\phi)$, for a single complex scalar field, $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$, uniquely defined by the parameters: μ^2 and λ. The plot on the left shows the scenario for $\mu^2 > 0$ and $\lambda > 0$, whilst the plot on the right shows the scenario for $\mu^2 < 0$ and $\lambda > 0$. It is the form on the right which is required for SSB via the BEH mechanism. Figure taken from Ref. [58].

the Higgs potential, $V(H)$,

$$V(H) = \mu^2 H^\dagger H + \frac{1}{4}\lambda(H^\dagger H)^2. \tag{2.27}$$

It is useful to begin by considering the possible forms of this potential for a *single* complex scalar field, $V(\phi)$, where $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$, as shown in Figure 2.5. In this example, the Lagrangian of the underlying theory is invariant under local U(1) gauge transformations. The parameter λ must be positive to ensure the minima of the potential are finite; however, there is no such requirement on μ^2. A value of $\mu^2 > 0$ (left) results in a symmetric potential with a minimum at $(\phi_1, \phi_2) = (0, 0)$. For $\mu^2 < 0$ (right) there is an infinite set of minima defined by $\phi_1^2 + \phi_2^2 = -\mu^2/\lambda = v^2$, as shown by the dashed circle in the plot. In other words, the lowest energy (vacuum) state does not occur at $(\phi_1, \phi_2) = (0, 0)$ but instead at one of the infinite points along $\phi_1^2 + \phi_2^2 = v^2$. It is the choice of the vacuum state which (spontaneously) breaks the local U(1) gauge symmetry of the Lagrangian [58].

We can extend this idea to the Higgs doublet with four degrees of freedom, where the non-zero vacuum states satisfy,

$$H^\dagger H = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}. \tag{2.28}$$

To achieve the SSB expressed in equation 2.24, a vacuum state is chosen to generate the masses of the $W^\pm$ and Z bosons, whilst leaving the symmetry of the electromagnetic (EM) interaction intact, and therefore the photon massless. The conventional choice [6] is,

$$\langle 0|H|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.29}$$

where $v = \sqrt{-\mu^2/\lambda}$ is the Higgs field vacuum expectation value. We then consider non-linear perturbations about this minimum of the form,

$$H = e^{iG^i(x)\frac{\sigma^i}{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.30}$$

where the Goldstone theorem [59] predicts the existence of three massless Goldstone boson fields, $G^i(x)$, for each broken generator of the symmetry group, and one massive scalar boson, $h(x)$, for the unbroken symmetry. The Goldstone bosons can be eliminated from the Lagrangian by making the appropriate local gauge transformation,

$$\mathrm{SU(2)_L} : H \mapsto H' = e^{i\theta^i(x)\frac{\sigma^i}{2}} H \tag{2.31}$$

and choosing $\theta^i(x) = \frac{-G(x)^i}{v}$.[8] This is known as the *unitary gauge*, where the Higgs doublet field takes the form,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.32}$$

Crucially, the three degrees of freedom corresponding to the massless Goldstone boson fields no longer appear in the Lagrangian. Instead, due to the covariant derivative in $\mathcal{L}_{\mathrm{Higgs}}$, they are absorbed as the longitudinal polarisation states of the weak vector bosons: $\mathrm{W}^{\pm}$ and Z. This is explicitly shown by inserting H in the unitary gauge into equation 2.26, where we find the following terms,

$$\mathcal{L}_{\mathrm{Higgs}} \supset \frac{1}{4} g^2 v^2 W^{+\mu} W^-_\mu + \frac{1}{8} v^2 (g^2 + g'^2) Z^\mu Z_\mu. \tag{2.33}$$

These are directly the weak boson mass terms that we set out to achieve. By introducing the BEH mechanism into the SM, we have successfully generated the masses of the weak bosons via the spontaneous symmetry breaking of the EW interaction. Using the known form of a mass term for a spin-1 gauge boson ($m^2_{\mathrm{W}} W_\mu W^\mu$), we read off the $\mathrm{W}^{\pm}$ and Z boson masses as,

$$m_{\mathrm{W}} = \frac{1}{2} g v = 80.4 \text{ GeV}, \qquad m_{\mathrm{Z}} = \frac{v}{2}\sqrt{g^2 + g'^2} = 91.2 \text{ GeV}, \tag{2.34}$$

which depend on the Higgs vacuum expectation value, $v = 246$ GeV, and the coupling strengths of the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ EW interaction. Remarkably, no equivalent mass term exists for the A^μ field: the photon remains massless. This is exactly what is required for a consistent description of the weak and electromagnetic interactions.

[8] It is important to keep in mind that we are free to make this choice, since the underlying Lagrangian is invariant under local gauge transformations.

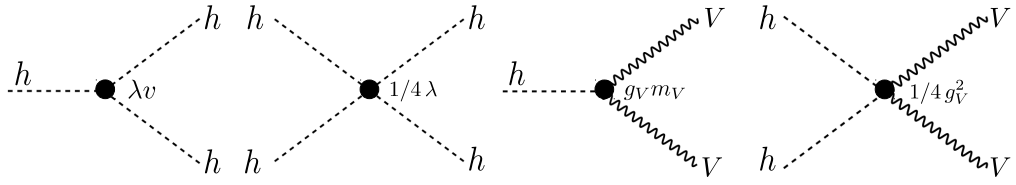


Figure 2.6: The Higgs boson self-interactions (left) and the Higgs boson interactions with the weak gauge bosons (right). The couplings $g_V = g_W$ and g_Z are equal to g and $g/\cos\theta_W$, respectively.

The Lagrangian also contains the following terms,

$$\mathcal{L}_{\text{Higgs}} \supset \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \lambda v^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4. \tag{2.35}$$

which describe an additional massive scalar boson field, h, with mass $m_{\text{H}} = \sqrt{2\lambda}v$.[9] The quantum excitation of this field, the Higgs boson, was discovered by the ATLAS and CMS Collaborations in 2012 [14–16], with a mass of around 125 GeV.

The final two terms in equation 2.35 describe the trilinear and quartic self-interactions of the Higgs boson, which are shown as Feynman diagrams on the left of Figure 2.6. Such interactions are extremely rare and have not yet been observed in high-energy physics experiments. In the SM, the value of λ is uniquely defined by the relation $m_{\text{H}} = \sqrt{2\lambda}v$, nevertheless, the presence of BSM physics can modify λ without changing the values of m_{H} and v. Measurements of the Higgs boson self-interactions are therefore in high demand as they provide an independent probe of λ; from which we can infer the shape of the Higgs potential, and subsequently the dynamics of EW symmetry breaking. Chapter 8 investigates an indirect method for probing λ at a future operation of the LHC machine.

Additionally, the $(D^\mu H)^\dagger(D_\mu H)$ term in $\mathcal{L}_{\text{Higgs}}$ gives rise to terms $\propto \text{VV}h$ and $\propto \text{VV}h^2$, where $\text{V} = \text{W}^\pm, \text{Z}$. These correspond to the triple and quartic couplings between the weak bosons and the Higgs boson, as shown on the right of Figure 2.6. The coupling strength at the $h\text{W}^+\text{W}^-$ vertex is equal to $\frac{1}{2}g^2 v \equiv g\, m_{\text{W}}$; the coupling of the Higgs boson to the W boson is proportional to the W boson mass. The equivalent is true for the $h\text{ZZ}$ vertex, where the coupling strength is $\propto m_{\text{Z}}$. No coupling exists between the h and the photon field, A_μ, nor does h couple to the gluon field, G^a_μ.

Yukawa interactions

The fermion masses are generated in the SM via the Yukawa interactions [58]. The following $\text{SU}(2)_{\text{L}} \otimes \text{U}(1)_{\text{Y}}$ gauge invariant terms are added to the SM Lagrangian,

[9] The (seemingly poor) choice in notation in using m_{H} instead of m_h, is simply made because m_{H} is more commonly used in the literature.

$$\mathcal{L}_{\text{SM}} \supset \mathcal{L}_{\text{Yukawa}} = -(\lambda_j \bar{L}_{jL} H \ell_{jR} + \lambda_j \bar{Q}^a_{jL} H d^a_{jR} + \lambda_j \bar{Q}^a_{jL} \tilde{H} u^a_{jR} + \text{h.c.}) \tag{2.36}$$

where λ_j are the Yukawa couplings for fermion flavour state, j, the notation $\tilde{H}^k = \epsilon_{kl}(H^l)^*$ is used, where ϵ_{jk} is totally antisymmetric in its indices with $\epsilon_{12} = +1$, and h.c. are the hermitian conjugates of the explicitly written terms. After SSB (substituting the unitary gauge H doublet into equation 2.36) we find terms of the form

$$\mathcal{L}_{\text{Yukawa}} \subset -\frac{\lambda_e}{\sqrt{2}} v(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{\lambda_e}{\sqrt{2}} h(\bar{e}_L e_R + \bar{e}_R e_L), \tag{2.37}$$

where for simplicity, only the terms relevant for the electron field are shown. The first term has exactly the form required for fermion masses, $m_j(\bar{\psi}_{jL}\psi_{jR} + \bar{\psi}_{jR}\psi_{jL})$, from which we can infer the relation, $\lambda_j = \sqrt{2}m_j/v$. The second term represents the couplings of the fermions fields to the Higgs boson: $h\bar{\psi}\psi$. These interactions have a coupling strength proportional to λ_j, and therefore the Higgs boson coupling to fermions is directly proportional to the mass of the fermion field. Whilst the BEH mechanism does not predict the values of λ_j, the masses of the fermions are measured experimentally, and hence provide a handle on the coupling properties of the Higgs boson.

2.1.4 The SM Lagrangian

Putting everything together, we arrive at an expression for the $\text{SU(3)}_\text{C} \otimes \text{SU(2)}_\text{L} \otimes \text{U(1)}_\text{Y}$ gauge invariant SM Lagrangian (before SSB) to be,

$$\begin{aligned}
\mathcal{L}_{\text{SM}} =& \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{Yukawa}} \\
=& -\frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu} - \frac{1}{4} W^i_{\mu\nu} W^{i,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\
&+ (D^\mu H)^\dagger (D_\mu H) - \mu^2 H^\dagger H - \frac{1}{4}\lambda (H^\dagger H)^2 \\
&+ i(\bar{L} \not{D} L + \bar{\ell} \not{D} \ell + \bar{Q} \not{D} Q + \bar{u} \not{D} u + \bar{d} \not{D} d) \\
&- (\lambda_\ell \bar{L} H \ell + \lambda_d \bar{Q} H d + \lambda_u \bar{Q} \tilde{H} u + \text{h.c.})
\end{aligned} \tag{2.38}$$

where the colour, flavour and chirality indices (L,R) have been dropped from the fermion fields in the notation, and $\lambda_{\ell,u,d}$ are now matrices in the flavour space. The covariant derivative, D_μ, which maintains the gauge invariance, is,

$$D_\mu = \partial_\mu + i g_s G^a_\mu \frac{\lambda^a}{2} + i g W^i_\mu \frac{\sigma^i}{2} + i g' B_\mu \text{Y} \tag{2.39}$$

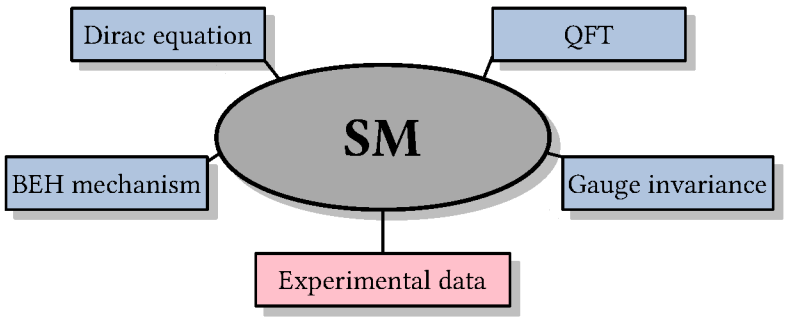


Figure 2.7: The pillars of SM theory.

Table 2.2: The 26 free parameters of the SM, which are input directly from experiment. References are provided for the parameters associated with topics that have not been addressed in the text.

Fermion masses	m_u, m_c, m_t, m_d, m_s, m_b, m_e, m_μ, m_τ, $m_{\nu 1}$, $m_{\nu 2}$, $m_{\nu 3}$
Gauge interaction coupling strengths	g, g', g_s
BEH mechanism	v, m_{H}
CKM and PMNS mixing angles [60–63]	λ, A, ρ, η, θ_{12}, θ_{13}, θ_{23}, δ
Strong CP phase [64]	$\theta^{\mathrm{CP}}(\approx 0)$

where the G^a_μ, W^i_μ and B_μ terms act on the fields with non-zero colour charge, non-zero weak isospin (t_3) and non-zero weak hypercharge (y), respectively. Incredibly, this Lagrangian density which barely spans four lines encompasses all known particle interactions via the electromagnetic, strong and weak forces, and with the inclusion of the BEH mechanism, explains how the weak gauge boson and fermion fields[10] acquire mass. The pillars of the theory that have lead to the construction of $\mathcal{L}_{\mathrm{SM}}$ are summarised in Figure 2.7.

One expects the ultimate theory of particle physics to be described by a simple equation with few free parameters [58]; plainly, the SM is *not* this theory. Despite being constructed from a number of profound theoretical ideas (blue boxes), it is put together in a somewhat ad-hoc manner. This becomes clear when considering the relatively large number of free parameters in the model (26 in total) which are independently tuned to match experimental observations (red box). Table 2.2 lists the free parameters of the SM.

Given this set of input parameters, all the SM particle masses and couplings are fully predicted. For example, the λ parameter which appears in the definition of the Higgs potential is equal to $m^2_{\mathrm{H}}/2v^2$. Nevertheless, the presence of new BSM fields can modify

[10] This does not include the generation of neutrino masses which requires a separate mechanism. The mechanism will not be discussed here since it is not relevant for this thesis.

the SM predictions, and thus leave an imprint on the masses and couplings of the SM fields. In turn, this affects not only the rate of interesting collision events at the LHC, but also their kinematic properties[11]. Altogether, this provides a concrete tool to search for new physics: by precisely measuring the masses and couplings of the SM fields and checking for consistency with the SM theory, we can indirectly infer the presence of new BSM states. In this thesis, we investigate the precision measurements of Higgs boson properties at the CMS experiment.

2.2 Higgs boson phenomenology

This section describes how the SM predictions manifest in terms of the Higgs boson phenomenology at a hadron collider. As mentioned at the beginning of this chapter, in high-energy particle collision experiments we do not observe the fundamental fields, but instead infer their interactions via production cross sections, σ, and decay rates, Γ. These observables are mapped back to the parameters of the model Lagrangian by computing the relevant matrix element, $\mathcal{M}_{i\to f}$: a quantity which encodes the probability of the transition from initial state i, to final state f, as a function of the particle couplings and masses. A useful introduction to the calculation of matrix elements (using the Feynman rules) is left to Ref. [58].

Higgs boson production

For proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, the four major Higgs boson production modes are (in order of decreasing cross section): gluon-gluon fusion (ggH), vector boson fusion (VBF), vector boson associated production (VH), and production in association with a top quark-antiquark pair (ttH).[12] All such production modes are now observed with significance $>5\sigma$ at the LHC [2, 65]. Additionally, there are a number of rarer production modes including single-top associated production (tH), gluon-initiated Z boson associated production (ggZH), and production in association with a bottom quark-antiquark pair (bbH). The leading-order[13] Feynman diagrams for all the aforementioned Higgs boson production modes are displayed in Figure 2.8, and the expected SM cross sections for each process are listed in Table 2.3. These values correspond to a nominal Higgs boson mass of $m_{\mathrm{H}} = 125.0$ GeV.

[11] For example, an additional high-mass state decaying to a Higgs boson will enhance the Higgs boson production rate at high momentum.

[12] From now on the distinction between fermions and anti-fermions is left out of the notation. For example, ttH strictly represents $\mathrm{t\bar{t}H}$ production. Additionally, the V in VBF and VH includes both W and Z bosons: $\mathrm{V} = \mathrm{W}^{\pm}, \mathrm{Z}$.

[13] See Section 3.5.

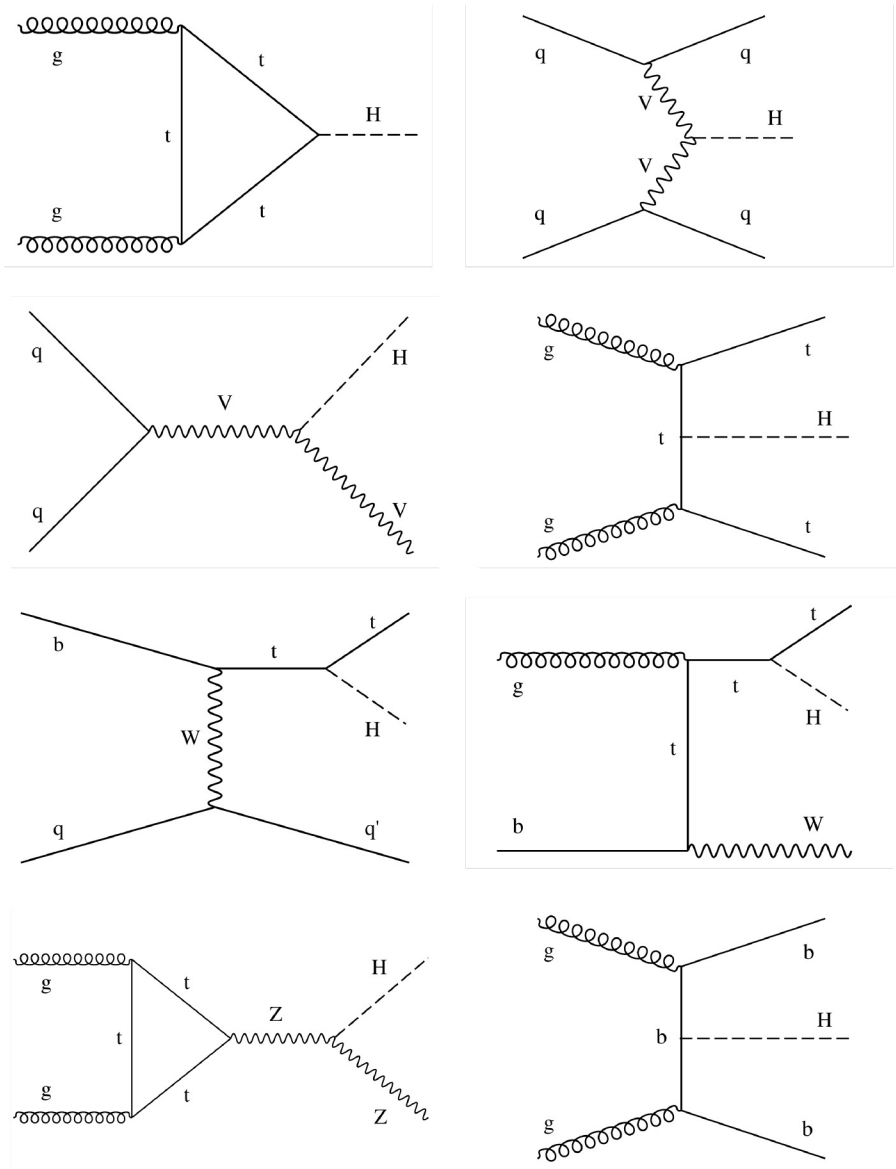


Figure 2.8: Example leading order Feynman diagrams for the major Higgs boson production modes at the LHC: ggH (top left), VBF (top right), VH (second left), ttH (second right), tHq (third left), tHW (third right), ggZH (bottom left) and bbH (bottom right).

Table 2.3: SM Higgs boson production cross sections at $\sqrt{s} = 13$ TeV, for $m_{\mathrm{H}} = 125.0$ GeV. The VH and tH production modes are separated into the contributions from the WH and ZH modes, and tHq and tHW modes, respectively. Values are taken from Ref. [42].

Production mode	ggH	VBF	WH	ZH	ggZH	ttH	tHq	tHW	bbH
Cross section [pb]	48.6	3.78	1.37	0.76	0.12	0.51	0.077	0.015	0.49

Table 2.4: Higgs boson decay branching fractions for $m_{\mathrm{H}} = 125.0$ GeV. Values are taken from Ref. [42].

Decay channel	bb	WW^*	gg	$\tau\tau$	cc	ZZ^*	$\gamma\gamma$	$\mu\mu$
Branching fraction [%]	58.2	21.4	8.2	6.3	2.8	2.6	0.23	0.022

The ggH production mode proceeds via an internal quark loop as the Higgs boson does not directly couple to the gluon (g). This makes ggH particularly sensitive to new physics where the quarks in the loop could in principle be replaced by a heavy BSM state that couples to the Higgs boson. In the SM, the cross section of ggH dominates by roughly an order of magnitude. Nevertheless, the other modes contain additional objects in their final states which can help identify them from other SM background processes. For example, VBF events are typically characterised by two additional jets, produced from the final-state quarks, with a large angular separation. Ultimately, the experimental sensitivity to a particular production-mode depends on both the cross section and the degree of background contamination mimicking the signal process.

Higgs boson decay

The Higgs boson is an unstable particle with a very short lifetime [24]. Consequently, it decays almost instantaneously and can only be inferred from the observation of its decay products. The principal Higgs boson decay channels are listed with their respective branching fractions, $\mathcal{B}^f$, for $m_{\mathrm{H}} = 125.0$ GeV in Table 2.4, where $\mathcal{B}^f = \Gamma^f/\Gamma^H$, and Γ^H is the total decay width of the Higgs boson.

The dominant Higgs boson decay channel is $\mathrm{H} \to \mathrm{bb}$, which occurs via the direct Yukawa coupling of the Higgs boson to the bottom quark. Nevertheless, the $\mathrm{H} \to \mathrm{bb}$ decay channel was only recently observed during Run 2 of the LHC [31, 34] due to the large hadronic background at the LHC. Other decay channels with relatively low branching fractions (e.g. $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$, $\mathrm{H} \to \gamma\gamma$) offer greater sensitivity to Higgs boson measurements. Of particular interest in this thesis is the $\mathrm{H} \to \gamma\gamma$ decay channel. This channel benefits from a clean final state topology of two well-reconstructed photons, that provide a narrow invariant mass ($m_{\gamma\gamma}$) peak about the Higgs boson mass, m_{H}. This property is used to effectively distinguish $\mathrm{H} \to \gamma\gamma$ events from the smoothly-falling background continuum. Moreover, it is one of the few channels that is sensitive to all of the major Higgs boson production modes.

Two of the leading-order Feynman diagrams for $\mathrm{H} \to \gamma\gamma$ in the SM are displayed in Figure 2.9. Like ggH production, the $\mathrm{H} \to \gamma\gamma$ decay proceeds via an internal loop as the Higgs boson does not couple directly to the photon field. This loop includes contributions

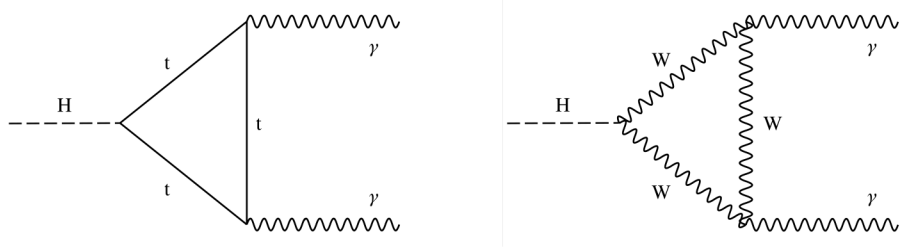


Figure 2.9: Leading order Feynman diagrams for the H $\rightarrow \gamma\gamma$ decay channel.

from both top quarks and $W^{\pm}$ bosons, meaning the H $\rightarrow \gamma\gamma$ decay rate is sensitive to both the top Yukawa coupling and the Higgs-W boson coupling, respectively.

Previous H $\rightarrow \gamma\gamma$ results at CMS

At this point, it would be cumbersome to review all results from the H $\rightarrow \gamma\gamma$ decay channel at CMS, let alone from all decay channels. In preparation for the following section, it is useful to introduce a single result from the H $\rightarrow \gamma\gamma$ decay channel, which uses proton-proton collision data collected during the 2016 period [66]. Figure 2.10 shows the cross section measurements of the dominant Higgs boson production modes, relative to their SM predictions. The VH production mode is separated according to the decay of the vector boson: leptonic ($Z \rightarrow \ell\ell/\nu\nu$ and $W \rightarrow \ell\nu$) and hadronic ($W/Z \rightarrow qq'/qq$), and the leptonic channel is further divided into the WH and ZH production contributions. With the 2016 data only, the ggH cross section is measured to within an uncertainty of around 20%. Other production modes are measured less well, with uncertainties ranging from 50-60% for VBF production to 250% for VH hadronic. Looking forward, the inclusion of more data will help reduce the statistical uncertainty in these measurements, and will enable new, more granular measurements to be made. One framework for performing such granular measurements is the *simplified template cross section* (STXS) framework.

2.3 Simplified template cross sections

The STXS framework [42] is a remarkably simple concept: as the available data increases, we divide the Higgs boson production phase-space into increasingly granular kinematic regions. These regions, or so-called *bins*, are split primarily by the Higgs boson production mode and subsequently by the kinematic quantities of the event. The framework provides a natural progression to inclusive production-mode measurements (see Figure 2.10), such that by measuring the cross section in the kinematic regions, we build up a more granular description of Higgs boson production, and gain an understanding of the event kinematics. Ultimately, this enhances the sensitivity to potential new-physics contributions that

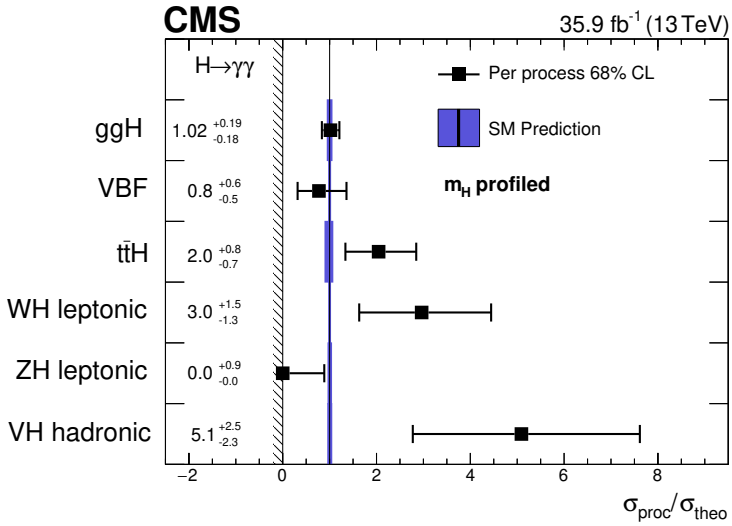


Figure 2.10: Measurements of the dominant Higgs boson production mode cross sections, relative to their SM predictions, in the H $\rightarrow \gamma\gamma$ decay channel. The results use data collected by the CMS experiment during the 2016 data taking period. The black points represent the best-fit values and 68% confidence intervals for the measured parameters of interest. The blue boxes demonstate the theoretical uncertainties in the SM predictions.

appear in specific regions of the Higgs boson production phase-space.

This coherent approach to precision Higgs boson measurements has been adopted by the ATLAS and CMS Collaborations for a number of years. The evolution of the framework with increasing statistics is defined in so-called *stages* [42, 67], which refer to different kinematic binning schemes of varying granularity. The H $\rightarrow \gamma\gamma$ analysis documented in Chapters 4–6 targets the most-granular stage 1.2 binning definition, which is shown schematically in Figure 2.11. The effective field theory interpretation of STXS measurements in Chapter 7 combines cross section measurements from different Higgs boson decay channels at stage 0, stage 1.0 and stage 1.1. All stages are displayed schematically in Appendix A.

There are many advantages to cross section measurements in the STXS framework. Firstly, splitting the phase space into different kinematic bins systematically reduces the theory dependence of the measurements, in the sense that we no longer rely on the SM to predict the relative compositions of two different physics processes (STXS bins). Reducing the theory dependence in this fashion makes the measurements easier to re-interpret, and also helps to preserve the usefulness of the measurements as they are less affected by future improvements to the theoretical predictions. Moreover, the splitting helps to isolate specific kinematic regions of phase space which are likely to be affected by BSM physics e.g. ggH production with high p_T^H.

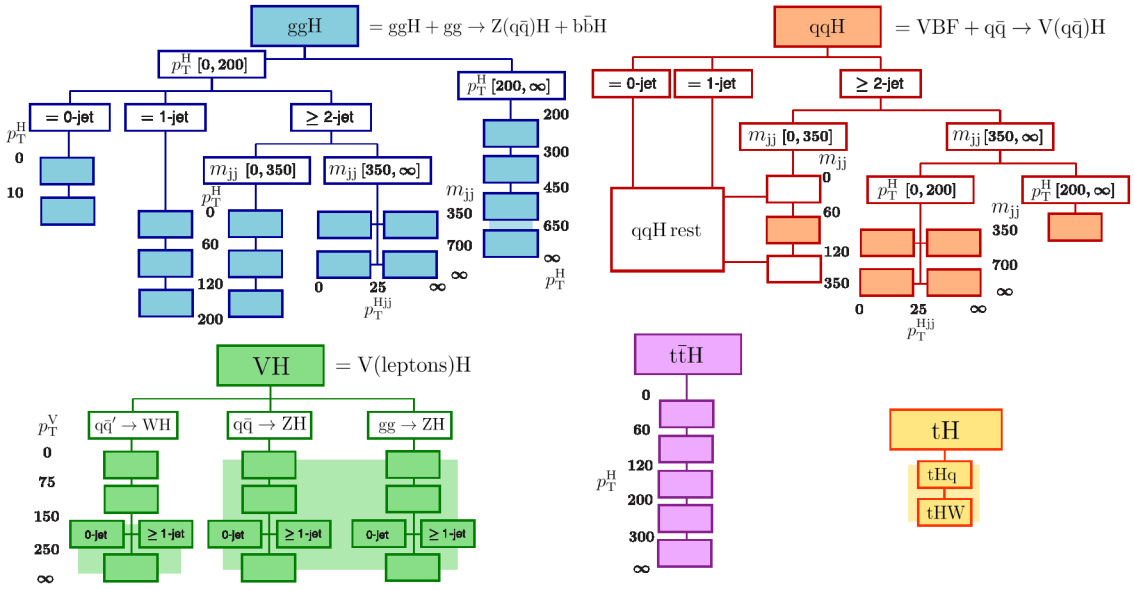


Figure 2.11: Schematic of the full STXS stage 1.2 binning scheme, adapted from Ref. [42]. This is defined for events with $|y_H|< 2.5$. The solid boxes represent each STXS stage 1.2 bin. The units of p_T^H, $m_{\gamma\gamma}$, p_T^{Hjj}, and p_T^V are in GeV. The shaded regions indicate the STXS bins which are divided at stage 1.2, but are not measured independently in the H $\rightarrow \gamma\gamma$ analysis described in Chapters 4–6.

Measurements in the STXS framework differ from standard differential cross sections measurements as they do not include a fiducial-region[14] definition of the final state particles in the event, including the Higgs boson decay products. This ultimately permits the use of sophisticated analysis techniques to optimise the sensitivity to the cross sections, such as the application of machine learning algorithms for event classification (see Appendix B). Additionally, by being agnostic to the Higgs boson decay, the STXS framework enables the smooth combination of measurements across different Higgs boson decay channels.

Whilst the benefits listed above are numerous, it is important to keep in mind that STXS measurements rely on SM simulation to model the experimental acceptance of events from a given STXS bin. This assumption may break down in the presence of BSM physics, which can modify the event kinematics and subsequently the within-bin experimental acceptance. In particular, new physics affecting the final-state kinematics of the Higgs boson decay goes unaccounted for in the STXS framework as there is no fiducial selection on the decay products. Consequently, BSM re-interpretations of STXS measurements must be approached with care.

The remainder of the section provides a description of the stage 1.2 binning scheme

[14] In this context, a *fiducial region* refers to a region of the final-state phase-space defined at truth-level (see Section 3.5), which corresponds closely to what can be measured experimentally in the detector.

used in the H $\rightarrow \gamma\gamma$ analysis of Chapters 4–6 (Figure 2.11). Throughout, the units of GeV are removed in the STXS bin-naming for brevity. All events require the absolute value of the Higgs boson rapidity, $|y_H|$, to be less than 2.5, as events with values above this threshold are typically outside of the experimental acceptance. The ggH production mode (blue) is split into bins according to the Higgs boson transverse momentum (p_T^H) and the number of jets[15]. The boosted kinematic region with $p_T^H > 200$ GeV is particularly sensitive to BSM physics appearing in the ggH loop. This region is further split according to additional p_T^H boundaries, which are measured for the first time in this thesis.

Additionally, the ggH binning scheme defines a VBF-like region with high dijet invariant mass (m_{jj}). This VBF-like region is split into four bins according to m_{jj} and the transverse momentum of the Higgs boson-plus-dijet system (p_T^{Hjj}). For the purpose of the H $\rightarrow \gamma\gamma$ analysis, events originating from bbH production and ggZH production in which the associated Z boson decays hadronically, are grouped with ggH events. This choice is made since the current available data set is not sensitive to the independent measurements of these rarer production modes. The SM predicted cross section breakdown for the ggH STXS bins is given in Table 2.5. These values are calculated by classifying events in SM simulation, as described in Section 4.2.2. The final column in the table lists the product of the cross section times H $\rightarrow \gamma\gamma$ branching fraction, $\sigma_{\text{SM}}\mathcal{B}$, for each bin. Ultimately, these are the observables that the H $\rightarrow \gamma\gamma$ analysis aims to measure, and by comparing to the SM predictions it is possible to constrain (or more optimistically infer) the presence of new physics.

The electroweak qqH production scheme (orange) considers both VBF production and VH production in which the vector boson decays hadronically. This reflects the fact that VBF and VH hadronic production are the t+u-channel and s-channel diagrams, respectively, of the same physics process: qq→Hqq, and therefore cannot be distinguished at higher orders. The kinematic bins are defined according to the number of jets, p_T^H, m_{jj}, and p_T^{Hjj}, in the attempt to isolate different topologies of qqH events. Firstly, the bin with the dijet invariant mass window $60 < m_{jj} < 120$ GeV specifically targets VH hadronic production. Akin to the ggH scheme, events with a VBF-like topology are defined by the region with $m_{jj} > 350$ GeV, which is further split into four bins according to boundaries in m_{jj} and p_T^{Hjj}. Finally, the BSM-sensitive region with a boosted Higgs boson is defined by a single bin with $p_T^H > 200$ GeV. The four STXS bins which define the "qqH rest" region (see Figure 2.11) are not explicitly probed in the H $\rightarrow \gamma\gamma$ analysis. Table 2.6 provides the SM predicted cross section breakdown for the qqH STXS bins.

[15] In the STXS event classification, jets are defined with the FASTJET package [68] using the anti-k_T algorithm with a distance parameter of 0.4 [69]. This is applied to the particle-level event constituents after parton showering and hadronisation. All jets are required to have a transverse momentum, $p_T^j > 30$ GeV. More detail concerning the definition of jets in simulation is provided in Section 3.5.

Table 2.5: Definition of the ggH STXS bins. The product of the cross section times H $\to \gamma\gamma$ branching fraction, $\sigma_{\mathrm{SM}}\mathcal{B}$, evaluated at $\sqrt{s} = 13$ TeV and $m_{\mathrm{H}} = 125$ GeV, is given for each bin in the final column. Additionally, the fraction of the total production mode cross section from each STXS bin is shown. Events originating from ggZH production, in which the Z decays hadronically, are grouped with ggH in the STXS measurements, and are shown as a separate column in the table. The bbH production mode, whose $\sigma_{\mathrm{SM}}\mathcal{B} = 1.054$ fb, is grouped together with the ggH 0J high p_T^H bin. Unless stated otherwise, the STXS bins are defined for $|y_H| < 2.5$. Events with $|y_H| > 2.5$ are mostly outside of experimental acceptance, and therefore make a negligible contribution to the H $\to \gamma\gamma$ analysis.

STXS bin	Definition units of p_T^H, m_{jj} and p_T^{Hjj} in GeV	Fraction of cross section ggH	 gg $\to$ Z(q$\bar{\mathrm{q}}$)H	$\sigma_{\mathrm{SM}}\mathcal{B}$ (fb)
ggH forward	$\lvert Y_H \rvert > 2.5$	8.09%	2.73%	8.93
ggH 0J low p_T^H	Exactly 0 jets, $p_T^H < 10$	13.87%	0.01%	15.30
ggH 0J high p_T^H	Exactly 0 jets, $10 < p_T^H < 200$	39.40%	0.29%	43.45
ggH 1J low p_T^H	Exactly 1 jet, $p_T^H < 60$	14.77%	2.00%	16.29
ggH 1J med p_T^H	Exactly 1 jet, $60 < p_T^H < 120$	10.23%	5.34%	11.29
ggH 1J high p_T^H	Exactly 1 jet, $120 < p_T^H < 200$	1.82%	3.53%	2.01
ggH $\geq$2J low p_T^H	At least 2 jets, $p_T^H < 60$, $m_{jj} < 350$	2.56%	5.74%	2.83
ggH $\geq$2J med p_T^H	At least 2 jets, $60 < p_T^H < 120$, $m_{jj} < 350$	4.10%	19.63%	4.56
ggH $\geq$2J high p_T^H	At least 2 jets, $120 < p_T^H < 200$, $m_{jj} < 350$	1.88%	29.55%	2.13
ggH BSM $200 < p_T^H < 300$	No jet requirements, $200 < p_T^H < 300$	0.98%	13.93%	1.11
ggH BSM $300 < p_T^H < 450$	No jet requirements, $300 < p_T^H < 450$	0.25%	3.86%	0.28
ggH BSM $450 < p_T^H < 650$	No jet requirements, $450 < p_T^H < 650$	0.03%	0.77%	0.03
ggH BSM $p_T^H > 650$	No jet requirements, $p_T^H > 650$	0.01%	0.20%	0.01
ggH VBF-like low m_{jj} low p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $350 < m_{jj} < 700$, $p_T^{Hjj} < 25$	0.63%	1.14%	0.70
ggH VBF-like low m_{jj} high p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $350 < m_{jj} < 700$, $p_T^{Hjj} > 25$	0.77%	8.06%	0.86
ggH VBF-like high m_{jj} low p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $m_{jj} > 700$, $p_T^{Hjj} < 25$	0.28%	0.36%	0.31
ggH VBF-like high m_{jj} high p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $m_{jj} > 700$, $p_T^{Hjj} > 25$	0.32%	2.85%	0.36

Events produced via the VH and ggZH production modes, in which the vector boson decays leptonically are categorised according to the VH leptonic binning scheme (green). Three equivalent regions are defined for WH, ZH and ggZH production, with kinematic boundaries in the transverse momentum of the vector boson (p_T^V) and the number of jets. The SM predicted cross section breakdown for the VH leptonic STXS bins is listed in Table 2.7.

The ttH production mode (pink) is split according to four boundaries in p_T^H. This splitting was first introduced at stage 1.2, and consequently the H $\rightarrow \gamma\gamma$ analysis documented here is the first to measure ttH production in different kinematic regions. Finally, there is an additional tH production STXS bin (yellow), which includes contributions from both the tHq and tHW production modes. This thesis also includes the first explicit measurement of tH production cross section in the H $\rightarrow \gamma\gamma$ decay channel. Table 2.8 gives the SM predicted cross section breakdown for the top-associated STXS bins.

Table 2.6: Definition of the qqH STXS bins. The product of the cross section times H $\rightarrow \gamma\gamma$ branching fraction, $\sigma_{\mathrm{SM}}\mathcal{B}$, evaluated at $\sqrt{s} = 13$ TeV and $m_{\mathrm{H}} = 125$ GeV, is given for each bin in the final column. Additionally, the fraction of the total production mode cross section from each STXS bin is shown. Unless stated otherwise, the STXS bins are defined for $|y_H| < 2.5$. Events with $|y_H| > 2.5$ are mostly outside of experimental acceptance, and therefore make a negligible contribution to the H $\rightarrow \gamma\gamma$ analysis.

STXS bin	Definition units of p_T^H, m_{jj} and p_T^{Hjj} in GeV	Fraction of cross section VBF	$q\bar{q}' \rightarrow W(q\bar{q}')H$	$q\bar{q} \rightarrow Z(q\bar{q})H$	$\sigma_{\mathrm{SM}}\mathcal{B}$ (fb)
qqH forward	$\lvert Y_H\rvert > 2.5$	6.69%	12.57%	9.84%	0.98
qqH 0J	Exactly 0 jets	6.95%	5.70%	3.73%	0.77
qqH 1J	Exactly 1 jet	32.83%	31.13%	25.03%	3.82
qqH $m_{jj} < 60$	At least 2 jets, $m_{jj} < 60$	1.36%	3.58%	2.72%	0.23
qqH VH-like	At least 2 jets, $60 < m_{jj} < 120$	2.40%	29.43%	28.94%	1.23
qqH $120 < m_{jj} < 350$	At least 2 jets, $120 < m_{jj} < 350$	12.34%	13.92%	12.59%	1.53
qqH VBF-like low m_{jj} low p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $350 < m_{jj} < 700$, $p_T^{Hjj} < 25$	10.26%	0.44%	0.35%	0.90
qqH VBF-like low m_{jj} high p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $350 < m_{jj} < 700$, $p_T^{Hjj} > 25$	3.85%	1.86%	1.74%	0.39
qqH VBF-like high m_{jj} low p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $m_{jj} > 700$, $p_T^{Hjj} < 25$	15.09%	0.09%	0.08%	1.30
qqH VBF-like high m_{jj} high p_T^{Hjj}	At least 2 jets, $p_T^H < 200$, $m_{jj} > 700$, $p_T^{Hjj} > 25$	4.25%	0.40%	0.39%	0.38
qqH BSM	At least 2 jets, $m_{jj} > 350$, $p_T^H > 200$	3.98%	0.88%	0.71%	0.37

Table 2.7: Definition of the VH leptonic STXS bins. The product of the cross section times $H \to \gamma\gamma$ branching fraction, $\sigma_{SM}\mathcal{B}$, evaluated at $\sqrt{s} = 13$ TeV and $m_H = 125$ GeV, is given for each bin in the final column. Additionally, the fraction of the total production mode cross section from each STXS bin is shown. Unless stated otherwise, the STXS bins are defined for $|y_H| < 2.5$. Events with $|y_H| > 2.5$ are mostly outside of experimental acceptance, and therefore make a negligible contribution to the $H \to \gamma\gamma$ analysis.

STXS bin	Definition units of p_T^V in GeV	Fraction of cross section $q\bar{q}' \to WH$	$q\bar{q}' \to ZH$	$gg \to ZH$	$\sigma_{SM}\mathcal{B}$ (fb)
WH lep forward	$\|Y_H\| > 2.5$	12.13%	-	-	0.123
ZH lep forward		-	11.21%	-	0.058
ggZH lep forward		-	-	2.71%	0.002
WH lep $p_T^V < 75$	No jet requirements, $p_T^V < 75$	46.55%	-	-	0.473
WH lep $75 < p_T^V < 150$	No jet requirements, $75 < p_T^V < 150$	29.30%	-	-	0.298
WH lep 0J $150 < p_T^V < 250$	Exactly 0 jets, $150 < p_T^V < 250$	5.10%	-	-	0.052
WH lep $\geq$1J $150 < p_T^V < 250$	At least 1 jet, $150 < p_T^V < 250$	3.97%	-	-	0.040
WH lep $p_T^V > 250$	No jet requirements, $p_T^V > 250$	2.95%	-	-	0.030
ZH lep $p_T^V < 75$	No jet requirements, $p_T^V < 75$	-	45.65%	-	0.237
ZH lep $75 < p_T^V < 150$	No jet requirements, $75 < p_T^V < 150$	-	30.70%	-	0.160
ZH lep 0J $150 < p_T^V < 250$	Exactly 0 jets, $150 < p_T^V < 250$	-	5.16%	-	0.027
ZH lep $\geq$1J $150 < p_T^V < 250$	At least 1 jet, $150 < p_T^V < 250$	-	4.27%	-	0.022
ZH lep $p_T^V > 250$	No jet requirements, $p_T^V > 250$	-	3.01%	-	0.016
ggZH lep $p_T^V < 75$	No jet requirements, $p_T^V < 75$	-	-	15.96%	0.013
ggZH lep $75 < p_T^V < 150$	No jet requirements, $75 < p_T^V < 150$	-	-	43.32%	0.036
ggZH lep 0J $150 < p_T^V < 250$	Exactly 0 jets, $150 < p_T^V < 250$	-	-	9.08%	0.008
ggZH lep $\geq$1J $150 < p_T^V < 250$	At least 1 jet, $150 < p_T^V < 250$	-	-	20.49%	0.017
ggZH lep $p_T^V > 250$	No jet requirements, $p_T^V > 250$	-	-	8.45%	0.007

Table 2.8: Definition of the top-associated STXS bins. The product of the cross section times $H \to \gamma\gamma$ branching fraction, $\sigma_{SM}\mathcal{B}$, evaluated at $\sqrt{s} = 13$ TeV and $m_H = 125$ GeV, is given for each bin in the final column. Additionally, the fraction of the total production mode cross section from each STXS bin is shown. Unless stated otherwise, the STXS bins are defined for $|y_H| < 2.5$. Events with $|y_H| > 2.5$ are mostly outside of experimental acceptance, and therefore make a negligible contribution to the $H \to \gamma\gamma$ analysis.

STXS bin	Definition units of p_T^H in GeV	Fraction of cross section $t\bar{t}H$	tHq	tHW	$\sigma_{SM}\mathcal{B}$ (fb)
$t\bar{t}H$ forward	$\|Y_H\| > 2.5$	1.35%	-	-	0.016
tH forward		-	2.79%	1.06%	0.005
$t\bar{t}H$ $p_T^H < 60$	No jet requirements, $p_T^H < 60$	22.42%	-	-	0.259
$t\bar{t}H$ $60 < p_T^H < 120$	No jet requirements, $60 < p_T^H < 120$	34.61%	-	-	0.400
$t\bar{t}H$ $120 < p_T^H < 200$	No jet requirements, $120 < p_T^H < 200$	25.60%	-	-	0.296
$t\bar{t}H$ $200 < p_T^H < 300$	No jet requirements, $200 < p_T^H < 300$	10.72%	-	-	0.124
$t\bar{t}H$ $p_T^H > 300$	No jet requirements, $p_T^H > 300$	5.31%	-	-	0.061
tH	No additional requirements	-	97.21%	98.94%	0.204

2.4 Effective field theory

The chapter concludes with an introduction to effective field theory (EFT), which is later applied in a BSM interpretation of STXS measurements in Chapter 7. EFT is by no means a new concept in particle physics. In 1933, Enrico Fermi formulated a low energy description of nuclear β-decay as a four-fermion contact interaction (an interaction which occurs at a single point in space-time) between a neutron, electron, (anti)-neutrino, and a proton [70]. This is expressed by the matrix element[16],

$$\mathcal{M}_{i\to f} = \frac{1}{\sqrt{2}} G_F \eta_{\mu\nu} (\bar{\psi}_3 \gamma^\mu (1-\gamma^5)\psi_1)(\bar{\psi}_4 \gamma^\nu (1-\gamma^5)\psi_2), \tag{2.40}$$

where G_F governs the strength of the contact interaction, and is referred to as the Fermi constant. It is now known that this process is described by the weak interaction, and occurs via the exchange of a W boson. This has the corresponding matrix element,

$$\mathcal{M}_{i\to f} = -\Big(\frac{1}{2\sqrt{2}} g \bar{\psi}_3 \gamma^\mu (1-\gamma^5)\psi_1\Big) \cdot \Big(\frac{\eta_{\mu\nu} - q_\mu q_\nu / m_{\mathrm{W}}^2}{q^2 - m_{\mathrm{W}}^2}\Big) \cdot \Big(\frac{1}{2\sqrt{2}} g \bar{\psi}_4 \gamma^\nu (1-\gamma^5)\psi_2\Big). \tag{2.41}$$

In the low energy ($q^2 \ll m_{\mathrm{W}}^2$) limit, this simply reduces to the matrix element from Fermi's theory (equation 2.40). In other words, Fermi's theory provides a low-energy *effective field theory* of the weak interaction. Moreover, we can *match* the contact interaction strength, G_F, to the parameters of the complete high-energy theory:

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8 m_{\mathrm{W}}^2}, \tag{2.42}$$

where g is the weak interaction coupling strength, and m_{W} is the mass of the W boson. Hence, by measuring G_F experimentally, it is possible to infer some knowledge concerning the high-energy theory (in this case being the weak interaction). The application of EFT in the weak interaction is shown diagrammatically in Figure 2.12.

EFT can be used to extend the SM Lagrangian to provide a (almost) model independent approach to search for BSM physics [71]. First, we assume that any new BSM states reside at an energy scale, Λ, far beyond the electroweak scale: $\Lambda \gg v$. This allows us to replace the non-local interactions involving the exchange of new particles, by contact interactions between the SM fields. In essence, the EFT describes the effect of the UV (ultraviolet) short-distance BSM physics on the IR (infrared) long-distance SM physics, without the need to construct a fully consistent BSM theory. The validity of this approach is restricted to energies below the new-physics mass scale, Λ.

[16] This expression includes a modification of the original Fermi theory to incorporate the effects of parity violation, observed experimentally by Wu *et. al.* in 1957 [57].

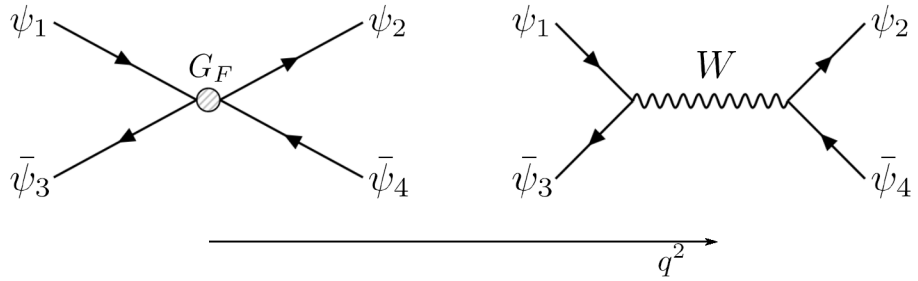


Figure 2.12: Fermi's theory, which models the weak interaction at low energies ($q^2 \ll m_{\mathrm{W}}^2$) as a contact interaction between four fermions (left). At energies approaching m_{W}, the exchange of the W boson is required to correctly describe the interaction (right).

The contact interactions are expressed mathematically as higher dimensional operators of the SM fields. The action of a Lagrangian, $S = \int \mathrm{d}^4x\,\mathcal{L}$, is a dimensionless quantity. Consequently, the operators in the SM Lagrangian, $\mathcal{L}_{\mathrm{SM}}$, must be of (energy) dimension-4,[17] and from this we infer the dimensions of the SM fields to be the values listed in Table 2.9. The EFT Lagrangian, $\mathcal{L}_{\mathrm{EFT}}$, extends $\mathcal{L}_{\mathrm{SM}}$ with higher-dimensional operators according to,

$$\mathcal{L}_{\mathrm{EFT}} = \mathcal{L}_{\mathrm{SM}}^{(4)} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} + \mathcal{L}^{(7)} + \mathcal{L}^{(8)} + \ldots, \tag{2.43}$$

where the higher-order terms are of the form,

$$\mathcal{L}^d = \sum_{p=1}^{N^{(d)}} \frac{w_p^{(d)}}{\Lambda^{d-4}} \mathcal{O}_p^{(d)}. \tag{2.44}$$

Here, $\mathcal{O}_p^{(d)}$ is a dimension d operator constructed from the SM fields (Table 2.9), and $w_p^{(d)}$ is the corresponding *Wilson coefficient*. These Wilson coefficients embed the influence of the UV BSM physics on the IR operator, $\mathcal{O}_p^{(d)}$, such that a deviation from zero in $w_p^{(d)}$ implies the existence of new physics. In the context of Fermi's theory, the operator $\mathcal{O} = \eta_{\mu\nu}(\bar{\psi}_3\gamma^\mu(1-\gamma^5)\psi_1)(\bar{\psi}_4\gamma^\nu(1-\gamma^5)\psi_2)$ describes the four-fermion contact interaction, whilst G_F is related to the corresponding Wilson coefficient, encoding the strength of the (high-energy) weak interaction. In summary, by measuring the set of $w_p^{(d)}$ coefficients in experiment, we can infer not only the strength of new physics interactions, but also the relevant processes which are affected.

In order to keep the action dimensionless, the higher-order contributions enter with an energy scale suppression of $\Lambda^{-(d-4)}$. For example, the dimension-6 terms and dimension-8 terms are suppressed by factors Λ^{-2} and Λ^{-4}, respectively. As a consequence, lower-dimensional operators have a larger impact on physics observables. The total number of operators at dimension d is expressed as $N^{(d)}$. At dimension-6 there are 2499 CP-

[17] $[\mathrm{d}^4x] = 4$ in terms of x, and $[x] = [E]^{-1}$, where E is energy.

Table 2.9: The dimension of the SM Lagrangian terms.

Term	Dimension
Field strength tensors: $G^a_{\mu\nu}$, $W^i_{\mu\nu}$, $B_{\mu\nu}$	2
Derivative: ∂_μ	1
Gauge fields: G^a_μ, W^i_μ, B_μ	1
Covariant derivative: D_μ	1
Higgs field: H	1
Higgs vacuum expectation value: v	1
Fermion fields: ψ	3/2
All couplings: λ, g, λ_i etc	0

even operators, which are typically reduced to a manageable number by assuming certain flavour symmetries in the model (with minimal flavour violation, the number of independent operators reduces to 59 [72]). This feature of EFT becomes increasingly problematic at higher dimensions, such that at dimension-8 there are 44,807 independent operators that enter the EFT Lagrangian [73]. Ultimately, when measuring EFT parameters in experiment, we typically consider only a subset of (the most relevant) operators.

Dimension-5 operators violate lepton number conservation [74], whilst all other odd-dimension operators violate the conservation of baryon-number minus lepton-number $(B - L)$ [75]. So far, all LHC measurements suggest these approximate symmetries are conserved, and therefore odd-dimension operators are not considered in this thesis. Moreover, due to the energy-scale suppression, all terms of dimension-8 and higher are ignored, leaving only dimension-6 contributions to the Lagrangian,

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}^{(4)}_{\text{SM}} + \mathcal{L}^{(6)}. \tag{2.45}$$

Let's introduce a concrete example. In the SM, the Higgs field does not couple to the gluon field. Moving to dimension-6, it is possible to construct a gauge-invariant operator of the form

$$\mathcal{O}^{(6)}_G = |H|^2 G^a_{\mu\nu} G^{a,\mu\nu}. \tag{2.46}$$

which describes the effective contact interaction between the Higgs and gluon fields (Figure 2.13, right). The corresponding Wilson coefficient, $w^{(6)}_G$, encodes the contribution from new physics in this interaction, such that a value different from zero would imply the existence of some new BSM state that couples to the Higgs and the gluon fields (Figure 2.13, left). Approaching the problem in this way means we are agnostic to the specifics of the UV-complete BSM theory, or in other words, it enables a model-independent method to search for BSM physics. If required, the measurements of the Wilson coefficients can be systematically *matched* to UV-complete BSM theories to place constraints on the parti-

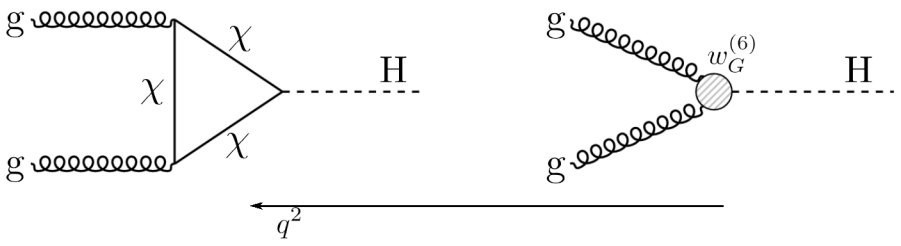


Figure 2.13: Modelling a BSM physics state, χ, at low energy ($q^2 \ll \Lambda^2$) with the dimension-6 contact interaction between the Higgs and gluon fields.

cle masses and couplings of the theory. Further detail on this matching procedure, with specific examples, is provided in Ref. [76].

The modified dynamics of the SM fields are derived by the applying the Euler-Lagrange equations to $\mathcal{L}_{\text{EFT}}$. This provides a new set of Feynman rules, which subsequently describe how the additional contact interactions affect the experimentally measured observables: σ and Γ. Crucially, the effects are not limited to the inclusive rates, but also modify the kinematic properties of the events. This makes STXS (or any differential) measurements an excellent candidate for probing EFT effects in the Higgs sector, as the additional kinematic information they provide can be used to further constrain the EFT Wilson coefficients. An EFT interpretation of STXS measurements is provided in Chapter 7.

2.4.1 SMEFT and the choice of basis

The EFT expansion of equation 2.45 can be defined under different assumptions. A Standard Model EFT (SMEFT) only considers higher-dimensional operators which obey the gauge symmetries of the SM: $\text{SU(3)}_\text{C} \otimes \text{SU(2)}_\text{L} \otimes \text{U(1)}_\text{Y}$ [71]. In this manner, the operators are constructed from the field definitions before SSB (the gauge eigenbasis): W_μ^i, B_μ and H, such that the Higgs field enters as the SU(2)_L doublet. Whilst this approach does not offer a perfect mapping to the observed physical states: $\text{W}^\pm$, Z, γ and h, it provides a fully consistent approach that can be combined with other high-energy physics results, such as top quark or electroweak measurements. The EFT used in this thesis is SMEFT.

Other options do exist, such as the Higgs EFT (HEFT) [42, 71], where the expansion is performed in terms of the Higgs boson singlet, h. As a result, HEFT allows for deviations from the SM SU(2)_L doublet structure of the scalar sector. This approach benefits from being defined in the terms of the physical states (mass eigenbasis) and therefore provides a more simple mapping to the experimentally measured observables. The downside of HEFT lies in the complicated matching to explicit UV-complete theories.

In addition, when performing an EFT interpretation there exists a choice in the expan-

sion *basis*, where a basis is defined as a complete and non-redundant set of EFT operators: $\mathcal{O}_p$ [42]. When listing EFT operators, we typically find redundant combinations of the fields that can be related by field redefinitions, equations of motion, integration by parts, or Fierz identities [77]. Depending on how these techniques are applied can lead to different basis definitions. Crucially, new physics appears equivalently in any complete basis definition. However, when considering only a subset of EFT operators in experiment, the choice of basis becomes important. Two commonly used bases in SMEFT are the SILH and Warsaw bases. The SILH basis is more suited for modified bosonic interactions and is used for the main interpretation described in Chapter 7. On the other hand, the Warsaw basis is more suited for modified fermionic interactions; this features in Chapter 7 when considering the progression of EFT measurements at CMS. More information regarding the two bases can be found in Refs. [46, 47, 78].

2.4.2 SM parameter redefinitions

One important consequence of using an EFT framework is the redefinition of the SM Lagrangian parameters. As mentioned above, the SM fields are redefined when constructing an EFT basis to remove redundant combinations of operators. Whilst this effect cancels out when calculating the relevant matrix elements, it results in a non-zero shift in the internal parameters of the SM Lagrangian: g, g', v, m_{H}, etc. These shifts, induced by the introduction of higher-dimensional operators, are accounted for by defining the Lagrangian parameters as functions of the EFT Wilson coefficients. The functions depend on the specified *input parameter scheme* of the SMEFT model. An example in the context of the Higgs boson mass[18] is described below; for a detailed description of the SMEFT parameter redefinitions see Ref. [79].

The dimension-6 operator, $\mathcal{O}_H^{(6)} = (H^\dagger H)^3$, changes the shape of the Higgs potential,

$$V(H) = \mu^2 H^\dagger H + \frac{1}{4}\lambda (H^\dagger H)^2 - w_H^{(6)} (H^\dagger H)^3, \tag{2.47}$$

where $w_H^{(6)}$ is the corresponding Wilson coefficient (which for the point of this discussion has absorbed the factor of $1/\Lambda^2$). This yields the new minimum,

$$H^\dagger H = \frac{v^2}{2}\left(1 + \frac{3 w_H^{(6)} v^2}{4\lambda}\right) \equiv \frac{1}{2} v_T^2, \tag{2.48}$$

i.e. the shift in the Higgs vacuum expectation value is proportional to $w_H^{(6)} v^2$. To provide a canonically normalised kinetic term for the Higgs boson field, h, when the Lagrangian ($\mathcal{L}_{\mathrm{EFT}} = \mathcal{L}_{\mathrm{SM}} + \mathcal{L}^{(6)}$) is expressed in terms of the mass eigenstates (after SSB), we redefine

[18] In the Warsaw basis.

the scalar doublet field as,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ [1 + w^{(6)}_{H,\text{kin}}]h + v_T \end{pmatrix}, \tag{2.49}$$

where,

$$w^{(6)}_{H,\text{kin}} = \Big(w^{(6)}_{H,\text{Box}} - \frac{1}{4} w^{(6)}_{HDD} \Big) v^2, \tag{2.50}$$

and $w^{(6)}_{H,\text{Box}}$ and $w^{(6)}_{HDD}$ are the Wilson coefficients for the operators, $\mathcal{O}^{(6)}_{H,\text{Box}} = (H^\dagger H)\Box(H^\dagger H)$ and $\mathcal{O}^{(6)}_{HDD} = (H^\dagger D^\mu H)^*(H^\dagger D_\mu H)$, respectively. The kinetic terms,

$$\mathcal{L}_{\text{EFT}} \supset (D^\mu H)^\dagger (D_\mu H) + w^{(6)}_{H,\text{Box}} (H^\dagger H)\Box(H^\dagger H) + w^{(6)}_{HDD} (H^\dagger D^\mu H)^* (H^\dagger D_\mu H), \tag{2.51}$$

and the modified Higgs potential of equation 2.47, yield a term in the EFT Lagrangian,

$$\mathcal{L}_{\text{EFT}} \supset \lambda v_T^2 \Big(1 - \frac{3 w^{(6)}_H v^2}{2\lambda} + 2 w^{(6)}_{H,\text{kin}} \Big) h^2, \tag{2.52}$$

when expressed in terms of the Higgs boson mass eigenstate field, h. This equates to a *redefinition* of the Higgs boson mass in the SMEFT framework, such that,

$$m^2_{\text{H}} = 2\lambda v_T^2 \Big(1 - \frac{3 w^{(6)}_H v^2}{2\lambda} + 2 w^{(6)}_{H,\text{kin}} \Big). \tag{2.53}$$

Chapter 3

The CMS experiment

3.1 Introduction

With a circumference of 27 km, the Large Hadron Collider (LHC) [19] at CERN is the largest particle accelerator in the world. The machine is designed to collide hadrons together with sufficiently high energy and frequency to enable stringent tests of the SM at the electroweak energy scale. The ATLAS [17], ALICE [80], CMS [18] and LHCb [81] experiments are situated at four independent locations along the LHC ring, at which the oppositely circulating hadron beams are focused and brought into collision. Each experiment consists of a particle detector apparatus to measure the products of the hadron collisions, where the design of the detector is chosen to facilitate the respective physics programme: ATLAS and CMS are general-purpose detectors designed to measure a wide range of physics processes, whereas LHCb and ALICE are more specialised, focusing on flavour physics and heavy-ion physics respectively.

The measurements presented in this thesis are performed using data collected by the CMS experiment. This chapter serves as an introduction to both the LHC and the CMS detector, and will help the reader understand how the design of these machines enables the predictions of the SM to be accurately probed using high-energy hadron collisions. After introducing the operation and design of the machines in Sections 3.2 and 3.3, the focus shifts towards the techniques used to reconstruct the collision products in the CMS detector, detailed in Section 3.4. Here, particular attention is given to the objects which are most relevant for the $H \rightarrow \gamma\gamma$ measurements outlined in Chapters 4–6. Following this, the use of Monte-Carlo simulation to accurately predict the behaviour of collision events in the detector is detailed.

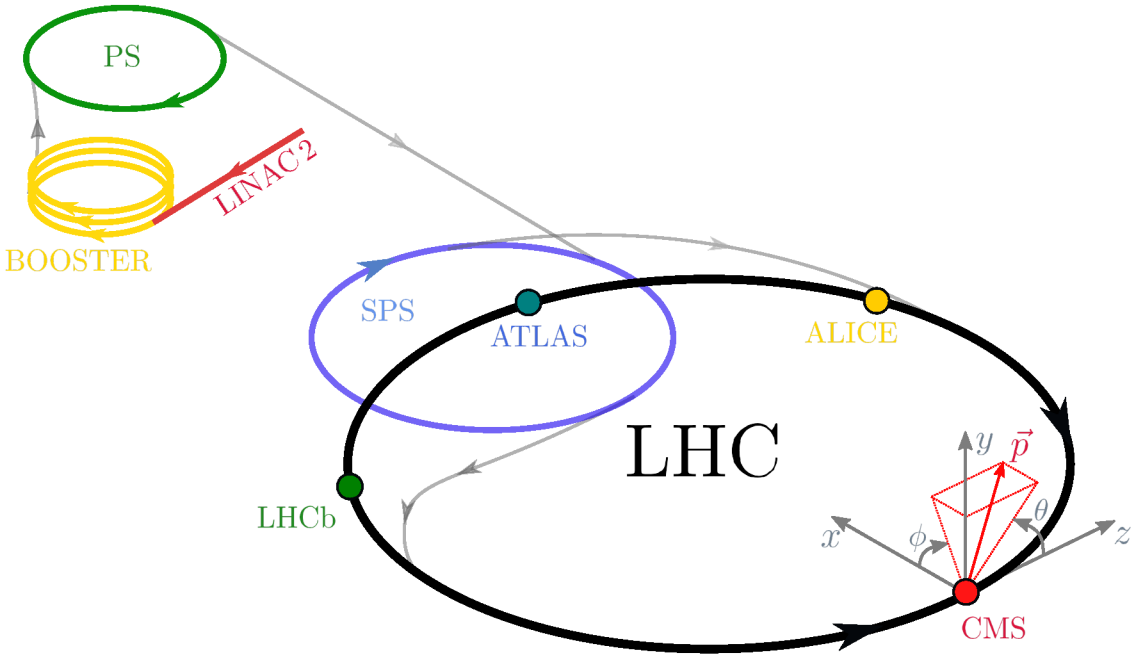


Figure 3.1: An illustration of the CERN accelerator complex, including the LINAC 2, BOOSTER, PS and SPS which sequentially accelerate the proton beams before injection into the LHC ring. The four main LHC experiments are also shown: ATLAS, ALICE, CMS, and LHCb. The co-ordinate system used by the CMS experiment is indicated for an outgoing particle with momentum, $\vec{p}$. The sizes of the rings are not to scale, and the physical positions serve as a rough guide to the actual layout at CERN.

3.2 Large Hadron Collider

The LHC crosses the French-Swiss border, and is situated 100 m underground in the tunnel which previously housed the LEP collider [82]. Both proton-proton (p-p) collisions and heavy ion collisions are performed, where the former is used for measurements at the electroweak scale such as those presented in this thesis. Therefore p-p collisions will be the focus of this section. A chain of machines, known as the CERN accelerator complex, are used to progressively accelerate protons to higher and higher energies, until they are eventually injected into the LHC ring and brought into collision. The CERN accelerator complex, including the LHC and its experiments are illustrated in Figure 3.1.

Firstly, protons are extracted by stripping electrons from hydrogen atoms using a strong electric field. The protons are sequentially accelerated to an energy of 50 MeV by the Linear Accelerator 2 (LINAC 2), to 1.4 GeV in the BOOSTER, and to 25 GeV in the Proton Synchotron (PS). Here, they are additionally spaced into bunches, with each bunch containing several billion protons. Following this, the bunched beams are fed into the Super Proton Synchotron (SPS), accelerated to an energy of 540 GeV and finally

injected into the two concentric LHC beam-pipes. The beam travels clockwise in the first pipe, and counter-clockwise in the second, producing two counter-circulating proton beams in the LHC ring. This injection is performed until each beam consists of 2808 proton bunches, with a spacing between them of around 25 ns.

A series of 1,232 super-conducting dipole magnets are located along the LHC ring to keep the beams in a circular orbit. These bending magnets are cooled to a temperature of 1.9 K using superfluid helium. Sixteen radiofrequency (RF) cavities are used to accelerate the beams from 540 GeV to the final beam energy. As the beam energy increases, the magnetic field delivered by the bending magnets is increased accordingly to maintain the circular trajectories of the beams. Currently, the highest energy reached for stable operation is 6.5 TeV per beam, which corresponds to a bending magnetic field of 8.3 T. Quadrapole magnets are then used to focus the proton beams at the four interaction points, where the beams are made to collide every 25 ns with a corresponding centre-of-mass energy of $\sqrt{s} = 13$ TeV. Note, this is slightly below the maximum LHC design energy of 14 TeV, which would require an energy of 7 TeV per beam; this is expected to be achieved either during Run 3 of the LHC (beginning 2022) or during the High-Luminosity LHC operation[1] (beginning 2027).

3.2.1 Luminosity

The rate of a particular physics process, R, in an LHC experiment is governed by the following relation,

$$R = \sigma(\sqrt{s}) \cdot \mathcal{L}_{\text{inst}}, \tag{3.1}$$

where σ is the cross section of the process of interest, and $\mathcal{L}_{\text{inst}}$ is the instantaneous luminosity of the LHC machine. The cross section depends on the collision centre-of-mass energy, $\sqrt{s}$, such that raising the collision energy can increase the probability of rare processes involving heavy particles e.g. the Higgs boson. The instantaneous luminosity depends solely on the beam parameters according to [19],

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

where n_b is the number of bunches per 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 reduction factor which accounts for the crossing angle of the beams at the collision point.

[1] See Chapter 8.

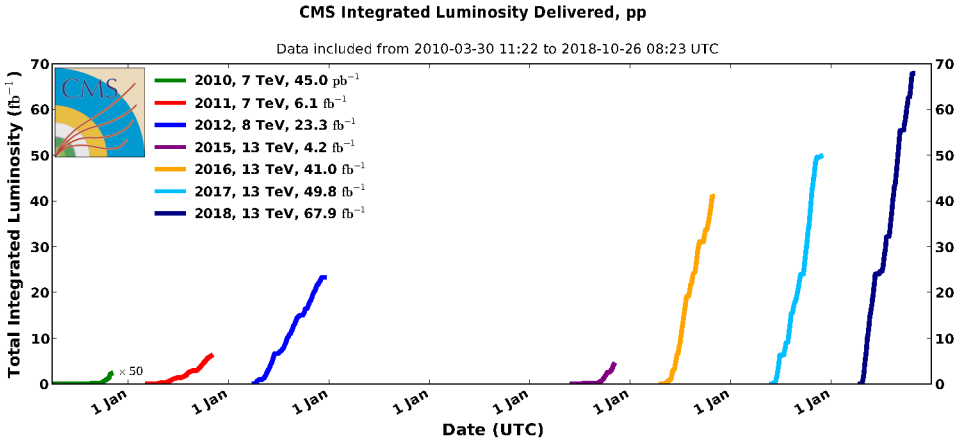


Figure 3.2: The total integrated luminosity delivered to the CMS experiment as a function of time, for each year of operation. Figure is taken from Ref. [83].

Ultimately, the exploration of rare events in an LHC experiment requires both high energy and high luminosity.

The LHC was initially designed to run with an instantaneous luminosity of 10^{34} cm^2 s^{-1}. During the 2016-2018 data-taking period this design luminosity was exceeded, eventually levelling at a value of 2×10^{34} cm^2 s^{-1} for most of the 2018 operation. By integrating the relation in equation 3.1, we arrive at an expression for the number of events of the process of interest, $N = \sigma \cdot \mathcal{L}$, where $\mathcal{L}$ is the time-*integrated luminosity*, and is a direct measure of the amount of p-p collision data delivered to a collider experiment. Figure 3.2 summarises the total integrated luminosity delivered to the CMS experiment throughout its operation. There have been two active phases of the LHC, separated by a shutdown period for upgrades. Run 1 began in 2010 with $\sqrt{s} = 7$ TeV, continuing to 2011, such that a total of 6.1 fb^{-1} of data were collected by CMS at this centre-of-mass energy. In 2012, the energy was increased to $\sqrt{s} = 8$ TeV, and a further 23.3 fb^{-1} of data were collected. This Run 1 data set was the one used for the Higgs boson discovery [14–16].

Run 2 commenced in 2015 and finished in 2018, where protons were collided with $\sqrt{s} = 13$ TeV for the full data-taking period. The increase in instantaneous luminosity (gradient of the lines in Figure 3.2) during this period has allowed for an extremely large p-p collision data set to be accumulated, therefore enabling a large improvement in the statistical precision of the measured processes of interest. The results presented in this thesis use data collected during the 2016-2018 data-taking period. In practice, the CMS experiment operates with a data-taking efficiency of less than 100%, such that the amount of data recorded by the experiment and available for the physics analyses shown in this thesis is approximately 137 fb^{-1}.

One of the drawbacks of increasing the instantaneous luminosity is the enhancement of

pileup, defined as the number of additional inelastic p-p collisions for each hard-scattering process of interest. As pileup increases, more sophisticated techniques are required to separate the rare process of interest from the objects originating from pileup interactions. In 2016, the mean number of pileup interactions was 23 per bunch crossing, rising to 32 in both the 2017 and 2018 periods. During the HL-LHC phase of operation, the pileup will increase up to a maximum value of around 200, which poses a major challenge to maintain the current excellent reconstruction performance of the CMS detector.

3.3 The CMS detector

CMS is one of two general purpose particle detectors at the LHC [18], located close to the French village of Cessy. It is over 28 m long, 15 m in diameter, and weighs approximately 14,000 tons. The detector is designed to overcome the experimental challenges that arise in a high-energy collision environment with $\mathcal{O}(1000)$ charged particles being produced every 25 ns. This includes a high level of spatial and timing granularity, with many synchronized detector electronic channels, to maintain a sufficiently low occupancy in these conditions. In addition, the detector and electronics must be sufficiently radiation-hard to accommodate the high flux of particles.

One of the main goals of the CMS physics programme was the discovery, and is now the measurement of the Higgs boson and its interactions with other particles. Moreover, the programme includes the precise measurement of other rare processes in the SM, and the search for new BSM physics such as supersymmetry or extra dimensions. To achieve these goals, the detector is designed to:

- Identify and reconstruct muons with excellent efficiency and precision. This must be achieved over a wide range of muon energies and angles. In addition, the charge of a muon must be ascertained to a high level of accuracy. The reconstruction of muons is central to Higgs boson measurements, specifically in the $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ decay channel.

- Achieve a good momentum resolution for charged particles, as well as have the ability to locate secondary interaction vertices consistent with the decay of B hadrons (hadrons containing b quarks) and τ leptons.

- Measure the energy of electrons and photons with excellent resolution over a wide geometrical coverage. Additionally, the detector is able to isolate photons and electrons efficiently in a high occupancy environment. These characteristics are key to the $\mathrm{H} \to \gamma\gamma$ measurement described in Chapters 4-6.

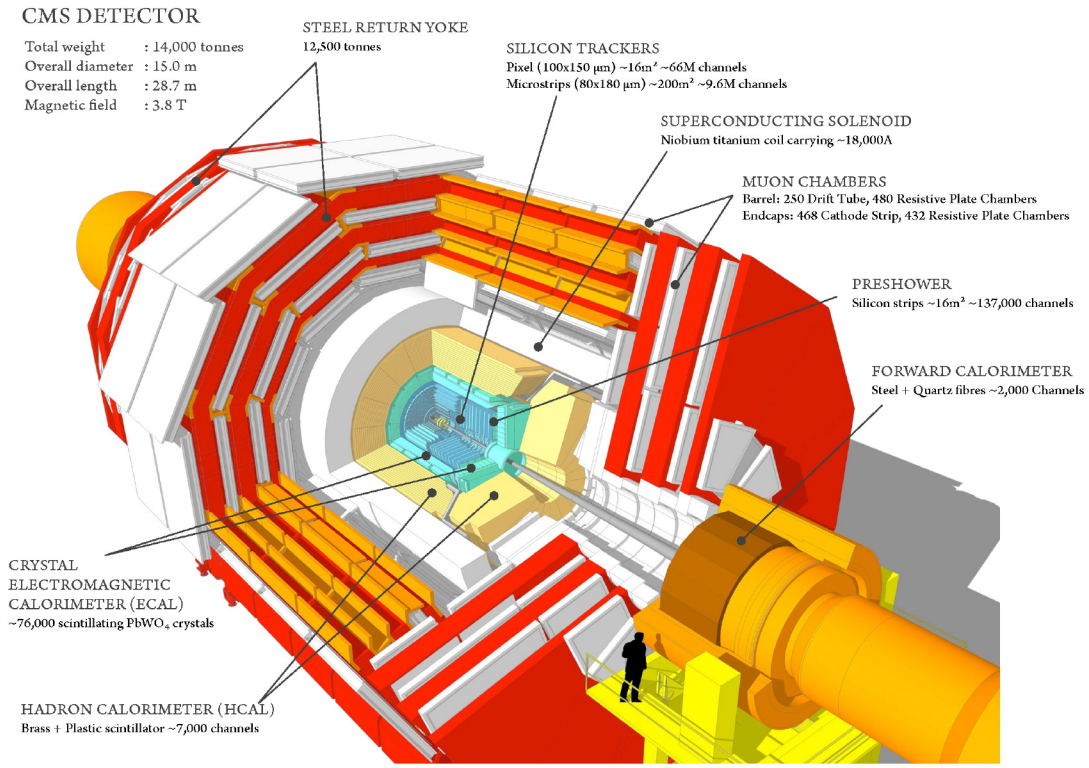


Figure 3.3: A schematic of the CMS detector. Part of the detector has been removed so that the layout is visible. Figure is taken from Ref. [84].

- Identify sprays of hadrons, known as jets, which originate from the hadronisation of quarks and gluons, and achieve a good dijet mass resolution.
- Accurately calculate the missing momentum in an event, which is the key signature of neutrinos or potential BSM particles which do not interact with the detector material.

A schematic of the CMS detector is provided in Figure 3.3. The detector consists of a number of components layered around the beam axis, where each component is comprised of a cylindrical *barrel* section and two *endcaps*. A key feature of this (almost) hermetic cylindrical design is the 3.8 T superconducting solenoid magnet, which provides an extremely high bending power for charged particles traversing the inner region of the detector. Within the coil of the 13 m long, 6 m in diameter solenoid lies the silicon tracker (Section 3.3.2), the homogeneous crystal electromagnetic calorimeter (Section 3.3.3), and the sampling hadronic calorimeter (Section 3.3.4), listed in increasing distance from the interaction point. The muon detection system (Section 3.3.5) is embedded within the iron return yoke of the solenoid. This system of subdetectors, and their respective layering,

enables the precise reconstruction of the wide range of final-state particles produced in hadron collisions.

3.3.1 Co-ordinate system

A right-handed Cartesian co-ordinate system is adopted, centred at the nominal interaction point, such that the x-axis points towards the centre of the LHC ring, the y-axis points vertically upwards, and the z-axis points along the beam pipe in the direction of the counter-clockwise beam. It is more convenient to use a cylindrical co-ordinate system where the direction of an outgoing particle is expressed using the angular quantities: ϕ and η. Here, $\phi \in [-\pi, \pi]$ is defined as the azimuthal angle in the $(x - y)$ plane, relative to the x-axis. The quantity η, referred to as the pseudorapidity, is a measure of the polar angle relative to the beam axis, θ, such that,

$$\eta = -\ln[\tan(\theta/2)]. \tag{3.3}$$

Particles with high values of η correspond to a direction close to the LHC beam pipe, and are said to be *forward*. The distance measure in the (η, ϕ) space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

In processes of interest, particles are generally produced with a high momentum in the plane perpendicular to the beam axis. As a result, a useful quantity to characterise a particle is the transverse momentum, $p_T = \sqrt{p_x^2 + p_y^2}$, defined as the projection of the particle's total momentum onto this transverse plane. Finally, the missing transverse momentum, p_T^{miss}, is defined as the magnitude of the negative vector sum of visible momenta in the transverse plane. The CMS co-ordinate system is shown for a particle with momentum, $\vec{p}$, in Figure 3.1.

3.3.2 Tracker

The tracker is the innermost component of the CMS detector [18, 85], and is designed to measure the trajectory of charged particles deflected by the 3.8 T magnetic field. It is also able to accurately locate the position of the primary hard-scattering interaction vertex, as well as identify secondary-interaction vertices which originate from the decay of τ leptons or B hadrons.

Being the closest subdetector to the interaction point (IP), the tracker experiences an intense particle flux of $\mathcal{O}(1000)$ particles every 25 ns. As a result, it must simultaneously be able to withstand severe radiation damage, whilst exhibiting excellent spatial and temporal granularity to correctly identify trajectories and attribute them to the correct proton bunch crossing. To achieve such levels of granularity, a complicated system of

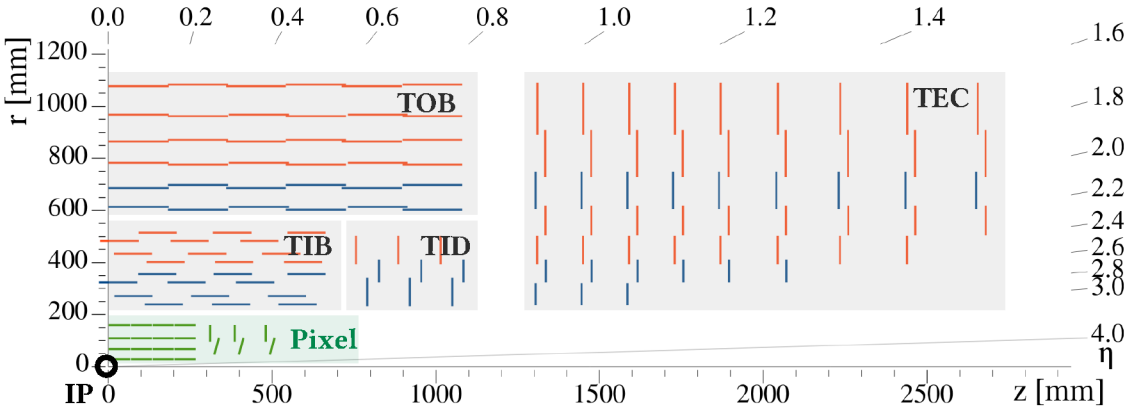


Figure 3.4: A diagram showing one quarter of the CMS tracker in the r-z view, where r is a measure of the radial distance in the (x-y)-plane. The pixel detector, closest to the interaction point (IP, black circle) is shown in green. The sections of the silicon strip tracker (TIB, TID, TOB, TEC) are also shown, where the red and blue lines signify one-sided and two-sided strips respectively. Figure has been adapted from the original in Ref. [86].

power-dense on-detector electronics are required, which in-turn require an efficient cooling system to obviate overheating. These technical requirements directly oppose the need to limit the amount of material in the tracker to mitigate unwanted interactions, such as bremsstrahlung and photon conversions. Effectively, a larger amount of material in the tracker leads to a worse energy measurement in the electromagnetic calorimeter, directly impacting the sensitivity of the $H \rightarrow \gamma\gamma$ analysis. Ultimately, to satisfy the requirements of radiation hardness and granularity, whilst keeping the material budget to a minimum, the tracker is composed entirely of silicon detector technology.

Figure 3.4 shows one quarter of the CMS tracker, where the full tracker layout is symmetric about the r and z-axes. A pixel detector is located closest to the beampipe. The original pixel detector was designed to operate for ten years at a maximum instantaneous luminosity of 10^{34} cm^{-2} s^{-1}, and consisted of three barrel layers at radii of 44, 73, and 102 mm, and two endcap disks at distances of 345 and 465 mm from the IP. To accommodate the enhancement of the LHC instantaneous luminosity, an upgraded pixel detector was installed during the end-of-year technical stop of the LHC in 2016/2017 [87]. The upgraded detector lies closer to the beampipe with *four* barrel layers at radii of 29, 68, 109, and 160 mm, and *three* endcap disks at distances of 291, 396, and 516 mm from the IP. This enables the measurement of four high precision space points (hits) on each charged particle trajectory, over the pseudorapidity range: $|\eta| < 2.5$. Additionally, the upgrade brings an improved performance at higher rates, increased radiation tolerance and offers more robust tracking. The upgraded pixel detector consists of approximately 124 million individual silicon pixels, each of size 100 µm × 150 µm, covering a total area

of approximately 1.9 m^2. This results in a spatial resolution of around 10 µm in the transverse (r-ϕ) direction and around 20 µm in the longitudinal (z) direction.

Beyond a radius of 200 mm, the reduced particle flux allows for the use of silicon strip detectors. The CMS strip tracker consists of three sub-systems. The Tracker Inner Barrel and Disks (TIB/TID) contains four layers of 320 µm thick silicon strip sensors in the barrel, supplemented by three disks of the same width at each end. The Tracker Outer Barrel (TOB) encompasses the TIB/TID, extending out to a radius of 1200 mm from the beam pipe, and ±1180 mm in the z-direction. It consists of six 500 µm thick layers, positioned with their strips parallel to the beam axis. Beyond this z-range, the Tracker Endcaps (TECs) extend out to ±2820 mm in z, providing a pseudorapidity coverage of $|\eta|< 2.4$. Each TEC is composed of 9 disks, carrying as many as seven rings of silicon strip sensors.

Each silicon strip sensor provides a one-dimensional measurement in ϕ of a point along a charged particles trajectory. In addition, the first two layers of the TIB/TID and the TOB, as well as the first, second and fifth rings of the TEC are supplemented with a second strip sensor to provide a measurement of a second spatial co-ordinate: z in the barrel and r in the disks. All in all, the CMS strip tracker has a total of 9.3 million strips, corresponding to 198 m^2 of active silicon.

Charged particles follow helical trajectories in the solenoidal field. By making several precise measurements of hits in both the pixel and strip tracker systems, these charged-particle trajectories (tracks) can be reconstructed (see Section 3.4). The momentum of the outgoing particle is then inferred from the curvature of the track, with a resolution of 2% for high-p_T (100 GeV) charged particles up to $|\eta| \approx 1.5$. This momentum resolution worsens as a function of charged particle p_T, as the curvature of the track decreases. All tracks are extrapolated back to a common point of origin, to identify the primary hard-scattering vertex and any secondary interaction vertices. The performance of this vertex location is driven by the excellent longitudinal resolution of the pixel detector.

3.3.3 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) [18, 88, 89] is used to reconstruct the energy of electromagnetic showers originating from electrons and photons, and therefore is the key sub-detector in the $H \rightarrow \gamma\gamma$ analysis. The overall structure of the CMS ECAL is shown in Figure 3.5.

The ECAL is a homogeneous calorimeter consisting of 75,848 lead-tungstage ($PbWO_4$) scintillating crystals. It is comprised of the ECAL barrel (EB) section, covering the pseudorapidity region, $|\eta| < 1.48$, and two ECAL endcaps (EE), which extend the coverage up to $|\eta| < 3.0$. The barrel and endcaps are separated by a transition region, $1.44 < |\eta| < 1.57$,

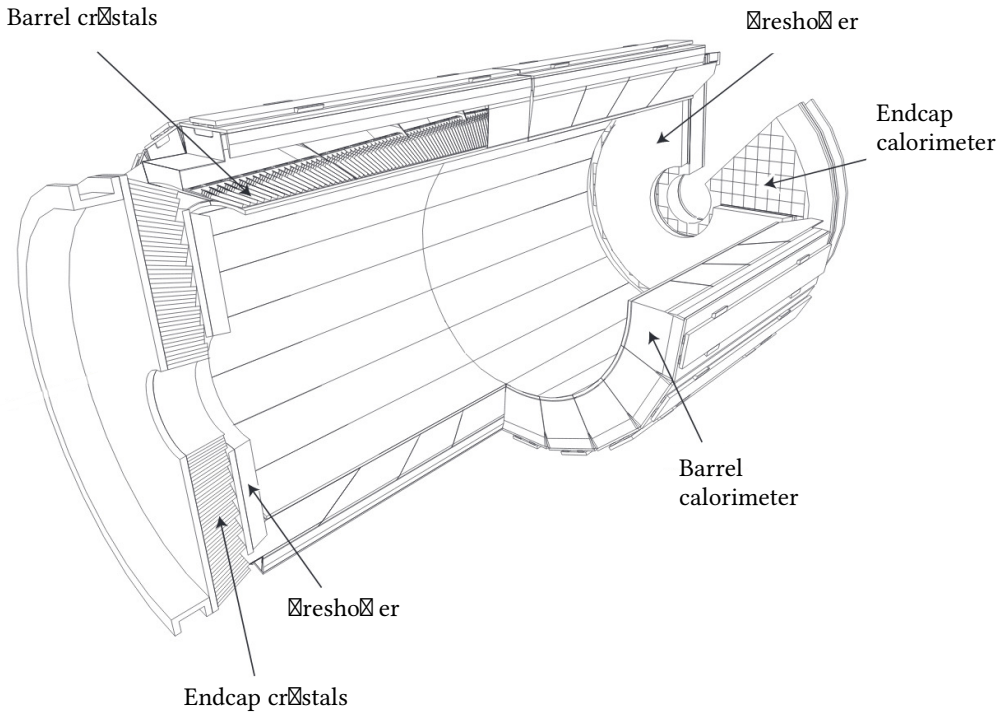


Figure 3.5: A schematic of the CMS ECAL. The tracker and LHC beampipe have been removed from the diagram, in addition to a section of the ECAL, for clarity. Figure has been adapted from that shown in Ref. [18].

in which electromagnetic showers cannot be reconstructed.

When photons and electrons interact with ECAL crystals, they produce scintillation light which is collected by photo-sensors to measure the energy of the incident particle. The choice of $PbWO_4$ is made due to the following properties of the material:

- Short radiation length, defined as the mean distance over which an electron loses all but $1/e$ of its energy due to bremsstrahlung, of approximately $X_0 = 0.89$ cm. This means the longitudinal extension of the electromagnetic shower is kept to a reasonable level.
- Narrow Molière radius, defined as the average radius containing on average 90% of a shower's total energy deposit, of $r_{\mathrm{M}} = 2.19$ cm. This means the lateral extension of the electromagnetic shower is kept to a reasonable level.
- Fast response time, such that approximately 80% of the total scintillation light is emitted by the crystals in 25 ns. This is necessary so that the majority of the energy is collected before the next proton bunch crossing.
- Hard radiation tolerance.

The crystals are arranged in a *quasi*-projective geometry, such that their axis makes a small angle (3°) with the vector pointing directly from the nominal IP. This ensures no particle trajectories are completely aligned with the cracks between crystals and therefore a large fraction of all electromagnetic showers are contained within the crystals. The EB (EE) crystals are ∼26 (∼25) X_0 long, meaning that electromagnetic showers up to an energy of approximately 500 GeV are fully contained within the ECAL. The front-face cross section of the EB (EE) crystals are 2.2×2.2 cm^2 (2.86×2.86 cm^2); this size is comparable to the Molière radius, and therefore provides a handle on the shape of the electromagnetic shower which can subsequently be used for photon/electron identification. In the current CMS ECAL, the crystals are arranged in a single layer. A future upgrade of the calorimeter endcaps, known as the HGCAL (Section 8.2), will also exhibit longitudinal segmentation, and therefore provide granular measurements of the electromagnetic shower in the direction of propagation.

One of the disadvantages of $PbWO_4$ is the relatively low light-yield, which demands the use of photo-sensors with internal amplification inside the CMS solenoidal magnetic field. Silicon avalanche photodiodes (APDs) operating with an amplification factor of around 50, and vacuum photo-triodes (VPTs) operating with an amplication factor of around 10, are used in the EB and EE, respectively. Both produce roughly 4,500 photo-electrons per GeV, which are subsequently digitised using a 12-bit analog-to-digital converter (ADC), and stored as discrete amplitude measurements in a buffer. For a given shower, each crystal stores ten consecutive amplitude measurements. If the event is deemed of interest and a trigger (Section 3.3.6) is received, the ten measurements pass to the off-detector electronics, where the amount of energy deposited in the crystal is inferred from the pulse shape.

An additional subdetector, referred to as the preshower detector (ES), is mounted in front of each endcap, providing a coverage of the pseudorapidity region: $1.65 < |\eta| < 2.6$. The ES is a 20 cm thick sampling calorimeter composed of two alternating layers of lead (to initiate the electromagnetic showers from incoming photons and electrons), and silicon strips (to measure the deposited energy and the transverse shower profile). The main aim of the ES is to distinguish neutral pions (π^0), from true photons in this η range.

Electromagnetic shower energy reconstruction

The total width of the Higgs boson, Γ^H, is many orders of magnitude smaller than its mass (∼4 MeV in the SM). This means the width of the Higgs boson decay-products invariant-mass distribution is entirely driven by the experimental resolution. Consequently, in the case of the $H \to \gamma\gamma$ analysis, the sensitivity is driven by the ECAL energy resolution.

The intrinsic energy resolution of the ECAL crystals is modelled according to

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{3.4}$$

where $S = 2.8\%$ is the stochastic term, $N = 12\%$ is the noise term, and $C = 0.3\%$ is the constant term, whose values have been derived using test-beam data [18]. The energy, E, is expressed in units of GeV, and corresponds to the sum of energy in a 5×5 array of ECAL crystals. The energy resolution can be improved using a series of corrections, described below.

A typical photon or electron shower is spread over many ECAL crystals. Showers which undergo photon conversions or bremsstrahlung in the material upstream of the ECAL typically extend in the ϕ-direction. This is a result of the charged particles (e^+/e^-) produced in the interaction being deflected in the azimuthal direction by the CMS solenoidal field. To encompass the energy deposited in the ECAL, a dynamic clustering algorithm is applied (Section 3.4), where clusters are extended in the ϕ-direction to form superclusters (SC), thus improving the containment for showers that have already undergone photon conversions or bremsstrahlung.

The reconstructed shower energy for photons or electrons in the ECAL, $E_{e,\gamma}$, is calculated according to the following equation,

$$E_{e,\gamma} = F_{e,\gamma} \cdot E_{\mathrm{SC}} = F_{e,\gamma} \cdot [G(\eta) \cdot \sum_i (C_i \cdot S_i(t) \cdot A_i) + E_{\mathrm{ES}}], \tag{3.5}$$

where the index i iterates over crystals in the SC. The individual channel amplitudes, A_i, are multiplied by a time-dependent crystal response correction, $S_i(t)$, and a channel calibration constant, C_i, before being summed and multiplied by the global ADC-to-GeV absolute energy scale factor, $G(\eta)$. Showers in the EE are supplemented with the energy measured in the preshower detector, E_{ES}. Finally, the energy of the SC, E_{SC}, is corrected by applying a multivariate regression, $F_{e,\gamma}$, trained separately for photon [90] and electron [91] showers. The regression is described in the context of photons in the following section, in preparation for the $\mathrm{H} \to \gamma\gamma$ analysis described in Chapters 4–6.

Photon energy regression

The photon energy-response of the ECAL ($E_{\mathrm{true}}/E_{\mathrm{SC}}$) is parameterised by a function with a Gaussian core and power law tails, where E_{true} is the true energy of the incident photon. Using simulated photons, a multivariate regressor is trained to estimate the shape parameters of this energy response function, thus providing a prediction of the full $E_{\mathrm{true}}/E_{\mathrm{SC}}$ probability density function for each photon [90]. The mode of this predicted

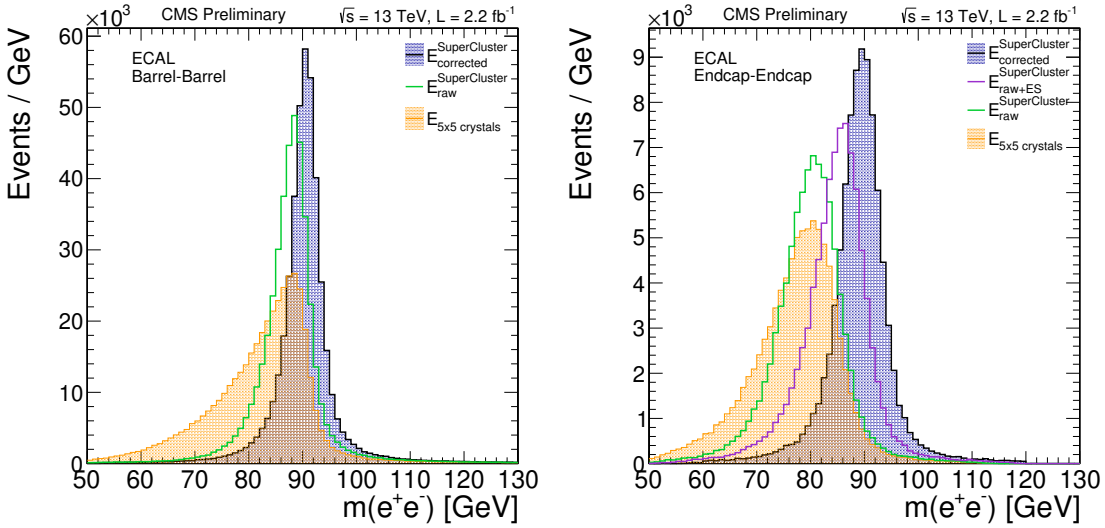


Figure 3.6: Corrections to the measured energy in the ECAL after applying the multivariate regression, illustrated using the invariant mass distribution of electron pairs from $Z \to e^+e^-$ events. The left (right) plot shows electron pairs reconstructed in the EB (EE). The yellow histogram uses the simple energy sum over the 5×5 array of ECAL crystals. The green line uses the supercluster energy after applying the clustering algorithm, and the purple line in the EE plot also includes the energy deposited in the ES detectors. Finally, the blue histogram uses the energy after all corrections are applied, including the multivariate regression. The data shown was collected during the 2015 period. Figure is taken from Ref [92].

function, F_γ, is used to correct E_{SC} to E_{true}, to give the corrected photon energy: $E_\gamma = F_\gamma \cdot E_{\text{SC}}$, whilst the shape provides a per-photon energy resolution estimate which is used in the $H \to \gamma\gamma$ analysis for categorising events. The regressor corrects not only for the imperfect shower containment arising from converting photons and electromagnetic showers that begin upstream of the ECAL, but also for the localised containment within the ECAL, where energy can be lost in the gaps between crystals. Input variables related to the SC shower shape provide information on the upstream showering and photon conversions, which combined with the SC η and ϕ values allows the regressor to learn variations in the ECAL geometry. On the other hand, the seed-crystal positions and seed cluster energy ratios enable the regressor to correct for the localised containment effects. In addition, the total number of primary interaction vertices and the total energy density, ρ, are included to account for systematic enhancements of E_{SC} due to pileup.

Final ECAL energy resolution

The impact of the full set of ECAL energy corrections (described by equation 3.5) is illustrated for electron showers in Figure 3.6. The final ECAL energy resolution after

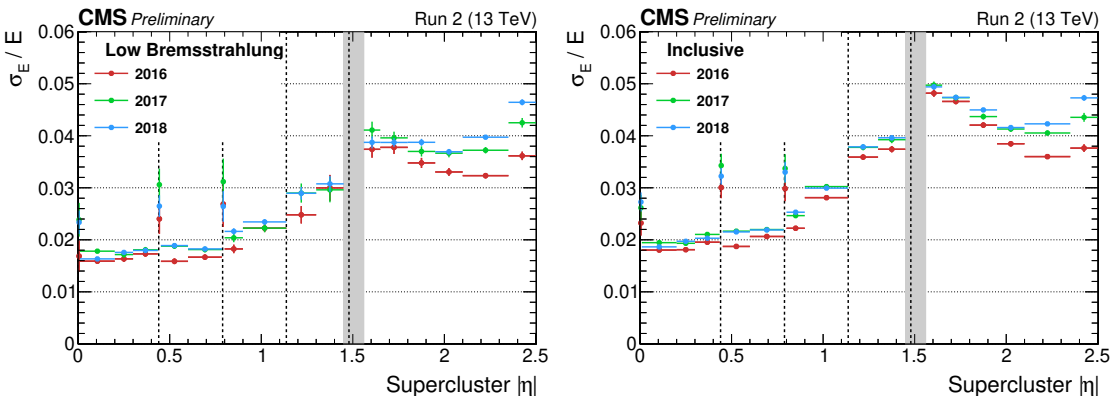


Figure 3.7: ECAL energy resolution as a function of η for electrons from $Z \to e^+e^-$ decays, after all corrections are applied. The resolutions are shown separately for each data-taking year, and both the EB and EE are shown, separated by the transition region (shaded grey). The left plot corresponds to low-bremsstrahlung electrons ($R_9 > 0.94$), where the variable R_9 is defined in Table 4.2, and the right plot to all electrons (inclusive). Figure is taken from Ref [93].

these corrections is shown as a function of η for $Z \to e^+e^-$ electrons, for each data-taking year separately, in Figure 3.7. The values are shown for low-bremsstrahlung electrons (left) and all electrons inclusively (right). A resolution of around 1.5% is observed for low-bremsstrahlung electrons ($R_9 > 0.94$) up to an $|\eta| = 1$, rising to around 3% towards the edge of the EB, and up to 4% in the EE. Inclusively, the resolution is about 2–3% up to $|\eta| = 1$, 3–4% up to $|\eta| = 1.44$, and up to 5% in the EE[2]. The plots demonstrate a (relatively) stable ECAL energy resolution over the course of Run 2, despite the increased instantaneous luminosity and the aging of the CMS detector. The worse energy resolution in the EE compared to the EB is a direct consequence of the higher radiation dose in the forward region, which affects the crystal transparency.

3.3.4 Hadronic calorimeter

Quarks and gluons produced in the proton collisions hadronise before reaching the detector, resulting in collimated sprays of particles known as jets. The CMS hadronic calorimeter (HCAL) [18, 94] is used to measure the position and energy of hadrons in jets. This is especially important for neutral hadrons which leave no track in the silicon tracker, and deposit little energy in the ECAL. Additionally, the HCAL is required for the reconstruction of the p_T^{miss}. The layout of the HCAL for one quarter of the CMS detector is displayed in Figure 3.8.

[2] As electron and photons showers are practically indistinguishable, the photon energy resolution in $H \to \gamma\gamma$ decays is approximately the same, with low-bremsstrahlung electrons mapping to unconverted photons, and high-bremsstrahlung electrons mapping to photons undergoing a conversion to e^+e^- pairs in the tracker.

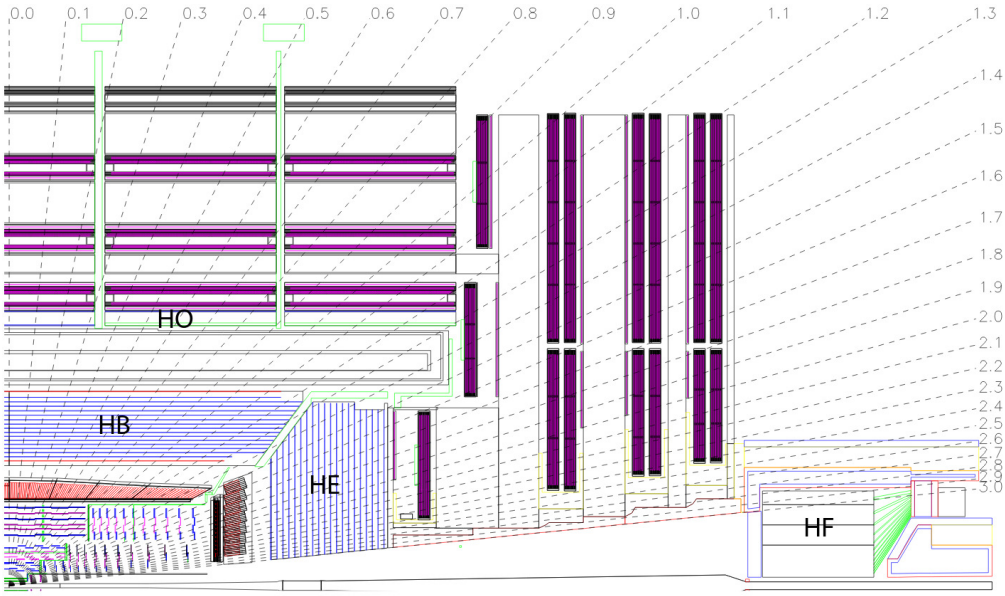


Figure 3.8: A schematic showing one quarter of the CMS detector. The dashed lines indicate lines of constant pseudorapidity. The locations of the HB, HO, HE and HF components of the HCAL are shown. Figure is taken from Ref. [18].

The HCAL is a sampling calorimeter consisting of absorber plates made from brass or steel, interleaved with active layers of plastic scintillator. Hadrons traversing the HCAL interact with the detector medium via nuclear interactions, producing hadronic showers. Light produced in the scintillator material from these showers is read out by wavelength-shifting plastic fibres, and is used to infer the energy of the incident hadron. It is important that the shower is fully contained to make an accurate measurement of the hadron energy. The HCAL is split up into four components:

- The barrel (HB) has coverage up to $|\eta| = 1.3$, and consists of 18 identical azimuthal wedges in both the $+z$ and $-z$ directions from the IP. The scintillator in each wedge is divided into 16 η-sectors, resulting in a spatial granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The HB is confined radially from the outer-edge of the EB at $r = 1.77$ m to the inner-coil of the CMS solenoid magnet at $r = 2.95$ m. This corresponds to a depth of between 5.8 and 10.6 nuclear interaction lengths (λ_I), increasing as a function of η. Here, λ_I is a property of the material defined as the mean distance a hadron travels before undergoing an inelastic nuclear interaction.

- To ensure hadronic showers in the central region are fully contained, an outer calorimeter (HO) is placed outside of the solenoid. This component treats the solenoidal coil as the absorber, and uses the same active scintillator as the HB,

extending the minimum depth to 11.8 λ_I.

- The endcap (HE) calorimeters cover the pseudorapidity range: $1.3 < |\eta| < 3$. They are designed to be particularly radiation-hard due to the increased particle flux at high η. The spatial granularity is equivalent to the HB for $|\eta| < 1.6$, reducing to $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for higher pseudorapidities. In the HE, the minimum depth is 10 λ_I.

- Additional forward calorimeters (HF) are placed 11.2 m from the IP, and extend the coverage up to $|\eta| = 5.2$. The particle flux in this region is extremely high, resulting in a very hostile environment; at the design luminosity approximately 800 GeV per p-p collision is deposited in the two HFs, compared to only 100 GeV across the rest of the detector. To withstand this extremely high radiation dose, steel-quartz fibres are chosen as the active material, encompassed by a steel absorber structure. Charged particles in the shower emit Cherenkov light in the fibres, which is read by photo-multiplier tubes. Ultimately, the HF are important for the measurement of forward jets, such as those produced in Higgs boson production via VBF, and for making the overall HCAL structure as hermetic as possible for the p_T^{miss} reconstruction. In the context of the H $\rightarrow \gamma\gamma$ analysis, the p_T^{miss} is a useful quantity for tagging events with additional neutrinos, such as those originating from Z($\rightarrow \nu\nu$)H and W($\rightarrow \ell\nu$)H production.

3.3.5 Muon chambers

Muons traverse the CMS calorimeters with few interactions. A dedicated muon tracking system [18, 95] is positioned furthest from the IP, built into the steel return yoke structure outside the solenoid magnet. Using a combination of information from the innermost silicon tracker and the muon tracking system, CMS is able to accurately identify muons, infer their charge, and measure their energy with excellent resolution. The muon tracking system is comprised of three different gaseous particle detector technologies, where the layout is shown for one quarter of the CMS detector in Figure 3.9.

Drift tube (DT) chambers are located in the barrel region, and detect muons for $|\eta| < 1.2$, whilst cathode strip chambers (CPS) cover the pseudorapidity range $0.9 < |\eta| < 2.4$. Both are complimented by a system of resistive plate chambers (RPC) for $|\eta| < 1.6$. All rely on gaseous detector technology. As a muon traverses the chamber it ionises gas molecules. The resulting ionisation electrons drift towards the anode of the detector producing an electric signal. The choice of detector technology in each region is driven by the properties of the return magnetic field at that point, the rate of muons, and the level of neutron-induced background.

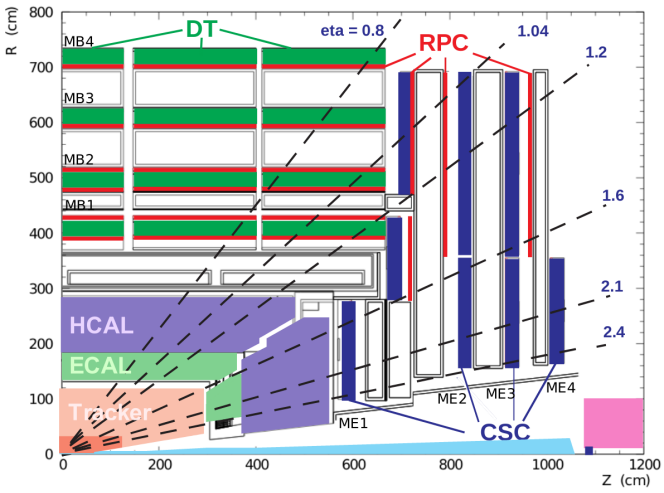


Figure 3.9: A schematic showing one quarter of the CMS detector. The dashed lines indicate lines of constant pseudorapidity. The locations of the DT, RPC and CSC of the muon tracking system are shown. Figure is taken from in Ref. [96].

Overall, muons with p_T larger than a few GeV are identified with an efficiency of above 95%; the corresponding misidentification rate is lower than 1% for a loose selection and 0.1% for a tight selection. The momentum resolution for muons with $20 < p_T < 100$ GeV is between 1.3 and 2.0% in the barrel ($|\eta| < 1.2$) and better than 6% in the endcaps ($1.2 < |\eta| < 2.4$). Over this p_T range, the momentum measurement is provided by the silicon tracker. For higher-p_T muons, the best momentum measurement is obtained using a combination of information from the silicon tracker and muon chambers, providing a resolution of better than 10% up to 1 TeV.

3.3.6 Trigger

The CMS detector operates at a proton bunch crossing frequency of 40 MHz, where each event contains of the order 1 Mb of data [97]. A two-tiered trigger system is implemented to manage this high collision rate, selecting only events of interest to be recorded. The Level-1 Trigger (L1T), composed of custom hardware processor boards, successfully reduces the output rate from 40 MHz to 100 kHz. This is compatible with the design readout rate of the CMS sub-detector electronics [98]. Selected events are then propagated to the software-based High-Level Trigger (HLT), where more sophisticated algorithms are applied using more detailed event information. The HLT further reduces the event rate to 1 kHz; an acceptable rate to be saved to disk. Crucially, this total reduction in the event rate by a factor of 40,000 is achieved whilst maintaining a high efficiency for the physics processes of interest.

The event detail used in the L1T is limited by the design latency; it must be decided

whether to keep or discard an event within a fixed time interval of 4 µs. As a result, the L1T decision is based on *coarse* measurements of the energy deposited in the calorimeters and muon chambers, and it is currently not possible to use information from the tracker. In the $H \to \gamma\gamma$ analysis, a *seed* at the L1T stage is defined as a deposit of energy in the ECAL, above a certain energy threshold. Despite signal events being characterised by two photon candidates, a higher overall efficiency is achieved if only one seed is required at the L1T stage, with a tight transverse energy threshold applied to the seed in order to limit the rate of events passing to a manageable level. The energy threshold is typically set at 40 GeV, lowering to 30-32 GeV for isolated L1T seeds in the ECAL. With the presence of this energy threshold, there is an unavoidable drop in efficiency for $H \to \gamma\gamma$ events at low transverse energy. To circumvent this effect, an additional double-seed selection is included at the L1T with lower energy thresholds than the single-seed trigger. In 2016, the thresholds were set to 23 and 10 GeV for the leading and subleading seeds respectively, rising to 25 and 14 GeV for the 2017 and 2018 data taking periods. During the processing time, the full event information is stored in a buffer. Upon reception of a L1T *accept signal*, this full information is read-out and passed to the HLT.

The HLT decision is based on more granular event information, including measurements from the tracker. A single farm of around 1000 commercially-available processors is used, where simplified versions of the full offline reconstruction algorithms are applied to the events. At this stage in the $H \to \gamma\gamma$ analysis, events are required to contain two SCs in the ECAL with invariant mass greater than 90 GeV and passing asymmetric p_T thresholds, initially set at 30 and 18 GeV. After the 2016 data-taking period, the lower threshold was raised to 22 GeV to counterbalance the increased instantaneous luminosity and hence maintain a constant event rate. In addition, a number of selection criteria are imposed on higher-level variables related to the SC shower shape, isolation and the ratio of the HCAL and ECAL deposits. If the event is deemed to be of interest, it is saved to disk and subsequently enters the $H \to \gamma\gamma$ analysis.

3.3.7 Detector summary

Figure 3.10 illustrates how different objects produced in a p-p collision interact with the CMS detector. Clearly, it is the combination of the different subdetectors that makes it possible to reconstruct the full range of final-state objects. The CMS approach to object reconstruction is detailed in the following section.

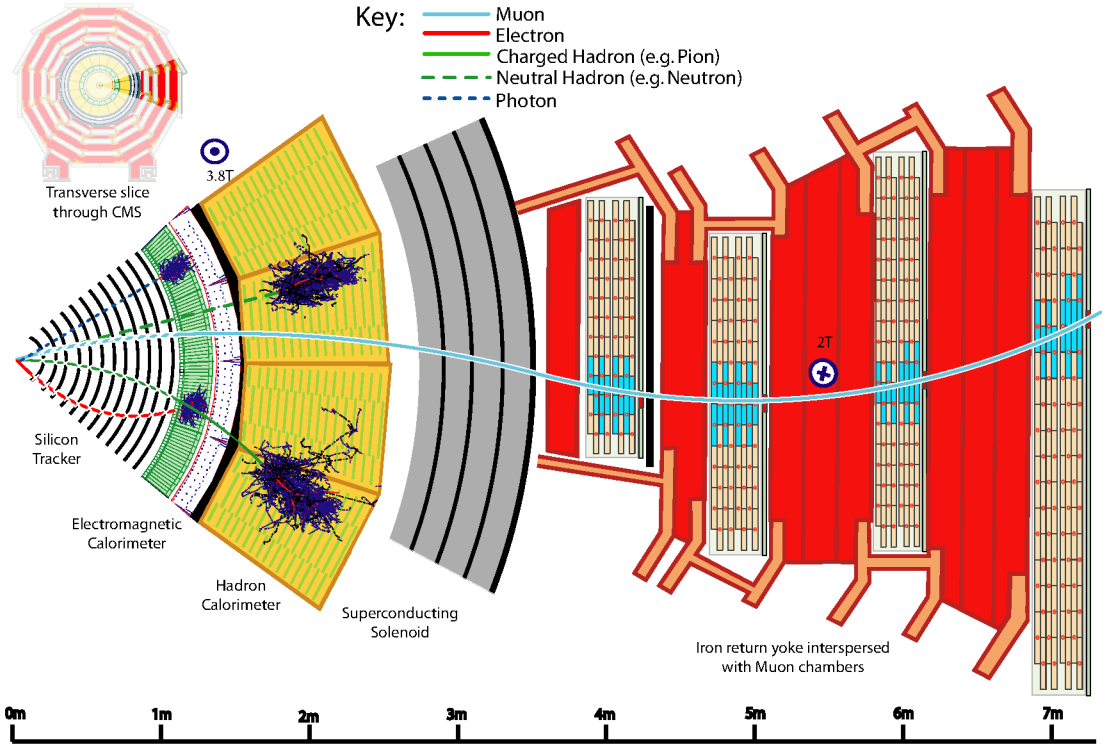


Figure 3.10: A transverse slice of the CMS detector. The diagram illustrates the typical interactions of multiple final-state objects, produced at the interaction point (furthest left), with the various CMS subdetectors. Charged particles, such as electrons, muons and charged hadrons, are deflected by the solenoidal field and leave hits in the silicon tracker. Electrons and photons create electromagnetic showers in the ECAL, whilst charged and neutral hadrons produce hadronic showers in the HCAL. Muons traverse through the detector, are deflected in the opposite direction by the return magnetic field, and leave hits in the muon chambers. Figure is taken from Ref. [99].

3.4 Object reconstruction: particle flow

CMS adopts a holistic approach to event reconstruction: the information from all subdetectors is correlated to identify each final-state particle, and the corresponding measurements are combined to reconstruct the particle properties based on this identification [99]. The comprehensive list of final-state particles produced by the algorithm then enters physics analyses. This approach, referred to as the *Particle-Flow* (PF) reconstruction, provides a global event description and offers unprecedented performance in terms of jet and p_T^{miss} reconstruction, and electron, photon, and muon identification. Furthermore, objects produced from pileup interactions can be identified efficiently, enabling powerful pileup mitigation.

The basic elements of the PF reconstruction are *tracks* in the silicon tracker and muon systems, and *clusters* of energy in the ECAL and HCAL. Tracks are built using a combinatorial track-finder algorithm [85], that proceeds in the following way:

- An initial track candidate (seed) is identified as a few (2-3) hits in the tracker, from which the initial trajectory parameters and their corresponding uncertainties are calculated.
- Tracks are located by applying a Kalman filter [100]: the expected flight path of a charged particle is extrapolated from the seed trajectory parameters, where additional hits along this path are assigned to the track candidate. Following this, a track-fitting module is used to calculate the trajectory parameters more precisely.
- The track candidate is accepted or rejected based on a set of quality criteria.

This is repeated for six iterations, in a process referred to as *iterative tracking*. The initial iterations locate the most prominent tracks, for example those with high p_T and lying close to the IP. After each iteration, hits associated with a track are removed from the process, and the quality criteria for forming seeds and building tracks are relaxed. Subsequent iterations then locate a more difficult class of tracks with low p_T and/or high displacement, in a less combinatorially-complex environment. Ultimately, this iterative procedure helps to increase the tracking efficiency, whilst keeping the misreconstructed track rate to a reasonable level. An additional procedure is used to build muon tracks from hits in the muon chambers [96].

Clusters are built by collecting the energy deposits in the calorimeters using a dedicated clustering algorithm [91]. The procedure is effectively the same for the ECAL and HCAL, but is introduced in the context of a photon shower in the ECAL here:

- A seed crystal is identified as a local energy maximum, above a given energy threshold.
- Clusters are built iteratively around the seed. This is done by aggregating crystals which share at least one corner in common with a crystal already in the cluster, and have an energy in excess of twice the noise level of the ECAL.
- If a given crystal can belong to multiple clusters, the crystal energy is shared between them assuming a Gaussian shower profile.
- Photons that convert to e^+e^- pairs in the tracker, typically have a wider shower profile. This results from the bending of the electrons/positrons in the solenoidal field, which radiate bremsstrahlung photons and thus deposit energy over a wider azimuthal (ϕ) range. Superclusters (SC) are built by merging together clusters. This

ensures good containment of the electromagnetic shower for converting photons. The spatial position of the SC (η,ϕ) is defined as the logarithmic energy-weighted average position of the individual crystals. The logarithmic energy-weighted standard deviations of the crystal η and ϕ-values, σ_η and σ_ϕ, provide an indication of the width of the SC. The median (σ_η,σ_ϕ) values for a SC in the EB are around (0.5,1) crystal lengths for an unconverted photon, and (0.5,2.5) crystal lengths for a converted photon, demonstrating a wider shower profile in the ϕ-direction for photons undergoing a conversion. The SC energy reconstruction was previously discussed in Section 3.3.3.

The procedure for reconstructing photon and electron showers in the ECAL is identical. This is an important feature of the H $\rightarrow$ $\gamma\gamma$ analysis described in Chapters 4-6, as it enables the use of Z $\rightarrow$ e^+e^- events for the photon energy calibration and for the validation of numerous multivariate algorithms.

A given final-state particle can give rise to several tracks and clusters in the various CMS subdetectors. The dedicated *link algorithm* is applied to connect these basic PF elements and output a PF candidate from the following classes:

- **Muons**: identified as a track in the muon chambers linked to a track in the tracker. Energy is calculated from the track curvature.
- **Electrons**: identified as an ECAL SC linked to a track in the tracker. PF electron energy is calculated from a combination of the SC energy and the track curvature, where the track curvature becomes useful at low p_T.
- **Photons**: identified as an ECAL SC with no associated track in the tracker. Energy is calculated from the SC energy only.
- **Neutral hadrons**: identified as linked clusters in the ECAL and HCAL, with no associated track in the tracker. Energy is calculated as the sum of cluster energies.
- **Charged hadrons**: identified as linked clusters in the ECAL, HCAL, and a track in the tracker. Energy is calculated from a combination of the track curvature and the sum of cluster energies.

Collections of PF candidates are then used in CMS physics analyses. Chapter 4 details how they are used in the H $\rightarrow$ $\gamma\gamma$ analysis.

3.5 Simulating hadron collisions

The goal of the CMS experiment is ultimately to use the final-state objects in the detector to infer some knowledge on the parameters of the model Lagrangian, $\vec{\alpha}$. This could be, for

example, the mass of the Higgs boson, the couplings of the Higgs boson to other SM fields, or even the Wilson coefficients of an EFT. We can establish a likelihood function [101] which quantifies the probability of observing an event, x, given the model parameters, $\vec{\alpha}$,

$$L(x \,|\, \vec{\alpha}) = \int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p \, p(x|z_d)p(z_d|z_s)p(z_s|z_p)p(z_p|\vec{\alpha}), \tag{3.6}$$

where x is a vector of observables such as the reconstructed energies, momenta, and angles of all final-state objects. The latent variables, z_i, are defined as follows:

- z_p: the four-momenta, charges and helicities of the partons in the hard-scattering event. The quantity $p(z_p|\vec{\alpha})$, which defines the probability for partons, z_p, given the model parameters, $\vec{\alpha}$, is related to the distribution of incoming partons in the colliding protons and the squared hard-scattering matrix element, $|\mathcal{M}|^2$. The collection of incoming and outgoing partons is referred to as the *parton-level* event.

- z_s: encodes the entire shower history of the partons. Effectively, $p(z_s|z_p)$ describes the transition from the parton-level event, z_p, to the *particle-level* event, z_s.

- z_d: describes the interactions of the particles with the detector, such that $p(z_d|z_s)$ describes the probability of observing a set of electronic signals in the detector given the particles, z_s. This is not a simple problem; the CMS detector has around 10^8 read-out channels. The final term, $p(x|z_d)$, explains how the detector signals are reconstructed into the final-state objects (Section 3.4).

The sheer complexity of the problem is illustrated in Figure 3.11. Clearly, evaluating the integral of equation 3.6 over the full parameter space is impossible: we cannot simply take the observed events, x, and the model parameters, $\vec{\alpha}$, and compute the likelihood, $L(x \,|\, \vec{\alpha})$.

To overcome this experimental challenge, we approximate the integral using a series of state-of-the-art Monte Carlo (MC) simulators. Firstly, the MC generation chain begins with simulating the hard-scattering process. This models the distribution of initial-state partons inside the colliding protons, computes the relevant matrix element, $\mathcal{M}_{i\to f}$, for certain values of the input model parameters, $\vec{\alpha}$, and generates parton-level final states with some probability. Following this, the parton shower is simulated, where initial and final-state coloured partons are showered and formed into hadrons, and any remaining unstable particles (e.g. H, $\mathrm{W}^{\pm}$, etc.) are decayed. Additional secondary interactions to the hard-scattering process are modelled as underlying events. Finally, the events are propagated through a simulation of the CMS detector, where the response of the detector is calibrated to match the performance and efficiency observed in data. The final output of this chain is a collection of simulated collision events which aims to match real events

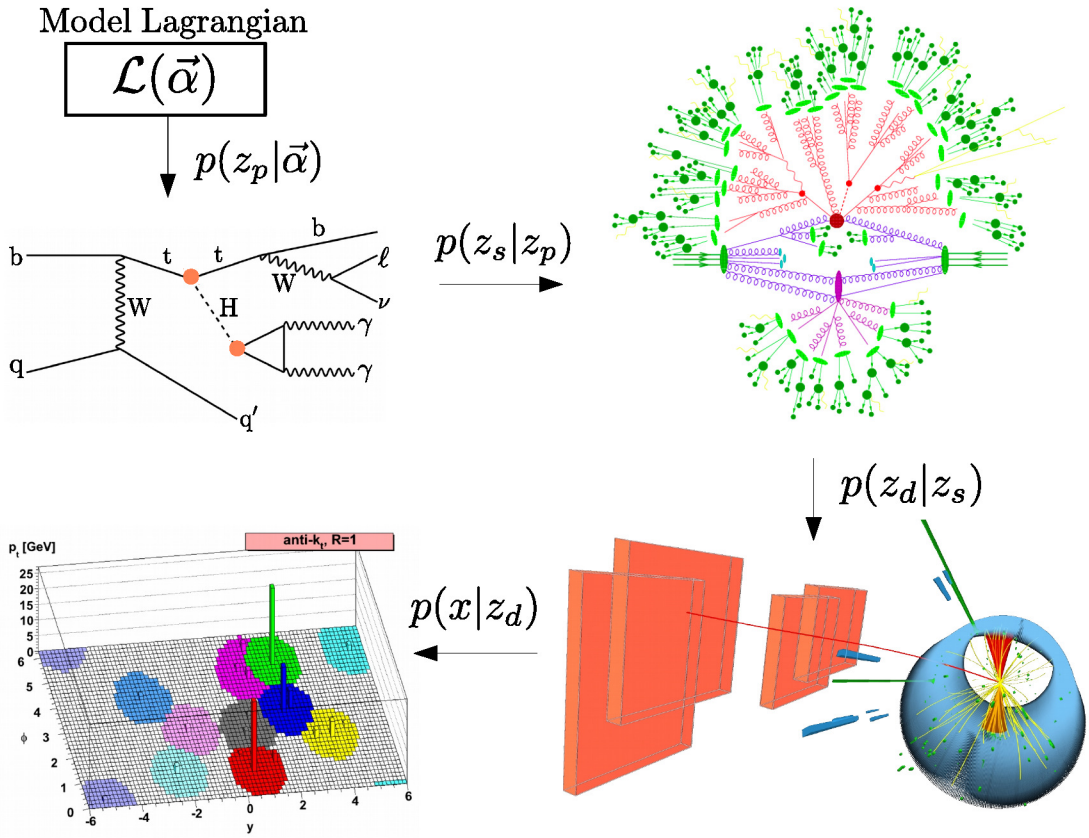


Figure 3.11: An illustration of the intractable likelihood function of the CMS experiment. The Feynman diagram (top left) corresponds to tHq production in the H $\to \gamma\gamma$ decay channel, in which the top quark decays leptonically. The event display (bottom right) corresponds to a candidate tHq, H $\to \gamma\gamma$ event in the CMS detector, adapted from the additional material of Ref. [102]. The parton shower diagram (top right) and the anti-k_T algorithm illustration (bottom left) are taken from Refs [103] and [69], respectively.

in data as closely as possible. Ultimately, we can then compare the MC simulation to data to infer knowledge on the model Lagrangian parameters, $\vec{\alpha}$.

MC simulation is used in many ways in this thesis, most notably for the SM simulation of events in the H $\to \gamma\gamma$ analysis (Section 4.2.2), and for deriving the EFT parametrisation of Higgs boson cross sections and decay widths (Section 7.4). The remainder of this section will help the reader understand some of the key concepts of MC event generation that are discussed throughout.

Parton-level event generation

The fully differential cross section for a hard-scatter process $(a, b \to f)$ in a hadron collision is calculated by factorising the initial-state physics from the hard-scatter, according to

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Phi_n} = \sum_{a,b} \int_0^1 \mathrm{d}x_a \mathrm{d}x_b f_a(x_a, \mu_F) f_b(x_b, \mu_F) \times \frac{1}{\hat{s}} |\mathcal{M}_{a,b\to f}|^2(\Phi_n; \mu_F, \mu_R, \vec{\alpha}), \tag{3.7}$$

where Φ_n represents the final-state phase space, $|\mathcal{M}_{a,b\to f}|^2$ is the squared-matrix element averaged over the initial-state spin and colour degrees of freedom, and $1/\hat{s} = 1/2x_a x_b s$ is the initial-state parton flux, where s is the squared hadronic centre-of-mass energy. The quantities $f_a(x_a, \mu_F)$ and $f_b(x_b, \mu_F)$ model the distribution of the initial-state partons, a and b, inside the colliding protons. The factorisation of the processes at different energy scales is possible since the hard-scatter occurs over a very short distance (high energy), whilst the interactions within the proton occur over much longer distances (low energy). This means that when computing the hard-scatter process, we can safely ignore the dynamics within the proton and treat the initial-state partons, a and b, as independent. As a result, we can use perturbation theory to calculate hard-scatter matrix elements.

In the parton-level event generation, events are sampled from the initial and final-state phase spaces, and a probability is assigned according to equation 3.7. By generating large numbers of events, we can successfully build up the kinematic distributions that one would expect to see in data (for some values of the model parameters, $\vec{\alpha}$). A cross section, σ^i, for a particular region of the final-state phase space, Φ_n^i can then be inferred by integrating equation 3.7 over this region. In the STXS framework, the regions are defined according to the bin boundaries introduced in Section 2.3.

Factorisation and renormalisation

Two of the main theoretical aspects of the MC simulation are factorisation and renormalisation. As introduced in the previous section, factorisation concerns the decoupling of the long-distance (low energy) physics from the short-distance (high energy) hard-scattering process, such that the different energy regimes can be simulated separately [104]. A factorisation scale, μ_F, is introduced to avoid IR divergences in the hard-scattering calculation. This scale effectively defines the borderline between the low energy and high-energy physics.

Renormalisation details the regularisation of UV divergences at the opposite end of the energy range [105]. These arise due to the unrestricted integration of the momentum flow through internal loops in the Feynman diagrams. Regularisation gives rise to a renormalisation scale, μ_R, which acts as an UV cut-off. The bare couplings of the theory

absorb all the very short-distance physics beyond this cut-off, and as a consequence, acquire a μ_R dependence. For example, in QCD the contribution of quarks and gluons in loops is absorbed into the strong coupling constant[3] as,

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \frac{\beta_0}{4\pi}\,\alpha_S(\mu_R^2)\,\ln\left(\frac{Q^2}{\mu_R^2}\right)}, \tag{3.8}$$

where Q^2 represents the energy scale of the process. The constant β_0 is positive (> 0), which confirms the notion of *asymptotic freedom* [55]: the strength decreases ($\alpha_S(Q^2) \to 0$) as the energy scale increases ($Q^2 \to \infty$). At very high-energy scales, the strongly-interacting partons can be treated as independent particles.

Crucially, the underlying physics does not depend on these scales: μ_R and μ_F. They are simply an artifact of using perturbation theory to calculate the process, and truncating the expansion to a particular number of terms. Nevertheless, there exists a freedom in the choice of their values, where some choices can be better than others. Since we know an observable (e.g. a cross section) should not depend on μ_R and μ_F, it is theoretically favourable to choose values where the observable varies least with the scales. This is typically of the order of the energy scale characterising the process (e.g. $\mu_R = m_{\mathrm{Z}}$ for Z-pole measurements) to avoid large logarithmic terms in the perturbative expansion [106]. The theoretical uncertainty originating from the series truncation can then be modelled by varying the scales about their nominal value.

Order in perturbation theory

The hard-scatter is defined by a high momentum transfer, Q^2, and as a result the strong coupling constant, $\alpha_s(Q^2)$, is small. The perturbative expansion of the matrix element, $|\mathcal{M}_{a,b\to f}|^2(\Phi_n; \mu_F, \mu_R, \vec{\alpha})$,[4] can therefore be performed in powers of α_s, which is done up to some finite order in the calculation. This can be pictured in terms of Feynman diagrams, where going to higher orders equates to including diagrams with additional internal loops. The leading order (LO) calculation corresponds to including only the first term in the expansion. The subsequent orders are referred to as next-to-leading order (NLO), next-to-next-to-leading order (NNLO) and so on. For example, the inclusive ggH cross section has been calculated at N^{3}LO [107].

As described above, the neglection of higher-order terms introduces some intrinsic uncertainty into the calculation. We model this uncertainty by varying the renormalisation and factorisation scales, μ_R and μ_F, by a factor of two around the nominal scale values. Ultimately, the more terms that are included in the perturbative expansion (going

[3] $\alpha_s = g_s^2/4\pi$, where g_s is the coupling strength of the strong interaction, introduced in Section 2.1.3.
[4] Note the dependence on both the factorisation and renormalisation scales: μ_F and μ_R.

to a higher order), the less the predictions depend on the scales, or in other words, the accuracy of the predictions increases.

Parton distribution function

Parton distribution functions (PDFs), $f_i(x, \mu_F)$, model the number density of parton type, $i = \{g, u, d, s, c, (b),$+anti-quarks$\}$, with momentum fraction, x, inside the colliding proton. These functions effectively absorb the long-distance physics of the internal proton, and therefore acquire a dependence on the factorisation scale, μ_F. The exact forms of the PDFs are extracted using data from lepton-hadron and hadron-hadron collisions, and are defined in so-called PDF sets to be used by the MC generators [108]. The evolution of the PDF with increasing μ_F is modelled with the DGLAP equations [109–111]. Ultimately, the choice of PDF affects not only the simulation of the hard-scattering, but also the parton showering and the modelling of multiple parton interactions in the underlying event. This influences both the predicted cross sections and the event kinematics.

Parton shower and hadronisation

Parton showering describes the evolution in momentum transfer from the hard-scattering process at high scales, to the formation of the final-state hadrons at low scales, typically $\mathcal{O}(1$ GeV). The evolution is performed by modelling the QCD radiation of initial-state and final-state coloured partons using perturbative QCD [106]. The radiated gluons can go on to radiate further, producing additional gluons and quark-antiquark pairs, which are typically soft or collinear with the outgoing partons. In this manner, the parton shower effectively *dresses* the event with additional QCD radiation, which is iterated until the momentum transfer reaches the scales associated with hadron formation. Free parameters of the parton showering model are tuned to improve the description of real data, where different values of these parameters are referred to as different parton-shower tunes [112, 113].

Hadronisation occurs at the end of the parton shower, and describes the confinement of the outgoing coloured partons into colourless hadrons. At this energy scale, QCD becomes strongly interacting and perturbation theory breaks down. Dedicated hadronisation models have been developed to deal with this non-perturbative regime, such as the string fragmentation and cluster models [114–116]. Fortunately, the hadronisation process can be effectively decoupled from the upstream hard-scatter process, and therefore once tuned, the hadronisation model can be applied successfully to different physics processes at different energies [106].

Jets

The final-state objects in hadron collisions typically include collimated sprays of particles, referred to as jets. These originate from the parton showering and hadronisation of both final-state partons and the QCD radiation of initial-state partons. Jets are not fundamental objects, and as a consequence they must be defined according to some jet algorithm [56]. These algorithms associate which particles belong to which jet according to some set of rules, and then reconstructs their total momentum. One such example used in this thesis is the anti-k_T algorithm, which repeatedly combines pairs of particles (i and j) according to the distance measure [69],

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2}, \tag{3.9}$$

where $p_{T,i}$ and $p_{T,j}$ are the transverse momentum of particles i and j, ΔR_{ij} is the angular separation between the particles, and R is the distance parameter for the algorithm, typically set to 0.4. Particles are combined into a single new particle (the jet) if,

$$d_{ij} < d_{iB}; \qquad d_{iB} = p_{T,i}^{-2}. \tag{3.10}$$

If this condition is not met, the particle i is declared as a final-state jet, and is removed from the list of particles for the subsequent jet-finding.

Jets can originate from both the hard-scatter matrix element and the parton shower, where the former is more accurate for modelling harder (high momentum) jets and the latter for softer (low momentum) jets. When a hard-process event is showered a double-counting can occur, leading to an overestimation of the jet multiplicity in the event. This is avoided by applying a jet matching or merging procedure [117]. These typically introduce some cut-off merging scale that defines whether the jets originating from the matrix element or the parton shower are used.

Underlying event

Secondary QCD interactions occur in each hadron collision in addition to the hard-scattering process, including beam remnant interactions, multiple parton interactions, and pileup. The additional activity which is not associated with the hard interaction is referred to as the underlying event [106]. These interactions typically involve coloured particles, which can "talk" to the particles in the hard-scattering. As a result, the underlying event can influence the properties of the process of interest. Like the parton shower, the free parameters of the underlying event model are tuned to match what is observed in data.

Detector simulation

The particle-level events are propagated through a simulation of the CMS detector. This is done using the GEANT4 package [118] which models the geometry of each subdetector, and describes how particles traverse through the CMS solenoidal field and subsequently interact with the detector material. The outputs are hits (deposits of energy) in the various subdetectors, which have been designed to match the true electronic read-out channels of the detector. Consequently, we can apply the same event reconstruction techniques on the MC simulation as to what is used on data.

To accurately model the detector response in terms of the efficiency and reconstruction performance, the detector simulation must be calibrated. This typically involves using a set of measurements e.g. the jet momentum distribution, and tuning the simulation to match what is observed in data. A mis-modelling of the detector response introduces (experimental) systematic uncertainties into the analysis. For example, in the context of the $H \to \gamma\gamma$ analysis, the mis-modelling of the photon energy resolution will affect the predicted diphoton invariant mass ($m_{\gamma\gamma}$) distribution in signal events, and consequently the extracted results.

It is worth introducing some terminology at this stage. Throughout this thesis, *truth-level* quantities correspond to the event properties in the simulation before the detector modelling i.e. their true values. These are the quantities which are used in the STXS event classification into the different bins. On the other hand, *reconstruction-level* quantities refer to the event properties reconstructed from the signals in the detector. For example, the Higgs boson transverse momentum, p_T^H, is a *truth-level* quantity, whilst the transverse momentum of the reconstructed photon pair, $p_T^{\gamma\gamma}$, is the equivalent *reconstruction-level* quantity. The difference between the two is purely a result of the detector efficiency and resolution. As discussed at the beginning of this section, the ultimate goal of the experiment is to infer knowledge about the parameters of the model Lagrangian. In the context of STXS measurements, this requires *unfolding* the response of the detector, or in other words using the reconstruction-level information to infer the truth-level information.

Generators

A number of tools are used to perform the different stages of MC event generation. In this thesis, the MG5_aMC@NLO [119], POWHEG [120–125] and SHERPA [126] generators are used to simulate the parton-level event. This includes the computation of the hard-scattering matrix element to some fixed order in the perturbative expansion. The MG5_aMC@NLO and POWHEG generators are interfaced with PYTHIA8 [127] for parton showering and hadronisation. Finally, the GEANT4 and DELPHES packages are used to perform detailed and fast-simulations of the CMS detector respectively.

Chapter 4

Measuring Higgs boson properties in the H $\rightarrow \gamma\gamma$ decay channel

4.1 Introduction

The following three chapters provide a detailed description of the CMS H $\rightarrow \gamma\gamma$ analysis documented in Ref. [1]. This analysis measures Higgs boson production cross sections and coupling-modifiers in the diphoton decay channel, using p-p collision data at $\sqrt{s} = 13$ TeV collected by the CMS experiment at the LHC from 2016 to 2018, corresponding to an integrated luminosity of 137 fb^{-1}. The increased statistical power which comes with using the full Run 2 data, leads to a reduction in the uncertainties of previous H $\rightarrow \gamma\gamma$ measurements, and enables the Higgs boson production phase space to be probed more finely.

Building upon the strategies developed in previous CMS H $\rightarrow \gamma\gamma$ analyses [66, 128, 129], a set of orthogonal event-categories are constructed to target kinematic regions (bins) of the STXS framework (see Figure 2.11). The analysis categories, or so-called "tags", are defined at *reconstruction-level* by placing a set of selection criteria on the *reconstructed* objects in an event. These selection criteria are chosen to closely align with the *truth-level* STXS bin boundary definitions. For example, there are two "tags" in the analysis which target events from the qqH VH-like STXS bin. This bin is defined at truth-level (see Table 2.6) by a dijet system with invariant mass in the window $60 < m_{jj} < 120$ GeV. The corresponding tags require at least two reconstructed-jets in the event, with a reconstructed dijet-invariant-mass value in the same window. Moreover, machine-learning (ML) algorithms are trained using reconstructed event quantities, to further isolate events from a given Higgs boson production-mode, and reject SM background processes.

In this manner, a total of 80 reconstruction-level analysis categories (tags) are de-

fined, which are each enriched in events from a given truth-level STXS bin (or group of bins). Similar to a fiducial-style analysis where the experimental selection is defined to closely match a pre-defined fiducial phase-space, we *unfold* the reconstruction-level tags back to the truth-level STXS bins. In other words, we undo the detector efficiency and experimental acceptance effects to measure the truth-level STXS bin cross sections. This *unfolding* procedure requires estimating the composition of each tag in terms of the STXS bins (using MC simulation), and is performed directly by a maximum-likelihood fit to data (see Chapter 5).

The fitted observable used to extract the truth-level cross sections is the diphoton invariant mass. This extremely powerful observable effectively distinguishes signal from background, where photon pairs produced via Higgs boson decay form a narrow signal peak centred around m_{H}, on top of a smoothly-falling background distribution from other SM processes. For events with two reconstructed photons, γ_1 and γ_2, the diphoton invariant mass, $m_{\gamma\gamma}$, is defined as,

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

where $E_{\gamma 1}$ and $E_{\gamma 2}$ are the measured energies of γ_1 and γ_2, respectively, and θ is the opening angle between the two photons. As displayed graphically in Figure 4.1, it is possible to isolate three aspects of the $m_{\gamma\gamma}$ spectrum which affect the sensitivity to Higgs boson properties:

1. The diphoton mass resolution: a narrower signal peak is more easily distinguished from the background continuum and therefore results in an enhanced sensitivity. The diphoton mass resolution is driven by the energy response of the ECAL (see Section 3.3.3). A good mass resolution also relies on the correct identification of photons from other objects in the detector such as jets. Furthermore, the precise assertion of the interaction vertex from which the two photons originate is crucial for accurately determining the opening angle, θ.

2. The signal-to-background ratio in an analysis category: the sensitivity is improved by reducing the background contamination under the Higgs boson peak. In this analysis, ML algorithms are used to better discriminate between signal and background. Since different background processes are important for different Higgs boson production modes, a number of signal-vs-background classifiers are trained and used in the relevant analysis categories.

3. The purity of the signal events in an analysis category: here, analysis categories are defined to target specific bins of the STXS framework. Therefore to improve the

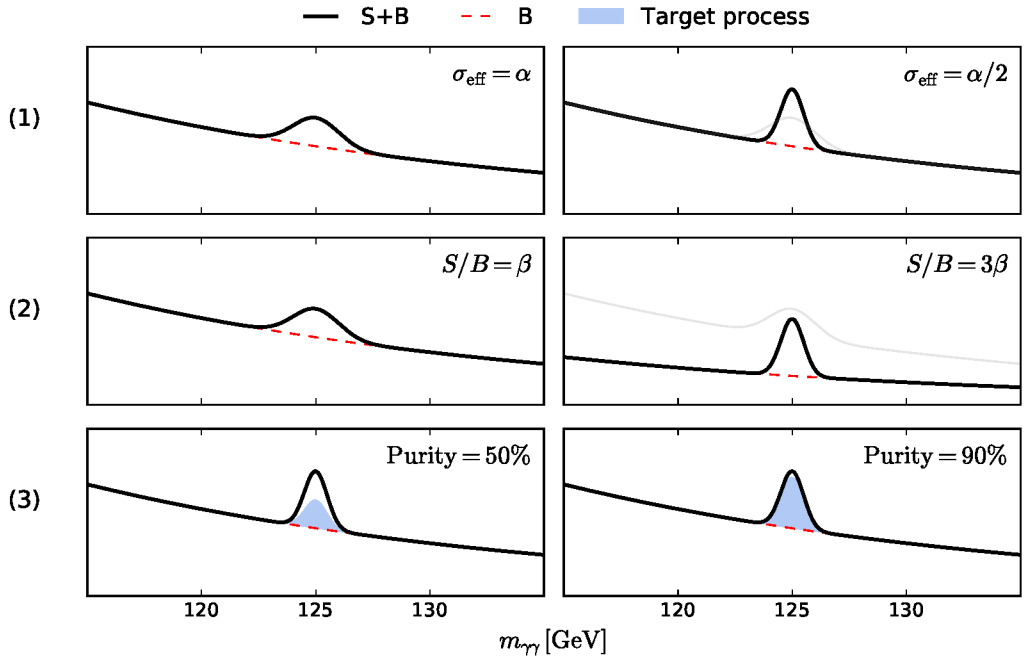


Figure 4.1: An illustration of the various means of improving the measurements of Higgs boson properties in the diphoton decay channel. The top row demonstrates the effect of improving the diphoton mass resolution, leading to a narrower signal peak. The middle row shows the effect of improving the signal-to-background ratio, which would result from using a more powerful signal-vs-background discriminator. Finally, the bottom row represents the improvement from increasing the target signal process purity.

sensitivity, we aim to maximise the purity of an analysis category with respect to the targeted STXS bin (or bins). In other words, make the confusion matrix which quantifies the fraction of each truth-level STXS bin in each reconstruction-level event category as diagonal as possible (see Figure 4.20). This helps the unfolding procedure, and reduces the correlations between the measured cross sections.

This chapter will focus on the experimental techniques used to maximise the sensitivity of the analysis, according to the three aspects stated above. Several applications of ML algorithms are introduced, to perform both regression and classification tasks. Appendix B provides an introduction to the foundations of ML, and will help the reader familiarise themselves with a number of recurrent concepts. Firstly, the data and MC simulation samples used in the analysis are introduced. Section 4.3 then details the procedure used to reconstruct candidate $H \to \gamma\gamma$ events, specifically the techniques which improve the diphoton mass resolution: the additional photon energy corrections, the photon identification and selection, and the primary vertex selection. Following this, the reconstruction-level categorisation of events is described in Section 4.4. Additional objects

in the events such as jets and charged leptons, and quantities such as the missing transverse momentum are used to define categories targeting the different truth-level STXS bins. This section covers the methods used to both reduce background contamination and to increase the purity of the targeted STXS bin (or bins) in such categories[1].

4.2 Samples

4.2.1 Data

This analysis uses p-p collision data collected by the CMS experiment at $\sqrt{s} = 13$ TeV during Run 2 of the LHC. The total integrated luminosity is 137 fb^{-1}, of which 35.9 fb^{-1} was collected in 2016, 41.5 fb^{-1} in 2017 and 59.4 fb^{-1} in 2018.

Events in data are selected using the two-tiered trigger system described in Section 3.3.6. The tag-and-probe method is used to evaluate the efficiency of the trigger selection [130]. This method exploits the decay of a known resonance such as the Z boson, where the *tag* is defined as one of the decay products passing very tight identification criteria, and the *probe* as the other decay product, subject to much looser identification requirements. Moreover, the combined invariant mass of the tag-and-probe pair is required to be consistent with the mass of the original resonance to ensure a high purity sample. For some selection criteria, $\mathcal{C}$, the efficiency, $\epsilon_{\mathcal{C}}$, is then defined as the fraction of probes passing $\mathcal{C}$. This method remains valid as long as the identification requirements on the probe do not affect the efficiency of $\mathcal{C}$.

Given the proximity of the Z boson and Higgs boson masses, as well as the fact that both electrons and photons are reconstructed as SCs in the ECAL, Z $\rightarrow e^+e^-$ events in data provide an excellent candidate for evaluating efficiencies in the H $\rightarrow \gamma\gamma$ analysis. Dielectrons with invariant mass close to the Z boson mass are used to define the tag and probe. After reweighting the Z $\rightarrow e^+e^-$ events to match the η and R_9 (see Section 4.3.1) distributions of H $\rightarrow \gamma\gamma$ events, the trigger efficiency is evaluated per SC in bins of the probe electron p_T, η and R_9. For an SC in the EB and EE, the average trigger efficiencies are above 97% and 95%, respectively. The product of the two per-SC efficiencies of the diphoton is then used to weight simulated events to replicate the trigger efficiency observed in data. Ultimately, the trigger selection criteria are significantly looser than the offline selection criteria applied in the H $\rightarrow \gamma\gamma$ analysis, ensuring that the trigger has a negligible effect on the H $\rightarrow \gamma\gamma$ selection efficiency.

[1]For clarity, throughout the following three chapters the terms "analysis category" or "tag" refer to the *reconstruction-level* event categories that are constructed for this analysis. On the other hand, the term "STXS bin" refers to a kinematic region of phase-space defined at *truth-level*. It is the STXS bin cross sections that the analysis aims to measure.

4.2.2 Simulation

Monte Carlo (MC) simulated events are used for both training event classifiers and constructing the final signal model. The simulated events are subject to the exact same event reconstruction and categorisation procedure as used for data.

Signal samples were simulated for the different Higgs boson production mechanisms at next-to-leading order (NLO) in perturbative QCD using the MG5_aMC@NLO (version 2.4.2) [119], and Powheg (version 2.0) generators [120–125]. When possible, an independent event sample from the alternative generator is used for training event classifiers, thus ensuring the event categorisation and the construction of the final signal model are independent. Events produced via ggH production are weighted as a function of the Higgs boson p_T and the number of jets to match the predictions of the NNLOPS program [131]. Parton distribution functions (PDFs) used to model the distribution of colliding partons inside the initial-state protons are taken from the NNPDF 3.0 (NNPDF 3.1) set [132, 133], when simulating 2016 (2017/2018) data. The parton-level events are subsequently interfaced with Pythia8 (version 8.226 for 2016 MC and version 8.230 for 2017/2018 MC) for decaying the Higgs boson to photons, parton showering and hadronisation [127]. The Pythia8 CUETP8M1 [112] and CP5 [113] tunes are used for the simulation of 2016 data and 2017/2018 data, respectively.

The signal samples are normalised according to the production cross sections and the $H \to \gamma\gamma$ branching fraction (0.227%) recommendations by the LHCHWG [42]. The fractional breakdown of each production mode in the STXS bins is computed directly from the signal MC samples, and serves as the SM prediction of the cross section in each bin. Table 4.1 summarises the event generators used for the signal simulation, as well as the total cross section times branching fraction, $\sigma_{\mathrm{SM}}\mathcal{B}$, for each production mode with details on the order of the calculation. The fractional breakdowns of each production mode into the STXS bins are shown in Chapter 2 in Tables 2.5–2.8.

The final background model used for the extraction of results is derived directly from data. Nevertheless, simulated background events are required for training the multivariate event classifiers. For inclusive production, the dominant source of background is SM diphoton production, which is simulated using the Sherpa (version 2.2.4) generator [126]. In this sample, matrix elements are calculated at NLO and LO for up to one and three additional partons, respectively, which are subsequently matched with the Sherpa generator parton showering. A subdominant background originates from γ+jet or jet+jet events, where the jets are misidentified by the PF algorithm as isolated photons. These backgrounds are simulated with Pythia8, applying a filter in the generation to enrich the production of jets with high electromagnetic activity. Furthermore, other sources of background become important for categories targeting the sub-dominant production modes,

Table 4.1: Details of the signal simulation. For each production mode, the generator used for the final signal-modelling is listed. If available, an independent sample is used from the alternative generator when training the event classifiers. In addition, the cross sections times branching fraction, $\sigma_{\mathrm{SM}}\mathcal{B}$, are provided for a nominal Higgs boson mass, $m_{\mathrm{H}} = 125.0$ GeV, at $\sqrt{s} = 13$ TeV. The final column details the order of the cross section calculation. For the tHq, tHW and bbH production modes, the flavour scheme (FS) used in the calculation is specified, where 5FS (4FS) includes (does not include) the bottom quark/anti-quark components in the colliding protons.

Production mechanism	Naming convention	Generator	$\sigma_{\mathrm{SM}}\mathcal{B}$ [fb]	Order of σ_{SM} calc
ggH	ggH	MG5_aMC@NLO	110.27	N^{3}LO(QCD)+NLO(EW)
gg→ZH, Z→qq	ggZH had	Powheg	0.19	NNLO(QCD)+NLO(EW)
VBF	VBF	MG5_aMC@NLO	8.59	NNLO(QCD)+NLO(EW)
qq/qg→WH, W→qq	WH had	MG5_aMC@NLO	2.10	NNLO(QCD)+NLO(EW)
qq/qg→ZH, Z→qq	ZH had	MG5_aMC@NLO	1.40	NNLO(QCD)+NLO(EW)
qq/qg→WH, W→ $\ell\nu$	WH lep	MG5_aMC@NLO	1.016	NNLO(QCD)+NLO(EW)
qq/qg→ZH, Z→ $\ell\ell/\nu\nu$	ZH lep	MG5_aMC@NLO	0.520	NNLO(QCD)+NLO(EW)
gg→ZH, Z→ $\ell\ell/\nu\nu$	ggZH lep	Powheg	0.084	NNLO(QCD)+NLO(EW)
ttH	ttH	MG5_aMC@NLO	1.155	NLO(QCD)+NLO(EW)
tHq	tHq	MG5_aMC@NLO	0.175	NLO(QCD) in 5FS
tHW	tHW	MG5_aMC@NLO	0.034	NLO(QCD) in 5FS
bbH	bbH	MG5_aMC@NLO	1.108	NNLO(5FS)+NLO(4FS)

such as diboson production in the VH leptonic categories and tt+$\gamma\gamma$ in the top-associated categories. Additional MC samples simulated with the MG5_aMC@NLO and Powheg generators are used to model such backgrounds. Finally, Drell-Yan events with leptonic final-states and tt+Z events are used for validation purposes. Both are simulated with the MG5_aMC@NLO generator and Pythia8 for parton showering and hadronisation.

Each particle-level sample is propagated through the Geant4 package to model the response of the CMS detector [118]. Separate MC samples are produced for each year to account for the variations in the detector conditions and the LHC beam parameters. This modelling includes the effect of pileup interactions originating from both the nominal bunch-crossing (in-time pileup) and the crossing of previous and subsequent proton bunches (out-of-time pileup). The simulation is weighted to match the distribution of the number of interaction vertices in data, which corresponds to an average pileup of 23 in 2016, and 32 in 2017 and 2018.

4.3 Event reconstruction

This section describes the offline reconstruction of events passing the trigger selection. The data and the corresponding simulation are reconstructed separately for each year to account for the differences in the detector performance and LHC beam parameters.

4.3.1 Photon reconstruction

Photons are defined using the set of photon candidates from the PF algorithm (see Section 3.4). In the algorithm, SCs are formed by clustering together deposits of energy in the ECAL crystals, consistent with originating from the same electromagnetic shower. Due to imperfect shower containment in the crystals and shower losses for photons which convert to e^+e^- pairs before reaching the ECAL, the SC energy, E_{SC}, can often differ from the true initial photon energy, E_{true}. As described in Section 3.3.3, a multivariate regression technique is applied to all photons to correct for such losses, estimating both the value of E_{true} and its uncertainty. Remaining differences in the photon-energy scale and resolution between data and simulation are accounted for using a series of additional scale and smearing corrections, derived using $Z \to e^+e^-$ events. Photons are then subject to a set of selection criteria (pre-selection) concerning the photon shower-shape, kinematic, and isolation variables. In addition, a ML algorithm known as a Boosted Decision Tree (BDT, see Appendix B) is trained to separate genuine photons from fake photons; this is referred to as the photon-identification BDT. The selection includes a minimum requirement on the output score of this BDT, which reduces the contribution from background processes with hadronic jets mimicking a photon signature. Before describing the photon reconstruction techniques in more detail, it is useful to list the photon variables used in the $H \to \gamma\gamma$ analysis in Table 4.2. Two of these variables, namely R_9 and $\mathcal{I}_{\text{ph}}$, are shown in Figure 4.3 when describing the photon-identification BDT. The distributions correspond to probe electrons from $Z \to e^+e^-$ decays, however the photon distributions are effectively the same since photons and electrons have almost identical shower profiles in the ECAL.

Photon energy

After applying the photon energy regression described in Section 3.3.3, a number of residual discrepancies between data and simulation remain that cannot be derived from simulation alone. Using $Z \to e^+e^-$ events in data, in which the electrons are reconstructed as photons[2], a series of scale and smearing corrections are derived to correct the photon-energy scale and resolution [49]. To account for the degradation of the ECAL crystal

[2]Meaning no track information is used in the reconstruction, and the energy is determined using the algorithm and corrections corresponding to photons rather than electrons.

Table 4.2: A summary of the photon variables used in this analysis. The shower-shape variables are used to both correct the photon energy in the regressor (see Section 3.3.3) and to discriminate between real and fake photons. The isolation variables help to identify real photons from other objects such as jets mimicking a photon signature.

Shower shape variables	
R_9	($=E_{3\times3}/E_{\mathrm{SC}}$) The ratio of the energy sum in the 3×3 grid surrounding the SC seed to the energy of the SC before corrections. The value of R_9 is typically high (> 0.85) for unconverted photons, and typically lower (< 0.85) for photons that have undergone a conversion upstream of the ECAL.
$E_{2\times2}/E_{5\times5}$	The ratio of the energy sum in the 2×2 grid containing the most energetic crystals in the SC, to the energy in the 5×5 grid surrounding the SC seed.
σ_η	A measure of the lateral extension of the shower, defined as the standard deviation of single crystal η values within the SC, weighted by the logarithm of the crystal energy.
$\sigma_{i\eta i\eta}$	The standard deviation of the shower in η in terms of the absolute number of crystal cells.
σ_ϕ	A measure of the lateral extension of the shower, defined as the standard deviation of single crystal ϕ values within the SC, weighted by the logarithm of the crystal energy.
$\mathrm{cov}_{i\eta i\phi}$	The covariance of the single crystal η and ϕ values for the 5×5 grid centred around the crystal with the most energy.
σ_{RR}	For photons in the ECAL endcaps only. The standard deviation of the shower spread in the x-y plane of the preshower detector.
n_{clusters}	The number of clusters in the SC.
Isolation variables	
H/E_{SC}	Ratio of the energy in the HCAL cells directly behind the SC to the energy of the SC.
$\mathcal{I}_{\mathrm{ph}}$	Photon isolation, defined as the sum of transverse energy of PF photons falling inside a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ around the SC. The transverse energies are corrected using ρ to mitigate the effect of pileup.
$\mathcal{I}_{\mathrm{ch}}$	Charged-hadron isolation, defined as the sum of transverse energy of the PF charged hadrons falling inside a cone of radius $\Delta R = 0.3$ around the SC. This is measured with respect to both the selected and the worst vertex; the benefit of this is that true photons are generally isolated from other vertices, but fake photons are not.
$\mathcal{I}_{\mathrm{tk}}$	Track isolation, defined as the sum of transverse energy of all tracks in a hollow cone with a smaller (larger) annulus of $\Delta R = 0.04$ ($\Delta R = 0.3$).
Other variables	
Electron veto	Boolean flag variable which is set to false if the supercluster is matched to an electron track.
ρ	Median energy density per unit area in the event (sensitive to pileup).
E_{SC}	Uncorrected supercluster energy.
E_{true}	True photon energy.
E_γ	Reconstructed photon energy.

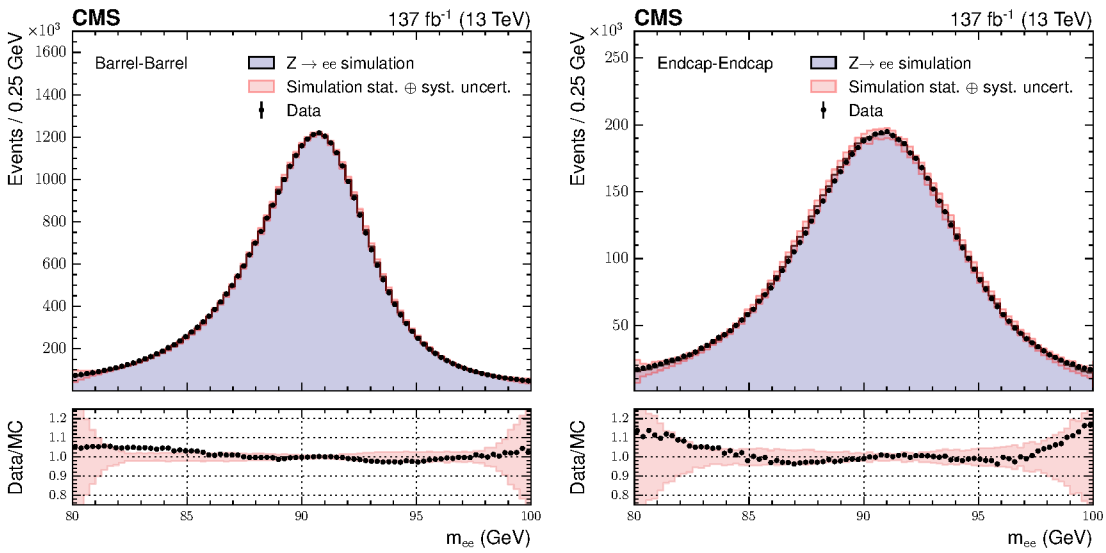


Figure 4.2: Comparison of the dielectron mass spectrum for $Z \to e^+e^-$ events in data (black points) and simulation (filled histogram), after the full set of energy scale and smearing corrections are applied. The total uncertainty in the simulated events is shown by the pink bands. The plots show the full data set collected during the 2016-2018 data-taking period, and the corresponding simulation, where the left-hand (right-hand) plot corresponds to events in which both electrons are reconstructed in the ECAL barrel (endcaps).

performance over time, and the subsequent drift in energy scale this causes, a time-dependent correction is applied in bins which equate roughly to the duration of one LHC fill. This ensures the energy scale of the ECAL is constant over time. Subsequently, the dielectron mass spectrum, m_{ee}, is used to simultaneously shift the peak in data to match that in simulation, and smear the resolution in simulation to match that in data. These scale and smearing corrections are applied differentially in bins of SC η and R_9.

Referring back to equation 4.1, all techniques[3] used to improve the photon-energy resolution, σ_E/E, in turn improve the diphoton mass resolution, $\sigma_m/m_{\gamma\gamma}$, and thus lead to an enhanced sensitivity to Higgs boson production. Figure 4.2 shows the m_{ee} distributions for $Z \to e^+e^-$ events in both data and simulation, after the full set of energy corrections are applied. The left-hand panel shows the situation where both electrons are reconstructed in the ECAL barrel, and the right-hand panel where both are reconstructed in the ECAL endcaps. The pink bands indicate the statistical and systematic uncertainty in the simulation, where the systematic component includes contributions which are related to the measurement and reconstruction of the photon energy (see Section 5.4). Clearly, after the corrections are applied, the data and simulation are in good agreement, within

[3]Including the energy corrections described in Section 3.3.3.

the uncertainty bands.

Photon pre-selection

Table 4.3 provides a schema of the selection criteria applied to the photon candidates. The pre-selection efficiency is calculated using the tag-and-probe method on $Z \rightarrow e^+e^-$ events for all selection criteria, barring the requirement on the electron veto, which is computed using $Z \rightarrow \mu^+\mu^-\gamma$ events. For the latter, the $\mu^+\mu^-\gamma$ system is required to be consistent with the decay of a Z boson, with a three-body invariant mass between 60 and 120 GeV. Probes are then defined as the set of photons passing all pre-selection criteria except the electron veto, and the efficiency is simply calculated as the fraction of these probes which pass the veto. The total pre-selection efficiency is computed for both data and simulation in bins of photon η and R_9, and is typically above 95% for photons with high values of R_9 (> 0.85) and around 90% for photons with lower R_9 (< 0.85). Weights are then applied differentially per photon to match the pre-selection efficiencies in simulation to those observed in data.

Photon identification

The photon-identification (photon-ID) BDT aims to distinguish between real photons in the CMS detector and hadronic jets mimicking a photon signature. The BDT is trained using the γ+jet simulation sample, where the true photon is used as signal and the fake photon from the jet as background. Photon shower-shape, kinematic, and isolation variables are used as input features to the BDT, along with parameters sensitive to pileup, such as the median-energy-density per unit area, ρ.

One of the dominant sources of systematic uncertainty in the H $\rightarrow \gamma\gamma$ analysis arises from the modelling of the electromagnetic shower in simulation, in particular the variables describing the shower-shape and isolation. Since these variables are direct inputs to the photon-ID BDT, any discrepancies between data and simulation are propagated to the output BDT score, and thus introduce a systematic uncertainty into the analysis. To mitigate this, a chained quantile regression (CQR) method [134] is applied, which sequentially corrects the photon-ID BDT input variables in simulation.

The corrections are derived using a set of probe electrons in $Z \rightarrow e^+e^-$ events. The CQR method morphs the distribution of some photon-ID BDT input variable, y_i (e.g. R_9), in simulation to match that in data by training a series of 21 BDTs to predict points along the cumulative distribution function, CDF(y_i): [0.01,0.05,0.10,...,0.95,0.99]. For example, the CQR BDTs predict the values of R_9 at which the CDF(R_9) is equal to the 21 points ranging from 0.01 to 0.99. The full CDF is determined by linearly interpolating between these 21 points, and is extracted separately for both data and simulation. A

Table 4.3: Schema of the photon pre-selection criteria. The shower-shape and isolation requirements are different for photons in the ECAL barrel and for photons in the ECAL endcaps. These are then split into regions of different R_9 criteria, with varying levels of additional selection on $\sigma_{i\eta i\eta}$, $\mathcal{I}_{\text{ph}}$ and $\mathcal{I}_{\text{tk}}$.

Minimum p_T	$p_T^{\gamma 1} > 35$ GeV (leading), $p_T^{\gamma 2} > 25$ GeV (subleading)				
↓					
Geometrical acceptance	$\lvert\eta\rvert < 2.5$, excluding barrel-endcap transition region $1.44 < \lvert\eta\rvert < 1.57$				
↓					
Electron veto	True				
↓					
Hadronic shower rejection	$\text{H}/E_{\text{SC}} < 0.08$				
↓					
Shower shape and isolation requirements	For photons in the EB:	R_9	$\sigma_{i\eta i\eta}$	$\mathcal{I}_{\text{ph}}$ (GeV)	$\mathcal{I}_{\text{tk}}$ (GeV)
		> 0.85	-	-	-
		$[0.50, 0.85]$	< 0.015	< 4.0	< 6.0
	For photons in the EE:	R_9	$\sigma_{i\eta i\eta}$	$\mathcal{I}_{\text{ph}}$ (GeV)	$\mathcal{I}_{\text{tk}}$ (GeV)
		> 0.90	-	-	-
		$[0.80, 0.90]$	< 0.035	< 4.0	< 6.0
	All photons: at least one of $R_9 > 0.8$, $\mathcal{I}_{\text{ch}} < 20$ GeV, and $\mathcal{I}_{\text{ch}}/p_T^{\gamma} < 0.3$				

correction is then applied to the simulated variable, per electron, according to:

$$y_i \longrightarrow y_i^{\text{corr}} = \text{CDF}_{\text{data}}^{-1}(\text{CDF}_{\text{simulation}}(y_i)), \tag{4.2}$$

which successfully morphs the distribution of y_i in simulation to match that in data.

Shower-shape variables are first ordered into a chain: $[S_1,S_2,...S_N]$. The CDF for the first variable, S_1, is predicted using solely the electron p_T, η, ϕ and global energy density, ρ, as input features to the 21 BDTs. For variable, S_i, the input features also include the previously processed shower-shape variables: $[S_1^{\text{corr}},...,S_{i-1}^{\text{corr}}]$ for simulation and $[S_1,...,S_{i-1}]$ for data. By deriving the corrections in this manner, correlations between the photon-ID BDT input variables are accounted for. The ordering of the chain is optimised to minimise the final discrepancy between data and simulation in the photon-ID BDT output score. For example, the shower-shape variables are corrected in 2018 simulation according to the chain ordering,

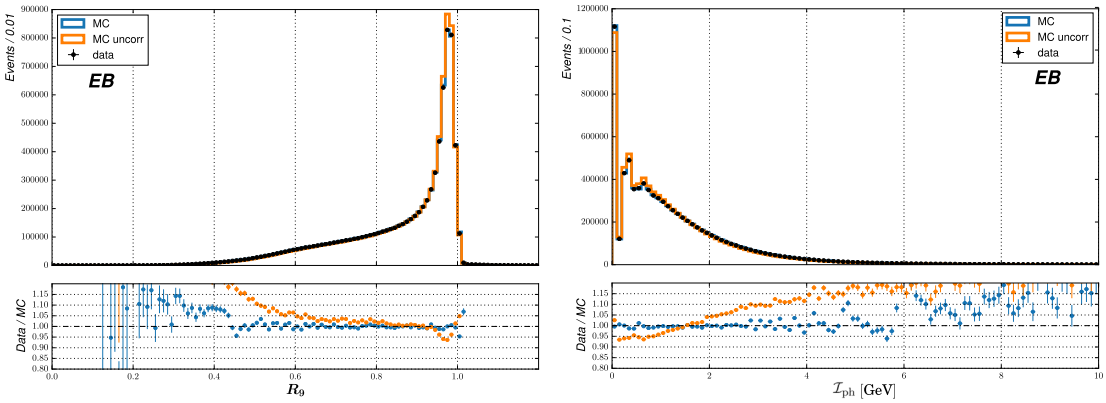


Figure 4.3: Probe electron R_9 (left) and $\mathcal{I}_{\mathrm{ph}}$ (right) distributions from $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ events, where the probe electron is reconstructed in the ECAL barrel. The black points correspond to data taken in 2017. The corresponding simulation is shown without corrections in orange, and with the CQR corrections in blue. The bottom panel shows the ratio of data to simulation, clearly demonstrating the improved agreement after applying the CQR corrections.

$$[\,\mathrm{cov}_{i\eta i\phi},\ E_{2\times2}/E_{5\times5},\ R_9,\ \sigma_\phi,\ \sigma_{i\eta i\eta},\ \sigma_\phi\,]. \tag{4.3}$$

For isolation variables, $\mathcal{I}_i$, an additional stochastic-shifting procedure is applied to account for electrons migrating across the discontinuity in the distributions i.e. from the peak at $\mathcal{I}_i = 0$, to the tail at $\mathcal{I}_i > 0$. Two additional classifiers are trained to predict the probabilities for electrons to fall in the peak or tail, given the electron p_T, η and ϕ, and the energy density, ρ. Using the output probabilities, the simulated electrons are then migrated between the peak and tail to match the relative compositions in data. Finally, the distribution of electrons in the tail is then corrected using the quantile morphing technique described above. Figure 4.3 demonstrates the performance of the CQR method for shower-shape variable, R_9 (left) and isolation variable, $\mathcal{I}_{\mathrm{ph}}$ (right).

The systematic uncertainty originating from these corrections is derived by splitting the original $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ samples in half, and re-calculating the corrections using the two independent event sets. The magnitude of the uncertainty is defined per-bin of the photon-ID BDT output score, as the standard deviation of the event-by-event corrected output score values. This assumes the major source of uncertainty in the method is originating from the limited size of the training samples. All in all, the CQR method provides a vastly improved technique for calculating the shower-shape and isolation corrections, reducing the impact of a dominant systematic uncertainty from around 5% in previous H → γγ analyses to roughly 2%.

The output score of the photon-ID BDT, for the lowest-scoring photon in the diphoton

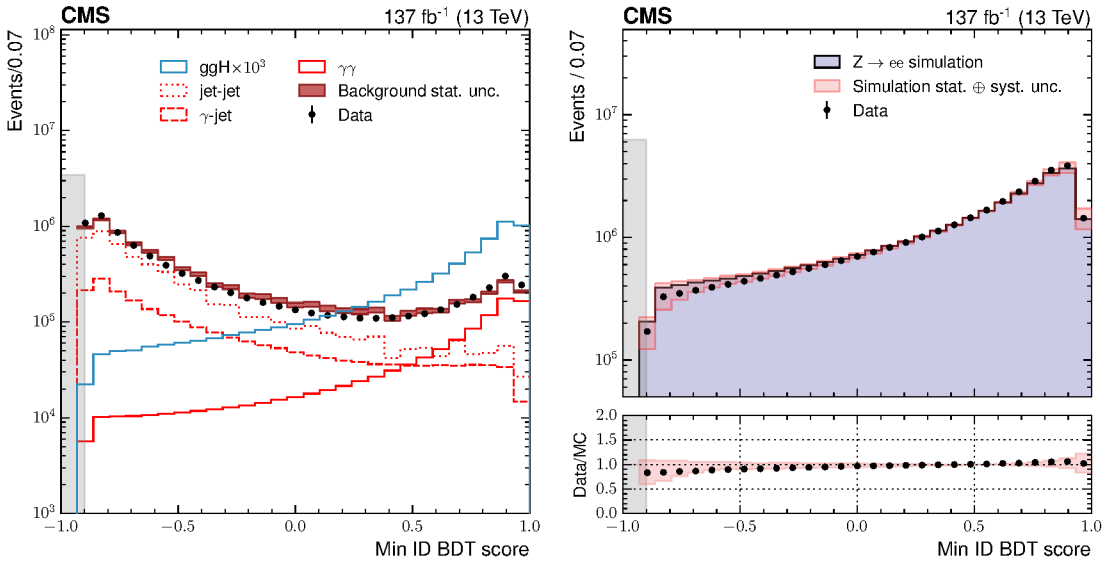


Figure 4.4: The left plot shows the distribution of the photon-ID BDT score of the lowest scoring photon in diphoton pairs with an invariant mass in the range $100 < m_{\gamma\gamma} < 180$ GeV. Data events passing pre-selection are shown by the black points. Simulated background events with the corresponding statistical uncertainty are shown by the pink band. The different components of the background and the Higgs boson signal events are also plotted, shown by the red and blue histograms respectively. The right plot shows the same distribution for $Z \to e^+e^-$ events in data and simulation, where the electrons are reconstructed as photons. The total uncertainty in the simulation, calculated as the quadrature sum of the statistical and systematic components, is shown by the pink band. The shaded grey regions demonstrate the minimum threshold applied for photons in this analysis. The full data set collected in the period 2016-2018 and the corresponding simulation is shown.

pair, is shown in Figure 4.4. The left-hand plot shows the distribution for signal and background processes, highlighting the impressive discriminating power between events with two real photons and the γ+jet and jet+jet backgrounds. The right-hand plot shows the same distribution for $Z \to e^+e^-$ events in both data and simulation. The two are in excellent agreement, within the calculated uncertainties. For completeness, Figure 4.5 shows the photon-ID BDT score for probe electrons ($Z \to e^+e^-$) reconstructed in the EB, with and without the CQR corrections applied to the BDT input features. Clearly, the improved agreement between data and simulation in the input features propagates to the output score, thus demonstrating the power of the CQR method.

Each photon in an event is required to have a photon-ID BDT score > -0.9. Events which do not satisfy this criteria have already been removed from the plots in Figure 4.4, hence why the γ+jet and jet+jet distributions do not peak at -1. This selection criteria reduces the contribution from the γ+jet and jet+jet background processes by over 70%, whilst maintaining at least 99% of signal events. The cut is purposely chosen to be

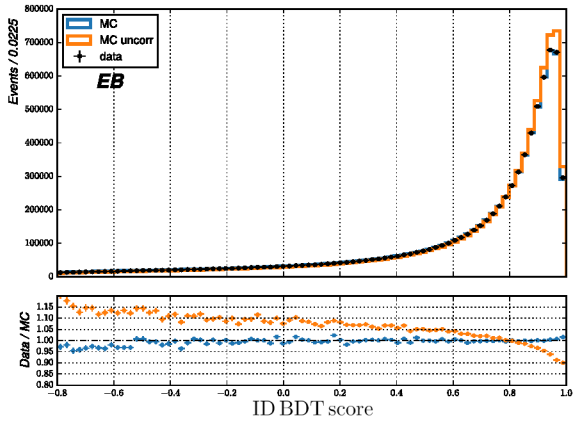


Figure 4.5: Photon-ID BDT score distributions for probe electrons in $Z \to e^+e^-$ events, where the probe electrons are reconstructed in the ECAL barrel. The black points correspond to data taken in 2017. The corresponding simulation is shown without corrections in orange, and with the CQR corrections in blue. The bottom panel shows the ratio of data to simulation. The improved agreement from applying the CQR corrections is shown to propagate to the photon-ID BDT output score.

loose in order to maintain reasonable statistics in the background simulation for training the event categorisation discriminants. Akin to the pre-selection criteria, the efficiency of this requirement is derived in simulation and data using the tag-and-probe method on $Z \to e^+e^-$ events, and the events in simulation are scaled to match the efficiency in data. In addition, the photon-ID BDT score is used in numerous ways in the event categorisation.

4.3.2 Vertex selection

Equation 4.1 shows that the diphoton mass resolution depends not only on the per-photon-energy resolutions, but also on the determination of the diphoton opening angle, θ. For this, it is necessary to know the precise location of the primary hard-scattering vertex from which the photons originate. If the location is correctly determined within 1 cm of the true vertex position, the diphoton mass resolution is dominated by the photon-energy resolution. For events with additional objects such as charged leptons and jets, it is relatively easy to assign the primary vertex, due to the presence of distinct charged-particle tracks. On the other hand, for $H \to \gamma\gamma$ events with no additional objects (e.g. ggH 0J), it becomes a more difficult problem. In the $H \to \gamma\gamma$ analysis, the diphoton vertex is assigned from a collection of candidate primary vertices using a dedicated BDT, which is trained using ggH $H \to \gamma\gamma$ simulated events and takes as input features related to the tracks recoiling against the diphoton system:

- $\sum_i |\vec{p}_T^{\,i}|^2$,

- $-\sum_i(\vec{p}_T^{\,i}\cdot\vec{p}_T^{\,\gamma\gamma}/|\vec{p}_T^{\,\gamma\gamma}|)$,
- $(|\sum_i\vec{p}_T^{\,i}|-p_T^{\gamma\gamma})/(|\sum_i\vec{p}_T^{\,i}|+p_T^{\gamma\gamma})$,
- the total number of photon conversions in the tracker,
- the pull $(|z_{\text{vtx}}-z_{\text{e}}|/\sigma_z)$ between the longitudinal positions of the reconstructed vertex, z_{vtx}, and the vertex estimated using tracks from photon conversions, z_e, where σ_z is the uncertainty in z_{e}.

In these observable definitions, the sums run over all PF tracks associated with a given vertex, labelled by the index i. The quantity $\vec{p}_T^{\,\gamma\gamma}$ corresponds to the transverse momentum of the diphoton system, measured with respect to the same vertex. The final two BDT input variables in the list are only included for events which contain tracks originating from photon conversions.

The performance of the vertex assignment BDT is evaluated using $\text{Z}\to\mu^+\mu^-$ events. Here, the vertices are reconstructed after omitting the muon tracks, in order to imitate a diphoton-like system. The vertex is then said to be correctly assigned by the BDT if the location is within 1 cm of the true vertex position. Figure 4.6 (left) shows the efficiency of the vertex assignment as a function of the dimuon p_T, for both simulation and data, demonstrating an agreement within 2% across the whole p_T range. Correction factors are applied to simulation to match the fraction of correctly assigned vertices observed in data, keeping the total number of events constant. For inclusive $\text{H}\to\gamma\gamma$ events, the efficiency of correctly assigning the diphoton vertex to be within 1 cm of the true vertex is roughly 79%. This value rises for events with additional objects such as jets and charged leptons.

An additional vertex-related BDT is trained to estimate the *probability* that the vertex assignment is within 1 cm of the interaction point from which the diphoton originates. The input features are the total number of reconstructed vertices in an event, the relative positions and respective vertex assignment BDT scores for the three highest scoring vertices, $|p_T^{\gamma\gamma}|$, as well as the number of converted photons in the tracker. Akin to the vertex-assignment BDT, this vertex-probability BDT is trained using ggH, $\text{H}\to\gamma\gamma$ events. The right-hand plot of Figure 4.6 demonstrates the agreement between the output probability and the true vertex assignment efficiency in simulated events.

Finally, the width of the z distribution of reconstructed vertices, known as the beamspot width, is measured to be between 3.4–3.6 cm in data. A year-dependent correction is applied to simulated events to ensure the beamspot width matches that observed in data.

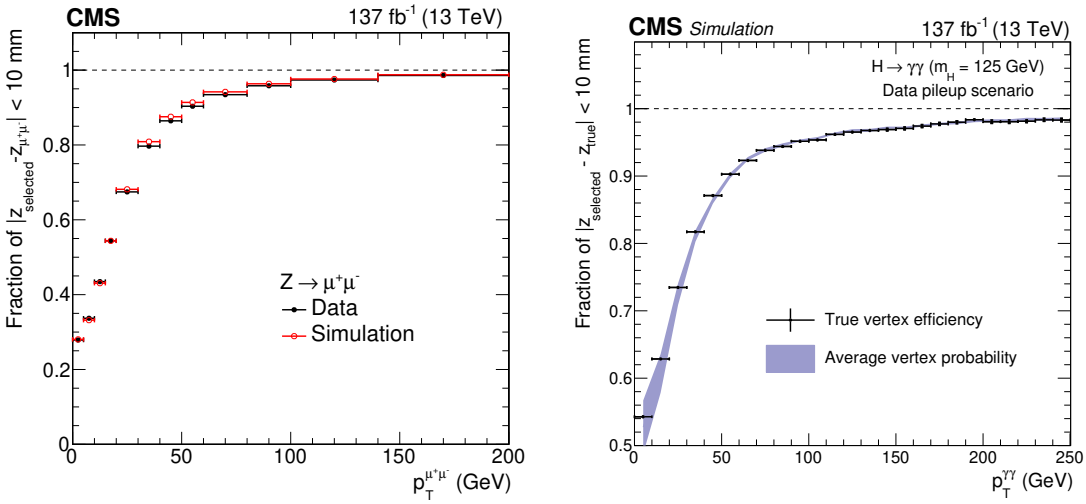


Figure 4.6: The left plot shows the fraction of $Z \to \mu^+\mu^-$ events where the assigned vertex is within 1 cm of the true vertex position as a function of $p_T^{\mu\mu}$, for simulated events in red and data events in black. Here, the muon tracks are omitted from the event reconstruction to mimic a $H \to \gamma\gamma$ system. Simulated events are weighted to match the pileup and beamspot width distributions observed in data. The right plot demonstrates that the average vertex-probability BDT score agrees with the true vertex-assignment efficiency in simulated events. The full data set collected in the period 2016-2018 and the corresponding simulation are shown in the plots.

4.3.3 Reconstruction of other objects

The analysis categories are defined to target events from different STXS bins. In order to define such categories it is necessary to place requirements on additional, reconstructed objects in the event, such as two forward-jets in VBF production or two-same flavour, oppositely-charged leptons in $Z(\ell\ell)H$ production. The following section briefly describes the reconstruction of these additional objects from the PF collections, described in Section 3.4.

Jets

The precise reconstruction of jets is particularly important in this analysis, not only for differentiating between different Higgs boson production modes, but also for splitting events according to the jet-related STXS boundaries: the number of jets, m_{jj} and p_T^{Hjj}. On average, a jet carries 65% of its total energy in charged hadrons, 25% in photons, and the final 10% in neutral hadrons. By providing separate collections of such final-state particles with independent measurements of their energies, the PF algorithm offers unprecedented performance in terms of the jet reconstruction at CMS.

Jets are built by clustering the reconstructed PF candidates using the infrared and

collinear-safe anti-k_T algorithm [68, 69], with a distance parameter of 0.4. A Charged-Hadron Subtraction (CHS) technique is applied to remove charged hadrons that are associated with vertices other than the primary vertex (as chosen by the vertex-assignment BDT) from the clustering procedure. This helps to mitigate the contribution to the jet momentum from pileup. The algorithm operates by defining a distance measure between PF candidates that depends on their p_T and angular parameters (see equation 3.9), and iteratively clusters the four-momentum vectors with the shortest distance measure. Once the shortest distance measure is between the cluster and the LHC beam, the procedure terminates and the clustered objects are defined as a jet. This is repeated until all PF candidates have been clustered. The jet momentum is then calculated as the vectorial sum of all the PF candidate momenta in the jet. From simulation, this is found to be, on average, within 5 to 10% of the truth-level jet momentum over the whole p_T spectrum and detector acceptance.

After building the jets, a set of corrections are applied to correct the jet-energy scale and resolution in both data and simulation [135]. Firstly, an offset correction is applied to account for any remaining contributions from pileup. This depends on the event energy-density, ρ, and the jet p_T^j, η^j and angular spread, ΔR_j. Following this, corrections are derived from simulation to (on average) match the reconstruction-level jet energy to the truth-level jet energy. These are calculated in bins of p_T^j and η^j to account for variations in the response across the detector. Any residual discrepancies in the jet-energy scale between data and simulation are rectified using in situ measurements of the momentum balance in dijet, γ+jet, Z+jet, and multijet events. The jet energy resolution after the corrections is typically 15–20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV, for jets in the central region of the detector [135].

Finally, a pileup-identification BDT is trained to reduce the number of pileup-induced jets entering the analysis [136]. Such jets typically contain PF candidates from multiple pileup collisions, and therefore tend to be more broad and diffuse than jets originating from the hard process. In addition, they usually contain tracks which are not associated with the primary vertex. To make use of these observations, the input features to the BDT are variables related to the jet shape, and additional track variables related to the interaction vertex. A p_T and η-dependent threshold is then placed on the output score of this BDT to reject pileup jets. On top of this, to be considered in the event selection, jets are required to have $p_T^j > 25$ GeV, fall within $|\eta^j| < 4.7$, and have an angular separation with respect to both photons, $\Delta R_{j,\gamma} > 0.4$.

b tagging

Jets originating from b quarks typically contain tracks associated with secondary vertices, which are displaced from the hard-scatter primary vertex by a few mm to 1 cm. This is a result of the non-negligible lifetime of B hadrons (1.5 ps), which traverse a measurable distance in the detector before decaying. The identification of such jets (b tagging) is crucial for isolating ttH and tH production events, since the top quarks almost always decay to b quarks. A deep neural network (DNN, see Refs. [137–139]) algorithm has been trained [140], which takes secondary vertices and PF candidates as input features, and outputs a probability that a given jet has originated from the decay of a B hadron. The per-jet output probabilities are used in the event categorisation, both for the pre-selection of ttH and tH events, and as input features to numerous ML classification algorithms.

Charged leptons

Tagging on additional charged leptons is crucial for isolating VH and top-associated Higgs boson production events, where at least one vector boson decays leptonically: $W \rightarrow \ell\nu$ and $Z \rightarrow \ell\ell$. Electrons and muons are taken from the PF electron and muon collections respectively, and additional isolation and identification requirements are placed on both [91, 141]. To be considered in the event categorisation, electrons are required to have $p_T^e > 10$ GeV and be within $|\eta^e| < 2.4$, excluding the barrel-endcap transition region, whilst muons must have $p_T^\mu > 5$ GeV and fall within $|\eta^\mu| < 2.4$.

Missing transverse momentum

The missing transverse momentum vector, $\vec{p}_T^{\,miss}$, is calculated as the negative vector p_T sum of all PF candidates in an event, and its magnitude is denoted as p_T^{miss} [142]. The reconstruction of $\vec{p}_T^{\,miss}$ accounts for the energy scale corrections applied to the reconstructed jets in the event. This is a useful quantity for selecting signal processes with additional neutrinos, such as WH (ZH) production where the vector boson decays into a neutrino and a lepton (two neutrinos).

4.4 Event categorisation

4.4.1 Overview

After the event reconstruction, a total of 80 event categories are constructed in which the $m_{\gamma\gamma}$ spectrum is fitted simultaneously to measure Higgs boson production cross sections and couplings. Each category imposes a set of selection criteria related to the features of both the reconstructed diphoton system and the additional objects in the event. The criteria of a given category are defined to maximise the sensitivity to events from a particular STXS region[4].

Before applying the dedicated per-category selection criteria, all events are required to have at least two pre-selected photon candidates, satisfying $p_T^{\gamma 1} > m_{\gamma\gamma}/3$ and $p_T^{\gamma 2} > m_{\gamma\gamma}/4$, with a diphoton invariant mass in the range $100 < m_{\gamma\gamma} < 180$ GeV. By scaling the transverse momentum by $m_{\gamma\gamma}$ in the selection criteria, distortions at the lower end of the mass spectrum are prevented. Furthermore, both photons are required to be in the geometrical acceptance of the ECAL: $|\eta| < 2.5$, excluding the transition region $1.44 < |\eta| < 1.57$ between the barrel and endcaps. We then construct the analysis categories by following the procedure described below.

1. *Global categories* are defined to be enriched with events originating from the different Higgs boson production modes. Dedicated selection cuts are defined based on the different event topologies. For example, VBF-like events are isolated by requiring a pair of jets (dijet) with an invariant mass, $m_{jj} > 350$ GeV. To further boost the ratio of signal-to-background events, ML classification algorithms are trained to isolate the targeted signal process from both SM background processes and the other Higgs boson production modes. At least one of these so-called *background-rejection discriminants* are used in each global category. In this analysis, global categories are defined to target the tHq, ttH, VH, VBF and ggH production modes, in a variety of final states.

2. The next step is to divide the per-production-mode global categories into *analysis regions* to differentiate between the different kinematic regions (bins) of the STXS framework. For ggH production, excluding the kinematic region with high dijet mass ($m_{jj} > 350$ GeV), the division is performed using a dedicated ML classifier with output classes corresponding to the ggH STXS bins. For the other production modes, the global categories are split according to the reconstructed equivalents of the truth-level variables defining the STXS boundaries e.g. $p_T^{\gamma\gamma}$ for p_T^H. These

[4]This can be a single bin or group of bins if the statistics are insufficient.

divisions are performed as long as there exists some statistical sensitivity to the split STXS bins.

3. An analysis region can be further sub-divided according to the signal-to-background purity. This is performed by placing boundaries on the relevant background-rejection discriminant(s). The position of these boundaries is optimised independently in each region to maximise the sensitivity to the targeted STXS bin (or bins). The usual metric for this optimisation is the approximate median significance (AMS),

$$\text{AMS} = \sqrt{2\Big((\text{S}_{68} + \text{B}_{68})\ln(1 + \text{S}_{68}/\text{B}_{68}) - \text{S}_{68}\Big)}, \tag{4.4}$$

where S_{68} is the number of signal events from the targeted STXS bin (or bins) and B_{68} the number of background events, both calculated within a $\pm 1\sigma_{\text{eff}}$ window of m_{H}. Here, σ_{eff} is defined as the shortest interval containing 68% of the total number of targeted signal events. The background estimate, B_{68}, is determined by fitting an exponential function to background events, either directly to data sidebands or to the MC simulated background events. In the limit of small S_{68}/B_{68}, the AMS reduces to the familiar $S_{68}/\sqrt{S_{68} + B_{68}}$ formula.

The final *analysis categories* designed to target the *truth-level* STXS bins are referred to as "tags", which are defined in decreasing order of the expected signal-to-background ratio. For example, the tag with the highest S_{68}/B_{68}, targeting the ggH 0J low p_T^H STXS bin, is denoted as "0J low $p_T^{\gamma\gamma}$ Tag0". The total number of tags is realised by a stopping criterion, where the expected AMS gain by introducing an additional boundary is below a threshold e.g. 5% improvement.

The data and simulation from all three years are merged together in each category. This provides larger statistics for the training of the numerous multivariate discriminants, and the subsequent optimisation of the selection criteria. In addition, the larger yields allow to better constrain the shape parameters of the per-category background models (see Section 5.3). The potential gain in splitting each category by year, and thus exploiting the year-dependent mass-resolution information in the final fit, was found to be negligible.

It is possible for any given event to pass the selection criteria for multiple analysis categories. A priority sequence is defined to ensure the analysis categories are truly orthogonal: each event is assigned to at most one category, choosing the category with the highest priority. The sequence is defined to prioritise rarer Higgs boson production modes, by ordering according to the expected number of signal events. In other words, global categories targeting a Higgs boson production mode with a lower expected signal-yield (rarer process) are given a higher priority.

A summary of all analysis categories is provided in Table 4.4, ordered according to decreasing priority in the category sequence. Example Feynman diagrams are shown for the targeted Higgs boson production modes. The background-rejection discriminants are listed, as well as the methods for splitting the global categories to target the kinematic regions of the STXS framework. Finally, the number of tags for each targeted STXS bin (or bins) are shown, defined by placing boundaries on the respective background-rejection discriminants. A graphical schema of the analysis event categorisation procedure is provided in Appendix C.

The remainder of this chapter provides additional detail concerning the construction, optimisation, and validation of the event categorisation workflow, and shows the performance in terms of the purity of the final analysis categories. For reference, Table 4.5 provides a summary of all the ML classifiers used in the categorisation. The classifiers mainly use the BDT algorithm, barring the DNN which is trained to discriminate ttH production from tHq production. In the table, the training samples, training software, and final outputs are listed. The full set of input features for each classifier are listed in Appendix C. When training the various classifiers, the samples are first weighted according to their respective SM cross sections, and are subject to the initial selection cuts of the relevant global category. Often, the weights of each output class (e.g. signal and background) are then equalised in the training to improve the classifier performance.

Table 4.4: A summary of the analysis event categorisation, ordered according to decreasing priority in the category sequence. For each global category, an example Feynman diagram of the targeted process is shown. In addition, the background-rejection discriminants and the methods used to split the global category to target different kinematic regions of the STXS framework are listed. The final column shows the number of tags defined to target each bin or group of bins.

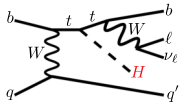
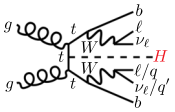
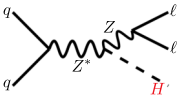
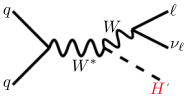
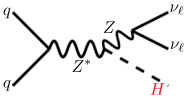
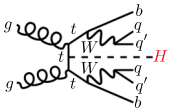
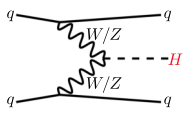
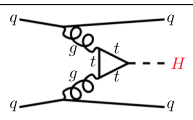
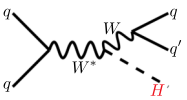
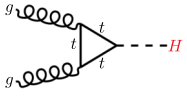
Global category	Example diagram	Bkg rej. discriminant	Splitting method to target STXS bins	Targeted STXS bins (# of bins)	# of tags
tHq leptonic		tHq leptonic BDT Top DNN (tH-vs-ttH)	None	tHq	1
ttH leptonic		ttH leptonic BDT Top DNN (tH-vs-ttH)	Reco $p_T^{\gamma\gamma}$	ttH $p_T^H < 60$ ttH $120 < p_T^H < 120$ ttH $120 < p_T^H < 200$ ttH $200 < p_T^H < 300$ ttH $p_T^H > 300$	3 3 2 1 1
ZH leptonic		ZH leptonic BDT	None	All ZH lep and ggZH lep bins (10)	2
WH leptonic		WH leptonic BDT	Reco $p_T^{\gamma\gamma}$	WH lep $p_T^V < 75$ WH lep $75 < p_T^V < 150$ WH lep $p_T^V > 150$ (3)	2 2 1
VH MET		VH MET BDT	None	All VH lep bins (15)	3
ttH hadronic		ttH hadronic BDT	Reco $p_T^{\gamma\gamma}$	ttH $p_T^H < 60$ ttH $120 < p_T^H < 120$ ttH $120 < p_T^H < 200$ ttH $200 < p_T^H < 300$ ttH $p_T^H > 300$	3 3 4 3 2
VBF-like		Dijet BDT Diphoton BDT	Reco $p_T^{\gamma\gamma}$, $p_T^{\gamma\gamma jj}$, m_{jj}	qqH BSM qqH VBF-like low m_{jj} low p_T^{Hjj} qqH VBF-like low m_{jj} high p_T^{Hjj} qqH VBF-like high m_{jj} low p_T^{Hjj} qqH VBF-like high m_{jj} high p_T^{Hjj}	2 2 2 2 2
			None	All ggH VBF-like bins (4)	2
VH hadronic		VH hadronic BDT Diphoton BDT	None	qqH VH-like	2
ggH		Diphoton BDT	ggH BDT + Reco $p_T^{\gamma\gamma}$ for high p_T^H region	ggH 0J low p_T^H ggH 0J high p_T^H ggH 1J low p_T^H ggH 1J med p_T^H ggH 1J high p_T^H ggH ≥2J low p_T^H ggH ≥2J med p_T^H ggH ≥2J high p_T^H ggH BSM $200 < p_T^H < 300$ ggH BSM $300 < p_T^H < 450$ ggH BSM $450 < p_T^H < 650$ ggH BSM $p_T^H > 650$	3 3 3 3 3 3 3 3 2 2 1 1
No dedicated category	-	-	-	qqH 0J, 1J, ≥2J $m_{jj} < 60$ qqH ≥2J $120 < m_{jj} < 350$ bbH, tHW	0
				Total	80

Table 4.5: A summary of the ML classifiers used in this analysis. Each classifier is trained using the listed samples after applying the initial selection criteria of the relevant global category. In the list of samples, top-associated backgrounds correspond to events containing at least one top quark and at least two reconstructed photons, such as tt+$\gamma\gamma$ or tt+γ+jet where the jet is misidentified as a photon. The software used to train the classifiers, and their respective outputs are also listed. Usually, the weights of each output class are equalised in training to improve the performance of the classifiers. The full set of input features for each classifier are provided in Appendix C.

Discriminant/Classifier	Software	Training samples (simulation unless stated)	Output
ggH BDT	XGBoost	ggH	9 ggH STXS class probabilities: 0J low p_T^H, 0J high p_T^H, 1J low p_T^H, 1J med p_T^H, 1J high p_T^H, $\geq$2J low p_T^H, $\geq$2J med p_T^H, $\geq$2J high p_T^H, $p_T^H > 200$
Diphoton BDT	XGBoost	S: all Higgs boson events B: SM $\gamma\gamma$	S-vs-B score
Dijet BDT	XGBoost	VBF, ggH, SM $\gamma\gamma$ + data-driven estimate for γ+jet & jet+jet	3 output class probabilities: VBF, ggH, bkg
VH hadronic BDT	XGBoost	S: VH hadronic B: ggH, SM $\gamma\gamma$ + data-driven estimate for γ+jet & jet+jet	S-vs-B score
VH MET BDT	TMVA	S: VH 0-leptons B: other H prod modes, SM $\gamma\gamma$, Drell-Yan, diboson, top quark prod. + data-driven estimate for γ+jet	S-vs-B score
WH leptonic BDT	TMVA	S: VH 1-lepton B: other H prod modes, SM $\gamma\gamma$, γ+jet, Drell-Yan, diboson, top-associated bkgs	S-vs-B score
ZH leptonic BDT	TMVA	S: VH $\geq$2-leptons B: other H prod modes, SM $\gamma\gamma$, γ+jet, Drell-Yan, diboson, top-associated bkgs	S-vs-B score
ttH hadronic BDT	XGBoost	S: ttH 0-leptons, $\geq$3-jets ($\geq$1 b-tagged) B: other H prod modes, SM $\gamma\gamma$, Drell-Yan, V+γ, diboson, top-associated bkgs + data-driven estimate for γ+jet & jet+jet	S-vs-B score
ttH leptonic BDT	XGBoost	S: ttH $\geq$0-leptons, $\geq$1-jet B: other H prod modes, SM $\gamma\gamma$, γ+jet, Drell-Yan, V+γ, diboson, top-associated bkgs	S-vs-B score
tHq leptonic BDT	TMVA	S: tHq leptonic B: SM $\gamma\gamma$, γ+jet, top-associated bkgs	S-vs-B score
Top DNN	Keras + TensorFlow	tHq and ttH	tHq-vs-ttH score

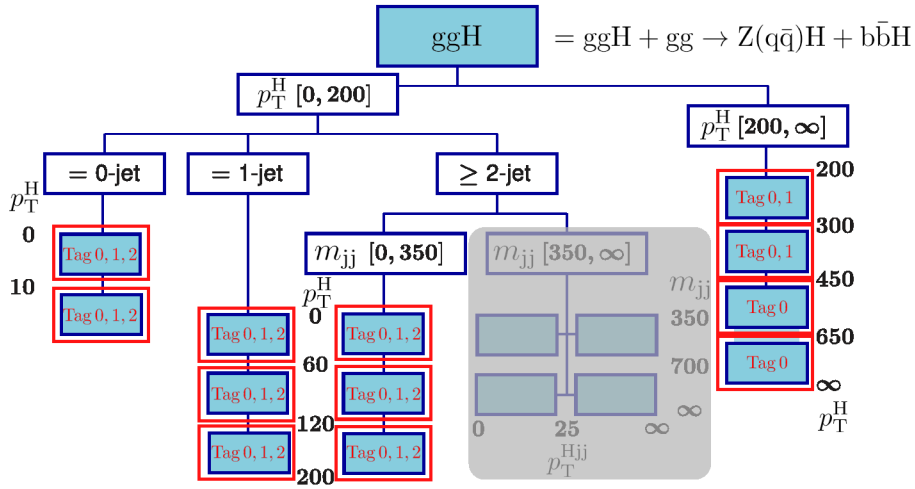


Figure 4.7: A schematic of the ggH categorisation scheme. The analysis categories, defined by the red boxes, surround their targeted STXS bin. The number of tags for each analysis region are highlighted in the plot. A shaded-grey region is shown for the STXS bins with a dijet invariant mass, $m_{jj} > 350$ GeV, as they are not targeted in the ggH categorisation scheme.

4.4.2 ggH categorisation

The final global category in the sequence targets the ggH production mode and considers all events that pass the photon pre-selection but are not selected by higher-priority categories. Firstly, the ggH BDT is used to predict the most probable ggH STXS region from which the event originates. The BDT outputs nine class-probabilities corresponding to the region with $p_T^H > 200$ GeV and the two 0-jet, three 1-jet and three ≥2-jet STXS bins with $p_T^H < 200$ GeV and $m_{jj} < 350$ GeV. To minimise the dependence on the modelling of events with high p_T^H, the BDT is not trained to distinguish between the four bins with $p_T^H > 200$ GeV; instead events in this class are further split into the final STXS bins using the reconstructed $p_T^{\gamma\gamma}$. Events with a reconstructed dijet invariant mass $m_{jj} > 350$ GeV are not assigned by the ggH BDT and are instead classified using the dijet BDT, as described in Section 4.4.3.

The nine ggH classes are uniquely defined by the Higgs boson transverse-momentum and the number of jets. It was found that the ggH BDT outperforms the assignment from using the reconstructed $p_T^{\gamma\gamma}$ and number of jets alone. This improvement arises as the BDT can exploit the correlations between the photon and jet kinematic properties, and thus use the well-measured photon quantities to infer information related to the less well-measured jets. As a result, the correct assignment of the per-event number of jets increases from 77% to 82%; the improvement is negligible for the $p_T^{\gamma\gamma}$ assignment since this is already a well-measured quantity with little migration across the p_T^H boundaries. Propagating this

to the final results leads to an improvement in the measured cross sections, particularly for the 0J and 1J bins ($\sim 10\%$ reduction in the uncertainty), and reduces the correlations between measured quantities.

The nine output class-probabilities are calculated for each event, and the event is subsequently assigned to the STXS region with the highest probability. The final 12 analysis regions (9 output ggH BDT classes, where the $p_T^H > 200$ GeV class is further split using boundaries in the reconstructed $p_T^{\gamma\gamma}$ at 300, 450 and 650 GeV) are then divided into the final tags using the diphoton BDT. This BDT is trained to discriminate all Higgs boson signal from all other modes of SM diphoton production, meaning the output score, as shown in Figure 4.8, can be used to reduce the background contamination in each of the analysis regions. The boundaries on the diphoton BDT output are optimised independently in each region to maximise the expected AMS. Three tags are defined for each of the eight STXS bins with $p_T^H < 200$ GeV, two for the two bins with $200 < p_T^H < 450$ GeV, and one for the two bins with $p_T^H > 450$ GeV. The final set of analysis categories, with their respective targeted STXS bins, are shown by the red boxes in Figure 4.7. Events that fail the lowest diphoton BDT output threshold in the respective region are discarded from the analysis.

Clearly, the agreement between data and the background simulation is reasonably poor in certain bins of the diphoton BDT output distribution. This mis-modelling is likely to be improved in the future when samples generated at a higher order in perturbation theory are used for estimating the γ+jet and jet+jet contributions to the background. Nevertheless, the poor agreement does not manifest as a systematic uncertainty in the analysis, as the background model is estimated directly from data. Instead, it leads only to a sub-optimal performance of the diphoton BDT. This is a common trait for all ML classifiers used in the event categorisation, and is highlighted again when discussing the validation of the event categorisation in Section 4.4.6.

Table 4.6 presents the expected signal and background yields in each of the ggH analysis categories. The signal yields are broken down into the fractional contribution from the target STXS bin, as well as the fractional contributions from the different Higgs boson production modes.

4.4.3 EW qqH categorisation

As described in Section 2.3, the definition of qqH production in the STXS framework includes both the t and u-channel contributions (VBF production), and the s-channel contribution (VH production in which the vector boson decays hadronically). Characteristics of the dijet system are used to construct separate orthogonal global categories, effectively distinguishing between the VBF and VH hadronic topologies. Of these two,

Table 4.6: The expected number of events for $m_{\mathrm{H}} = 125$ GeV in the analysis categories targeting the ggH production mode, shown for an integrated luminosity of 137 fb^{-1}. The fraction of the total number of events arising from each production mode in each analysis category is provided, as is the fraction of events originating from the targeted STXS bin (or bins). Here, ggH includes contributions from the sub-dominant ggZ(q$\bar{\mathrm{q}}$)H production mode, qqH includes both VBF and V(q$\bar{\mathrm{q}}$)H production, and "Top" represents both ttH and tH production together. The σ_{eff}, defined as the smallest interval containing 68.3% of the $m_{\gamma\gamma}$ distribution provides an indication of the mass resolution in each category. Also provided are the estimated number of background events-per-GeV in the signal peak region, the quantity $F_{68} = S_{68}/(S_{68}+B_{68})$, where S_{68} and B_{68} are the expected number of signal and background events in a $\pm 1\sigma_{\mathrm{eff}}$ window centred on m_{H}, respectively, and the approximate significance, $Z_{68} = S_{68}/\sqrt{S_{68}+B_{68}}$. The final column shows the significance for the targeted STXS bin (or bins) only, Z_{68}^{target}, where other Higgs boson signal events are considered as background.

Analysis categories	SM 125 GeV Higgs boson expected signal								Bkg	F_{68}	Z_{68}	Z_{68}^{target}
	Total	Target	Fraction of total events					σ_{eff}	(GeV^{-1})			
		STXS bin(s)	ggH	bbH	qqH	VH lep	Top	(GeV)				
0J low $p_T^{\gamma\gamma}$ Tag0	296.2	86.6%	97.9%	1.1%	0.8%	0.1%	-	1.89	1760	0.06	3.37	2.92
0J low $p_T^{\gamma\gamma}$ Tag1	340.0	88.5%	98.0%	1.0%	0.8%	0.1%	-	2.31	3140	0.03	2.66	2.35
0J low $p_T^{\gamma\gamma}$ Tag2	279.6	89.3%	98.1%	1.0%	0.8%	0.1%	-	2.53	4980	0.01	1.67	1.49
0J high $p_T^{\gamma\gamma}$ Tag0	612.4	81.9%	95.6%	1.4%	2.6%	0.4%	-	1.64	2720	0.08	5.89	4.82
0J high $p_T^{\gamma\gamma}$ Tag1	1114.6	79.4%	95.4%	1.3%	2.8%	0.4%	-	2.19	7610	0.04	5.66	4.50
0J high $p_T^{\gamma\gamma}$ Tag2	1162.6	78.3%	95.3%	1.4%	2.7%	0.5%	-	2.56	19500	0.02	3.46	2.71
1J low $p_T^{\gamma\gamma}$ Tag0	132.0	66.2%	88.8%	0.8%	9.4%	0.8%	0.1%	1.53	487	0.11	3.06	2.03
1J low $p_T^{\gamma\gamma}$ Tag1	340.0	66.3%	88.6%	0.8%	9.6%	0.9%	0.1%	1.95	2300	0.05	3.31	2.19
1J low $p_T^{\gamma\gamma}$ Tag2	260.6	66.2%	88.3%	0.8%	9.7%	1.0%	0.1%	2.37	3860	0.02	1.80	1.19
1J med $p_T^{\gamma\gamma}$ Tag0	184.1	65.2%	81.7%	0.5%	16.3%	1.4%	0.2%	1.65	443	0.14	4.21	2.74
1J med $p_T^{\gamma\gamma}$ Tag1	310.2	66.3%	83.6%	0.4%	14.3%	1.6%	0.1%	1.91	1410	0.07	3.82	2.53
1J med $p_T^{\gamma\gamma}$ Tag2	291.4	65.0%	83.7%	0.5%	13.8%	1.8%	0.2%	2.13	2960	0.03	2.40	1.56
1J high $p_T^{\gamma\gamma}$ Tag0	37.3	61.9%	75.7%	0.2%	22.8%	1.0%	0.2%	1.55	43	0.27	2.61	1.61
1J high $p_T^{\gamma\gamma}$ Tag1	31.2	61.7%	75.0%	0.3%	23.4%	1.1%	0.2%	1.73	67	0.15	1.78	1.10
1J high $p_T^{\gamma\gamma}$ Tag2	80.9	62.2%	76.5%	0.2%	21.5%	1.6%	0.2%	1.97	388	0.07	1.87	1.16
$\geq$2J low $p_T^{\gamma\gamma}$ Tag0	17.7	52.7%	76.7%	0.6%	19.0%	1.3%	2.4%	1.56	120	0.06	0.84	0.44
$\geq$2J low $p_T^{\gamma\gamma}$ Tag1	57.6	54.0%	74.4%	0.6%	20.5%	1.4%	3.0%	1.88	664	0.03	1.08	0.58
$\geq$2J low $p_T^{\gamma\gamma}$ Tag2	43.9	50.5%	72.7%	0.6%	20.8%	1.7%	4.2%	2.46	1130	0.01	0.56	0.28
$\geq$2J med $p_T^{\gamma\gamma}$ Tag0	21.2	64.9%	80.6%	0.3%	16.3%	1.0%	1.8%	1.42	52	0.16	1.51	0.98
$\geq$2J med $p_T^{\gamma\gamma}$ Tag1	70.1	61.4%	77.9%	0.3%	18.1%	1.1%	2.6%	1.82	350	0.07	1.79	1.10
$\geq$2J med $p_T^{\gamma\gamma}$ Tag2	135.4	57.5%	74.8%	0.4%	19.7%	1.4%	3.8%	2.08	1630	0.03	1.54	0.88
$\geq$2J high $p_T^{\gamma\gamma}$ Tag0	29.0	65.5%	77.8%	0.2%	18.7%	1.3%	2.1%	1.48	47	0.22	2.05	1.34
$\geq$2J high $p_T^{\gamma\gamma}$ Tag1	52.5	62.3%	76.1%	0.2%	19.6%	1.5%	2.6%	1.76	172	0.10	1.90	1.19
$\geq$2J high $p_T^{\gamma\gamma}$ Tag2	45.5	58.4%	73.8%	0.2%	20.4%	1.9%	3.7%	1.92	306	0.05	1.22	0.71
BSM $200 < p_T^{\gamma\gamma} < 300$ Tag0	30.7	75.8%	77.5%	0.2%	19.4%	1.2%	1.6%	1.41	24	0.38	2.78	2.10
BSM $200 < p_T^{\gamma\gamma} < 300$ Tag1	39.6	69.9%	73.8%	0.1%	21.5%	1.7%	2.8%	1.90	123	0.10	1.63	1.14
BSM $300 < p_T^{\gamma\gamma} < 450$ Tag0	15.5	74.8%	76.3%	0.1%	19.7%	1.7%	2.2%	1.53	14	0.33	1.85	1.39
BSM $300 < p_T^{\gamma\gamma} < 450$ Tag1	2.6	66.3%	67.9%	0.1%	22.5%	2.6%	7.0%	1.42	13	0.08	0.38	0.25
BSM $450 < p_T^{\gamma\gamma} < 650$	3.1	58.1%	61.8%	0.1%	30.0%	2.4%	5.6%	1.55	5.5	0.20	0.65	0.38
BSM $p_T^{\gamma\gamma} > 650$	0.9	72.5%	72.3%	0.1%	21.0%	2.9%	3.8%	1.21	0.95	0.36	0.48	0.34

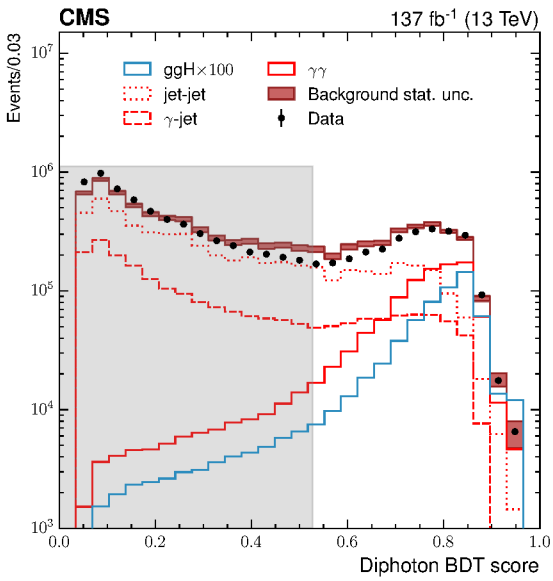


Figure 4.8: The diphoton BDT score distribution for all events passing pre-selection and satisfying $100 < m_{\gamma\gamma} < 180$ GeV. Events from simulation and data are shown by the red band and the black points respectively. Histograms are shown for the different contributions of the simulated background in red. The blue histogram corresponds to simulated Higgs boson signal events, produced via ggH, multiplied by a factor of 100 to ease the comparison with the background distributions. The shaded grey region represents scores below the lowest diphoton BDT threshold used to define the final analysis tags. The full data set collected in the period 2016-2018 and the corresponding simulation is shown.

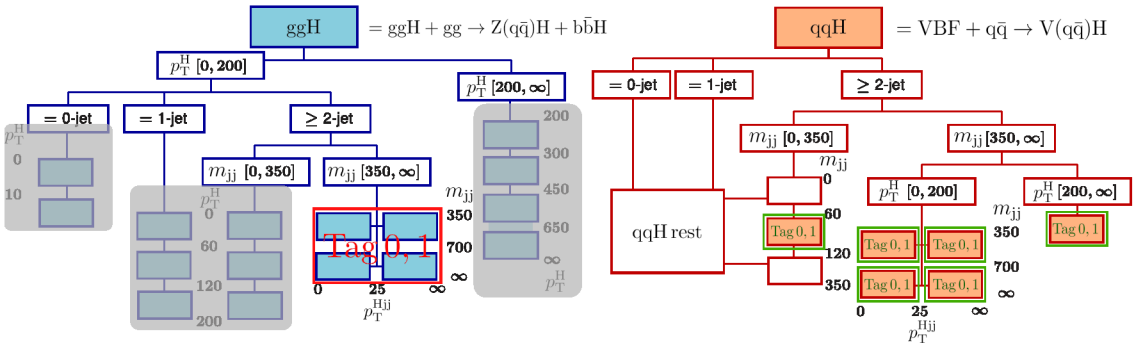


Figure 4.9: A schematic of the qqH and ggH VBF-like categorisation scheme. The analysis categories, defined by the red and green boxes, surround their targeted STXS bin (or bins). The number of tags for each analysis region are highlighted in the plot. The shaded-grey regions indicate the ggH STXS bins which are targeted in the ggH categorisation scheme.

the global category targeting the VBF-like topology is described first.

VBF-like topology

The VBF-like STXS bins are specified at truth-level as events containing a dijet system, with a dijet invariant mass, $m_{jj} > 350$ GeV. At reconstruction-level, the signal region requires events to have at least two jets, with $p_T^j > 40$ and 30 GeV for the highest-p_T (leading) and second highest p_T (sub-leading) jets, respectively, $|\eta^j| < 4.7$, and a reconstructed $m_{jj} > 350$ GeV. In addition, all jets are required to pass a threshold on the pileup-identification BDT score and photons are required to have a photon-ID BDT score of greater than -0.2. Note, the definition of this signal region includes contributions from both the VBF and ggH production modes (shown by the Feynman diagrams in the VBF-like row of Table 4.4), where ggH events with a VBF-like topology are defined by a set of STXS bins in the ggH scheme with equivalent boundaries on the truth-level kinematic quantities. Events are subsequently classified as originating from either VBF, ggH, or from other SM backgrounds using the dijet BDT, where the outputs of the BDT are per-event probability estimates for each of the three classes: p_{VBF}, p_{ggH}, and p_{bkg}.

The dijet BDT is trained using simulated VBF, ggH and SM diphoton-production events. Background events in which at least one of the photons originates from a misreconstructed jet are poorly modelled in simulation. This is a result of the difficult-to-model quark/gluon-fragmentation processes, in addition to the fact that very few events pass the VBF-like signal region selection criteria defined above. To improve the prediction of this background contribution when training the dijet BDT, simulated events are replaced with data from a dedicated control region, defined to be enriched with γ+jet and jet+jet events. The procedure is as follows:

- Events are split into the signal and control regions. In the control region, at least one of the photons is required to have a photon-ID BDT score less than -0.5. This defines two contributions in the control region: fake-prompt (FP) where one of the reconstructed photons has a score less than -0.5 and the other greater than -0.2, and fake-fake (FF) where both the photons have a score less than -0.5. The events in the signal region are defined by requiring both photons to have photon-ID BDT scores greater than -0.2. Since the photon-ID BDT score is used to define the control region, it is not included as an input feature to the dijet BDT.

- The number of events differs in the control region and signal region. Moreover, events in the two regions can have different kinematic properties. To correct this, a fake-factor is derived in bins of reconstructed photon p_T^γ and η^γ using simulated events, as the ratio of the expected number of events in the signal region,

$N^{\mathrm{SR}}(p_T^\gamma, \eta^\gamma)$, to the expected number of events in the control region, $N^{\mathrm{CR}}(p_T^\gamma, \eta^\gamma)$:

$$p_{\mathrm{fake}}(p_T^\gamma, \eta^\gamma) = \Big(\frac{N^{\mathrm{SR}}(p_T^\gamma, \eta^\gamma)}{N^{\mathrm{CR}}(p_T^\gamma, \eta^\gamma)}\Big)_{\mathrm{MC}}. \tag{4.5}$$

- As the SM diphoton contribution to the background is predicted directly from simulation, it is necessary to remove this contribution from the control regions to avoid double-counting. Again using simulation, the fraction of background events in the control region which originate from a fake photon (γ+jet and jet+jet) is calculated in bins of p_T^γ and η^γ as,

$$p_{\mathrm{QCD}}(p_T^\gamma, \eta^\gamma) = \Big(\frac{N_{\gamma j}^{\mathrm{CR}}(p_T^\gamma, \eta^\gamma) + N_{jj}^{\mathrm{CR}}(p_T^\gamma, \eta^\gamma)}{N_{\gamma\gamma}^{\mathrm{CR}}(p_T^\gamma, \eta^\gamma) + N_{\gamma j}^{\mathrm{CR}}(p_T^\gamma, \eta^\gamma) + N_{jj}^{\mathrm{CR}}(p_T^\gamma, \eta^\gamma)}\Big)_{\mathrm{MC}}. \tag{4.6}$$

- The total transfer factor, evaluated for each photon with photon-ID BDT score less than -0.5, is defined as,

$$f(p_T^\gamma, \eta^\gamma) = p_{\mathrm{fake}} \times p_{\mathrm{QCD}}. \tag{4.7}$$

  These transfer factors are applied as weights to the data events in the control region to estimate the contribution in the signal region, according to the illustration in Figure 4.10. The negative sign for events in the FF region originates by considering the contribution which enters the FP region from the FF region when applying the transfer factors. It is this sample of reweighted data events from the control region which is used to train the dijet BDT.

The data-driven approach is validated by comparing the dijet-BDT input-feature distributions from the background prediction (simulated diphoton component plus the data-driven γ+jet and jet+jet components) to data from sidebands in the diphoton-mass distribution. This sideband data is subject to the usual VBF-like signal region selection, except the diphoton mass is required to be outside of the range $115 < m_{\gamma\gamma} < 135$ GeV. Figure 4.11 demonstrates good agreement in the m_{jj} and $\Delta\phi_{jj,\gamma\gamma}$ distributions with the data-driven background estimate (middle panel). In addition, the figure highlights the increase in the training statistics with respect to using simulated events only (bottom panel), thus leading to an improved performance of the classifier.

Figure 4.12 shows the two independent output probabilities of the dijet BDT: the VBF probability, p_{VBF} (left), and the ggH probability, p_{ggH} (right). The background probability is simply realised according to the equation: $p_{\mathrm{bkg}} = 1 - p_{\mathrm{VBF}} - p_{\mathrm{ggH}}$.

The VBF-like global category is split to target different kinematic bins of the STXS

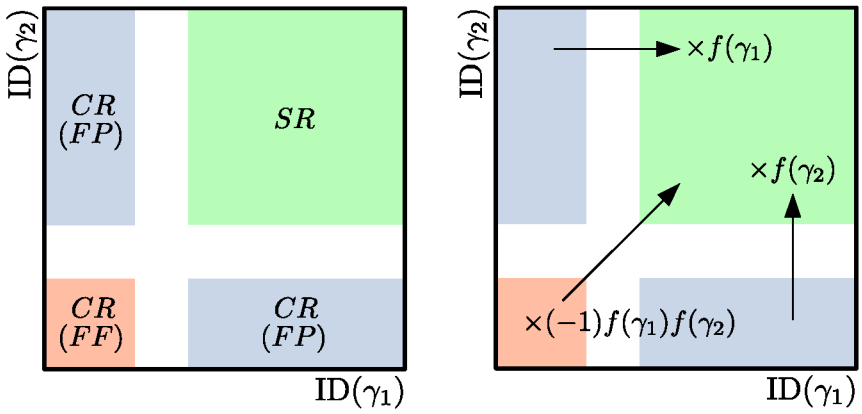


Figure 4.10: A schematic showing the application of transfer factors to estimate the fake component of the background in the signal region (SR) directly from data events in dedicated control regions (CR).

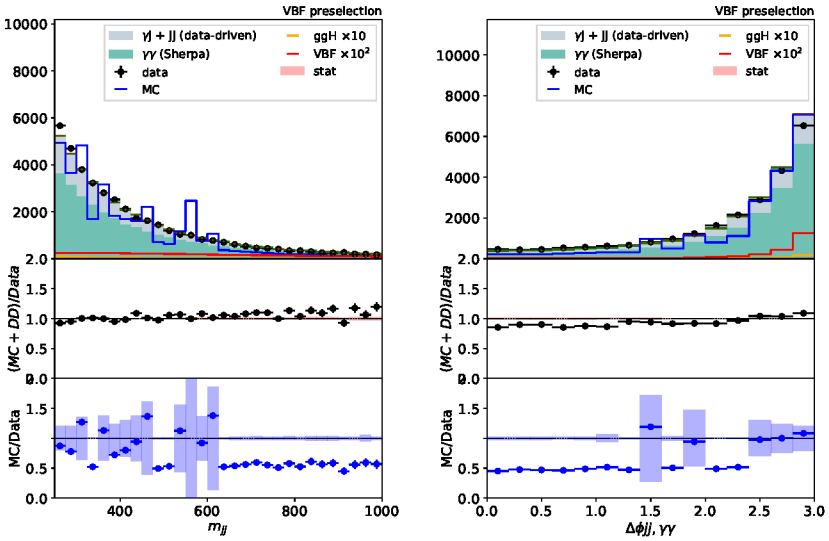


Figure 4.11: Input features to the dijet BDT: m_{jj} and $\Delta\phi_{jj,\gamma\gamma}$. The data events shown by the black points were collected in the 2018 period. The simulated $\gamma\gamma$ background estimate and the data-driven γ+jet and jet+jet background estimates are shown by the filled green and blue histograms, respectively. In addition, the simulated γ+jet and jet+jet background estimate is shown by the blue line. The bottom panel shows the ratio of the purely simulated background estimate to data, where the agreement clearly suffers from a lack of MC statistics in the γ+jet and jet+jet component of the background. The middle panel shows the ratio including the data-driven estimate of the background, where the agreement is much improved.

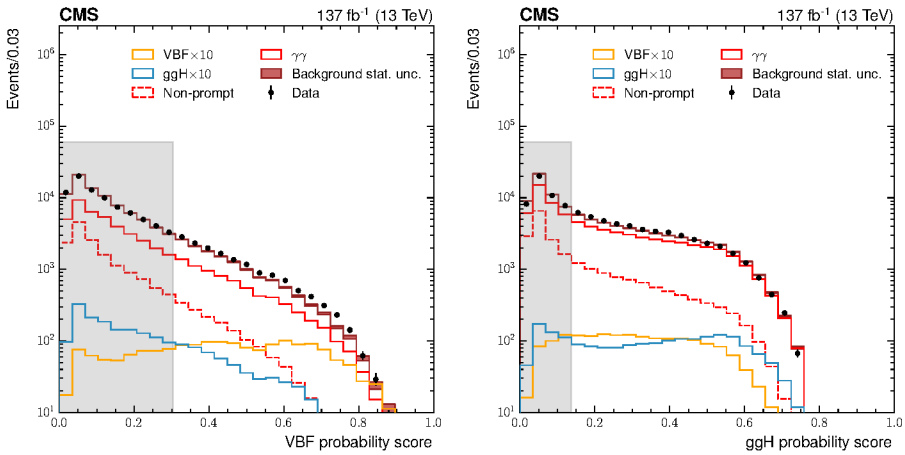


Figure 4.12: The dijet BDT output probabilities, p_{VBF} (left) and p_{ggH} (right), for events with an invariant mass in the range $100 < m_{\gamma\gamma} < 180$ GeV, passing the VBF-like signal region selection. Data events are shown by the black points, and the corresponding background simulation with statistical uncertainty is shown by the red band. The components of the background are plotted in separate histograms in red. In addition the orange and blue histograms show the simulated VBF and ggH signal events, respectively. The shaded grey regions contain p_{VBF} values below the lowest threshold used to define the VBF production tags, and p_{ggH} values below the lowest threshold used to define the ggH VBF-like tags. The full data set collected in the period 2016-2018 and the corresponding simulation is shown.

framework. For the qqH STXS bins, events are assigned to one of five regions using the reconstructed equivalents of the truth-level variables. Firstly, the qqH BSM STXS bin ($p_T^H > 200$ GeV) is targeted by requiring the reconstructed $p_T^{\gamma\gamma} > 200$ GeV. The four remaining VBF-like bins are targeted according to a boundary in the reconstructed m_{jj} at 700 GeV, and a boundary in the transverse momentum of the dijet-plus-diphoton system, $p_T^{\gamma\gamma jj}$, at 25 GeV. In each region, two tags are constructed by placing optimised boundaries on the dijet-BDT output probabilities. This optimisation considers VBF production as signal and groups ggH production with other background events. In this manner, a lower bound is placed on p_{VBF}, and an upper bound is placed on p_{ggH}. On top of this, the optimisation procedure includes a threshold on the diphoton BDT score to further reduce the number of background events.

Events failing the above criteria are then considered for an additional two tags, designed to be enriched in events from the four ggH VBF-like bins. This optimisation considers ggH as signal, and VBF as background. This amounts to placing a lower bound on p_{ggH} and an upper bound on p_{VBF}, in addition to a lower threshold on the diphoton BDT score. Events with $p_T^{\gamma\gamma} > 200$ GeV that have failed the qqH BSM selection criteria are not considered in the ggH VBF-like tags as they are specifically targeted further down the priority sequence by the ggH BSM analysis categories.

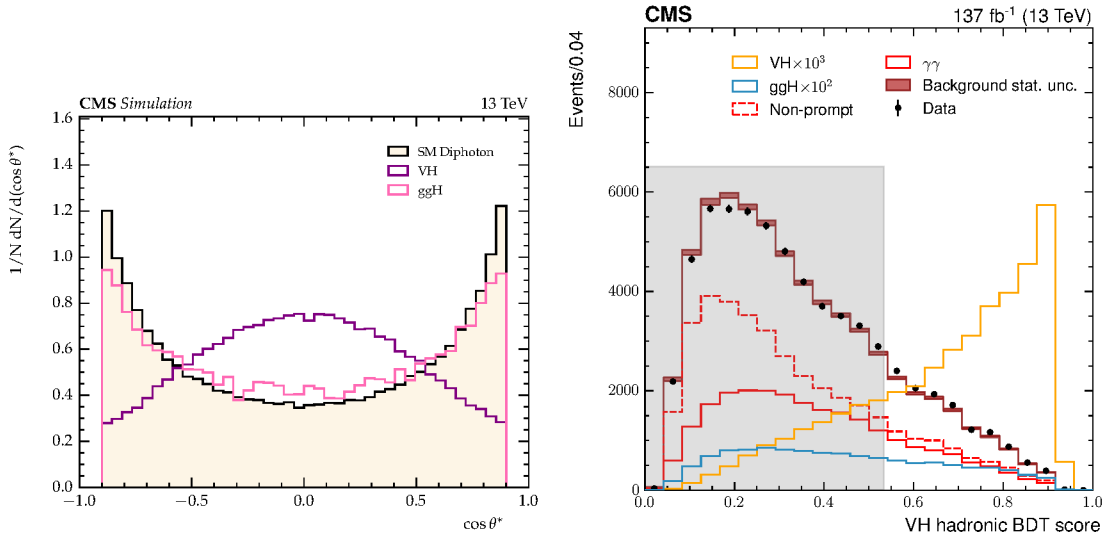


Figure 4.13: Distribution of the $\cos\theta^*$ variable for simulated VH, ggH, and SM-diphoton production events (left). The output score for the VH hadronic BDT (right), for events with an invariant mass in the range $100 < m_{\gamma\gamma} < 180$ GeV, passing the VH-like signal region selection. Data events are shown by the black points, and the corresponding background simulation with statistical uncertainty is shown by the red band. The components of the background are plotted in separate histograms in red. In addition the orange and blue histograms show the simulated VH and ggH signal events, respectively. The shaded grey region illustrates the lowest threshold used to define the VH hadronic analysis categories, The full data set collected in the period 2016-2018 and the corresponding simulation is shown.

VH hadronic topology

VH hadronic events enrich a single STXS bin (qqH VH-like) in the EW qqH scheme. This is defined at truth-level with the requirement, $60 < m_{jj} < 120$ GeV, to be consistent with a dijet system originating from the decay of a vector boson. To target this bin, an orthogonal signal region is defined with a different selection on the reconstructed dijet-system: events are required to have two jets within $|\eta^j| < 2.4$ and with $p_T^j > 30$ GeV, satisfying the same pileup-identification BDT and photon-ID BDT score requirements as the VBF-topology signal region, with a reconstructed m_{jj} between 60 and 120 GeV.

The VH hadronic BDT is then trained to isolate VH signal events, and subsequently reduce the contribution from ggH and other background processes in the signal region. One of the key input features which helps identify dijets consistent with a vector boson decay is the $\cos\theta^*$ variable, where θ^* is the angle that the diphoton system makes in the diphoton-dijet centre-of-mass frame, with respect to the direction of motion of the diphoton-dijet system in the lab frame. The distribution of this variable is reasonably uniform for VH events, while it is strongly peaked at ± 1 for background and ggH events, as

Table 4.7: The expected number of events for $m_{\mathrm{H}} = 125$ GeV in the analysis categories targeting the EW qqH STXS regions, in addition to the two tags targeting ggH production with a VBF-like topology, shown for an integrated luminosity of 137 fb^{-1}. The fraction of the total number of events arising from each production mode in each analysis category is provided, as is the fraction of events originating from the targeted STXS bin (or bins). Here, ggH includes contributions from the sub-dominant ggZ($q\bar{q}$)H and bbH production modes, and "Top" represents both ttH and tH production together. The σ_{eff}, defined as the smallest interval containing 68.3% of the $m_{\gamma\gamma}$ distribution provides an indication of the mass resolution in each category. Also provided are the estimated number of background events-per-GeV in the signal peak region, the quantity $F_{68} = S_{68}/(S_{68} + B_{68})$, where S_{68} and B_{68} are the expected number of signal and background events in a $\pm 1\sigma_{\mathrm{eff}}$ window centred on m_{H}, respectively, and the approximate significance, $Z_{68} = S_{68}/\sqrt{S_{68} + B_{68}}$. The final column shows the significance for the targeted STXS bin (or bins) only, Z_{68}^{target}, where other Higgs boson signal events are considered as background.

Analysis categories	SM 125 GeV Higgs boson expected signal								Bkg	F_{68}	Z_{68}	Z_{68}^{target}
	Total	Target STXS bin(s)	Fraction of total events ggH	VBF	VH had	VH lep	Top	σ_{eff} (GeV)	(GeV^{-1})			
ggH VBF-like Tag0	14.1	37.7%	65.9%	27.3%	3.8%	0.8%	2.3%	1.85	35	0.13	1.09	0.41
ggH VBF-like Tag1	32.5	30.2%	61.3%	29.8%	4.1%	1.1%	3.7%	1.83	117	0.09	1.39	0.42
qqH low m_{jj} low p_T^{Hjj} Tag0	17.2	48.2%	36.6%	62.6%	0.4%	0.1%	0.3%	1.89	26	0.19	1.48	0.71
qqH low m_{jj} low p_T^{Hjj} Tag1	13.5	48.5%	35.5%	63.4%	0.6%	0.1%	0.3%	1.74	24	0.18	1.27	0.62
qqH high m_{jj} low p_T^{Hjj} Tag0	27.0	70.4%	17.1%	82.7%	0.2%	-	0.1%	1.78	11	0.48	2.94	2.07
qqH high m_{jj} low p_T^{Hjj} Tag1	12.9	58.2%	20.8%	78.7%	0.3%	0.1%	0.2%	1.99	13	0.25	1.48	0.86
qqH low m_{jj} high p_T^{Hjj} Tag0	10.4	15.0%	56.0%	41.3%	1.3%	0.4%	1.0%	1.92	27	0.12	0.90	0.14
qqH low m_{jj} high p_T^{Hjj} Tag1	20.2	17.0%	57.9%	36.9%	2.4%	0.7%	2.1%	1.74	94	0.08	1.01	0.17
qqH high m_{jj} high p_T^{Hjj} Tag0	18.1	25.6%	28.1%	70.8%	0.4%	0.1%	0.5%	1.88	17	0.27	1.82	0.47
qqH high m_{jj} high p_T^{Hjj} Tag1	17.5	23.8%	39.5%	57.8%	0.9%	0.3%	1.5%	1.98	42	0.12	1.20	0.29
qqH BSM Tag0	11.2	71.2%	24.4%	74.8%	0.1%	0.1%	0.6%	1.62	3.9	0.54	2.02	1.44
qqH BSM Tag1	6.8	56.4%	36.9%	59.9%	1.1%	0.4%	1.7%	1.67	4.6	0.37	1.31	0.74
qqH VH-like Tag0	16.3	55.8%	36.5%	2.8%	55.0%	1.4%	4.2%	1.72	20	0.24	1.64	0.91
qqH VH-like Tag1	47.1	26.8%	64.9%	4.7%	26.4%	1.2%	2.9%	1.66	135	0.12	1.97	0.53

shown in Figure 4.13 (left). To train the BDT, the VH, ggH, and SM diphoton samples are taken from simulation, whereas the background with jets faking photons is again derived from a data control region, in an analogous manner to that used for the dijet BDT. The output score for the VH hadronic BDT is shown in Figure 4.13 (right). Two analysis tags are defined by placing optimised boundaries on both the VH hadronic BDT and diphoton BDT output scores.

Table 4.7 provides a summary of the expected signal and background yields in each of the analysis categories described in this section: those targeting the VBF-topology from both the ggH and qqH production modes, and the VH hadronic topology. The final set of analysis categories described in this section, along with their targeted STXS bin (or bins), are shown in Figure 4.9.

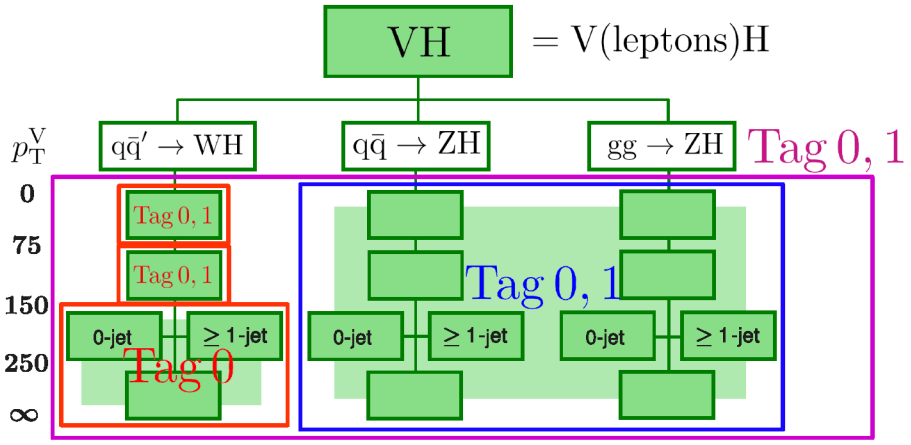


Figure 4.14: A schematic of the VH leptonic categorisation scheme. The WH leptonic, ZH leptonic and VH MET analysis categories are defined by the red, blue, and purple boxes, respectively. The analysis categories surround the targeted STXS bin (or bins). The number of tags for each analysis region are highlighted in the plot.

4.4.4 VH leptonic categorisation

Three global categories are constructed to target Higgs boson production in association with a leptonically-decaying vector boson. These categories are orthogonal to each other by construction due to requiring a different number of isolated charged leptons in the event: two for ZH leptonic, one for WH leptonic and zero for the VH MET category. The ZH leptonic category imposes a requirement on the dilepton invariant mass, m_{ll}, to be between 60 and 120 GeV. This ensures the two leptons are consistent with originating from the decay of a Z boson: $Z(\to \ell\ell)H$. The WH leptonic category, which targets $W(\to \ell\nu)H$ production, places additional requirements on the photon-ID BDT score to further reject backgrounds from jets faking photons, and on the invariant mass of the reconstructed lepton with each photon to reduce the contribution from the Drell-Yan background in which an electron has been misidentified as a photon. The VH MET category is constructed to target events from $Z(\to \nu\nu)H$ production, and $W(\to \ell\nu)H$ production where the charged lepton is not reconstructed in the detector. Here, events are required to have $p_T^{miss} > 50$ GeV and the azimuthal angle between $\vec{p}_T^{miss}$ and the diphoton system must be greater than two radians. The categories targeting VH leptonic production are placed between the ttH leptonic and ttH hadronic global categories in the priority sequence.

To further distinguish between VH leptonic signal and background events, a BDT is trained in each global category; the details of which are provided in Table 4.5. The simulated background processes include SM diphoton production, γ+jets, Drell-Yan, diboson production, and tt production. In addition, the Higgs boson production modes other than VH are treated as backgrounds. Due to the large contamination of the γ+jet events

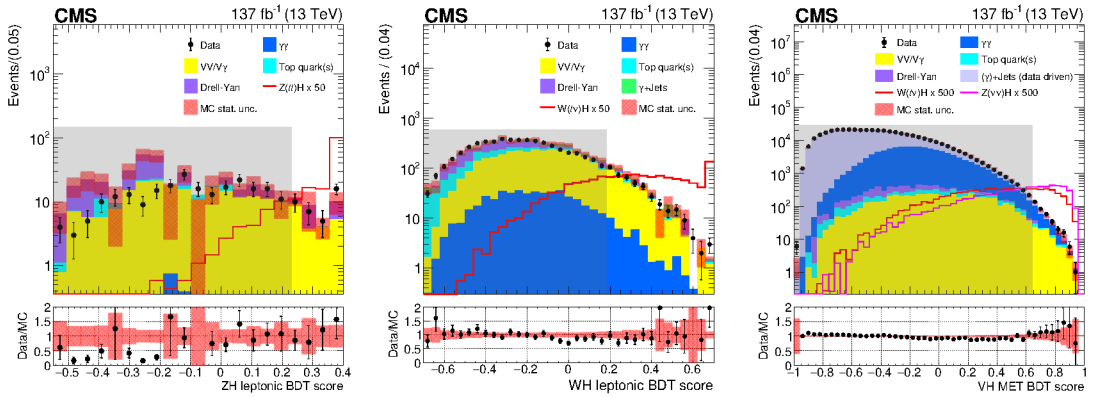


Figure 4.15: Distributions of the ZH leptonic (left), WH leptonic (centre) and VH MET (right) BDT output scores, for both data (black points) and simulated signal and background events (histograms). For the VH MET BDT, the γ+jet component of the background is derived directly from data. The statistical uncertainty in the data (simulation) is shown by the black error bars (shaded pink band). Events in the regions shaded grey are not considered for the VH leptonic categories. The full data set collected in the period 2016-2018 and the corresponding simulation is shown.

in the VH MET category, a data-driven approach is used to estimate this component of the background. The precise details of this approach are described in more detail in the context of the dijet BDT in the previous section. Figure 4.15 shows the output scores of the VH leptonic BDTs.

The WH leptonic category is sensitive to three kinematic regions in the STXS framework. Boundaries are placed on the reconstructed $p_T^{\gamma\gamma}$ at 75 and 150 GeV, to target the equivalent truth-level splittings in the p_T^V variable. Note, the reconstructed $p_T^{\gamma\gamma}$ provides the best handle on the truth-level p_T^V, since the presence of a neutrino in the final-state means that the vector boson cannot be fully reconstructed. No splitting of the ZH leptonic or the VH MET categories is performed. The final analysis categories are constructed by placing optimised boundaries on the respective BDT output scores. In total, two tags are defined for ZH leptonic, five for WH leptonic (two for both the $p_T^{\gamma\gamma} < 75$ GeV and $75 < p_T^{\gamma\gamma} < 150$ GeV regions, and one for the $p_T^{\gamma\gamma} > 150$ GeV region), and three for VH MET. This set of analysis categories are shown in the schematic in Figure **??**. The expected signal and background yields for the categories targeting VH leptonic production are shown in Table 4.8.

4.4.5 Top-associated categorisation

The two global categories with the highest priority in the sequence are constructed to be enriched with events from the tHq and ttH production modes, where at least one top quark in the event decays leptonically. In both, events are required to have at least one

Table 4.8: The expected number of events for $m_{\mathrm{H}} = 125$ GeV in the analysis categories targeting the VH leptonic production modes, shown for an integrated luminosity of 137 fb^{-1}. The fraction of the total number of events arising from each production mode in each analysis category is provided, as is the fraction of events originating from the targeted STXS bin (or bins). Here, ggH includes contributions from the sub-dominant ggZ(q$\bar{\mathrm{q}}$)H and bbH production modes, qqH includes both VBF and V(q$\bar{\mathrm{q}}$)H production, and "Top" represents both ttH and tH production together. The σ_{eff}, defined as the smallest interval containing 68.3% of the $m_{\gamma\gamma}$ distribution provides an indication of the mass resolution in each category. Also provided are the estimated number of background events-per-GeV in the signal peak region, the quantity $F_{68} = S_{68}/(S_{68}+B_{68})$, where S_{68} and B_{68} are the expected number of signal and background events in a $\pm 1\sigma_{\mathrm{eff}}$ window centred on m_{H}, respectively, and the approximate significance, $Z_{68} = S_{68}/\sqrt{S_{68}+B_{68}}$. The final column shows the significance for the targeted STXS bin (or bins) only, Z_{68}^{target}, where other Higgs boson signal events are considered as background.

Analysis categories	SM 125 GeV Higgs boson expected signal												
	Total	Target STXS bin(s)	Fraction of total events ggH	qqH	WH lep	ZH lep	ggZH lep	Top	σ_{eff} (GeV)	Bkg (GeV^{-1})	F_{68}	Z_{68}	Z_{68}^{target}
ZH lep Tag0	2.4	99.6%	-	-	-	82.0%	17.7%	0.4%	1.67	0.74	0.57	0.97	0.97
ZH lep Tag1	0.9	97.5%	0.1%	-	0.2%	80.7%	16.9%	2.2%	1.85	0.74	0.30	0.43	0.41
WH lep $p_T^V < 75$ Tag0	2.0	81.1%	-	0.2%	95.0%	3.3%	0.2%	1.3%	1.89	0.99	0.42	0.75	0.61
WH lep $p_T^V < 75$ Tag1	4.5	75.7%	2.6%	0.5%	87.2%	7.0%	0.3%	2.4%	1.85	7.5	0.18	0.74	0.56
WH lep $75 < p_T^V < 150$ Tag0	3.0	77.7%	0.7%	0.3%	93.2%	3.4%	0.8%	1.6%	1.94	0.89	0.54	1.04	0.81
WH lep $75 < p_T^V < 150$ Tag1	3.3	60.8%	1.7%	1.4%	83.1%	7.7%	1.6%	4.4%	2.02	2.4	0.31	0.83	0.51
WH lep $p_T^V > 150$ Tag0	3.5	79.9%	0.5%	0.4%	91.5%	3.6%	1.1%	2.8%	1.84	0.33	0.80	1.37	1.10
VH MET Tag0	2.2	97.9%	0.4%	0.9%	23.5%	56.9%	17.6%	0.8%	2.22	0.78	0.46	0.84	0.82
VH MET Tag1	3.6	90.5%	4.6%	3.1%	28.8%	46.0%	15.7%	1.9%	2.30	2.2	0.33	0.89	0.81
VH MET Tag2	6.6	72.2%	15.5%	8.8%	27.7%	33.5%	11.0%	3.5%	2.15	10	0.17	0.87	0.63

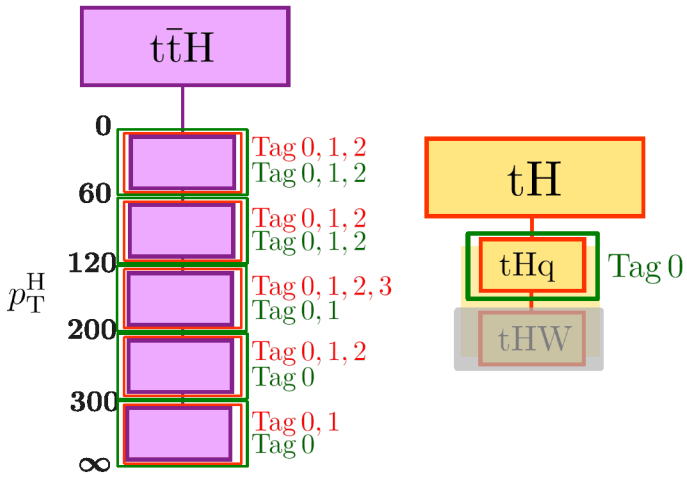


Figure 4.16: A schematic of the top-associated categorisation scheme. The hadronic and leptonic analysis categories are defined by the red and green boxes, respectively. The analysis categories surround the targeted STXS bin (or bins). The number of tags for each analysis region are highlighted in the plot. The shaded-grey region indicates there is no analysis category explicitly targeting the tHW production mode.

isolated charged lepton, as well as one or more jets. Moreover, the tHq leptonic selection requires an additional jet, tagged as originating from the hadronisation of a b quark.

Due to their similar final-state topology, a dedicated deep neural network, referred to as the Top DNN, is trained to differentiate between the tHq and ttH production modes. Figure 4.17 (top left) shows the output score of the Top DNN for tHq, ttH and the relevant SM background processes. A requirement is placed on this score in both the tHq leptonic and ttH leptonic categories to minimise the contamination from the opposing production mode. An additional discriminant is then trained in each global category using simulated events to reduce the contamination from other non-Higgs boson background processes. The simulated backgrounds include tt+$\gamma\gamma$, tt+γ+jet, tt+jets, γ+jets, V+γ, Drell-Yan, diboson and t+V production. The output scores for the so-called tHq leptonic BDT and ttH leptonic BDT are presented in Figure 4.17 (top right and bottom left, respectively).

The ttH leptonic global category is split using the reconstructed $p_T^{\gamma\gamma}$, with boundaries matching the truth-level p_T^H splittings of the ttH STXS bins. Boundaries on the ttH leptonic BDT score are then optimised separately in each region to maximise the expected sensitivity to the STXS bins: three tags are defined for the two bins with $p_T^H < 120$ GeV, two tags for the bin with $120 < p_T^H < 200$ GeV, and one tag for the two bins with $p_T^H > 200$ GeV. Due to the low expected tHq signal yield, a single tag is defined for tHq leptonic.

The ttH hadronic global category targets ttH production in which both top quarks decay hadronically; entering the category sequence after those which target VH leptonic production. Initially, events are required to have zero isolated leptons and three or more jets, where at least one of the jets is b tagged. Another discriminant, named the ttH hadronic BDT, is then trained to further suppress the contamination from non-Higgs boson SM backgrounds. The simulated backgrounds are the same as for the ttH leptonic BDT, except a data-driven technique, analogous to that described for the dijet BDT in Section 4.4.3, is used to achieve a better estimate of the γ+jet process. This improves both the description of the input features and provides a greater number of events with which to train the discriminant. Since the photon-ID BDT score is an input feature to the ttH hadronic BDT, each event in the data-driven training sample is given a new score value, randomly drawn from the respective distribution of simulated γ+jet events which pass the full set of selection criteria. The output score for the ttH hadronic BDT is shown in the bottom right of Figure 4.17.

Akin to the ttH leptonic categorisation, the ttH hadronic global category is split into five regions according to the reconstructed $p_T^{\gamma\gamma}$, and independently optimised boundaries are placed on the ttH hadronic BDT output score. Here, four tags are defined for the $120 < p_T^H < 200$ GeV STXS bin, three for the $p_T^H < 60$ GeV, $60 < p_T^H < 120$ GeV and

$200 < p_T^H < 300$ GeV bins, and two for the $p_T^H > 300$ GeV bin.

Table 4.9 presents the expected signal and background yields in each of the analysis categories targeting the top-associated Higgs boson production modes. This set of analysis categories are shown schematically in Figure 4.16.

4.4.6 Validation

It is necessary to validate the modelling of the large number of ML classifiers used in this analysis. In this context, validation means requiring a good agreement between data and MC simulation in the outputs of the numerous classifiers, particularly in the signal-like regions where events enter the analysis.

In Sections 4.4.2–4.4.5, the output-score distributions of all ML classifiers used in the event categorisation are shown for simulated signal and background events (used to train the ML classifiers), and for the corresponding data which can be compared to the simulated background. A good agreement here gives confidence that the background processes are accurately modelled, and therefore that the ML algorithm performs well in the classification task. Since the background model used in the final results extraction is derived directly from data, poor agreement in the background-like regions cannot introduce bias, but will only lead to a sub-optimal performance of the ML classifier.

A second form of validation is used for the signal-like regions. This involves finding an independent sample of events with signal-like characteristics and comparing the output score of the ML classifiers in both data and simulation. Since the signal model is derived from simulation, a disagreement in the classifier output may introduce a systematic uncertainty into the signal estimate. Therefore, we require a good agreement between data and simulation, within the statistical and systematic uncertainties in the MC, to instil confidence that the simulated signal events are sufficiently well-modelled.

The plots in Figure 4.18 show the validation of the ggH BDT, the diphoton BDT, the dijet BDT (via p_{VBF}), and the VH hadronic BDT, all using $\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ events in simulation and data. The uncertainty band on the simulation demonstrates the addition in quadrature of the statistical and systematic components, where the systematic component includes uncertainties originating from the photon-ID BDT, photon-energy resolution, and the jet-energy scale and resolution (see Section 5.4). The ggH BDT distribution shows the number of events attributed to each output class. It can be seen that there is an excellent agreement between data and simulation, well within uncertainties, particularly in the signal regions which enter the analysis categories. All in all, this means the included systematic uncertainties in the signal estimates are sufficient to cover the discrepancies in the output classifier scores, and no additional sources of systematic uncertainty are required.

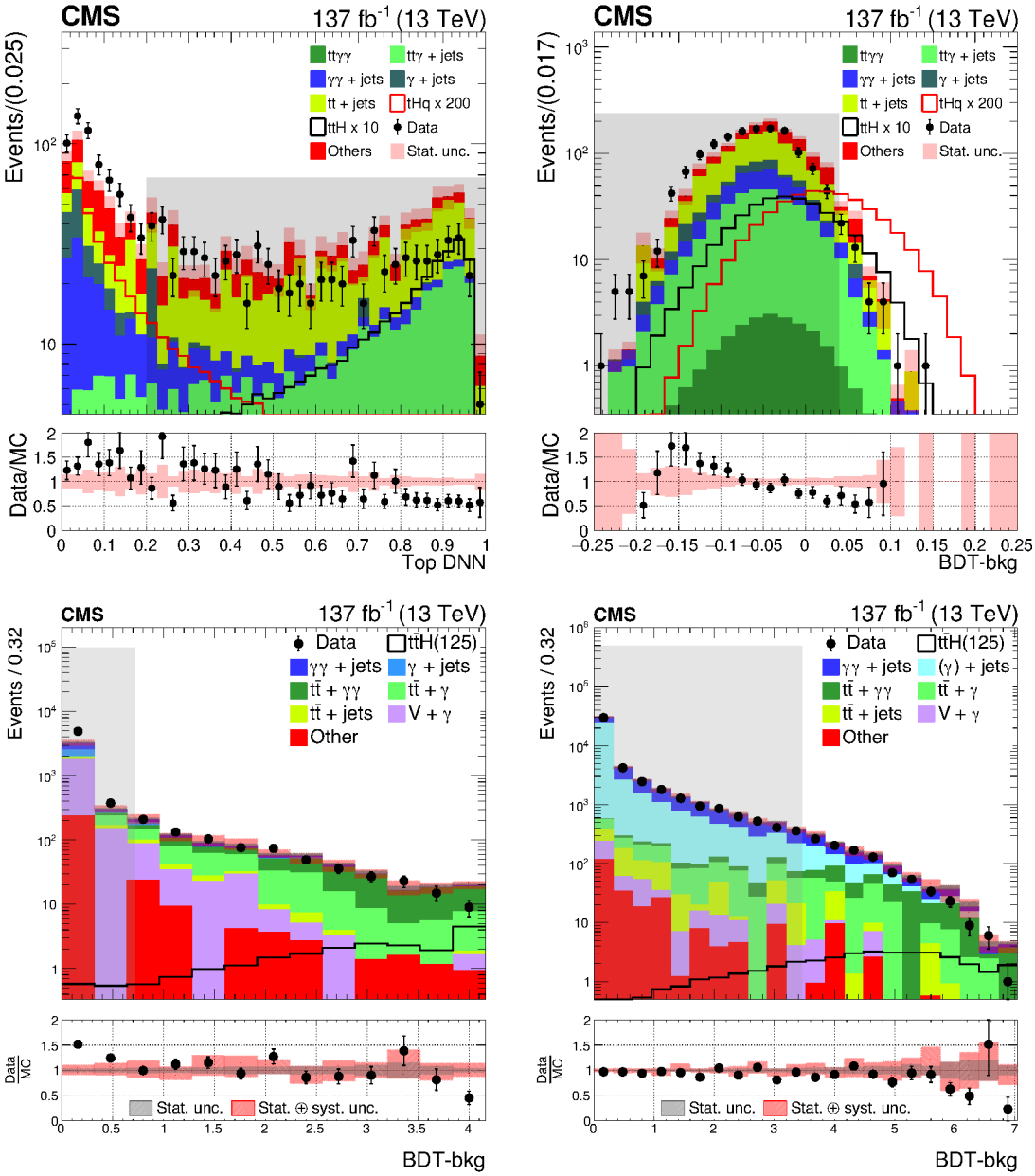


Figure 4.17: Distributions of the Top DNN (top left), tHq leptonic BDT (top right), ttH leptonic BDT (bottom left), and ttH hadronic BDT (bottom right) output scores, for both data and simulated signal and background events. The data events are taken from the $m_{\gamma\gamma}$ sidebands: $[100, 120] \cup [130, 180]$ GeV. The statistical (statistical $\oplus$ systematic) component of the uncertainty in the background estimate is shown by the red shaded band in the top two (bottom two) score distributions. The grey shaded region for the Top DNN score indicates the upper threshold for the tHq leptonic category. For the other output scores, the shaded regions illustrate the lowest threshold, beyond which events are not considered for the corresponding analysis categories. The full data set collected in the period 2016-2018 and the corresponding simulation is shown.

Table 4.9: The expected number of events for $m_{\mathrm{H}} = 125$ GeV in the analysis categories targeting the top-associated production modes, shown for an integrated luminosity of 137 fb^{-1}. The fraction of the total number of events arising from each production mode in each analysis category is provided, as is the fraction of events originating from the targeted STXS bin (or bins). Here, ggH includes contributions from the sub-dominant ggZ(q$\bar{\mathrm{q}}$)H and bbH production modes, and qqH includes both VBF and V(q$\bar{\mathrm{q}}$)H production. The σ_{eff}, defined as the smallest interval containing 68.3% of the $m_{\gamma\gamma}$ distribution provides an indication of the mass resolution in each category. Also provided are the estimated number of background events-per-GeV in the signal peak region, the quantity $F_{68} = S_{68}/(S_{68}+B_{68})$, where S_{68} and B_{68} are the expected number of signal and background events in a $\pm1\sigma_{\mathrm{eff}}$ window centred on m_{H}, respectively, and the approximate significance, $Z_{68} = S_{68}/\sqrt{S_{68}+B_{68}}$. The final column shows the significance for the targeted STXS bin (or bins) only, Z_{68}^{target}, where other Higgs boson signal events are considered as background.

Analysis categories	SM 125 GeV Higgs boson expected signal									Bkg	F_{68}	Z_{68}	Z_{68}^{target}
	Total	Target	Fraction of total events						σ_{eff}	(GeV^{-1})			
		STXS bin(s)	ggH	qqH	VH lep	ttH	tHq	tHW	(GeV)				
tHq lep	1.8	23.9%	3.5%	3.7%	34.0%	28.8%	23.9%	6.0%	1.62	1.1	0.40	0.70	0.17
ttH lep $p_T^{\gamma\gamma} < 60$ Tag0	0.8	93.8%	-	-	0.7%	98.2%	0.7%	0.5%	1.71	0.19	0.63	0.58	0.55
ttH lep $p_T^{\gamma\gamma} < 60$ Tag1	1.0	94.4%	-	-	0.5%	97.9%	1.5%	0.7%	1.69	0.38	0.51	0.58	0.55
ttH lep $p_T^{\gamma\gamma} < 60$ Tag2	1.8	87.7%	-	0.5%	5.1%	90.7%	3.2%	1.1%	1.94	3.0	0.17	0.46	0.40
ttH lep $60 < p_T^{\gamma\gamma} < 120$ Tag0	1.4	95.0%	-	-	1.0%	97.3%	1.0%	0.8%	1.60	0.34	0.64	0.78	0.75
ttH lep $60 < p_T^{\gamma\gamma} < 120$ Tag1	0.6	90.8%	-	0.7%	1.0%	95.6%	1.6%	1.1%	1.61	0.21	0.55	0.49	0.44
ttH lep $60 < p_T^{\gamma\gamma} < 120$ Tag2	2.1	90.9%	-	0.1%	2.8%	93.7%	2.5%	1.3%	1.92	1.2	0.39	0.74	0.67
ttH lep $120 < p_T^{\gamma\gamma} < 200$ Tag0	3.6	90.1%	0.3%	0.2%	2.7%	92.8%	2.0%	2.0%	1.63	0.59	0.71	1.31	1.18
ttH lep $120 < p_T^{\gamma\gamma} < 200$ Tag1	0.8	77.9%	2.0%	0.5%	11.3%	80.6%	3.2%	2.5%	1.72	0.44	0.43	0.50	0.39
ttH lep $200 < p_T^{\gamma\gamma} < 300$ Tag0	2.5	85.9%	0.1%	-	4.1%	88.1%	3.0%	4.8%	1.54	0.51	0.68	1.08	0.93
ttH lep $p_T^{\gamma\gamma} > 300$ Tag0	2.1	61.7%	1.0%	-	18.0%	69.3%	3.0%	8.7%	1.57	0.53	0.64	0.96	0.59
ttH had $p_T^{\gamma\gamma} < 60$ Tag0	1.2	94.2%	1.7%	0.2%	-	96.6%	0.9%	0.4%	1.68	0.50	0.49	0.64	0.60
ttH had $p_T^{\gamma\gamma} < 60$ Tag1	0.4	93.5%	0.1%	0.9%	-	96.7%	1.7%	0.6%	1.66	0.26	0.38	0.31	0.29
ttH had $p_T^{\gamma\gamma} < 60$ Tag2	3.1	89.8%	1.6%	1.5%	0.3%	92.9%	3.0%	0.7%	1.88	6.6	0.14	0.54	0.49
ttH had $60 < p_T^{\gamma\gamma} < 120$ Tag0	1.8	92.6%	0.6%	-	0.1%	97.6%	1.1%	0.6%	1.55	0.24	0.77	0.97	0.90
ttH had $60 < p_T^{\gamma\gamma} < 120$ Tag1	0.4	90.8%	4.6%	0.8%	-	91.9%	1.9%	0.8%	1.35	0.33	0.39	0.33	0.30
ttH had $60 < p_T^{\gamma\gamma} < 120$ Tag2	5.2	88.7%	1.0%	2.2%	0.5%	91.8%	3.5%	1.0%	1.90	6.6	0.22	0.88	0.78
ttH had $120 < p_T^{\gamma\gamma} < 200$ Tag0	3.6	91.4%	1.5%	0.4%	0.1%	94.7%	2.2%	1.3%	1.53	0.87	0.65	1.25	1.14
ttH had $120 < p_T^{\gamma\gamma} < 200$ Tag1	2.1	83.3%	4.6%	2.9%	0.5%	86.2%	4.2%	1.7%	1.76	1.3	0.38	0.74	0.61
ttH had $120 < p_T^{\gamma\gamma} < 200$ Tag2	1.7	74.3%	10.0%	4.6%	0.6%	76.5%	6.3%	2.0%	1.65	1.9	0.26	0.55	0.41
ttH had $120 < p_T^{\gamma\gamma} < 200$ Tag3	2.6	62.2%	15.4%	8.4%	1.2%	64.7%	8.5%	1.9%	1.73	6.6	0.13	0.49	0.30
ttH had $200 < p_T^{\gamma\gamma} < 300$ Tag0	2.0	90.1%	0.5%	0.4%	0.1%	92.3%	3.8%	2.9%	1.44	0.37	0.72	1.00	0.90
ttH had $200 < p_T^{\gamma\gamma} < 300$ Tag1	1.5	74.6%	8.8%	3.1%	0.7%	77.0%	6.8%	3.5%	1.47	0.59	0.54	0.74	0.55
ttH had $200 < p_T^{\gamma\gamma} < 300$ Tag2	1.7	56.5%	18.8%	8.4%	0.4%	58.0%	10.5%	3.8%	1.59	1.8	0.30	0.59	0.33
ttH had $p_T^{\gamma\gamma} > 300$ Tag0	2.5	73.8%	8.3%	1.6%	0.8%	74.9%	7.7%	6.8%	1.44	0.39	0.75	1.13	0.84
ttH had $p_T^{\gamma\gamma} > 300$ Tag1	1.9	45.6%	27.1%	7.3%	1.4%	46.0%	11.4%	6.7%	1.56	0.74	0.52	0.82	0.37

Figure 4.19 shows the equivalent plots for the ttH leptonic BDT and ttH hadronic BDT classifiers; here ttZ, $Z \to e^+e^-$ events are used for the validation as they share similar kinematic properties with the targeted ttH signal process. Again, good agreement is observed between data and simulation, particularly in the signal-like regions with high BDT output score.

4.5 Summary

This chapter has described the selection, reconstruction and subsequent categorisation of p-p collision events recorded by the CMS detector in order to measure Higgs boson production cross sections and couplings in the $H \to \gamma\gamma$ decay channel. The confusion matrix in Figure 4.20 displays the expected signal-composition of the final *reconstruction-level* analysis categories in terms of the *truth-level* STXS bins. In the plot, the analysis categories targeting a common STXS region (tags) are summed, such that the signal compositions of the individual analysis categories are weighted according to $F_{68} = S_{68}/(S_{68} + B_{68})$. The numbers in the matrix represent the fractional contribution of the total signal yield in a given analysis category group arising from each STXS bin, such that each row sums to 100%. Ultimately, by making this matrix as diagonal as possible, we improve the sensitivity to the individual STXS regions, and reduce the correlations between the measured quantities. The full unweighted confusion matrix, split into the individual analysis categories, is provided in Appendix D.

The next chapter explains the procedure for unfolding this matrix by fitting the diphoton mass distributions in each analysis category, in order to extract the measurements of the truth-level production cross sections and other parameters of interest.

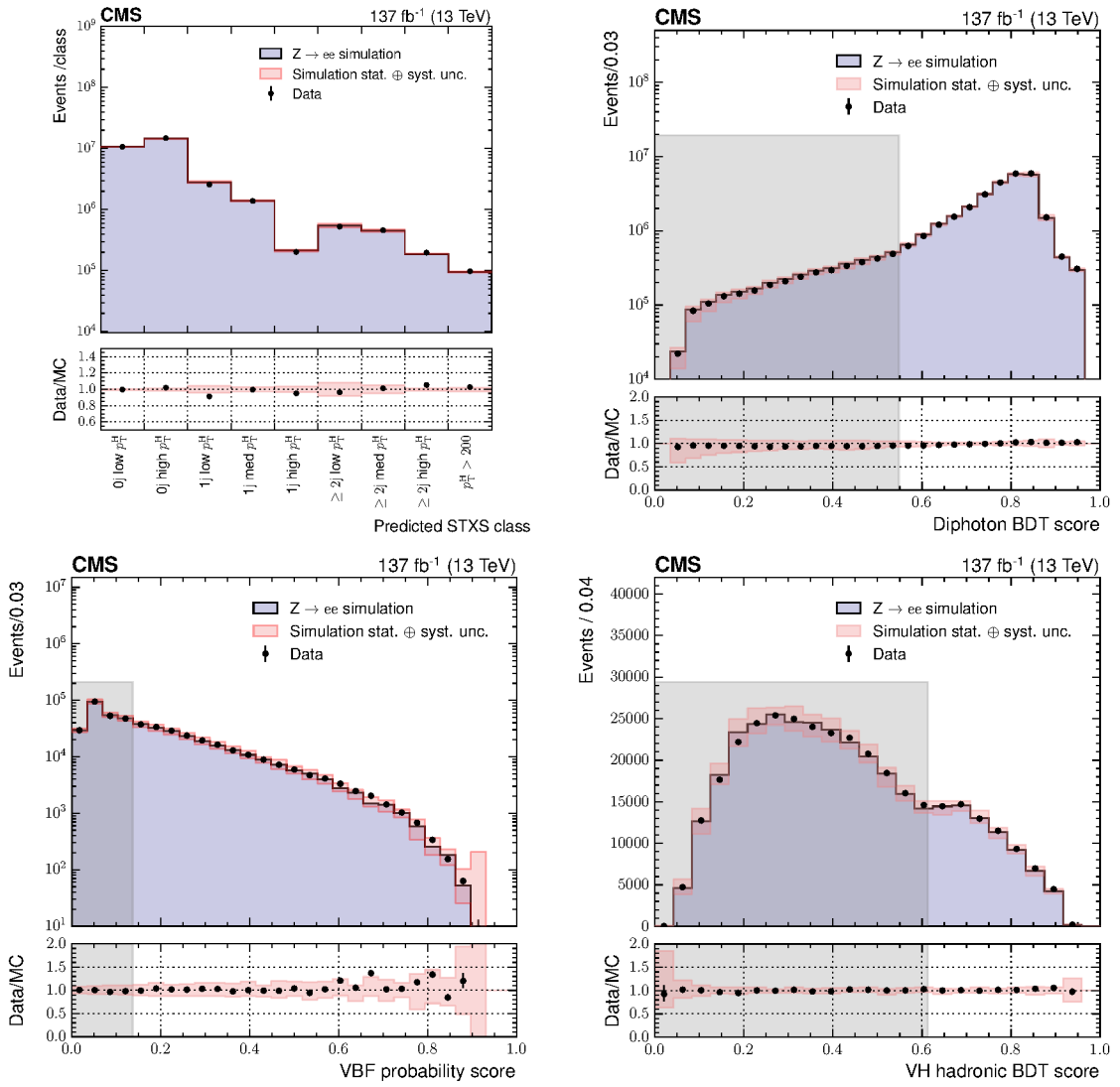


Figure 4.18: The validation of the ggH BDT (top left), diphoton BDT (top right), dijet BDT via p_{VBF} (bottom left) and the VH hadronic BDT (bottom right), where the distribution of the outputs are shown for $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ events in data (black points) and simulation (filled histogram). The pink bands represent the quadrature sum of the statistical and systematic uncertainties in the simulated events. The regions shaded grey show the thresholds in the scores, below which no events enter the relevant analysis categories.

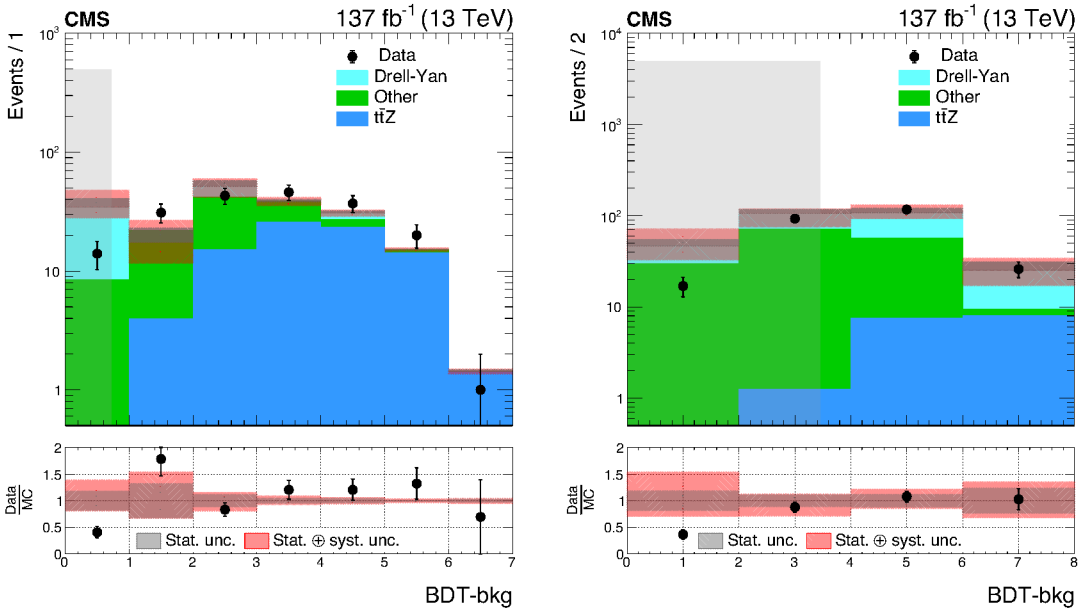


Figure 4.19: The validation of the ttH leptonic BDT (left) and ttH hadronic BDT (right) classifiers, where the distribution of the outputs are shown for ttZ events in data (black points) and simulation (filled histogram). The ttZ events are required to pass the same initial selection criteria for the corresponding global category, with additional requirements on the dielectron kinematics, number of jets, and number of b tagged jets to increase the ttZ purity. The pink (black) bands represent the quadrature sum of the statistical and systematic uncertainties (statistical uncertainty only) in the simulated background events. The shaded-grey regions indicate the scores which do not enter the relevant analysis categories.

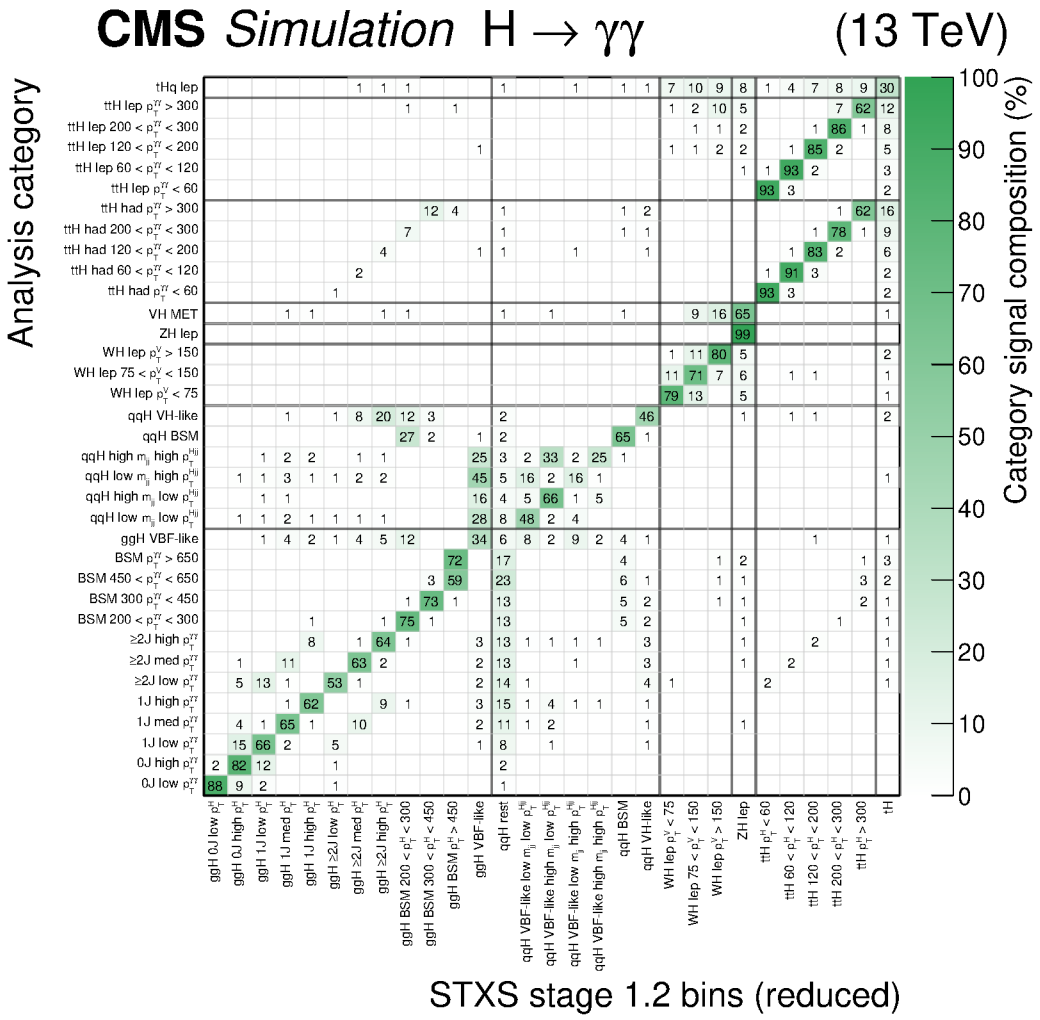


Figure 4.20: Confusion matrix displaying the composition of the analysis categories in terms of a merged set of STXS bins. The granularity of the STXS bin merging corresponds to the finest granularity used for the cross section measurements. Analysis categories targeting a common STXS region are summed, where the signal compositions of the individual analysis categories are weighted in the sum by $F_{68} = S_{68}/(S_{68} + B_{68})$. The colour scale indicates the fractional yield in each analysis category group (rows) accounted for by each STXS process (columns). Each row therefore sums to 100%. Entries with values less than 0.5% are not shown. Simulated events for each year in the period 2016-2018 are combined with appropriate weights corresponding to their relative integrated luminosity in data. The column labelled as qqH rest includes the contributions from the qqH 0J, qqH 1J, qqH $m_{jj} < 60$ GeV and the qqH $120 < m_{jj} < 350$ GeV STXS bins.

Chapter 5

Statistical inference

5.1 Introduction

The statistical methodology used to extract the results follows the procedure developed by the ATLAS and CMS Collaborations, documented in Ref. [143]. The following section uses the definition,

$$\text{Poisson}\Big(n \,\Big|\, s+b\Big) = \frac{(s+b)^n}{n!} \cdot e^{-(s+b)}. \tag{5.1}$$

5.1.1 Construction of the likelihood

A simultaneous binned maximum-likelihood fit is performed to the $m_{\gamma\gamma}$ distributions of all analysis categories. This requires the construction of a likelihood function for each analysis category, k, of the form,

$$\begin{aligned} &L_k(\text{data}\,|\,\mu^{i,\gamma\gamma}, m_{\text{H}}, \vec{\theta}_s, \vec{\theta}_b) = \\ &\prod_{X=1}^{N_{\text{bins}}=320} \text{Poisson}\Big(N_{k,X}^{\text{data}} \,\Big|\, \Big[\sum_i S_{k,X}^{i,\gamma\gamma}(\mu^{i,\gamma\gamma}, m_{\text{H}}, \vec{\theta}_s)\Big] + B_{k,X}(\vec{\theta}_b)\Big), \end{aligned} \tag{5.2}$$

where the index X runs over bins in the $m_{\gamma\gamma}$ distribution in the range $100 < m_{\gamma\gamma} < 180$ GeV with a bin width of 250 MeV; this choice is sufficiently small compared to the diphoton-mass resolution to ensure that a negligible amount of information is lost. The likelihood itself is a function of the signal parameters, $\mu^{i,\gamma\gamma}$, the Higgs boson mass, m_{H}, and nuisance parameters, $\vec{\theta} = \{\vec{\theta}_s, \vec{\theta}_b\}$, which account for systematic uncertainties in the signal and background estimates. The nuisance parameters are grouped according to their effect, as shown in equation 5.3,

$$\{\vec{\theta}_s, \vec{\theta}_b\} = \{\vec{\theta}_s^{\,\mathrm{th}}, \vec{\theta}_s^{\,\epsilon,\mathrm{th}}, \vec{\theta}_s^{\,\epsilon,\mathrm{exp}}, \vec{\theta}_s^{\,\mathrm{shape}}, \vec{\theta}_s^{\,\mathrm{lumi}}, \vec{\theta}_b^{\,\mathrm{shape}}, \vec{\theta}_b^{\,\mathrm{discrete}}\}, \tag{5.3}$$

where $\vec{\theta}_s^{\,\mathrm{th}}$ are the uncertainties in the SM prediction of the cross sections times branching fraction, $[\sigma^i \cdot \mathcal{B}^{\gamma\gamma}]_{\mathrm{SM}}$. The $\vec{\theta}_s^{\,\epsilon,\mathrm{th}}$ and $\vec{\theta}_s^{\,\epsilon,\mathrm{exp}}$ terms correspond to systematic uncertainties in the efficiency-times-acceptance of the final analysis categories, originating from theoretical and experimental sources, respectively. Nuisance parameters affecting the shape of the analytic signal-model, described in more detail in Section 5.2, are labelled by $\vec{\theta}_s^{\,\mathrm{shape}}$. The uncertainties in the luminosity are referred to as $\vec{\theta}_s^{\,\mathrm{lumi}}$; these affect only the signal estimate as the background estimate is derived directly from data. Finally, the background-model shape parameters are labelled as $\vec{\theta}_b^{\,\mathrm{shape}}$, whilst the discrete nuisance parameters corresponding to the uncertainty in the choice of background function are labelled as $\vec{\theta}_b^{\,\mathrm{discrete}}$. More detail concerning the individual sources of the systematic uncertainties is provided in Section 5.4.

In the Poisson term of the likelihood, $N_{k,X}^{\mathrm{data}}$, corresponds to the number of data events in bin X of category k, and $S_{k,X}^{i,\gamma\gamma}$ and $B_{k,X}$ are the signal and background estimates in the same bin. The index, i, labels the particular *signal process*, which in this analysis corresponds to the STXS bins. Equation 5.4 shows the total signal-yield for process i in analysis category k, integrated over all $m_{\gamma\gamma}$ bins,

$$S_k^{i,\gamma\gamma} = \mu^{i,\gamma\gamma} \times [\sigma^i \cdot \mathcal{B}^{\gamma\gamma}]_{\mathrm{SM}}(m_{\mathrm{H}}, \vec{\theta}_s^{\,\mathrm{th}}) \times \epsilon_k^{i,\gamma\gamma}(m_{\mathrm{H}}, \vec{\theta}_s^{\,\epsilon,\mathrm{th}}, \vec{\theta}_s^{\,\epsilon,\mathrm{exp}}) \times \mathcal{L}(\vec{\theta}_s^{\,\mathrm{lumi}}). \tag{5.4}$$

Here $[\sigma^i \cdot \mathcal{B}^{\gamma\gamma}]_{\mathrm{SM}}$ is the SM prediction for the cross section times branching fraction for process i, listed in Tables 2.5–2.8 for $m_{\mathrm{H}} = 125.0$ GeV. The product of the detector efficiency and the analysis acceptance is represented by $\epsilon_k^{i,\gamma\gamma}$, which effectively encodes the fraction of the total yield of process i landing in analysis category k. The luminosity estimate is represented by $\mathcal{L}$. The signal parameters, $\mu^{i,\gamma\gamma}$, define the *parameters of interest* (POI). For example, when measuring cross sections in the STXS framework,

$$\mu^{i,\gamma\gamma} = \frac{[\sigma^i \cdot \mathcal{B}^{\gamma\gamma}]_{\mathrm{obs}}}{[\sigma^i \cdot \mathcal{B}^{\gamma\gamma}]_{\mathrm{SM}}(m_{\mathrm{H}}, \vec{\theta}_s^{\,\mathrm{th}})}. \tag{5.5}$$

In this signal parametrisation, the theory systematic uncertainties, $\vec{\theta}_s^{\,\mathrm{th}}$, in the denominator cancel out the same terms in equation 5.4. As a result, $\vec{\theta}_s^{\,\mathrm{th}}$ do not enter the cross section measurements, but are instead attributed to the uncertainty in the SM predictions of the measured quantities. This property of the measurements has the benefit that they remain useful in the long-term, as they can accommodate future improvements in the SM theoretical predictions.

Other signal parametrisations are considered. The most-constraining fit that is per-

formed introduces a single inclusive signal-strength modifier, μ, which scales all Higgs boson signal processes equally. Adding more degrees of freedom, the per-production-mode signal-strength parametrisation defines four POIs: μ_{ggH}, μ_{VBF}, μ_{VH} and μ_{top}, which act as global scaling factors for the corresponding Higgs boson production modes. The κ-framework [144] replaces $\mu^{i,\gamma\gamma}$ with functions of Higgs boson coupling-modifiers (κ-parameters), $\mu^{i,\gamma\gamma}(\vec{\kappa})$, where the form of the function depends on the signal process i (see Section 6.4)[1]. In such *interpretations*, there is no cancellation of $\vec{\theta}_s^{\text{th}}$, meaning these nuisance parameters are directly folded into the measurements. In general, we can write $\mu^{i,\gamma\gamma}$ as a function of some set of parameters of interest, $\vec{\alpha}$,

$$\mu^{i,\gamma\gamma} \equiv \mu^{i,\gamma\gamma}(\vec{\alpha}). \tag{5.6}$$

To determine $S_{k,X}^{i,\gamma\gamma}$ (i.e the subset of $S_k^{i,\gamma\gamma}$ that falls in bin X of the $m_{\gamma\gamma}$ distribution), it is necessary to model the functional form of the signal peak in the diphoton invariant-mass distribution. Analytic models are constructed for each process i in each reconstructed analysis category k, using simulated events. More information regarding the signal modelling is provided in Section 5.2.

The background model is derived directly from the observed diphoton invariant-mass distribution in data. Described in more detail in Section 5.3, the form of the analytic model in each analysis category is treated as a discrete nuisance parameter in the fit, with options coming from a number of different functions. The background estimate in bin X, $B_{k,X}$, is calculated using the analytic background model.

The total likelihood function is defined as the product over all 80 per-category likelihoods, multiplied by a constraint term, $\mathcal{C}$,

$$L(\text{data}\,|\,\vec{\alpha}, m_{\text{H}}, \vec{\theta}) = \prod_{k=1}^{N_{\text{cats}}=80} \Big[L_k(\text{data}\,|\,\vec{\alpha}, m_{\text{H}}, \vec{\theta}) \Big] \times \mathcal{C}(\vec{\theta}), \tag{5.7}$$

where the signal parameters, $\mu^{i,\gamma\gamma}$, have been expressed in terms of the general parameters of interest, $\vec{\alpha}$, and $\vec{\theta} = \{\vec{\theta}_s, \vec{\theta}_b\}$. In all fits, the Higgs boson mass, m_{H}, is fixed to its most precisely measured value of 125.38 GeV [49]. This ensures all measurements are reported with respect to the theoretical predictions consistent with the best available knowledge of m_{H}. It should be made clear that despite fixing m_{H}, the mean position of the signal models can vary from 125.38 GeV due to the inclusion of the $\vec{\theta}_s^{\text{shape}}$ nuisance parameters, which describe the uncertainties in the photon energy measurements (see Section 5.4.2). In other words, the experimental uncertainty in the measured Higgs boson mass is folded

[1] Looking ahead to Section 7, here the signal yields are parametrised in an EFT framework, $\mu^{i,\gamma\gamma}(w_p)$, so that the same statistical procedure can be used to extract constraints on the EFT Wilson coefficients, w_p.

into the fit. All in all, the fixing of the Higgs boson mass means the dependence of the likelihood on m_{H} is absorbed into the $\vec{\theta}_s^{\,\mathrm{shape}}$ nuisance parameters.

The constraint term, $\mathcal{C}$, applies a penalty for deviations from the expected values of the nuisance parameters. The form of this penalty depends on the choice of probability density function (pdf) for a given nuisance. In this analysis all nuisance parameters affecting the signal-shape are associated with a Gaussian pdf, whilst those affecting the signal-yield estimates are associated with a log-normal pdf. The latter is useful for avoiding the possibility of the signal-yield becoming negative. For nuisance parameters with small impacts on the signal-yields, the log-normal and Gaussian pdfs are practically identical. If necessary, nuisance parameters with larger impacts can be modelled with an asymmetric log-normal distribution (see Section 5.4.2) to avoid distorting the signal-yield estimates[2]. The background nuisance parameters, $\vec{\theta}_b$, are instead associated with a flat pdf, since there is no a-priori knowledge of their values, and therefore changes in their value are not explicitly penalised by the constraint term. However, an additional penalty is included according to the total number of degrees of freedom in the background-model function to penalise unnecessarily complex functions.

5.1.2 Extraction of results

In practice, the fit is performed by minimising the value of $-2\ln L(\mathrm{data}\,|\,\vec{\alpha},\vec{\theta})$. This is done numerically using the RooFit software package [145]. The values of the parameters of interest which minimise this quantity are described as the "best-fit" values, and are labelled as the point in the parameter space, $\hat{\vec{\alpha}}$. The values of the nuisance parameters at this point, $\hat{\vec{\theta}}$, are referred to as the unconditional maximum-likelihood estimates of $\vec{\theta}$.

To calculate the confidence intervals for the parameters of interest, a profile likelihood test statistic, $q(\vec{\alpha})$ is constructed as shown in equation 5.8,

$$q(\vec{\alpha}) = -2\ln\left(\frac{L(\mathrm{data}\,|\,\vec{\alpha},\hat{\vec{\theta}}_{\vec{\alpha}})}{L(\mathrm{data}\,|\,\hat{\vec{\alpha}},\hat{\vec{\theta}})}\right). \tag{5.8}$$

The quantity $\hat{\vec{\theta}}_{\vec{\alpha}}$ corresponds to the conditional maximum-likelihood estimates of the nuisance parameters, for fixed values of the parameters of interest, $\vec{\alpha}$. In the asymptotic regime, the value of $q(\vec{\alpha})$ follows a χ^2_n distribution with n degrees of freedom, where n is the dimensionality of the parameters of interest vector, $\vec{\alpha}$ [146]. Confidence intervals are then derived for some confidence level, $1-p$, as the regions in $\vec{\alpha}$ for which the value of $q(\vec{\alpha})$ is below $F^{-1}_{\chi^2_n}(1-p)$, where $F_{\chi^2_n}$ is the cumulative function of the χ^2_n distribution.

For one-dimensional measurements, such as the signal-strength and cross section fits,

[2] Asymmetric log-normal distributions can look significantly more Gaussian.

the 68% and 95% confidence intervals are defined by the union of intervals for which $q(\alpha) < 0.99$ and $q(\alpha) < 3.84$, respectively. In the case where there are multiple parameters of interest in the signal parametrisation, the intervals are determined by treating the other parameters as nuisance parameters i.e. profiling them in the minimisation. In practice, for each parameter of interest, α, the minimisation is performed for a discrete set of points, and the full $q(\alpha)$ function is determined by interpolating between these points. The number of points is chosen to sufficiently cover the shape of the $q(\alpha)$ function.

For two-dimensional measurements, such as those performed in the κ-framework (see Figure 6.17), the 68% and 95% confidence regions are defined by the set of parameter values for which $q(\alpha_1, \alpha_2) < 2.30$ and $q(\alpha_1, \alpha_2) < 5.99$, respectively. Again, the full $q(\alpha_1, \alpha_2)$ surface is determined by performing the numerical minimisation for a discrete grid of parameter points, (α_1, α_2), and interpolating between these values.

It is possible to compute the compatibility of the results with the SM hypothesis, by evaluating the test-statistic, $q(\vec{\alpha}_{\text{SM}})$, at the point in parameter space where all parameters of interest take their SM expected values, $\vec{\alpha}_{\text{SM}}$. The probability of compatibility is expressed as a p-value, p_{SM}, computed as,

$$p_{\text{SM}} = 1 - F_{\chi^2_n}\Big(q(\vec{\alpha}_{\text{SM}})\Big). \tag{5.9}$$

The correlations between the fitted parameters are derived under the assumption of symmetric uncertainties about the best-fit parameter point, $\hat{\vec{\alpha}}$, by using the second derivatives of $q(\vec{\alpha})$. In practice this is performed numerically by stepping around the $q(\vec{\alpha})$ minimum to estimate the Hessian matrix, from which the correlation coefficients between the fitted parameters can be extracted. Providing the correlation coefficients in addition to the best-fit values and uncertainties, enables the future re-interpretation of the cross section measurements in terms of other signal parametrisations. Nevertheless, this only serves as an approximation of the full likelihood surface, since the observed uncertainties in the measurements are asymmetric. Ultimately, this fact motivates experiments to perform interpretations *in-house*, as they have access to the full likelihood surface.

Finally, in addition to the observed results, it is useful to compute the results one would expect to obtain given the SM hypothesis. These so-called *expected results* are determined by replacing the observed data with an Asimov toy data set, in which all parameters take their SM expected values and all statistical fluctuations are suppressed [146].

5.2 Signal modelling

The analytic signal model, derived using MC simulated events, is constructed to fit the $m_{\gamma\gamma}$ spectrum of each STXS bin in each analysis category: $\mathcal{S}_k^{i,\gamma\gamma}(m_{\gamma\gamma}; m_{\text{H}}, \vec{\theta}_s^{\text{shape}})$. The

distribution of events depends on whether the selected vertex (Section 4.3.2) is correctly identified within 1 cm of the true diphoton vertex. For this reason, the right-vertex (RV) and wrong-vertex (WV) scenarios, defined according to the 1 cm threshold, are considered separately when building the signal shape. The final model is the weighted sum of the RV and WV contributions, where f_{RV} is the fraction of simulated events with the selected vertex within 1 cm of the true vertex, calculated separately for each (i,k) combination,

$$\mathcal{S}_k^{i,\gamma\gamma} = f_{\mathrm{RV}} \cdot \mathcal{S}_{k,\mathrm{RV}}^{i,\gamma\gamma} + (1 - f_{\mathrm{RV}}) \cdot \mathcal{S}_{k,\mathrm{WV}}^{i,\gamma\gamma}. \tag{5.10}$$

To account for the variation in detector performance, the signal models are constructed separately for each data-taking year i.e. using the independent MC samples which correspond to the 2016, 2017 and 2018 data-taking conditions. In this approach, the variation in the diphoton mass resolution is incorporated into the model, and year-dependent systematic uncertainties in the signal estimate can be propagated to the final fit. The index i used in Section 5.1.1, is effectively extended to label each signal process from each data-taking year e.g. (ggH 0J high p_T^H, 2016). This is such that the efficiency-times-acceptance factor in equation 5.4, $\epsilon_k^{i,\gamma\gamma}$, is derived separately for each year and the luminosity estimate, $\mathcal{L}$, takes the relevant year-dependent value: 35.9 fb^{-1} for 2016, 41.5 fb^{-1} for 2017, and 59.4 fb^{-1} for 2018. Clearly, the signal parametrisation, $\mu^{i,\gamma\gamma}$, and theory predictions, $[\sigma^i \cdot \mathcal{B}^{\gamma\gamma}]_{\mathrm{SM}}$, are the same for each year.

Each model, $\mathcal{S}_{k,V}^{i,\gamma\gamma}$, consists of a sum of up to five Gaussians, where $V = \{\mathrm{RV}, \mathrm{WV}\}$ labels the vertex scenario. The parameters of each Gaussian function, namely the mean, width, and the relative contribution to the total model are extracted by performing a fit to the $m_{\gamma\gamma}$ spectrum of simulated events, generated with a nominal Higgs boson mass of $m_{\mathrm{H}} = 125.0$ GeV. Since m_{H} is fixed to 125.38 GeV in the results extraction, the Gaussian function mean values are shifted up by 380 MeV. On the other hand, the width and relative contributions of each Gaussian are assumed to be constant with respect to m_{H}. This approach relies on the fact that the variation in the signal shape is small when moving from 125.0 to 125.38 GeV and is well covered by the shape systematic uncertainties, $\vec{\theta}_s^{\,\mathrm{shape}}$, introduced in Section 5.4.2.

The number of Gaussian functions to fit each (i,k,V) combination depends on the shape of the $m_{\gamma\gamma}$ distribution, and is selected as that which minimises the χ^2/n_{dof}, where n_{dof} is equal to the number of $m_{\gamma\gamma}$ bins (with at least one event) minus the number of shape parameters in the fitted function. If the choice is observed to over-fit statistical fluctuations in the simulation, then the number of Gaussians is reduced. Figure 5.1 shows fits with a different number of Gaussians for ggH 0J high p_T^H events from the 2016 simulation, in the 0J high $p_T^{\gamma\gamma}$ Tag0 analysis category, for the RV (left) and WV (right) scenarios. In this case, the optimal choice which minimises the χ^2/n_{dof} is 5 Gaussians for

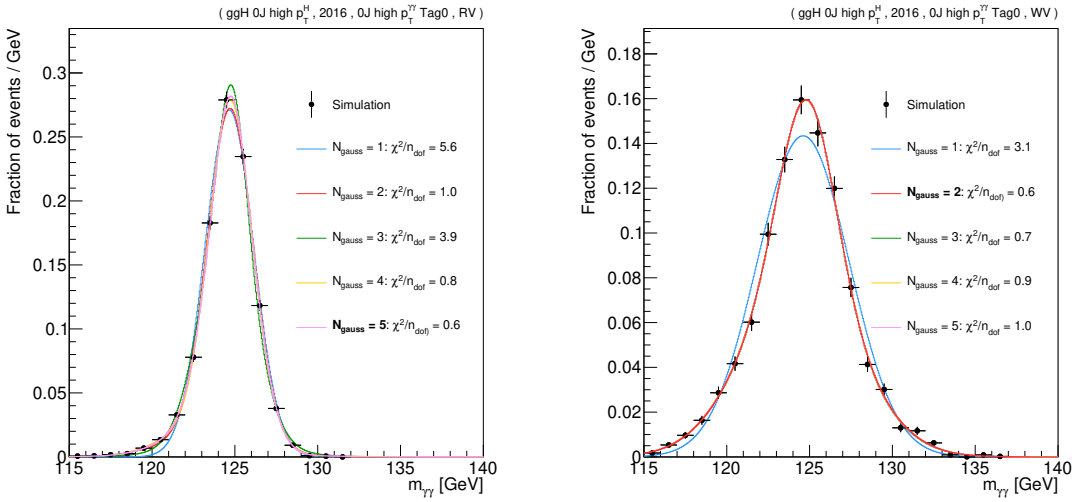


Figure 5.1: Finding the optimal number of Gaussian functions to fit the signal peak for ggH 0J high p_T^H events in the 0J high $p_T^{\gamma\gamma}$ Tag0 category, from 2016 simulation with $m_{\mathrm{H}} = 125.0$ GeV. Events in the RV and WV scenarios are shown in the left and right plots, respectively. Up to five Gaussian functions are trialled for each vertex scenario. The optimal choices are 5 Gaussians for the RV events and 2 Gaussians for the WV events.

RV events, and 2 Gaussians for WV events. The final models, $\mathcal{S}_{k,V}^{i,\gamma\gamma}$, decomposed into the contributions from the individual Gaussian functions, are shown in Figure 5.2.

For a number of (i,k,V) combinations, particularly those corresponding to the far off-diagonal elements in Figure 4.20, there is an insufficient number of simulated events to accurately model the signal shape. In this case, the shape is replaced with the shape of the STXS bin with the highest yield in analysis category k. This replacement is motivated by the fact that events in the same region of experimental phase space i.e subject to the same selection criteria, tend to have similar values of the diphoton mass resolution.

The signal models are normalised according to equation 5.4. The SM predictions of the cross sections and the $\mathrm{H} \to \gamma\gamma$ branching fraction, evaluated at $m_{\mathrm{H}} = 125.38$ GeV, are taken from Ref. [42], and the fractional breakdowns into the respective STXS bins are calculated directly from the MC samples (see Tables 2.5–2.8). As mentioned above, the efficiency-times-acceptance terms, $\epsilon_k^{i,\gamma\gamma}$, are derived separately for each year. The values are calculated directly from SM MC simulation with $m_{\mathrm{H}} = 125.0$ GeV, as the fraction of the total yield of STXS bin i that lands in analysis category k. Again, this approach works under the fact that the variation in $\epsilon_k^{i,\gamma\gamma}$ is negligible in going from $m_{\mathrm{H}} = 125.0$ GeV to 125.38 GeV, and is well below the yield systematic uncertainties introduced in Section 5.4.2. This was explicitly checked using dedicated signal samples at $m_{\mathrm{H}} = 120$ and 130 GeV, and interpolating between mass points to obtain the efficiency-

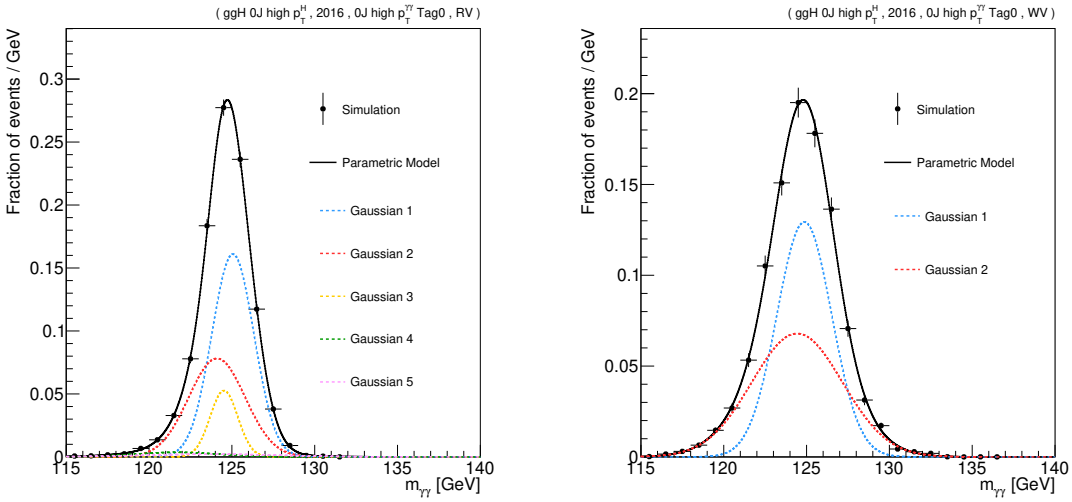


Figure 5.2: The final signal model shapes (black lines) for ggH 0J high p_T^H events in the 0J high $p_T^{\gamma\gamma}$ Tag0 category, from 2016 simulation with $m_{\mathrm{H}} = 125.0$ GeV (black points). Events in the RV and WV scenarios are shown in the left and right plots, respectively. In each case, the total shape is decomposed into the individual Gaussian functions, shown by the coloured lines.

times-acceptance values at $m_{\mathrm{H}} = 125.38$ GeV. The inclusive efficiency-times-acceptance, defined as the fraction of the total signal yield landing in any category, was found to vary by less than 0.1%. The $\epsilon_k^{i,\gamma\gamma}$ are shown separately for each year in Appendix D.

At this point it is worth reminding the reader of a particular property of STXS measurements: the experimental acceptance of events within a given STXS bin are assumed to take their SM values (as the $\epsilon_k^{i,\gamma\gamma}$ terms are extracted with SM simulation). In the presence of BSM physics, the event kinematics can be modified which in turn affects experimental acceptance. This assumption, which is explicitly discussed in the context of EFT in Section 7.7.4, should be treated carefully when performing BSM interpretations of STXS measurements.

The per-category signal models are defined by summing the individual models of each STXS bin,

$$\mathcal{S}_k = \sum_i \mathcal{S}_k^{i,\gamma\gamma} = \sum_i f_{\mathrm{RV}} \cdot \mathcal{S}_{k,\mathrm{RV}}^{i,\gamma\gamma} + (1 - f_{\mathrm{RV}}) \cdot \mathcal{S}_{k,\mathrm{WV}}^{i,\gamma\gamma}. \tag{5.11}$$

Figure 5.3 shows the per-category signal models for the 0J high p_T^H Tag0 and qqH VH-like Tag0 analysis categories. The σ_{eff}, defined as half of the smallest interval containing 68.3% of the invariant mass distribution, is used to quantify the diphoton mass resolution. In the plots, the models are split into the contributions from each year, and the respective σ_{eff} values are displayed. In general, the detector performance, and therefore the diphoton mass resolution, worsen as a function of time due to radiation damage in

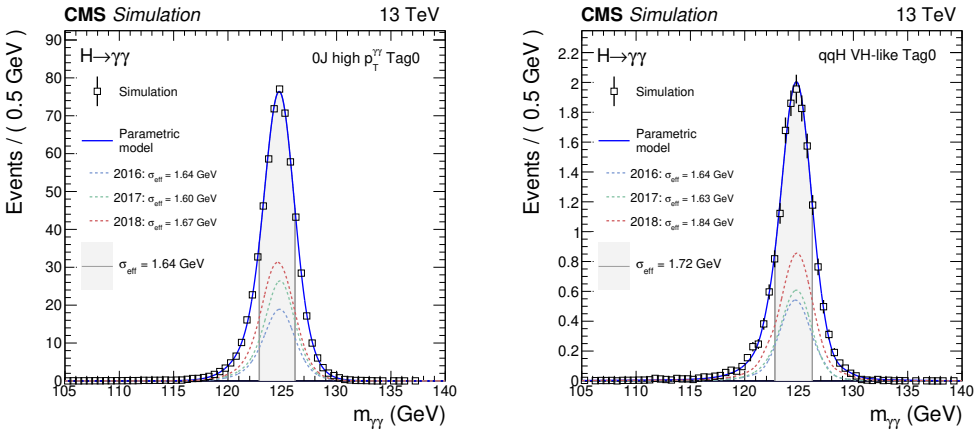


Figure 5.3: Per-category signal models for the 0J high p_T^H Tag0 (left) and qqH VH-like Tag0 (right) categories, shown for $m_{\mathrm{H}} = 125.0$ GeV. The model is shown for each year of simulated data (dashed lines) and for the sum of all years (solid line), normalised to the expected signal yield in each category. The open squares correspond to the simulated signal events. Also shown is the σ_{eff} in the grey shaded area.

the electromagnetic calorimeter. However, an extensive and thorough offline reconstruction program was developed specifically for the 2017 data, resulting in improved diphoton mass resolution and therefore smaller values of σ_{eff}. Such offline reconstruction programs will be performed by the CMS Collaboration for the 2016 and 2018 data in the future.

Going further, we can plot the sum of all per-category models to show the total signal model, $\mathcal{S}$. In Figure 5.4, each per-category model is weighted in the sum according to the $F_{68} = S_{68}/(S_{68} + B_{68})$ values displayed in Tables 4.6–4.9, such that the total signal yield remains constant. The explicit form of the weighted sum is

$$\mathcal{S}' = \sum_k w_k \, \mathcal{S}_k; \qquad w_k = (F_{68}^k) \times \frac{\sum_l S_{68}^l}{\sum_l F_{68}^l S_{68}^l}. \tag{5.12}$$

where the index l runs over all analysis categories.

5.3 Background modelling

Events entering the analysis categories that do not originate from Higgs boson production form a smoothly falling $m_{\gamma\gamma}$ distribution, on top of which the signal peak resides. To derive the $B_{k,X}$ terms of equation 5.2, an analytic model is constructed in each analysis category to describe the distribution of background events. These models are extracted directly from data.

The underlying shape of the background distribution is not *a priori* known, and therefore a number of functional forms must be considered. Ultimately, this analysis amounts

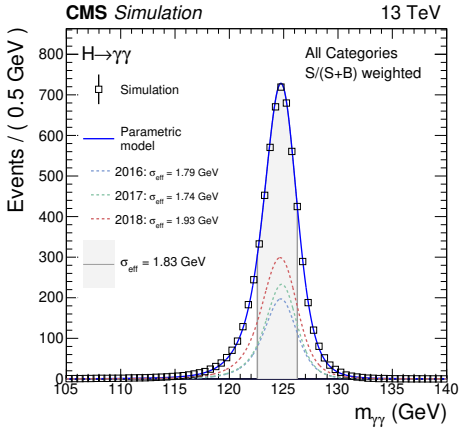


Figure 5.4: The total signal model is shown for $m_{\mathrm{H}} = 125.0$ GeV, defined as the weighted sum of all per-category signal models. In the sum each category is weighted according to $F_{68} = S_{68}/(S_{68} + B_{68})$, such that the total signal yield remains constant.

to measuring a small Higgs boson signal sitting on top of a larger background. As a result, relatively small changes in the background model shape and thus the estimated background contribution under the signal peak, may incur a large variation in the measured signal parameters of interest. The uncertainty in the choice of background function must therefore be accounted for.

5.3.1 The effect of nuisance parameters on the likelihood

Before introducing the background-modelling procedure in detail, it is worth taking time to understand the effect of nuisance parameters on the likelihood. When minimising the quantity $-2\ln L$, the nuisance parameters representing systematic uncertainties are profiled: their value is free to vary during the minimisation, in accordance with the specified constraint $\mathcal{C}$, but their final value is not of interest. The increased freedom in the fit means that for a given point in parameter space, $\vec{\alpha}$, a configuration of the nuisance parameters can be found, $\vec{\theta}_{\vec{\alpha}}$, which increases the likelihood, $L|_{\vec{\alpha}}$, or equivalently decreases the value of $-2\ln L|_{\vec{\alpha}}$. This manifests as a widening of the $q(\vec{\alpha})$ curve, and therefore an increase of the uncertainty in the fitted parameters of interest.

The contribution to the total uncertainty from a set of nuisance parameters, $\vec{\theta}'$, is realised by fixing the nuisance parameters to their best-fit values, $\hat{\vec{\theta}}'$, in the fit. The width of the resulting $q(\vec{\alpha})$ curve represents the total uncertainty without the effect of $\vec{\theta}'$, and therefore will be narrower than the curve in which the nuisance parameters in $\vec{\theta}'$ are profiled. Ignoring correlation effects between sets of nuisance parameters[3], we can define

[3]This is valid when considering the theoretical and experimental uncertainties as separate nuisance parameter sets, as the correlations between them are negligible.

the contribution to the uncertainty from $\vec{\theta}'$ as the quadrature-difference in the curve widths. In this analysis, we use this technique to decompose the total uncertainty into contributions from theoretical sources ($\vec{\theta}_s^{\text{th}}$, $\vec{\theta}_s^{\epsilon,\text{th}}$), experimental sources ($\vec{\theta}_s^{\epsilon,\text{exp}}$, $\vec{\theta}_s^{\text{shape}}$, $\vec{\theta}_s^{\text{lumi}}$), and the statistical component[4]. The *impact* of a single nuisance parameter, θ', on parameter of interest, α, is evaluated using a different technique. Two fits are performed with θ' fixed to its $+1\sigma$ and -1σ post-fit values, and the impact is defined as the shift in the parameter of interest value, $\Delta\alpha$. In the fits, the other nuisance parameters are profiled as normal such that their correlations are accounted for. The impact effectively provides a measure of the correlation between θ' and α, and therefore determines which nuisance parameters have the largest effect on the α uncertainty.

Figure 5.5 illustrates a different approach to the effect of nuisance parameters on the likelihood. The blue curve represents a fit in which the nuisance parameter in question, θ, is fixed to its best-fit value. By fixing θ to different values, it is possible to build up a series of likelihood curves, shown by the dashed red lines, with minima shifted from the unconditional best-fit point. The minimum envelope of these likelihood curves, shown by the dashed green line, can be used to approximate the contribution to the uncertainty from θ. In the limit of infinitesimally small steps in θ, the envelope converges to the fully-profiled likelihood-curve (black line).

5.3.2 Discrete profiling method

The example described above was introduced in the context of a continuous nuisance parameter. This same approach of building the envelope can be extended to the discrete case, where the nuisance parameter is limited to discrete values, albeit provided there is sufficient freedom in the allowed values to provide good coverage of the uncertainty.

In this analysis, the so-called *discrete profiling method* is applied to model the uncertainty in the choice of background function [147]. A number of candidate functions are considered to fit the background in each analysis category, and a discrete nuisance parameter is introduced to label the choice of function. In theory, the complete set of all analytic functions should be considered to obtain the exact uncertainty; in practice it is sufficient to consider a subset of functions which provide a reasonable fit to data. This keeps the computing time required for the full minimisation to a tolerable level. By allowing the value of the discrete nuisance parameter, and thus the functional form of the background to vary, an envelope of likelihoods is constructed (as in Figure 5.5) which successfully approximates the uncertainty in the choice of background function.

The different families of background-shape functions considered are listed in Table 5.1.

[4]Here, the background model nuisance parameters, $\vec{\theta}_b^{\text{shape}}$, $\vec{\theta}_b^{\text{discrete}}$, are grouped with the statistical uncertainty.

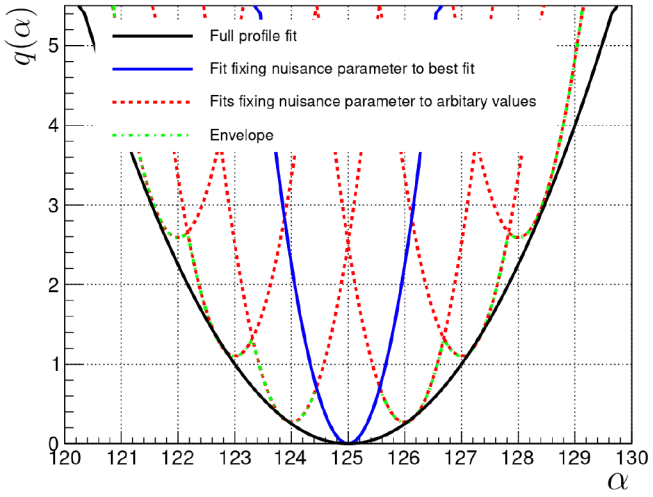


Figure 5.5: An illustration of constructing the envelope. The blue $q(\alpha)$ curve corresponds to a fit in which a nuisance parameter is fixed to its best-fit value. The red dashed lines show the curves obtained from fits where the nuisance parameter is fixed to different values. The minimum envelope of these curves is shown by the dashed green line. By sampling a sufficient number of different nuisance parameter values, the envelope approximately models the fully profiled likelihood curve (black line). Figure taken from Ref. [147].

Table 5.1: Different function families considered for modelling the background with the discrete profiling method. Each function is shown for order, N.

Sum of exponentials	$f_N(x) = \sum_{i=0}^{N} p_{2i} \exp(p_{2i+1} x)$
Sum of power law functions	$f_N(x) = \sum_{i=0}^{N} p_{2i} x^{-p_{2i+1}}$
Bernstein polynomials	$f_N(x) = \sum_{i=0}^{N} p_i \binom{N}{i} x^i (1-x)^{N-i}$
Laurent series	$f_N(x) = \sum_{i=0}^{N} p_i x^{-4+g(i)}; \qquad g(i) = \sum_{j=0}^{i} (-1)^j j$

For each family, the form is shown for order N, uniquely described by parameters: p_0, p_1, ..., p_N. The following procedure is used to determine the final set of candidate functions for a given analysis category, $\mathcal{F}^k$:

- A background-only fit is performed to the $m_{\gamma\gamma}$ distribution in data, using the lowest-order function in a given family. This is performed by minimising the value of $-2\ln L_b$, allowing the parameters of the function to vary, where the subscript b has been added to indicate this is a background-only fit ($S = 0$). The process is repeated, incrementally raising the order, until a minimum goodness-of-fit criteria is reached. All orders below are not considered in $\mathcal{F}^k$ as they do not fit the data well.

- For all subsequent orders, an F-test is performed to decide if the improvement in fit quality warrants the increase in function complexity [148]. Using the same procedure, the $-2\ln L_b$ value is determined for the next-highest-order function in the family. Given a large enough sample size, the difference in $-2\ln L_b$ values for successive orders, Δ, is distributed according to the χ^2 distribution with m degrees of freedom, where m is the difference in the number of parameters between the two orders. A p-value is calculated as,

$$p = \text{Prob}\Big(\Delta > \Delta_{\text{obs}} \,\Big|\, \chi^2(m)\Big), \tag{5.13}$$

  where Δ_{obs} is the observed value in data. If the p-value is less than 0.05 then the improvement in fit quality is deemed worthwhile, and the higher-order function is added to $\mathcal{F}^k$. This step is repeated for successive orders until the calculated p-value is larger than 0.05. In this scenario, the higher-order function is deemed to be unnecessarily complex given the data and the procedure terminates.

- This process is repeated for each of the four families listed in Table 5.1, where each order passing the above selection procedure enters $\mathcal{F}^k$.

The set of candidate functions, $\mathcal{F}^k$, are shown for the ggH 1J high $p_T^{\gamma\gamma}$ Tag0 and ggH BSM $200 < p_T^{\gamma\gamma} < 300$ Tag0 analysis categories in Figure 5.6. The different functional forms provide a different background estimate when integrating under the signal peak. This sometimes large variation in the background estimate gives rise to an uncertainty in the fitted signal parameters of interest, originating from the lack of knowledge of the true background functional form. Clearly some of the functions give a reasonably poor fit to data, such as the order-1 Laurent series function in the ggH 1J high $p_T^{\gamma\gamma}$ Tag0 analysis category (yellow). Despite entering $\mathcal{F}^k$, such functions will never be chosen by the discrete profiling method in the maximum-likelihood fit and therefore will not impact the results.

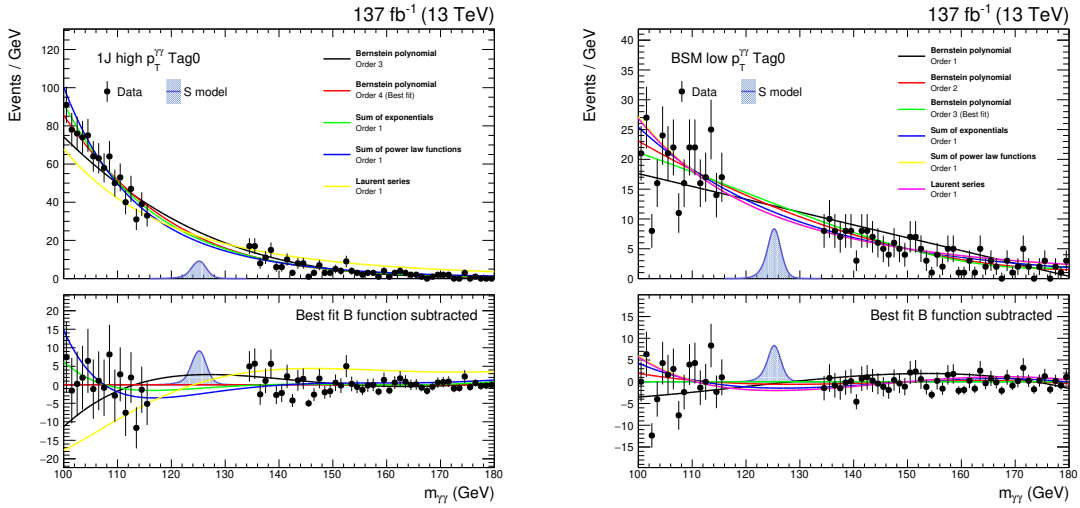


Figure 5.6: The set of candidate functions, $\mathcal{F}^k$, considered in the final fit for the ggH 1J high $p_T^{\gamma\gamma}$ Tag0 (left) and ggH BSM $200 < p_T^{\gamma\gamma} < 300$ Tag0 (right) categories, shown by the coloured lines. Data events entering these analysis categories are shown by the black points, for diphoton invariant mass outside of the range $115 < m_{\gamma\gamma} < 135$ GeV. The corresponding signal models for the two categories are overlaid (filled blue histograms) to gain an idea of the impact of the uncertainty in the choice of background function. The bottom panel in each plot shows the residuals after subtracting the best-fit background function in the background-only fit.

In the final results extraction (see Section 5.1.2), both the discrete nuisance parameters describing the choice of background functions ($\vec{\theta}_b^{\text{discrete}}$), and the parameters of the functions themselves ($\vec{\theta}_b^{\text{shape}} \equiv p_0, ..., p_N$) are free to vary. In accordance with the procedure described above, a penalty term is added to the value of $-2\ln L$ equal to the number of parameters in the chosen function, thus penalizing functions with high complexity. Further details concerning the discrete profiling method are provided in Ref. [147]. This includes a series of tests to show the method provides good coverage of the uncertainty in the choice of background function and leads to unbiased estimates of the parameters of interest.

5.4 Systematic uncertainties

This section provides further detail on the sources of systematic uncertainty in the signal estimate, and how they are modelled using different types nuisance parameters. The uncertainties are divided into experimental sources and theoretical sources, described in Sections 5.4.2 and 5.4.3, respectively. For each source, the magnitude of the uncertainty's impact is calculated separately per-year for each STXS bin in each analysis category.

5.4.1 Uncertainty correlation scheme

Before listing the various sources, it is important to introduce the concept of correlating nuisance parameters. As mentioned in Section 5.2, the signal is modelled independently for the 2016, 2017, and 2018 data. This allows for a choice in the modelling of each uncertainty source. A correlated uncertainty means defining a single nuisance parameter in the final fit which affects the signal estimate in multiple years simultaneously. Note, this correlation may be defined for a pair of years, or can extend to a correlation across all three data-taking years. Alternatively, uncorrelated means defining a separate nuisance parameter for each year, such that the impacts on the signal estimates from each year are independent.

Theoretical uncertainties are fully correlated across years. This somewhat trivial assignment is made since the underlying theoretical predictions are constant and do not depend on the data-taking conditions. In general, experimental sources are uncorrelated across years. This choice reflects the difference in data-taking conditions and data reconstructions for each year. Exceptions are the uncertainties in the luminosity estimate and the jet-energy scale. For these, a partial-correlation scheme is used i.e. a combination of uncorrelated and correlated nuisance parameters are defined, where the correlated parameters correspond to uncertainty sources that are common across years.

At least for the STXS measurements, in which the dominant uncertainties are statistical in origin, the choice of uncertainty correlation scheme is a subdominant feature of the analysis. Nevertheless, the importance will increase as more data is recorded or when the results are combined with other decay channels, as the systematic uncertainties become comparable in size to the statistical component.

5.4.2 Experimental uncertainties

Experimental uncertainties can be separated into those which affect the shape of the signal $m_{\gamma\gamma}$ peak, and those that do not. Uncertainties that affect the signal shape are modelled using Gaussian-constrained nuisance parameters, $\vec{\theta}_s^{\text{shape}}$, that can simultaneously shift the mean, width, and normalisation of the Gaussian functions in the signal models. Typically, the sources for $\vec{\theta}_s^{\text{shape}}$ are related to the measurement and reconstruction of the photon energy. On the other hand, uncertainties which do not affect the signal shape are modelled as log-normal variations in the signal-yield estimates,

$$S_k^{i,\gamma\gamma}(\vec{\theta}) = S_k^{i,\gamma\gamma} \cdot \prod_a \left(1 + \frac{\Delta_{a,k}^{i,\gamma\gamma}}{S_k^{i,\gamma\gamma}}\right)^{\theta_a}, \tag{5.14}$$

where $\Delta_{a,k}^{i,\gamma\gamma}$ is the variation in signal yield $S_k^{i,\gamma\gamma}$, due to the uncertainty source en-

coded by the nuisance parameter, θ_a. For example, a 2% variation corresponds to $\Delta_{a,k}^{i,\gamma\gamma}/S_k^{i,\gamma\gamma} = 0.02$. Two values can be defined for $\Delta_{a,k}^{i,\gamma\gamma}$ (one for positive values of θ_a and one for negative values) to account for asymmetric variations caused by the signal yield uncertainties. The product in equation 5.14 is calculated over all nuisance parameters, θ_a, which are defined as log-normal variations in the signal yield.

Signal-shape uncertainties

The effects of the signal-shape uncertainties are directly encoded into the signal models themselves as variations in the Gaussian function parameters. The general approach here is to vary the source of uncertainty and compare the mean, width, and normalisation of the resulting signal $m_{\gamma\gamma}$ peak to the nominal. Uncertainties concerning the energy scale predominantly affect the mean values, whilst uncertainties in the energy resolution mostly affect the widths. Nevertheless, for each source of uncertainty the combined effects on the mean, width, and normalisation of the signal peak are correlated into a single nuisance parameter. Figure 5.7 shows the maximum variation in the signal shape of the 0J high $p_T^{\gamma\gamma}$ Tag0 category, by deviating the nuisance parameters to $\pm 1\sigma$. For illustration purposes the total effect has been decomposed into the maximum variation in the mean (left) and width (right). In general, the effect on the total normalisation is negligible. Ultimately, the total impact from the signal-shape uncertainties on the inclusive Higgs boson signal-strength measurement is approximately 2%. The experimental sources of the signal-shape uncertainties are as follows.

- *Photon energy scale and resolution*: the residual corrections to the photon-energy scale in data and photon energy resolution in simulation, defined in Section 4.3.1, introduce uncertainties into the analysis. These uncertainties are evaluated using $Z \to e^+e^-$ events, by varying the energy-regression training scheme, the distribution of shower-shape variable R_9, and the electron selection-criteria, and re-deriving the scale and smearing corrections. Typically, the resulting uncertainty in the photon-energy scale is 0.05–0.1%, but can rise to be 0.5–3% for very high p_T photons.

- *Nonlinearity of the photon-energy scale*: an additional source of uncertainty covering differences in the linearity of the photon-energy scale between data and simulation arises since the corrections are estimated using $Z \to e^+e^-$ events with electron p_T typically around 45 GeV, but are applied to photons with typical p_T around 60 GeV. The magnitude of this uncertainty is estimated using boosted (high-p_T) $Z \to e^+e^-$ events to provide a sample of electrons with p_T values that are more representative of the photons which enter the analysis. The uncertainty in the photon-energy scale is calculated to be 0.2% for the full range of photon p_T values [149].

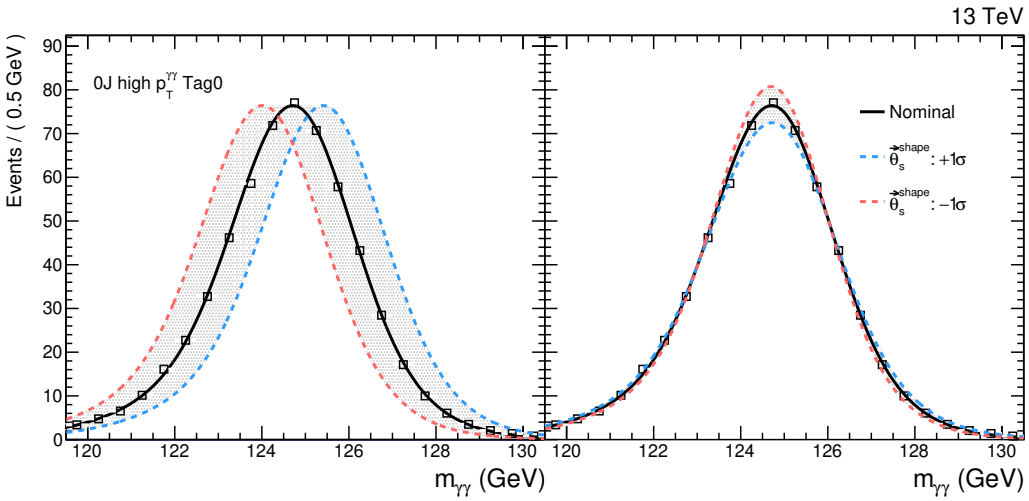


Figure 5.7: The impact of the signal-shape uncertainties on the 0J high $p_T^{\gamma\gamma}$ Tag0 per-category model, decomposed into the maximum effect on the mean (left) and width (right). All nuisance parameters, $\vec{\theta}_s^{\text{shape}}$, are shifted to plus 1σ (blue dashed line) and minus 1σ (red dashed line) of their nominal values. The maximum absolute variation in the mean is around 0.7 GeV, whilst the maximum relative variation in the width is approximately 5%.

- *Shower shape corrections*: imperfect modelling of electromagnetic shower-shapes in simulation is a source of uncertainty. The impact of this uncertainty is derived by evaluating the energy scale in simulation before and after the CQR corrections to the shower-shape variables are applied. The magnitude of the uncertainty in the energy scale depends on the photon $|\eta|$ and R_9 values, ranging from 0.01–0.15%.

- *Non-uniformity of light collection in the ECAL crystals*: the maximum shower-length is deeper for photons than electrons by approximately one radiation-length. Again, as the corrections are derived using electrons, an uncertainty is introduced concerning the modelling of light-collection as a function of emission depth within a given ECAL crystal. The calculation of this uncertainty is described in detail in Ref [49]. For photons with $R_9 > 0.96$, the uncertainty in energy is between 0.16–0.25%, whereas the magnitude is below 0.07% for low-R_9 photons, which are more likely to have undergone a conversion in the tracker.

- *Modelling of material in front of the ECAL*: the amount of material upstream of the ECAL affects the properties of the electromagnetic shower, and may not be perfectly modelled in simulation. Dedicated simulation samples are used with differing amounts of upstream material to evaluate the impact on the photon energy measurement. For the most central photons, the uncertainty in the energy ranges

from 0.02–0.05%, but increases to a maximum of 0.24% for photons in the endcaps.

- *Vertex assignment*: somewhat different to the sources described above, a nuisance parameter is introduced to model the uncertainty in the vertex scenario assignment, directly modifying the value of f_{RV}. The largest contribution originates from the modelling of the underlying event i.e. everything in the event that is not associated with the hard-scattering process. An additional contribution comes from the vertex assignment BDT correction factors due to differences between $\mathrm{Z} \to \mu^+\mu^-$ events in data and simulation. The total uncertainty allows f_{RV} to vary by $\pm 2\%$.

Yield uncertainties

The experimental uncertainties that only modify the signal-yield estimates include the uncertainties in the luminosity, $\vec{\theta}_s^{\,\mathrm{lumi}}$, and those affecting the efficiency-times-acceptance of the event selection, $\vec{\theta}_s^{\,\epsilon,\mathrm{exp}}$. In general, the magnitude of the yield variation is calculated by varying the uncertainty source in simulation, propagating the events to the final analysis categories, and comparing the systematic-varied yield to the nominal. The sources of uncertainty are listed below.

- *Integrated luminosity*: uncertainties of 2.5%, 2.3%, and 2.5% are determined by the CMS luminosity monitoring for the 2016, 2017, and 2018 data sets respectively [150–152], whilst the uncertainty in the total integrated luminosity of all three years is 1.8%. A partial-correlation scheme is introduced to account for common sources of uncertainty in the luminosity measurements of each year. In total, this amounts to defining three uncorrelated and six correlated nuisance parameters.

- *Photon pre-selection*: the photon pre-selection efficiency scale factors are derived using the tag-and-probe method on $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ events. This amounts to estimating the number of probes passing and failing selection by fitting a signal-plus-background model to the dielectron invariant-mass distribution. The largest source of uncertainty is from the choice of signal-shape in the fit, and is estimated by fitting an alternative shape. Propagating to the category yields, the impact is less than 1%. An additional uncertainty is included for the electron-veto scale factor, calculated using $\mathrm{Z} \to \mu^+\mu^-\gamma$ events, which has an even smaller impact on the yield, typically less than 0.5%.

- *Trigger efficiency*: again the trigger efficiency scale factors are measured with $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ events using the tag-and-probe technique, and the size of the uncertainty is also less than 1%. An additional source is included to account for a gradual shift in the timing of the inputs of the ECAL L1 trigger in the region $|\eta| > 2.0$, which

caused a specific trigger inefficiency during the 2016 and 2017 data-taking. Both photons and to a greater extent jets can be affected by this inefficiency. The resulting uncertainty is largest for the categories targeting VBF production, with a maximum impact on the yield of around 1.4%.

- *Photon identification BDT score*: as discussed in Section 4.3.1, the largest uncertainty in the photon identification originates from the limited size of the training sample in the CQR method. The size of the uncertainty is estimated by splitting the training sample in half and re-calculating the shower-shape and isolation corrections. Figure 4.4 shows this uncertainty (pink band) successfully covers the residual discrepancies between data and simulation. The effect on the yield is then calculated by propagating this source of uncertainty through the full event categorisation, resulting in a variation of around 3% for the most sensitive categories.

- *Per-photon energy resolution estimate*: the per-photon energy resolution estimate is an output of the photo-energy regression and is used as an input feature in the event classifiers. The uncertainty in the resolution is parametrised as a $\pm 5\%$ variation about the nominal value, chosen to sufficiently cover all differences between data and simulation in the per-photon energy resolution distribution. The maximum yield variation in an analysis category from this source is around 5%, however for most categories the impact is below the per-cent level.

- *Jet-energy scale and smearing corrections*: the jet-energy scale is calculated using the p_T balance of jets with Z bosons and photons in $Z \to e^+e^-$, $Z \to \mu^+\mu^-$ and γ+jet events, as well as the p_T balance between jets in dijet and multijet events [135]. This energy scale is then used to correct the jet energies in simulation and data as a function of jet p_T^j and $|\eta^j|$. The sources of uncertainty in this calculation arise from the absolute value of the scale, the relative $|\eta|$-dependence of the scale, pile-up mitigation and the detector response to different jet flavours. Over the full jet phase-space considered in this analysis, the final uncertainties in the jet-energy scale are below 3%. Similar to the luminosity uncertainties, a partial-correlation scheme is introduced, with correlations ranging between 0 and 100%, to account for common sources of uncertainty in the jet-energy scale measurement. Propagating this to the event yields, the impact is largest in categories targeting VBF, VH hadronic and top-associated production, and can be as high as 22%. A separate nuisance parameter is introduced to account for the uncertainty in the jet-energy resolution calculation. The impact on the event yields is in general smaller than the jet-energy scale uncertainties, but can be as high as 8% for the most-affected categories.

- *Pileup jet identification*: the uncertainty in the pileup-identification BDT score is determined by comparing the scores of jets in events with a Z boson plus one balanced jet, in data and simulation. The effect on event categories with jet requirements is of the order of 1%.

- *Missing transverse momentum*: the p_T^{miss} is used as an input variable in a number of event classifiers. The uncertainty in the p_T^{miss} is derived by shifting the reconstructed p_T of the physics objects entering the p_T^{miss} calculation, within the momentum scale and resolution uncertainties appropriate to each type of reconstructed object, as described in Ref [153]. Independent nuisance parameters are defined for the p_T shifting of jets, photons, and unclustered objects. The impact on the category yields is never larger than 5%, even for analysis categories that explicitly use the p_T^{miss} in the respective classifier.

- *Lepton isolation and identification*: for electrons (muons), the efficiencies of the lepton isolation and identification cuts are calculated using the tag-and-probe technique on $Z \to e^+e^-$ ($Z \to \mu^+\mu^-$) events. The corresponding impact on the signal-yield estimates are computed by varying the ratio of the efficiencies measured in data and simulation, within their uncertainties. The variations in the yields are less than 0.7% in the ttH leptonic and tHq leptonic categories, 0.6% in the WH leptonic categories and 1% in the ZH leptonic categories.

- *Jet b tagging*: uncertainties in the b tagging efficiency are evaluated by comparing b tag discriminator output distributions in data and simulation. The uncertainties include a statistical component on the estimate of heavy and light-flavour jets in data and simulation. For top-associated categories, which make use of the b tagging discriminant, the variations in the yields are around 3%.

All in all, the combined impact of these experimental yield uncertainties on the inclusive Higgs boson signal-strength measurement is between 3 and 4%.

5.4.3 Theoretical uncertainties

The effect of all theory uncertainties are modelled using log-normal variations in the signal yield estimates (see equation 5.14). Introduced in Section 5.1.1, theory uncertainties contribute to the likelihood in two ways. Firstly, there are the uncertainties in the SM predictions of the cross sections and the branching fraction, $\vec{\theta}_s^{\,\mathrm{th}}$. The cross section uncertainties include both the inclusive per-production mode effects and the uncertainties in the kinematic distributions, which can migrate events across STXS bin boundaries. As a result of the signal parametrisation used for the cross section measurements, shown

in equation 5.5, uncertainties of this first type cancel in the ratio with equation 5.4, and therefore do not enter the final-fit likelihood. Instead, they are attributed to the uncertainties in the SM predictions of the measured quantities. Conversely, for all interpretations of cross section measurements, e.g. signal-strengths, coupling-modifiers or EFT parameters, the $\vec{\theta}_s^{\,\mathrm{th}}$ nuisance parameters are folded into the measurement, i.e. included in the likelihood.

Originating from our lack of understanding of the underlying theory when simulating events, the second type of uncertainty, $\vec{\theta}_s^{\,\epsilon,\mathrm{th}}$, affects the acceptance of the analysis categories. For each uncertainty source, the total effect on the normalisation of each STXS bin is removed by setting the sum of the yield variations across all analysis categories (including the out-of-acceptance events) to zero,

$$\sum_k \Delta_{a,k}^{i,\gamma\gamma} = 0. \tag{5.15}$$

The variation in the yield of STXS bin i in analysis category k is subsequently measured relative to the total variation of that STXS bin. This provides a nuisance parameter which models the migration of events from a given STXS bin between analysis categories, and does not effect the overall normalisation of that bin. In other words, this models the within-STXS-bin shape effects. When calculating the impact of these nuisance parameters, it is imperative to include events that do not enter any analysis category to ensure that the migration of events in and out-of-acceptance is accurately modelled. In contrast to the first type, these uncertainties enter the cross section measurements as variations in the calculated $\epsilon_k^{i,\gamma\gamma}$ values.

The different sources of theoretical uncertainty and their application via $\vec{\theta}_s^{\,\mathrm{th}}$ and $\vec{\theta}_s^{\,\epsilon,\mathrm{th}}$ nuisance parameters are described below.

- *Renormalisation and factorisation scales*: the uncertainty arising from varying the renormalisation and factorisation scales from the nominal value used to calculate the SM predictions for the cross sections, and to simulate the kinematic properties of the events. These uncertainties account for missing higher-order terms in perturbative calculations, and reduce in size as the order of the calculation is increased. The recommendations from Ref. [42] are used, where the $\pm 1\sigma$ variations are displayed for the different Higgs boson production modes in Table 5.2. The effect on the ggH cross section depends on the number of jets in the event.

  An additional uncertainty is included for ggH production originating from the resummation of divergent terms in the perturbative expansion. The magnitude of this uncertainty is 0.1% for 0J events, 4.5% for 1J events, and 8.9% for events with at-least two jets.

Table 5.2: The QCD scale uncertainties in the production mode cross sections, expressed as percentage variations.

ggH 0J	ggH 1J	ggH ≥2J	VBF	WH	ZH	ggZH	ttH	tHq	tHW	bbH
±3.8%	±5.2%	±8.9%	±0.4%	$^{+0.5}_{-0.7}$%	$^{+3.8}_{-3.1}$%	$^{+25.1}_{-18.9}$%	$^{+5.8}_{-9.2}$%	$^{+6.5}_{-14.9}$%	$^{+4.9}_{-6.7}$%	$^{+20.2}_{-23.9}$%

The within-bin shape variations in the event kinematic properties are modelled by recalculating the fraction of events in each category, $\epsilon_k^{i,\gamma\gamma}$, when changing the renormalisation and factorisation scales by a factor of two in the same direction simultaneously. These uncertainties, of type $\vec{\theta}_s^{\epsilon,\text{th}}$, are decorrelated for different production modes and different kinematic regions of the Higgs boson phase space, resulting in 22 independent nuisance parameters. The magnitude of these within-bin uncertainties accounting for migrations between analysis categories are in general around 1%, but can be as large as 5% for off-diagonal elements in Figure 4.20.

- *Modelling of ggH STXS fractions*: additional sources of uncertainty of type $\vec{\theta}_s^{\text{th}}$ are included to account for the migration of events *across* STXS bin boundaries, which therefore affect the SM predictions of the cross sections in each bin. These migration uncertainties are defined such that the total cross section of the summed bins, either side of the boundary, remains constant. In other words, the combined impact of such across-bin migration uncertainties cancels in the sum.

  In total, fifteen of such nuisance parameters are defined for ggH. Nine account for the uncertainty in the modelling of the p_T^H distribution: four for migrations across the $p_T^H = 200$, 300, 450, and 650 GeV boundaries, one for the migration of 0J events across the $p_T^H = 10$ GeV, and four nuisances for migrations across the $p_T^H = 60$ and 120 GeV, defined independently for 1J and ≥2J events. Two nuisance parameters are defined to reflect migrations across the m_{jj} boundaries at 350 and 700 GeV, and a further two to account for the variations in the poorly-modelled p_T^{Hjj} variable, covering migrations across the 25 GeV boundary for events in different m_{jj} bins. Finally, two nuisances are included to account for the migration between the zero, one, and two-or-more jet bins.

  An additional source is included to model the uncertainty of ggH in the boosted regime. The flat value of 15% is taken directly from Ref. [154], and affects all ggH STXS bins with $p_T^H > 200$ GeV. Figure 5.8 shows the full set of $\vec{\theta}_s^{\text{th}}$ nuisance parameters considered for ggH, and their relative impact on each STXS bin. The total magnitude of these uncertainties ranges from 6.5% for ggH 0J high p_T^H, to around 40% for bins in the VBF-like region.

- *Modelling of qqH STXS fractions*: in a similar fashion, a set of nine nuisance parameters are included to account for migrations across bin boundaries in the qqH STXS scheme. Six are used to model the uncertainty in the m_{jj} distribution, explicitly covering the migrations at m_{jj} = 60, 120, 350, 700, 1000 and 1500 GeV. A further three nuisances are introduced for the boundaries at p_T^H = 200 GeV and p_T^{Hjj} = 25 GeV for high dijet-mass events, and the migration of 0J and 1J events into the $\geq$2J region. The variations in cross sections arising from these migration uncertainties are typically below 3%.

- *Modelling of VH leptonic STXS fractions*: mis-modelling of the p_T^V spectrum is covered by migration uncertainties defined for the boundaries at 75, 150 and 250 GeV. An additional nuisance parameter is included for events with p_T^V between 150 and 250 GeV, to cover migrations across the 0J to $\geq$1J boundary. These nuisances are defined independently for the WH, ZH, and ggZH production modes, resulting in 12 independent nuisance parameters. The values are similar for WH and ZH, being consistently lower than 6%, but can be as large as 90% for ggZH [155].

- *Modelling of ttH STXS fractions*: four nuisance parameters are introduced to account for variations in the p_T^H spectrum, defined for the boundaries at p_T^H = 60, 120, 200, and 300 GeV. The magnitude of these migration uncertainties are typically around 5%, but can be as large as 10% for the ttH $p_T^H < 60$ STXS bin.

- *Parton density functions*: uncertainties are included to reflect the limited precision in the composition of the colliding protons. As introduced in Section 3.5, the PDFs describe the probability of finding a parton in the proton with a certain momentum fraction, x. Therefore, uncertainties in the parton density functions affect the likelihood of colliding partons, and in turn the cross section calculations and the event kinematics. The overall normalisation effect for each Higgs boson production mode is taken from Ref. [42]. The impact is largest for the top-associated production modes, ranging from 3% for ttH production to 6% for tHW production.

  The PDF4LHC recommendations [108, 132, 156, 157] are followed to evaluate the impact on the event kinematics. Here, additional PDF sets are calculated from the NNPDF 3.0 [132] PDF set using the MC2HESSIAN procedure [158, 159]. Events are reweighted for each additional PDF set, and these weights are propagated through the analysis to calculate the impact on the signal-yield estimates. The nuisance parameters are of type, $\vec{\theta}_s^{\epsilon,\text{th}}$ i.e. the overall normalisation effect on each STXS bin is integrated out to leave nuisance parameters describing the migration of events between analysis categories. In total, 60 independent nuisance parameters are introduced, which account for relatively small (<1%) migrations between analysis

categories.

- *Strong force coupling constant, α_s*: the uncertainty in α_s is treated along with the PDF uncertainties in the PDF4LHC prescription. In the same way, an overall normalisation effect is included for each Higgs boson production mode, and an additional $\vec{\theta}_s^{\epsilon,\text{th}}$ nuisance parameter is included to model the variation in the event kinematics from changing the α_s value. The normalisation uncertainty is largest for ggH production, with a value of 2.6%, whilst the within-bin migration uncertainty is in general smaller.

- *Underlying event*: dedicated simulation samples are produced, where the underlying event settings of the Pythia8 CUETP8M1 and CP5 tunes are modified. This varies the structure of the underlying event beneath the hard scattering process. This source of uncertainty is treated as a variation in the STXS bin cross sections, derived by comparing the STXS bin fractions to the nominal values. In general, the magnitude is less than 5% but can be as high as 30% for bins with high p_T and a high number of jets.

- *Parton shower*: in a similar way to the underlying event uncertainty, dedicated samples are produced to model the uncertainties in the parton shower. In the samples, the renormalisation scale for QCD emissions in initial-state and final-state radiation is varied by a factor of 2 and 0.5. Again, treated as an uncertainty in the cross section predictions, the impact is typically below 10% but increases with the event p_T and number of jets. In particular, the qqH STXS bin boundaries that are sensitive to the presence of a third jet ($p_T^{Hjj} = 25$ GeV) are greatly affected by this source of uncertainty, with impacts as high as 40%. This is because the third jet in the event always originates from the parton showering, and not the matrix element of the hard process.

- *Modelling of ggH contamination in the top-associated categories*: the theoretical predictions for ggH are less reliable in the region of phase space where the Higgs boson is produced in association with a large number of jets. Three independent nuisance parameters are introduced with a combined effect of around 50% on the ggH yield in the top-associated categories. Firstly, the uncertainty in the parton shower is modelled by comparing the difference in jet multiplicity between tt+jets events in data and MG5_aMC@NLO simulation [160]. A second source of uncertainty originates from the gluon-splitting modelling, estimated by measuring the quantity $\sigma(\text{ttbb})/\sigma(\text{ttjj})$ in both data and simulation, and scaling the fraction of ggH events with real b jets in simulation by their ratio [161]. The final source of uncertainty

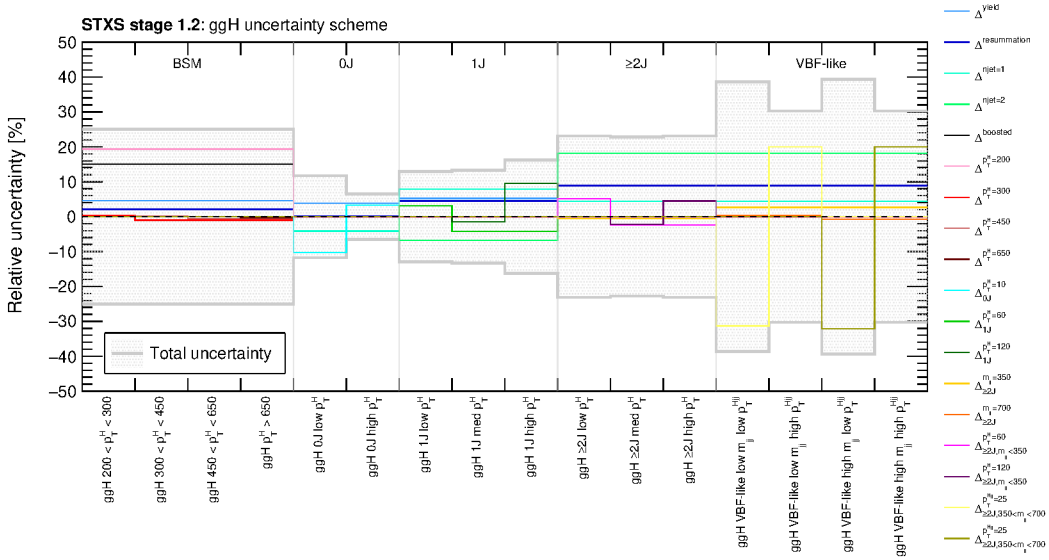


Figure 5.8: The impact of the ggH theoretical uncertainties on each STXS stage 1.2 bin, expressed as a percentage variation in the nominal cross sections. The coloured lines show the impact from each individual nuisance parameter, whilst the filled histogram shows the total uncertainty, defined as the quadrature sum of all contributions.

originates from the limited size of the simulated sample, providing a small number of ggH events which enter the top-associated categories. The combined impact of these uncertainties in the top-associated signal-strength measurement is about 2%.

- $H \to \gamma\gamma$ *branching fraction*: the probability of the Higgs boson decaying to photons is only known to within an uncertainty of 2% [42]. This uncertainty affects the inclusive signal-yield estimate and hence is of type $\vec{\theta}_s^{\,\mathrm{th}}$.

As described later in Section 6.3, it is necessary to merge groups of STXS bins in the measurement to avoid large uncertainties or very-high correlations between the measured cross sections. This induces an additional complication into the theory uncertainty scheme, such that the nuisance parameter representing the migration of events across the merged boundary must be introduced into the cross section measurement. Effectively, this act of merging re-defines the signal process and changes the across-bin migration uncertainty ($\vec{\theta}_s^{\,\mathrm{th}}$), to a within-merged-bin shape effect ($\vec{\theta}_s^{\,\epsilon,\mathrm{th}}$). For example, if the WH lep 0J $150 < p_T^V < 250$, WH lep $\geq$1J $150 < p_T^V < 250$ and WH lep $p_T^V > 250$ STXS bins are merged into a single parameter of interest, the nuisance parameters representing migrations across the $p_T^V = 250$ GeV boundary and the $N_{\mathrm{jets}} = 1$ for $150 < p_T^V < 250$ GeV boundary are included in the definition of the likelihood for the cross section measurement.

5.5 Summary

This chapter has introduced the method for the statistical inference of Higgs boson properties using high-energy p-p collision data, in particular for events consistent with the diphoton decay of the Higgs boson. Firstly, the construction of the per-category likelihood function and the subsequent method for extracting the results were described in detail. Following this, the various inputs to the likelihood function were covered. The signal model is constructed per-year for each STXS bin in each analysis category, modelled as a sum of up to five Gaussian functions. The uncertainty in the choice of background functions is accounted for using the discrete profile likelihood method, where the exact form of the function in each analysis category is modelled using a discrete nuisance parameter. Systematic uncertainties regarding the signal estimate are incorporated as constrained nuisance parameters in the likelihood. Those which affect the signal shape are encoded directly as variations in the mean, width, and normalisation of the signal-model Gaussian functions. The remaining experimental and theoretical uncertainties are modelled as log-normal variations in the yield estimates, calculated separately for each STXS bin in each analysis category.

Chapter 6

Results: H $\rightarrow \gamma\gamma$

6.1 Introduction

This chapter presents the results of the H $\rightarrow \gamma\gamma$ analysis, extracted using the statistical inference techniques introduced in Chapter 5. The measurements of Higgs boson production cross sections and couplings are presented in a number of parametrisations, corresponding to different definitions of the parameters, $\mu^{i,\gamma\gamma}$. The simplest of such parametrisations are the signal-strength modifiers, which act as global scaling factors either inclusively or per-Higgs boson production mode. The results of the signal-strength modifier fits are shown in Section 6.2. Following this, the STXS fits for which the analysis is configured, are shown in Section 6.3. To ensure a reasonable sensitivity to the measured parameters, it is necessary to merge groups of STXS bins in the cross section measurements. The results of three different merging schemes with varying degrees of granularity are presented. Finally, Section 6.4 shows the fits to Higgs boson coupling-modifiers in the κ-framework [144].

6.2 Signal-strengths

The most-constraining fit that can be performed introduces a common signal-strength modifier, μ, for all signal processes. This is defined as the ratio of the observed product of the Higgs boson cross section and diphoton branching-fraction to the SM expectation, such that a value of one corresponds directly to the SM prediction. The observed test-statistic curve, $q(\mu)$, is shown by the solid line in Figure 6.1, in addition to the dashed lines representing the $q(\mu)$ curves when groups of nuisance parameters are fixed to their best-fit values. The best-fit value of μ and the corresponding 68% confidence intervals are inferred from the curves to be

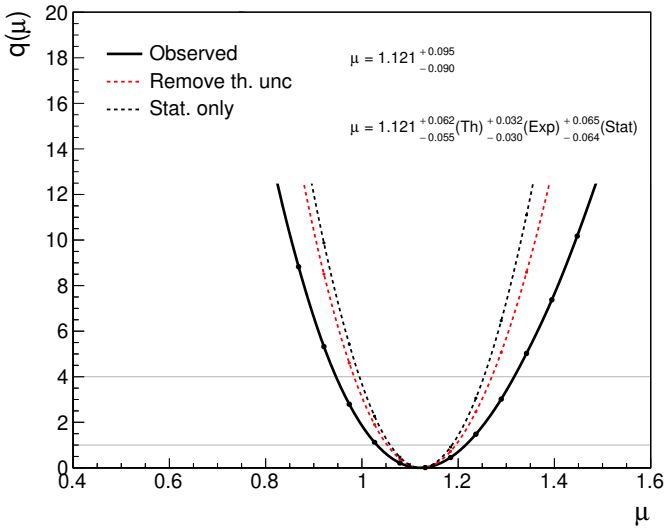


Figure 6.1: Observed likelihood curve as a function of the inclusive signal-strength modifier, μ. The likelihood in which all nuisance parameters are profiled is shown by the solid black line. The likelihood curves shown by the dashed lines represent the fits in which all nuisances are fixed to their best-fit values in black (stat-only), and only the theoretical uncertainties are fixed to their best-fit values in red.

$$\mu = 1.12^{+0.09}_{-0.09} = 1.12^{+0.06}_{-0.06}(\text{th.})^{+0.03}_{-0.03}(\text{exp.})^{+0.07}_{-0.06}(\text{stat.}),$$

where the total uncertainty has been decomposed into contributions from theoretical systematic, experimental systematic and statistical uncertainties, using the method described in Section 5.3.1. As a result of the increased integrated luminosity that comes with using the full Run 2 data, CMS are now approaching the realm of systematics-limited measurements in Higgs boson physics, where the systematic uncertainties are comparable to, if not larger than, the statistical component. The ±9% total uncertainty represents the current best constraint on inclusive Higgs boson production from a single decay channel at CMS. The compatibility of this fit with the SM prediction is approximately $p_{\text{SM}} = 17\%$, which describes the probability of repeating the measurement and obtaining a value further from the SM prediction ($\mu = 1$) than the observed value.

The observed diphoton-mass distribution in data is shown for the sum of all analysis categories in Figure 6.2. In the sum, each category is weighted using the approach detailed in equation 5.12, i.e. weight each category by the ratio of signal to signal-plus-background events, keeping the total signal-yield constant. The best-fit signal-plus-background model obtained in the inclusive signal-strength fit is overlaid. In addition, the uncertainty in the background estimate is shown by the uncertainty bands. These bands are populated by first performing the signal-plus-background fit to data, and then generating 500 toy data

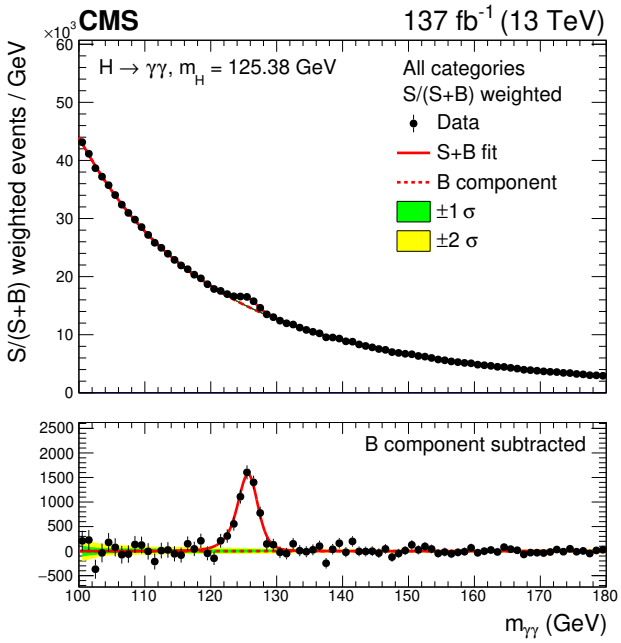


Figure 6.2: Data points (black) and the best-fit signal-plus-background model for the sum of all analysis categories is shown. The best-fit model corresponds to the inclusive signal-strength fit. In the plot, the events in each category are weighted according to $S_{68}/(S_{68} + B_{68})$, where S_{68} and B_{68} are the expected signal and background estimates, respectively, in a $\pm 1\sigma_{\text{eff}}$ window centred on m_{H}, such that the total signal yield remains constant. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panel shows the residuals after subtraction of this background component.

sets from the best-fit background-only model. The green and yellow bands signify the regions in which 68.3% (1σ) and 95.4% (2σ) of the generated toy data sets lie, respectively. The signal peak is extremely clear above the falling background distribution.

A more granular fit is performed in the signal-strength modifier parametrisation, introducing a separate μ_i parameter for each Higgs boson production mode. Unlike, the subsequent STXS fits described in Section 6.3, the VH hadronic and VH leptonic processes are grouped to scale according to μ_{VH}, whereas the VBF production mode scales with μ_{VBF}. The parameter μ_{top} scales the ttH, tHq, and tHW production modes equally, and μ_{ggH} scales both ggH and bbH production. This defines four parameters of interest in total.

The best-fit signal-plus-background model in this signal parametrisation is shown with data in the $m_{\gamma\gamma}$ distributions in Figure 6.3. Here, the analysis categories have been divided into four groups, corresponding to those targeting the ggH, VBF, VH, and top-

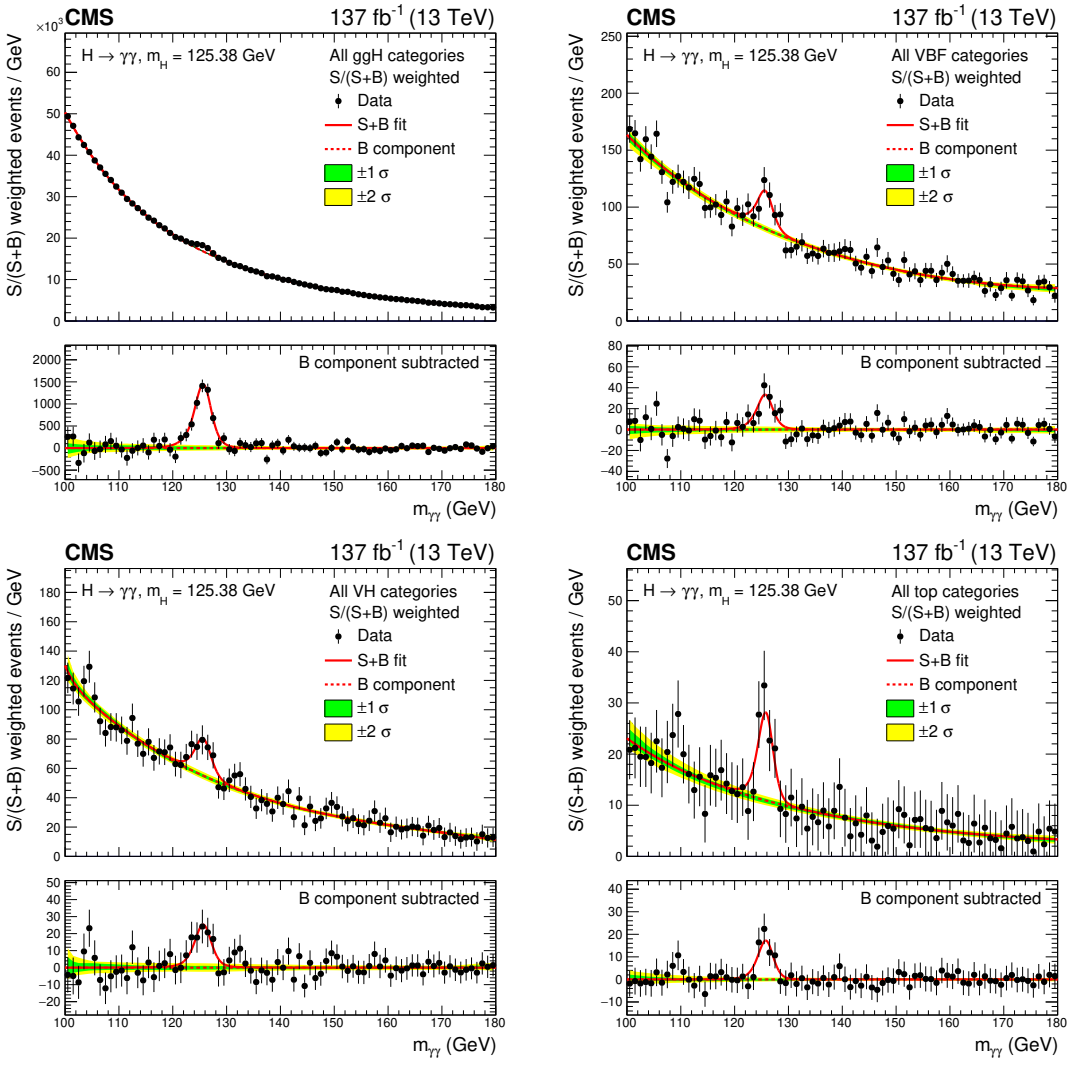


Figure 6.3: Data points (black) and the best-fit signal-plus-background model for groups of analysis categories targeting ggH (top-left), VBF (top-right), VH (bottom-left) and top-associated (bottom-right) production. The best-fit model corresponds to the per-production mode signal-strength fit. In each category group, the events in each individual category are weighted according to $S_{68}/(S_{68}+B_{68})$, where S_{68} and B_{68} are the expected signal and background estimates, respectively, in a $\pm 1\sigma_{\text{eff}}$ window centred on m_{H}, such that the total signal yield in each category group remains constant. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

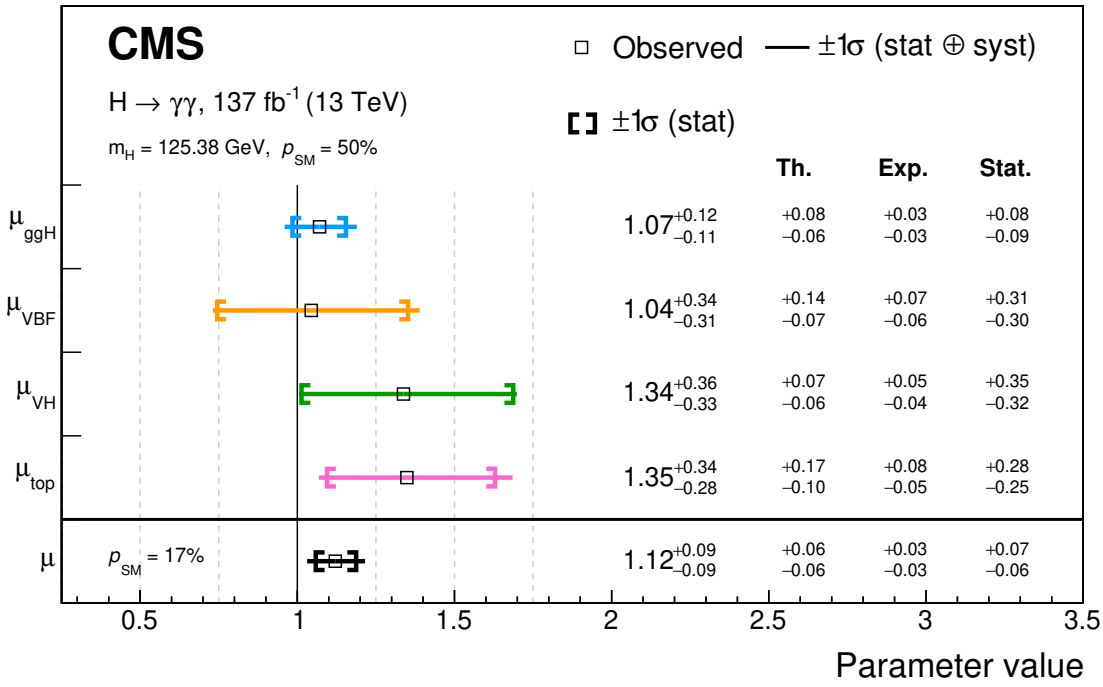


Figure 6.4: Observed best-fit values and confidence intervals for the per-production mode signal-strength parameters. The uncertainty bands are shown for including all systematic uncertainties, and the statistical uncertainty only. In addition, the results are tabulated, where the systematic uncertainty is further decomposed into the contributions from theoretical and experimental sources. Also shown in black is the result of the inclusive signal-strength modifier fit.

associated production modes. In each group, the individual categories are again summed and weighted according to the procedure defined by equation 5.12. In all cases, the signal peak is clearly visible amongst the smoothly falling background distribution. The equivalent plots for each individual analysis category are presented in Appendix E.

The best-fit values of the per-production mode signal-strength modifiers and their respective uncertainties are summarised in Figure 6.4. Again, the uncertainties are decomposed into the theoretical systematic, experimental systematic, and statistical components. For μ_{ggH}, the systematic and statistical uncertainties are comparable, whereas for other parameters the dominant source of uncertainty is statistical in origin. All production modes are observed to have a signal-strength larger than unity. Nevertheless, the results are compatible with the SM prediction, corresponding to a p-value of $p_{\mathrm{SM}} = 50\%$.

The correlation coefficients between the fitted parameters are displayed in Figure 6.5. The coefficients give a measure of the relationship between two signal-strengths in the likelihood, and are calculated using the second derivatives of $q(\vec{\alpha})$, assuming symmetric uncertainties. The correlations arise as events from one production mode migrate

into analysis categories targeting other production modes (see confusion matrix in Figure 4.20). As a result, the measurements are not completely orthogonal and there exists some correlation between the parameters of interest. In general, the correlation coefficients are small, with the largest value of -0.20 occurring between the μ_{ggH} and μ_{VBF} parameters due to the contamination of VBF events in the analysis categories targeting ggH and vice versa.

Finally, the main sources of systematic uncertainty affecting each production mode signal-strength modifier are presented in Figure 6.6. The impact from each source, or group of sources, is derived using the procedure detailed in Section 5.3.1. The top half of the plot shows the experimental sources of uncertainty, whereas the bottom half shows the theoretical sources of uncertainty, including the impact of the $\vec{\theta}_s^{\text{th}}$ nuisance parameters which are directly folded into the measurement in the signal-strength parametrisation. The dominant contributions to the uncertainties in the measured parameters are theoretical in origin. For the μ_{ggH}, μ_{VH} and μ_{top} parameters, the largest impact comes from the corresponding renormalisation and factorisation scale uncertainties. For μ_{VBF}, the dominant source of uncertainty is in the modelling of the underlying event and parton shower. These are particularly important for VBF production due to the presence of additional jets in the events. The largest experimental uncertainties originate from the integrated luminosity, the photon identification, and the photon-energy measurement for the μ_{ggH} and μ_{VH} parameters. The uncertainties in the jet-energy scale and resolution have a larger impact on μ_{VBF} and μ_{top}, where μ_{top} has an additional large contribution from the uncertainty in the b tagging.

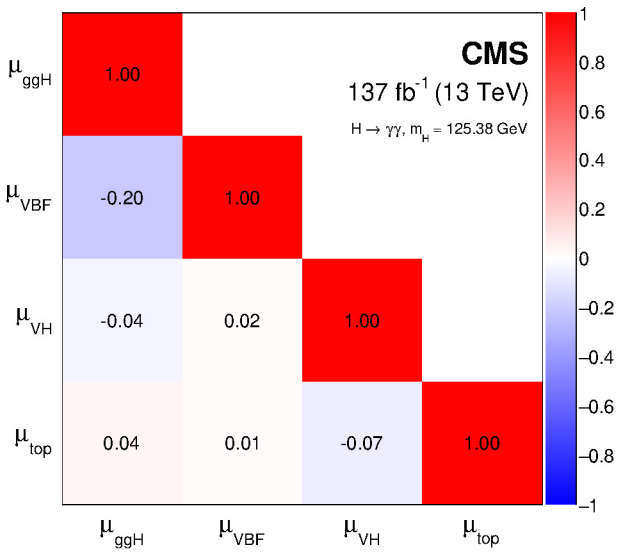


Figure 6.5: Observed correlations between the parameters in the per-production mode signal-strength fit. The size of the correlations is indicated by the colour scale.

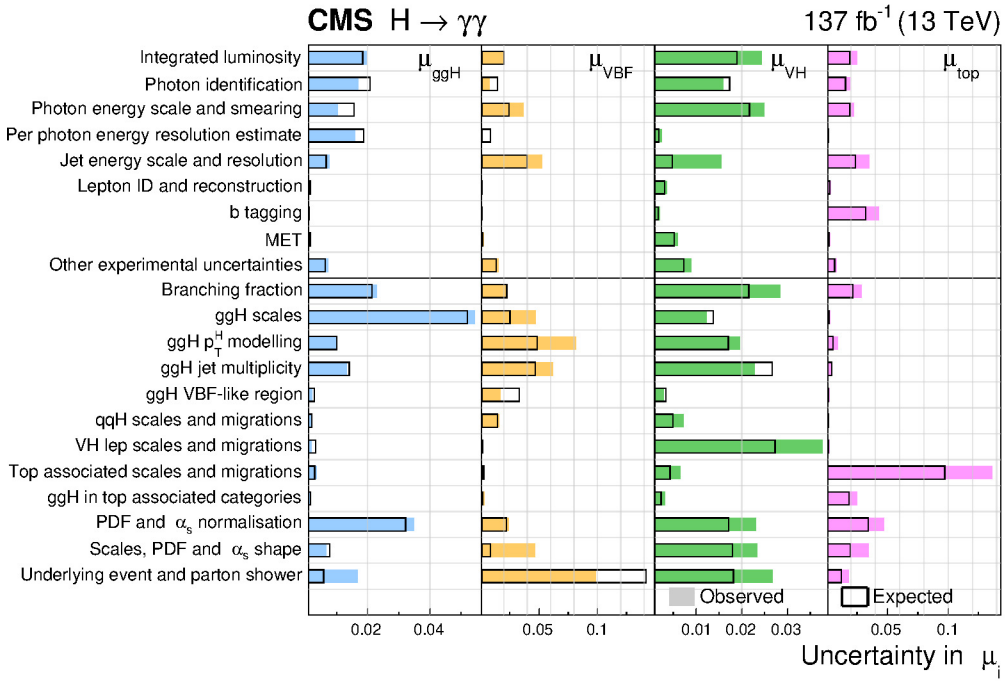


Figure 6.6: A summary of the absolute impact of the main sources of systematic uncertainty in the four parameters: μ_{ggH}, μ_{VBF}, μ_{VH}, μ_{top}. The uncertainties are split into experimental (top half) and theoretical (bottom half) sources. The observed (expected) impacts are shown by the solid (empty) bars.

6.3 STXS measurements

This section describes the extraction of cross sections in the STXS framework and their respective 68% confidence intervals and correlation coefficients. The event categorisation, described in Section 4.4, is optimised to measure as many STXS stage 1.2 bins as possible. Nevertheless, given the current available statistics, it is not possible to accurately measure all bins simultaneously using the H $\rightarrow \gamma\gamma$ decay channel alone. Three STXS bin-merging schemes are defined with varying levels of granularity to ensure reasonable sensitivity to the measured parameters, and avoid very-high correlations between them. In the construction of the schemes there is a trade-off between model-dependence and the size of uncertainties. Merging fewer bins keeps the model-dependence as low as possible, as no additional assumptions are made about the relative contributions of different STXS bins. However, this reduced model-dependence comes at the cost of larger uncertainties in the measured cross section parameters.

In contrast to the signal-strength fits, the VH hadronic processes are grouped with VBF production to define the qqH parameters (see Section 2.3), and bbH and ggZH production in which the Z boson decays hadronically are grouped with ggH. Furthermore, the theory uncertainty treatment is different to the signal-strength and coupling-modifier fits. This difference reflects the distinction between cross section *measurements* and *interpretations*. Nuisance parameters that directly affect the SM predictions of the cross sections and branching fraction, $\vec{\theta}_s^{\,\mathrm{th}}$, cancel in equation 5.4, and therefore are not included in the measurements. That said, when merging STXS bins and thus re-defining the signal processes, it is necessary to introduce the relevant θ_s^{th} nuisance parameter that accounts for the migration of events across the merged boundary.

In each fit the cross section parameters of interest are limited to the positive domain: $\sigma_{\mathrm{obs}} \geq 0$ fb. This eliminates the possibility of the signal-plus-background model going below zero in some bins of the $m_{\gamma\gamma}$ distributions, which subsequently causes the minimisation procedure to fail.

6.3.1 STXS stage 0

A fully-merged fit is performed, corresponding to the STXS stage 0 bin definitions [42]. The parameters of interest roughly correspond to the different Higgs boson production modes, such that all kinematic boundaries for a given production-mode are merged. This scheme defines six parameters in total: ggH, qqH, WH lep, ZH lep, ttH, and tH, where tH includes the contributions from both tHq and tHW production. The best-fit values of the cross sections times branching fraction, $\sigma_{\mathrm{obs}}\,\mathcal{B}$, and their respective 68% confidence levels are shown in Figure 6.7. In the plot, the colour scheme has been chosen to match

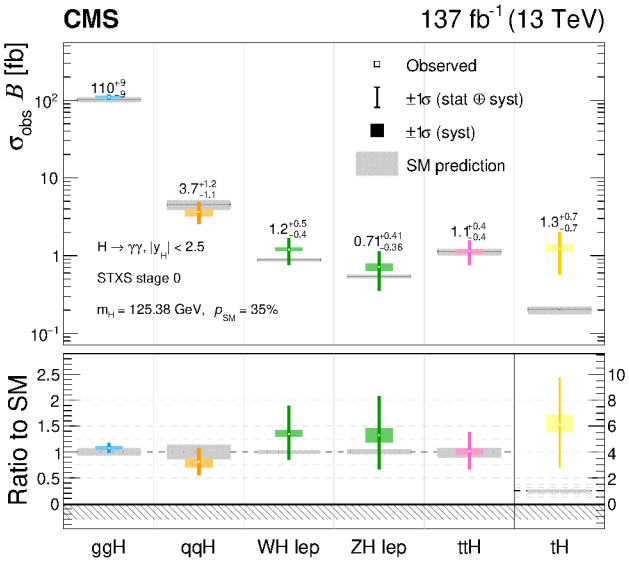


Figure 6.7: Observed results of the STXS stage 0 fit. The best-fit cross sections times branching fraction are plotted along with the respective 68% confidence intervals. The systematic components of the uncertainty in each parameter are shown by the coloured boxes. The hatched grey boxes represent the theoretical uncertainty in the SM predictions. The bottom panel shows the ratio of the fitted values to the SM predictions. The compatibility of this fit with the SM prediction is approximately $p_{\mathrm{SM}} = 35\%$.

the STXS bin schematic in Figure 2.11. The systematic components of the measured uncertainties are shown by the coloured boxes, highlighting that for all parameters except ggH, the statistical uncertainty dominates. The 68% uncertainties in the SM predicted values are shown by the hatched grey boxes, where the size of the uncertainty is computed by adding the effect of the $\vec{\theta}_s^{\mathrm{th}}$ nuisance parameters (which do not enter the measurement) in quadrature.

All measurements are consistent with the SM predicted values within one standard deviation, barring the tH parameter which is measured to have an excess of around six times the SM prediction. Despite this, the low expected tH event yield results in a large statistical uncertainty, meaning the excess corresponds to a less than 2σ deviation from the SM. The overall compatibility with the SM is measured to be $p_{\mathrm{SM}} = 35\%$. The correlation coefficients between parameters are displayed in Figure 6.8. The largest (anti-)correlation exists between the ttH and tH parameters, due to the significant contamination of ttH events in the tHq leptonic analysis category.

Importantly, this is the first dedicated measurement of single-top associated production in the $\mathrm{H} \to \gamma\gamma$ channel at CMS. The observed (expected) limit on tH production at 95% CL, using the $\mathrm{CL_s}$ procedure detailed in Ref. [162], is 14 (8) times the SM expectation. This represents one of the most stringent constraints on tH production to-date.

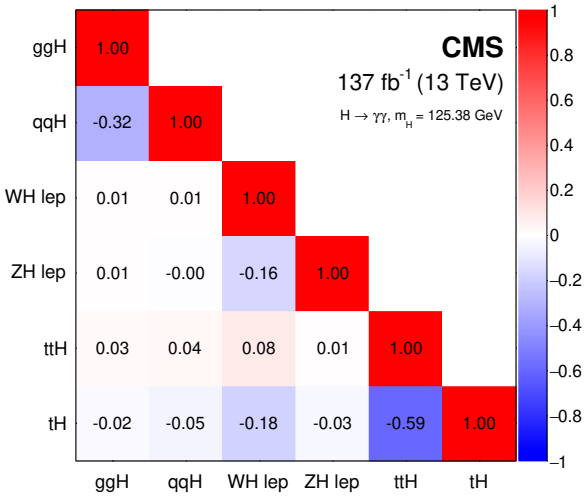


Figure 6.8: Observed correlations between the six parameters in the STXS stage 0 fit. The size of the correlations is indicated by the colour scale.

6.3.2 Maximal-merging scheme

Increasing in granularity, the *maximal-merging scheme* defines 17 parameters of interest which begin to target a number of the STXS kinematic splittings. In the scheme, bins are merged until their expected uncertainty is less than 150% of the SM prediction. This is true for all parameters except tH, which is again measured separately. The VBF-like regions (≥2J, $m_{jj} > 350$ GeV) in the ggH and qqH schemes are merged to define the ggH VBF-like and qqH VBF-like parameters, respectively. The four bins with $p_T^H > 200$ GeV in the ggH scheme are merged into a single bin, labelled as ggH BSM. Additionally, the WH leptonic, ZH leptonic, and ttH bins are fully merged into single parameters. The ZH leptonic parameter groups both (qq-initiated) ZH lep and ggZH lep production. Due to a lack of constraining power, in this fit the 0J, 1J, $m_{jj} < 60$ and $120 < m_{jj} < 350$ bins in the qqH binning scheme are constrained to their SM prediction within theory uncertainties. A schematic representation of the merging scheme is presented in Figure 6.9, whilst Appendix F tabulates the STXS bins which contribute to each parameter.

Figure 6.10 shows the $\sigma_{\text{obs}}\,\mathcal{B}$ best-fit values and 68% confidence intervals in the maximal-merging scheme, plotted in the same style as the STXS stage 0 results. With this level of kinematic splitting, the statistical component of the uncertainty dominates for all parameters. This motivates the need to increase the size of the data set, which will in turn provide substantial improvements in the precision of the measured quantities. Moreover, for a number of parameters, the uncertainty in the measurement is becoming comparable

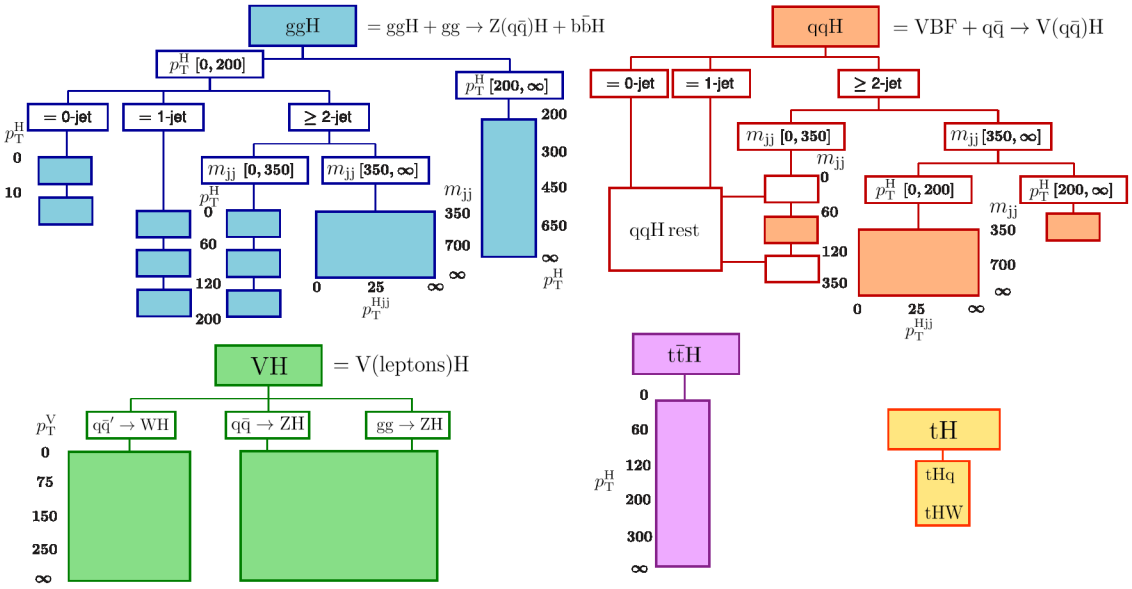


Figure 6.9: Schematic to show the maximal-merging scheme which defines 17 parameters of interest. Each parameter of interest is shown as a single coloured box.

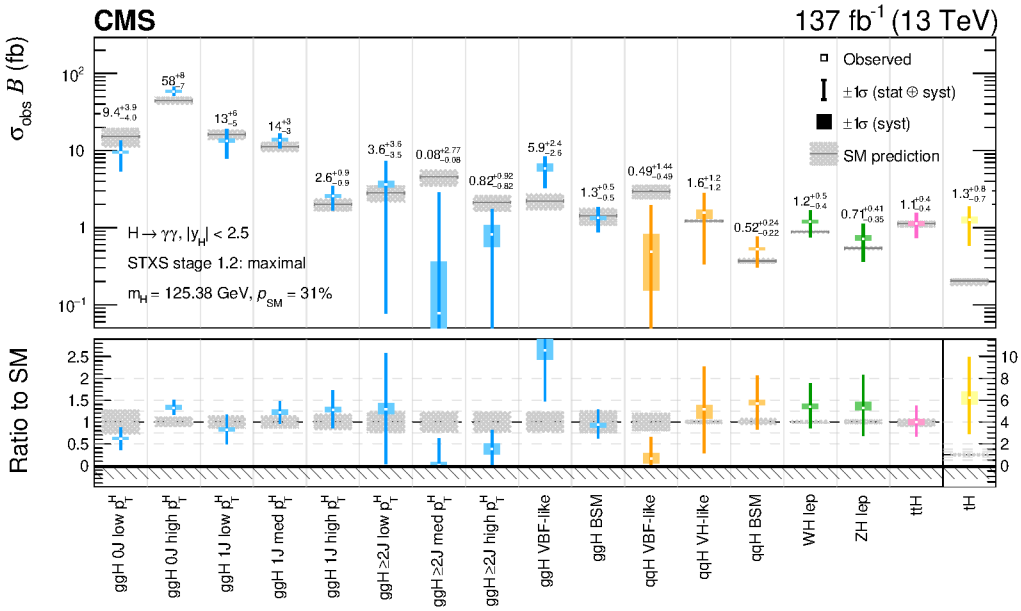


Figure 6.10: Observed results of the STXS maximal-merging fit. The best-fit cross sections times branching fraction are plotted along with the respective 68% confidence intervals. The systematic components of the uncertainty in each parameter are shown by the coloured boxes. The hatched grey boxes represent the theoretical uncertainty in the SM predictions. The bottom panel shows the ratio of the fitted values to the SM predictions. The compatibility of this fit with the SM prediction is approximately $p_{\mathrm{SM}} = 31\%$.

to the uncertainty in the SM prediction, meaning the possibility of constraining Higgs boson theory using experimental measurements is approaching. This is especially interesting for the ggH BSM parameter, which is particularly sensitive to new physics in the ggH loop. That said, the measured value of the ggH BSM $\sigma_{\text{obs}}\,\mathcal{B}$ is in excellent agreement with the SM, with a measured value of $0.9^{+0.4}_{-0.3}$, relative to the SM prediction. The overall compatibility with the SM is $p_{\text{SM}} = 31\%$. Table 6.1 summarises these results, providing in addition the expected uncertainties in the parameters, derived using the Asimov data set.

The correlation coefficients between the maximal-merging parameters are shown in Figure 6.11. As expected, the larger correlations appear for pairs of parameters which correspond to the significantly populated off-diagonal bins in the confusion matrix of Figure 4.20. This effect arises as it is difficult to unfold the signal in a given analysis category, when there are significant contributions from multiple truth-level STXS bins. For the ggH parameters, the correlations are small for parameters representing adjacent p_T^H bins but larger for parameters representing adjacent N_{jet} bins. This stems from the fact that $p_T^{\gamma\gamma}$ is a well-measured quantity, whereas reconstructing the number of jets in an event is a more difficult problem, and hence there are larger migrations between jet-multiplicity bins. Nevertheless, the application of the ggH BDT in the event categorisation helps to reduce these correlations. The largest correlations exist between the qqH VBF-like and ggH VBF-like parameters (-0.76), and the tH and ttH parameters (-0.59). This results from the difficulty in distinguishing qqH from ggH production in the VBF-like phase-space, and in distinguishing tH from ttH, attempted by the dijet BDT and Top DNN, respectively. Two-dimensional likelihood scans are performed for each pair of highly-correlated parameters to gain a better understanding of the impact of their correlations on the total $q(\vec{\alpha})$ surface. The 68% and 95% confidence regions are plotted for ggH VBF-like vs qqH VBF-like (left) and tH vs ttH (right) in Figure 6.12. In the scans, all other maximal-merging parameters are profiled. Both scans show compatibility with the SM within the 95% confidence contour.

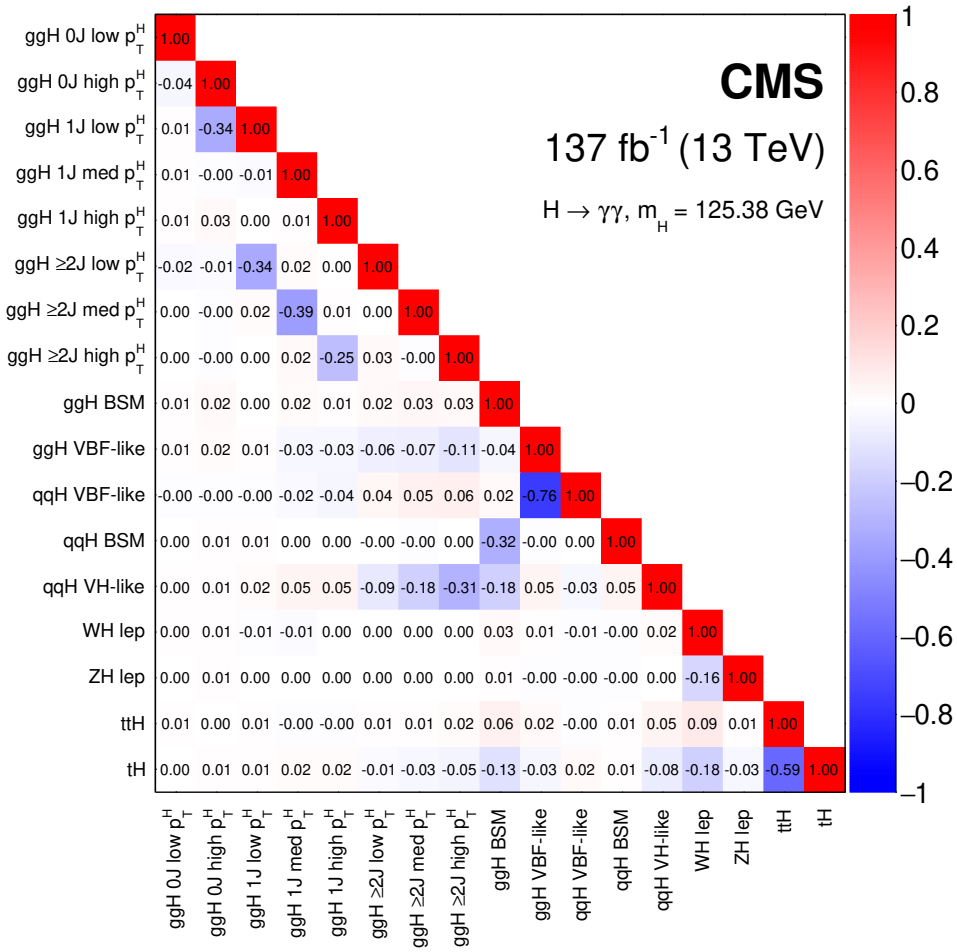


Figure 6.11: Observed correlations between the 17 parameters in the STXS maximal-merging fit. The size of the correlations is indicated by the colour scale.

Table 6.1: The best-fit cross sections times branching fraction with 68% confidence intervals for the STXS maximal-merging fit. The uncertainty is decomposed into the systematic and statistical components. The expected uncertainties in the fitted parameters are given in brackets. Also listed are the SM predictions for the cross sections times branching fraction and the theoretical uncertainty in these predictions. The final column shows the ratio of the observed value to the SM prediction.

Parameters	$\sigma\mathcal{B}$ (fb)				$\sigma\mathcal{B}/(\sigma\mathcal{B})_{\text{SM}}$
	SM prediction	Observed (Expected)			Observed (Expected)
	($m_H = 125.38$)	Best fit	Stat. unc.	Syst. unc.	Best fit
ggH 0J low p_T^H	$15.21^{+4.14}_{-4.18}$	$9.41^{+3.92}_{-3.99}\left(^{+4.20}_{-4.06}\right)$	$^{+3.90}_{-3.98}\left(^{+4.16}_{-4.05}\right)$	$^{+0.44}_{-0.25}\left(^{+0.51}_{-0.33}\right)$	$0.62^{+0.26}_{-0.26}\left(^{+0.28}_{-0.27}\right)$
ggH 0J high p_T^H	$44.25^{+4.84}_{-4.61}$	$58.50^{+8.10}_{-7.17}\left(^{+7.87}_{-7.77}\right)$	$^{+7.70}_{-6.91}\left(^{+7.67}_{-7.63}\right)$	$^{+2.50}_{-1.92}\left(^{+1.78}_{-1.42}\right)$	$1.32^{+0.18}_{-0.16}\left(^{+0.18}_{-0.18}\right)$
ggH 1J low p_T^H	$16.20^{+2.25}_{-2.27}$	$13.39^{+5.58}_{-5.49}\left(^{+5.67}_{-5.59}\right)$	$^{+5.52}_{-5.45}\left(^{+5.61}_{-5.56}\right)$	$^{+0.80}_{-0.63}\left(^{+0.77}_{-0.48}\right)$	$0.83^{+0.34}_{-0.34}\left(^{+0.35}_{-0.34}\right)$
ggH 1J med p_T^H	$11.23^{+1.56}_{-1.55}$	$13.66^{+2.91}_{-2.96}\left(^{+3.15}_{-3.39}\right)$	$^{+2.83}_{-2.92}\left(^{+3.09}_{-3.36}\right)$	$^{+0.70}_{-0.50}\left(^{+0.59}_{-0.45}\right)$	$1.22^{+0.26}_{-0.26}\left(^{+0.28}_{-0.30}\right)$
ggH 1J high p_T^H	$2.00^{+0.36}_{-0.36}$	$2.56^{+0.90}_{-0.87}\left(^{+0.91}_{-0.92}\right)$	$^{+0.90}_{-0.87}\left(^{+0.90}_{-0.90}\right)$	$^{+0.11}_{-0.11}\left(^{+0.15}_{-0.19}\right)$	$1.28^{+0.45}_{-0.44}\left(^{+0.46}_{-0.46}\right)$
ggH $\geq$2J low p_T^H	$2.82^{+0.68}_{-0.68}$	$3.62^{+3.65}_{-3.55}\left(^{+3.73}_{-2.82}\right)$	$^{+3.62}_{-3.53}\left(^{+3.69}_{-2.82}\right)$	$^{+0.41}_{-0.31}\left(^{+0.55}_{-0.55}\right)$	$1.29^{+1.29}_{-1.26}\left(^{+1.32}_{-1.00}\right)$
ggH $\geq$2J med p_T^H	$4.53^{+1.07}_{-1.07}$	$0.08^{+2.77}_{-0.08}\left(^{+2.87}_{-2.82}\right)$	$^{+2.76}_{-0.08}\left(^{+2.84}_{-2.82}\right)$	$^{+0.28}_{-0.08}\left(^{+0.38}_{-0.14}\right)$	$0.02^{+0.61}_{-0.02}\left(^{+0.63}_{-0.62}\right)$
ggH $\geq$2J high p_T^H	$2.12^{+0.49}_{-0.50}$	$0.82^{+0.92}_{-0.82}\left(^{+1.15}_{-1.10}\right)$	$^{+0.88}_{-0.82}\left(^{+1.11}_{-1.09}\right)$	$^{+0.26}_{-0.26}\left(^{+0.31}_{-0.14}\right)$	$0.39^{+0.43}_{-0.39}\left(^{+0.54}_{-0.52}\right)$
ggH VBF-like	$2.22^{+0.52}_{-0.52}$	$5.86^{+2.45}_{-2.59}\left(^{+2.90}_{-2.22}\right)$	$^{+2.27}_{-2.55}\left(^{+2.81}_{-2.22}\right)$	$^{+0.92}_{-0.48}\left(^{+0.71}_{-0.71}\right)$	$2.64^{+1.10}_{-1.17}\left(^{+1.31}_{-1.00}\right)$
ggH BSM	$1.43^{+0.36}_{-0.35}$	$1.34^{+0.50}_{-0.47}\left(^{+0.59}_{-0.49}\right)$	$^{+0.49}_{-0.46}\left(^{+0.58}_{-0.49}\right)$	$^{+0.05}_{-0.09}\left(^{+0.09}_{-0.05}\right)$	$0.94^{+0.35}_{-0.33}\left(^{+0.41}_{-0.35}\right)$
qqH VBF-like	$2.96^{+0.59}_{-0.59}$	$0.49^{+1.44}_{-0.49}\left(^{+1.49}_{-1.53}\right)$	$^{+1.40}_{-0.49}\left(^{+1.47}_{-1.47}\right)$	$^{+0.34}_{-0.34}\left(^{+0.25}_{-0.43}\right)$	$0.17^{+0.49}_{-0.17}\left(^{+0.50}_{-0.52}\right)$
qqH VH-like	$1.22^{+0.05}_{-0.04}$	$1.57^{+1.20}_{-1.24}\left(^{+1.15}_{-1.23}\right)$	$^{+1.19}_{-1.21}\left(^{+1.15}_{-1.23}\right)$	$^{+0.13}_{-0.26}\left(^{+0.07}_{-0.04}\right)$	$1.29^{+0.98}_{-1.01}\left(^{+0.94}_{-1.01}\right)$
qqH BSM	$0.37^{+0.03}_{-0.02}$	$0.52^{+0.24}_{-0.22}\left(^{+0.26}_{-0.23}\right)$	$^{+0.24}_{-0.22}\left(^{+0.25}_{-0.23}\right)$	$^{+0.03}_{-0.01}\left(^{+0.03}_{-0.01}\right)$	$1.42^{+0.65}_{-0.59}\left(^{+0.69}_{-0.62}\right)$
WH lep	$0.88^{+0.03}_{-0.03}$	$1.19^{+0.49}_{-0.44}\left(^{+0.51}_{-0.42}\right)$	$^{+0.48}_{-0.43}\left(^{+0.50}_{-0.41}\right)$	$^{+0.07}_{-0.04}\left(^{+0.05}_{-0.05}\right)$	$1.35^{+0.55}_{-0.49}\left(^{+0.57}_{-0.47}\right)$
ZH lep	$0.54^{+0.03}_{-0.02}$	$0.71^{+0.41}_{-0.35}\left(^{+0.42}_{-0.35}\right)$	$^{+0.40}_{-0.35}\left(^{+0.41}_{-0.35}\right)$	$^{+0.07}_{-0.03}\left(^{+0.06}_{-0.03}\right)$	$1.32^{+0.76}_{-0.65}\left(^{+0.78}_{-0.65}\right)$
ttH	$1.13^{+0.08}_{-0.11}$	$1.13^{+0.42}_{-0.39}\left(^{+0.42}_{-0.41}\right)$	$^{+0.42}_{-0.38}\left(^{+0.41}_{-0.40}\right)$	$^{+0.07}_{-0.07}\left(^{+0.09}_{-0.05}\right)$	$1.00^{+0.37}_{-0.35}\left(^{+0.37}_{-0.36}\right)$
tH	$0.20^{+0.01}_{-0.03}$	$1.27^{+0.76}_{-0.69}\left(^{+0.76}_{-0.20}\right)$	$^{+0.75}_{-0.68}\left(^{+0.76}_{-0.20}\right)$	$^{+0.10}_{-0.13}\left(^{+0.08}_{-0.08}\right)$	$6.24^{+3.72}_{-3.37}\left(^{+3.73}_{-1.00}\right)$

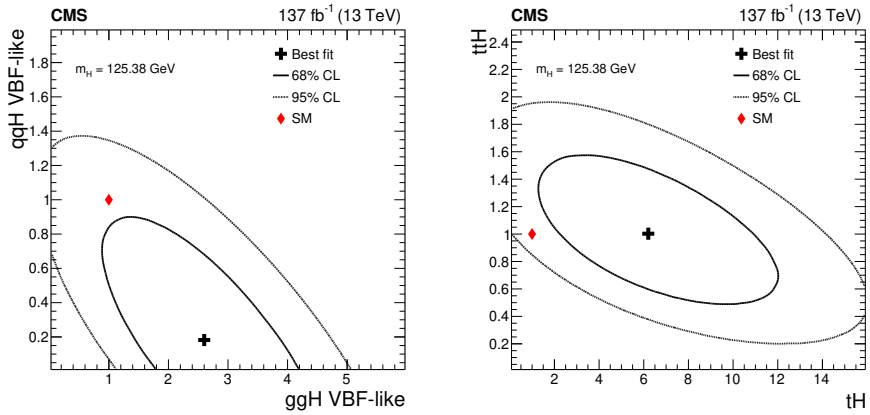


Figure 6.12: Two-dimensional $q(\vec{\alpha})$ surfaces for the pairs of parameters in the maximal-merging scheme with the largest correlations: ggH VBF-like vs qqH VBF-like (left) and tH vs ttH (right). The best-fit value along with the 68% and 95% confidence interval contours are shown by the black cross, solid line and dashed line, respectively. The parameters are plotted as a ratio with respect to their SM prediction.

6.3.3 Minimal-merging scheme

The most-granular fit performed in this analysis is in the so-called *minimal-merging scheme*. In this scheme a total of 27 parameters of interest are defined, with the aim to merge as few bins as possible whilst ensuring that the correlations between parameters remain smaller than approximately 0.9.[1] In contrast to the maximal scheme, the qqH VBF-like region is fully split into the four STXS bins defined by the boundaries at $m_{jj} = 700$ GeV and $p_T^{Hjj} = 25$ GeV. To avoid large correlations, the four ggH VBF-like bins are merged with the corresponding bins in the qqH scheme. This approach is more model-independent, as the fit makes no attempt to separate ggH and VBF production in a very similar phase-space. Additional splittings are introduced in the ggH scheme at $p_T^H = 300$ and 450 GeV, and the WH leptonic scheme at $p_T^V = 75$ and 150 GeV. Furthermore, the ttH region is fully split into five parameters according to the boundaries at $p_T^H = 60$, 120, 200 and 300 GeV. Again the 0J, 1J, $m_{jj} < 60$ and $120 < m_{jj} < 350$ bins in the qqH binning scheme are constrained to their SM prediction within theory uncertainties. The full list of STXS bins which contribute to each parameter are provided in Appendix F, with the respective schematic for the minimal-merging scheme shown in Figure 6.13.

The $\sigma_{\text{obs}}\,\mathcal{B}$ best-fit values and 68% confidence intervals in the minimal-merging fit are shown in Figure 6.14, and are listed in Table 6.2. This result represents the most granular STXS measurement performed by the CMS experiment to-date, showing reasonable

[1] This threshold was chosen as the fits became unstable for correlations above 0.9.

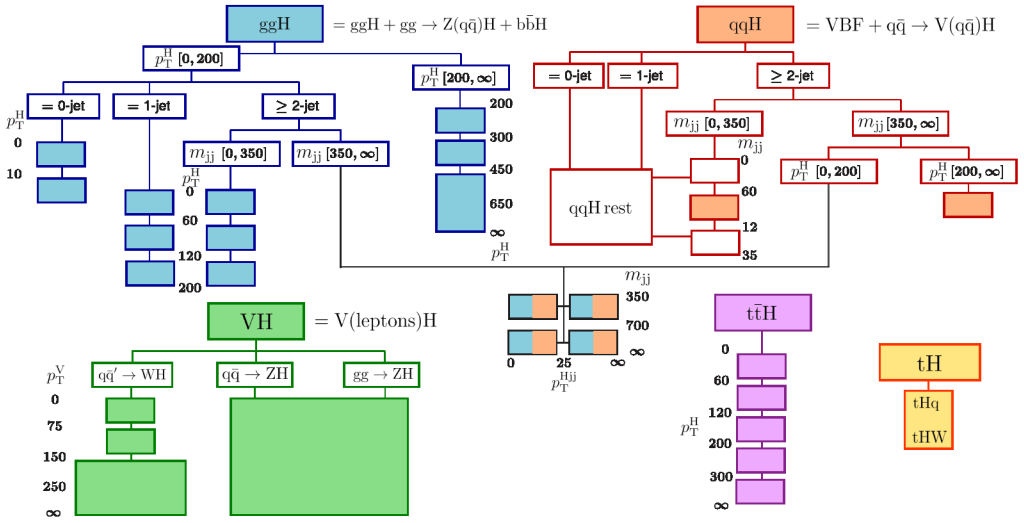


Figure 6.13: Schematic to show the minimal-merging scheme which defines 27 parameters of interest. Each parameter of interest is shown as a single coloured box. The VBF-like bins in the ggH and qqH binning schemes are merged and are represented in the diagram by the half-blue half-orange boxes.

sensitivity to many different kinematic regions of Higgs boson production phase-space. In particular, the result contains the first measurements of ttH production in different p_T^H bins, where the size of the uncertainty in each of the four bins with $p_T^H < 300$ GeV is less than 100% of the SM prediction. Since the uncertainties in all measured parameters are dominated by the statistical component, there is considerable room for improvement by taking more data and combining with the results from other Higgs boson decay channels. Ultimately, this divide-and-measure approach of the STXS framework allows to systematically constrain increasingly granular regions of phase space, providing sensitivity to BSM physics which may appear in specific kinematic bins. The results here are highly compatible with the SM hypothesis, with a corresponding p-value of $p_{\text{SM}} = 70\%$.

The correlations between parameters are shown in Figure 6.15. By merging the qqH VBF-like STXS bins with the equivalent bins in the ggH scheme, the correlations in this region of phase space are kept to an acceptable level. The region with the highest correlations is that defined by the high p_T^H ttH and tH parameters. With such high correlations, the act of splitting ttH into five separate parameters forces the tH $\sigma_{\text{obs}}\,\mathcal{B}$ to an even higher value, corresponding to an excess of approximately eight times the SM prediction.

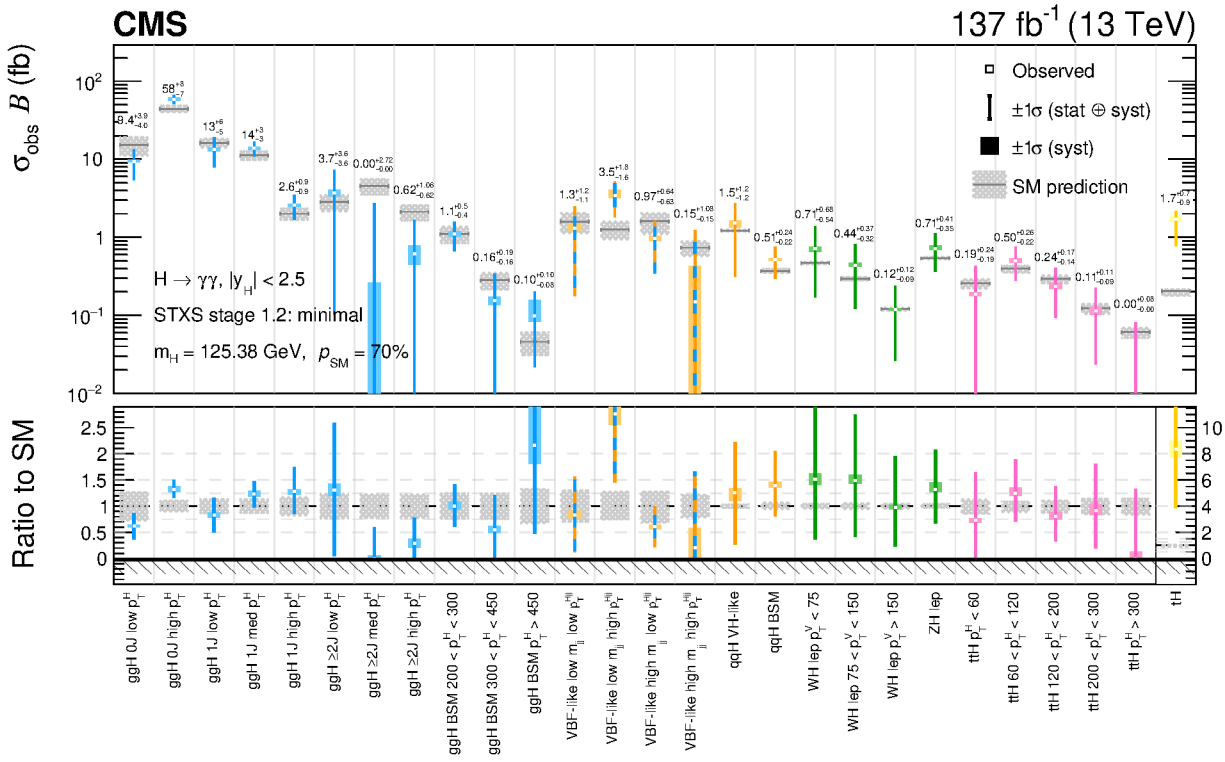


Figure 6.14: Observed results of the STXS minimal-merging fit. The best-fit cross sections times branching fraction are plotted along with the respective 68% confidence intervals. The systematic components of the uncertainty in each parameter are shown by the coloured boxes. The hatched grey boxes represent the theoretical uncertainty in the SM predictions. The bottom panel shows the ratio of the fitted values to the SM predictions. The colour scheme has been chosen to match the STXS schematic in Figure 2.11, such that the orange and blue dashed lines for the VBF-like parameters represent the contributions from both ggH and qqH STXS bins. The compatibility of this fit with the SM prediction is approximately $p_{\mathrm{SM}} = 70\%$.

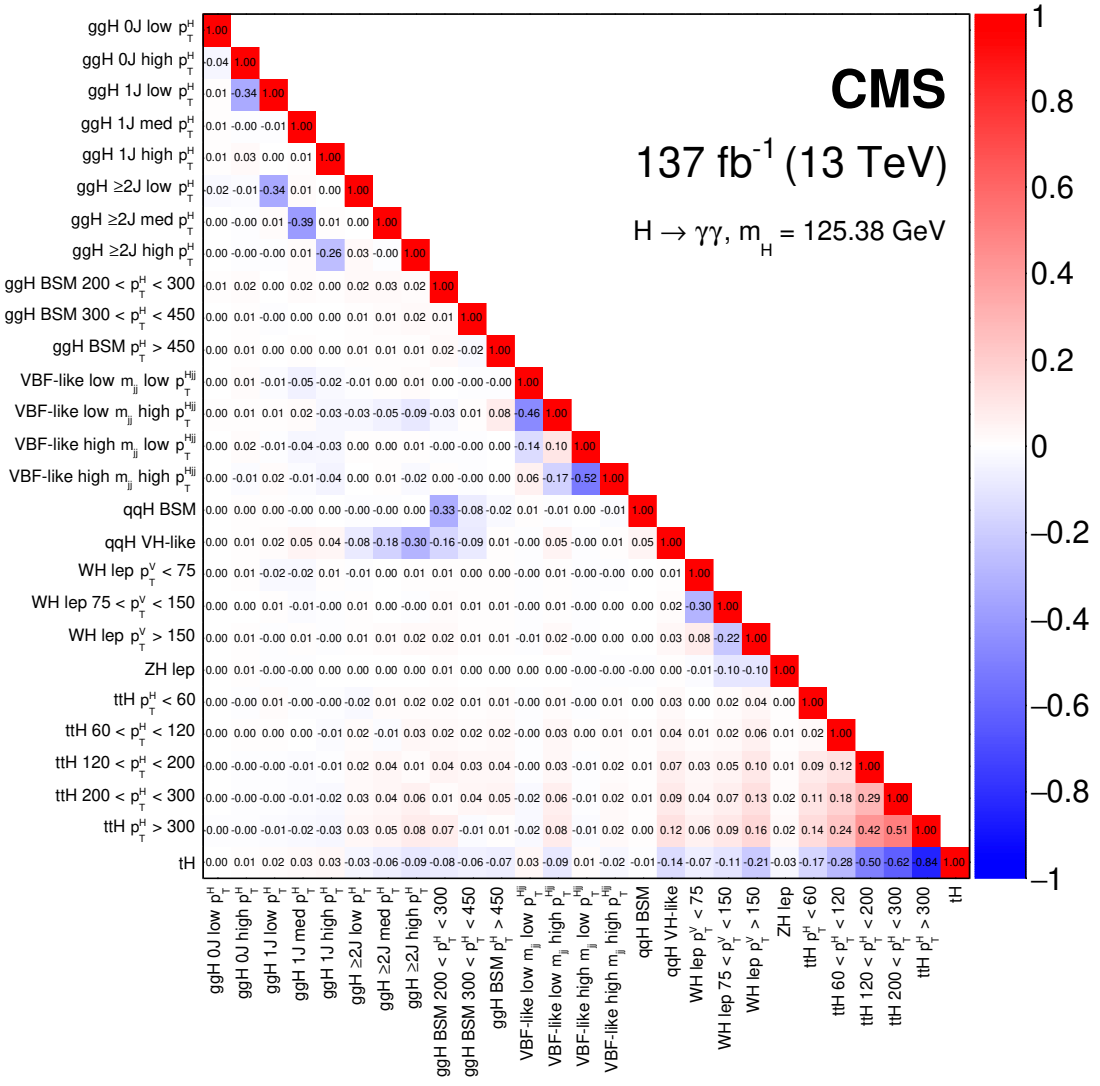


Figure 6.15: Observed correlations between the 27 parameters in the STXS minimal-merging fit. The size of the correlations is indicated by the colour scale.

Table 6.2: The best-fit cross sections times branching fraction with 68% confidence intervals for the STXS minimal-merging fit. The uncertainty is decomposed into the systematic and statistical components. The expected uncertainties in the fitted parameters are given in brackets. Also listed are the SM predictions for the cross sections times branching fraction and the theoretical uncertainty in these predictions. The final column shows the ratio of the observed value to the SM prediction.

Parameters	SM prediction	$\sigma\mathcal{B}$ (fb) Observed (Expected)			$\sigma\mathcal{B}/(\sigma\mathcal{B})_{\mathrm{SM}}$ Observed (Expected)
	($m_H = 125.38$)	Best fit	Stat. unc.	Syst. unc.	Best fit
ggH 0J low p_T^H	$15.21^{+4.14}_{-4.18}$	$9.41^{+3.91}_{-4.00}\left(^{+4.19}_{-4.06}\right)$	$^{+3.90}_{-3.99}\left(^{+4.16}_{-4.05}\right)$	$^{+0.37}_{-0.30}\left(^{+0.50}_{-0.36}\right)$	$0.62^{+0.26}_{-0.26}\left(^{+0.28}_{-0.27}\right)$
ggH 0J high p_T^H	$44.25^{+4.84}_{-4.61}$	$58.46^{+8.12}_{-7.17}\left(^{+7.87}_{-7.78}\right)$	$^{+7.69}_{-6.91}\left(^{+7.66}_{-7.63}\right)$	$^{+2.60}_{-1.94}\left(^{+1.78}_{-1.50}\right)$	$1.32^{+0.18}_{-0.16}\left(^{+0.18}_{-0.18}\right)$
ggH 1J low p_T^H	$16.20^{+2.25}_{-2.27}$	$13.40^{+5.59}_{-5.50}\left(^{+5.70}_{-5.58}\right)$	$^{+5.53}_{-5.46}\left(^{+5.64}_{-5.55}\right)$	$^{+0.79}_{-0.67}\left(^{+0.77}_{-0.56}\right)$	$0.83^{+0.34}_{-0.34}\left(^{+0.35}_{-0.34}\right)$
ggH 1J med p_T^H	$11.23^{+1.56}_{-1.55}$	$13.80^{+2.90}_{-2.94}\left(^{+3.14}_{-3.41}\right)$	$^{+2.82}_{-2.90}\left(^{+3.08}_{-3.37}\right)$	$^{+0.68}_{-0.51}\left(^{+0.59}_{-0.50}\right)$	$1.23^{+0.26}_{-0.26}\left(^{+0.28}_{-0.30}\right)$
ggH 1J high p_T^H	$2.00^{+0.36}_{-0.36}$	$2.57^{+0.94}_{-0.88}\left(^{+0.92}_{-0.90}\right)$	$^{+0.94}_{-0.87}\left(^{+0.91}_{-0.88}\right)$	$^{+0.08}_{-0.12}\left(^{+0.13}_{-0.16}\right)$	$1.28^{+0.47}_{-0.44}\left(^{+0.46}_{-0.45}\right)$
ggH $\geq$ 2J low p_T^H	$2.82^{+0.68}_{-0.68}$	$3.67^{+3.63}_{-3.57}\left(^{+3.74}_{-2.82}\right)$	$^{+3.62}_{-3.56}\left(^{+3.71}_{-2.82}\right)$	$^{+0.34}_{-0.30}\left(^{+0.49}_{-0.49}\right)$	$1.30^{+1.29}_{-1.27}\left(^{+1.33}_{-1.00}\right)$
ggH $\geq$ 2J med p_T^H	$4.53^{+1.07}_{-1.07}$	$0.00^{+2.72}_{-0.00}\left(^{+2.90}_{-2.80}\right)$	$^{+2.71}_{-0.00}\left(^{+2.86}_{-2.78}\right)$	$^{+0.26}_{-0.00}\left(^{+0.45}_{-0.27}\right)$	$0.00^{+0.60}_{-0.00}\left(^{+0.64}_{-0.62}\right)$
ggH $\geq$ 2J high p_T^H	$2.12^{+0.49}_{-0.50}$	$0.62^{+1.06}_{-0.62}\left(^{+1.15}_{-1.10}\right)$	$^{+1.04}_{-0.62}\left(^{+1.11}_{-1.10}\right)$	$^{+0.17}_{-0.17}\left(^{+0.30}_{-0.13}\right)$	$0.29^{+0.50}_{-0.29}\left(^{+0.54}_{-0.52}\right)$
ggH BSM $200 < p_T^H < 300$	$1.10^{+0.28}_{-0.27}$	$1.11^{+0.47}_{-0.44}\left(^{+0.56}_{-0.45}\right)$	$^{+0.46}_{-0.43}\left(^{+0.56}_{-0.45}\right)$	$^{+0.08}_{-0.07}\left(^{+0.05}_{-0.03}\right)$	$1.00^{+0.42}_{-0.40}\left(^{+0.51}_{-0.41}\right)$
ggH BSM $300 < p_T^H < 450$	$0.28^{+0.07}_{-0.07}$	$0.16^{+0.19}_{-0.16}\left(^{+0.20}_{-0.18}\right)$	$^{+0.18}_{-0.16}\left(^{+0.19}_{-0.18}\right)$	$^{+0.02}_{-0.02}\left(^{+0.03}_{-0.01}\right)$	$0.55^{+0.66}_{-0.55}\left(^{+0.69}_{-0.65}\right)$
ggH BSM $p_T^H > 450$	$0.05^{+0.02}_{-0.02}$	$0.10^{+0.10}_{-0.08}\left(^{+0.10}_{-0.05}\right)$	$^{+0.09}_{-0.08}\left(^{+0.09}_{-0.05}\right)$	$^{+0.06}_{-0.02}\left(^{+0.04}_{-0.04}\right)$	$2.16^{+2.25}_{-1.69}\left(^{+2.19}_{-1.00}\right)$
VBF-like low m_{jj} low p_T^{Hjj}	$1.59^{+0.49}_{-0.48}$	$1.31^{+1.19}_{-1.13}\left(^{+1.22}_{-1.16}\right)$	$^{+1.18}_{-1.13}\left(^{+1.21}_{-1.16}\right)$	$^{+0.14}_{-0.09}\left(^{+0.13}_{-0.05}\right)$	$0.82^{+0.75}_{-0.71}\left(^{+0.77}_{-0.73}\right)$
VBF-like low m_{jj} high p_T^{Hjj}	$1.25^{+0.35}_{-0.32}$	$3.46^{+1.76}_{-1.64}\left(^{+1.79}_{-1.25}\right)$	$^{+1.65}_{-1.62}\left(^{+1.76}_{-1.25}\right)$	$^{+0.61}_{-0.25}\left(^{+0.32}_{-0.32}\right)$	$2.76^{+1.40}_{-1.31}\left(^{+1.43}_{-1.00}\right)$
VBF-like high m_{jj} low p_T^{Hjj}	$1.60^{+0.45}_{-0.51}$	$0.97^{+0.64}_{-0.63}\left(^{+0.72}_{-0.63}\right)$	$^{+0.63}_{-0.62}\left(^{+0.71}_{-0.63}\right)$	$^{+0.07}_{-0.07}\left(^{+0.11}_{-0.06}\right)$	$0.61^{+0.40}_{-0.39}\left(^{+0.45}_{-0.39}\right)$
VBF-like high m_{jj} high p_T^{Hjj}	$0.73^{+0.16}_{-0.16}$	$0.15^{+1.08}_{-0.15}\left(^{+0.93}_{-0.73}\right)$	$^{+1.04}_{-0.15}\left(^{+0.92}_{-0.73}\right)$	$^{+0.28}_{-0.15}\left(^{+0.14}_{-0.14}\right)$	$0.20^{+1.47}_{-0.20}\left(^{+1.26}_{-1.00}\right)$
qqH VH-like	$1.22^{+0.05}_{-0.05}$	$1.53^{+1.20}_{-1.21}\left(^{+1.13}_{-1.27}\right)$	$^{+1.20}_{-1.20}\left(^{+1.12}_{-1.27}\right)$	$^{+0.11}_{-0.19}\left(^{+0.05}_{-0.08}\right)$	$1.25^{+0.98}_{-0.99}\left(^{+0.92}_{-1.04}\right)$
qqH BSM	$0.37^{+0.03}_{-0.02}$	$0.51^{+0.24}_{-0.22}\left(^{+0.24}_{-0.24}\right)$	$^{+0.24}_{-0.22}\left(^{+0.24}_{-0.24}\right)$	$^{+0.03}_{-0.01}\left(^{+0.03}_{-0.02}\right)$	$1.40^{+0.66}_{-0.60}\left(^{+0.66}_{-0.65}\right)$
WH lep $p_T^V < 75$	$0.47^{+0.02}_{-0.02}$	$0.71^{+0.68}_{-0.54}\left(^{+0.75}_{-0.47}\right)$	$^{+0.68}_{-0.54}\left(^{+0.75}_{-0.47}\right)$	$^{+0.05}_{-0.05}\left(^{+0.05}_{-0.05}\right)$	$1.51^{+1.45}_{-1.15}\left(^{+1.60}_{-1.00}\right)$
WH lep $75 < p_T^V < 150$	$0.29^{+0.02}_{-0.02}$	$0.44^{+0.37}_{-0.32}\left(^{+0.37}_{-0.29}\right)$	$^{+0.37}_{-0.32}\left(^{+0.37}_{-0.29}\right)$	$^{+0.03}_{-0.02}\left(^{+0.02}_{-0.02}\right)$	$1.49^{+1.26}_{-1.08}\left(^{+1.25}_{-1.00}\right)$
WH lep $p_T^V > 150$	$0.12^{+0.01}_{-0.01}$	$0.12^{+0.12}_{-0.09}\left(^{+0.13}_{-0.10}\right)$	$^{+0.12}_{-0.09}\left(^{+0.13}_{-0.10}\right)$	$^{+0.01}_{-0.00}\left(^{+0.01}_{-0.01}\right)$	$0.98^{+0.98}_{-0.76}\left(^{+1.05}_{-0.79}\right)$
ZH lep	$0.54^{+0.03}_{-0.02}$	$0.71^{+0.41}_{-0.35}\left(^{+0.42}_{-0.35}\right)$	$^{+0.40}_{-0.35}\left(^{+0.41}_{-0.35}\right)$	$^{+0.07}_{-0.03}\left(^{+0.04}_{-0.04}\right)$	$1.32^{+0.76}_{-0.65}\left(^{+0.77}_{-0.65}\right)$
ttH $p_T^H < 60$	$0.26^{+0.03}_{-0.04}$	$0.19^{+0.24}_{-0.19}\left(^{+0.23}_{-0.19}\right)$	$^{+0.24}_{-0.19}\left(^{+0.23}_{-0.19}\right)$	$^{+0.01}_{-0.01}\left(^{+0.03}_{-0.03}\right)$	$0.73^{+0.92}_{-0.73}\left(^{+0.89}_{-0.75}\right)$
ttH $60 < p_T^H < 120$	$0.40^{+0.04}_{-0.05}$	$0.50^{+0.26}_{-0.22}\left(^{+0.28}_{-0.22}\right)$	$^{+0.25}_{-0.22}\left(^{+0.28}_{-0.22}\right)$	$^{+0.04}_{-0.02}\left(^{+0.03}_{-0.01}\right)$	$1.25^{+0.65}_{-0.55}\left(^{+0.72}_{-0.55}\right)$
ttH $120 < p_T^H < 200$	$0.29^{+0.03}_{-0.04}$	$0.24^{+0.17}_{-0.14}\left(^{+0.17}_{-0.16}\right)$	$^{+0.17}_{-0.14}\left(^{+0.17}_{-0.16}\right)$	$^{+0.02}_{-0.02}\left(^{+0.02}_{-0.01}\right)$	$0.80^{+0.58}_{-0.49}\left(^{+0.58}_{-0.53}\right)$
ttH $200 < p_T^H < 300$	$0.12^{+0.02}_{-0.02}$	$0.11^{+0.11}_{-0.09}\left(^{+0.10}_{-0.09}\right)$	$^{+0.11}_{-0.09}\left(^{+0.10}_{-0.09}\right)$	$^{+0.01}_{-0.01}\left(^{+0.01}_{-0.01}\right)$	$0.92^{+0.89}_{-0.73}\left(^{+0.81}_{-0.75}\right)$
ttH $p_T^H > 300$	$0.06^{+0.01}_{-0.01}$	$0.00^{+0.08}_{-0.00}\left(^{+0.07}_{-0.06}\right)$	$^{+0.08}_{-0.00}\left(^{+0.07}_{-0.06}\right)$	$^{+0.01}_{-0.00}\left(^{+0.02}_{-0.02}\right)$	$0.00^{+1.34}_{-0.00}\left(^{+1.21}_{-1.00}\right)$
tH	$0.20^{+0.01}_{-0.03}$	$1.71^{+0.71}_{-0.93}\left(^{+1.01}_{-0.20}\right)$	$^{+0.70}_{-0.92}\left(^{+1.00}_{-0.20}\right)$	$^{+0.13}_{-0.13}\left(^{+0.11}_{-0.11}\right)$	$8.38^{+3.48}_{-4.55}\left(^{+4.93}_{-1.00}\right)$

6.4 Coupling modifiers in the κ-framework

In the κ-framework, coupling-modifiers are introduced to directly parametrise deviations from the SM expectation in the couplings of the Higgs boson to other particles [144]. Under the assumption that there are no additional Higgs boson decays to BSM particles, the cross section times branching fraction for production mode i and decay channel f can be expressed as

$$\sigma^i \cdot \mathcal{B}^f = \sigma^i(\vec{\kappa}) \cdot \frac{\Gamma^f(\vec{\kappa})}{\Gamma^H(\vec{\kappa})}, \tag{6.1}$$

where Γ^H is the total Higgs boson decay width, and Γ^f is the partial Higgs boson decay width to final state f. Each term in equation 6.1 is expressed as a function of multiplicative coupling-modifiers, $\vec{\kappa}$, as shown in Table 6.3, where in the SM all κ values are positive and equal to unity. For example, a value of $\kappa_W = 1.2$ represents a 20% enhancement in the strength of the Higgs boson to W boson coupling.

Using the notation introduced in Section 5.1.1, the signal yields are parametrised as the product of scaling functions for each term in equation 6.1,

$$\mu^{i,\gamma\gamma} = \frac{\sigma^i(\vec{\kappa})}{\sigma^i_{\text{SM}}} \cdot \frac{\Gamma^{\gamma\gamma}(\vec{\kappa})}{\Gamma^{\gamma\gamma}_{\text{SM}}} \cdot \frac{\Gamma^H_{\text{SM}}}{\Gamma^H(\vec{\kappa})}. \tag{6.2}$$

This approach works under the narrow Higgs boson width assumption, effectively factoring the total signal parametrisation into the effect at Higgs boson production and Higgs boson decay; a crucial concept in the EFT parametrisation discussed in the following chapter.

In this analysis two independent parametrisations are considered. The first uses the resolved scaling functions listed in Table 6.3, introducing universal coupling-modifiers, κ_V and κ_F, which modify the Higgs boson couplings to vector bosons and fermions, respectively, such that,

$$\begin{aligned} \kappa_V &= \kappa_W = \kappa_Z, \\ \kappa_F &= \kappa_t = \kappa_b = \kappa_\tau = \kappa_\mu. \end{aligned} \tag{6.3}$$

As shown in Table 6.3, the tHq and tHW production mode cross sections include an interference term proportional to the product, $\kappa_V\kappa_F$. This means that by measuring single-top associated production, the analysis gains sensitivity to the relative sign of the t-H (fermion) and V-H (vector boson) couplings, unlike when measuring ttH alone. Figure 6.16 shows both the observed (black) and expected (red) $q(\kappa_F)$ curves, where the value of κ_V is profiled in the fit. In addition, the dashed red curve represents the expected likelihood that would be obtained if the tHq leptonic category is removed from

Table 6.3: Scaling functions for all the major Higgs boson production modes and decay channels. The effective κ parameters representing deviations in loop processes are provided, as well as the fully resolved scaling functions into the fundamental SM couplings.

	Loops	Interference	Effective scaling factor	Resolved scaling factor
Production				
$\sigma(ggH)$	✓	g-t	κ_g^2	$1.04\kappa_t^2 + 0.002\kappa_b^2 - 0.038\kappa_t\kappa_b$
$\sigma(VBF)$	—	—		$0.73\kappa_W^2 + 0.27\kappa_Z^2$
$\sigma(WH)$	—	—		κ_W^2
$\sigma(qq/qg \to ZH)$	—	—		κ_Z^2
$\sigma(gg \to ZH)$	✓	Z-t		$2.46\kappa_Z^2 + 0.47\kappa_t^2 - 1.94\kappa_Z\kappa_t$
$\sigma(ttH)$	—	—		κ_t^2
$\sigma(gb \to WtH)$	—	W-t		$2.91\kappa_t^2 + 2.31\kappa_W^2 - 4.22\kappa_t\kappa_W$
$\sigma(qb \to tHq)$	—	W-t		$2.63\kappa_t^2 + 3.58\kappa_W^2 - 5.21\kappa_t\kappa_W$
$\sigma(bbH)$	—	—		κ_b^2
Partial decay width				
Γ^{ZZ}	—	—		κ_Z^2
Γ^{WW}	—	—		κ_W^2
$\Gamma^{\gamma\gamma}$	✓	W-t	κ_γ^2	$1.59\kappa_W^2 + 0.07\kappa_t^2 - 0.67\kappa_W\kappa_t$
$\Gamma^{\tau\tau}$	—	—		κ_τ^2
Γ^{bb}	—	—		κ_b^2
$\Gamma^{\mu\mu}$	—	—		κ_μ^2
Total width for $\mathcal{B}_{BSM} = 0$				
Γ_H	✓	—	κ_H^2	$0.58\kappa_b^2 + 0.22\kappa_W^2 + 0.08\kappa_g^2 +$ $+ 0.06\kappa_\tau^2 + 0.026\kappa_Z^2 + 0.029\kappa_c^2 +$ $+ 0.0023\kappa_\gamma^2 + 0.0015\kappa_{Z\gamma}^2 +$ $+ 0.00025\kappa_s^2 + 0.00022\kappa_\mu^2$

the analysis. Clearly, the inclusion of the tHq leptonic category successfully reduces the degeneracy between positive and negative κ_F values. The observed likelihood shows a slight favouring for negative κ_F values with respect to the expected likelihood, due to the observed excess in the tH production cross section. Furthermore, the results of a two-dimensional likelihood scan in κ_V and κ_F are presented in the upper plot of Figure 6.17. The region with negative values of κ_F is observed (expected) to be excluded with a significance of 0.5σ (2.4σ).

The second parametrisation considered uses the effective coupling-modifiers to gluons and photons, κ_g and κ_γ, to measure potential deviations in the ggH and H $\to \gamma\gamma$ loops. The observed results of a two-dimensional likelihood scan in these two parameters is shown in the lower plot of Figure 6.17. In the scan, all other κ parameters (κ_t,κ_b,κ_W,κ_Z,κ_τ,κ_μ) are fixed to unity. The κ_g and κ_γ parameters are particularly sensitive to additional heavy BSM particles, that would contribute to the rate of Higgs boson production and decay via loop processes. The observed best-fit point is consistent with the SM expectation at around the 68% confidence level, suggesting there are no new states that add major

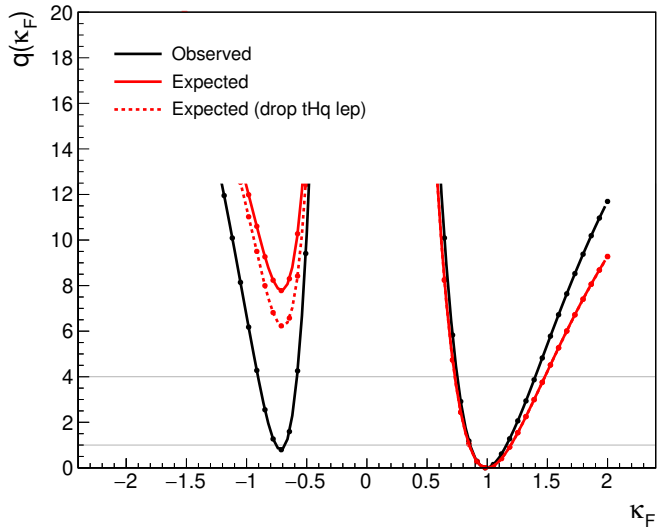


Figure 6.16: The observed (solid black) and expected (solid red) $q(\kappa_F)$ curves, where κ_V is profiled in the fit. The expected curve that would be obtained by removing the tHq leptonic event category from the analysis is shown by the dashed red line.

contributions to the loops, or that the masses of any new states are significantly higher than the electroweak energy scale.

Ultimately, the coupling-modifier scaling functions are defined inclusively for each production-mode cross section. As a result, this parametrisation does not make use of the kinematic information available in the STXS measurements. Introduced in the next chapter, the EFT parametrisation extends upon the κ-framework by defining scaling functions for each individual STXS bin. In this approach, the fit is able to use the kinematic information to more tightly constrain BSM physics.

6.5 Summary

This section concludes the description of Higgs boson production cross sections and couplings measurements using the diphoton decay channel at CMS. Chapter 4 detailed the techniques used to reconstruct events consistent with the H → γγ decay in p-p collision data, and subsequently categorise them to become sensitive to the different kinematic regions of the STXS framework. In Chapter 5, the statistical inference methods were described, which amount to performing a likelihood fit to the observed $m_{\gamma\gamma}$ distribution in each analysis category. This chapter presented the results of the analysis which represent the most comprehensive study of Higgs boson production cross sections to-date, using data collected by the CMS experiment.

A range of measurements were presented using different signal parametrisations. The

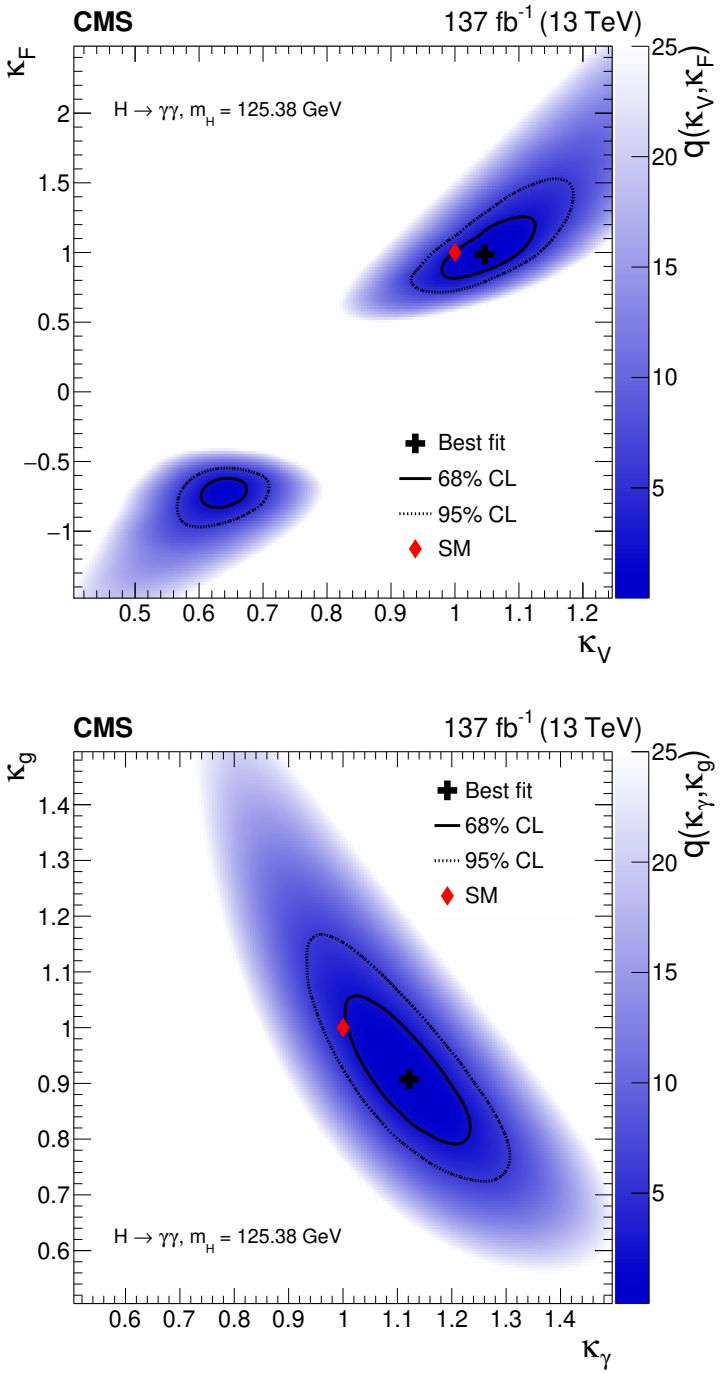


Figure 6.17: Observed two-dimensional likelihood scans performed in the κ-framework: κ_V vs κ_F using the resolved scaling functions (top), and κ_γ vs κ_g using the effective scaling functions (bottom). The 68% and 95% confidence-level regions are represented by the solid and dashed contours, respectively. The best-fit and SM expected points are shown by the black cross and red diamond, respectively. The colour scale indicates the value of the test statistic, $q(\alpha_1, \alpha_2)$.

inclusive Higgs boson signal-strength relative to the SM prediction was measured to be 1.12, with a $\pm 9\%$ uncertainty. This represents the most accurate measurement of this particular quantity in a single decay channel at CMS. A second fit was performed in the signal-strength parametrisation, where the four principal Higgs boson production modes are scaled by a separate parameter. The simultaneous fit of these four signal-strengths was found to be compatible with the SM prediction, with a p-value of 50%. Moreover, each of the four production modes are now measured with uncertainties ranging from 10–35%.

Production cross sections were measured in the STXS framework, using merging schemes of different granularities. In the maximal (minimal)-merging scheme fit, a total of 17 (27) independent kinematic regions are measured simultaneously. The minimal-merging scheme fit represents the most-granular measurement of Higgs boson production performed in a single decay channel to-date. Both fits were compatible with SM predictions, with p-values of 31% and 70% for the maximal and minimal schemes, respectively. Many of the kinematic regions were measured here for the first time, including the measurement of ttH production in five different p_T^H regions. Ultimately, by probing the different kinematic regions, we become sensitive to new physics which appears in specific kinematic regions of production phase-space, hints of which would be washed out when measuring the equivalent processes inclusively.

Furthermore, the measurement of ggH production with $p_T^H > 200$ GeV represents the most precise measurement of this particular kinematic region; a region which is sensitive to heavy BSM states appearing in the ggH loop. The cross section was found to be compatible with the SM, with a measured value of $0.9^{+0.4}_{-0.3}$, relative to the SM prediction. Also, an upper limit was placed on the rate of tH production for the first time in the H $\rightarrow \gamma\gamma$ decay channel at CMS. The observed (expected) limit at the 95% C.L. was found to be 14 (8) times the SM prediction. All other results, including measurements of the Higgs boson coupling-modifiers, were in agreement with the SM expectations.

Chapter 7

An EFT interpretation of STXS measurements

7.1 Introduction

Effective field theories were introduced in Section 2.4 as a model-independent approach to constraining BSM physics. In EFT, new BSM states are assumed to exist with masses at an energy scale, Λ, far beyond the electroweak energy scale, $v = 246$ GeV. The dynamics introduced by the BSM states can be parametrised at low energies ($E \sim v$) using higher-dimensional operators built up from the SM fields, where the operators are confined to respect both the symmetries and gauge-invariance of the SM. This expansion of the SM Lagrangian, shown explicitly in equation 2.45, is fully general and thus can be used to constrain a wide class of BSM theories that reduce to the SM at low energies. The Wilson coefficients, w_p, directly parametrise the contribution from operator, $\mathcal{O}_p$, and by constraining these coefficients it is possible to infer both the coupling strength and potential type of new BSM interactions. Ultimately, the final constraints on the Wilson coefficients can then be systematically matched to explicit UV-complete BSM theories [76].

This chapter details the application of EFT to Higgs boson property measurements at CMS. The Higgs Effective Lagrangian (HEL) is used as the language to encode modifications to Higgs boson properties from BSM physics [47, 163]. This interpretation is applied to the most recent CMS Higgs boson combination, documented in Ref. [2], which combines the measurements of cross sections in the STXS framework from multiple decay-channels. In doing so, a more complete set of EFT operators, affecting multiple Higgs boson interactions with other particles, can be constrained.

In contrast to the κ-framework discussed in Section 6.4, the EFT approach is based on a fully consistent expansion of the SM Lagrangian. As a result, the EFT-dependence

can be extended from simple normalisation effects on inclusive production-mode cross sections, to also capture shape variations in the kinematic distributions e.g. p_T^H, m_{jj}, N_{jets} etc. In this interpretation, the parametrisation is defined at the granularity of the STXS framework. This ensures that the kinematic information available in STXS measurements is used to better-constrain BSM physics. Additionally, the κ-framework defines coupling modifiers at LO. The EFT approach on the other hand is systematically improvable by computing higher-order contributions to the EFT predictions [164]. The interpretation shown in this chapter only considers EFT effects at LO, however the future transition to higher-orders is briefly discussed in Section 7.7.

The structure of this chapter is as follows: firstly the HEL model is described in Section 7.2 and the choice of operators considered in the fit is motivated. Following this, the CMS Higgs combination [2] is detailed in Section 7.3, including the full set of input analyses and the extension to the statistical inference techniques introduced in Chapter 5. Section 7.4 discusses the EFT parametrisation of the signal yield, where the cross sections and branching fractions are expressed as functions of the Wilson coefficients, $\mu^{i,f}(\vec{w})$. These functions are then used to fit the Wilson coefficients to Higgs boson measurements, and extract their respective confidence intervals. This is performed using both a simplified re-interpretation procedure and using the full likelihood fit, described in Sections 7.5 and 7.6, respectively. Finally, future improvements to EFT measurements at CMS are discussed in Section 7.7.

7.2 Higgs Effective Lagrangian

The Higgs Effective Lagrangian (HEL) model [163] is a SMEFT, which corresponds to a partial implementation of the complete SILH basis [47, 78], encompassing all operators at dimension-6 related to the Higgs sector. The perturbative expansion is defined in terms of the CP-even complex scalar Higgs field, H, and therefore assumes that the scalar sector follows the SU(2)_L doublet nature in the SM. The model is constructed by extending the SM Lagrangian with 39 flavour independent dimension-6 operators, $\mathcal{O}_p^{(6)}$,

$$\mathcal{L}_{\text{HEL}} = \mathcal{L}_{\text{SM}} + \sum_p \frac{w_p}{\Lambda^2}\mathcal{O}_p^{(6)}, \tag{7.1}$$

such that new BSM dynamics in the Higgs sector would manifest itself as deviations from zero in the HEL Wilson coefficients, w_p. Operators representing four-fermion interactions are not included.

Currently, there is insufficient data to constrain all 39 directions of the parameter space, $\vec{w}$. As a result, a subset of operators, $\{\mathcal{O}\}$ most relevant to the available Higgs

boson measurements at CMS is considered. This choice ultimately introduces a model-dependence into the interpretation, assuming the contribution from other operators is zero:

$$w_p = 0 \; \forall \; \mathcal{O}_p \notin \{\mathcal{O}\}. \tag{7.2}$$

The subset of operators considered in this interpretation is motivated below.

7.2.1 Operator selection

Non-zero contributions are considered in a total of eight dimension-6 operators,

$$\{\mathcal{O}\} \equiv \{\mathcal{O}_G, \mathcal{O}_A, \mathcal{O}_u, \mathcal{O}_d, \mathcal{O}_\ell, \mathcal{O}_{HW}, \mathcal{O}_{WW}, \mathcal{O}_B\}, \tag{7.3}$$

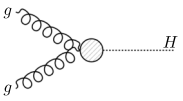
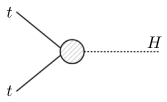
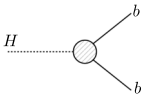
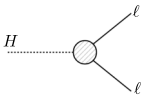
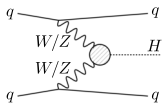
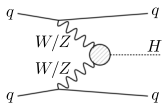
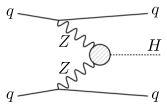
where the explicit form of these operators in terms of the SM fields and the relevant Higgs boson interaction vertices are listed in Table 7.1. In this analysis, the nominal Wilson coefficients, $\vec{w}$, are redefined as dimensionless *HEL parameters*, $\vec{c}$, which absorb the factor of Λ^{-2}. The definitions of the HEL parameters for each operator are also provided in Table 7.1. The operators $\mathcal{O}_G$ and $\mathcal{O}_A$ correspond to the effective Hgg and H$\gamma\gamma$ vertices, respectively. Since the interpretation presented here is a LO implementation of the HEL, the ggH and H $\rightarrow \gamma\gamma$ loops are not resolved into their SM structures and therefore only depend on $\mathcal{O}_G$ and $\mathcal{O}_A$.

This set of operators is chosen since the HEL parameters account for the leading CP-even terms in the scaling functions for the measured cross sections and branching fractions, and are not tightly constrained by existing measurements. CP-odd parameters are neglected as they do not enter the parametrisation at leading order in $1/\Lambda^2$, and since there is no splitting in the STXS framework that is sensitive to the Higgs boson CP (e.g. a splitting in $\Delta\phi_{jj}$) the dependence is completely degenerate with the corresponding CP-even terms at $1/\Lambda^4$.

The parameters, c_{WW} and c_B are fitted together in the combination, $c_{WW} - c_B$, since the orthogonal combination ($S = c_{WW} + c_B$) is strongly constrained at zero by electroweak precision data [165]. Finally, the operator, $\mathcal{O}_{HB}$, is neglected. Despite having a sizeable impact on the measured quantities, the effects of $\mathcal{O}_{HB}$ are degenerate with $\mathcal{O}_{HW}$, without including additional differential measurements of the VH production mode cross section or measurements of the H $\rightarrow$ Zγ decay channel. In conclusion, seven parameters of interest are defined:

$$c_G, c_A, c_u, c_d, c_\ell, c_{HW}, c_{WW} - c_B. \tag{7.4}$$

Table 7.1: The dimension-6 operator subset, $\{\mathcal{O}\}$, considered in the HEL interpretation. The definition of each operator is provided in terms of the SM field tensors. In addition, the corresponding HEL parameter is defined in terms of the nominal EFT Wilson coefficients. The final two columns show the affected Higgs boson interaction vertices and an example Feynman diagram of the EFT interaction.

Operator	Definition	HEL Parameter	Relevant vertices	Example diagrams
$\mathcal{O}_G$	$\lvert H\rvert^2 G^a_{\mu\nu} G^{a,\mu\nu}$	$c_G = \frac{m_W^2}{g_s^2}\frac{w_G}{\Lambda^2}$	Hgg	
$\mathcal{O}_A$	$\lvert H\rvert^2 B_{\mu\nu} B^{\mu\nu}$	$c_A = \frac{m_W^2}{g'^2}\frac{w_A}{\Lambda^2}$	Hγγ, HZZ	
$\mathcal{O}_u$	$\lambda_u \lvert H\rvert^2 \bar{Q}_L H^\dagger u_R + \text{h.c.}$	$c_u = -v^2\frac{w_u}{\Lambda^2}$	Htt	
$\mathcal{O}_d$	$\lambda_d \lvert H\rvert^2 \bar{Q}_L H^\dagger d_R + \text{h.c.}$	$c_d = -v^2\frac{w_d}{\Lambda^2}$	Hbb	
$\mathcal{O}_\ell$	$\lambda_\ell \lvert H\rvert^2 \bar{L}_L H^\dagger \ell_R + \text{h.c.}$	$c_\ell = -v^2\frac{w_\ell}{\Lambda^2}$	Hττ	
$\mathcal{O}_{HW}$	$i(D^\mu H)^\dagger \sigma^a (D^\nu H) W^a_{\mu\nu}$	$c_{HW} = \frac{m_W^2}{2g}\frac{w_{HW}}{\Lambda^2}$	HWW, HZZ	
$\mathcal{O}_{WW}$	$i(H^\dagger \sigma^a D^\mu H) D^\nu W^a_{\mu\nu}$	$c_{WW} = \frac{m_W^2}{g}\frac{w_{WW}}{\Lambda^2}$	HWW, HZZ	
$\mathcal{O}_B$	$i(H^\dagger D^\mu H)\partial^\nu B_{\mu\nu}$	$c_{WW} = \frac{2m_W^2}{g'}\frac{w_B}{\Lambda^2}$	HZZ	

7.3 CMS Higgs combination

The HEL interpretation is applied to the latest Higgs boson combination performed by the CMS experiment, documented in Ref. [2]. The combination includes analyses targeting all the major Higgs boson decay channels, with integrated luminosities ranging from 35.9 fb^{-1} to 137 fb^{-1}, depending on the analysis. By targeting different final states, most analyses are orthogonal in terms of the selected events. For similar final states, the number of events common to both analyses has been checked and found to be negligible.

The inclusion of different decay channels ensures sensitivity to a larger subset of operators (see Table 7.1). Each input analysis measures cross sections in the STXS framework, however, these measurements are performed using different STXS binning schemes with varying granularity. Hence the EFT parametrisation of the signal yield is defined at the granularity of all binning schemes which enter the combination: stages 0, 1.0 and 1.1 (see Appendix A). For stage 1.0 and above, the bins are split according to the event kinematics (e.g. p_T^H, N_{jets} etc). As a result, the kinematic information available in these measurements helps to further constrain BSM effects beyond simple inclusive effects, and allows more directions in the parameter space to be probed simultaneously.

The full list of input analyses is provided in Table 7.2. This combination was performed before the $\text{H} \to \gamma\gamma$ analysis described in Chapters 4–6, and so the $\text{H} \to \gamma\gamma$ inputs are taken from previous analyses. For each analysis, the targeted production-modes and final-states are listed, as well as the STXS stage of the measurement and the corresponding integrated luminosity. More detailed information on each analysis can be found in the listed references.

7.3.1 Statistical procedure

The statistical inference procedure used in the combination extends upon the procedure described in Sections 5.1.1 and 5.1.2. A likelihood function[1] is constructed for each analysis category or *region*, k, now defined for a generic final state f,

$$L_k(\text{data} \,|\, \vec{\alpha}, \vec{\theta_s}, \vec{\theta_b}) = \mathcal{P}_k(\text{data} \,|\, \vec{\alpha}, \vec{\theta_s}, \vec{\theta_b}), \tag{7.5}$$

where the likelihood-term, $\mathcal{P}_k$, takes the following form for binned analysis regions,

$$\mathcal{P}_k^{\text{binned}} = \prod_X^{N_{\text{bins}}} \text{Poisson}\Big(N_{k,X}^{\text{data}} \,\Big|\, \Big[\sum_{i,f} S_{k,X}^{i,f}(\vec{\alpha}, \vec{\theta_s})\Big] + B_{k,X}(\vec{\theta_b})\Big). \tag{7.6}$$

Here, the index X runs over bins of some observable(s) e.g. $m_{\gamma\gamma}$ for the $\text{H} \to \gamma\gamma$

[1] The dependence of the likelihood on the Higgs boson mass, m_{H}, has been dropped from the notation. For the form of the Poisson terms, please refer to equation 5.1.

Table 7.2: Input analyses to the CMS Higgs boson combination, documented in Ref. [2]. The integrated luminosity used in each analysis is listed, as well as the targeted final states and production modes. No input analysis explicitly targets single-top associated production (tH). Also listed are the STXS binning schemes in which the measurements are performed. Despite entering the combination, the ggH H $\to$ bb boosted analysis is left out of the EFT interpretation as the LO approximation of the EFT breaks down at very-high p_T^H. Also, the H $\to$ $\mu\mu$ measurements are not included as no branching-fraction scaling function was available for the H $\to$ $\mu\mu$ decay channel in Ref. [166].

Decay channel	Final states	Targeted production modes				STXS stage	$\mathcal{L}$ [fb^{-1}]	Ref.
		ggH	qqH	VH lep	ttH			
H $\to$ $\gamma\gamma$	$\gamma\gamma$	X	X	-	-	1.0	77.4	[128]
		-	-	-	X	0	77.4	[66, 167]
H $\to$ ZZ $\to$ 4ℓ	4μ, 2e2μ, 4e	X	X	X	X	1.1	137	[168]
H $\to$ WW	$e\mu$, 2e, 2μ $e\mu$ + jj, 3ℓ, 4ℓ	X	X	X	-	0	35.9	[169]
H $\to$ $\tau\tau$	$e\mu$, $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$	X	X	-	-	1.0	77.4	[170]
		-	-	X	-	0	35.9	[171]
H $\to$ bb	$\ell\ell$ + bb, $\ell\nu$ + bb, $\nu\nu$ + bb	-	-	X	-	0	77.4	[34, 172]
	bb	-	-	-	X	0	77.4	[173]
		X	-	-	-	ggH high p_T^H	35.9	[174]
ttH($\to$ leptons)	2ℓss, 3ℓ, 4ℓ 1ℓ + 2τ_h, 2ℓss + 1τ_h, 3ℓ + 1τ_h	-	-	-	X	0	77.4	[175, 176]
H $\to$ $\mu\mu$	$\mu\mu$	X	X	-	-	0	35.9	[177]

analyses. The quantity $S_{k,X}^{i,f}$ corresponds to the signal estimate in bin X, of analysis region k, originating from STXS bin i, and decaying to final state, f. The background estimate and number of data events in the same observable bin are referred to as $B_{k,X}$ and $N_{k,X}^{\text{data}}$, respectively. For unbinned analysis regions, the likelihood-term is defined as,

$$\mathcal{P}_k^{\text{unbinned}} = \frac{1}{z}\prod_j^z \text{Poisson}\Big(1 \,\Big|\, \Big[\sum_{i,f} S_k^{i,f}(\vec{\alpha},\vec{\theta}_s)\cdot\rho_{k,\text{sig}}^{i,f}(x_j|\vec{\alpha},\vec{\theta}_s)\Big] + B_k(\vec{\theta}_b)\cdot\rho_{k,\text{bkg}}(x_j|\vec{\theta}_b)\Big), \tag{7.7}$$

for z events in data landing in region k, where each event is labelled by the index j. The terms $\rho_{k,\text{sig}}^{i,f}(x)$ and $\rho_{k,\text{bkg}}(x)$ are the probability density functions of some observable(s) x, for signal and background, respectively. The total signal and background yield estimates in region k are expressed by $S_k^{i,f}$ and B_k. Comparing to the binned scenario, $S_k^{i,f}$ and B_k, are equal to the sum of the $S_{k,X}^{i,f}$ and $B_{k,X}$ terms over all observable bins, X. In all

equations, it is assumed that the background estimate does not depend on the parameters of interest, $\vec{\alpha}$, which may not always be the case in a fully consistent EFT framework (see Section 7.7). Also, the modelling of the signal in terms of the observable(s) is extracted using the SM template, and is assumed to be independent of $\vec{\alpha}$: $\rho^{i,f}_{k,\mathrm{sig}}(x|\vec{\alpha},\vec{\theta}) = \rho^{i,f}_{k,\mathrm{sig}}(x|\vec{\theta})$. For example, the shape of the signal $m_{\gamma\gamma}$ peak in the $\mathrm{H} \to \gamma\gamma$ analysis is not parametrised as a function of the HEL parameters.

In the EFT interpretation, the signal-yield for STXS bin i, in final state f, landing in analysis region k is expressed as,

$$S^{i,f}_k = \mu^{i,f}(\vec{c}) \times [\sigma^i \cdot \mathcal{B}^f]_{\mathrm{SM}} \times \epsilon^{i,f}_k(\vec{c}) \times \mathcal{L}. \tag{7.8}$$

This is effectively an extension of equation 5.4, where the explicit dependence on the HEL parameters, $\vec{c}(\equiv \vec{\alpha})$ is stated. The dependence on the nuisance parameters, $\vec{\theta}$, has been dropped from the notation for simplicity. The extraction of the cross section times branching-fraction scaling functions, $\mu^{i,f}(\vec{c})$, is described in detail in Section 7.4. Interestingly, since the EFT operators can distort the event kinematics away from the SM hypothesis, the efficiency-times-acceptance values, $\epsilon^{i,f}_k(\vec{c})$ become dependent on the HEL parameters. This is especially true for measurements in the STXS framework, where the products of the Higgs boson decay are not restricted to fiducial phase space definitions. Nevertheless, in this interpretation, the so-called *acceptance effects* are ignored,

$$\epsilon^{i,f}_k(\vec{c}) = \epsilon^{i,f}_k. \tag{7.9}$$

The potential impact of fully accounting for the detector efficiencies and analysis acceptance is investigated in Section 7.7.4.

The total likelihood is now defined as the product of the analysis-region likelihood functions, taken over all analyses included in the combination

$$L(\mathrm{data}\,|\,\vec{\alpha},\vec{\theta}) = \prod_k^{N_{\mathrm{regions}}} \Big[L_k(\mathrm{data}\,|\,\vec{\alpha},\vec{\theta})\Big] \times \mathcal{C}(\vec{\theta}). \tag{7.10}$$

Here the constraint term, $\mathcal{C}$, takes the same form of that shown in Section 5.1.1, such that deviations from the expected values of the nuisance parameters are penalised according to the associated pdf. In the combination, the mass of the Higgs boson, m_{H}, is taken to be $m_{\mathrm{H}} = 125.09$ GeV. This represents the most precise measurement of m_{H} at the time the combination was performed[2], determined from the combined LHC Run 1 ATLAS

[2]The combination was performed before the $\mathrm{H} \to \gamma\gamma$ STXS analysis documented in Chapters 4–6. This explains why the m_{H} value is different between the analyses, where the $\mathrm{H} \to \gamma\gamma$ STXS analysis uses the more recent value of 125.38 GeV [49].

and CMS measurement, using the high resolution $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ decay channels [178].

Since the different input analyses can have common sources of systematic uncertainty, the corresponding nuisance parameters must be correlated across decay channels. This follows the same procedure as described in Section 5.4.1, where in addition to defining per-year correlations, a single nuisance parameter is defined in the construction of the likelihood which affects the yield estimate in multiple decay channels simultaneously.

All theoretical uncertainties arising from the renormalisation and factorisation scales used in the cross section calculations[3], the parton distribution functions, and the branching-fraction predictions are treated as fully correlated across decay channels. In addition, since other analyses use MC to estimate background contributions, it is necessary in these cases to introduce the uncertainties in the theoretical predictions of the background cross sections. These uncertainties are correlated between channels in which the same background appears. The results presented in this section are an *interpretation* of cross section measurements, therefore the theory uncertainties directly affecting $[\sigma^i \cdot \mathcal{B}^f]_{\text{SM}}$ (i.e. type $\vec{\theta}_s^{\text{th}}$) are folded into the measurement.

Most experimental uncertainties are analysis-specific (e.g. $m_{\gamma\gamma}$ signal shape uncertainties), and are therefore uncorrelated. However, there are a number of exceptions. This includes the uncertainties in the luminosity estimates, the lepton-efficiency scale-factors, the jet-energy scale and resolution, and the b tagging efficiency. For most channels, which do not use a specific treatment of the aforementioned uncertainty sources, they are treated as correlated nuisance parameters. More information regarding the experimental-uncertainty correlation scheme is provided in Ref. [2].

After the total likelihood function has been constructed, the method of extracting the final results is identical to that described in Section 5.1.2.

7.4 Signal-yield parametrization

The scaling functions, $\mu^{i,f}(\vec{c})$, shown in equation 7.8, parametrise the signal cross section times branching fractions as a function of HEL parameters, and are derived as follows. Within the HEL framework (equation 7.1) the amplitude for each Higgs boson production and decay process can be described as,

$$|\mathcal{M}_{\text{HEL}}|^2 = \left|\mathcal{M}_{\text{SM}} + \mathcal{M}_{\text{BSM}}\right|^2 = |\mathcal{M}_{\text{SM}}|^2 + 2\text{Re}\{\mathcal{M}_{\text{SM}}\mathcal{M}_{\text{BSM}}^{\dagger}\} + |\mathcal{M}_{\text{BSM}}|^2, \tag{7.11}$$

where $\mathcal{M}_{\text{SM}}$ and $\mathcal{M}_{\text{BSM}}$ are the matrix elements originating from the SM and BSM parts of

[3] This includes both the inclusive effects and the migration of events *across* STXS bin boundaries.

the Lagrangian, respectively. The total amplitude now contains an SM-BSM interference term, suppressed by a factor of Λ^{-2}, and a purely-BSM term, suppressed by a factor Λ^{-4}. In this interpretation, only one BSM vertex is considered per Feynman diagram, which means $\mathcal{M}_{\text{BSM}}$ is linear in the HEL Wilson coefficients. Substituting this linearity condition into equation 7.11, and using the fact $\sigma \propto |\mathcal{M}|^2$, we arrive at an expression for the cross section of signal process, i,

$$\sigma^i_{\text{HEL}} = \sigma^i_{\text{SM}} + \sigma^i_{\text{int}} + \sigma^i_{\text{BSM}}. \tag{7.12}$$

This results in a scaling function relative to the SM prediction which is quadratic in the HEL parameters,

$$\mu^i_{\text{prod}}(\vec{c}) = \frac{\sigma^i_{\text{HEL}}}{\sigma^i_{\text{SM}}} = 1 + \sum_p A^i_p \, c_p + \sum_{pr} B^i_{pr} \, c_p \, c_r, \tag{7.13}$$

where the index i corresponds to the STXS bin. The terms, A^i_p and B^i_{pr}, are constant prefactors which encode the impact of the HEL parameters on each STXS bin. We can ignore the $|\mathcal{M}_{\text{BSM}}|^2$ term in the expansion simply by setting the B^i_{pr} prefactors to zero. These terms, despite having an energy-scale suppression of the same order as the leading dimension-8 SM-BSM interference contributions (Λ^{-4}), are kept in this interpretation since they are the leading purely-BSM terms and they prevent the scaling functions from going negative.

Applying the same reasoning, the partial Higgs boson decay width to final state f scales relative to the SM prediction as,

$$\frac{\Gamma^f_{\text{HEL}}}{\Gamma^f_{\text{SM}}} = 1 + \sum_p A^f_p \, c_p + \sum_{pr} B^f_{pr} \, c_p \, c_r. \tag{7.14}$$

It is necessary to also consider the variation in the total Higgs boson decay width, Γ^H, such that the scaling function of the branching fraction to final state f is expressed as

$$\mu^f_{\text{decay}}(\vec{c}) = \frac{\mathcal{B}^f_{\text{HEL}}}{\mathcal{B}^f_{\text{SM}}} = \frac{\Gamma^f_{\text{HEL}}/\Gamma^f_{\text{SM}}}{\Gamma^H_{\text{HEL}}/\Gamma^H_{\text{SM}}} = \frac{1 + \sum_p A^f_p \, c_p + \sum_p \sum_r B^f_{pr} \, c_p \, c_r}{1 + \sum_p A^H_p \, c_p + \sum_p \sum_r B^H_{pr} \, c_p \, c_r}. \tag{7.15}$$

The total scaling function for signal events originating from STXS bin i and decaying to final state f, is the product of the individual cross section and branching-fraction scaling functions,

$$\mu^{i,f}(\vec{c}) = \mu^i_{\text{prod}}(\vec{c}) \cdot \mu^f_{\text{decay}}(\vec{c}). \tag{7.16}$$

This works under the narrow Higgs boson width assumption, such that the effects at production and decay have been factorised.

In summary, the Higgs boson signal scaling functions are uniquely described by the set of constant prefactors, $\{A_p^i, B_{pr}^i, A_p^f, B_{pr}^f, A_p^H, B_{pr}^H\}$. These prefactors are derived using LO MC samples with the reweighting procedure to be described in the following section. It should be stressed that the SM predictions for the cross sections and branching fractions in equation 7.8 are computed at the highest available order, however the EFT parametrisation is derived using LO MC samples. This strategy therefore assumes that the corrections to the cross sections and branching fractions from HEL operators is comparable at LO and higher orders [179]. Once defined, the scaling functions are then applied in the likelihood fit to extract the best-fit values and corresponding confidence intervals for the considered HEL parameters.

7.4.1 Derivation: Monte Carlo reweighting

The impact of the HEL operators is computed using the `HEL_UFO` model [163] in MadGraph [119], where the Higgs boson production and decay processes are generated at LO in both QCD and QED. The LO MadGraph reweighting functionality [180] is utilised, to reweight the generated events to different points in the HEL parameter space, according to,

$$W_{\vec{c}} = \frac{|\mathcal{M}_{\text{HEL}}^{\vec{c}}|^2}{|\mathcal{M}_{\text{HEL}}^{\text{nominal}}|^2} \cdot W_{\text{nominal}} \tag{7.17}$$

where $\mathcal{M}_{\text{HEL}}^{\vec{c}}$ is the matrix element at the point in parameter space, $\vec{c}$, $\mathcal{M}_{\text{HEL}}^{\text{nominal}}$ is the matrix element at the nominal point, and W_{nominal} is the corresponding event weight at that nominal point. Here, the nominal point in parameter space is chosen as the SM, i.e. $\vec{c} = (0, 0, ..., 0)$.

For each operator, $\mathcal{O}_p$, two weights are defined by setting c_p to two different values (a,$2a$), whilst all other HEL parameters are kept at 0. In doing so, simultaneous equations are constructed, as shown in equation 7.18, where the reweighted and SM values of the observable, X (σ^i for production, Γ^f for decay), can be used to infer the values of A_p and B_{pp},

$$\begin{aligned} \frac{X_{c_p=a}}{X_{\text{nominal}}} &= 1 + a \cdot A_p + a^2 \cdot B_{pp} \\ \frac{X_{c_p=2a}}{X_{\text{nominal}}} &= 1 + 2a \cdot A_p + 4a^2 \cdot B_{pp}. \end{aligned} \tag{7.18}$$

An additional weight is required to extract the cross-terms, B_{pr}, where $p \neq r$. This is

defined by setting $(c_p, c_r) = (a, a)$, and keeping all other HEL parameters at 0, such that

$$\frac{X_{(c_p,c_r)=(a,a)}}{X_{\text{nominal}}} = 1 + a \cdot (A_p + A_r) + a^2 \cdot (B_{pp} + B_{rr} + B_{pr}). \tag{7.19}$$

The value of B_{pr} can then be inferred by using the previously calculated prefactors, $\{A_p, A_r, B_{pp}, B_{rr}\}$, from equation 7.18. In total, the number of weights required to fully specify the scaling functions is

$$N_{\text{weights}} = 1 + 2N + \frac{N(N-1)}{2}, \tag{7.20}$$

where N is the number of operators. This analysis considers 8 HEL operators and therefore requires 45 weights, including the nominal SM weight. The value of a is chosen to be small (0.01) to ensure the EFT effects do not blow up in the matrix-element calculations, and therefore do not invoke a large statistical uncertainty in the calculated prefactors[4].

No kinematic cuts are applied to the final-state objects in the generated events. All events are interfaced with Pythia8 for parton showering and hadronisation [127]. A matching is performed using the `MLM` algorithm [117] to remove phase-space overlap between jets specified in the matrix element and those originating from the parton shower. The generator-option choices used in the event simulation can affect the values of the prefactors, and therefore the scaling functions. For example, the EFT effects originate solely from the matrix element (in Madgraph) and not from the parton-showering (in Pythia8). As a result, the values of the prefactors can depend on the scheme and parameters used for the jet-matching. In general, these generator options have a small effect on the final parametrisation. Nevertheless it is important to specify the options used when reporting results; a summary of the MC options used in this interpretation is provided in Appendix G.

7.4.2 Effect at production

Each Higgs boson production mode is generated separately, according to the Madgraph process definitions listed in Appendix G. The Higgs boson decay is not specified in the process definition, such that EFT effects only enter in the Higgs boson production interaction vertices. The option `NP<=1` limits the number of BSM vertices to one per Feynman diagram. After interfacing with Pythia8, the particle-level events are propagated through the Rivet program [181], using the `HiggsTemplateCrossSections` routine. This routine sequentially extracts the simulated-event constituents, forms hadronic jets with

[4] If the value of a is large, the EFT kinematic phase-space can be sufficiently different to the SM kinematic phase-space. The reduced overlap in phase-space means that fewer events can be reweighted to describe the EFT effects, which explains why large values of a induce a large statistical uncertainty in the EFT parametrisation.

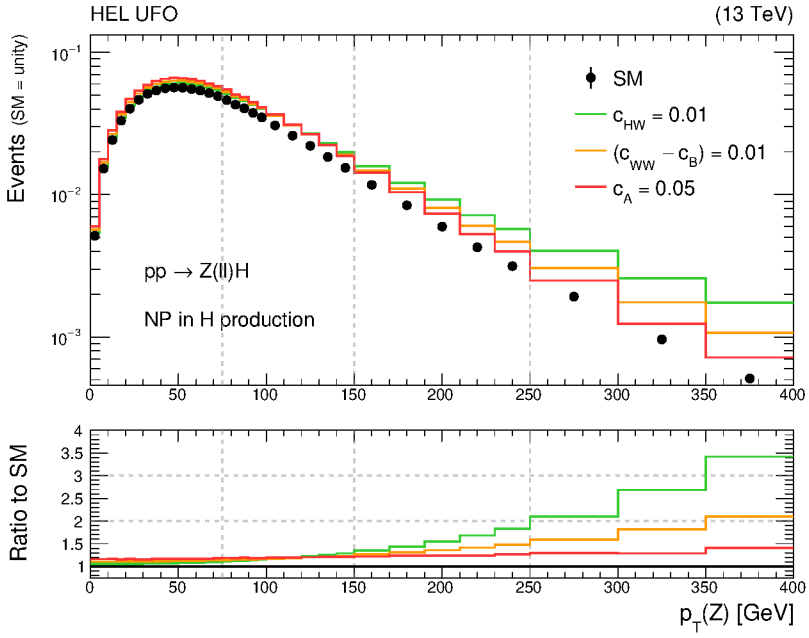


Figure 7.1: The Z boson transverse momentum, p_T^Z, spectrum for ZH leptonic events generated at LO. The black points correspond to the SM prediction, whilst the coloured lines show the distribution when various EFT contributions are introduced.

the FastJet packagage [68] using the anti-k_T algorithm with a distance parameter of 0.4 [69], calculates high-level kinematic quantities (e.g. p_T^H, p_T^V, p_T^{Hjj}, m_{jj}, N_{jets}) and, finally, *classifies* the simulated events according to their truth-level STXS bin. In all steps including the classification, the Higgs boson decay products formed by Pythia8 are neglected. The routine has been modified to output the bin-classification at each stage of the STXS framework considered in the combination: stage 0, 1.0 and 1.1.

A large number of events are generated for each production mode (10^6) to ensure each STXS bin is sufficiently populated. As a result the uncertainty in the scaling function prefactors arising from the limited MC statistics is small (typically below 1%). The set of 45 event weights are then applied to extract the SM and EFT-reweighted cross sections of each individual STXS bin, which are subsequently used to calculate the relevant prefactors: $\{A_p^i, B_{pr}^i\}$, according to the prescription described above.

Figure 7.1 shows the p_T^Z distribution for ZH leptonic events in the SM (black points) and when turning on various HEL parameters (coloured lines). The enhancement due to the c_{HW} and $c_{WW} - c_B$ parameters grows with increasing p_T^Z; improving the measurements of this high-p_T region of phase space would allow for tighter constraints on these parameters. The dashed lines in the plot indicate the boundaries in p_T^Z which define the

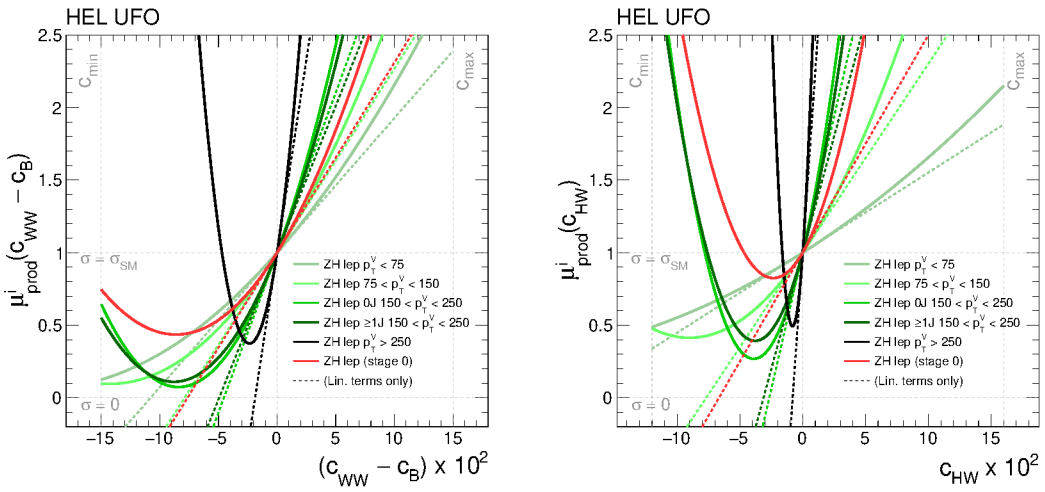


Figure 7.2: Cross-section scaling functions, $\mu^i_{\text{prod}}(\vec{c})$, for the ZH leptonic STXS stage 1.1 bins in terms of $c_{WW} - c_B$ (left) and c_{HW} (right). The red line shows the scaling behaviour for inclusive ZH leptonic production (stage 0). The dashed lines indicate the scaling functions when only the linear terms are considered ($B_{pr} = 0$).

ZH leptonic STXS stage 1.1 bins[5]. The cross-section scaling functions, $\mu^i_{\text{prod}}(\vec{c})$, for each of these bins are shown as a function of $c_{WW} - c_B$ and c_{HW} in Figure 7.2. In the plots, all other HEL parameters apart from the dependent variable are set to zero. Clearly the dependence on the HEL parameters increases as a function of p_T^Z. Going further, Figure 7.3, shows $\mu^i_{\text{prod}}(\vec{c})$ for the ZH leptonic STXS stage 0 bin, considering variations in the pair of HEL parameters: $c_{WW} - c_B$ and c_{HW}. The tilt in the distribution arises from the relevant cross-term, $B_{WW-B,HW}$.

The scaling function for the inclusive ttH STXS bin[6] is shown as a function of c_u in Figure 7.4. Interestingly, the scaling function is equal to unity for two distinct values of c_u within the allowed range of the parameter: at the SM point ($c_u = 0$) and at $c_u = -4/3$. Without additional measurements sensitive to c_u entering the likelihood, it is impossible to distinguish between these two points in parameter space. As a result the $q(c_u)$ likelihood curve will exhibit a double-minimum structure, with minima situated around $c_u = 0$ and $c_u = -4/3$. The inclusion of the tH production cross section measurement shown in Chapter 6 would help alleviate this degeneracy. A similar degeneracy is observed in the ZH leptonic scaling functions in Figure 7.2. This however, is broken by including other measurements in the combination which depend differently on c_{HW} and $c_{WW} - c_B$: the qqH and WH leptonic production modes, and the H $\to$ ZZ and H $\to$ WW decay channels.

[5]The ZH leptonic $150 < p_T^V < 250$ GeV region is further split into bins with zero jets and ≥1 jet.
[6]The ttH production mode is only split (according to p_T^H) at stage 1.2.

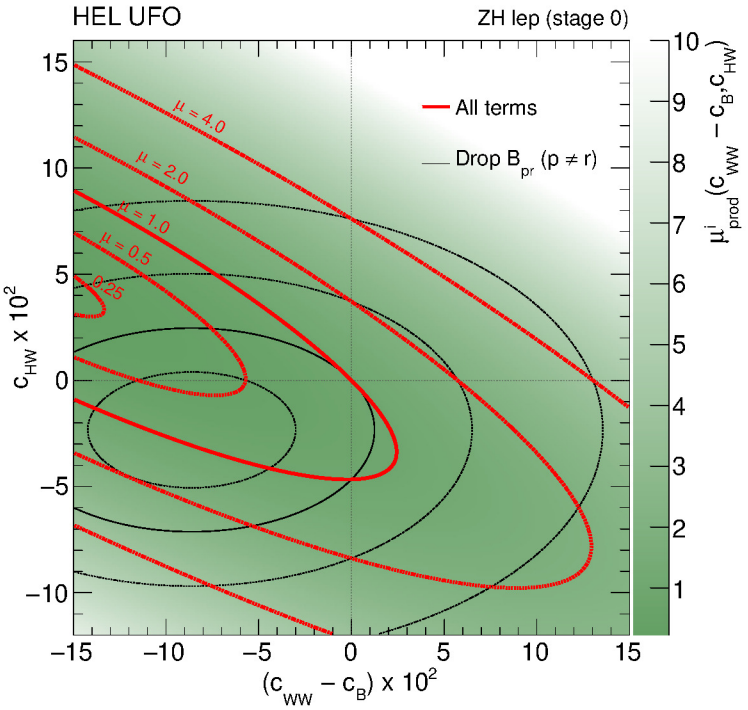


Figure 7.3: Cross-section scaling function, $\mu^{i}_{\text{prod}}(\vec{c})$, for inclusive ZH leptonic production (stage 0), considering variations in the pair of HEL parameters: $c_{WW} - c_B$ and c_{HW}. The contours indicate lines of constant μ. The black lines show the scaling function when the cross-term, $B_{WW-B,HW}$, is neglected.

7.4.3 Effect at decay

The decay-mode parametrisation is taken directly from Ref. [166]. Higgs bosons are generated at rest and are made to decay in a particular channel (see Appendix G), where the EFT effects enter the Higgs boson decay vertices. The STXS framework does not include a fiducial region definition for the Higgs boson decay products, and therefore the EFT effects at decay are defined for the inclusive Higgs boson decay phase-space.

The partial-width scaling functions are derived using a similar approach to the cross-section scaling functions. This amounts to extracting the SM prediction of the partial width at LO and the partial width with EFT effects turned on, and following the derivation procedure outlined in Section 7.4.1. The total Higgs boson decay-width parametrisation is then inferred from the sum of all partial-width scaling functions, where the decay channels are weighted in the sum according to their total contribution to Γ^{H}. Figure 7.5 summarises the branching-fraction scaling functions, $\mu^{f}_{\text{decay}}(\vec{c})$, for the decay channels considered in the CMS Higgs boson combination. As the H $\rightarrow$ bb decay channel has the largest branching fraction, the greatest impact on the total decay-width is from the c_d

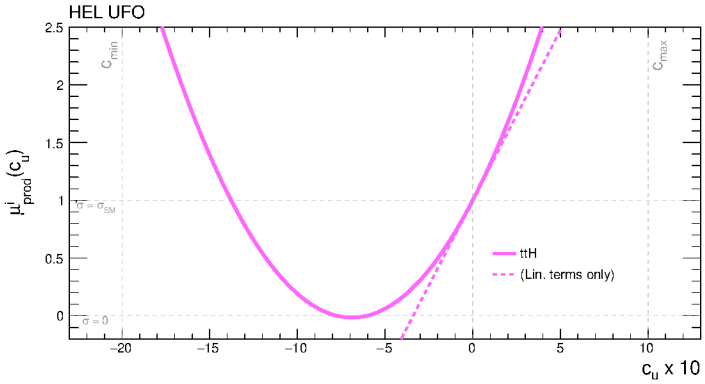


Figure 7.4: Cross-section scaling function, $\mu^i_{\text{prod}}(\vec{c})$, for ttH production in terms of c_u. The dashed line indicates the scaling function when only the linear term is considered ($B_{uu} = 0$).

parameter. Also, in accordance with ttH production, the H $\to$ bb and H $\to \tau\tau$ scaling functions have two points in c_d and c_ℓ respectively, which correspond to the SM prediction (unity). Again, since no additional measurements are included in the combination which are sensitive to these parameters, the corresponding $q(c_d)$ and $q(c_\ell)$ likelihood curves will exhibit a double-minimum structure.

7.4.4 Summary

The impacts of each HEL parameter on the stage 0, 1.0 and 1.1 cross sections, and the branching fractions are displayed in Figure 7.6, relative to the SM predictions. Each parameter is varied to its expected upper and lower 1σ confidence level values to indicate the size of variations to which the combination measurements are sensitive to[7].

The full HEL parametrisation, $\mu^i_{\text{prod}}(\vec{c})$ and $\mu^f_{\text{decay}}(\vec{c})$, is presented in Tables H.1–H.5 of Appendix H. For each signal process, the total scaling function is defined as the product of the corresponding cross-section and branching-fraction scaling functions. Figure 7.7 shows the example for the qqH BSM STXS stage 1.1 bin in the H $\to$ ZZ $\to 4\ell$ decay channel, plotted as a function of $(c_{WW} - c_B)$-vs-c_{HW}. These total scaling functions, $\mu^{i,f}(\vec{c})$, are then applied to the signal-yield estimates when constructing the likelihood, as shown in equation 7.8, and constraints on the HEL parameters, $\vec{c}$, are extracted using the techniques described in Section 5.1.2.

[7] These are taken from the fit in which only variations in a single HEL parameter are considered (see Section 7.6). For parameters which demonstrate a double-minimum in the likelihood (c_u,c_d,c_ℓ), the confidence levels are taken from the minimum around 0.

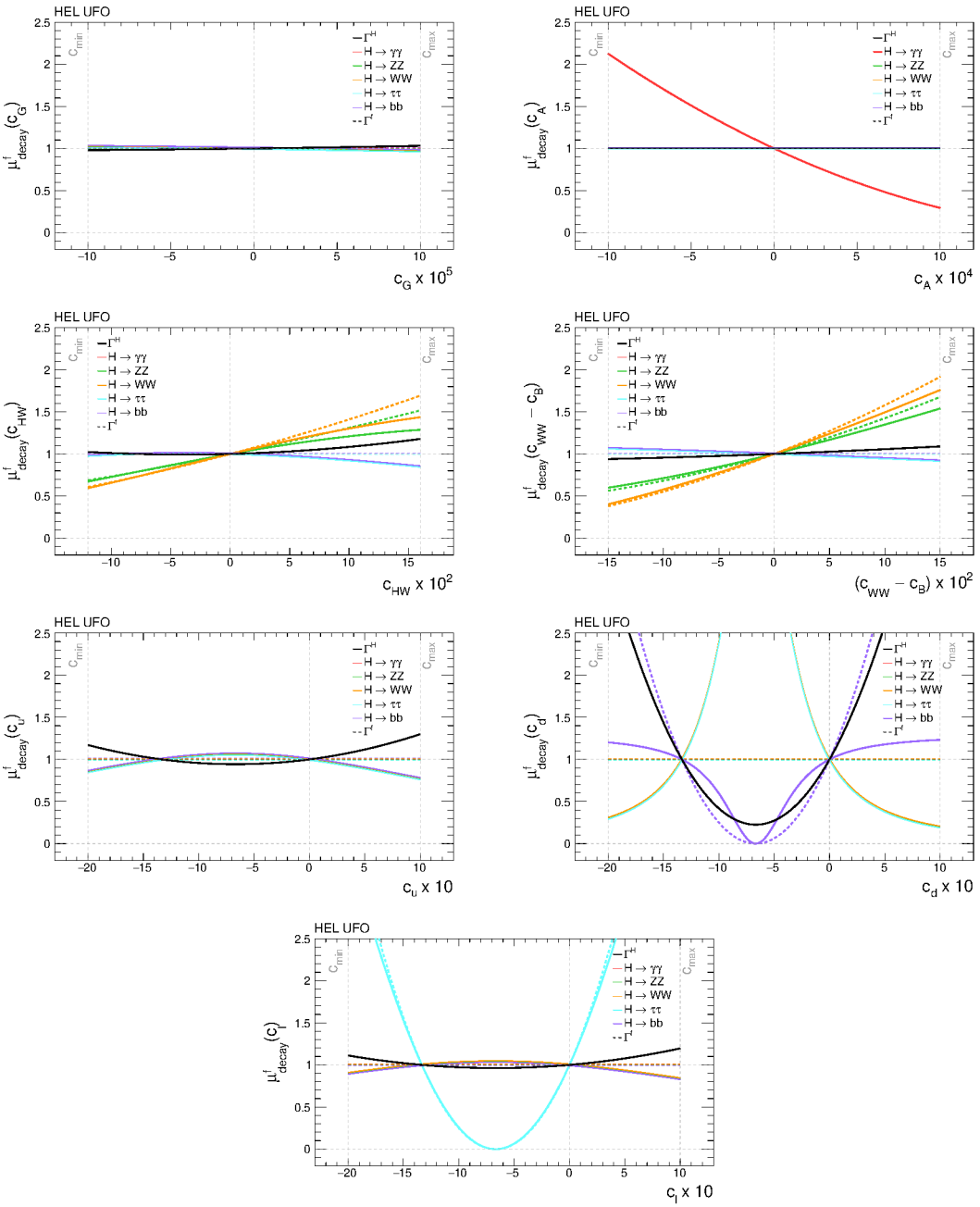


Figure 7.5: Branching-fraction scaling functions, $\mu^f_{\text{decay}}(\vec{c})$ (solid coloured lines) for each decay channel considered in the CMS Higgs boson combination. The dashed lines represent the corresponding partial-width scaling functions, and the solid black lines shows the Higgs boson total-width scaling functions. It is the ratio of these two which provide the branching-fraction scaling functions.

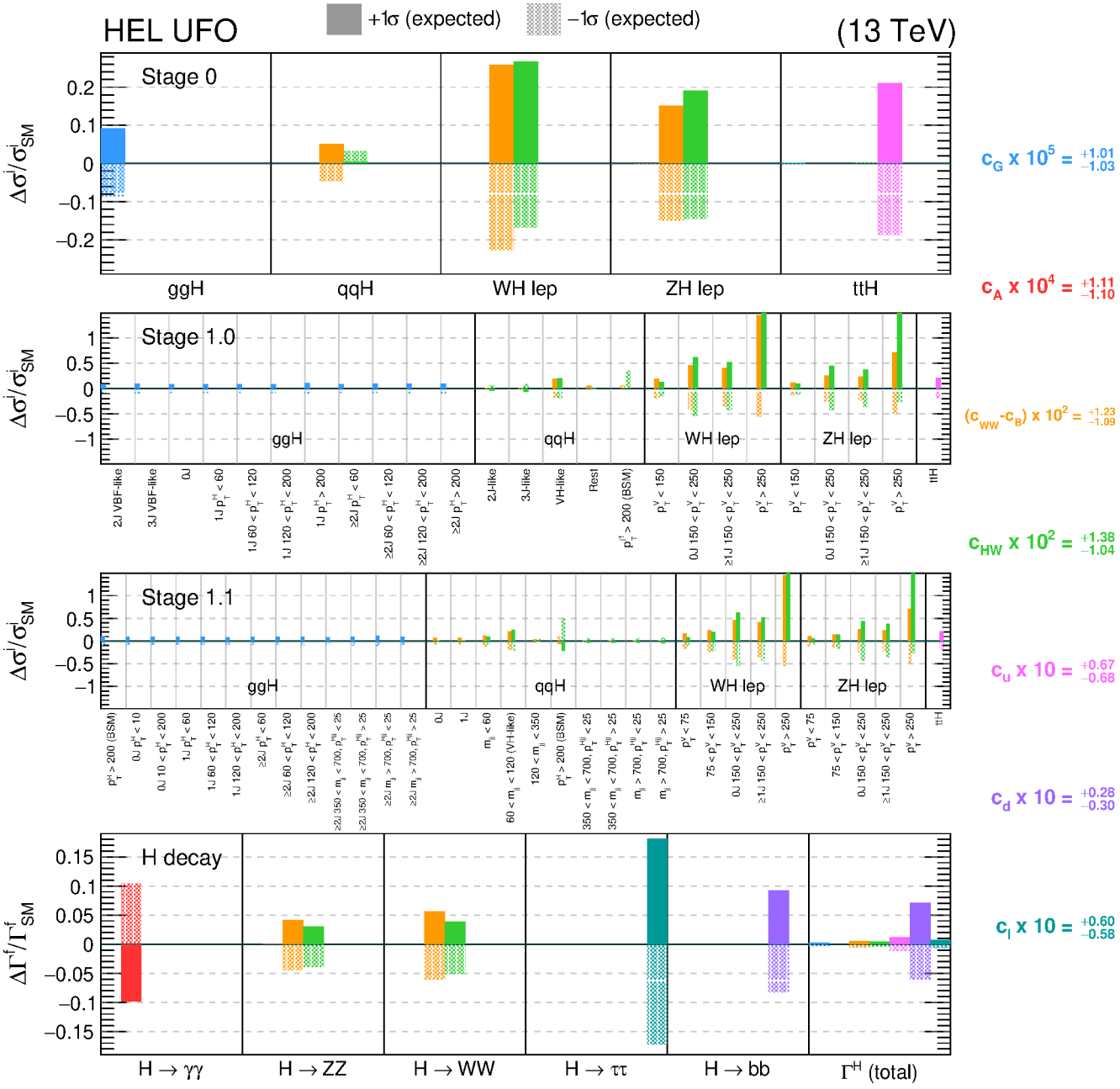


Figure 7.6: The impact of each HEL parameter on all STXS stage 0, stage 1.0 and stage 1.1 cross sections, and the relevant partial and total Higgs boson decay widths. Each parameter is varied to its corresponding expected upper and lower 1σ confidence level values to indicate the size of variations to which the CMS Higgs boson combination is sensitive.

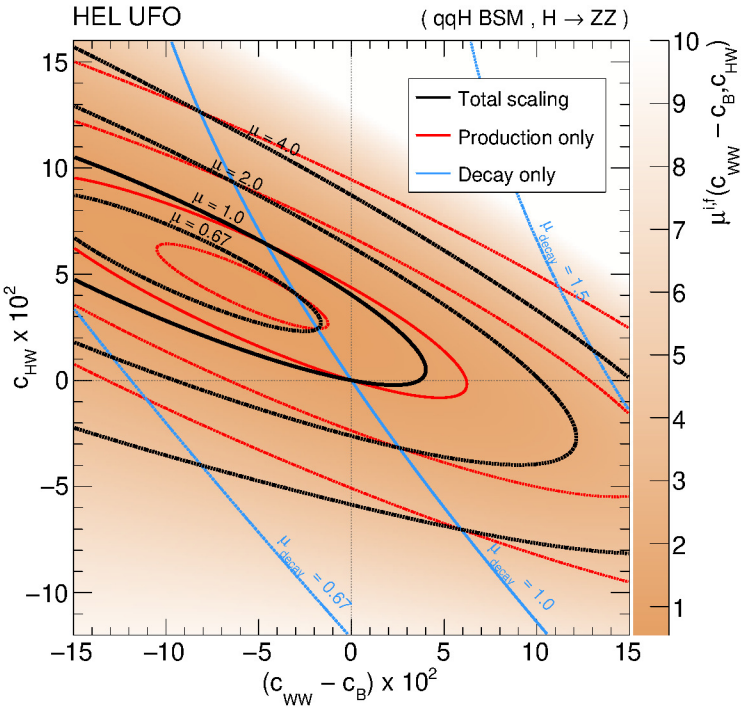


Figure 7.7: The total scaling function, $\mu^{i,f}(\vec{c})$, for qqH BSM events in the H $\rightarrow$ ZZ $\rightarrow$ 4ℓ decay channel, considering variations in the pair of HEL parameters: $c_{WW} - c_B$ and c_{HW}. The contours indicate lines of constant μ. The effect is decomposed into the cross-section scaling function, $\mu^{i}_{\text{prod}}(\vec{c})$ (red), and branching-fraction scaling function $\mu^{f}_{\text{decay}}(\vec{c})$ (blue).

7.4.5 Validation of the scaling functions

The cross-section scaling functions have been calculated previously for the STXS stage 1.0 bin definitions in Ref. [166]. This enables a direct comparison of the prefactors, A_p and B_{pr}, calculated there (HXSWG parametrisation) to those derived for this thesis (CMS parametrisation). For most stage 1.0 bins the two parametrisations are in excellent agreement, with the calculated prefactors exhibiting differences below the 10% level. This is not the case for the linear terms for a number of bins in the qqH binning scheme, particularly the qqH bin with p_T of the leading jet greater than 200 GeV, the qqH VH-like bin, and the qqH rest bin. For such bins, the A_{HW} and A_{WW-B} values were found to differ by at least 50%, with some even showing opposite signs.

An extensive series of tests were performed to try and reconcile these differences. These include changing the nominal point in parameter space away from the SM, altering the MG5_aMC@NLO and Pythia8 options, and generating separate samples for each point in parameter space as opposed to using the reweighting procedure described in

Section 7.4.1. In all cases, the calculated prefactors were consistent within statistical fluctuations, demonstrating the CMS parametrisation is stable to variations in the extraction method.

Nevertheless, the effects of the parametrisation differences have been investigated. This was done by using the HXSWG parametrisation for the STXS stage 1.0 bins when performing the results extraction. It was found that the constraints on $c_{WW} - c_B$ and c_{HW} improve by roughly 20% and 25% respectively, compared to the results when using the CMS parametrisation. The constraints on all other HEL parameters were in excellent agreement.

7.5 Simplified likelihood re-interpretation procedure

Before showing the results extracted using the full likelihood, it is useful to introduce a simplified approach to *re-interpreting* cross section measurements. This has been used as a tool to investigate particular properties of the HEL interpretation, such as the most important operators for the combination input analyses, and to gain an estimate of their respective sensitivity. A χ^2 function is constructed using measurements from different input analyses,

$$\chi^2(\vec{c}) = \sum_a (\mathbf{X}_a - \boldsymbol{\mu})^T \mathbf{V}_a^{-1} (\mathbf{X}_a - \boldsymbol{\mu}), \tag{7.21}$$

with the following inputs:

- a: index to label the input analysis.
- $\mathbf{X}_a$: a vector of cross section times branching fraction measurements from analysis a. The elements of the vector are the best-fit values of $[\sigma^i \cdot \mathcal{B}^f]_{\text{obs}}$, relative to the SM prediction: $x_a^{i,f} = [\sigma^i \cdot \mathcal{B}^f]_{\text{obs}}/[\sigma^i \cdot \mathcal{B}^f]_{\text{SM}}$. For example, to use the $\text{H} \to \gamma\gamma$ minimal-merging results shown in Chapter 6, $\mathbf{X}_a$ would be a vector of the best-fit values shown in the final column of Table 6.2.
- $\boldsymbol{\mu}$: a vector of EFT scaling functions, $\mu^{i,f}(\vec{c})$, where the elements match the corresponding measurement in the $\mathbf{X}_a$ vector: $x_a^{i,f}$. In this manner, the element-wise subtraction is minimised for the HEL parameter point in which, $\mu^{i,f}(\vec{c}) = [\sigma^i \cdot \mathcal{B}^f]_{\text{obs}}/[\sigma^i \cdot \mathcal{B}^f]_{\text{SM}}$.
- $\mathbf{V}_a$: covariance matrix for the cross section times branching fraction measurements from analysis, a, with elements: $V_a^{(i,f),(j,g)} = \rho_{(i,f),(j,g)} \Sigma_{i,f} \Sigma_{j,g}$. The terms $\Sigma_{i,f}$ and $\Sigma_{j,g}$ are the *symmetrised* 68% confidence intervals in the measurements $x_a^{i,f}$ and

$x_a^{j,g}$, respectively. The term $\rho_{(i,f),(j,g)}$ refers to the correlation coefficient between $x_a^{i,f}$ and $x_a^{j,g}$. Note, if the input analysis corresponds to measurements in a single decay channel then $f = g$. To use the H $\rightarrow \gamma\gamma$ minimal-merging example, the $\Sigma_{i,f}$ and $\Sigma_{j,g}$ would be symmetrised values of the 68% confidence intervals shown in the final column of Table 6.2, and $\rho_{(i,f),(j,g)}$ would be taken from the correlation matrix shown in Figure 6.15.

The χ^2 value is minimised with respect to the HEL parameters, $\vec{c}$. This is done numerically using the `scipy.optimize` package [182]. The point in HEL parameter space which minimises χ^2 corresponds to the best-fit point, whilst the points which incur a change $\Delta\chi^2 = 1$ and 4, correspond to the $\pm 1\sigma$ ($\sim$68%) and $\pm 2\sigma$ ($\sim$95%) confidence intervals. This minimisation is performed for two scenarios. The first scenario, only considers variations in a single HEL parameter, whilst the other parameters are fixed to 0. The second scenario allows variations in all parameters simultaneously, performed by scanning over one parameter and profiling the other parameters in the minimisation. From a physical perspective, the first approach corresponds to considering BSM effects in a single EFT operator, whilst the second approach is more general and allows BSM effects in a number of operators simultaneously.

In summary, the $\Delta\chi^2(\vec{c})$ surface is a simplified approximation of the $q(\vec{c})$ surface, derived from the full-combination likelihood. In this approximation, the likelihoods of the input analyses are assumed to be Gaussian in nature, such that the uncertainties in the measurements are symmetric. In addition, it is assumed that the results of different analyses are completely independent i.e. the correlation coefficients between them are 0. This assumption completely ignores the common sources of systematic uncertainty between input analyses.

7.5.1 Re-interpreting CMS STXS measurements

The simplified re-interpretation procedure is applied to the full set of analyses listed in Table 7.2. For both fitting scenarios, the χ^2 minimisation is performed when using the quadratic scaling functions, and when considering only the linear terms in the parametrisation ($B_{pr} = 0$). A comparison between the two $\Delta\chi^2$ curves demonstrates the impact of including the purely-BSM terms, and therefore indicates the sensitivity of the measurement to terms suppressed by a factor Λ^{-4}.

Figures 7.8 and 7.9 show the $\Delta\chi^2(c_p)$ functions for each considered HEL parameter. The black and purple lines represent the fits in which the other parameters are profiled and fixed to zero, respectively. All results are in agreement with the SM ($c_p = 0$) within the 2σ confidence intervals. In the plots, solid lines are used to show the results from using the

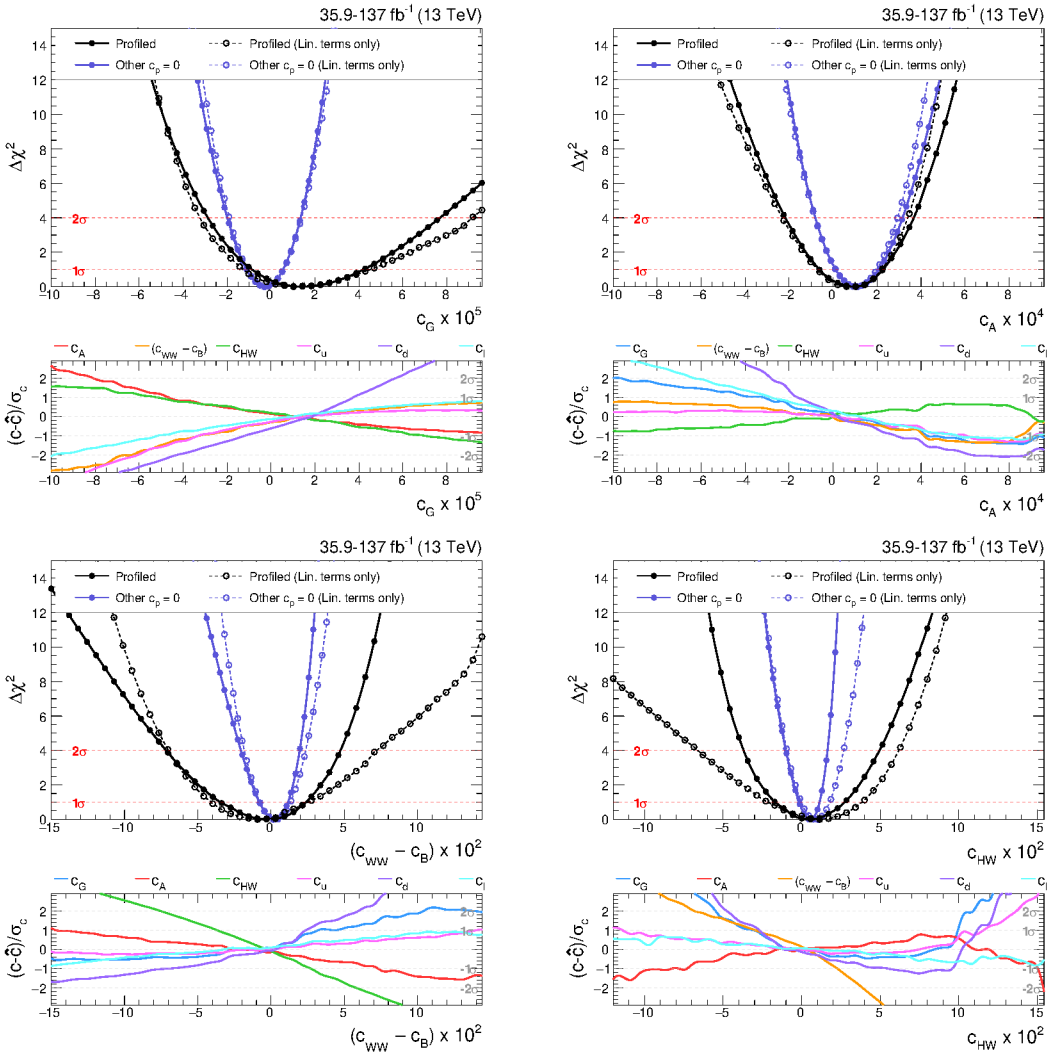


Figure 7.8: The $\Delta\chi^2(c_p)$ curves for the HEL parameters: c_G, c_A, $c_{WW} - c_B$ and c_{HW}. The black and purple lines in the top panels correspond to the fits in which the other parameters are profiled and fixed to the SM, respectively. The dashed lines indicate the fits when only linear terms are considered in the parametrisation. The points in each curve show the values of c_p where the minimisation is performed; the lines are extracted by interpolating between these points. The horizontal red lines at $\Delta\chi^2(c_p) = 1$ and 4 indicate the 1σ and 2σ confidence intervals in c_p, respectively. The bottom panels show the pull of the profiled parameters with respect to the parameter of interest.

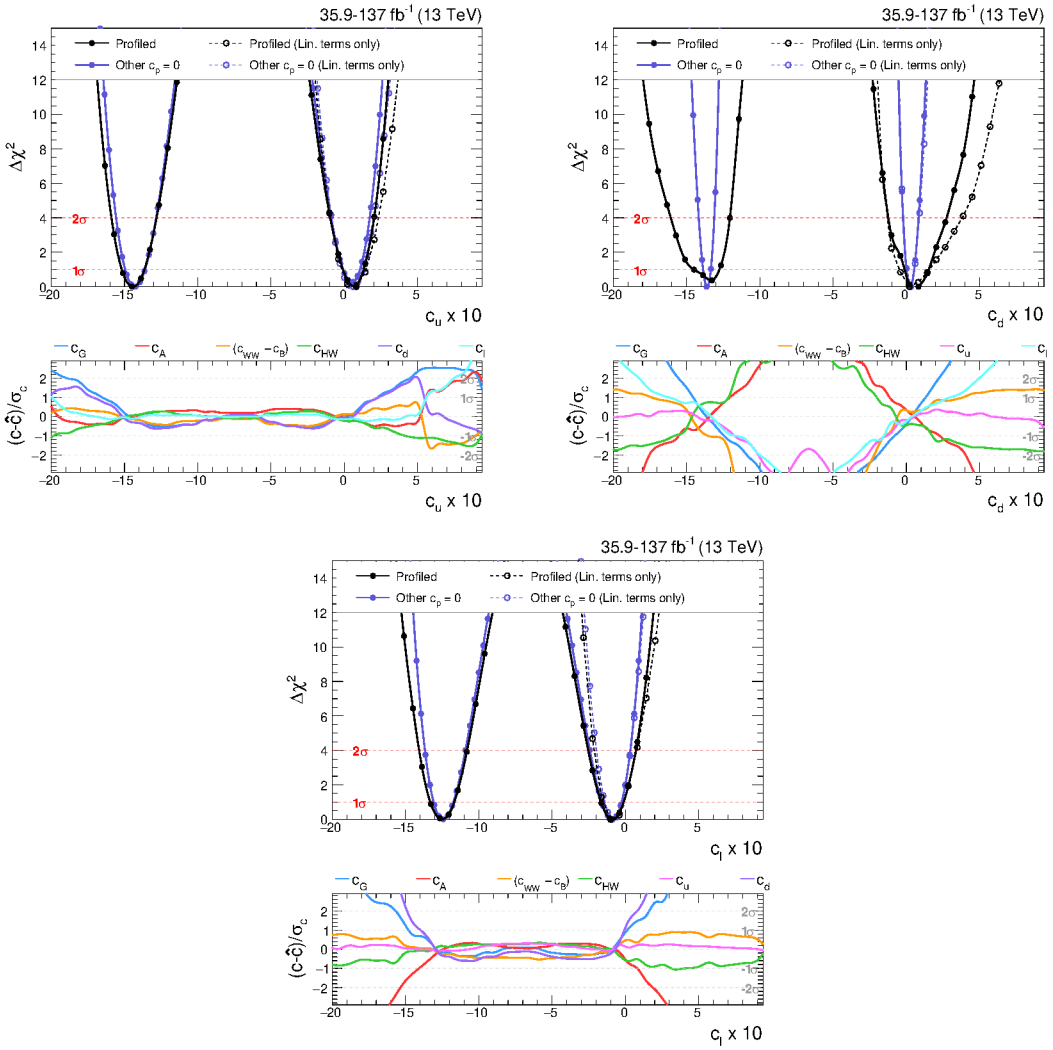


Figure 7.9: The $\Delta\chi^2(c_p)$ curves for the HEL parameters: c_u, c_d and c_ℓ. The black and purple lines in the top panels correspond to the fits in which the other parameters are profiled and fixed to the SM, respectively. The dashed lines indicate the fits when only linear terms are considered in the parametrisation. The points in each curve show the values of c_p where the minimisation is performed; the lines are extracted by interpolating between these points. The horizontal red lines at $\Delta\chi^2(c_p) = 1$ and 4 indicate the 1σ and 2σ confidence intervals in c_p, respectively. The bottom panels show the pull of the profiled parameters with respect to the parameter of interest.

full quadratic scaling functions, whilst the dashed lines represent the results from using the linear terms only. The inclusion of the quadratic terms is observed to have a particularly large impact on the c_{HW} and $c_{WW} - c_B$ constraints. This effect can be inferred from the difference between the quadratic and linear scaling functions in Figure 7.2, where for example, the ZH leptonic stage 0 bin (red line) has a steeper dependence for $c_{HW} > 0$ when the quadratic terms are included. Consequently, the upper constraint on c_{HW} in Figure 7.8 is tighter compared to the linear-terms-only case. In addition, the c_u, c_d and c_ℓ $\Delta\chi^2$ curves exhibit a double-minimum structure when the quadratic terms are included. All in all, these stark differences demonstrate the importance of including the Λ^{-4}-suppressed terms in the parametrisation[8].

The bottom panels of each plot show the pulls of the profiled HEL parameters as a function of the parameter of interest. These are taken from the fully quadratic fit (solid black line). The pulls can be used as an indication of the correlation between HEL parameters. For example, the c_{HW} and $c_{WW} - c_B$ exhibit a strong anti-correlation due to their similar effects on the HWW and HZZ interaction vertices. The largest pulls are observed for the c_d parameter. This is due to the sizeable impact on the total Higgs boson decay width from c_d, which in turn incurs a larger variation in the other profiled parameters.

Additionally, many of the profiled $\Delta\chi^2(c_p)$ functions demonstrate an asymmetry for positive and negative values of the HEL parameter, even when considering the linear-terms only. There are two effects here which can cause this behaviour. Firstly, the total scaling functions are the product of the cross-section and branching-fraction scaling functions (equation 7.16). Therefore, even when considering only the linear terms ($A_p\, c_p$) in $\mu^i_{\text{prod}}(\vec{c})$ and $\mu^f_{\text{decay}}(\vec{c})$, the product of the two can introduce non-linear effects. Secondly, it can be seen that the fits in which the other parameters are fixed to zero (purple lines) are generally more symmetric about the best-fit point. An additional factor causing the asymmetry is therefore the correlation between HEL parameters. This is confirmed by the flattening of the linear-only profiled $\Delta\chi^2(c_p)$ functions (black dashed lines) for opposite sign values of the strongly anti-correlated c_{HW} and $c_{WW} - c_B$ parameters. Ultimately, as the re-interpretation shown here includes many independent measurements which depend differently on the HEL parameters, it is difficult to pinpoint the exact origins of the asymmetries. Such a study would require fitting only a subset of the input measurements to understand which production modes and decay channels create the asymmetry.

[8]This finding questions the validity of neglecting dimension-8 operators, where the leading interference terms enter at $\mathcal{O}(\Lambda^{-4})$. This choice of only including dimension-6 operators effectively introduces some model-dependence into the interpretation. The definition of a fully-consistent SMEFT up to dimension-8 has recently been achieved [73]. However, going to dimension-8 introduces an additional 44,807 operators and it is not yet known how we will approach constraining such a high number of independent parameters.

In summary, the simplified re-interpretation procedure is a useful tool for extracting the approximate constraints on the parameters of interest, and testing the properties of the parametrisation, such as the impact of including the B_{pr} terms. In addition, the pulls of the profiled parameters can be used to estimate their respective correlations. It should be stressed that this method is not particular to EFT. It can be used to re-interpret a wide range of LHC measurements in terms of some other BSM model, as long as the measurements, $\mathbf{X}_a$, can be expressed as functions, $\mu^{i,f}$, of the parameters of the underlying theory.

7.6 Full likelihood results and discussion

This section presents the results using the full combination likelihood, according to the procedure detailed in Section 5.1.2. Performing the interpretation in this way, with access to the full likelihood, is unique to experiments. The theoretical uncertainties in the SM predictions of the cross sections and branching fractions ($\vec{\theta}_s^{\,\mathrm{th}}$) are directly folded into the measurement. Uncertainties in the HEL parametrisation, namely the uncertainties in the A_p and B_{pr} prefactors arising from missing higher-order corrections and limited MC-statistics are neglected. Furthermore, the bbH, tH and ggZH processes are fixed to their SM predictions in the fit, within theoretical uncertainties. This reflects the fact that there are no dedicated analysis categories targeting such production modes in the combination; although conservative this ensures there is no gain in constraining-power from production modes that are not explicitly probed.

Figure 7.10 shows the resulting $q(c_p)$ curves. Two likelihood-scans are performed for each of the seven HEL parameters following the same scenarios used in Section 7.5.1. The first scenario, represented by the dashed lines, corresponds to considering BSM effects in a single EFT operator, such that the other HEL parameters are fixed to 0 in the scan. The second scenario considers BSM effects in all HEL parameters simultaneously, and is shown by the solid lines. In practice, this is performed by scanning over the parameter of interest and profiling the other HEL parameters in the minimisation. As expected, the constraints are tighter for the first scenario, however the act of setting the other HEL parameters to zero introduces a higher degree of model dependence into the interpretation. In all fits, the HEL parameters have been scaled by a constant multiplier as the minimiser is more stable for parameters of order 1.

The best-fit values of the HEL parameters, and the corresponding confidence intervals are summarised in Figures 7.11 and Table 7.3. The double-minimum in the c_u, c_d, and c_ℓ likelihood scans originates from the degeneracy in the relevant scaling functions, where two points in the parameter space correspond to the SM prediction. For example, the c_d

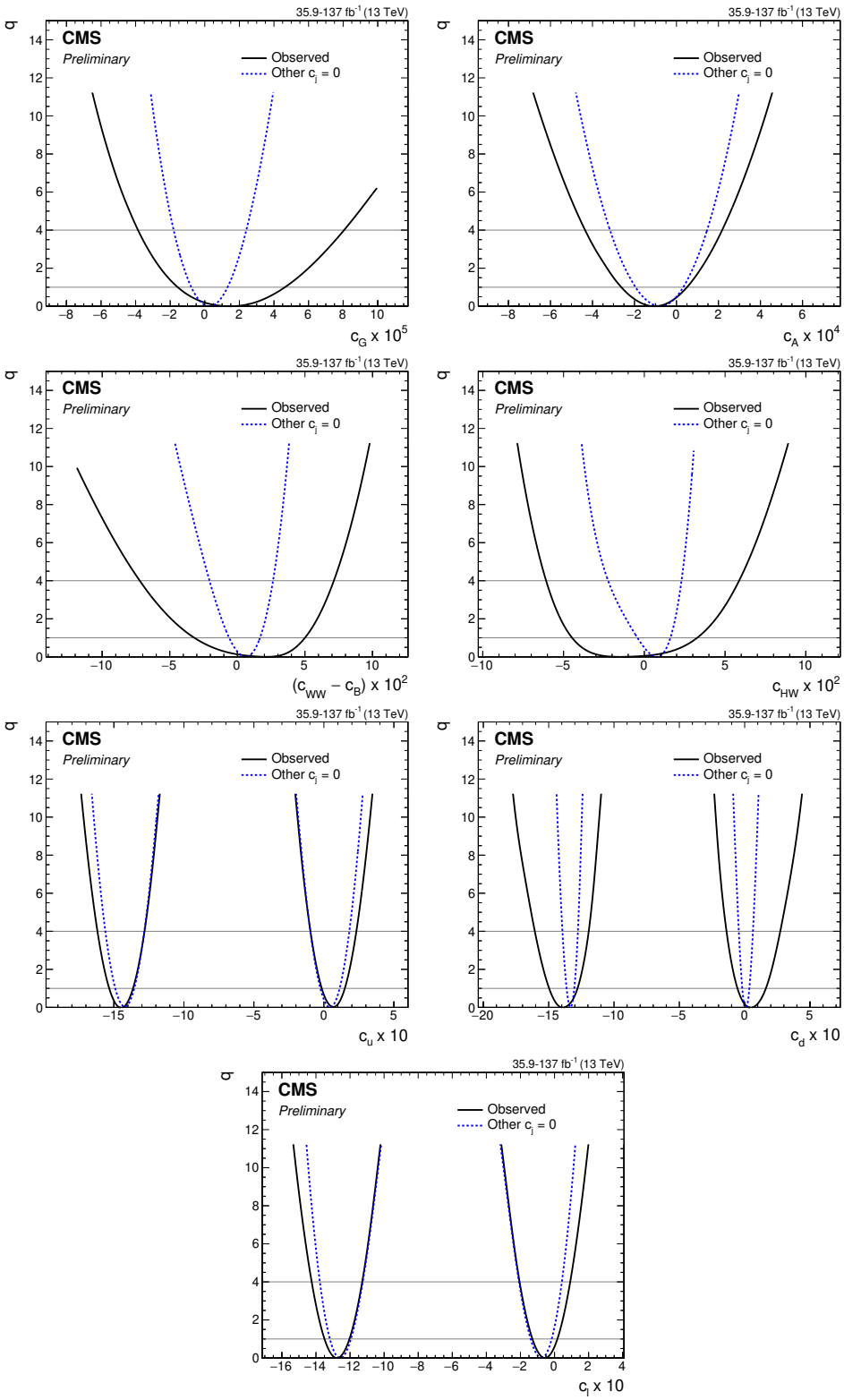


Figure 7.10: Plots of the $q(c_p)$ curves extracted in the HEL interpretation. The solid black (dashed blue) lines correspond to the fits in which the other HEL parameters are profiled (fixed to the SM). The horizontal lines at $q(c_p) = 1$ and 4 indicate the 1σ ($\sim$68%) and 2σ ($\sim$95%) confidence intervals in c_p.

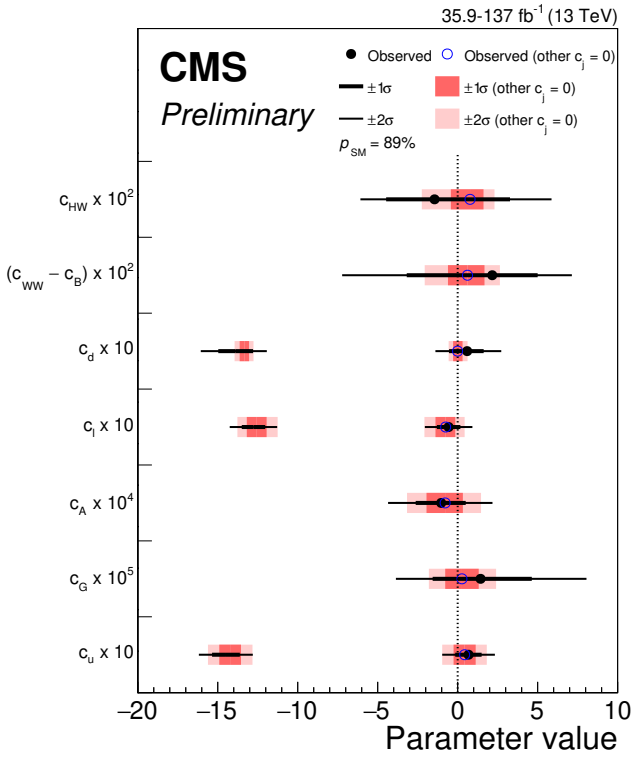


Figure 7.11: Observed best-fit values and confidence intervals for the parameters of interest in the HEL interpretation. The black circles, black thick lines, and black thin lines correspond to the best-fit values, the $\pm1\sigma$ (68%) confidence intervals, and the $\pm2\sigma$ (95%) confidence intervals, respectively, taken from the fits in which the other HEL parameters are profiled. The corresponding results from the fits in which the other HEL parameters are fixed to zero are shown by the hollow blue circles, the dark red bands and the light red bands respectively. The compatibility of the profiled fit with respect to the SM prediction is approximately $p_{\mathrm{SM}} = 89\%$.

constraint is driven by the measurement of the H $\rightarrow$ bb branching fraction, which has a scaling function equal to unity for $c_d = 0$ and $c_d = -4/3$. Including EFT variations in bbH production would help alleviate this degeneracy due to the introduction of a term $\propto c_d\, c_G$ in the bbH scaling function. Nevertheless, as described above, the bbH production mode is constrained to the SM prediction within theory uncertainties, since there is no dedicated analysis category targeting this production mode.

It is observed that both the best-fit values and confidence intervals can vary dramatically between the two fitting scenarios. This is especially true for the parameters which have sizeable correlations e.g. $c_{WW} - c_B$ and c_{HW}, and also for c_d due to its large impact on the Higgs boson total width. The differences arise since the values of the other HEL parameters in the profiled fit can counter the impact from varying the parameter of interest, thus leading to a wider $q(c_p)$ curve. In contrast, this is not possible in the fit

Table 7.3: The best-fit values with 1σ ($\sim$68%) confidence intervals for the parameters of interest in the HEL interpretation. The results from both fitting scenarios are listed: profiling the other HEL parameters in the minimisation, and fixing the other HEL parameters to zero. The expected confidence intervals derived using the Asimov data set are given in brackets. For the c_u, c_d and c_ℓ parameters, the best-fit values and confidence intervals are stated for the minimum closest to zero in the respective $q(c_p)$ curves.

HEL interpretation		
Parameter	Others profiled	Fix others to SM
$c_G \times 10^5$	$1.43^{+3.20}_{-3.00}\left(^{+3.13}_{-2.74}\right)$	$0.27^{+1.05}_{-1.05}\left(^{+1.03}_{-1.01}\right)$
$c_A \times 10^4$	$-1.03^{+1.53}_{-1.59}\left(^{+1.59}_{-1.56}\right)$	$-0.78^{+1.11}_{-1.16}\left(^{+1.10}_{-1.11}\right)$
$(c_{WW} - c_B) \times 10^2$	$2.16^{+2.84}_{-5.35}\left(^{+3.46}_{-5.00}\right)$	$0.62^{+1.06}_{-1.22}\left(^{+1.09}_{-1.23}\right)$
$c_{HW} \times 10^2$	$-1.45^{+4.72}_{-3.03}\left(^{+3.93}_{-3.27}\right)$	$0.77^{+0.84}_{-1.20}\left(^{+1.04}_{-1.38}\right)$
$c_u \times 10$	$0.68^{+0.82}_{-0.83}\left(^{+0.83}_{-0.79}\right)$	$0.43^{+0.69}_{-0.69}\left(^{+0.68}_{-0.67}\right)$
$c_d \times 10$	$0.59^{+1.03}_{-1.13}\left(^{+1.08}_{-1.05}\right)$	$-0.01^{+0.31}_{-0.28}\left(^{+0.30}_{-0.28}\right)$
$c_\ell \times 10$	$-0.57^{+0.74}_{-0.73}\left(^{+0.72}_{-0.77}\right)$	$-0.75^{+0.60}_{-0.64}\left(^{+0.58}_{-0.60}\right)$

in which the other parameters are fixed to 0, resulting in narrower curves and therefore tighter constraints. This effect is less pronounced for the parameters which have smaller correlations and smaller effects on Γ^H: c_u, c_A and c_ℓ.

The correlation coefficients between the HEL parameters are displayed in Figure 7.12. As mentioned above, a large anti-correlation is observed between the pair of HEL parameters which mostly affect the HWW and HZZ vertices, namely c_{HW} and $c_{WW} - c_B$. Including more granular measurements of the VH leptonic and qqH STXS bins will help reduce these correlations, since the c_{HW} and $c_{WW} - c_B$ dependence varies for different kinematic regions of phase space. Moreover, large correlations are observed between c_G and several other parameters. This results from the fact that the parametrisation is defined at LO. As a result the ggH production mode depends solely on c_G, and this dependence is relatively flat across all STXS bins. The c_G parameter is therefore constrained by the total ggH production rate, which cannot be easily distinguished from an overall increase in the total Higgs boson decay width. The exceptions are c_A and c_l, which show a small correlation with c_G as they do not contribute significantly towards Γ^H. This

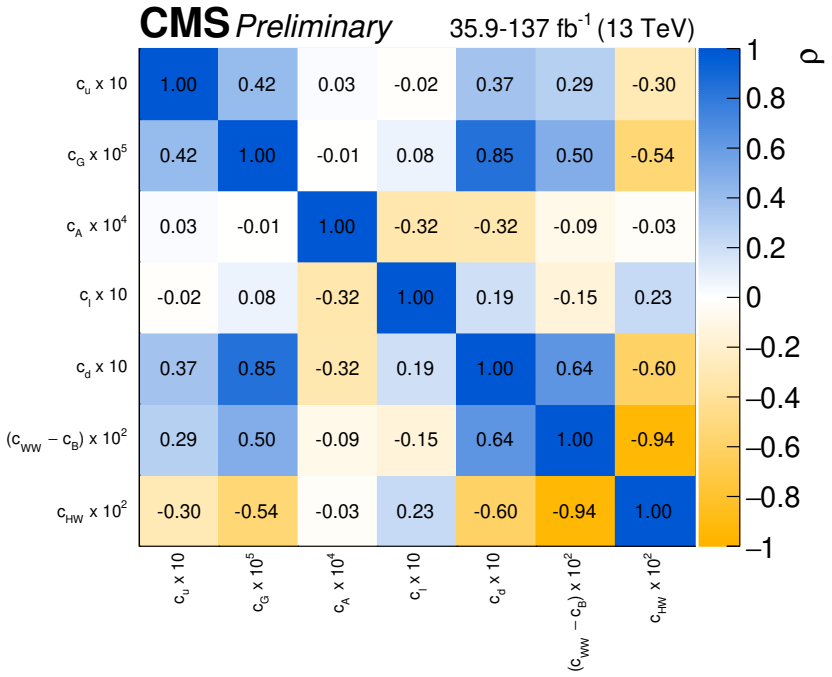


Figure 7.12: Observed correlations between the parameters of interest in the HEL interpretation. The size of the correlations is indicated by the colour scale.

property also explains the sizeable difference between the fixed and profiled $q(c_G)$ curves in Figure 7.10.

All in all, the results are extremely compatible with the SM prediction. This is clearly visible in the $q(c_p)$ curves, such that all parameters in the profiled fit are in agreement with the SM prediction ($c_p = 0$) within the 68% confidence intervals. The corresponding p-value from the profiled fit, with respect to the SM hypothesis, is $p_{\text{SM}} = 89\%$.

The uncertainties in the HEL parameters are amongst the most powerful constraints from Higgs boson measurements, thereby reducing the possible parameter-space for BSM physics in the Higgs sector. Naively, one can convert the constraints on the dimensionless HEL parameters to a lower limit on the energy scale of new physics, Λ, by inverting the relationships defined in Table 7.1. For example, the constraint on c_G,

$$|c_G| \lesssim 5 \times 10^{-5} \text{ @ } 68\% \text{ C.L.}, \tag{7.22}$$

roughly corresponds to the following bound on the scale of new physics in the effective Hgg interaction vertex, assuming the nominal Wilson coefficient, w_G, is of order one,

$$\Lambda^2 = \frac{m_{\text{W}}^2}{g_s^2}\left|\frac{w_G}{c_G}\right| \quad \Rightarrow \quad \Lambda \gtrsim 11 \text{ TeV @ } 68\% \text{ C.L.}. \tag{7.23}$$

Table 7.4: The expected 1σ (∼68%) confidence intervals for the HEL parameters in the profiled fit. The equivalent constraints are shown from the ATLAS result documented in Ref. [183]. The final three rows show parameters which are not included in both the CMS and ATLAS operator subsets.

Parameter	ATLAS result [183]	CMS result
$c_G \times 10^5$	$^{+3.8}_{-2.6}$	$^{+3.1}_{-2.7}$
$c_A \times 10^4$	$^{+2.8}_{-2.2}$	$^{+1.6}_{-1.6}$
$(c_{WW} - c_B) \times 10^2$	$^{+5.7}_{-7.4}$	$^{+3.4}_{-5.0}$
$c_{HW} \times 10^2$	$^{+4.1}_{-4.3}$	$^{+3.9}_{-3.3}$
$c_u \times 10$	$^{+2.4}_{-2.8}$	$^{+0.83}_{-0.79}$
$c_d \times 10$	-	$^{+1.1}_{-1.1}$
$c_\ell \times 10$	-	$^{+0.7}_{-0.8}$
$c_{HB} \times 10^1$	$^{+1.4}_{-1.6}$	-

Ultimately, the HEL model can be matched to UV-complete BSM theories. In doing so, the constraints on the HEL parameters shown here can be re-interpreted as constraints on the BSM couplings and masses of a UV-complete theory. A number of EFT matching examples are provided in Ref. [76].

7.6.1 Comparison to ATLAS result

Table 7.4 compares the expected constraints[9] extracted in this analysis to those from a previous result by the ATLAS Collaboration [183], which combined measurements from the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ decay channels using 2016 data only (36.1 fb^{-1}). The improvements with respect to the ATLAS results are explained below for the common parameters of interest. It is important to keep in mind that this is not a perfect comparison since the ATLAS result considers a different HEL operator subset, $\{\mathcal{O}'\}$.

- The constraint on c_A improves by approximately 40% due to the inclusion of the 2017 data (41.5 fb^{-1}) on top of the 2016 data (35.9 fb^{-1}) in the $H \to \gamma\gamma$ channel. The improvement from the increased luminosity is not as large for c_G since the constraint comes predominantly from ggH production which is limited by systematic uncertainties. Hence, there is a smaller gain from the increase in statistics.

- The constraint on c_u improves dramatically (∼70%) due to the inclusion of analyses which specifically target ttH production in the CMS combination.

[9] The expected constraints offer the best point of comparison in terms of the sensitivity, as they are not affected by statistical fluctuations in data.

- The inclusion of the H $\to$ WW and VH, H $\to$ bb measurements in the combination, in addition to the increased statistics in the H $\to$ ZZ $\to 4\ell$ channel, leads to an improvement in the constraint on $c_{WW} - c_B$ by approximately 35%. Nevertheless, the improvement is somewhat hampered by the differences between the CMS and HXSWG parametrisation, described in Section 7.4.5. For c_{HW}, the parametrisation differences have a larger effect, resulting in a smaller improvement in the constraint ($\sim$15%).

- The c_d and c_ℓ parameters were not considered in Ref [183], since their constraints are driven mainly from the H $\to$ bb and H $\to \tau\tau$ channels. The c_{HB} parameter is not considered in the CMS interpretation.

Recently, the ATLAS Collaboration have superseded the result of Ref. [183] with an EFT interpretation of STXS measurements in the H $\to$ bb, H $\to \gamma\gamma$ and H $\to$ ZZ $\to 4\ell$ decay channels, using data taken during the 2016-2018 periods [184]. This corresponds to a total integrated luminosity of 139 fb^{-1} in each of the input channels. The interpretation is performed in the Warsaw basis [46] (see Section 7.7.1), where the `SMEFTsim` model [79] is utilised to derive the EFT parametrisation. Due to the difference in EFT bases, it is difficult to make a direct comparison with the results of the HEL interpretation shown here. The next CMS Higgs boson combination will include an EFT interpretation in the Warsaw basis, based on the combination of STXS stage 1.2 measurements in all of the major Higgs boson decay channels. More detail regarding this future CMS analysis is provided in Section 7.7.1.

7.6.2 Comparison to the simplified re-interpretation procedure

The differences between the $q(c_p)$ curves in Figure 7.10 and the $\Delta\chi^2(c_p)$ curves in Figures 7.8 and 7.9 directly reflect the assumptions of the simplified re-interpretation procedure. Namely, assuming the likelihood surface is Gaussian and therefore the measurement uncertainties are symmetric, and also treating the input analyses as statistically independent. The results from both approaches are summarised in Table 7.5.

Most of the HEL parameters show a reasonable level of agreement between the two approaches. This is particularly true for the expected sensitivities, given in brackets, and for (the majority of) the fits in which the other parameters are fixed to zero. For the profiled fits, the agreement is less good. This is because the profiled fit represents a more complicated problem, where certain nuances of the likelihood, such as the correlations between input measurements, become increasingly important. The complexity of the CMS Higgs boson combination likelihood makes it extremely difficult to pinpoint the

Table 7.5: The best-fit values and 1σ ($\sim$68%) confidence intervals for the HEL parameters extracted using the full likelihood results extraction and the simplified re-interpretation procedure. The expected 1σ confidence intervals are given in brackets.

Parameter	Other profiled		Fix others to SM	
	Full q	$\Delta\chi^2$	Full q	$\Delta\chi^2$
$c_G \times 10^5$	$1.43^{+3.20}_{-3.00}\left(^{+3.13}_{-2.74}\right)$	$1.44^{+2.77}_{-2.31}\left(^{+2.92}_{-2.41}\right)$	$0.27^{+1.05}_{-1.05}\left(^{+1.03}_{-1.01}\right)$	$-0.27^{+0.82}_{-0.85}\left(^{+0.82}_{-0.85}\right)$
$c_A \times 10^4$	$-1.03^{+1.53}_{-1.59}\left(^{+1.59}_{-1.56}\right)$	$0.97^{+1.31}_{-1.60}\left(^{+1.49}_{-1.55}\right)$	$-0.78^{+1.11}_{-1.16}\left(^{+1.10}_{-1.11}\right)$	$1.08^{+1.01}_{-0.97}\left(^{+1.06}_{-1.01}\right)$
$(c_{WW} - c_B) \times 10^2$	$2.16^{+2.84}_{-5.35}\left(^{+3.46}_{-5.00}\right)$	$-0.11^{+2.49}_{-3.37}\left(^{+2.87}_{-3.76}\right)$	$0.62^{+1.06}_{-1.22}\left(^{+1.09}_{-1.23}\right)$	$0.25^{+0.91}_{-1.06}\left(^{+0.98}_{-1.16}\right)$
$c_{HW} \times 10^2$	$-1.45^{+4.72}_{-3.03}\left(^{+3.93}_{-3.27}\right)$	$0.54^{+2.19}_{-2.24}\left(^{+2.71}_{-2.67}\right)$	$0.77^{+0.84}_{-1.20}\left(^{+1.04}_{-1.38}\right)$	$0.55^{+0.60}_{-0.71}\left(^{+0.80}_{-1.09}\right)$
$c_u \times 10$	$0.68^{+0.82}_{-0.83}\left(^{+0.83}_{-0.79}\right)$	$0.64^{+0.68}_{-0.74}\left(^{+0.76}_{-0.76}\right)$	$0.43^{+0.69}_{-0.69}\left(^{+0.68}_{-0.67}\right)$	$0.50^{+0.57}_{-0.60}\left(^{+0.65}_{-0.70}\right)$
$c_d \times 10$	$0.59^{+1.03}_{-1.13}\left(^{+1.08}_{-1.05}\right)$	$0.52^{+1.02}_{-0.64}\left(^{+1.38}_{-0.81}\right)$	$-0.01^{+0.31}_{-0.28}\left(^{+0.30}_{-0.28}\right)$	$0.26^{+0.28}_{-0.26}\left(^{+0.26}_{-0.24}\right)$
$c_\ell \times 10$	$-0.57^{+0.74}_{-0.73}\left(^{+0.72}_{-0.77}\right)$	$-0.89^{+0.78}_{-0.77}\left(^{+0.73}_{-0.67}\right)$	$-0.75^{+0.60}_{-0.64}\left(^{+0.58}_{-0.60}\right)$	$-0.89^{+0.64}_{-0.71}\left(^{+0.57}_{-0.62}\right)$

origins of the differences. This ultimately motivates *in-house* EFT interpretations, such as the one presented here, since the experiments have access to the full likelihood function.

It should be stressed that there exist more in-depth methods for performing re-interpretations. An extension to the simplified re-interpretation procedure used here is described in Ref. [185], which introduces a new method to deal with asymmetric uncertainties. Furthermore, experiments are beginning to publish simplified or full likelihood functions to accompany their measurements [186]. Such approaches will systematically improve the accuracy of the re-interpretations.

7.7 The future of Higgs EFT measurements in CMS

The interest in using SMEFT to parametrise the effects of many candidate BSM theories has increased substantially in recent years. The HEL interpretation discussed in this chapter was the first application of a SMEFT to STXS measurements at CMS. Going forward, the obvious means to tightening the constraints on the parameters of interest in this approach, is to use more data, thereby decreasing the statistical uncertainties. In addition, increasing the granularity of the measurements will help disentangle regions of phase space that depend differently on the EFT parameters. Not only does this result in tighter constraints and smaller correlations due to the kinematic information available in the measurements, but it also allows more directions (more operators) in the EFT to be probed.

The CMS Collaboration will perform a combination of Higgs boson measurements from all the major decay channels, using the full data set collected in the period 2016-

2018.[10] This corresponds to an integrated luminosity of 137 fb^{-1} in all input analyses. In the combination, the signal processes will be defined at the granularity of the STXS stage 1.2 binning scheme, thus enabling a highly granular EFT interpretation of STXS measurements, akin to the ATLAS result of Ref. [184]. Of course, this will only be possible when all of the individual input analyses to the combination have been finalised.

The HEL interpretation discussed in this chapter provides a solid platform on which to build EFT interpretations of STXS measurements at CMS. Nevertheless, there are a number of caveats associated with the result. These include the choice of operator basis in the EFT expansion, the effect of higher-order corrections in the parametrisation, and the impact of neglecting acceptance effects. This section discusses the implication of these caveats, and how they may be addressed in future Higgs EFT measurements at CMS.

7.7.1 Warsaw basis: SMEFTsim

The first item to be addressed is the choice of operator basis. Ultimately, the choice of basis only becomes relevant when the effect of certain operators are set to zero; new physics will appear equivalently in any complete operator expansion. However, given the limited available statistics and limited breadth of measurements, it is infeasible to constrain all EFT operators simultaneously. As a result, some operators must be fixed to zero, and the choice of basis becomes important.

The interpretation described in this chapter has been performed using a partial implementation of the SILH basis, known as the Higgs Effective Lagrangian [163]. Whilst the SILH basis benefits from being relatively straightforward to match to UV-complete models with modified bosonic interactions, and is easily propagated to the impact on LEP observables, it suffers from the fact that the same operators contribute to both anomalous triple gauge couplings (aTGC) in EW measurements and Higgs observables [187]. Since performing this interpretation, the emphasis in the high-energy physics community has shifted towards the *Warsaw basis* [46]. The Warsaw basis is more appropriate for BSM physics with modified fermionic interactions, and has already been used in a number of Higgs, EW and top-quark measurements[11]. By adopting a common language in the EFT interpretations, it becomes possible to combine results from multiple disciplines (Higgs, EW, top), thus enabling the ultimate consistency test of the SM in which many directions of the EFT parameter space can be probed simultaneously. Motivated by this notion of a cross-discipline *global combination*, future EFT interpretations of STXS measurements at CMS will be performed in the Warsaw basis.

The `SMEFTsim` package [79] provides a number of LO implementations of the com-

[10] Including the $\text{H} \to \gamma\gamma$ measurements described in Chapters 4–6.
[11] Some examples are provided in Refs. [184, 188, 189].

plete Warsaw basis, up to operators of dimension-6, with different assumptions regarding the flavour symmetry. The implementation discussed in the following section assumes a $U(3)^5$ flavour symmetry i.e. unbroken global flavour symmetry present in the SM outside the Yukawa sector, and includes the possibility of CP-violating phases. The model is configured using the theoretically-favoured m_W, m_Z and G_F input-parameter scheme. Additionally, the Warsaw-basis parameters, $\vec{C}$ (capitalised to help differentiate from the HEL parameters, $\vec{c}$), are defined for a nominal energy scale of $\Lambda = 1$ TeV. This differs from the HEL parameters, which absorbed the Λ^{-2} factor into their definition. Nevertheless, the Warsaw-basis parameters can be obtained for alternate values of $\Lambda = \Lambda^*$ by scaling with the factor $(\Lambda^*/1\text{ TeV})^2$.

7.7.2 Warsaw parametrisation and expected sensitivities

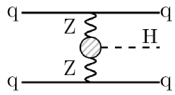
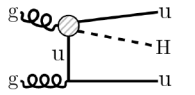
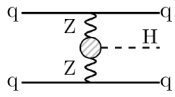
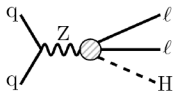
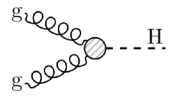
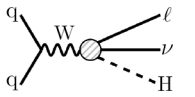
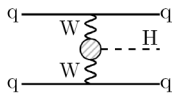
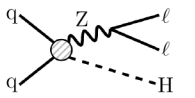
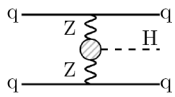
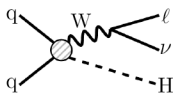
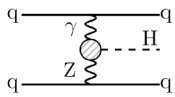
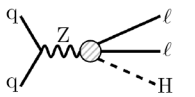
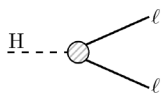
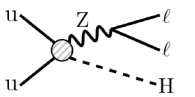
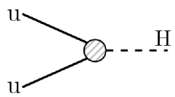
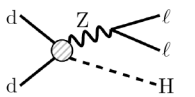
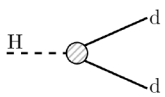
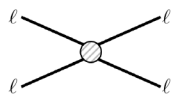
The signal-yield parametrisation has been derived in the Warsaw basis using the same procedure detailed in Section 7.4.1. All MC-generator options are identical to those defined in Appendix G, except the `HEL_UFO` model has been exchanged with the `SMEFTsim` model. The parametrisation is defined at the granularity of the STXS stage 1.2 binning scheme to match the input analyses in the future CMS Higgs boson combination.

Again, non-zero effects are considered in only a subset of EFT operators, since it is not possible to constrain all directions of parameter space using Higgs boson measurements alone. A larger number of operators, listed in Table 7.6, are included in this parametrisation to account for the enhanced sensitivity from the increased integrated luminosity and more granular measurements. The full set of cross-section and branching-fraction scaling functions are derived in the Warsaw basis including both the linear and quadratic terms, and are tabulated in Appendix I. As an example, the scaling terms for the ZH leptonic and ttH stage 1.2 bins are shown as functions of the C_{HW} and $|C_{uG}|$ Warsaw-basis parameters, respectively, in Figure 7.13.[12]

The simplified re-interpretation procedure, introduced in Section 7.5, is used to provide an estimate of the sensitivity to the Warsaw-basis parameters, applying the preliminary parametrisation of Appendix I. The measurements, $\mathbf{X}_a$, entering the χ^2 function are the current public STXS stage 1.2 measurements performed by the CMS experiment: the $H \to \gamma\gamma$ analysis described in Chapters 4–6, and the $H \to ZZ \to 4\ell$ analysis of Ref. [190]. The results from the $H \to \tau\tau$, $H \to bb$ and $H \to WW$ decay channels are not available at the time of writing this thesis. For STXS bins that are merged in the measurement, the corresponding scaling function is taken as the weighted sum of the individual scaling functions, where each bin is weighted according to its relative SM

[12] It should be stressed that this parametrisation, particularly the chosen operator subset, is not yet final and is likely to change before the future CMS Higgs boson combination.

Table 7.6: The dimension-6 operator subset, $\{\mathcal{O}\}$, considered in the Warsaw-basis parametrisation shown in Appendix I. An example Feynman diagram of the corresponding contact interaction is shown for each operator. The quantity, $\sigma^{\mu\nu}$, is defined by the gamma matrices relation: $\sigma^{\mu\nu} = i[\gamma_\mu, \gamma_\nu]/2$. A U(3)^5 flavour symmetry is assumed, such that in the diagrams, u, d and ℓ represent all up-type quarks, all down-type quarks, and all charged leptons, respectively.

Parameter	Operator definition	Example diagram	Parameter	Operator definition	Example diagram
$C_{H\text{Box}}$	$(H^\dagger H)\Box(H^\dagger H)$		$\lvert C_{uG}\rvert$	$(\bar{Q}_L\sigma^{\mu\nu}T^a u_R)(\tilde{H}G^{a,\mu\nu})$	
C_{HDD}	$(H^\dagger D^\mu H)^*(H^\dagger D_\mu H)$		$C_{H\ell}^{(1)}$	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{L}_L\gamma^\mu L_L)$	
C_{HG}	$(H^\dagger H)(G^a_{\mu\nu}G^{a,\mu\nu})$		$C_{H\ell}^{(3)}$	$(H^\dagger i\overleftrightarrow{D}^i_\mu H)(\bar{L}_L\sigma^i\gamma^\mu L_L)$	
C_{HW}	$(H^\dagger H)(W^i_{\mu\nu}W^{i,\mu\nu})$		$C_{Hq}^{(1)}$	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{Q}_L\gamma^\mu Q_L)$	
C_{HB}	$(H^\dagger H)(B_{\mu\nu}B^{\mu\nu})$		$C_{Hq}^{(3)}$	$(H^\dagger i\overleftrightarrow{D}^i_\mu H)(\bar{Q}_L\sigma^i\gamma^\mu Q_L)$	
C_{HWB}	$(H^\dagger\sigma^i H)(W^i_{\mu\nu}B^{\mu\nu})$		C_{He}	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{\ell}_R\gamma^\mu\ell_R)$	
$\lvert C_{eH}\rvert$	$(H^\dagger H)(\bar{L}_L\ell_R H)$		C_{Hu}	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{u}_R\gamma^\mu u_R)$	
$\lvert C_{uH}\rvert$	$(H^\dagger H)(\bar{Q}_L u_R\tilde{H})$		C_{Hd}	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{d}_R\gamma^\mu d_R)$	
$\lvert C_{dH}\rvert$	$(H^\dagger H)(\bar{Q}_L d_R H)$		$C_{\ell\ell}^{(1)}$	$(\bar{L}_L\gamma_\mu L_L)(\bar{L}_L\gamma^\mu L_L)$	

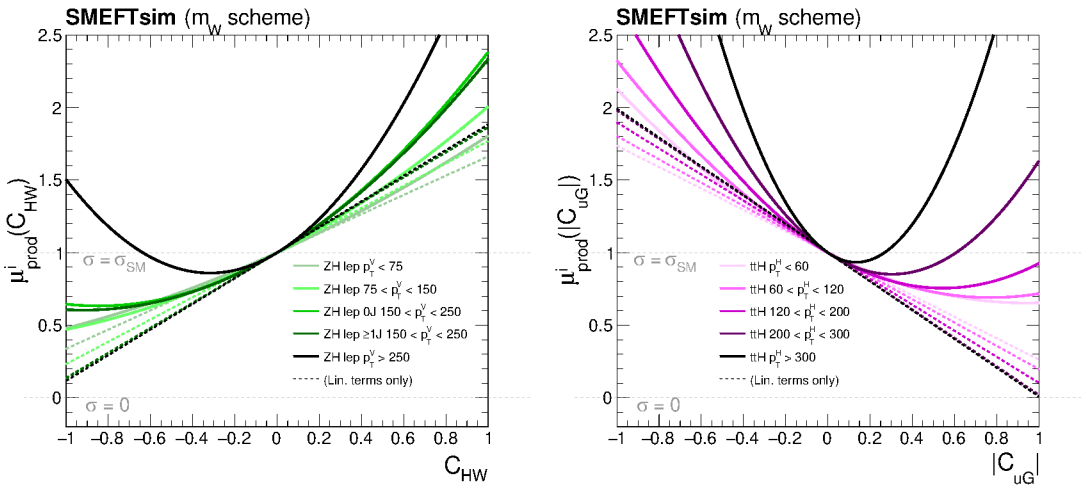


Figure 7.13: Cross-section scaling functions, $\mu^i_{\text{prod}}(\vec{c})$, for the ZH leptonic (left) and ttH (right) STXS stage 1.2 bins in terms of C_{HW} and $|C_{uG}|$, respectively. The dashed lines indicate the scaling functions when only the linear terms are considered ($B_{pr} = 0$).

fraction in the merged bin[13]. Due to the limited number of input measurements, only variations in pairs of Warsaw-basis parameters are considered.

The best-fit values and the 68% and 95% confidence level contours are shown for four pairs of Warsaw-basis parameters in Figure 7.14. This by no means constitutes the full set of parameters which Higgs boson measurements are sensitive to; the results are shown to simply give a flavour of what kind of constraints will be possible in future CMS EFT interpretations. In the plots, the observed and expected contours are shown in red and grey, respectively. The observed best-fit points in the $C_{H\text{Box}}$-vs-C_{HW} and $C^{(1)}_{Hq}$-vs-$C^{(3)}_{Hq}$ fits are consistent with the SM, (0,0), within the 68% confidence level contour. Interestingly, the C_{HG}-vs-$|C_{uG}|$ and $C^{(1)}_{H\ell}$-vs-$C^{(3)}_{H\ell}$ fits show discrepancies with the SM at around the 95% confidence level. It is crucial that more input measurements are included in the future interpretation to see if these discrepancies remain.

Figure 7.15 compares the constraints in the C_{HG}-vs-$|C_{uG}|$ fit, using different combinations of input measurements. The left-hand plot shows the 68% confidence level contours when only including the $H \to \gamma\gamma$ results (purple) and only including the $H \to ZZ \to 4\ell$ results (green), and compares to the combination of the two results (red). Clearly, both input channels have a significant impact on the sensitivity. From the expected contours (dashed lines), we see the $H \to ZZ \to 4\ell$ channel is more sensitive to the C_{HG} parameter than the $H \to \gamma\gamma$ channel, whilst the $H \to \gamma\gamma$ channel is more sensitive to $|C_{uG}|$. This latter observation is due to the relatively poor measurement of ttH production ($\propto |C_{uG}|$)

[13]This will not be necessary in the full likelihood results extraction since each signal processes is defined at the full granularity of the stage 1.2 binning scheme.

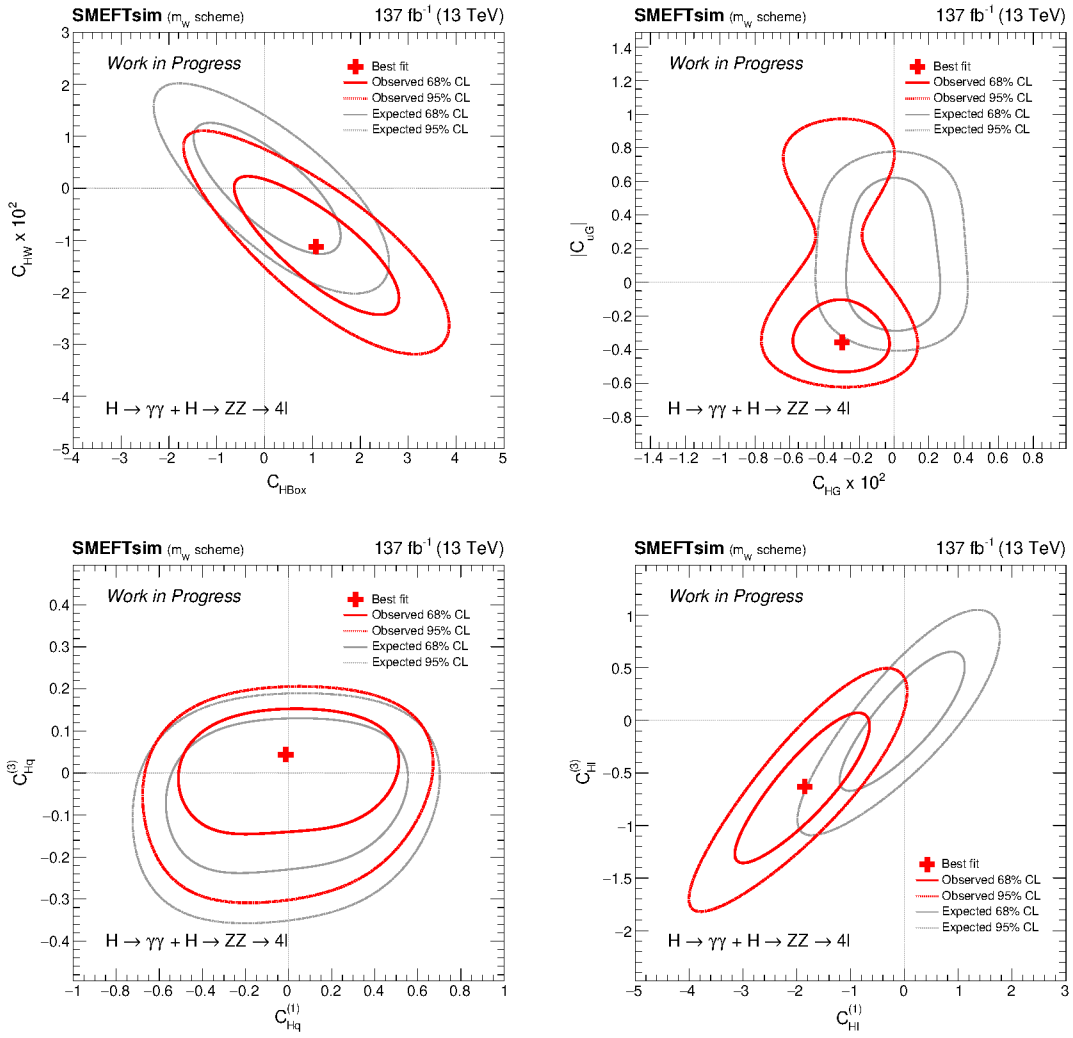


Figure 7.14: The best-fit values, 68% and 95% confidence level contours for pairs of Warsaw-basis parameters, extracted using the simplified re-interpretation procedure on STXS stage 1.2 measurements from the CMS H $\to \gamma\gamma$ and H $\to$ ZZ $\to 4\ell$ analyses. The red and grey lines show the observed and expected results, respectively.

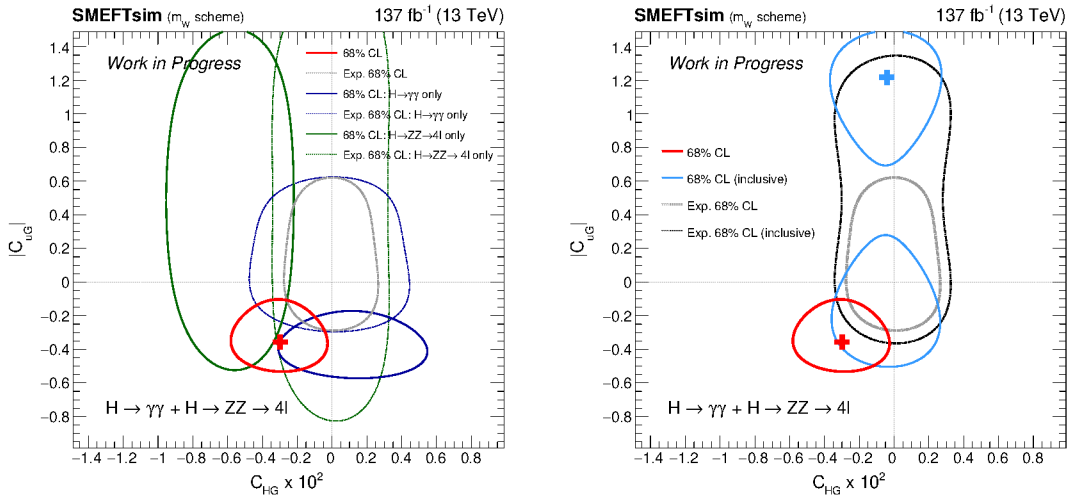


Figure 7.15: Constraints in the C_{HG}-vs-$|C_{uG}|$ plane, using different combinations of input measurements. The left-hand plot shows the 68% confidence level contours when using only the H $\rightarrow$ $\gamma\gamma$ results (purple), only the H $\rightarrow$ ZZ $\rightarrow$ 4ℓ results (green), and their combination (red). The right-hand plot compares the 68% confidence level contours when using the STXS stage 1.2 measurements (red), compared to using only the inclusive production mode measurements (blue), from both decay channels. In both plots, the expected contours are shown by the dashed lines.

in the H $\rightarrow$ ZZ $\rightarrow$ 4ℓ channel, which suffers from extremely low statistics. The plot on the right shows the 68% confidence level contours when using the highly granular STXS stage 1.2 measurements in the fit (red), compared to using the inclusive (STXS stage 0) measurements (blue). The sensitivity to C_{HG} does not significantly improve as the C_{HG} dependence is relatively flat across all ggH stage 1.2 bins. On the other hand, the $|C_{uG}|$ constraint improves dramatically: the 68% confidence contour tightens and the degeneracy is alleviated. This is a direct consequence of measuring ttH production in different kinematic regions in the H $\rightarrow$ $\gamma\gamma$ analysis of Chapters 4–6. Ultimately, the results shown here highlight the importance of including the kinematic information, which is made available in STXS measurements, across multiple Higgs boson decay channels.

Before finishing the discussion on the Warsaw basis, it is important to highlight one consequence of building a fully general SMEFT framework, which is that EFT effects are not limited to the "signal" processes, but can also impact the "backgrounds". In fact, it is conceivable to imagine a signal process in one analysis may be considered as a background in another e.g. ttZ production in a top quark analysis versus ttZ production in a Higgs boson analysis. To construct a fully consistent EFT interpretation, especially when combining different analyses (Higgs, top and EW), it will be necessary to consider EFT effects in all processes simultaneously.

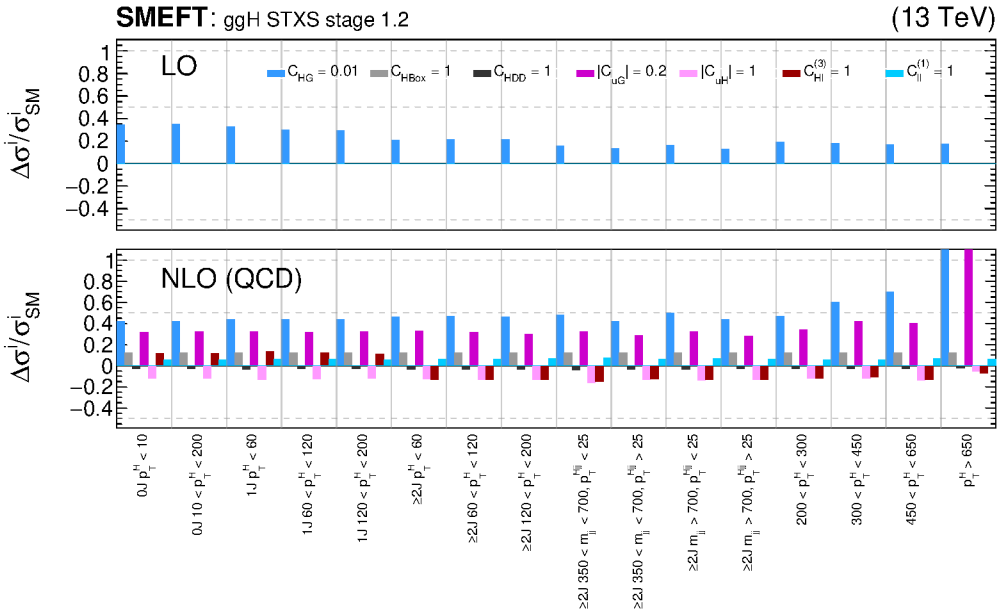


Figure 7.16: The impact of a set of Warsaw-basis operators on the ggH STXS stage 1.2 bins at LO (top panel) and at NLO QCD (bottom panel).

7.7.3 SMEFT at NLO

There are three independent means for improving EFT interpretations. The first is in the experimental determination of the observables e.g. $[\sigma^i \cdot \mathcal{B}^f]_{\text{obs}}$, which can be improved by taking more data and reducing the sources of systematic uncertainty in the measurement. Second, is the theoretical predictions of the observables in the SM e.g. $[\sigma^i \cdot \mathcal{B}^f]_{\text{SM}}$, which are improved by going to higher orders in their respective calculations. Finally, the third rests in how precisely we can parametrise the deviations between the previous two; in the interpretation described here it is the uncertainties in the derived A_p and B_{pr} prefactors.

The precision of the parametrisation is systematically improved by including higher-order contributions to the EFT predictions. An important milestone in this direction is documented in Ref. [164], where higher-order QCD corrections to the SMEFT predictions are computed for any observable. This is particularly important for processes which occur via loops in the lowest order diagrams, such as ggH production. In the LO EFT implementations, ggH depends solely on the effective vertex operator, $\mathcal{O}_{HG}$, however, going to higher-orders in the parametrisation calculation (i.e. including additional loops in the diagrams) introduces dependencies on other operators.

Figure 7.16 shows the impact on the ggH STXS stage 1.2 bins from a subset of operators in the Warsaw basis. In moving from the LO (top panel) to the NLO (bottom panel) parametrisation, not only do the C_{HG} effects become increasingly dependent on

the Higgs boson kinematics, but additional dependencies are observed. For example, by resolving the top quark loop, we see that ggH gains a dependency on the $|C_{uG}|$ Warsaw-basis parameter (purple). This plot highlights the importance of including higher-order contributions in future CMS EFT interpretations, particularly when parametrising the loop-induced processes.

7.7.4 EFT after the detector

Throughout this chapter, we have seen that EFT effects modify not only inclusive event rates, but also the event kinematics. This property means that the constraints on EFT parameters benefit from measuring kinematic distributions, for instance in the STXS framework or in differential cross section measurements. However, modifications to the event kinematics can directly impact both the detector efficiency and the analysis acceptance, meaning the $\epsilon_k^{i,f}$ factors become dependent on the EFT parameters, $\vec{C}$. In the interpretation described in this chapter, these factors are derived using SM MC, and are assumed to take their SM predicted values. This assumption is particularly relevant in the STXS framework, where the decay products of the Higgs boson are not restricted to a fiducial phase space[14]. As a result, the extrapolation from the experimental phase space in which the events are measured, to the full kinematic phase space can be large. If the events in the experimental phase space depend differently on the EFT parameters to the total phase space, then the SM assumption of the $\epsilon_k^{i,f}$ factors break down and the validity of the interpretation is questioned.

The impact is especially severe for the four-body $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ decay channel, where at least one of the Z bosons must be off-shell. Figure 7.17 shows the truth-level dilepton invariant-mass distributions, m_{Z1} (m_{Z2}), for the Z candidate with invariant mass closest to (furthest from) the nominal Z boson mass, $m_{\mathrm{Z}} = 91.188$ GeV [24]. Events are required to have a dilepton angular separation of $\Delta R_{\ell\ell} > 0.05$ to avoid divergences in the collinear limit. The distributions are shown for the SM prediction and when turning on various contributions from EFT operators in the Warsaw basis. In the corresponding analysis selection [190], all reconstructed pairs of leptons that form a Z candidate are required to have an invariant mass greater than 12 GeV, shown by the dashed vertical line in the m_{Z2} plot. Clearly, the EFT effects are dramatically reduced in the experimental phase space ($m_{\ell\ell} > 12$ GeV), compared to the full kinematic distribution, thereby reducing the sensitivity to the relevant EFT parameters. Figure 7.18 demonstrates the impact on the $\mathrm{H} \to \mathrm{ZZ}$ partial-width scaling as a function of C_{HW}, C_{HB} and $C_{H\ell}^{(3)}$, when including the $m_{\ell\ell} > 12$ GeV requirement at truth-level. Large discrepancies are observed in the scaling functions for the Warsaw-basis parameters that exhibit a strong kinematic dependence i.e.

[14] This may change in the future, where discussions regarding a "STXS at decay" scheme are ongoing.

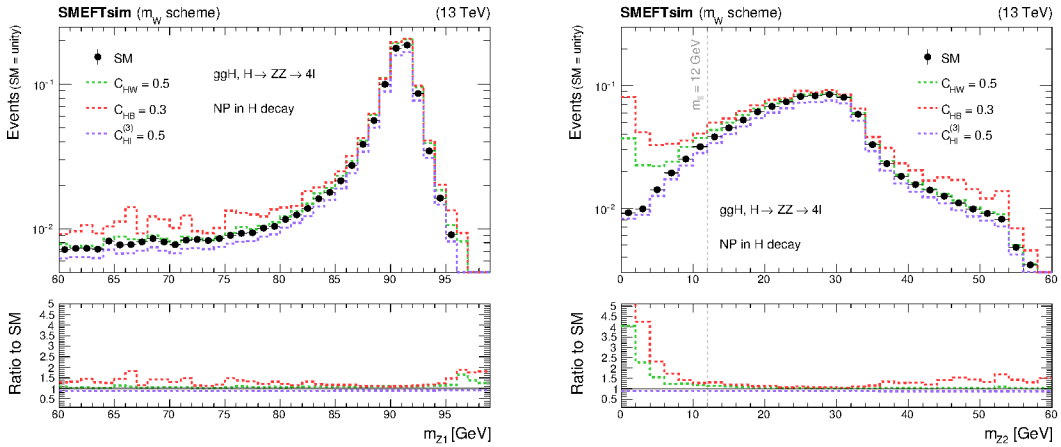


Figure 7.17: The m_{Z1} and m_{Z2} distributions for H $\rightarrow$ ZZ $\rightarrow$ 4ℓ events at LO, where m_{Z1} (m_{Z2}) is the invariant mass of the Z boson candidate closest to (furthest from) the nominal Z boson mass. The black points correspond to the SM prediction, whilst the coloured lines show the distributions when various EFT contributions are introduced. The dashed vertical line in the m_{Z2} distribution indicates the selection criteria used in the H $\rightarrow$ ZZ $\rightarrow$ 4ℓ analysis, which demands all reconstructed pairs of leptons that form a Z candidate have an invariant mass greater than 12 GeV.

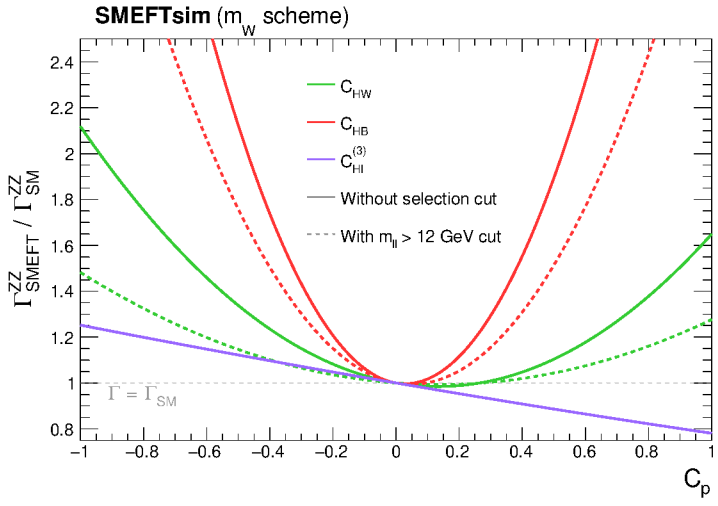


Figure 7.18: The impact of including the selection criteria ($m_{\ell\ell} > 12$ GeV) at truth-level, on the H $\rightarrow$ ZZ $\rightarrow$ 4ℓ partial-width scaling functions, in terms of the Warsaw-basis parameters: C_{HW}, C_{HB} and $C^{(3)}_{H\ell}$.

C_{HW} and C_{HB}, for with and without the selection requirement. This suggests that going forward it will be crucial to account for these so-called *acceptance effects*, particularly for processes which are significantly impacted.

There are two approaches to including the acceptance effects in the signal parametrisation. The first, demonstrated in Figure 7.17, is to apply the analysis-like selection cuts at truth-level and re-derive the relevant scaling functions. This somewhat lightweight approach is quick, but may oversimplify the problem at hand by not accounting for the detector efficiencies, nor would it be possible to calculate the impact for higher-level analysis selection techniques, such as using BDTs or DNNs. The second approach is to include the EFT effects in the full MC simulation samples used in the individual analyses, and subsequently propagate these effects through the whole analysis chain i.e. the event reconstruction and selection. The correct signal-yield parametrisation would then be derived using only the subset of events which enter the respective analysis region. Although this approach is the most thorough, generating separate samples for each EFT effect can be cumbersome and extremely time-consuming. A workaround for this is discussed in the following section.

7.7.5 Standalone reweighting procedure

The MADGRAPH matrix-element reweighting functionality [180] used for the signal parametrisation derivation in Section 7.4.1, can be exported as a standalone reweighting package. Subsequently, the reweighting procedure can be applied to existing MC events, provided all truth-level event information is available to fully specify the per-event matrix-element, $\mathcal{M}$. In doing so, events in the nominal SM simulation can be reweighted to any point in the EFT parameter space, such that the corresponding event weights can be propagated through the analysis framework. The obvious benefit here is that only one simulation sample, the nominal SM sample, is required per process. An example is provided in Figure 7.19, which shows the *reconstructed* p_T and η distributions for the leading lepton in ZH leptonic events, where the nominal SM events have been reweighted to account for variations in a number of Warsaw-basis parameters. This effectively shows the impact of EFT *after the CMS detector*.

The application of this standalone reweighting procedure is still in its preliminary stages, and there are a number of obstacles to overcome before integrating into a full CMS analysis. These include the reweighting of higher-order SM events (e.g. NNLO) using a lower-order reweighting scheme (e.g LO), and generating enough events to sufficiently populate all regions of phase space, including those regions which are not particularly favoured in the SM but might be enhanced in the EFT.

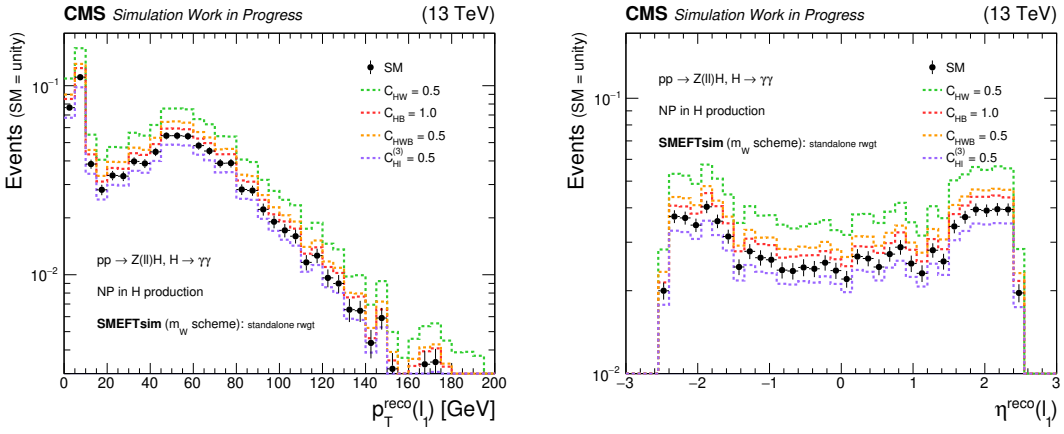


Figure 7.19: The reconstructed p_T and η distributions for the leading lepton in ZH leptonic events. The black points correspond to the SM prediction, whilst the coloured lines show the distributions when various EFT contributions are introduced. These events are taken from a nominal SM simulation sample, and are reweighted to different points in the EFT parameter space using the standalone reweighting procedure.

7.8 Summary

This section has detailed an EFT interpretation of STXS measurements at CMS. One of the main benefits in using an EFT framework is that the results are agnostic to a wide class of potential BSM theories that reduce to the SM in the IR limit. As a result, the EFT interpretation offers a (almost) model independent framework on which to search for BSM physics.

Higgs boson production cross sections and branching fractions have been parametrised as functions of EFT Wilson coefficients, such that deviations from zero in these coefficients signal new physics. This parametrisation was extracted using the Higgs Effective Lagrangian (HEL) model, at the granularity of the STXS framework. By doing so, the kinematic information available in STXS measurements was utilised to further constrain EFT operators. The parametrisation was applied to the most recent CMS Higgs boson combination, which combines STXS measurements from all the major Higgs boson decay channels. All results were found to be consistent with SM expectations, where the corresponding constraints are amongst the most powerful constraints on EFT parameters from Higgs boson measurements to-date.

The section concluded with a look to the future of Higgs EFT measurements at CMS. The transition towards the Warsaw basis was discussed, motivated by the possibility of a global cross-discipline combination. This included a sensitivity estimate in a number of Warsaw-basis parameters from a combination of STXS stage 1.2 measurements in the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ decay channels. Following this, the systematic improvement

of the EFT parametrisation by including higher-order corrections was addressed. Finally, the importance of including the effects of the CMS detector in future EFT analyses was highlighted.

Chapter 8

The High-Luminosity LHC

8.1 Introduction

The obvious means for improving precision measurements is to take more data. Run 2 of the LHC finished in 2018, delivering p-p collisions at $\sqrt{s} = 13$ TeV and reaching a maximum instantaneous luminosity of around 2×10^{34} cm^{-2} s^{-1}. After a shutdown period for upgrades and maintenance, Run 3 is expected to commence in 2022 and finish in 2024. Here, the LHC machine will operate with an instantaneous luminosity of 2×10^{34} cm^{-2} s^{-1} over the full data-taking period, at $\sqrt{s} = 13$ or 14 TeV. By the end of Run 3, over $\mathcal{L} = 300$ fb^{-1} of p-p collision data will have been collected by the CMS experiment.

Since the statistical uncertainty in a measurement scales according to $1/\sqrt{\mathcal{L}}$, simply operating with the same beam-conditions beyond Run 3 is not particularly interesting; the machine would have to run for fifteen years to see a factor of two improvement in the precision. A future operation of the LHC machine, referred to as the High-Luminosity LHC (HL-LHC) [191], will upgrade the LHC beam and maximise the physics potential. During this phase, proton bunches will be collided at $\sqrt{s} = 14$ TeV, with a nominal instantaneous luminosity of 5×10^{34} cm^{-2} s^{-1} (2.5 times the Run 3 instantaneous luminosity), rising to as high as 7.5×10^{34} cm^{-2} s^{-1} towards the end of operation. Scheduled to begin in 2027, this means by the mid-2030s an integrated luminosity of around 3000 fb^{-1} (3 ab^{-1}) will be available for physics analysis. Not only will this dramatically reduce the statistical uncertainty in existing measurements, but it will also open the door for new measurements and analyses that are not possible with a limited data set. One such example is provided in Section 8.4, looking at the potential of constraining the Higgs boson self-coupling using ttH + tH differential measurements.

Unfortunately, the increase in the instantaneous luminosity comes at a price. At the HL-LHC, the mean pileup-per-bunch-crossing is expected to be as high as 200. This poses a major challenge both in terms of the higher radiation levels delivered to the LHC exper-

iments, and the ability to trigger on and reconstruct physics of interest in a high-pileup environment. As a result, all LHC experiments will undergo major upgrade programmes to accommodate the HL-LHC conditions. The operation of the CMS detector during the HL-LHC is referred to as CMS Phase-2. For this operation, the CMS experiment will not only replace the existing parts of the detector with high levels of radiation damage, but will improve the functionality of these parts in terms of the radiation-hardness and the granularity of the readout channels [192]. Amongst the numerous upgrades, two of the most substantial developments are with respect to the L1T and the endcap calorimeters [3, 193].

The design latency of the CMS Phase-2 L1T will be extended to 12.5 µs. This enables the use of more granular information from the various subdectors in the trigger decision including, for the first time, hits in the inner tracker. Advances in Field Programmable Gate Array (FPGA) technology [194] will also facilitate more complex algorithms at the L1T stage, including the application of a Particle-Flow-like algorithm to link the various subdetectors, as well as ML algorithms such as Boosted Decision Trees (BDTs). Furthermore, the combined improvements in the front-end electronics and the data-acquisition (DAQ) system [195] allow the maximum output event rate of the L1T to be increased from 100 kHz to 500 kHz. All in all, these upgrades are crucial for maintaining an excellent trigger efficiency in a high-pileup environment.

By the end of Run 3, the CMS endcap calorimeters will be significantly radiation-damaged, such that their performance will be substantially reduced. For Phase-2, the ECAL and HCAL endcaps will be completely replaced by a single subdetector, known as the High-Granularity Calorimeter (HGCAL), which will have both electromagnetic and hadronic components. The HGCAL will exhibit a fine segmentation in both the transverse and longitudinal directions, as well as timing capabilities, to enable the precise reconstruction of both electromagnetic and hadronic showers in four dimensions (three spatial and one temporal). This is an extremely exciting prospect for studying physics processes in the forward region, such as reactions initiated by vector-boson fusion (VBF).

8.2 The High Granularity Calorimeter

The design of the HGCAL [193] is driven mainly by the need to be radiation tolerant and maintain a good energy resolution over the full lifetime of the HL-LHC project, as well as exhibiting sufficient granularity to perform calorimetry in the forward region; a schematic of this design is provided in Figure 8.1. It will cover the pseudorapidity region: $1.5 < |\eta| < 3$, and consist of an electromagnetic compartment (CE-E), followed by a hadronic compartment (CE-H). The design is comprised mainly of silicon detector

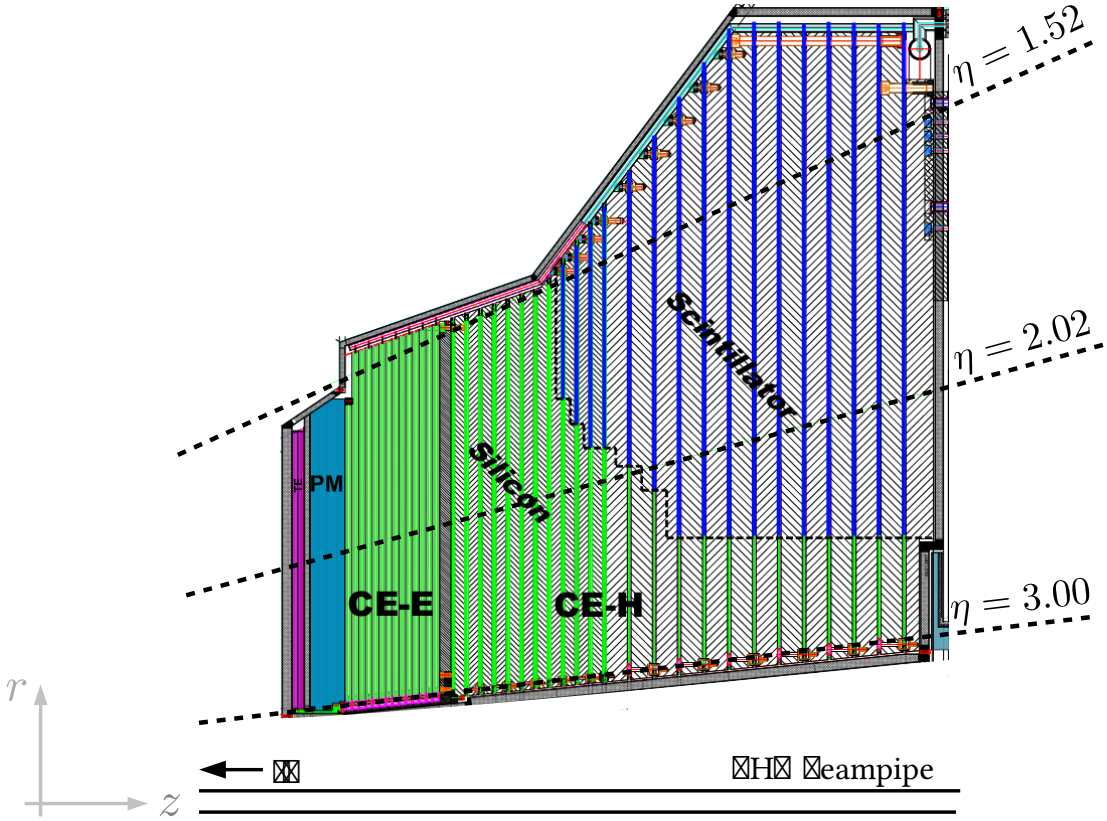


Figure 8.1: Longitudinal structure of the HGCAL. The electromagnetic compartment (CE-E) consists of 28 sampling layers of hexagonal silicon sensors (green), interleaved with absorber layers. The total depth of the CE-E is 26 radiation lengths (X_0) and 1.7 nuclear interaction lengths (λ_I). The hadronic compartment (CE-H) is formed of 24 sampling layers, with an increased thickness for the rear 12 layers. The active material in this region is composed of silicon sensors (green) and plastic scintillator (blue), for the region of the detector with lower levels of radiation. The CE-H extends the total depth of the HGCAL to 10.7 λ_I. In front of the CE-E lies the endcap timing layer (TE, purple) which aids in the mitigation of pileup, and a polythene neutron moderator (PM) layer to reduce the neutron flux in the CE-E. Figure has been adapted from that shown in Ref. [193].

technology, which has been shown to withstand high-radiation levels and perform well in the high particle-flux environment. This is supplemented with plastic scintillator tiles towards the rear of the detector, where the scintillation light is read out by silicon photo-multiplier tubes. The active material of the detector is interleaved with layers of lead and stainless-steel absorber, which increases the effective depth of the calorimeter and thus provides good containment of the particle showers.

The HGCAL will feature unprecedented transverse and longitudinal segmentation. Both the silicon sensors and the plastic scintillator tiles will be highly segmented, with a size of ≈ 0.5 to 1 cm^2 and ≈ 4 to 30 cm^2, respectively. Coupling this property with the fact

that the highly-dense material in the HGCAL leads to laterally-compact showers enables excellent shower separation. Moreover, the fine lateral granularity limits the region used for the shower-energy measurement and thus minimises the energy contamination from particles originating in pileup interactions. The longitudinal segmentation of 28 layers in the CE-E and 24 layers in the CE-H provides a handle on the longitudinal development of a particle shower. This capability, which is not possible in the current CMS calorimeters, improves the electromagnetic energy resolution, enables pattern recognition, and helps to mitigate showers originating from pileup. Finally, the intrinsic timing capabilities of the silicon sensors mean that each energy deposit can be given a precise time stamp. This timing information is especially useful for pileup rejection, identification of the interaction vertex, and for the PF reconstruction.

An extra design requirement of the HGCAL is the ability to contribute to the L1T decision. As introduced in the previous section, recent advances in FPGA technology in both the available hardware resources and algorithm implementation tools have enabled the application of more complex and powerful algorithms at the L1T stage. The following section is dedicated to an ML algorithm designed to differentiate electrons and photons from jets in the HGCAL L1T. More detail regarding the foundations of ML algorithms is provided in Appendix B.

8.3 Electron and photon identification in the HGCAL L1T

The high granularity of the HGCAL enables electromagnetic showers originating from single electrons or photons (e/γ) to be resolved, even in the very high-occupancy environment of the HL-LHC. To successfully reconstruct events containing such objects, it is necessary to correctly identify e/γ showers at the L1T decision stage. This section investigates the application of a BDT to distinguish e/γ candidates (signal) from pileup-induced clusters (background) in the HGCAL L1T. For the studies, dedicated MC-simulation samples are used which correspond to collisions in the CMS Phase-2 detector with a centre-of-mass energy of $\sqrt{s} = 14$ TeV and an average of 200 pileup interactions per event.

Despite the increased total latency of the CMS Phase-2 L1T allowing HGCAL information to be used, it is not possible to read out all the data with full granularity. To reduce the data, only alternate layers in the CE-E are used, and neighbouring silicon sensors (scintillator tiles) are summed into so-called *trigger cells* with a granularity of approximately 4 cm^2 (16 to 100 cm^2). Additionally, a reasonably tight energy threshold is placed on the trigger cells and no timing information is stored. A clustering algorithm [3] is applied to the selected trigger cells, which (in a similar fashion to the algorithm described in Section 3.4) first seeds the clusters and then builds topological clusters around

these identified seeds. The resulting 3D clusters form the collection of HGCAL *trigger primitives*, on which the L1T decision is based. Even with the data-reduction techniques, the trigger primitives contain sufficient information regarding the 3D development of the particle shower to efficiently identify e/γ candidates and reject clusters originating from pileup.

The XGBoost software package [196] is used for BDT training, where the input data are simulated HGCAL trigger primitives (3D clusters). Input features are the five longitudinal and four lateral shower-shape variables listed in Table 8.1. The energy-weighted RMS features are defined for generic trigger cell co-ordinate, p, as,

$$\text{Weighted RMS}(p) = \sqrt{\frac{1}{E_{\text{tot}}} \sum_{i}^{N_{\text{tc}}} E_i (p_i - \langle p \rangle)^2} \tag{8.1}$$

where the sum is over a collection of trigger cells, each with energy E_i, and co-ordinate p_i. The quantity $\langle p \rangle$ is the energy-weighted mean of p over the whole collection, whilst $E_{\text{tot}} = \sum_i^{N_{\text{tc}}} E_i$. Features of this type give an indication of the shower spread in the p co-ordinate direction.

In the training, signal clusters are identified as those consistent with originating from a truth-level electron of $p_T > 20$ GeV,[1] where the cluster is required to pass a minimum-p_T threshold of 10 GeV. Note, only electron clusters are required for training the e/γ identifier since both photon and electron showers have almost identical features in the HGCAL. Background clusters (pileup) are all clusters with $p_T > 20$ GeV that are not matched to a truth-level electron.

Two separate BDTs are trained in the pseudorapidity regions $1.5 < |\eta| < 2.7$ and $2.7 < |\eta| < 3.0$, to account for the fact that the e/γ shower-shape features evolve rapidly as a function of η. This improves the overall background rejection with respect to training a single BDT inclusive in η, particularly in the high $|\eta|$ region. Figure 8.2 shows the Max layer (longitudinal) and Weighted RMS(η) (lateral) distributions for signal and background clusters in each η region. Both features show good discriminating power. The Max layer distribution demonstrates that most e/γ showers deposit their maximum energy in the CE-E compartment (first 28 layers, where only alternate layers contribute to the trigger primitives), whilst clusters originating from pileup jets are more likely to deposit their maximum energy in the CE-H compartment (back 24 layers). The distributions of all input features are shown in Appendix J.

[1] This is done by requiring the reconstructed cluster to be within an angular separation of $\Delta R < 0.2$ with a truth-level electron.

Table 8.1: Input features to the HGCAL L1T e/γ identification BDT.

Longitudinal shower shape variables	
Weighted RMS(z)	Energy weighted RMS of trigger cell z co-ordinate, evaluated over the whole cluster. Measure of the longitudinal spread of the shower.
First layer	First layer of the HGCAL with an energy deposit (above the trigger cell threshold).
Max layer	Layer of the HGCAL with maximum cluster energy deposit.
Shower length	Total length of the cluster calculated as the difference between the first layer and the last layer with an energy deposit.
Core shower length	Maximum number of consecutive layers with energy deposits in the cluster.
Lateral shower shape variables	
Weighted RMS(r)	Energy weighted RMS of trigger cell r co-ordinate, evaluated over the whole cluster. The r co-ordinate is divided by the z co-ordinate in the calculation to account for the spreading out of the shower as it propagates through the HGCAL. Measure of the radial spread of the shower.
Mean layer weighted RMS(r)	Energy weighted RMS of trigger cell r co-ordinate, evaluated for each layer separately, and averaged over the whole cluster. Again, the r co-ordinate is divided by the z co-ordinate to account for the spreading out of the shower. Measure of the radial spread of the shower.
Weighted RMS(η)	Energy weighted RMS of trigger cell η co-ordinate, evaluated over the whole cluster. Measure of the polar angle spread of the shower.
Weighted RMS(ϕ)	Energy weighted RMS of trigger cell ϕ co-ordinate, evaluated over the whole cluster. Measure of the azimuthal spread of the shower.

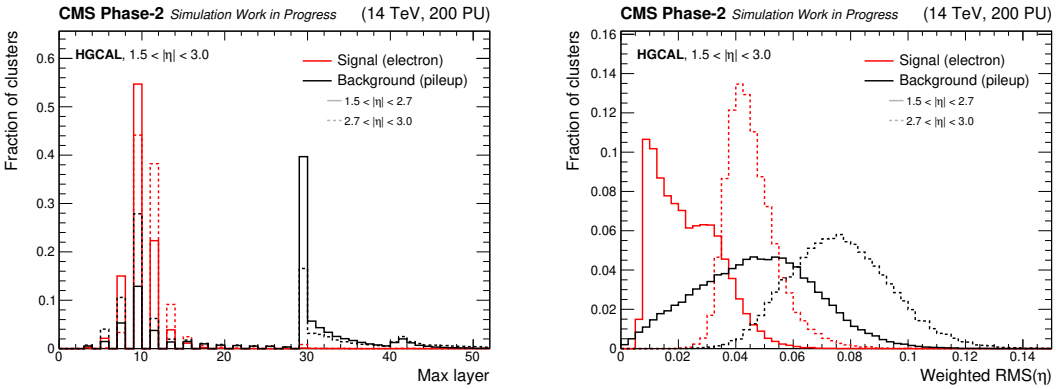


Figure 8.2: The Max layer (left) and Weighted RMS(η) (right) distributions for signal and background clusters, separated into the two pseudorapidity regions.

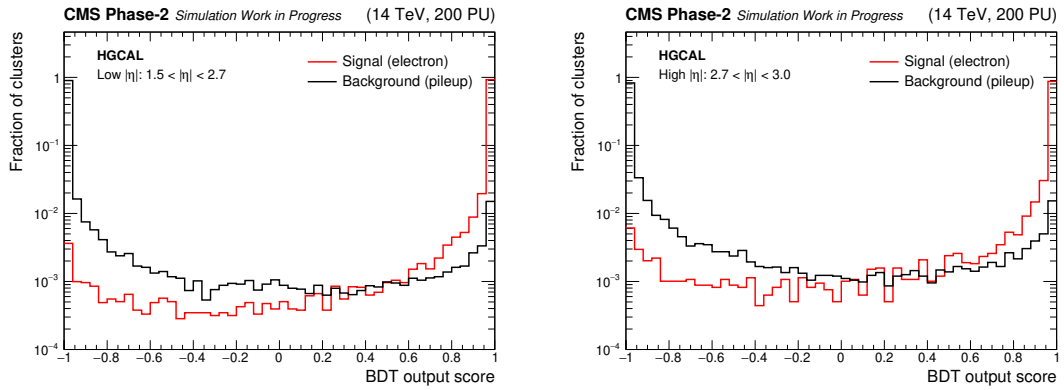


Figure 8.3: BDT output-score distributions for signal and background clusters. The BDT trained in the low $|\eta|$ region ($1.5 < |\eta| < 2.7$) and the BDT trained in the high $|\eta|$ region ($2.7 < |\eta| < 3.0$) are shown in the left and right plots, respectively. The outputs show excellent discrimination between signal and background clusters.

8.3.1 Performance

The output-score distributions of the two BDTs are shown for signal and background clusters in Figure 8.3; the scores are effectively a measure of how signal-like (1) or background-like (-1) the clusters are based on the input shower-shape features, $\vec{x}$.

The performance of the classifier is evaluated using the area under the receiver operating characteristic (ROC) curve [197]. Each point in the ROC curve corresponds to the signal efficiency and background rejection evaluated at a given threshold on the BDT output score. Here, the signal efficiency is defined as the fraction of truth-matched electron clusters above the BDT output-score threshold, whilst the background rejection is defined as the fraction of pileup clusters rejected at the same threshold. The ROC curves, evaluated using an independent test sample, are shown for both BDTs in Figure 8.4. The performance is shown to be slightly better for the low η region.

Baseline thresholds on the output scores (working points) are chosen for an inclusive signal efficiency of 97.5% in the $1.5 < |\eta| < 2.7$ region and 90.0% in the $2.7 < |\eta| < 3.0$ region. These correspond to background rejections of 96.7% and 97.3%, respectively. The tighter working point for high $|\eta|$ is chosen to combat the increased levels of pileup in this region. Ultimately, the excellent discriminating power is a result of the highly segmented design of the HGCAL, which provides a powerful handle on the lateral and longitudinal development of particle showers.

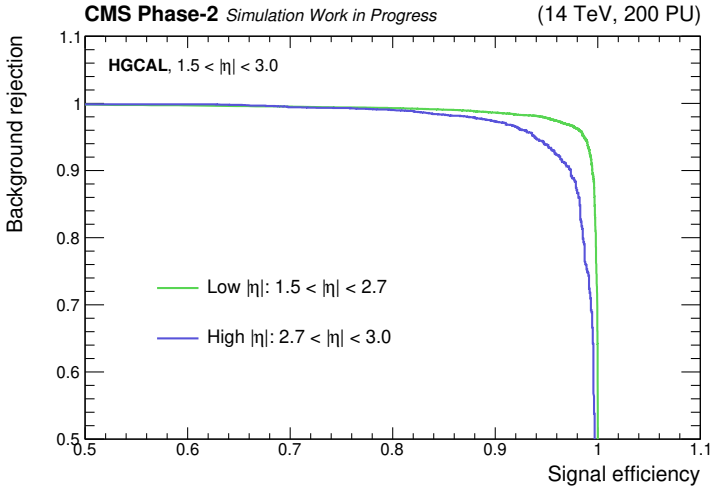


Figure 8.4: ROC curves for the e/γ identification BDTs, trained in the low $|\eta|$ region ($1.5 < |\eta| < 2.7$, green) and the high $|\eta|$ region ($2.7 < |\eta| < 3.0$, blue).

The trigger efficiency of the algorithm is shown as a function of the truth-level electron p_T and $|\eta|$ in Figure 8.5. This efficiency is defined as the fraction of truth-level electrons with $p_T > 30$ GeV that:

- have a matching trigger primitive cluster separated by an angle $\Delta R < 0.2$ with respect to the truth-level electron, where the cluster is required to have a reconstructed $p_T > 20$ GeV;
- *and* pass the aforementioned working points on the e/γ identification BDT.

In the plots, the grey lines indicate the fraction of electrons with a matching cluster; this is practically 100% for all pseudorapidity bins, except the two at the HGCAL edges where the electron can fall outside of acceptance. The blue lines then indicate the efficiency after applying the e/γ identification working points. It is shown to increase as a function of electron p_T, rising from 94% at $p_T = 30$ GeV to around 99% for $p_T = 100$ GeV. Also, the efficiency is shown to decrease with increasing electron $|\eta|$, barring the first pseudorapidity bin. This is especially noticeable in the high-$|\eta|$ region ($2.7 < |\eta| < 3.0$) where a tighter working point is applied on the BDT output score.

Clusters passing the e/γ identification BDT working point are subsequently promoted to calorimeter-only e/γ candidates. Those in the acceptance region of the tracker ($|\eta| < 2.4$) are combined with track finder trigger primitives to build track-matched objects for electrons and isolated showers (without a matching track) for photons. The objects then enter the L1T decision process, where an *accept signal* is sent to the detector read-out electronics if the event is deemed to be of interest [3].

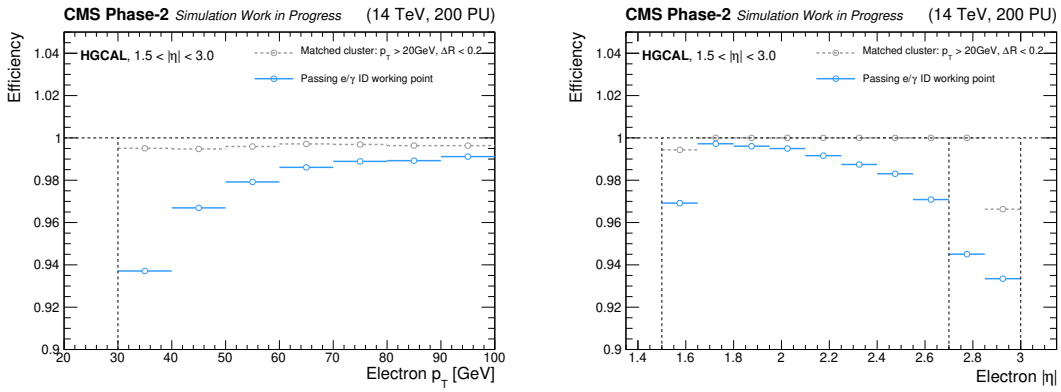


Figure 8.5: Trigger efficiency as a function of the truth-level electron p_T (left) and η (right). The grey lines indicate the fraction of electrons that contain a matching cluster with reconstructed $p_T > 20$ GeV, and $\Delta R < 0.2$. The blue lines indicate the fraction of those electrons where the cluster passes the e/γ identification working points.

The BDT algorithm described in this section was developed in offline software. In practice, the L1T operates in real time during data-taking (online) and therefore the algorithm must be implemented in firmware. Recent advances in FPGA technology have enabled the implementation of this particular e/γ identification algorithm in firmware, using the HLS4ML library [194]. Crucially, when developing such algorithms it is essential that the resources needed for running are consistent with the design constraints of the Phase-2 L1T. Studies looking at more complex algorithms for the e/γ identification, such as neural networks, have been discussed. To be successful, it must be feasible to implement these algorithms in firmware, and the resources needed for running must be compatible with the constraints of the CMS Phase-2 L1T architecture.

8.4 Higgs boson physics at the HL-LHC

The HL-LHC offers a wide and diverse physics programme over the coming decades [198, 199]. In particular, the potential gains from using the HL-LHC data set for Higgs boson physics are striking [4]. For example, the combination of future ATLAS and CMS measurements with 3 ab^{-1} of p-p collision data is expected to achieve uncertainties $\mathcal{O}$(1.5–2.5%) in the Higgs boson couplings to vector bosons, and $\mathcal{O}$(2–4%) in the Higgs boson couplings to third-generation fermions; where the dominant component of the uncertainties in all cases arises from the theoretical-prediction projections. In comparison, the current best measurements from the CMS experiment are $\mathcal{O}$(8–11%) and $\mathcal{O}$(10–17%), respectively [36]. Moreover, the increased data-set will enable more differential measurements, probing increasingly granular regions of the Higgs boson phase space, and will shed light on rarer Higgs boson interactions including the Higgs boson couplings to

second-generation fermions and the Higgs boson self-coupling. Altogether, the Higgs boson physics programme at the HL-LHC will go a long way towards elucidating the origins of electroweak symmetry breaking.

The analysis described in this section extracts the expected sensitivity of differential p_T^H cross section measurements for Higgs boson production in association with at least one top quark, with the Higgs boson decaying to photons (ttH + tH, H $\rightarrow \gamma\gamma$), using the CMS Phase-2 detector at the HL-LHC [48]. It is important to keep in mind that the observation of the ttH production mode was only made in 2018, using a combination of Higgs boson decay modes [30, 33]. The results presented here show that by the mid-2030s, we will be able to measure this production mode *differentially*, demonstrating impressive sensitivity ($\mathcal{O}$(15–40%) uncertainties in different p_T^H bins) in a single Higgs boson decay channel. These measurements can be used to indirectly constrain the trilinear Higgs boson self-coupling (λ_3). Variations in $\kappa_\lambda = \lambda_3/\lambda_3^{\mathrm{SM}}$ from unity affects the values of the differential cross sections due to NLO corrections in electroweak theory. The expected constraints on λ_3 from this indirect approach are determined. The analysis uses many of the same techniques as the H $\rightarrow \gamma\gamma$ analysis described in Chapters 4–6.

8.4.1 Top-associated differential p_T^H cross sections

Signal and background events are simulated with $\sqrt{s} = 14$ TeV using a combination of the MG5_aMC@NLO (version 2.2.2) [119], Powheg (version 2.0) [120–125], and sherpa (version 2.2.5) [126] generators, interfaced with Pythia8 (version 8.205) [127] for parton showering and hadronisation. The events are subsequently propagated through the delphes framework [200] to perform a *fast simulation* of the CMS Phase-2 detector response under HL-LHC conditions. This works by parametrising the detector efficiency and resolution of the various upgraded Phase-2 subdetectors as a function of the different final-state objects properties (e.g. p_T, η), where the exact forms of these parametrisations have been derived using detailed simulations [192]. The outputs of delphes are then collections of jets, b tagged jets, photons, charged leptons and p_T^{miss} for each generated event, which approximately match the expected performance of the CMS Phase-2 detector. All samples are normalised to the expected yields at an integrated luminosity of 3 ab^{-1}.

Events are required to contain two photons with $|\eta^\gamma| < 2.5$, excluding the barrel-endcap transition region ($1.44 < |\eta^\gamma| < 1.57$), with a diphoton invariant mass satisfying $100 < m_{\gamma\gamma} < 180$ GeV. The leading (sub-leading) photon is also required to have $p_T^\gamma/m_{\gamma\gamma} > 1/3$ (1/4). Additionally, the two photons are required to have an angular separation, $\Delta R_{\gamma\gamma} > 0.4$, and each photon must satisfy an isolation requirement which demands the sum of charged particle p_T in a cone of radius $\Delta R_\gamma = 0.4$, centred on the photon direction, is less than 30% of the photon p_T^γ. For events with multiple photon

pairs passing this selection, the pair with $m_{\gamma\gamma}$ closest to the Higgs boson mass are chosen.

Top quarks almost always decay to a W boson and a bottom quark. Therefore, to isolate events consistent with Higgs boson production in association with top quarks, all events are required to contain at least one b tagged jet (see Section 4.3.3). Events are then separated into two orthogonal global categories depending on the decay products of the W boson: a hadronic global category (W→qq) and a leptonic global category (W→ $\ell\nu$). In the hadronic selection, events are required to contain at least three jets, clustered using the anti-k_T algorithm [68, 69] with a distance parameter of $R = 0.4$, where each jet must satisfy $p_T^j > 30$ GeV and $|\eta^j| < 4$, and be separated by $\Delta R_{j,\gamma} > 0.4$ with respect to both photon candidates. The leptonic selection requires at least two jets, in addition to at least one isolated muon or electron. The muon or electron must satisfy $p_T^\ell > 20$ GeV and $|\eta^\ell| < 2.4$, excluding the barrel-endcap transition region for electrons. Muons are required to pass an isolation criteria, such that the sum of all particles p_T in a cone of radius $\Delta R_\mu = 0.4$, centred on the muon direction, is less than 25% of the muon p_T^μ. For electrons, the invariant mass of pairs formed from the electron and either photon, $m_{e\gamma}$, is required to be greater than 5 GeV from the nominal Z boson mass to reduce the contamination from $Z \to e^+e^-$ decays. Events passing the leptonic selection are excluded from the hadronic selection to ensure the two categories are orthogonal.

To improve the signal-vs-background discrimination, a BDT is trained independently for each global category using events passing the aforementioned selection criteria. The input features are the photon, jet and lepton p_T and η values, the photon isolation variables, the p_T^{miss}, the scalar sum of all final state objects p_T (mitigating the effects of pileup), the azimuthal separation between the photon pair and the closest jet/leading lepton, and the total number of jets, b tagged jets, and leptons in the event. The BDT output score distributions for the hadronic and leptonic global categories are shown in Figure 8.6. The plots indicate the important background processes for this study, and also show the contributions from other Higgs boson production modes: ggH + VH. The contamination from VBF production is negligible, and is therefore ignored in this analysis.

Table 8.2 shows the bin boundaries for which the differential p_T^H cross sections are measured. The hadronic and leptonic global categories are split according to equivalent boundaries in the reconstructed diphoton transverse momentum, $p_T^{\gamma\gamma}$. Events in each $p_T^{\gamma\gamma}$ bin are required to have BDT output score values greater than fixed thresholds[2], shown by the dashed lines in Figure 8.6. In the hadronic channel, the five bins with $p_T^{\gamma\gamma} < 350$ GeV are further split into low signal-purity and high signal-purity regions according to a second threshold on the BDT output score at a value of 0.61. This helps reduce the contamination from ggH production. In total, this corresponds to 17 analysis categories targeting the

[2]The position of these thresholds are chosen to maximise the sensitivity to κ_λ.

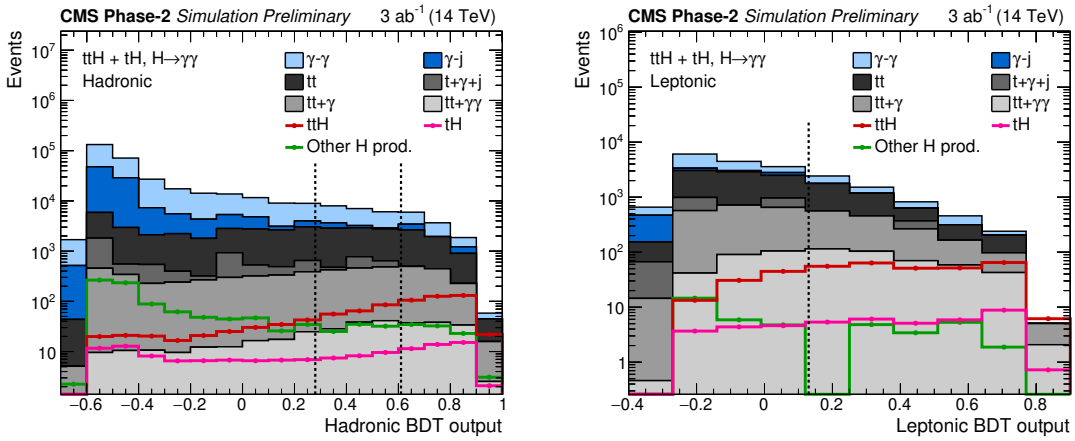


Figure 8.6: The BDT output score distributions for the hadronic (left) and leptonic (right) categories, after the preselection criteria are applied. The background processes are shown by the filled histograms, whilst the Higgs boson production modes are shown by the coloured lines. The dashed vertical lines indicate the positions of the BDT output score thresholds in the event categorisation.

Table 8.2: Bin boundaries for which the differential p_T^H cross sections are measured. To target these bins, the hadronic and leptonic categories are sub-divided by equivalent boundaries on the reconstructed $p_T^{\gamma\gamma}$.

p_T^H or $p_T^{\gamma\gamma}$ bin boundaries [GeV]						
0	45	80	120	200	350	∞

six p_T^H bins with different requirements on the BDT output scores: 11 for the hadronic channel and six for the leptonic.

The cross sections are extracted using a simultaneous binned maximum-likelihood fit to the $m_{\gamma\gamma}$ distribution in all analysis categories. The signal models are built for each production mode using a sum of Gaussian functions to fit the $m_{\gamma\gamma}$ peak. In order to account for detector resolution effects, a separate model is constructed for events from each truth-level p_T^H bin in each reconstruction-level $p_T^{\gamma\gamma}$ event category. The background models are a set of smoothly falling functions to fit the sum of simulated background events in each event category, where the choice of function is left free to vary in the likelihood fit. This procedure, known as the discrete profiling method [147], was previously described in more detail in Section 5.3. The final signal-plus-background models are shown for two example analysis categories in Figure 8.7. The black points, shown purely for illustration purposes, represent a possible HL-LHC data set and are extracted by throwing random toy data from the signal-plus-background model. The diphoton mass resolution corresponds to what is expected to be achieved during the HL-LHC operation.

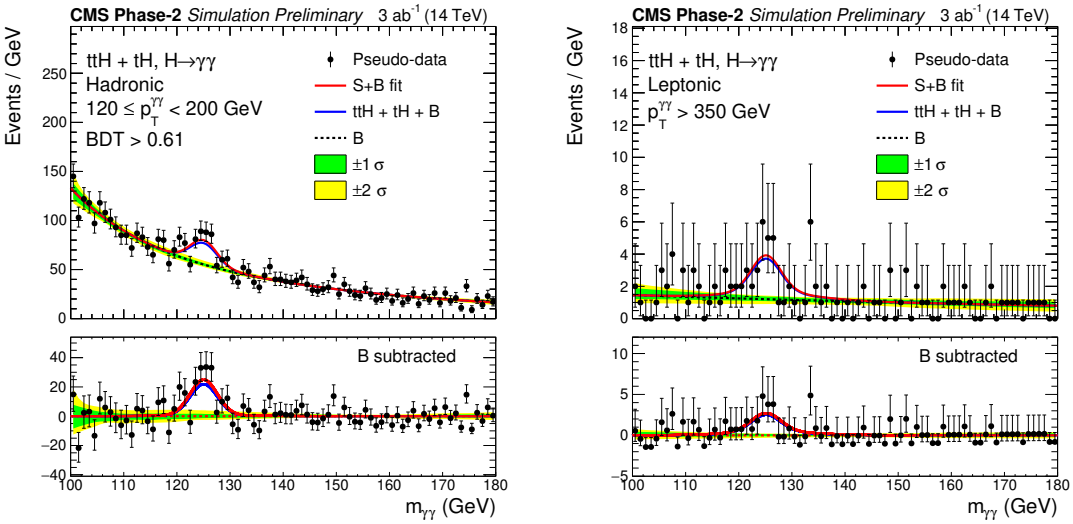


Figure 8.7: Best-fit signal-plus-background models for the high purity $120 < p_T^{\gamma\gamma} < 200$ GeV hadronic event category (left) and the $p_T^{\gamma\gamma} > 350$ GeV leptonic event category (right). An illustrative pseudo-data set is thrown from the best-fit models. The one (green) and two (yellow) standard deviation bands show the uncertainties in the background component of the fit. The residuals minus the background component are shown in the lower panels.

A likelihood function is constructed for each analysis category following the procedure introduced in Section 5.1.1. This uses the corresponding signal and background models, and an Asimov data set [146]. The parameters of interest, $\mu^{i,\gamma\gamma}$, are defined to scale the ttH + tH production cross section for each truth-level p_T^H bin, i. Defining the parameters in this way enables a likelihood unfolding of the detector resolution effects i.e. the fit accounts for the migrations between the truth-level p_T^H and reconstruction-level $p_T^{\gamma\gamma}$ bins. As this study concerns the expected sensitivity, the Asimov data set corresponds to the SM prediction (all $\mu^{i,\gamma\gamma} = 1$). The product over all per-category likelihoods is used to construct a profiled likelihood-ratio test-statistic to determine the expected uncertainties in each $\mu^{i,\gamma\gamma}$; a procedure described in detail in Section 5.1.2.

Systematic uncertainties affecting the signal-yield estimates are included as nuisance parameters in the likelihood function. Experimental uncertainties originating from the reconstruction and identification efficiencies for photons and b jets, as well as the energy scale and resolution of jets, are modelled as log-normal variations in the signal yields (see equation 5.14). Theoretical uncertainties which cause the migration of signal events between event categories are calculated using renormalisation and factorisation scale variations, whilst keeping the normalisation of each p_T^H bin constant. Additionally, theoretical uncertainties which modify the overall rates of ggH and VH production are included, as these production modes are not explicitly extracted in the fit. Parameters of the back-

ground model functions are free to vary in the fit, and are therefore constrained directly from data. This means the uncertainties in the background estimation are statistical in nature.

The $\mu^{i,\gamma\gamma}$ parameters and their uncertainties are converted to fiducial cross sections times branching fraction, $\sigma_{\text{fid}}^{\text{ttH+tH}} \cdot \mathcal{B}^{\gamma\gamma}$, by correcting for the event selection efficiencies. The fiducial region is common to both the hadronic and leptonic selections, and is defined according to the truth-level events as follows:

- Higgs boson rapidity: $|Y_H| < 2.5$.
- Two photons from the Higgs boson decay: $p_T^{\gamma} > 20$ GeV and $|\eta^{\gamma}| < 2.5$.
- At least two jets: $p_T^{\gamma} > 25$ GeV and $|\eta^{j}| < 4$.
- At least one of the jets, satisfying the above criteria, originates from a b quark.

A small fraction of the events passing the full selection (0.7% in the hadronic selection, and 0.4% in the leptonic selection) are not contained in the fiducial region. Although these events are included in the construction of the likelihood, they are subtracted when calculating the fiducial cross sections.

Figure 8.8 shows the expected differential fiducial cross sections times branching fraction, for Higgs boson production in association with at least one top quark, in bins of p_T^H. The error bars indicate the combined statistical and systematic uncertainties in the measurements using 3 ab^{-1} of HL-LHC data. Analogous likelihood fits are performed using only the hadronic event categories and only the leptonic event categories, shown by the red and purple error bars, respectively. In general, the hadronic channel is observed to provide greater sensitivity. This is a result of the larger absolute signal yield after selection, compared to the leptonic channel. The theoretical uncertainties in the predicted ttH + tH cross sections, displayed by the yellow boxes in the plot, include the projected uncertainties in the renormalisation and factorisation scales, PDFs and α_s. Additionally, the boxes include the uncertainty in the shape of the ttH + tH p_T^H spectrum, estimated by modifying the renormalisation and factorisation scales up and down by a factor of 2 and factoring out the inclusive effect.

The expected sensitivities are summarised in Table 8.3. Many extensions to the SM predict modifications to the Higgs boson interaction with the top quark. By measuring the differential cross sections within uncertainties $\mathcal{O}$(15–40%), and therefore gaining a handle on the kinematic spectrum of top-associated production, we will be able to tightly constrain potential new physics affecting the top-Higgs sector. One such example concerning anomalous values of the Higgs boson self-coupling, λ_3, is provided below.

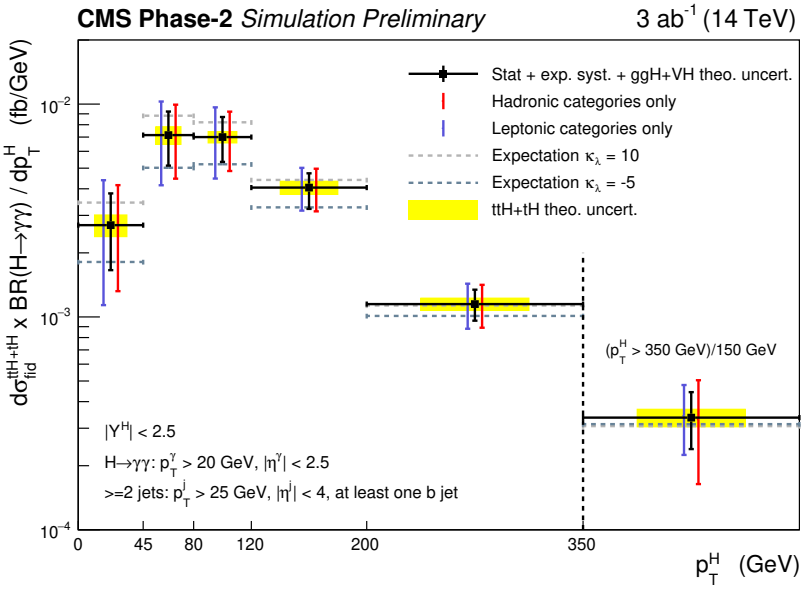


Figure 8.8: Expected ttH + tH differential cross sections times branching fraction, in bins of p_T^H, for 3 ab^{-1} of HL-LHC data. These are for the fiducial region of phase space defined in the bottom left of the plot. The error bars include the statistical, experimental systematic and ggH + VH theoretical systematic uncertainties. The theoretical uncertainties on the ttH + tH cross section predictions, originating from the uncertainty in the renormalisaion and factorisation scales, are shown by the shaded yellow boxes. The sensitivities extracted from the hadronic and leptonic categories alone, are indicated by the red and purple error bars respectively. The cross section for the $p_T^H = [350, \infty]$ GeV bin is scaled by the width of the previous bin. Additionally, the expected cross sections for anomalous values of the Higgs boson self-coupling ($\kappa_\lambda = 10$ and $\kappa_\lambda = -5$) are shown by the horizontal dashed lines.

Table 8.3: Expected uncertainties in the ttH + tH differential p_T^H fiducial cross sections times branching fraction for 3 ab^{-1} of data collected at the HL-LHC. The uncertainty is decomposed into the statistical and systematic components.

p_T^H bin	Expected $\pm1\sigma$ uncertainties		
	Total	Stat unc.	Syst unc.
$[0, 45]$	$^{+41\%}_{-39\%}$	$^{+41\%}_{-39\%}$	$^{+4\%}_{-2\%}$
$[45, 80]$	$^{+29\%}_{-28\%}$	$^{+29\%}_{-28\%}$	$^{+3\%}_{-2\%}$
$[80, 120]$	$^{+24\%}_{-24\%}$	$^{+24\%}_{-24\%}$	$^{+3\%}_{-2\%}$
$[120, 200]$	$^{+17\%}_{-21\%}$	$^{+16\%}_{-20\%}$	$^{+3\%}_{-3\%}$
$[200, 350]$	$^{+17\%}_{-17\%}$	$^{+16\%}_{-16\%}$	$^{+5\%}_{-5\%}$
$[350, \infty]$	$^{+33\%}_{-30\%}$	$^{+30\%}_{-28\%}$	$^{+14\%}_{-13\%}$

8.4.2 Constraining κ_λ

Measurements of the trilinear self-interaction of the Higgs boson are of upmost priority in future physics programmes [4]; they provide constraints on the shape of the Higgs potential close to the minimum, and will shed light on the dynamics of electroweak-symmetry breaking, including the order of the electroweak phase transition [21, 22, 25, 26]. In the SM, the trilinear-coupling strength, $\lambda_3^{\mathrm{SM}} = \lambda = m_{\mathrm{H}}^2/2v^2$, is fixed according to the Higgs boson mass, m_{H}, and the vacuum expectation value, v. BSM physics, such as an extended scalar sector, can modify the value of λ_3 without affecting m_{H} and v.

The direct approach to constraining λ_3 is via searches for di-Higgs production (HH), which depends on λ_3 at LO. A number of HH final states have been explored by ATLAS and CMS at $\sqrt{s} = 13$ TeV [201, 202]. The current best constraints on λ_3 come from the full Run 2 CMS HH $\to$ bb$\gamma\gamma$ analysis [203], which excludes $\kappa_\lambda = \lambda_3/\lambda_3^{\mathrm{SM}}$ values outside of the range $-3.3 < \kappa_\lambda < 8.5$ at the 95% confidence level. Despite this impressive result, HH production is not expected to be observed at 5σ until after the HL-LHC operation [4]. This is due to the small SM cross section ($31.1^{+1.4}_{-2.0}$ fb at $\sqrt{s} = 13$ TeV), which suffers from destructive interference amongst diagrams [204]. Consequently, alternative strategies for probing λ_3 are in high demand.

One such approach is to exploit radiative corrections to inclusive and differential single-Higgs boson production-rates [205–209]. At NLO in electroweak theory, single-Higgs boson production includes diagrams with the trilinear self-interaction, such as that shown in Figure 8.9. The effects of a modified λ_3 are sizeable for Higgs boson production in association with top quarks (ttH and tH) or a vector boson (VH). This is due to the large mass of the associated particles providing a larger coupling to the virtual Higgs boson. Moreover, the deformations to the Higgs boson rates are shown to have a kinematic dependence on λ_3 [206, 209]. As a result, differential cross section measurements can disentangle the effects of a modified λ_3 from other effects such as the presence of an anomalous top-Higgs coupling. Altogether, these features mean the ttH + tH differential cross section measurements introduced in the previous section provide an excellent candidate for indirectly probing λ_3.

The effect of anomalous $\kappa_\lambda = \lambda_3/\lambda_3^{\mathrm{SM}}$ values on the single-Higgs boson production cross sections and decay widths have been predicted [206]. The cross section is parametrised as a function of κ_λ according to the following function,

$$\mu(\kappa_\lambda, C_1) = \frac{\sigma}{\sigma_{\mathrm{SM}}} = \frac{1 + \kappa_\lambda C_1 + \delta Z_H}{(1 - (\kappa_\lambda^2 - 1)\delta Z_H)(1 + C_1 + \delta Z_H)}. \tag{8.2}$$

The $\delta Z_H = -1.536 \times 10^{-3}$ component originates from the Higgs boson wave-function renormalisation and is universal to all production modes. The C_1 parameter is defined as

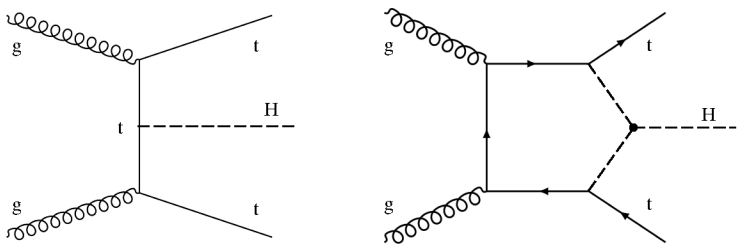


Figure 8.9: Feynman diagrams for ttH production at LO (left) and a λ_3-dependent correction at NLO (right).

the interference between the LO Born matrix element, $\mathcal{M}^{\mathrm{LO}}$, and the virtual λ_3-dependent SM matrix element at one-loop, $\mathcal{M}^{\mathrm{NLO}}_{\lambda_3^{\mathrm{SM}}}$,

$$C_1(\{p\}) = \frac{2\mathrm{Re}(\mathcal{M}^{\mathrm{LO}*}\mathcal{M}^{\mathrm{NLO}}_{\lambda_3^{\mathrm{SM}}})}{|\mathcal{M}^{\mathrm{LO}}|^2}. \tag{8.3}$$

Crucially, C_1, depends both on the Higgs boson production mode, and on some final-state observable, p. The C_1 factors relevant for this analysis are derived using the electroweak reweighting tool described in Ref. [210]. Leading order (LO) parton-level ttH, tH and VH events are simulated using the MG5_aMC@NLO (version 2.5.5) generator [119]. The tool then calculates λ_3-dependent corrections at NLO ($\mathcal{O}(\lambda_3)$) by reweighting events on an event-by-event basis. A diagram filter is applied to select only the relevant one-loop matrix elements which feature the trilinear coupling. The C_1 factors are then extracted by taking the ratio of the $\mathcal{O}(\lambda_3)$ to LO contributions in bins of the truth-level p_T^H spectrum. These C_1 factors are plugged into equation 8.2 to determine the differential cross section scaling functions. It should be noted that this parametrisation relies on the assumption that higher-order QCD effects and other NLO EW contributions factorise from the anomalous λ_3 effects. The validity of this assumption has been studied in detail in Ref. [206].

The reweighting tool does not accommodate ggH production, due to the presence of the ggH loop at LO. For this reason, a common scaling function is used for ggH production in all p_T^H bins, where the inclusive value of $C_1 = 0.0066$ is taken directly from Ref. [205]; this is small compared to the C_1 values for the ttH and tH production modes. Additionally, there is a small correction to the H $\rightarrow \gamma\gamma$ decay rate from anomalous κ_λ values which is also taken from Ref. [205]. The final scaling functions, $\mu^{i,\gamma\gamma}(\kappa_\lambda)$, for each Higgs boson production mode, split into the different truth-level p_T^H bins, are shown in Figure 8.10.

The κ_λ dependence is largest (largest C_1 values) for ttH production at threshold (low-p_T^H). For example, a 20% enhancement to the ttH production rate for $p_T^H \in [0, 45]$ GeV is predicted for $\kappa_\lambda \sim 10$. The horizontal dashed lines in Figure 8.8 correspond to the predicted values of the differential cross sections for anomalous values of the Higgs boson

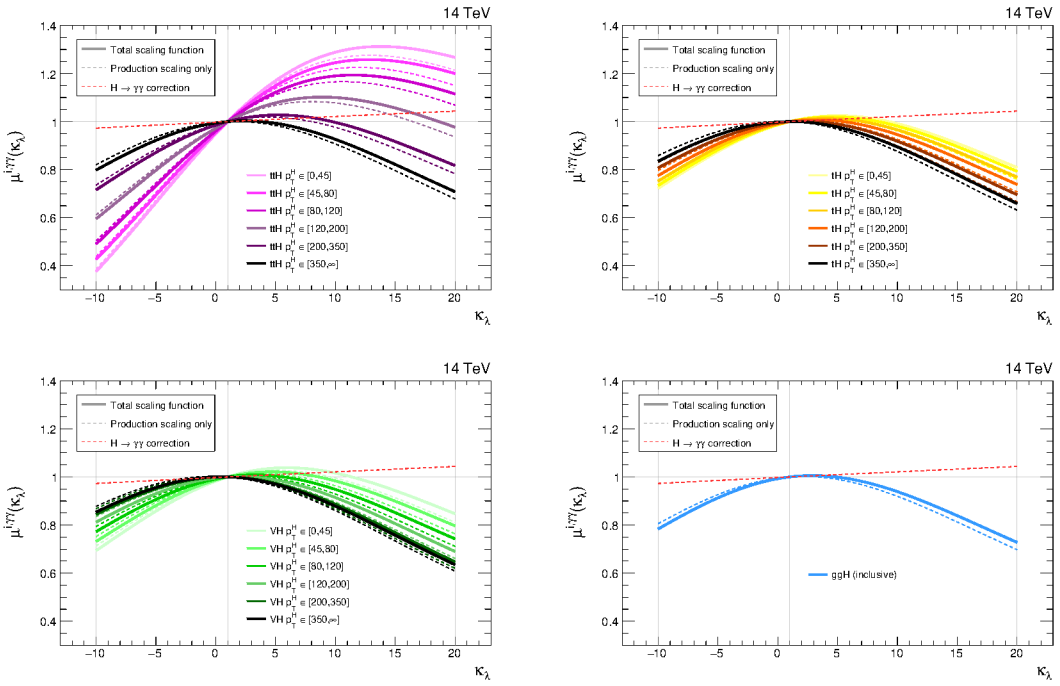


Figure 8.10: Scaling functions, $\mu^{i,\gamma\gamma}(\kappa_\lambda)$, for the differential cross sections times branching fraction targeted in this analysis. These are split into the different Higgs boson production modes. The ggH production mode uses an inclusive scaling function. In all plots, a lighter colour is used to represent lower p_T^H bins. The total scaling functions including the correction from the H $\rightarrow$ $\gamma\gamma$ decay rate (red dashed line) are shown by the solid lines, whereas the affect on the production cross sections alone are shown by the dashed lines. All functions are plotted for the region of validity: $\kappa_\lambda \in [-10, 20]$.

self-coupling: $\kappa_\lambda = 10$ and $\kappa_\lambda = -5$. It can be seen that the expected uncertainties in the differential cross section measurements will enable κ_λ to be constrained roughly between these values.

In order to extract the sensitivity to κ_λ, the substitution $\mu^{i,\gamma\gamma} \rightarrow \mu^{i,\gamma\gamma}(\kappa_\lambda)$ is made in the construction of the likelihood function, where $\mu^{i,\gamma\gamma}(\kappa_\lambda)$ are directly the scaling functions shown in Figure 8.10. As this represents an interpretation of cross section measurements, the theoretical uncertainties in the ttH + tH predictions ($\vec{\theta}_s^{\,\mathrm{th}}$) are directly folded into the measurement.

A scan of the profiled likelihood as a function of κ_λ ($\equiv \alpha$) is shown in Figure 8.11. The scan is performed in the region $\kappa_\lambda \in [-10, 20]$, beyond which the parametrisation becomes invalid as next-to-next-to-leading order (NNLO) effects become important. Additional likelihood scans are performed when only including the hadronic and leptonic analysis categories, shown in red and purple, respectively. Clearly, both channels con-

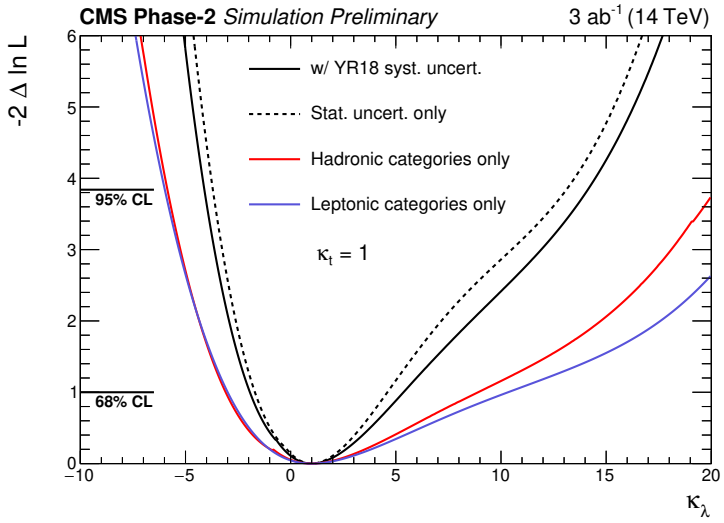


Figure 8.11: The $q(\kappa_\lambda) = -2\Delta \ln L$ likelihood curve for the κ_λ fit. The individual contributions of the statistical and systematic uncertainties are separated by performing a likelihood scan with all systematics removed. The observed deviation is dominated by the theoretical uncertainties in the Higgs boson cross section predictions. Additionally, the contributions from the hadronic and leptonic global categories have been separated, shown in red and purple, respectively.

tribute significantly towards the final sensitivity. The constraints are tighter for negative values of κ_λ since larger deviations in the ttH + tH differential cross sections are predicted, compared to positive values. The feature in the region around $5 < \kappa_\lambda < 15$ is a result of the turning points in the ttH scaling functions, which introduce a degeneracy into the parametrisation. This degeneracy is somewhat alleviated by the contamination of ggH in the signal model, which has a different scaling behaviour. Ultimately, the scan shows that with 3 ab^{-1} of HL-LHC data, we can expect to exclude κ_λ values outside of the range $-4.1 < \kappa_\lambda < 14.1$ at the 95% confidence level using ttH + tH differential measurements in the H $\rightarrow \gamma\gamma$ decay channel.

An additional fit is performed in which an overall normalisation parameter for Higgs boson signal processes, μ_H, is profiled. This parameter incorporates other BSM effects, such as an anomalous top-Higgs coupling, which in general cause an inclusive shift across the whole p_T^H spectrum. Figure 8.12 shows the results of a two-dimensional likelihood fit in the (κ_λ, μ_H)-plane, in terms of the 68% and 95% confidence-level contours. It can be seen that differential cross section measurements still provide sensitivity to κ_λ, without exploiting the overall normalisation of the p_T^H spectrum. In other words, the shape of the spectrum is used to constrain κ_λ.

All in all, this analysis indicates that additional sensitivity to the Higgs boson self-coupling is available through differential cross section measurements of single-Higgs boson

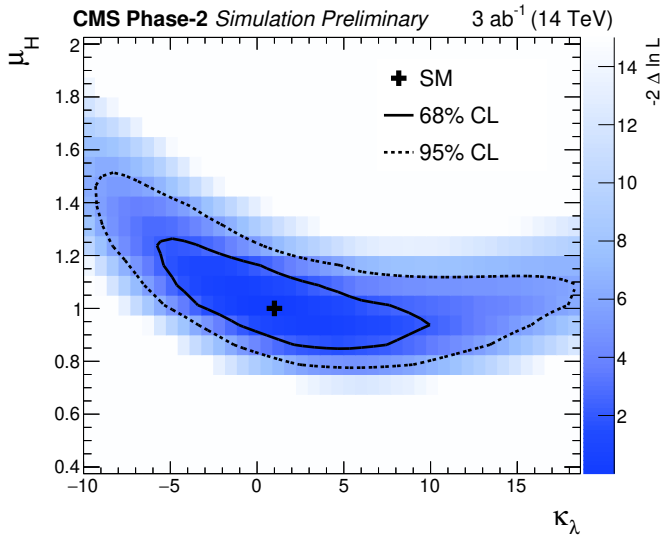


Figure 8.12: The two-dimensional $q(\kappa_\lambda, \mu_H) = -2\Delta \ln L$ likelihood surface, where μ_H is an inclusive scaling parameter for all Higgs boson production modes. The SM expectation, 68% confidence level contour and 95% confidence level contour are shown by the black cross, solid line, and dashed line, respectively.

production in association with top quarks. The expected sensitivity represents what could be achieved with 3 ab^{-1} of HL-LHC data, in a *single* Higgs boson decay channel ($H \to \gamma\gamma$). It should be acknowledged that this projected constraint on κ_λ is weaker than the current best constraint from the CMS HH $\to$ bb$\gamma\gamma$ analysis [203]. Nevertheless, it is imperative to investigate all methods for probing the Higgs boson self-coupling at the HL-LHC, as the ultimate sensitivity will be achieved from a combination of similar indirect analyses targeting the other Higgs boson decay channels and production modes, and with direct searches for HH production. The potential of these so-called *global fits* are studied in detail in Ref. [209], where the addition of single-Higgs measurements are particularly important for disentangling a modified κ_λ from deviations in other Higgs boson couplings.

8.5 Summary

The HL-LHC is a future operation of the LHC machine that will operate with instantaneous luminosities exceeding five times the nominal design value. The data-taking phase of the HL-LHC is scheduled to begin in 2027 and will run for at least a decade, during which the LHC experiments will collect a huge amount of p-p collision data. This chapter has provided an insight into the HL-LHC project, particularly focusing on the areas of research that the author has been involved in. Firstly, the HL-LHC project was motivated

and the experimental challenges from the increased levels of pileup were described. The CMS experiment will undergo a series of upgrades to maintain an excellent performance during the HL-LHC era. The upgrades to the CMS endcap calorimeters and the L1T were discussed. Section 8.3 introduced a ML algorithm to identify electron and photon showers from pileup-induced showers in the HGCAL L1T. The remainder of the chapter was dedicated to the physics reach of the HL-LHC, in terms of Higgs boson measurements. The expected sensitivity to differential cross section measurements of Higgs boson production in association with top quarks was outlined. These measurements were then shown to be sensitive to the Higgs boson self-coupling via radiative NLO corrections in single Higgs boson production. Measurements of this type will be complimentary to searches for HH production when extracting the ultimate constraints on λ_3 at the HL-LHC.

Chapter 9

Conclusions

The SM of particle physics has proved to be an extremely successful theory in describing the fundamental constituents of matter and their interactions via the strong, weak, and electromagnetic forces. Nevertheless, there are a number of physical observations that the SM does not account for including neutrino oscillations, dark matter, and most notably, gravity. This makes necessary the existence of BSM physics. At high-energy physics experiments, precision measurements offer an indirect approach to search for BSM physics, since as-yet-undetected new states can modify the predictions of the SM. In particular, precision measurements of Higgs boson properties will help to elucidate the origins of electroweak-symmetry breaking, and perhaps point to an extension of the scalar sector in the SM. This thesis reports the precision measurements of Higgs boson properties by the CMS experiment, using LHC Run 2 data.

Chapters 4–6 detailed the measurements of Higgs boson production cross sections and couplings in the $H \rightarrow \gamma\gamma$ decay channel. The results are based on 137 fb^{-1} of p-p collision data at $\sqrt{s} = 13$ TeV, and are published in Ref. [1]. Events with two reconstructed photons consistent with the decay of a Higgs boson were selected, and subsequently categorised to target different kinematic regions of the STXS framework. This was performed using a sophisticated chain of ML algorithms in order to maximise the sensitivity. A statistical inference procedure was applied to extract the Higgs boson properties from the diphoton invariant mass distribution in each analysis category. This involves the modelling of signal events as a peak around the Higgs boson mass, background events as a smoothly falling distribution, and accounting for the associated systematic uncertainties. The results were extracted using a maximum-likelihood fit under different signal hypotheses, providing measurements of signal-strength modifiers, coupling-modifiers, and production cross sections in the STXS framework.

All observed measurements were consistent with the SM predictions within uncertainties. The inclusive Higgs boson signal-strength, relative to the SM prediction, was

measured to be 1.12 ± 0.09. Three measurements were performed within the STXS framework, in which 6, 17, and 27 independent kinematic regions were measured simultaneously. The latter demonstrates the most granular fit of Higgs boson production cross sections in a single decay channel to-date. Moreover, many of the kinematic regions were measured here for the first time, including the splitting of the ttH production mode into five different p_T^H regions. Ultimately, this divide-and-measure approach of the STXS framework enhances the sensitivity to BSM physics which affects particular regions of the production phase-space. One region of interest is ggH production with $p_T^H > 200$ GeV, due to the sizeable enhancement that would arise from potential new physics states appearing in the ggH loop. The measured ggH cross section with $p_T^H > 200$ GeV is compatible with the SM, with an observed value of $0.9^{+0.4}_{-0.3}$ relative to the SM prediction. Finally, an upper limit was placed on single-top-associated Higgs boson production for the first time using $\mathrm{H} \to \gamma\gamma$ measurements at CMS. The observed (expected) limit at the 95% confidence level was found to be 14 (8) times the SM prediction.

Chapter 7 reported the BSM interpretation of STXS measurements using an EFT approach, which has been made public by the CMS Collaboration in Ref. [2]. This approach benefits from being agnostic to the specifics of the BSM theory, such that the short-range UV physics is integrated out and modelled as effective contact interactions between the SM fields. In the interpretation, STXS measurements were combined across all major Higgs boson decay channels to provide the ultimate sensitivity and enable multiple EFT operators to be probed simultaneously. The Higgs Effective Lagrangian (HEL) was used to parametrise deviations in the Higgs boson cross sections and branching fractions as functions of the EFT Wilson coefficients. This parametrisation was applied in a maximum-likelihood fit to extract constraints on seven independent EFT parameters. All results were found to be compatible with the SM expectation, and the confidence intervals are amongst the tightest constraints placed on this subset of EFT operators, thereby reducing the parameter space for potential BSM physics.

The final chapter investigated the physics potential of the HL-LHC programme. It is quite remarkable to consider the huge advances that have been made in characterising the Higgs boson and its interactions, all within ten years of the particle's discovery. A prime example are the $\mathrm{H} \to \gamma\gamma$ results in this thesis, which demonstrate a precision of around 9% on the inclusive Higgs boson production rate. Nevertheless, for a vast majority of measurements (e.g. in the STXS framework) the uncertainties are dominated by a lack of statistics; this means the accumulation of more high-energy collision data is of paramount importance. At the HL-LHC, we will collect at least ten times the amount of data expected at the end of Run 3 of the LHC. In doing so, it becomes possible to constrain BSM physics that introduces per-cent level modifications to the Higgs boson couplings.

Moreover, the increased data set facilitates new measurements, such as the indirect probe of the Higgs boson self-coupling via top-associated differential cross section measurements, as shown in Chapter 8. The results demonstrated that additional sensitivity to the Higgs boson self-coupling is available in single-Higgs measurements, which will become more apparent when combining with measurements from other Higgs boson production modes and decay channels. This analysis was published in a collection of HL-LHC projection studies in Ref. [4].

In summary, an extensive programme of work has been established at the LHC to characterise the Higgs boson. This thesis has presented a range of precision measurements of Higgs boson properties performed by the CMS experiment, where all measurements are found to be consistent with SM predictions. Despite this, there is increasing evidence from other areas of particle physics showing fundamental flaws in SM theory. With the Higgs boson lying at the centre of the SM, it is not unrealistic to assume that new BSM physics will interfere with the Higgs sector in some way. Looking to the future, it is therefore critical that we continue to improve the precision of Higgs boson measurements, both in terms of accumulating more data and using more sophisticated analysis techniques. In doing so, we will achieve a more fundamental understanding of our universe.

Appendix A

Evolution of the STXS framework

This Appendix demonstrates the evolution of the STXS framework by providing a schematic of each STXS stage. The stages are defined to accommodate increasing statistics, in the sense that with more p-p collision data, we become sensitive to increasingly granular regions of the Higgs boson production phase space. Therefore, subsequent stages are defined with more kinematic bins. The stage 0, stage 1.0 and stage 1.1 schemes are relevant for the EFT interpretation in Chapter 7, which combines cross sections measurements from multiple decay channels at these stages. The H $\rightarrow \gamma\gamma$ analysis in Chapters 4–6 is based on the stage 1.2 binning scheme.

A.1 Stage 0

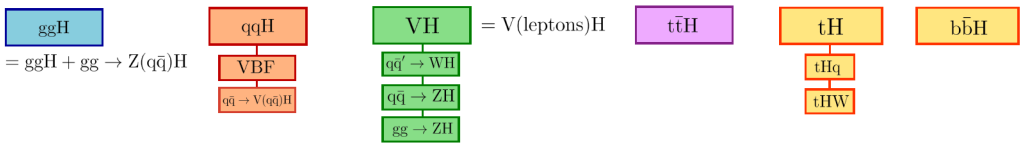


Figure A.1: Schematic showing the STXS stage 0 binning definition. The bins correspond to the different Higgs boson production modes at the LHC.

A.2 Stage 1.0

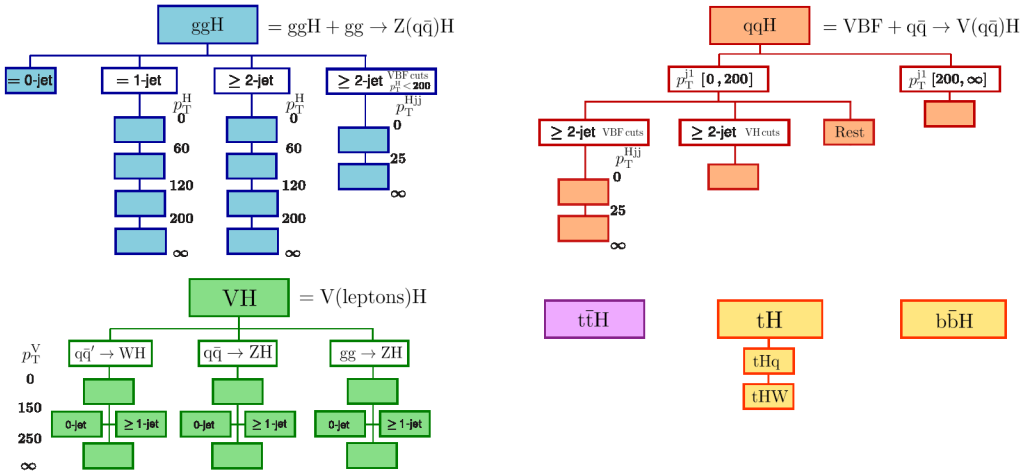


Figure A.2: Schematic showing the STXS stage 1.0 binning definition. Here, the ggH, qqH and VH leptonic production modes are split by the event kinematics for the first time. The "VBF cuts" require the dijet invariant mass, $m_{jj} > 400$ GeV, and the difference in pseudorapidity between the jets, $\Delta\eta_{jj} > 2.8$. The "VH cuts" require $60 < m_{jj} < 120$ GeV.

A.3 Stage 1.1

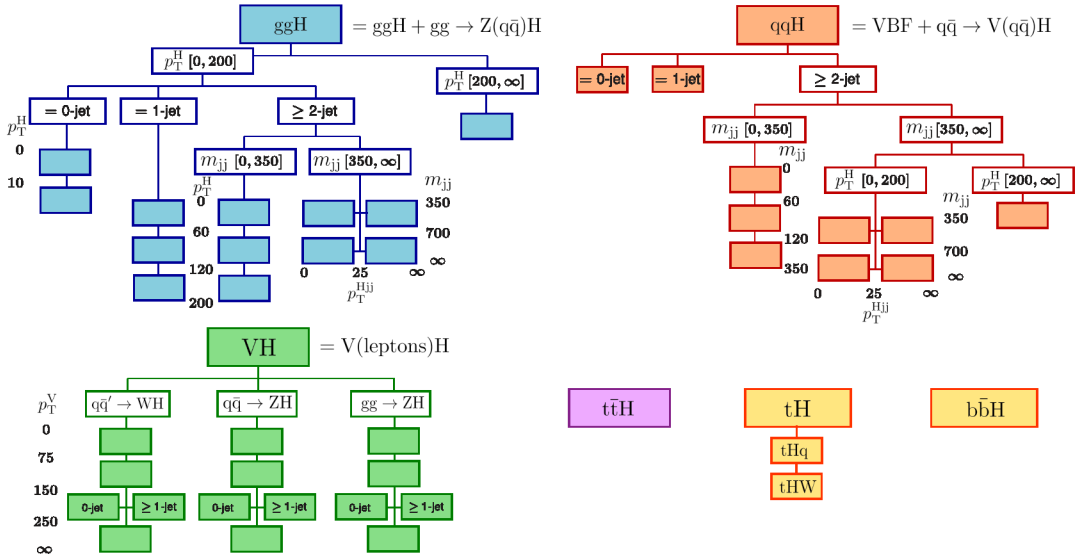


Figure A.3: Schematic showing the STXS stage 1.1 binning definition. The ggH and qqH binnings are significantly revised with respect to the stage 1.0 scheme. Additionally, the VH leptonic bins include an extra splitting at $p_T^V = 75$ GeV.

A.4 Stage 1.2

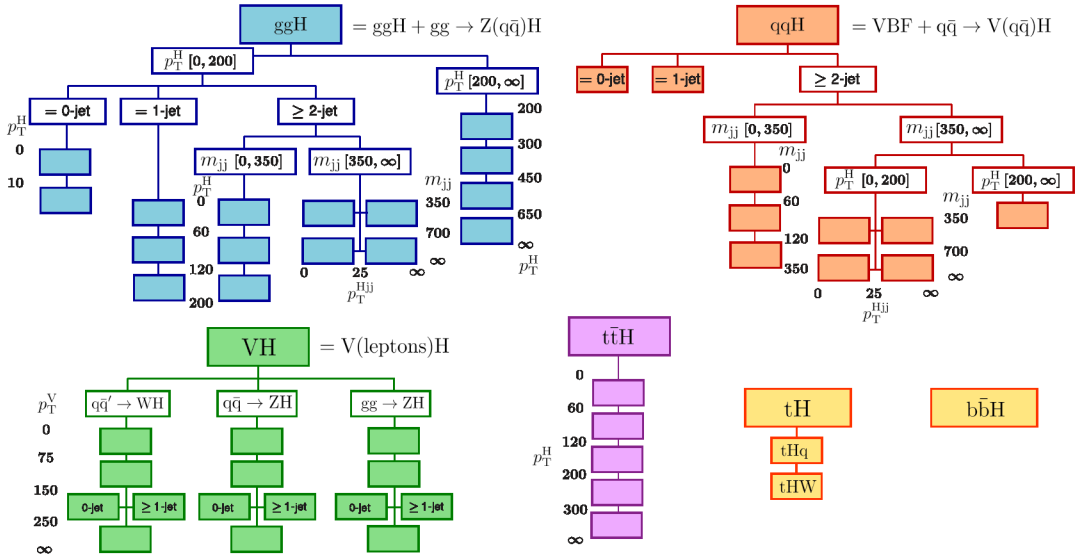


Figure A.4: Schematic showing the STXS stage 1.2 binning definition. With respect to the stage 1.1 scheme, the ggH BSM region with $p_T^H > 200$ GeV is split into four bins according to p_T^H boundaries. Also, the ttH production mode is split here for the first time, into five separate p_T^H bins.

Appendix B

Machine learning algorithms

Machine learning (ML) algorithms have become a widely used tool in high energy physics. This is particularly true for the H $\rightarrow \gamma\gamma$ analysis described in Chapters 4–6, where ML algorithms are used for a number of tasks including the photon energy regression and the event categorisation. Additionally, an ML algorithm is described in Section 8.3 to discriminate electromagnetic activity from hadronic activity in the HGCAL L1T. Because of this, it is worth providing an introduction to the foundations of ML in this Appendix.

The field of machine learning concerns developing sophisticated algorithms with the ability to *learn* from data, and subsequently apply the learnt information to solve complex problems. A generic ML problem can be formulated as follows [137]. The data are expressed as a vector space of dimension m: $X = \mathbb{R}^m$, where each dimension corresponds to an observable quantity referred to as a *feature*. An element of the data set, for example an event in a event classification task or a SC in an energy regression task, corresponds to a single feature vector, $\vec{x} \in X$. The full data set of N elements is defined by the set of feature vectors, $\vec{x}_i$, where $i = 1, ..., N$. The purpose of the ML algorithm is to develop a model,

$$f(\vec{x} \,|\, \vec{w}) \rightarrow Y, \tag{B.1}$$

to *predict an outcome*, Y, based on the feature vector, $\vec{x}$, given a set of model parameters, $\vec{w}$. We can identify two main types of ML algorithm based on the form of the predicted outcome:

- A *classification task* equates to predicting one of k possible *output classes*: $f(\vec{x} \,|\, \vec{w}) \rightarrow y$, where $y \in \{1, ..., k\}$. The obvious example here is a binary signal-vs-background event classifier, which aims to predict if an event looks signal-like or background-like based on a set of kinematic features. Many classification algorithms are introduced

throughout this thesis, including the HGCAL L1T algorithm introduced in Section 8.3.

- A *regression task* equates to predicting a quantitative outcome, or in other words a continuous value: $f(\vec{x}\,|\,\vec{w}) \to y$, where $y \in \mathbb{R}$. An example in this thesis is the SC energy regression described in Section 3.3.3, where the regressor predicts the true energy of the SC (and its uncertainty) using a combination of shower shape, seed crystal, and pileup-related features.

There are two stages when developing an ML algorithm. The learning process is referred to as *training* the model, where the values of the model parameters, $\vec{w}$, are optimised to maximise the performance. Following this, the performance of the model is evaluated using so-far unseen data; this is referred to as the *testing* stage.

A loss function, L, is constructed to measure the performance of the model, f, for a given set of input features, $\vec{x}$:

$$L[\, f(\vec{x}\,|\,\vec{w})\,] \to \mathbb{R}. \tag{B.2}$$

Supervised learning algorithms refer to the case where the target values, y, are known for each element of the training data set, $\vec{x}$. In this case, the loss function is constructed to minimise the discrepancy between the true and estimated values of the outcome, y. An additional class of algorithms where the target values are not known are referred to as *unsupervised* learning algorithms [138]; these do not feature in this thesis and are therefore not described further. Learning effectively corresponds to reducing the value of L by optimising the parameters, $\vec{w}$. In practice, for most ML algorithms this consists of some gradient based optimisation, where one descents the gradient of L with respect to $\vec{w}$ in order to find the minimum,

$$\vec{\nabla}_{\vec{w}}\, L = 0. \tag{B.3}$$

It is often not viable to evaluate this expression over the entire training dataset, especially in the case of large statistics with a high number of input features (dimensionality). A number of powerful optimisation algorithms have been developed to combat this [138, 211, 212]. These typically involve calculating the gradient for small batches of training data and optimising $\vec{w}$ iteratively, and extending this with the concept of *momentum*, where the update to the parameter vector, $\vec{w}$, depends on the size of the gradient at that point.

Crucially, it is not sufficient to simply find the configuration of $\vec{w}$ which minimises the loss for the training data set. In addition, the model is also required to generalise to new data. Therefore, the performance is evaluated on an independent test set, which is chosen to be representative of the whole data set. If the performance is significantly degraded

for the test set, then the model is said to have *over-trained* and has become specific to properties of the training set. One approach to controlling the level of over-training is to introduce regularisation terms into the loss function [138].

A large variety of ML algorithms are used in high energy physics. In this thesis, the most commonly used is the Boosted Decision Tree (BDT) algorithm [213], which is described in the remainder of this section. Neural networks are also used for discriminating between the ttH and tH production modes in the $H \to \gamma\gamma$ event classification (Section 4.4), as well as for identifying jets originating from the decay of b quarks (Section 4.3.3); further detail concerning neural networks can be found in Refs. [137–139].

BDTs are an example of *ensembling*, where multiple models are trained (base learners) and combined in some way to improve the overall performance of the algorithm. The base learners in this case are Decision Trees (DT) [214], which are built according to the following procedure:

- The feature space is partitioned into regions according to some selection (cut) on one or more of the input features. The choice and position of the cut is optimised according to a measure of purity for classification tasks, or a loss function such as the mean-squared error for regression tasks.
- This partitioning is repeated in each region, creating further subregions based on a new, optimised selection cut.
- The procedure terminates when a stopping criterion is reached. This can either be due to a predefined *max depth* (maximum number of splittings), or when a particular value of the splitting quantity (e.g. purity) has been reached. The final regions of the feature space that are not further split are referred to as *leaves*. Each leaf is assigned an output value according to the data points in that region: for classification, this is the most common output class; for regression, this is the mean of the data values.
- DTs are regularised by *pruning* branches which use unimportant features and give no performance improvement. This help mitigate over-training.

An ensemble of DTs is then constructed using a *boosting algorithm* [215, 216]. Here, multiple DTs are trained in succession, where each iteration aims to improve upon the weaknesses of the previous base learners. The final ensemble (BDT) is defined as a weighted linear combination of the individual DTs, $f_j(\vec{x}\,|\,\vec{w}_j)$, with corresponding selection cuts, $\vec{w}_j$, according to,

$$F(\vec{x}\,|\,\vec{\gamma},\vec{w}) = \sum_{j}^{N_{\mathrm{DT}}} \gamma_j \cdot f_j(\vec{x}\,|\,\vec{w}_j). \tag{B.4}$$

The set of coefficients, $\vec{\gamma} = (\gamma_1, ..., \gamma_{N_{\mathrm{DT}}})$, are determined by the boosting algorithm. Building an ensemble, F, in this way produces a more powerful predictor and helps to overcome the disadvantages of individual DTs. One important consequence for classification tasks is the BDT outputs are no longer restricted to discrete values, but become continuous variables representing the output class probabilities. For a binary signal-vs-background classifier, a value close to 1 corresponds to a signal-like event, whereas a value close to -1 corresponds to a background-like event. Selection criteria on these so-called *output scores* are a common feature of the H $\rightarrow \gamma\gamma$ analysis in Chapter 4.

Appendix C

The H $\rightarrow$ $\gamma\gamma$ event categorisation

This Appendix provides supplementary details regarding the event categorisation in the H $\rightarrow$ $\gamma\gamma$ analysis, described in Section 4.4. First, the input features ($\vec{x}$) which are used for each ML classifier are listed. Figure C.1 then provides a schematic of the event categorisation, highlighting the category priority sequence that is implemented.

C.1 ML classifier input features

Table C.1 provides a full list of the input features, $\vec{x}$, used for the ML classifiers in the H $\rightarrow$ $\gamma\gamma$ analysis. The definitions of the less obvious features are provided below:

- Δ: refers to the difference between two quantities. For example, $\Delta\eta_{jj}$, is the difference in pseudorapidity between the two jets in the dijet system.

- σ_{RV}: per-event relative diphoton mass resolution estimate, under the hypothesis that the mass is reconstructed with the correct primary vertex.

- σ_{WV}: per-event relative diphoton mass resolution estimate, under the hypothesis that the mass is reconstructed with the incorrect primary vertex.

- $C_{\gamma\gamma}$: dijet centrality defined as,

$$C_{\gamma\gamma} = \exp\left(- \frac{4}{(\eta_1 - \eta_2)^2} \left(\eta_{\gamma\gamma} - \frac{\eta_1 + \eta_2}{2} \right)^2 \right), \tag{C.1}$$

  where η_1, η_2, and $\eta_{\gamma\gamma}$, are the lead jet, sublead jet and diphoton η, respectively. This quantity is extremely useful for identifying the VBF-like topology with two forward jets.

- $\cos\theta^*$: θ^* is the angle that the diphoton system makes in the diphoton-dijet centre-of-mass frame, with respect to the direction of motion of the diphoton-dijet system in the lab frame
- H_T: scalar sum of the transverse energy of all reconstructed particles in the event.
- m_T: transverse mass, defined as,

$$m_T = \sqrt{2p_T^{\ell} p_T^{miss}(1 - \cos\Delta\phi_{\ell,p_T^{miss}})}, \tag{C.2}$$

  where $\Delta\phi_{\ell,p_T^{miss}}$ is the azimuthal separation between the lepton and p_T^{miss}.
- θ_H: angle between the two photons in the diphoton rest frame.
- DeepJet and DeepCSV: refer to the algorithm used for the CMS b tagging.
- Pixel seed veto: flag to veto photon objects with corresponding hits in the innermost tracker layers. Useful for rejecting electrons which can mimic the photon signal.
- $Y_{\gamma\gamma}$: rapidity of the diphoton.
- N_X: the multiplicity of object, X. For example, N_{jets}, is the number of reconstructed jets in the event.
- DNN scores in the ttH background rejection BDTs: additional DNNs are trained with ttH signal events against one source of background only. These discriminants first entered in the H $\rightarrow \gamma\gamma$ analysis of Ref. [129], which specifically targets the ttH production mode. Three DNNs are trained in total: one for each of the $\gamma\gamma$+jets and $tt + \gamma\gamma$ backgrounds in the hadronic channel, and one for the $tt + \gamma\gamma$ background in the leptonic channel. The performance of these DNNs benefits from the high number of simulated events on which to train, as well as the fact that both of the considered backgrounds are well modelled in simulation. Importantly, the DNNs use a combination of high-level and lower-level input features, where the latter includes the full four-momentum vectors of the reconstructed objects in the event. Adding this low-level information directly into the background rejection BDTs does not improve their performance. However, using a DNN as an intermediate step and feeding the output score of the DNN into the BDT, allows the low-level information to be utilised effectively.
- Top tagger BDT score: a ML algorithm developed for the analysis in Ref. [217], is used to distinguish events with top quarks decaying to three jets, from events that do not contain top quarks. The algorithm is trained on jet triplets from the simulation

of tt production events, with input features related to the event kinematics, b tag scores, and the jet shape. The signal is defined as a jet triplet which is matched at truth-level to a top quark, and the background is taken as random jet triplets.

Table C.1: Input features to the ML event classifiers used in the H $\to \gamma\gamma$ analysis. Photons, jets, b-tagged jets and leptons are labelled as γ, j, bj, and ℓ, respectively, and the numbers represent the p_T-ordered list of the respective objects e.g. $\gamma 1$ corresponds to the leading photon. The diphoton (dijet) variables are labelled by $\gamma\gamma$ (jj). In the final two classifiers, fwd corresponds to the jet with the highest $|\eta|$ value, which provides a useful handle on identifying events originating from tHq production. Definitions of the less obvious input features are provided in the main text of this Appendix.

Discriminant/Classifier	Input features, $\vec{x}$
ggH BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, $\cos\Delta\phi_{\gamma\gamma}$, $\gamma 1/\gamma 2$ ID BDT scores, σ_{RV}, σ_{WV}, vertex probability BDT score, $p_T^{\gamma\gamma}$, N_{jets}, m_{jj}, $\eta^{j1/j2/j3}$, $j1/j2/j3$ pile-up identification BDT scores, $\Delta\phi_{\gamma\gamma,j1/j2/j3}$, $\Delta\eta_{\gamma\gamma,j1/j2/j3}$
Diphoton BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, $\cos\Delta\phi_{\gamma\gamma}$, $\gamma 1/\gamma 2$ ID BDT scores, σ_{RV}, σ_{WV}, vertex probability BDT score
Dijet BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $p_T^{\gamma\gamma}/m_{\gamma\gamma}$, $\cos\Delta\phi_{\gamma\gamma}$, $p_T^{j1/j2}$, m_{jj}, $\Delta\phi_{\gamma\gamma,jj}$, $\min(\Delta R_{\gamma,j})$, $C_{\gamma\gamma}$, $\lvert\Delta\eta_{jj}\rvert$, $\Delta\phi_{jj}$
VH hadronic BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $p_T^{j1/j2}$, $\eta^{j1/j2}$, m_{jj}, $\lvert\Delta\eta_{jj}\rvert$, $\cos\theta^*$
VH MET BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, $\cos\Delta\phi_{\gamma\gamma}$, max/min γ ID BDT scores, p_T^{miss}, H_T, N_{jets}, p_T^{j1}, max jet b-tag score (deepCSV), $\Delta\phi_{\gamma\gamma,p_T^{miss}}$, $\min(\Delta\phi_{p_T^{miss},j})$, $(p_T^{\gamma\gamma} - p_T^{miss})/p_T^{\gamma\gamma}$
WH leptonic BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, $\cos\Delta\phi_{\gamma\gamma}$, max/min γ ID BDT scores, $\gamma 1/\gamma 2$ pixel seed veto, p_T^{ℓ}, η^{ℓ}, $\Delta R_{\gamma 1/\gamma 2,\ell}$, $\Delta\theta_{\gamma\gamma,\ell}$, p_T^{miss}, m_T, N_{jets}, p_T^{j1}, $j1/j2$ b-tag score (deepCSV)
ZH leptonic BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, $\cos\Delta\phi_{\gamma\gamma}$, max/min γ ID BDT scores, $\gamma 1/\gamma 2$ pixel seed veto, $p_T^{\ell 1/\ell 2}$, $\eta^{\ell 1/\ell 2}$, $\Delta R_{\gamma 1/\gamma 2,\ell 1/\ell 2}$, $\Delta\theta_{\gamma\gamma,\ell\ell}$, $m_{\ell\ell}$, N_{jets}, p_T^{j1}, $j1$ b-tag score (deepCSV)
ttH hadronic BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, max/min γ ID BDT scores, $\gamma 1/\gamma 2$ pixel seed veto, $p_T^{\gamma\gamma}/m_{\gamma\gamma}$, $Y_{\gamma\gamma}$, $\lvert\cos\Delta\phi_{\gamma\gamma}\rvert$, $\Delta R_{\gamma\gamma}$, $\cos\theta_H$, $p_T^{j1/j2/j3/j4}$, $\eta^{j1/j2/j3/j4}$, $j1/j2/j3/j4$ b-tag score (DeepJet), max/min b-tag score, N_{jets}, H_T, p_T^{miss}, DNN scores: ttH vs tt$\gamma\gamma$ (had) and ttH vs $\gamma\gamma$+jets (had), Top tagger BDT score
ttH leptonic BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, max/min γ ID BDT scores, $\gamma 1/\gamma 2$ pixel seed veto, $p_T^{\gamma\gamma}/m_{\gamma\gamma}$, $Y_{\gamma\gamma}$, $\lvert\cos\Delta\phi_{\gamma\gamma}\rvert$, $\Delta R_{\gamma\gamma}$, $\cos\theta_H$, $p_T^{j1/j2/j3}$, $\eta^{j1/j2/j3}$, $j1/j2/j3$ b-tag score (DeepJet), max/min b-tag score, N_{jets}, H_T, p_T^{miss}, p_T^{ℓ}, η^{ℓ}, N_{leptons} (tight ID), DNN scores: ttH vs tt$\gamma\gamma$ (lep)
tHq leptonic BDT	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, max/min γ ID BDT scores, $\gamma 1/\gamma 2$ pixel seed veto, N_{jets}, N_{bjets}, N_{jets} with $\lvert\eta\rvert < 1$, $p_T^{j1/j2/j3}$, $\eta^{j1/j2/j3}$, $p_T^{bj1/bj2/bj3}$, $\eta^{bj1/bj2/bj3}$, p_T^{fwd}, η^{fwd}, $\Delta\phi_{\gamma 1/\gamma 2,j1}$, $\Delta\phi_{\gamma 1/\gamma 2,\ell}$, $\Delta\phi_{\gamma 1/\gamma 2,bj1}$, $\Delta\phi_{\gamma 1/\gamma 2,\text{fwd}}$, p_T^{ℓ}, η^{ℓ}
Top DNN	$p_T^{\gamma 1/\gamma 2}/m_{\gamma\gamma}$, $\eta^{\gamma 1/\gamma 2}$, max/min γ ID BDT scores, $\gamma 1/\gamma 2$ pixel seed veto, $p_T^{\gamma\gamma}/m_{\gamma\gamma}$, $Y_{\gamma\gamma}$, $\lvert\cos\Delta\phi_{\gamma\gamma}\rvert$, $\Delta R_{\gamma\gamma}$, $\cos\theta_H$, $p_T^{j1/j2/j3}$, $\eta^{j1/j2/j3}$, $j1/j2/j3$ b-tag score (DeepJet), max/min b-tag score, N_{jets}, H_T, p_T^{miss}, p_T^{ℓ}, η^{ℓ}, N_{leptons} (tight ID), p_T^{fwd}, η^{fwd}, $\ell 1/\ell 2$ charge

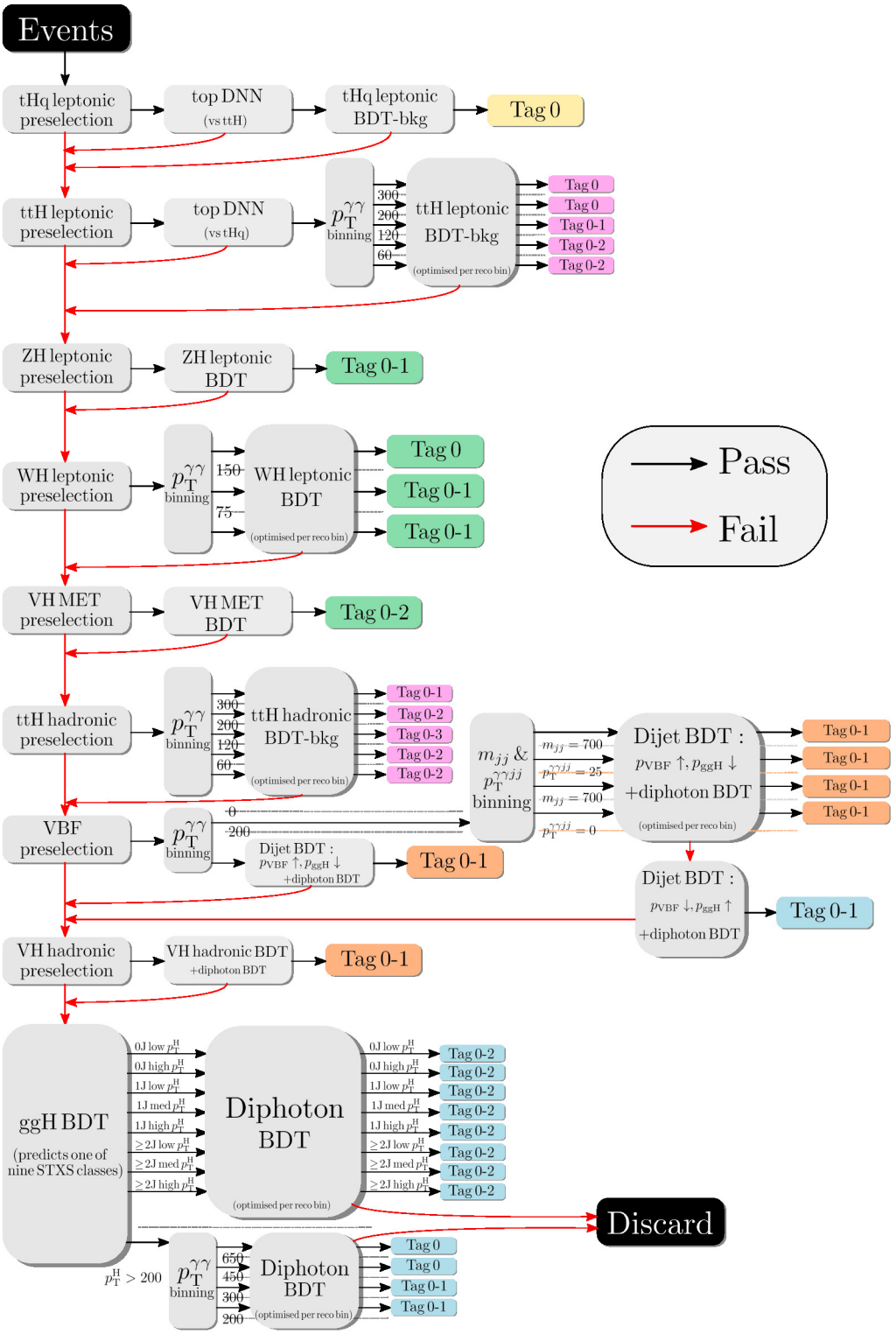


Figure C.1: A schematic of the event categorisation in the H $\rightarrow \gamma\gamma$ analysis. Reconstructed events passing the trigger selection are required to pass the photon pre-selection criteria. Those that do, enter the top left of the diagram. In the schematic, black and red lines signify events passing and failing the relevant criteria, respectively. Events that do not end up in any analysis category (coloured boxes) are discarded from the analysis.

Appendix D

Confusion and $\epsilon_k^{i,\gamma\gamma}$ matrices: H $\rightarrow$ $\gamma\gamma$

This Appendix contains the confusion matrix for the full set of analysis categories used in the H $\rightarrow$ $\gamma\gamma$ analysis. This differs from Figure 4.20 as the categories are fully split into the individual tags. Moreover, the detector efficiency times analysis acceptance terms, $\epsilon_k^{i,\gamma\gamma}$, used in the normalisation of the signal models are shown. These terms are derived separately for the 2016, 2017, and 2018 simulation.

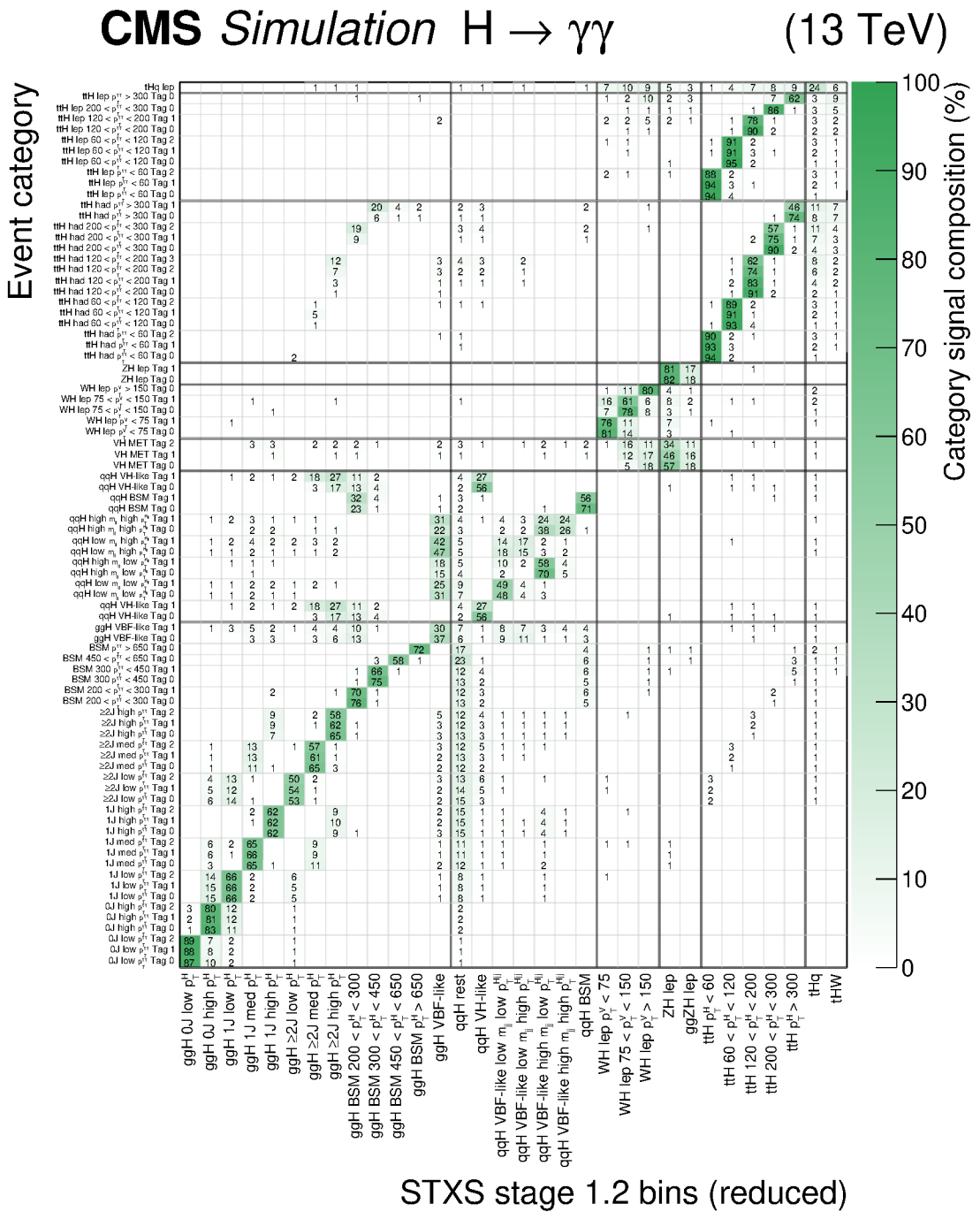


Figure D.1: Confusion matrix displaying the composition of each analysis category in terms of a merged set of STXS bins. The colour scale corresponds to the fractional yield in each analysis category (row), accounted for by each STXS process (column). Each row therefore sums to 100%. Entries with values less than 0.5% are not shown. Simulated events for each year in the period 2016-2018 are combined with appropriate weights corresponding to their relative integrated luminosity in data. The column labelled as qqH rest includes the contributions from the qqH 0J, qqH 1J, qqH $m_{jj} < 60$ GeV and the qqH $120 < m_{jj} < 350$ GeV STXS bins.

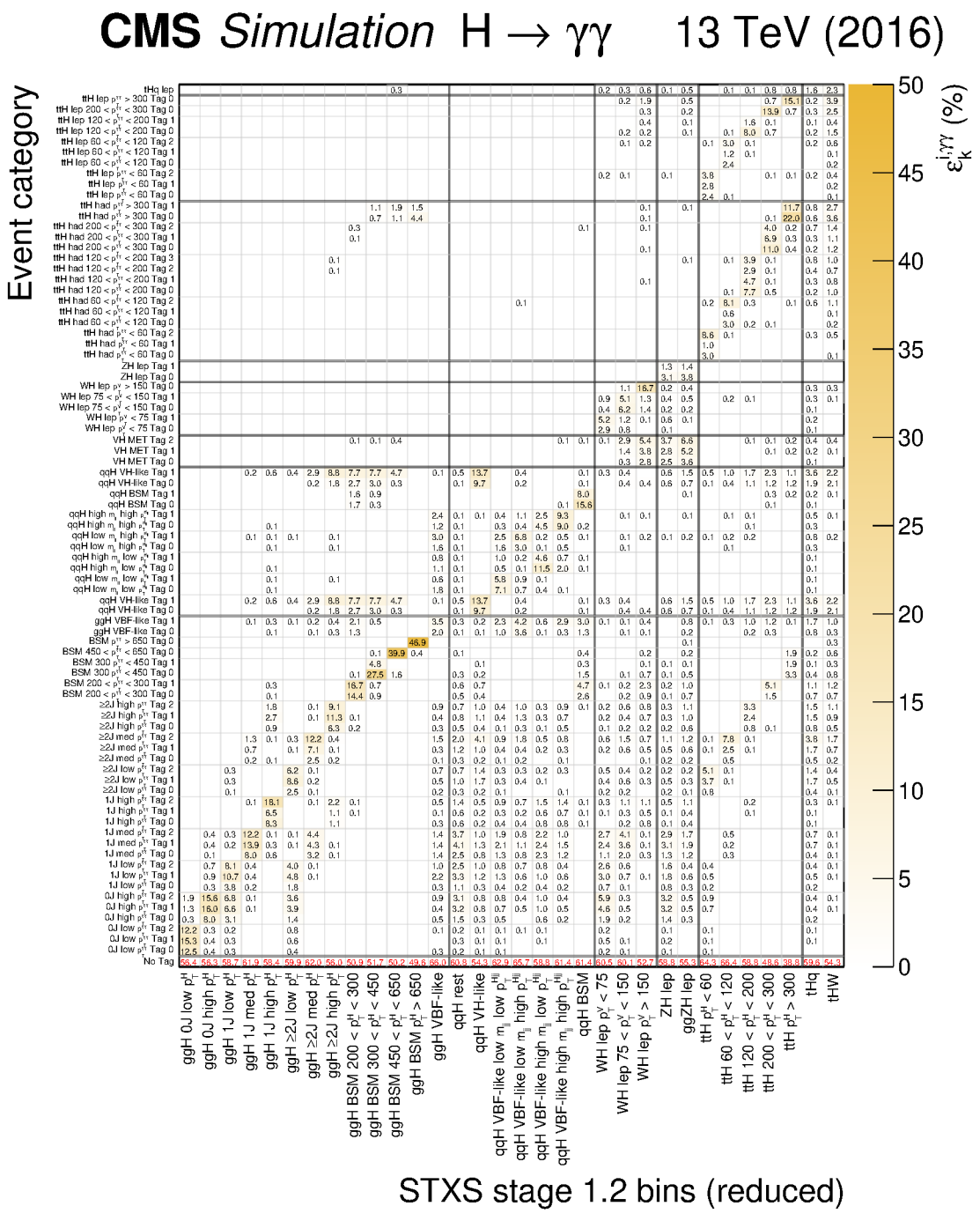


Figure D.2: A matrix showing the $\epsilon_k^{i,\gamma\gamma}$ terms for a merged set of STXS bins, derived from 2016 simulation. The bin merging is done purely for plotting purposes; in the analysis a separate $\epsilon_k^{i,\gamma\gamma}$ factor is calculated for each STXS bin. The numbers show the fraction of the total yield of each (merged) STXS bin, landing in each analysis category. Each column therefore sums to 100%. Entries with values less than 0.05% are not shown. The bottom row, referred to as "No Tag" indicates the fraction of events originating from the corresponding process which do not enter a single analysis category. The column labelled as qqH rest includes the contributions from the qqH 0J, qqH 1J, qqH $m_{jj} < 60$ GeV and the qqH $120 < m_{jj} < 350$ GeV STXS bins.

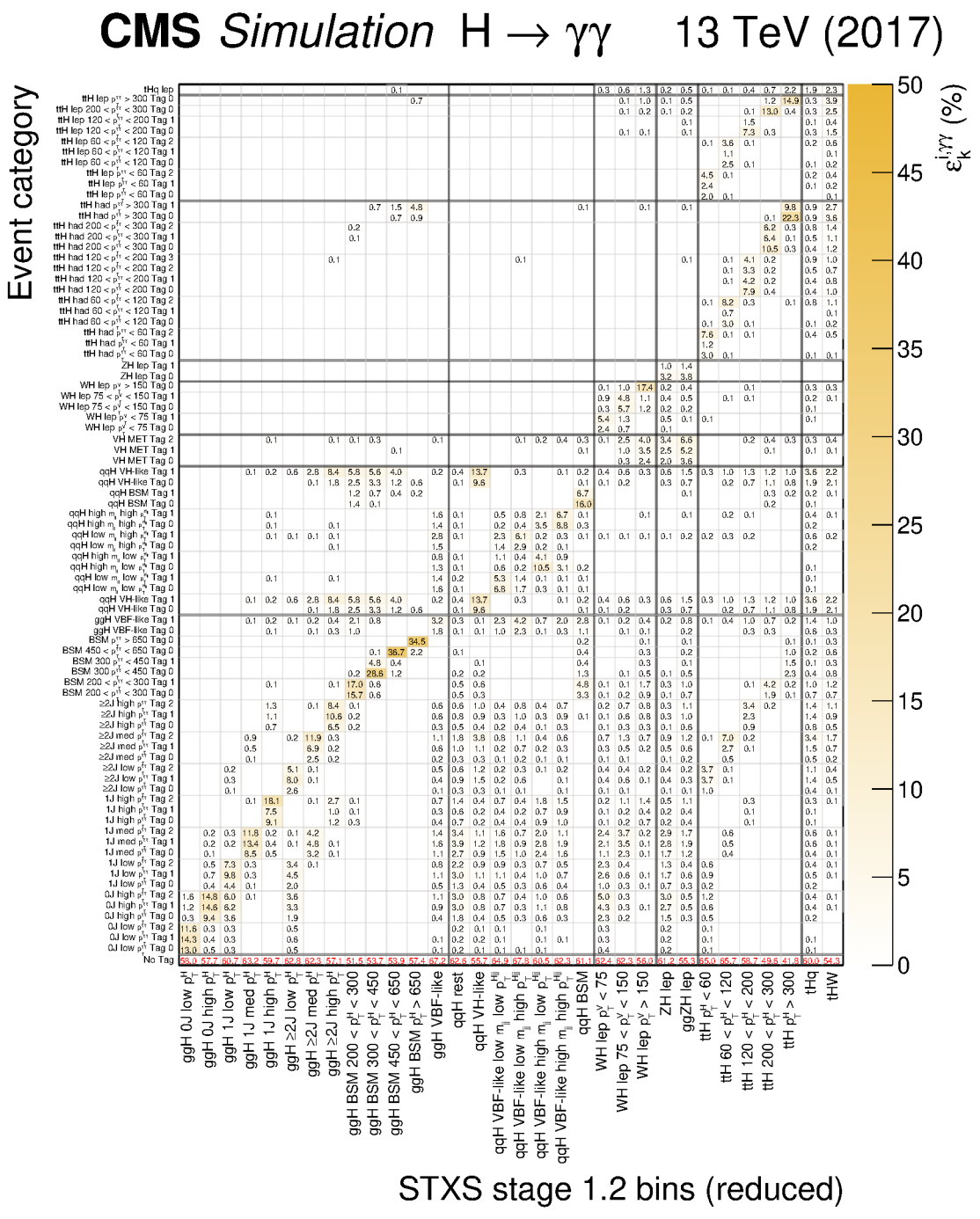


Figure D.3: A matrix showing the $\epsilon_k^{i,\gamma\gamma}$ terms for a merged set of STXS bins, derived from 2017 simulation. The bin merging is done purely for plotting purposes; in the analysis a separate $\epsilon_k^{i,\gamma\gamma}$ factor is calculated for each STXS bin. The numbers show the fraction of the total yield of each (merged) STXS bin, landing in each analysis category. Each column therefore sums to 100%. Entries with values less than 0.05% are not shown. The bottom row, referred to as "No Tag" indicates the fraction of events originating from the corresponding process which do not enter a single analysis category. The column labelled as qqH rest includes the contributions from the qqH 0J, qqH 1J, qqH $m_{jj} < 60$ GeV and the qqH $120 < m_{jj} < 350$ GeV STXS bins.

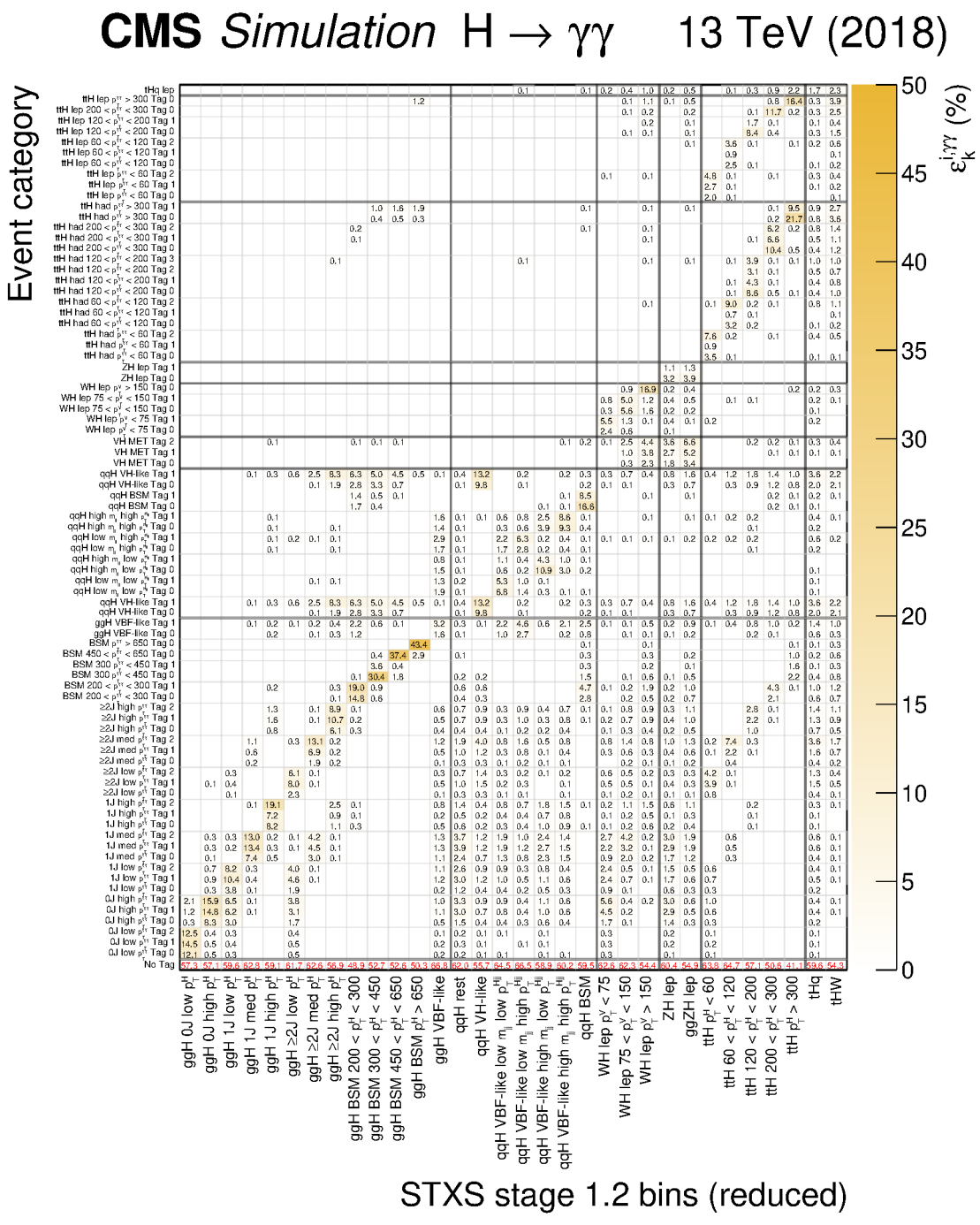


Figure D.4: A matrix showing the $\epsilon_k^{i,\gamma\gamma}$ terms for a merged set of STXS bins, derived from 2018 simulation. The bin merging is done purely for plotting purposes; in the analysis a separate $\epsilon_k^{i,\gamma\gamma}$ factor is calculated for each STXS bin. The numbers show the fraction of the total yield of each (merged) STXS bin, landing in each analysis category. Each column therefore sums to 100%. Entries with values less than 0.05% are not shown. The bottom row, referred to as "No Tag" indicates the fraction of events originating from the corresponding process which do not enter a single analysis category. The column labelled as qqH rest includes the contributions from the qqH 0J, qqH 1J, qqH $m_{jj} < 60$ GeV and the qqH $120 < m_{jj} < 350$ GeV STXS bins.

Appendix E

Observed diphoton-mass distributions

The observed diphoton invariant-mass distributions in the individual $H \to \gamma\gamma$ analysis categories are presented in this Appendix. Here, the best-fit signal-plus-background model corresponds to the per-production mode signal-strength fit.

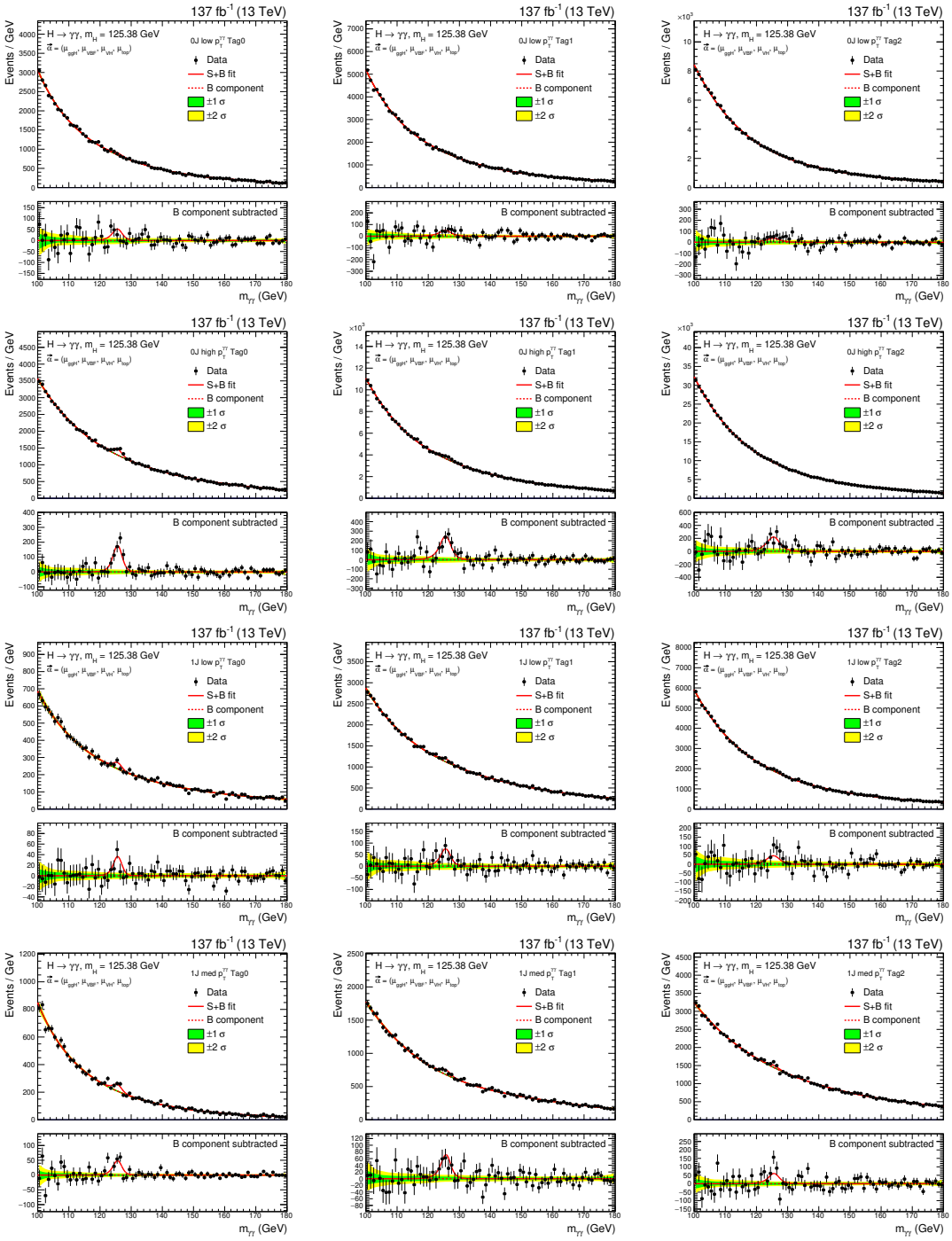


Figure E.1: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting a number of ggH STXS regions. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

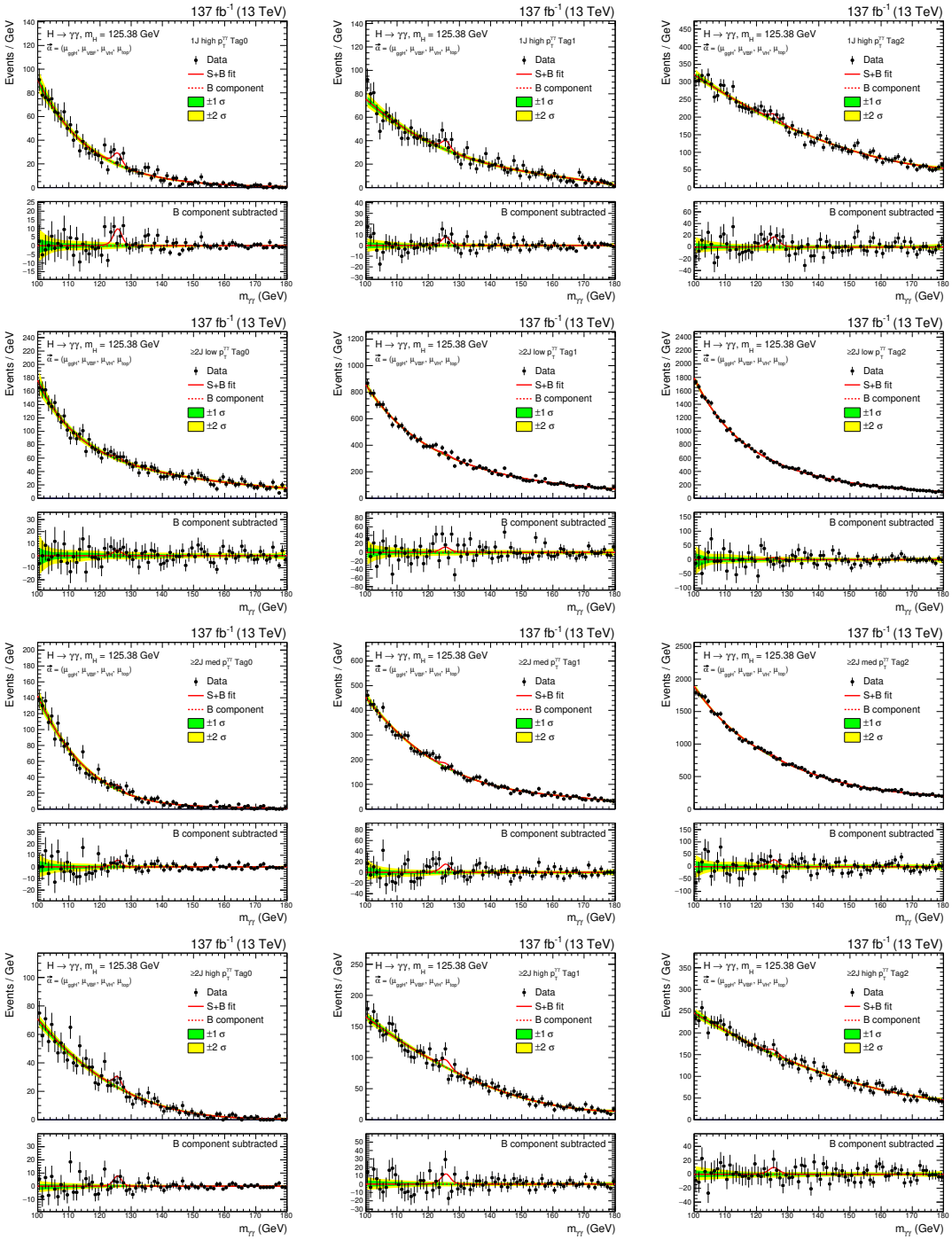


Figure E.2: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting a number of ggH STXS regions. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

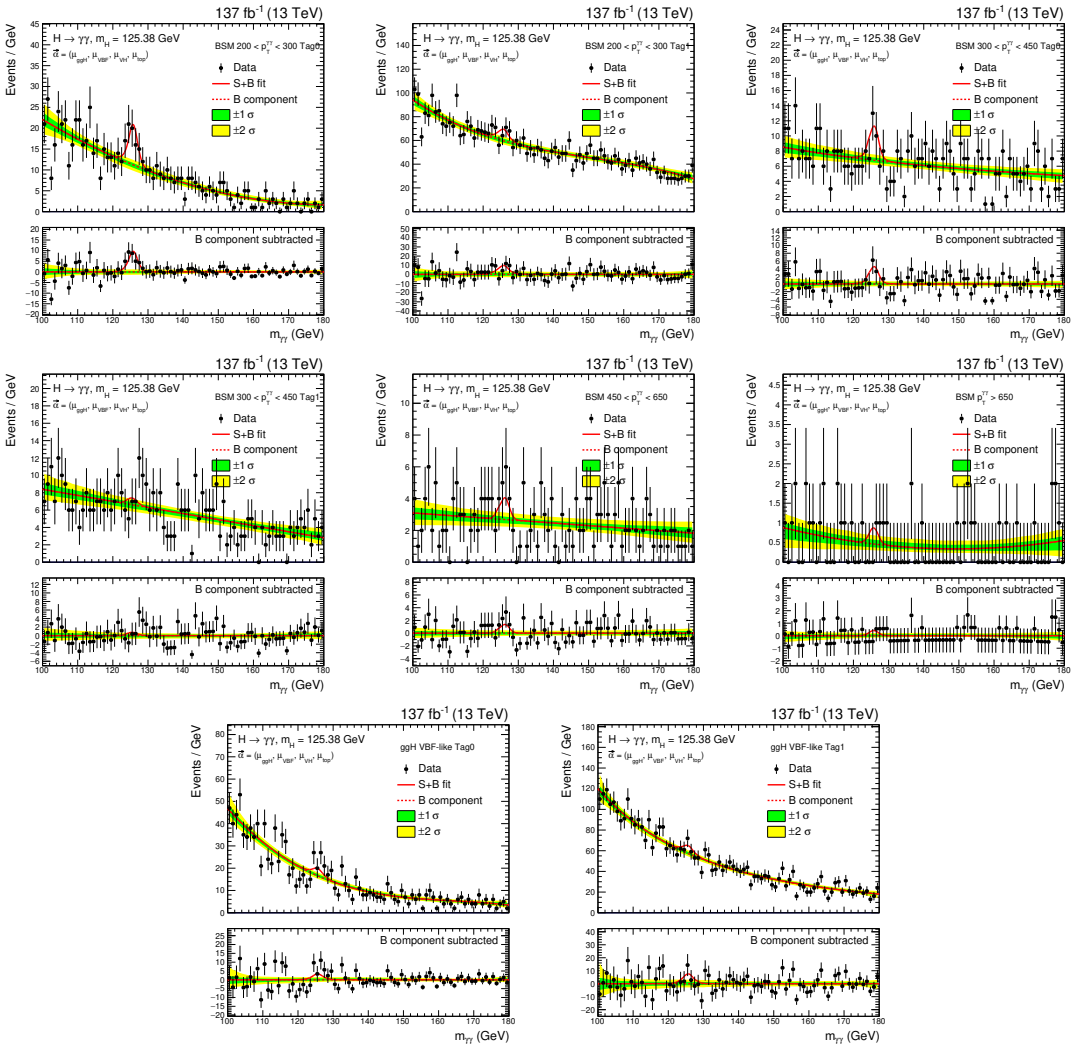


Figure E.3: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting the ggH BSM and ggH VBF-like STXS regions. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

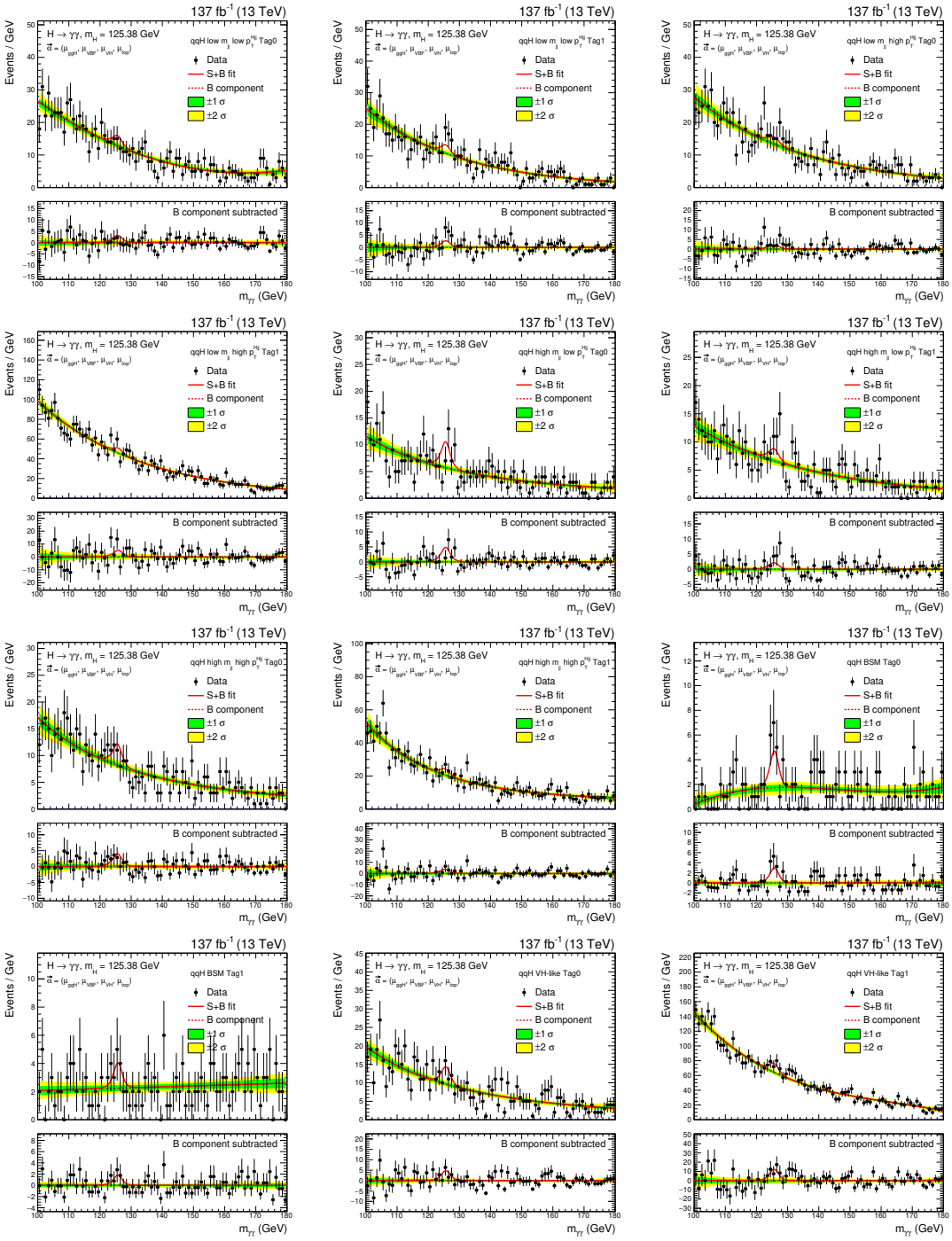


Figure E.4: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting the qqH STXS regions. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

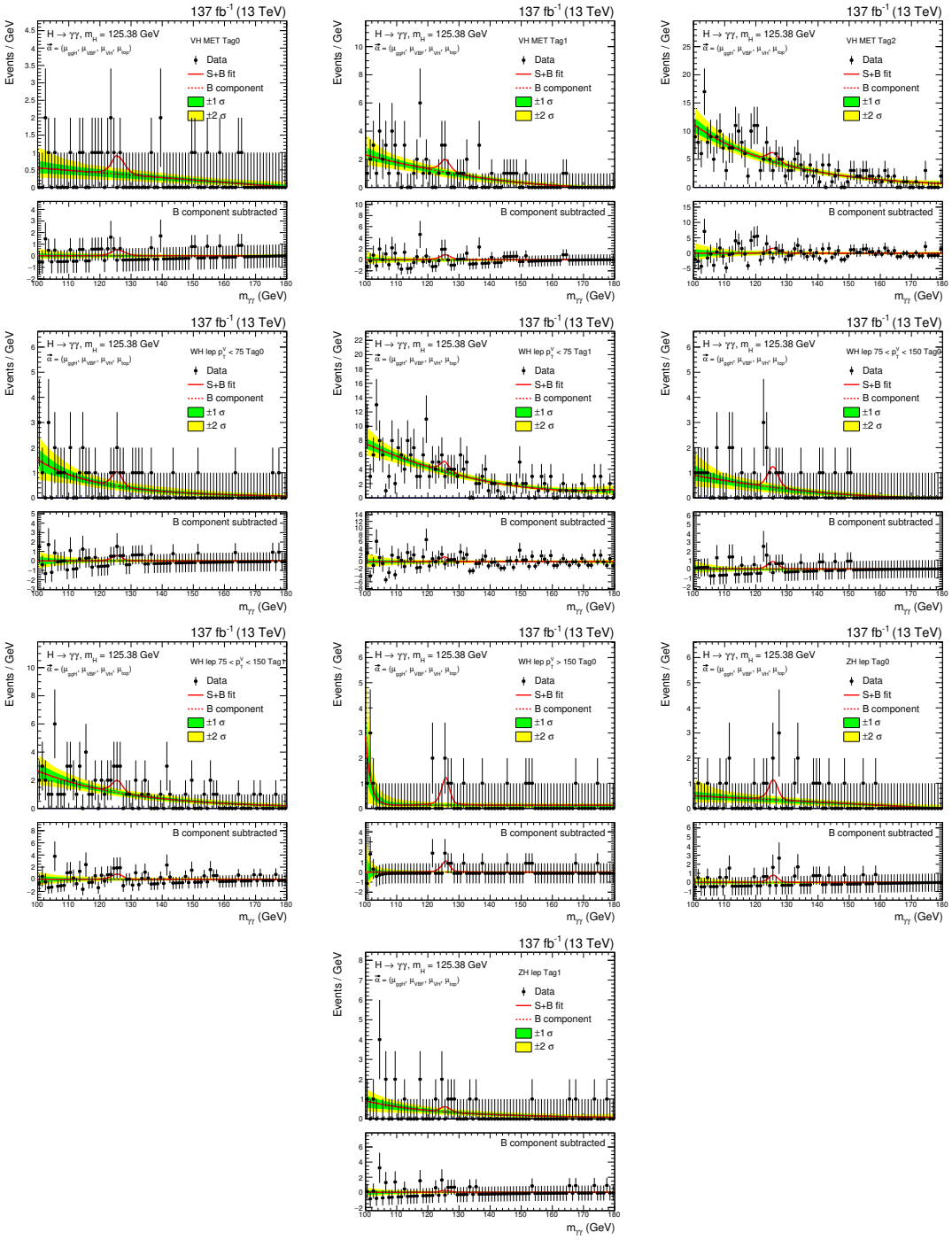


Figure E.5: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting the VH leptonic STXS regions. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

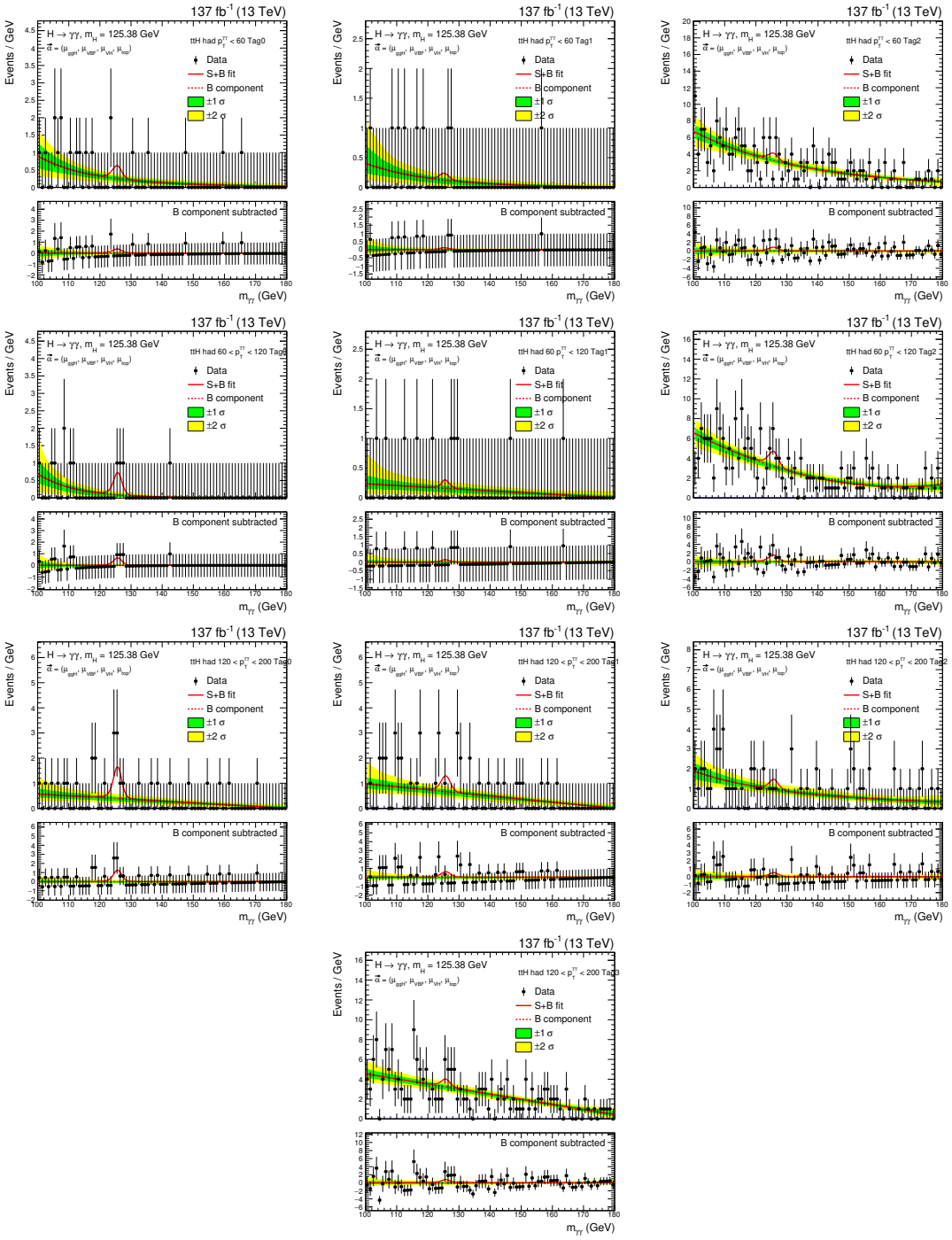


Figure E.6: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting the ttH STXS regions in the hadronic channel. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

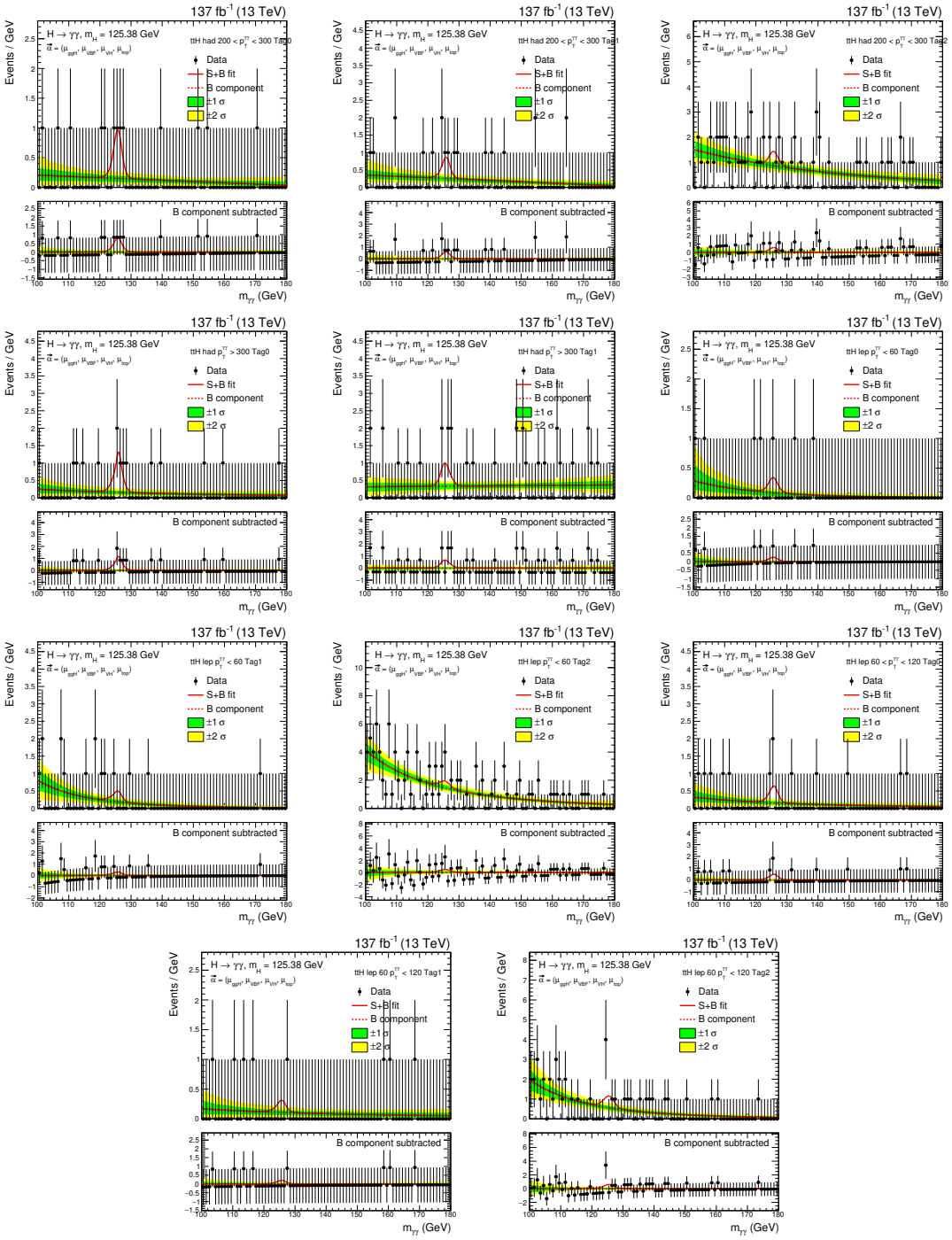


Figure E.7: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting the ttH STXS regions, in the hadronic and leptonic channels. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

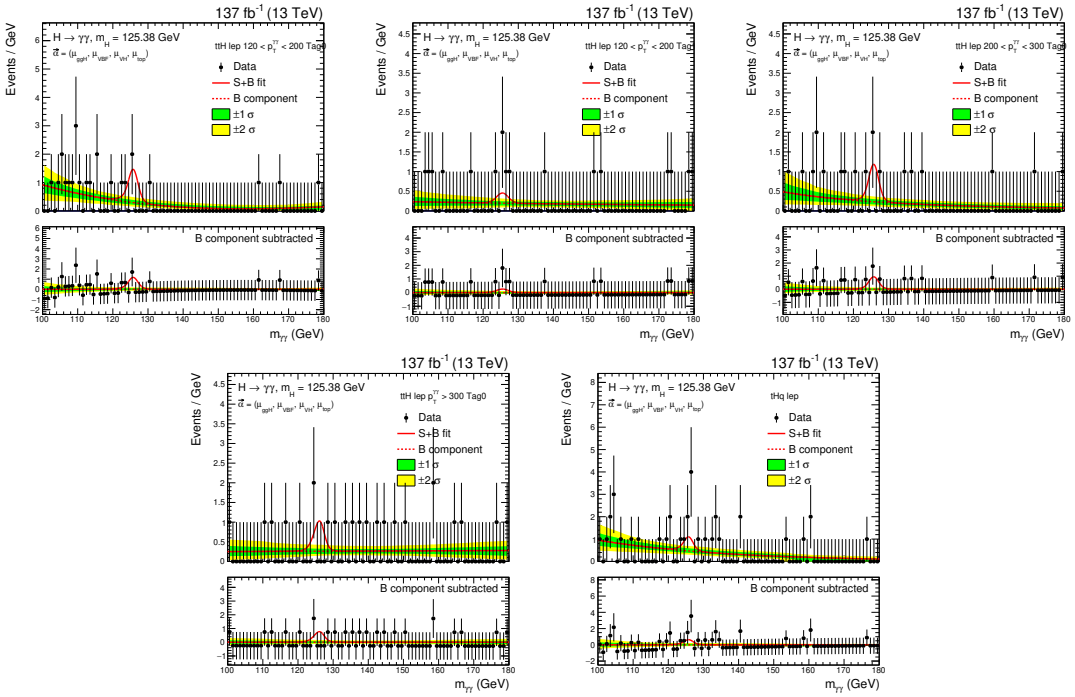


Figure E.8: Data points (black) and the best-fit signal-plus-background model for the individual analysis categories targeting the ttH and tH STXS regions, in the leptonic channel. The best-fit model corresponds to the per-production mode signal strength fit. The solid red line shows the best-fit signal-plus-background model, whereas the dashed line shows the background component only. The one standard deviation (green) and two standard deviation (yellow) bands show the uncertainties in the background component of the fit. The bottom panels in each plot show the residuals after subtraction of this background component.

Appendix F

$H \rightarrow \gamma\gamma$: STXS merging schemes

This Appendix tabulates the definition of the maximal and minimal merging schemes, introduced in Chapter 6. For each parameter of interest, the contributing STXS stage 1.2 bins are listed.

Table F.1: A summary of the maximal and minimal parameter merging schemes. The STXS bins that contribute to each parameter are listed. Furthermore, the bins that are constrained to their respective SM predictions in the fits are listed at the bottom.

Scheme	Parameters	STXS stage 1.2 bins (total number of bins)
Maximal (17 parameters)	ggH 0J low p_T^H	ggH 0J low p_T^H (1)
	ggH 0J high p_T^H	ggH 0J high p_T^H, bbH (2)
	ggH 1J low p_T^H	ggH 1J low p_T^H (1)
	ggH 1J med p_T^H	ggH 1J med p_T^H (1)
	ggH 1J high p_T^H	ggH 1J high p_T^H (1)
	ggH $\geq$ 2J low p_T^H	ggH $\geq$2J low p_T^H (1)
	ggH $\geq$ 2J med p_T^H	ggH $\geq$2J med p_T^H (1)
	ggH $\geq$ 2J high p_T^H	ggH $\geq$2J high p_T^H (1)
	ggH BSM	{ggH BSM $200 < p_T^H < 300$, ggH BSM $300 < p_T^H < 450$ / ggH BSM $450 < p_T^H < 650$, ggH BSM $p_T^H > 650$} (4)
	ggH VBF-like	{ggH VBF-like low m_{jj} low p_T^{Hjj}, ggH VBF-like low m_{jj} high p_T^{Hjj} / ggH VBF-like high m_{jj} low p_T^{Hjj}, ggH VBF-like high m_{jj} high p_T^{Hjj}} (4)
	qqH VBF-like	{qqH VBF-like low m_{jj} low p_T^{Hjj}, qqH VBF-like low m_{jj} high p_T^{Hjj} / qqH VBF-like high m_{jj} low p_T^{Hjj}, qqH VBF-like high m_{jj} high p_T^{Hjj}} (4)
	qqH VH-like	qqH VH-like (1)
	qqH BSM	qqH BSM (1)
	WH lep	All WH lep (5)
	ZH lep	All ZH lep and ggZH lep (10)
	ttH	All ttH (5)
	tH	tH = tHq + tHW (1)
Minimal (27 parameters)	ggH 0J low p_T^H	ggH 0J low p_T^H (1)
	ggH 0J high p_T^H	ggH 0J high p_T^H, bbH (2)
	ggH 1J low p_T^H	ggH 1J low p_T^H (1)
	ggH 1J med p_T^H	ggH 1J med p_T^H (1)
	ggH 1J high p_T^H	ggH 1J high p_T^H (1)
	ggH $\geq$ 2J low p_T^H	ggH $\geq$2J low p_T^H (1)
	ggH $\geq$ 2J med p_T^H	ggH $\geq$2J med p_T^H (1)
	ggH $\geq$ 2J high p_T^H	ggH $\geq$2J high p_T^H (1)
	ggH BSM $200 < p_T^H < 300$	ggH BSM $200 < p_T^H <$ 300 (1)
	ggH BSM $300 < p_T^H < 450$	ggH BSM $300 < p_T^H <$ 450 (1)
	ggH BSM $p_T^H > 450$	ggH BSM $450 < p_T^H < 650$, ggH BSM $p_T^H > 650$ (2)
	VBF-like low m_{jj} low p_T^{Hjj}	ggH + qqH VBF-like low m_{jj} low p_T^{Hjj} (2)
	VBF-like low m_{jj} high p_T^{Hjj}	ggH + qqH VBF-like low m_{jj} high p_T^{Hjj} (2)
	VBF-like high m_{jj} low p_T^{Hjj}	ggH + qqH VBF-like high m_{jj} low p_T^{Hjj} (2)
	VBF-like high m_{jj} high p_T^{Hjj}	ggH + qqH VBF-like high m_{jj} high p_T^{Hjj} (2)
	qqH VH-like	qqH VH-like (1)
	qqH BSM	qqH BSM (1)
	WH lep $p_T^V < 75$	WH lep $p_T^V < 75$ (1)
	WH lep $75 < p_T^V < 150$	WH lep $75 < p_T^V < 150$ (1)
	WH lep $p_T^V > 150$	{WH lep 0J $150 < p_T^V < 250$, WH lep $\geq$1J $150 < p_T^V < 250$ / WH lep $p_T^V > 250$} (3)
	ZH lep	All ZH lep and ggZH lep (10)
	ttH $p_T^H < 60$	ttH $p_T^H < 60$ (1)
	ttH $60 < p_T^H < 120$	ttH $60 < p_T^H < 120$ (1)
	ttH $120 < p_T^H < 200$	ttH $120 < p_T^H < 200$ (1)
	ttH $200 < p_T^H < 300$	ttH $200 < p_T^H < 300$ (1)
	ttH $p_T^H > 300$	ttH $p_T^H > 300$ (1)
	tH	tH = tHq + tHW (1)
Constrained to SM prediction		qqH 0J, qqH 1J, qqH $m_{jj} < 60$, qqH $120 < m_{jj} < 350$ (4)

Appendix G

EFT parametrisation: MC generator details

This Appendix describes the technical details in deriving the EFT parametrisations, namely the options used to configure the MC generators. Firstly, the treatment of m_{Z} in the HEL interpretation is described. Following this, a crib sheet is provided in Table G.1, which lists all the relevant details.

G.1 The treatment of m_{Z} in the HEL interpretation

The nominal HEL model is uniquely defined by the input parameters: m_{W}, α_{EM} and G_F, where m_{W} is the W boson mass, α_{EM} is the electromagnetic coupling constant, and G_F is the Fermi constant. In this input parameter scheme, the Z boson mass, m_{Z}, is dependent on the HEL parameters, c_T, c_{WW}, c_B and c_A, according to equation G.1,

$$m_{\mathrm{Z}}^2(\vec{c}) = m_{Z,\mathrm{SM}}^2\Big[1 - c_T + \frac{8\,c_A\,\sin^4(\theta_W) + 2\,c_{WW}\,\cos^2(\theta_W) + c_B\,\sin^2(\theta_W)}{\cos^2(\theta_W)}\Big], \qquad \text{(G.1)}$$

where $m_{Z,\mathrm{SM}}$ is the Z boson mass in the SM, and θ_W is the Weinberg angle. In the interpretation documented in this thesis, the HEL model has been adapted to remove the $\vec{c}$ dependence of m_{Z}, and instead fix its value to the SM prediction. This reflects the fact that, although there is no explicit Z boson mass measurement entering the combination, it is well measured experimentally and therefore it is not physical to consider large variations in its value. In fact, the chosen operator subset, $\{\mathcal{O}\}$, does not include $\mathcal{O}_T$, and $\mathcal{O}_{WW}$ and $\mathcal{O}_B$ are fit together in the combination of parameters, $c_{WW} - c_B$, which equation G.1 does not explicitly depend on. The operator $\mathcal{O}_A$ is included in the subset, however the c_A dependence is small ($\propto \sin^4(\theta_W)$), such that over the allowed range in c_A there is a negligible shift in the m_{Z} value, well within its measured uncertainty. All in all, this means the treatment of m_{Z} in the interpretation has a negligible effect.

Table G.1: Crib sheet for the MC generator details used when deriving the EFT parametrisation.

Generator	**Parton showering**	**Event classification**	**Perturbative order** (QCD and QED)
MG5_aMC@NLO [119] (v2.6.5 for HEL) (v2.6.7 for Warsaw)	Pythia8 [127] (v8.201 for HEL) (v8.244 for Warsaw)	Rivet [181]: `HiggsTemplateCrossSections`	LO

FeynRules `UFO` model		
SILH basis (HEL)	`HEL_UFO` [163]	Input parameter scheme: m_{W}, α_{EM}, G_F. Model has been adapted to fix the Z boson mass to the SM prediction.
Warsaw basis	`SMEFTsim` [79]	Input parameter scheme: m_{W}, m_{Z}, G_F. Assuming $\mathrm{U}(3)^5$ flavour symmetry.

Madgraph process definitions, $\mathcal{M}$

New physics option: `NP<=1` (added to every line in process definition)

Production Mode	Syntax

ggH	`generate p p > h QED=1`
	`add process p p > h j QED=1`
	`add process p p > h j j QED=1`

qqH	`generate p p > h j j QCD=0`
WH lep	`generate p p > h l+ vl` `add process p p > h l- vl~`
ZH lep	`generate p p > h l+ l-` `add process p p > h vl vl~`
ttH	`generate p p > h t t~`
tHq	`generate p p > h t b~ j` `add process p p > h t~ b j`

tHW	`define p = p b b~`
	`generate p p > h t w-`
	`add process p p > h t~ w+`

bbH	`generate p p > h b b~`

Decay channel	Syntax
$\mathrm{H} \to \gamma\gamma$	`generate h > a a`
$\mathrm{H} \to \mathrm{ZZ} \to 4\ell$	`generate h > l+ l- l+ l-`
$\mathrm{H} \to \mathrm{WW} \to \ell\nu\ell\nu$	`generate h > l+ vl l- vl`
$\mathrm{H} \to \tau\tau$	`generate h > ta+ ta-`
$\mathrm{H} \to \mathrm{bb}$	`generate h > b b~`
For total width, Γ^H	
$\mathrm{H} \to \mathrm{cc}$	`h > c c~`
$\mathrm{H} \to \gamma\gamma$	`h > g g`

+ others, same as in Table 4 of Ref. [218]

Generator, parton shower and classification options

All Madgraph options are set to their default values

Jet matching algorithm	`MLM`	Matching parameter: `xqcut=30`.
Kinematic cuts	None	A couple of cuts are applied at the classification stage: jets are required to have $p_T^j > 30$ GeV in the Rivet routine. Lepton pairs are required to have $\Delta R_{\ell\ell} > 0.05$ when deriving the $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ parametrisation to avoid divergences in the colinear limit.
Parton distribution function	`lhaid=230000` `NNPDF2.3` (LO)	-

Pythia options are same as used for the standard SM simulation in CMS

Appendix H

HEL parametrisation

This Appendix contains the cross section and branching fraction scaling functions used in the HEL parametrisation. The cross section scaling functions are defined at the granularity of the STXS stage 0, 1.0 and 1.1 binning schemes (see Appendix A). Only prefactors (A_p, B_{pr}) which are at least 0.1% of the leading prefactor are shown in the tables. Prefactors calculated with an MC statistical uncertainty $> 50\%$ are dropped from the parametrisation; these are typically the cross terms, B_{pr} (for $p \neq r$), for two subleading operators. The considered operators are those shown in Table 7.1.

Table H.1: Scaling functions for the STXS stage 0 bins in the HEL parametrisation.

STXS bin	Scaling function, μ^i_{prod}
ggH	$1 + 8.73 \times 10^3\, c_G + 1.95 \times 10^7\, c_G^2$
qqH	$1 + 9.02\, c_{WW} + 0.6\, c_B - 0.797\, c_{HW} + 0.474\, c_A + 171\, c_{WW}^2 + 3.42\, c_B^2 + 114\, c_{HW}^2 + 23.1\, c_{WW} c_B + 233\, c_{WW} c_{HW} + 6.22\, c_{WW} c_A + 15.3\, c_B c_{HW} + 2.02\, c_B c_A + 0.679\, c_{HW} c_A$
WH lep	$1 + 42.5\, c_{WW} + 19.9\, c_{HW} + 912\, c_{WW}^2 + 558\, c_{HW}^2 + 1.3 \times 10^3\, c_{WW} c_{HW}$
ZH lep	$1 + 36.6\, c_{WW} + 10.5\, c_B + 15\, c_{HW} + 5.21\, c_A + 602\, c_{WW}^2 + 51.7\, c_B^2 + 321\, c_{HW}^2 + 350\, c_{WW} c_B + 772\, c_{WW} c_{HW} + 102\, c_{WW} c_A + 227\, c_B c_{HW} + 31.4\, c_B c_A + 29.7\, c_{HW} c_A$
ttH	$1 + 2.95\, c_u + 115\, c_G + 2.14\, c_u^2 + 6.13\, c_{WW}^2 + 1\, c_B^2 + 5.87\, c_{HW}^2 + 2.97 \times 10^4\, c_G^2 + 167\, c_u c_G - 0.31\, c_{WW} c_B + 11.9\, c_{WW} c_{HW} - 0.318\, c_B c_{HW}$

Table H.2: Scaling functions for the STXS stage 1.0 bins in the HEL parametrisation. Units of p_T^H, m_{jj}, p_T^{Hjj} and p_T^V in GeV.

STXS bin	Scaling function, μ^i_{prod}
ggH 0J	$1 + 8.69 \times 10^3\, c_G + 1.9 \times 10^7\, c_G^2$
ggH 1J $p_T^H < 60$	$1 + 8.65 \times 10^3\, c_G + 1.93 \times 10^7\, c_G^2$
ggH 1J $60 < p_T^H < 120$	$1 + 8.39 \times 10^3\, c_G + 1.91 \times 10^7\, c_G^2$
ggH 1J $120 < p_T^H < 200$	$1 + 8.1 \times 10^3\, c_G + 1.97 \times 10^7\, c_G^2$
ggH 1J $p_T^H > 200$	$1 + 9.88 \times 10^3\, c_G + 2.01 \times 10^7\, c_G^2$
ggH $\geq$2J $p_T^H < 60$	$1 + 8.62 \times 10^3\, c_G + 2 \times 10^7\, c_G^2$
ggH $\geq$2J $60 < p_T^H < 120$	$1 + 8.94 \times 10^3\, c_G + 1.97 \times 10^7\, c_G^2$
ggH $\geq$2J $120 < p_T^H < 200$	$1 + 9.29 \times 10^3\, c_G + 2.01 \times 10^7\, c_G^2$
ggH $\geq$2J $p_T^H > 200$	$1 + 9.31 \times 10^3\, c_G + 2.11 \times 10^7\, c_G^2$
ggH VBF-like $p_T^{Hjj} < 25$	$1 + 8.59 \times 10^3\, c_G + 2.05 \times 10^7\, c_G^2$
ggH VBF-like $p_T^{Hjj} > 25$	$1 + 9.28 \times 10^3\, c_G + 2.24 \times 10^7\, c_G^2$
qqH VBF-like $p_T^{Hjj} < 25$	$1 + 2.55\, c_{WW} + 0.078\, c_B - 4.58\, c_{HW} + 0.158\, c_A + 8.63\, c_{WW}^2 + 0.117\, c_B^2 + 13.2\, c_{HW}^2 + 0.834\, c_{WW} c_B + 2.39\, c_{WW} c_{HW} - 0.253\, c_{WW} c_A + 0.341\, c_B c_{HW} - 0.711\, c_{HW} c_A$
qqH VBF-like $p_T^{Hjj} > 25$	$1 + 1.51\, c_{WW} + 0.029\, c_B - 5.88\, c_{HW} + 0.316\, c_A + 14.2\, c_{WW}^2 + 0.213\, c_B^2 + 20.9\, c_{HW}^2 + 0.961\, c_{WW} c_B + 6.94\, c_{WW} c_{HW} + 0.363\, c_{WW} c_A + 0.597\, c_B c_{HW} + 0.332\, c_B c_A - 0.9\, c_{HW} c_A$
qqH VH-like	$1 + 34.9\, c_{WW} + 2.87\, c_B + 17\, c_{HW} + 1.11\, c_A + 499\, c_{WW}^2 + 9.75\, c_B^2 + 235\, c_{HW}^2 + 68.1\, c_{WW} c_B + 619\, c_{WW} c_{HW} + 22.7\, c_{WW} c_A + 38.4\, c_B c_{HW} + 6.79\, c_B c_A + 7.74\, c_{HW} c_A$
qqH Rest	$1 + 9.51\, c_{WW} + 0.585\, c_B - 0.801\, c_{HW} + 0.316\, c_A + 77.9\, c_{WW}^2 + 1.55\, c_B^2 + 31.9\, c_{HW}^2 + 10.6\, c_{WW} c_B + 69.4\, c_{WW} c_{HW} + 5.04\, c_{WW} c_A + 4.48\, c_B c_{HW} + 1.56\, c_B c_A + 0.647\, c_{HW} c_A$
qqH $p_T^{j1} > 200$ (BSM)	$1 - 4.89\, c_{WW} - 0.566\, c_B - 10\, c_{HW} + 1.22 \times 10^3\, c_{WW}^2 + 25.4\, c_B^2 + 1.12 \times 10^3\, c_{HW}^2 + 166\, c_{WW} c_B + 2.14 \times 10^3\, c_{WW} c_{HW} + 19.8\, c_{WW} c_A + 146\, c_B c_{HW} + 6.95\, c_B c_A - 4.37\, c_{HW} c_A$
WH lep $p_T^V < 150$	$1 + 33.7\, c_{WW} + 11.5\, c_{HW} + 312\, c_{WW}^2 + 62.3\, c_{HW}^2 + 230\, c_{WW} c_{HW}$
WH lep 0J $150 < p_T^V < 250$	$1 + 76.3\, c_{WW} + 51.1\, c_{HW} + 1.56 \times 10^3\, c_{WW}^2 + 841\, c_{HW}^2 + 2.14 \times 10^3\, c_{WW} c_{HW}$
WH lep $\geq$1J $150 < p_T^V < 250$	$1 + 67.4\, c_{WW} + 42.2\, c_{HW} + 1.44 \times 10^3\, c_{WW}^2 + 752\, c_{HW}^2 + 1.92 \times 10^3\, c_{WW} c_{HW}$
WH lep $p_T^V > 250$	$1 + 183\, c_{WW} + 157\, c_{HW} + 1.52 \times 10^4\, c_{WW}^2 + 1.32 \times 10^4\, c_{HW}^2 + 2.78 \times 10^4\, c_{WW} c_{HW}$
ZH lep $p_T^V < 150$	$1 + 29.6\, c_{WW} + 8.35\, c_B + 8.48\, c_{HW} + 5.05\, c_A + 242\, c_{WW}^2 + 19.9\, c_B^2 + 34.8\, c_{HW}^2 + 136\, c_{WW} c_B + 148\, c_{WW} c_{HW} + 78.8\, c_{WW} c_A + 42.2\, c_B c_{HW} + 24.4\, c_B c_A + 18.2\, c_{HW} c_A$
ZH lep 0J $150 < p_T^V < 250$	$1 + 62.2\, c_{WW} + 18.1\, c_B + 37.6\, c_{HW} + 6\, c_A + 1.04 \times 10^3\, c_{WW}^2 + 88\, c_B^2 + 482\, c_{HW}^2 + 602\, c_{WW} c_B + 1.29 \times 10^3\, c_{WW} c_{HW} + 185\, c_{WW} c_A + 374\, c_B c_{HW} + 55.8\, c_B c_A + 70.7\, c_{HW} c_A$
ZH lep $\geq$1J $150 < p_T^V < 250$	$1 + 56.9\, c_{WW} + 16.5\, c_B + 32\, c_{HW} + 6\, c_A + 914\, c_{WW}^2 + 78.6\, c_B^2 + 422\, c_{HW}^2 + 533\, c_{WW} c_B + 1.11 \times 10^3\, c_{WW} c_{HW} + 176\, c_{WW} c_A + 324\, c_B c_{HW} + 53.5\, c_B c_A + 61\, c_{HW} c_A$
ZH lep $p_T^V > 250$	$1 + 151\, c_{WW} + 44.7\, c_B + 125\, c_{HW} + 6.32\, c_A + 9.16 \times 10^3\, c_{WW}^2 + 811\, c_B^2 + 7.68 \times 10^3\, c_{HW}^2 + 37.4\, c_A^2 + 5.45 \times 10^3\, c_{WW} c_B + 1.64 \times 10^4\, c_{WW} c_{HW} + 479\, c_{WW} c_A + 4.87 \times 10^3\, c_B c_{HW} + 144\, c_B c_A + 219\, c_{HW} c_A$
ttH	$1 + 2.95\, c_u + 115\, c_G + 2.14\, c_u^2 + 6.13\, c_{WW}^2 + 1\, c_B^2 + 5.87\, c_{HW}^2 + 2.97 \times 10^4\, c_G^2 + 167\, c_u c_G - 0.31\, c_{WW} c_B + 11.9\, c_{WW} c_{HW} - 0.318\, c_B c_{HW}$

Table H.3: Scaling functions for the ggH and qqH STXS stage 1.1 bins in the HEL parametrisation. Units of p_T^H, m_{jj}, p_T^{Hjj} and p_T^V in GeV.

STXS bin	Scaling function, μ^i_{prod}
ggH BSM $p_T^H > 200$)	$1 + 9.41 \times 10^3\, c_G + 2.09 \times 10^7\, c_G^2$
ggH 0J $p_T^H < 10$	$1 + 8.54 \times 10^3\, c_G + 1.88 \times 10^7\, c_G^2$
ggH 0J $10 < p_T^H < 200$	$1 + 8.72 \times 10^3\, c_G + 1.9 \times 10^7\, c_G^2$
ggH 1J $p_T^H < 60$	$1 + 8.65 \times 10^3\, c_G + 1.93 \times 10^7\, c_G^2$
ggH 1J $60 < p_T^H < 120$	$1 + 8.39 \times 10^3\, c_G + 1.91 \times 10^7\, c_G^2$
ggH 1J $120 < p_T^H < 200$	$1 + 8.1 \times 10^3\, c_G + 1.97 \times 10^7\, c_G^2$
ggH $\geq$2J $p_T^H < 60$	$1 + 8.74 \times 10^3\, c_G + 2.03 \times 10^7\, c_G^2$
ggH $\geq$2J $60 < p_T^H < 120$	$1 + 9.02 \times 10^3\, c_G + 1.98 \times 10^7\, c_G^2$
ggH $\geq$2J $120 < p_T^H < 200$	$1 + 9.24 \times 10^3\, c_G + 2.01 \times 10^7\, c_G^2$
ggH $\geq$2J $350 < m_{jj} < 700, p_T^{Hjj} < 25$	$1 + 7.52 \times 10^3\, c_G + 1.95 \times 10^7\, c_G^2$
ggH $\geq$2J $350 < m_{jj} < 700, p_T^{Hjj} > 25$	$1 + 9.24 \times 10^3\, c_G + 2.03 \times 10^7\, c_G^2$
ggH $\geq$2J $m_{jj} > 700, p_T^{Hjj} < 25$	$1 + 1.08 \times 10^4\, c_G + 2.13 \times 10^7\, c_G^2$
ggH $\geq$2J $m_{jj} > 700, p_T^{Hjj} > 25$	$1 + 8.34 \times 10^3\, c_G + 2.28 \times 10^7\, c_G^2$
qqH 0J	$1 + 12.8\, c_{WW} + 0.609\, c_B + 1.15\, c_{HW} + 0.316\, c_A + 57.5\, c_{WW}^2 + 3.67\, c_{HW}^2 + 5.6\, c_{WW} c_B + 25.8\, c_{WW} c_{HW} + 3.62\, c_{WW} c_A + 1.72\, c_B c_{HW} + 2.18\, c_B c_A + 1.04\, c_{HW} c_A$
qqH 1J	$1 + 12.5\, c_{WW} + 0.849\, c_B + 0.464\, c_{HW} + 0.474\, c_A + 110\, c_{WW}^2 + 2.09\, c_B^2 + 50\, c_{HW}^2 + 14.7\, c_{WW} c_B + 117\, c_{WW} c_{HW} + 6.21\, c_{WW} c_A + 6.74\, c_B c_{HW} + 1.85\, c_B c_A + 1.18\, c_{HW} c_A$
qqH $m_{jj} < 60$	$1 + 21.9\, c_{WW} + 1.58\, c_B + 7.77\, c_{HW} + 0.947\, c_A + 249\, c_{WW}^2 + 3.54\, c_B^2 + 108\, c_{HW}^2 + 24.3\, c_{WW} c_B + 283\, c_{WW} c_{HW} + 9.79\, c_{WW} c_A + 13.8\, c_B c_{HW} + 3.52\, c_B c_A + 2.81\, c_{HW} c_A$
qqH $60 < m_{jj} < 120$ (VH-like)	$1 + 38\, c_{WW} + 3.16\, c_B + 19.9\, c_{HW} + 1.26\, c_A + 668\, c_{WW}^2 + 13.7\, c_B^2 + 376\, c_{HW}^2 + 94.9\, c_{WW} c_B + 919\, c_{WW} c_{HW} + 25.3\, c_{WW} c_A + 62\, c_B c_{HW} + 7.56\, c_B c_A + 8.87\, c_{HW} c_A$
qqH $120 < m_{jj} < 350$	$1 + 6.29\, c_{WW} + 0.335\, c_B - 2.55\, c_{HW} + 0.316\, c_A + 120\, c_{WW}^2 + 2.53\, c_B^2 + 86.2\, c_{HW}^2 + 16.9\, c_{WW} c_B + 160\, c_{WW} c_{HW} + 5.32\, c_{WW} c_A + 10.7\, c_B c_{HW} + 1.75\, c_B c_A$
qqH $p_T^H > 200$ (BSM)	$1 - 16.1\, c_{WW} - 1.53\, c_B - 27.3\, c_{HW} + 0.474\, c_A + 526\, c_{WW}^2 + 10.7\, c_B^2 + 624\, c_{HW}^2 + 68.1\, c_{WW} c_B + 1.02 \times 10^3\, c_{WW} c_{HW} + 4.15\, c_{WW} c_A + 66.8\, c_B c_{HW} + 1.8\, c_B c_A - 12.2\, c_{HW} c_A$
qqH $350 < m_{jj} < 700, p_T^{Hjj} < 25$	$1 + 2.09\, c_{WW} - 0.012\, c_B - 4.06\, c_{HW} + 0.158\, c_A + 12.5\, c_{WW}^2 + 0.331\, c_B^2 + 10.1\, c_{HW}^2 + 1.27\, c_{WW} c_B + 6.63\, c_{WW} c_{HW} + 0.947\, c_{WW} c_A + 0.772\, c_B c_{HW} + 0.6\, c_B c_A$
qqH $350 < m_{jj} < 700, p_T^{Hjj} > 25$	$1 + 2.19\, c_{WW} - 0.028\, c_B - 3.91\, c_{HW} + 44.9\, c_{WW}^2 + 0.713\, c_B^2 + 27\, c_{HW}^2 + 4.35\, c_{WW} c_B + 22\, c_{WW} c_{HW} + 1.96\, c_{WW} c_A + 2.01\, c_B c_{HW} + 0.837\, c_B c_A - 0.505\, c_{HW} c_A$
qqH $m_{jj} > 700, p_T^{Hjj} < 25$	$1 + 2.5\, c_{WW} + 0.087\, c_B - 3.7\, c_{HW} + 14.4\, c_{WW}^2 + 0.186\, c_B^2 + 8.53\, c_{HW}^2 + 1.7\, c_{WW} c_B + 0.42\, c_{WW} c_{HW} + 0.33\, c_B c_{HW} + 0.284\, c_B c_A + 0.158\, c_{HW} c_A$
qqH $m_{jj} > 700, p_T^{Hjj} > 25$	$1 + 1.12\, c_{WW} + 0.058\, c_B - 4.47\, c_{HW} + 0.316\, c_A + 17.7\, c_{WW}^2 + 0.414\, c_B^2 + 10.2\, c_{HW}^2 + 2.08\, c_{WW} c_B + 12.3\, c_{WW} c_{HW} + 1.14\, c_{WW} c_A + 0.975\, c_B c_{HW} + 0.632\, c_B c_A$

Table H.4: Scaling functions for the VH lep and ttH STXS stage 1.1 bins in the HEL parametrisation. Units of p_T^H, m_{jj}, p_T^{Hjj} and p_T^V in GeV.

STXS bin	Scaling function, μ^i_{prod}
WH lep $p_T^V < 75$	$1 + 28.9\,c_{WW} + 7.38\,c_{HW} + 219\,c_{WW}^2 + 19.5\,c_{HW}^2 + 114\,c_{WW}c_{HW}$
WH lep $75 < p_T^V < 150$	$1 + 41.4\,c_{WW} + 18\,c_{HW} + 457\,c_{WW}^2 + 129\,c_{HW}^2 + 412\,c_{WW}c_{HW}$
WH lep 0J $150 < p_T^V < 250$	$1 + 76.3\,c_{WW} + 51.1\,c_{HW} + 1.56\times10^3\,c_{WW}^2 + 841\,c_{HW}^2 + 2.14\times10^3\,c_{WW}c_{HW}$
WH lep $\geq$1J $150 < p_T^V < 250$	$1 + 67.4\,c_{WW} + 42.2\,c_{HW} + 1.44\times10^3\,c_{WW}^2 + 752\,c_{HW}^2 + 1.92\times10^3\,c_{WW}c_{HW}$
WH lep $p_T^V > 250$	$1 + 183\,c_{WW} + 157\,c_{HW} + 1.52\times10^4\,c_{WW}^2 + 1.32\times10^4\,c_{HW}^2 + 2.78\times10^4\,c_{WW}c_{HW}$
ZH lep $p_T^V < 75$	$1 + 25.7\,c_{WW} + 7.2\,c_B + 5.52\,c_{HW} + 4.74\,c_A + 175\,c_{WW}^2 + 14.3\,c_B^2 + 10.4\,c_{HW}^2 + 97.5\,c_{WW}c_B + 74.5\,c_{WW}c_{HW} + 64.2\,c_{WW}c_A + 21\,c_Bc_{HW} + 20.1\,c_Bc_A + 12.8\,c_{HW}c_A$
ZH lep $75 < p_T^V < 150$	$1 + 35.5\,c_{WW} + 10.1\,c_B + 13\,c_{HW} + 5.37\,c_A + 345\,c_{WW}^2 + 28.6\,c_B^2 + 72.2\,c_{HW}^2 + 196\,c_{WW}c_B + 260\,c_{WW}c_{HW} + 101\,c_{WW}c_A + 74.6\,c_Bc_{HW} + 31\,c_Bc_A + 26.5\,c_{HW}c_A$
ZH lep 0J $150 < p_T^V < 250$	$1 + 62.2\,c_{WW} + 18.1\,c_B + 37.6\,c_{HW} + 6\,c_A + 1.04\times10^3\,c_{WW}^2 + 88\,c_B^2 + 482\,c_{HW}^2 + 602\,c_{WW}c_B + 1.29\times10^3\,c_{WW}c_{HW} + 185\,c_{WW}c_A + 374\,c_Bc_{HW} + 55.8\,c_Bc_A + 70.7\,c_{HW}c_A$
ZH lep $\geq$1J $150 < p_T^V < 250$	$1 + 56.9\,c_{WW} + 16.5\,c_B + 32\,c_{HW} + 6\,c_A + 914\,c_{WW}^2 + 78.6\,c_B^2 + 422\,c_{HW}^2 + 533\,c_{WW}c_B + 1.11\times10^3\,c_{WW}c_{HW} + 176\,c_{WW}c_A + 324\,c_Bc_{HW} + 53.5\,c_Bc_A + 61\,c_{HW}c_A$
ZH lep $p_T^V > 250$	$1 + 151\,c_{WW} + 44.7\,c_B + 125\,c_{HW} + 6.32\,c_A + 9.16\times10^3\,c_{WW}^2 + 811\,c_B^2 + 7.68\times10^3\,c_{HW}^2 + 37.4\,c_A^2 + 5.45\times10^3\,c_{WW}c_B + 1.64\times10^4\,c_{WW}c_{HW} + 479\,c_{WW}c_A + 4.87\times10^3\,c_Bc_{HW} + 144\,c_Bc_A + 219\,c_{HW}c_A$
ttH	$1 + 2.95\,c_u + 115\,c_G + 2.14\,c_u^2 + 6.13\,c_{WW}^2 + 1\,c_B^2 + 5.87\,c_{HW}^2 + 2.97\times10^4\,c_G^2 + 167\,c_uc_G - 0.31\,c_{WW}c_B + 11.9\,c_{WW}c_{HW} - 0.318\,c_Bc_{HW}$

Table H.5: Scaling functions for the partial and total Higgs boson width in the HEL parametrisation.

Decay channel	Partial width scaling, $\Gamma^f/\Gamma^f_{\text{SM}}$
$H \to \gamma\gamma$	$1 - 916\,c_A + 2.09\times10^5\,c_A^2$
$H \to ZZ \to 4\ell$	$1 + 10.2\,c_{WW} + 2.77\,c_B + 2.87\,c_{HW} + 2.88\,c_A + 27.8\,c_{WW}^2 + 3.52\,c_B^2 + 2.21\,c_{HW}^2 + 10\,c_{WW}c_B + 14.7\,c_{WW}c_{HW} + 3.84\,c_Bc_{HW} + 8.26\,c_Bc_A$
$H \to WW \to 2\ell2\nu$	$1 + 10.3\,c_{WW} + 3.72\,c_{HW} + 26.3\,c_{WW}^2 + 3.77\,c_{HW}^2 + 19.2\,c_{WW}c_{HW}$
$H \to \tau\tau$	$1 + 3\,c_\ell + 2.25\,c_\ell^2$
$H \to bb$	$1 + 3\,c_d + 2.25\,c_d^2$
Total width scaling	$1 + 0.171\,c_u + 2.33\,c_d + 0.112\,c_\ell + 1.03\,c_{WW} + 0.0228\,c_B + 0.366\,c_{HW} + 259\,c_G + 0.128\,c_u^2 + 1.74\,c_d^2 + 0.084\,c_\ell^2 + 2.63\,c_{WW}^2 + 4.67\,c_{HW}^2 + 5.73\times10^5\,c_G^2 + 0.11\,c_{WW}c_B + 1.88\,c_{WW}c_{HW} + 0.03\,c_Bc_{HW} + 15.8\,c_{HW}c_A$

Appendix I

Warsaw-basis parametrisation

This Appendix contains the cross section and branching fraction scaling functions used in the Warsaw-basis parametrisation. The cross section scaling functions are defined at the granularity of the STXS stage 1.2 binning scheme (see Appendix A). Only prefactors (A_p, B_{pr}) which are at least 5% of the leading prefactor are shown in the tables. This is a tighter threshold than shown for the HEL parametrisation due to the larger number of terms. In the fits, the full parametrisation without this threshold is used. Prefactors calculated with an MC statistical uncertainty $> 50\%$ are dropped from the parametrisation; these are typically the cross terms, B_{pr} (for $p \neq r$), for two subleading operators. The considered operators are those shown in Table 7.6. The parametrisation is defined for a nominal new physics energy scale of $\Lambda = 1$ TeV.

Table I.1: Scaling functions for the ggH STXS stage 1.2 bins in the Warsaw basis. Units of p_T^H, m_{jj}, p_T^{Hjj} and p_T^V in GeV.

STXS bin	Scaling function, μ^i_{prod}
ggH BSM $200 < p_T^H < 300$	$1 + 19.1\,C_{HG} + 164\,C_{HG}^2$
ggH BSM $300 < p_T^H < 450$	$1 + 18\,C_{HG} + 161\,C_{HG}^2$
ggH BSM $450 < p_T^H < 650$	$1 + 16.9\,C_{HG} + 161\,C_{HG}^2$
ggH BSM $p_T^H > 650$	$1 + 17.7\,C_{HG} + 172\,C_{HG}^2$
ggH 0J $p_T^H < 10$	$1 + 34.4\,C_{HG} + 308\,C_{HG}^2$
ggH 0J $10 < p_T^H < 200$	$1 + 35\,C_{HG} + 313\,C_{HG}^2$
ggH 1J $p_T^H < 60$	$1 + 32.7\,C_{HG} + 275\,C_{HG}^2$
ggH 1J $60 < p_T^H < 120$	$1 + 30.2\,C_{HG} + 238\,C_{HG}^2$
ggH 1J $120 < p_T^H < 200$	$1 + 29.5\,C_{HG} + 235\,C_{HG}^2$
ggH $\geq$2J $p_T^H < 60$	$1 + 20.9\,C_{HG} + 132\,C_{HG}^2$
ggH $\geq$2J $60 < p_T^H < 120$	$1 + 21.4\,C_{HG} + 150\,C_{HG}^2$
ggH $\geq$2J $120 < p_T^H < 200$	$1 + 21.3\,C_{HG} + 167\,C_{HG}^2$
ggH $\geq$2J $350 < m_{jj} < 700, p_T^{Hjj} < 25$	$1 + 15.9\,C_{HG} + 109\,C_{HG}^2$
ggH $\geq$2J $350 < m_{jj} < 700, p_T^{Hjj} > 25$	$1 + 13.6\,C_{HG} + 81.6\,C_{HG}^2$
ggH $\geq$2J $m_{jj} > 700, p_T^{Hjj} < 25$	$1 + 16.3\,C_{HG} + 119\,C_{HG}^2$
ggH $\geq$2J $m_{jj} > 700, p_T^{Hjj} > 25$	$1 + 12.7\,C_{HG} + 75.3\,C_{HG}^2$

Table I.2: Scaling functions for the qqH STXS stage 1.2 bins in the Warsaw basis. Units of p_T^H, m_{jj}, p_T^{Hjj} and p_T^V in GeV.

STXS bin	Scaling function, μ^i_{prod}
qqH 0J	$1 + 0.121\,C_{H\text{Box}} + 0.239\,C_{HW} + 0.0393\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} + 0.34\,C^{(3)}_{Hq} + 0.182\,C^{(1)}_{\ell\ell} + 0.0603\,C^2_{HW} + 0.0279\,C^2_{HB} + 0.0147\,C^2_{HWB} + 0.0331\,C^{(3)\,2}_{H\ell} + 0.0269\,C^{(1)\,2}_{Hq} + 0.125\,C^{(3)\,2}_{Hq} + 0.0138\,C^2_{Hu} + 0.0128\,C^2_{Hd} + 0.0083\,C^{(1)\,2}_{\ell\ell} + 0.0146\,C_{H\text{Box}}C_{HW} - 0.0221\,C_{H\text{Box}}C^{(3)}_{H\ell} + 0.0207\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.011\,C_{H\text{Box}}C^{(1)}_{\ell\ell} + 0.0205\,C_{HW}C_{HB} - 0.0435\,C_{HW}C^{(3)}_{H\ell} + 0.157\,C_{HW}C^{(3)}_{Hq} + 0.0218\,C_{HW}C^{(1)}_{\ell\ell} - 0.0314\,C_{HB}C_{HWB} + 0.0152\,C_{HWB}C^{(3)}_{Hq} - 0.062\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.0331\,C^{(3)}_{H\ell}C^{(1)}_{\ell\ell} + 0.031\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
qqH 1J	$1 + 0.121\,C_{H\text{Box}} + 0.223\,C_{HW} + 0.0482\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} + 0.354\,C^{(3)}_{Hq} + 0.0261\,C_{Hu} + 0.182\,C^{(1)}_{\ell\ell} + 0.102\,C^2_{HW} + 0.241\,C^{(1)\,2}_{Hq} + 1.01\,C^{(3)\,2}_{Hq} + 0.084\,C^2_{Hu} + 0.252\,C_{HW}C^{(3)}_{Hq} - 0.0646\,C^{(3)}_{H\ell}C^{(3)}_{Hq}$
qqH $m_{jj} < 60$	$1 + 0.121\,C_{H\text{Box}} + 0.355\,C_{HW} + 0.0577\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} + 0.778\,C^{(3)}_{Hq} + 0.0548\,C_{Hu} + 0.182\,C^{(1)}_{\ell\ell} + 0.155\,C^2_{HW} + 0.179\,C^{(1)\,2}_{Hq} + 0.737\,C^{(3)\,2}_{Hq} + 0.132\,C^2_{Hu} + 0.0616\,C^2_{Hd} + 0.0474\,C_{H\text{Box}}C^{(3)}_{Hq} - 0.0649\,C_{HW}C^{(3)}_{H\ell} + 0.443\,C_{HW}C^{(3)}_{Hq} + 0.0425\,C_{HWB}C^{(3)}_{Hq} - 0.142\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.0548\,C^{(1)}_{Hq}C^{(3)}_{Hq} + 0.0709\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
qqH $60 < m_{jj} < 120$ (VH-like)	$1 + 0.121\,C_{H\text{Box}} + 0.552\,C_{HW} + 0.088\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} + 1.72\,C^{(3)}_{Hq} + 0.13\,C_{Hu} + 0.182\,C^{(1)}_{\ell\ell} + 0.256\,C^2_{HW} + 0.667\,C^{(1)\,2}_{Hq} + 2.24\,C^{(3)\,2}_{Hq} + 0.398\,C^2_{Hu} + 0.245\,C^2_{Hd} + 0.883\,C_{HW}C^{(3)}_{Hq} - 0.313\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.317\,C^{(1)}_{Hq}C^{(3)}_{Hq} + 0.156\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
qqH $120 < m_{jj} < 350$	$1 + 0.121\,C_{H\text{Box}} + 0.0213\,C_{HW} + 0.0354\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} - 0.0845\,C^{(3)}_{Hq} + 0.182\,C^{(1)}_{\ell\ell} + 0.118\,C^2_{HW} + 0.0331\,C^{(3)\,2}_{H\ell} + 0.121\,C^{(1)\,2}_{Hq} + 0.538\,C^{(3)\,2}_{Hq} + 0.0725\,C^2_{Hu} + 0.203\,C_{HW}C^{(3)}_{Hq} - 0.0331\,C^{(3)}_{H\ell}C^{(1)}_{\ell\ell}$
qqH $p_T^H > 200$ (BSM)	$1 + 0.121\,C_{H\text{Box}} + 0.191\,C_{HW} - 0.364\,C^{(3)}_{H\ell} + 0.0635\,C^{(1)}_{Hq} - 1.25\,C^{(3)}_{Hq} - 0.0838\,C_{Hu} + 0.182\,C^{(1)}_{\ell\ell} + 0.503\,C^2_{HW} + 1.19\,C^{(1)\,2}_{Hq} + 4.48\,C^{(3)\,2}_{Hq} + 0.821\,C^2_{Hu} + 0.274\,C^2_{Hd} + 0.464\,C_{HW}C^{(3)}_{Hq} + 0.227\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.866\,C^{(1)}_{Hq}C^{(3)}_{Hq}$
qqH $350 < m_{jj} < 700, p_T^{Hjj} < 25$	$1 + 0.121\,C_{H\text{Box}} - 0.126\,C_{HW} + 0.0242\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} - 0.361\,C^{(3)}_{Hq} - 0.0195\,C_{Hu} + 0.182\,C^{(1)}_{\ell\ell} + 0.0518\,C^2_{HW} + 0.0122\,C^2_{HB} + 0.0138\,C^2_{HWB} + 0.0331\,C^{(3)\,2}_{H\ell} + 0.0184\,C^{(1)\,2}_{Hq} + 0.0856\,C^{(3)\,2}_{Hq} + 0.0117\,C^2_{Hu} + 0.0078\,C^2_{Hd} + 0.0083\,C^{(1)\,2}_{\ell\ell} - 0.0074\,C_{H\text{Box}}C_{HW} - 0.022\,C_{H\text{Box}}C^{(3)}_{H\ell} - 0.0218\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.011\,C_{H\text{Box}}C^{(1)}_{\ell\ell} + 0.0073\,C_{HW}C_{HB} - 0.0179\,C_{HW}C_{HWB} + 0.0227\,C_{HW}C^{(3)}_{H\ell} + 0.0846\,C_{HW}C^{(3)}_{Hq} - 0.0113\,C_{HW}C^{(1)}_{\ell\ell} - 0.0137\,C_{HB}C_{HWB} + 0.0105\,C_{HWB}C_{Hu} + 0.0655\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.0331\,C^{(3)}_{H\ell}C^{(1)}_{\ell\ell} - 0.0045\,C^{(1)}_{Hq}C^{(3)}_{Hq} + 0.0046\,C^{(3)}_{Hq}C_{Hu} - 0.0328\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
qqH $350 < m_{jj} < 700, p_T^{Hjj} > 25$	$1 + 0.121\,C_{H\text{Box}} - 0.12\,C_{HW} + 0.0289\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} - 0.325\,C^{(3)}_{Hq} + 0.182\,C^{(1)}_{\ell\ell} + 0.0827\,C^2_{HW} + 0.0141\,C^2_{HB} + 0.0141\,C^2_{HWB} + 0.033\,C^{(3)\,2}_{H\ell} + 0.0472\,C^{(1)\,2}_{Hq} + 0.184\,C^{(3)\,2}_{Hq} + 0.0223\,C^2_{Hu} + 0.0151\,C^2_{Hd} - 0.022\,C_{H\text{Box}}C^{(3)}_{H\ell} - 0.0196\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.011\,C_{H\text{Box}}C^{(1)}_{\ell\ell} - 0.0185\,C_{HW}C_{HWB} + 0.0215\,C_{HW}C^{(3)}_{H\ell} + 0.16\,C_{HW}C^{(3)}_{Hq} - 0.0108\,C_{HW}C^{(1)}_{\ell\ell} - 0.0127\,C_{HB}C_{HWB} + 0.0106\,C_{HWB}C_{Hu} + 0.059\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.033\,C^{(3)}_{H\ell}C^{(1)}_{\ell\ell} - 0.0205\,C^{(1)}_{Hq}C^{(3)}_{Hq} - 0.0295\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
qqH $m_{jj} > 700, p_T^{Hjj} < 25$	$1 + 0.121\,C_{H\text{Box}} - 0.124\,C_{HW} + 0.0229\,C_{HWB} - 0.364\,C^{(3)}_{H\ell} - 0.368\,C^{(3)}_{Hq} - 0.023\,C_{Hu} + 0.182\,C^{(1)}_{\ell\ell} + 0.0457\,C^2_{HW} + 0.0125\,C^2_{HB} + 0.0122\,C^2_{HWB} + 0.0331\,C^{(3)\,2}_{H\ell} + 0.0154\,C^{(1)\,2}_{Hq} + 0.0798\,C^{(3)\,2}_{Hq} + 0.0089\,C^2_{Hu} + 0.0085\,C^2_{Hd} + 0.0083\,C^{(1)\,2}_{\ell\ell} - 0.0073\,C_{H\text{Box}}C_{HW} - 0.022\,C_{H\text{Box}}C^{(3)}_{H\ell} - 0.0222\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.011\,C_{H\text{Box}}C^{(1)}_{\ell\ell} + 0.0074\,C_{HW}C_{HB} - 0.0177\,C_{HW}C_{HWB} + 0.0224\,C_{HW}C^{(3)}_{H\ell} - 0.0065\,C_{HW}C^{(1)}_{Hq} + 0.081\,C_{HW}C^{(3)}_{Hq} - 0.0112\,C_{HW}C^{(1)}_{\ell\ell} - 0.0134\,C_{HB}C_{HWB} + 0.0047\,C_{HWB}C^{(1)}_{Hq} - 0.0062\,C_{HWB}C^{(3)}_{Hq} + 0.0083\,C_{HWB}C_{Hu} - 0.0044\,C_{HWB}C_{Hd} + 0.0668\,C^{(3)}_{H\ell}C^{(3)}_{Hq} + 0.0042\,C^{(3)}_{H\ell}C_{Hu} - 0.0331\,C^{(3)}_{H\ell}C^{(1)}_{\ell\ell} - 0.0082\,C^{(1)}_{Hq}C^{(3)}_{Hq} - 0.0334\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
qqH $m_{jj} > 700, p_T^{Hjj} > 25$	$1 + 0.121\,C_{H\text{Box}} - 0.165\,C_{HW} + 0.0247\,C_{HWB} - 0.363\,C^{(3)}_{H\ell} - 0.427\,C^{(3)}_{Hq} - 0.026\,C_{Hu} + 0.182\,C^{(1)}_{\ell\ell} + 0.0737\,C^2_{HW} + 0.0147\,C^2_{HB} + 0.0143\,C^2_{HWB} + 0.033\,C^{(3)\,2}_{H\ell} + 0.0264\,C^{(1)\,2}_{Hq} + 0.128\,C^{(3)\,2}_{Hq} + 0.0142\,C^2_{Hu} + 0.0118\,C^2_{Hd} + 0.0083\,C^{(1)\,2}_{\ell\ell} - 0.0097\,C_{H\text{Box}}C_{HW} - 0.022\,C_{H\text{Box}}C^{(3)}_{H\ell} - 0.0257\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.011\,C_{H\text{Box}}C^{(1)}_{\ell\ell} + 0.0077\,C_{HW}C_{HB} - 0.0214\,C_{HW}C_{HWB} + 0.0298\,C_{HW}C^{(3)}_{H\ell} - 0.0114\,C_{HW}C^{(1)}_{Hq} + 0.141\,C_{HW}C^{(3)}_{Hq} - 0.0149\,C_{HW}C^{(1)}_{\ell\ell} - 0.0139\,C_{HB}C_{HWB} + 0.0079\,C_{HWB}C^{(1)}_{Hq} - 0.0079\,C_{HWB}C^{(3)}_{Hq} + 0.0118\,C_{HWB}C_{Hu} + 0.0774\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.033\,C^{(3)}_{H\ell}C^{(1)}_{\ell\ell} - 0.0156\,C^{(1)}_{Hq}C^{(3)}_{Hq} - 0.0387\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$

Table I.3: Scaling functions for the VH leptonic STXS stage 1.2 bins in the Warsaw basis. Units of p_T^H, m_{jj}, p_T^{Hjj} and p_T^V in GeV.

STXS bin	Scaling function, μ^i_{prod}
WH lep $p_T^V < 75$	$1 + 0.121\,C_{H\text{Box}} + 0.804\,C_{HW} - 0.243\,C^{(3)}_{H\ell} + 1.12\,C^{(3)}_{Hq} + 0.182\,C^{(1)}_{\ell\ell} + 0.192\,C^2_{HW} + 0.36\,C^{(3)\,2}_{Hq} + 0.0488\,C_{H\text{Box}}C_{HW} + 0.0679\,C_{H\text{Box}}C^{(3)}_{Hq} - 0.0971\,C_{HW}C^{(3)}_{H\ell} + 0.491\,C_{HW}C^{(3)}_{Hq} + 0.0731\,C_{HW}C^{(1)}_{\ell\ell} - 0.136\,C^{(3)}_{H\ell}C^{(3)}_{Hq} + 0.102\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
WH lep $75 < p_T^V < 150$	$1 + 0.121\,C_{H\text{Box}} + 0.922\,C_{HW} - 0.242\,C^{(3)}_{H\ell} + 1.87\,C^{(3)}_{Hq} + 0.182\,C^{(1)}_{\ell\ell} + 0.324\,C^2_{HW} + 1.04\,C^{(3)\,2}_{Hq} + 0.0559\,C_{H\text{Box}}C_{HW} + 0.113\,C_{H\text{Box}}C^{(3)}_{Hq} - 0.111\,C_{HW}C^{(3)}_{H\ell} + 0.938\,C_{HW}C^{(3)}_{Hq} + 0.0838\,C_{HW}C^{(1)}_{\ell\ell} - 0.226\,C^{(3)}_{H\ell}C^{(3)}_{Hq} + 0.17\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
WH lep 0J $150 < p_T^V < 250$	$1 + 1.02\,C_{HW} - 0.242\,C^{(3)}_{H\ell} + 3.96\,C^{(3)}_{Hq} + 0.684\,C^2_{HW} + 4.53\,C^{(3)\,2}_{Hq} + 0.24\,C_{H\text{Box}}C^{(3)}_{Hq} + 2.17\,C_{HW}C^{(3)}_{Hq} - 0.479\,C^{(3)}_{H\ell}C^{(3)}_{Hq} + 0.36\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
WH lep $\geq$1J $150 < p_T^V < 250$	$1 + 1.02\,C_{HW} - 0.242\,C^{(3)}_{H\ell} + 3.5\,C^{(3)}_{Hq} + 0.182\,C^{(1)}_{\ell\ell} + 0.632\,C^2_{HW} + 4.01\,C^{(3)\,2}_{Hq} + 0.212\,C_{H\text{Box}}C^{(3)}_{Hq} + 1.96\,C_{HW}C^{(3)}_{Hq} - 0.424\,C^{(3)}_{H\ell}C^{(3)}_{Hq} + 0.318\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
WH lep $p_T^V > 250$	$1 + 1.07\,C_{HW} + 10.9\,C^{(3)}_{Hq} + 58.2\,C^{(3)\,2}_{Hq} + 6.21\,C_{HW}C^{(3)}_{Hq}$
ZH lep $p_T^V < 75$	$1 + 0.122\,C_{H\text{Box}} + 0.663\,C_{HW} + 0.0817\,C_{HB} + 0.297\,C_{HWB} - 0.0747\,C^{(1)}_{H\ell} - 0.241\,C^{(3)}_{H\ell} + 1.13\,C^{(3)}_{Hq} + 0.255\,C_{Hu} - 0.101\,C_{Hd} + 0.182\,C^{(1)}_{\ell\ell} + 0.142\,C^2_{HW} + 0.0279\,C^2_{HWB} + 0.394\,C^{(1)\,2}_{Hq} + 0.394\,C^{(3)\,2}_{Hq} + 0.221\,C^2_{Hu} + 0.172\,C^2_{Hd} + 0.0407\,C_{H\text{Box}}C_{HW} + 0.0693\,C_{H\text{Box}}C^{(3)}_{Hq} - 0.0299\,C_{HDD}C_{Hu} + 0.106\,C_{HW}C_{HWB} - 0.0251\,C_{HW}C^{(1)}_{H\ell} - 0.0798\,C_{HW}C^{(3)}_{H\ell} - 0.0494\,C_{HW}C^{(1)}_{Hq} + 0.456\,C_{HW}C^{(3)}_{Hq} + 0.0605\,C_{HW}C^{(1)}_{\ell\ell} + 0.0394\,C_{HB}C^{(1)}_{Hq} + 0.0914\,C_{HB}C_{Hu} - 0.0357\,C_{HB}C_{Hd} - 0.0338\,C_{HWB}C^{(3)}_{H\ell} + 0.171\,C_{HWB}C^{(3)}_{Hq} + 0.0594\,C_{HWB}C_{Hu} - 0.023\,C_{HWB}C_{Hd} + 0.0271\,C_{HWB}C^{(1)}_{\ell\ell} - 0.0425\,C^{(1)}_{H\ell}C^{(3)}_{Hq} - 0.136\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.03\,C^{(3)}_{H\ell}C_{Hu} - 0.0922\,C^{(1)}_{Hq}C^{(3)}_{Hq} + 0.103\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell} + 0.0229\,C_{Hu}C^{(1)}_{\ell\ell}$
ZH lep $75 < p_T^V < 150$	$1 + 0.122\,C_{H\text{Box}} + 0.768\,C_{HW} + 0.0912\,C_{HB} + 0.335\,C_{HWB} - 0.241\,C^{(3)}_{H\ell} + 1.8\,C^{(3)}_{Hq} + 0.425\,C_{Hu} - 0.157\,C_{Hd} + 0.182\,C^{(1)}_{\ell\ell} + 0.239\,C^2_{HW} + 1.05\,C^{(1)\,2}_{Hq} + 1.05\,C^{(3)\,2}_{Hq} + 0.608\,C^2_{Hu} + 0.441\,C^2_{Hd} + 0.111\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.16\,C_{HW}C_{HWB} - 0.0925\,C_{HW}C^{(3)}_{H\ell} - 0.114\,C_{HW}C^{(1)}_{Hq} + 0.853\,C_{HW}C^{(3)}_{Hq} + 0.0703\,C_{HW}C^{(1)}_{\ell\ell} + 0.0765\,C_{HB}C^{(1)}_{Hq} + 0.182\,C_{HB}C_{Hu} - 0.067\,C_{HB}C_{Hd} + 0.301\,C_{HWB}C^{(3)}_{Hq} + 0.126\,C_{HWB}C_{Hu} - 0.0644\,C^{(1)}_{H\ell}C^{(3)}_{Hq} - 0.218\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 0.291\,C^{(1)}_{Hq}C^{(3)}_{Hq} + 0.165\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
ZH lep 0J $150 < p_T^V < 250$	$1 + 0.868\,C_{HW} + 0.373\,C_{HWB} - 0.241\,C^{(3)}_{H\ell} - 0.232\,C^{(1)}_{Hq} + 3.71\,C^{(3)}_{Hq} + 0.912\,C_{Hu} - 0.322\,C_{Hd} + 0.513\,C^2_{HW} + 4.39\,C^{(1)\,2}_{Hq} + 4.39\,C^{(3)\,2}_{Hq} + 2.62\,C^2_{Hu} + 1.8\,C^2_{Hd} + 0.229\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.304\,C_{HW}C_{HWB} - 0.337\,C_{HW}C^{(1)}_{Hq} + 2\,C_{HW}C^{(3)}_{Hq} + 0.441\,C_{HB}C_{Hu} + 0.666\,C_{HWB}C^{(3)}_{Hq} + 0.317\,C_{HWB}C_{Hu} - 0.449\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 1.51\,C^{(1)}_{Hq}C^{(3)}_{Hq} + 0.34\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
ZH lep $\geq$1J $150 < p_T^V < 250$	$1 + 0.863\,C_{HW} + 0.372\,C_{HWB} - 0.241\,C^{(3)}_{H\ell} - 0.293\,C^{(1)}_{Hq} + 3.27\,C^{(3)}_{Hq} + 0.818\,C_{Hu} - 0.278\,C_{Hd} + 0.182\,C^{(1)}_{\ell\ell} + 0.47\,C^2_{HW} + 3.81\,C^{(1)\,2}_{Hq} + 3.81\,C^{(3)\,2}_{Hq} + 2.28\,C^2_{Hu} + 1.46\,C^2_{Hd} + 0.202\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.279\,C_{HW}C_{HWB} - 0.314\,C_{HW}C^{(1)}_{Hq} + 1.79\,C_{HW}C^{(3)}_{Hq} + 0.42\,C_{HB}C_{Hu} + 0.595\,C_{HWB}C^{(3)}_{Hq} + 0.303\,C_{HWB}C_{Hu} - 0.395\,C^{(3)}_{H\ell}C^{(3)}_{Hq} - 1.51\,C^{(1)}_{Hq}C^{(3)}_{Hq} + 0.299\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
ZH lep $p_T^V > 250$	$1 + 0.883\,C_{HW} - 1.08\,C^{(1)}_{Hq} + 9.57\,C^{(3)}_{Hq} + 2.37\,C_{Hu} - 0.835\,C_{Hd} + 45\,C^{(1)\,2}_{Hq} + 45\,C^{(3)\,2}_{Hq} + 28\,C^2_{Hu} + 18.6\,C^2_{Hd} + 5.64\,C_{HW}C^{(3)}_{Hq} - 23\,C^{(1)}_{Hq}C^{(3)}_{Hq}$

Table I.4: Scaling functions for the ttH, tH and bbH STXS stage 1.2 bins in the Warsaw basis. Units of p_T^H, m_{jj}, p_T^{Hjj} and p_T^V in GeV.

STXS bin	Scaling function, μ^i_{prod}
ttH $p_T^H < 60$	$1 + 0.119\,C_{H\text{Box}} + 0.417\,C_{HG} - 0.119\,\lvert C_{uH}\rvert - 0.737\,\lvert C_{uG}\rvert - 0.12\,C^{(3)}_{H\ell} + 0.0599\,C^{(1)}_{\ell\ell} + 0.162\,C^2_{HG} + 0.391\,\lvert C_{uG}\rvert^2 + 0.026\,C_{H\text{Box}}C_{HG} - 0.0447\,C_{H\text{Box}}\lvert C_{uG}\rvert - 0.0255\,C_{HG}\lvert C_{uH}\rvert - 0.291\,C_{HG}\lvert C_{uG}\rvert - 0.025\,C_{HG}C^{(3)}_{H\ell} + 0.0446\,\lvert C_{uH}\rvert\lvert C_{uG}\rvert + 0.0444\,\lvert C_{uG}\rvert C^{(3)}_{H\ell} - 0.0222\,\lvert C_{uG}\rvert C^{(1)}_{\ell\ell}$
ttH $60 < p_T^H < 120$	$1 + 0.119\,C_{H\text{Box}} + 0.483\,C_{HG} - 0.119\,\lvert C_{uH}\rvert - 0.803\,\lvert C_{uG}\rvert - 0.12\,C^{(3)}_{H\ell} + 0.0598\,C^{(1)}_{\ell\ell} + 0.242\,C^2_{HG} + 0.52\,\lvert C_{uG}\rvert^2 + 0.0303\,C_{H\text{Box}}C_{HG} - 0.0487\,C_{H\text{Box}}\lvert C_{uG}\rvert - 0.0295\,C_{HG}\lvert C_{uH}\rvert - 0.397\,C_{HG}\lvert C_{uG}\rvert - 0.0288\,C_{HG}C^{(3)}_{H\ell} + 0.0485\,\lvert C_{uH}\rvert\lvert C_{uG}\rvert + 0.0483\,\lvert C_{uG}\rvert C^{(3)}_{H\ell}$
ttH $120 < p_T^H < 200$	$1 + 0.118\,C_{H\text{Box}} + 0.594\,C_{HG} - 0.118\,\lvert C_{uH}\rvert - 0.897\,\lvert C_{uG}\rvert - 0.119\,C^{(3)}_{H\ell} + 0.0596\,C^{(1)}_{\ell\ell} + 0.47\,C^2_{HG} + 0.823\,\lvert C_{uG}\rvert^2 - 0.0544\,C_{H\text{Box}}\lvert C_{uG}\rvert - 0.651\,C_{HG}\lvert C_{uG}\rvert + 0.054\,\lvert C_{uH}\rvert\lvert C_{uG}\rvert + 0.0538\,\lvert C_{uG}\rvert C^{(3)}_{H\ell}$
ttH $200 < p_T^H < 300$	$1 + 0.117\,C_{H\text{Box}} + 0.719\,C_{HG} - 0.117\,\lvert C_{uH}\rvert - 0.979\,\lvert C_{uG}\rvert - 0.119\,C^{(3)}_{H\ell} + 0.0593\,C^{(1)}_{\ell\ell} + 1.11\,C^2_{HG} + 1.61\,\lvert C_{uG}\rvert^2 - 1.25\,C_{HG}\lvert C_{uG}\rvert$
ttH $p_T^H > 300$	$1 + 0.115\,C_{H\text{Box}} + 0.874\,C_{HG} - 0.116\,\lvert C_{uH}\rvert - 0.991\,\lvert C_{uG}\rvert - 0.118\,C^{(3)}_{H\ell} + 0.0589\,C^{(1)}_{\ell\ell} + 4.38\,C^2_{HG} + 3.69\,\lvert C_{uG}\rvert^2 + 0.44\,C^{(1)\,2}_{Hq} + 0.441\,C^{(3)\,2}_{Hq} + 0.278\,C^2_{Hu} - 2.78\,C_{HG}\lvert C_{uG}\rvert - 0.372\,C^{(1)}_{Hq}C^{(3)}_{Hq}$
tHq	$1 + 0.121\,C_{H\text{Box}} - 0.0303\,C_{HDD} + 0.214\,C_{HW} - 0.364\,C^{(3)}_{H\ell} + 0.118\,C^{(3)}_{Hq} + 0.182\,C^{(1)}_{\ell\ell} + 2.72\,C^{(3)\,2}_{Hq} + 0.281\,C_{HW}C^{(3)}_{Hq}$
tHW	$1 + 0.12\,C_{H\text{Box}} - 0.03\,C_{HDD} + 0.48\,C_{HG} + 0.159\,C_{HW} - 0.226\,\lvert C_{uG}\rvert - 0.241\,C^{(3)}_{H\ell} + 0.358\,C^{(3)}_{Hq} + 0.12\,C^{(1)}_{\ell\ell} + 0.355\,C^2_{HG} + 0.0279\,C^2_{HW} + 0.391\,\lvert C_{uG}\rvert^2 + 0.194\,C^{(3)\,2}_{Hq} + 0.027\,C_{H\text{Box}}C_{HG} + 0.021\,C_{H\text{Box}}C^{(3)}_{Hq} + 0.1\,C_{HG}C_{HW} - 0.314\,C_{HG}\lvert C_{uG}\rvert - 0.0561\,C_{HG}C^{(3)}_{H\ell} + 0.262\,C_{HG}C^{(3)}_{Hq} + 0.0281\,C_{HG}C^{(1)}_{\ell\ell} - 0.0723\,C_{HW}\lvert C_{uG}\rvert + 0.0867\,C_{HW}C^{(3)}_{Hq} + 0.0265\,\lvert C_{uG}\rvert C^{(3)}_{H\ell} - 0.205\,\lvert C_{uG}\rvert C^{(3)}_{Hq} - 0.0427\,C^{(3)}_{H\ell}C^{(3)}_{Hq} + 0.0213\,C^{(3)}_{Hq}C^{(1)}_{\ell\ell}$
bbH	$1 + 0.0453\,C_{H\text{Box}} - 0.008\,C_{HDD} + 8.88\,C_{HG} + 0.036\,C_{HW} + 0.0174\,C_{HWB} - 0.0569\,C^{(3)}_{H\ell} + 0.103\,C^{(3)}_{Hq} + 0.0241\,C_{Hu} - 0.008\,C_{Hd} + 0.0284\,C^{(1)}_{\ell\ell} + 68.1\,C^2_{HG} + 0.014\,C^2_{HW} + 0.216\,C^{(1)\,2}_{Hq} + 0.224\,C^{(3)\,2}_{Hq} + 0.0646\,C^2_{Hu} + 0.0285\,C^2_{Hd} + 0.0233\,C_{HG}C^{(3)}_{H\ell} + 0.066\,C_{HW}C^{(3)}_{Hq} + 0.0226\,C_{HWB}C^{(3)}_{Hq} - 0.0189\,C^{(3)}_{H\ell}C^{(3)}_{Hq}$

Table I.5: Scaling functions for the partial and total Higgs boson width in the Warsaw basis. Prefactors (A_p, B_{pr}) which are at least 1% of the leading prefactor are shown in the table.

Decay channel	Partial width scaling, $\Gamma^f/\Gamma^f_{\rm SM}$
$H \to \gamma\gamma$	$1 - 14\,C_{HW} - 48.8\,C_{HB} + 26.1\,C_{HWB} + 49\,C_{HW}^2 + 596\,C_{HB}^2 + 171\,C_{HWB}^2 + 342\,C_{HW}C_{HB} - 183\,C_{HW}C_{HWB} - 638\,C_{HB}C_{HWB}$
$H \to ZZ \to 4\ell$	$1 + 0.12\,C_{H\Box} - 0.234\,C_{HW} - 0.23\,C_{HB} + 0.244\,C_{HWB} + 0.126\,C_{H\ell}^{(1)} - 0.236\,C_{H\ell}^{(3)} - 0.101\,C_{He} + 0.181\,C_{\ell\ell}^{(1)} + 0.0184\,C_{HDD}^2 + 0.883\,C_{HW}^2 + 4.03\,C_{HB}^2 + 1.18\,C_{HWB}^2 + 0.0164\,C_{H\ell}^{(3)2} + 0.0082\,C_{\ell\ell}^{(1)2} - 0.0142\,C_{H\Box}C_{H\ell}^{(3)} + 0.0109\,C_{H\Box}C_{\ell\ell}^{(1)} + 0.0711\,C_{HDD}C_{HW} + 0.741\,C_{HW}C_{HB} - 1.76\,C_{HW}C_{HWB} - 2.86\,C_{HB}C_{HWB} - 0.0099\,C_{H\ell}^{(1)}C_{H\ell}^{(3)} + 0.0115\,C_{H\ell}^{(1)}C_{\ell\ell}^{(1)} + 0.0173\,C_{H\ell}^{(3)}C_{He} - 0.0214\,C_{H\ell}^{(3)}C_{\ell\ell}^{(1)}$
$H \to WW \to 2\ell 2\nu$	$1 + 0.121\,C_{H\Box} - 0.0314\,C_{HDD} - 0.0889\,C_{HW} + 0.0028\,C_{H\ell}^{(1)} - 0.228\,C_{H\ell}^{(3)} - 0.0046\,C_{He} + 0.182\,C_{\ell\ell}^{(1)} + 0.0037\,C_{H\Box}^2 + 0.0048\,C_{HW}^2 + 0.001\,C_{HWB}^2 + 0.0132\,C_{H\ell}^{(3)2} + 0.0083\,C_{\ell\ell}^{(1)2} - 0.0019\,C_{H\Box}C_{HDD} - 0.0054\,C_{H\Box}C_{HW} - 0.0138\,C_{H\Box}C_{H\ell}^{(3)} + 0.011\,C_{H\Box}C_{\ell\ell}^{(1)} + 0.0023\,C_{HDD}C_{HW} + 0.004\,C_{HDD}C_{H\ell}^{(3)} - 0.0028\,C_{HDD}C_{\ell\ell}^{(1)} + 0.0099\,C_{HW}C_{H\ell}^{(3)} - 0.0081\,C_{HW}C_{\ell\ell}^{(1)} - 0.0207\,C_{H\ell}^{(3)}C_{\ell\ell}^{(1)}$
$H \to bb$	$1 + 0.121\,C_{H\Box} - 0.0303\,C_{HDD} - 0.121\,\lvert C_{dH}\rvert - 0.121\,C_{H\ell}^{(3)} + 0.0606\,C_{\ell\ell}^{(1)} + 0.0037\,C_{H\Box}^2 + 0.0037\,\lvert C_{dH}\rvert^2 + 0.0037\,C_{H\ell}^{(3)2} - 0.0018\,C_{H\Box}C_{HDD} - 0.0074\,C_{H\Box}\lvert C_{dH}\rvert - 0.0074\,C_{H\Box}C_{H\ell}^{(3)} + 0.0037\,C_{H\Box}C_{\ell\ell}^{(1)} + 0.0018\,C_{HDD}\lvert C_{dH}\rvert + 0.0018\,C_{HDD}C_{H\ell}^{(3)} + 0.0074\,\lvert C_{dH}\rvert C_{H\ell}^{(3)} - 0.0037\,\lvert C_{dH}\rvert C_{\ell\ell}^{(1)} - 0.0037\,C_{H\ell}^{(3)}C_{\ell\ell}^{(1)}$
$H \to \tau\tau$	$1 + 0.121\,C_{H\Box} - 0.0303\,C_{HDD} - 0.121\,\lvert C_{eH}\rvert - 0.121\,C_{H\ell}^{(3)} + 0.0606\,C_{\ell\ell}^{(1)} + 0.0037\,C_{H\Box}^2 + 0.0037\,\lvert C_{eH}\rvert^2 + 0.0037\,C_{H\ell}^{(3)2} - 0.0018\,C_{H\Box}C_{HDD} - 0.0074\,C_{H\Box}\lvert C_{eH}\rvert - 0.0074\,C_{H\Box}C_{H\ell}^{(3)} + 0.0037\,C_{H\Box}C_{\ell\ell}^{(1)} + 0.0018\,C_{HDD}\lvert C_{eH}\rvert + 0.0018\,C_{HDD}C_{H\ell}^{(3)} + 0.0074\,\lvert C_{eH}\rvert C_{H\ell}^{(3)} - 0.0037\,\lvert C_{eH}\rvert C_{\ell\ell}^{(1)} - 0.0037\,C_{H\ell}^{(3)}C_{\ell\ell}^{(1)}$
$H \to \mu\mu$	$1 + 0.121\,C_{H\Box} - 0.0303\,C_{HDD} - 0.121\,C_{H\ell}^{(3)} + 0.0606\,C_{\ell\ell}^{(1)} + 0.0037\,C_{H\Box}^2 + 0.0037\,C_{H\ell}^{(3)2} - 0.0018\,C_{H\Box}C_{HDD} - 0.0074\,C_{H\Box}C_{H\ell}^{(3)} + 0.0037\,C_{H\Box}C_{\ell\ell}^{(1)} + 0.0018\,C_{HDD}C_{H\ell}^{(3)} - 0.0037\,C_{H\ell}^{(3)}C_{\ell\ell}^{(1)}$
$H \to Z\gamma$	$1 - 14.7\,C_{HW} + 14.7\,C_{HB} + 9.8\,C_{HWB} + 54.2\,C_{HW}^2 + 54.2\,C_{HB}^2 + 24\,C_{HWB}^2 - 108\,C_{HW}C_{HB} - 72.2\,C_{HW}C_{HWB} + 72.2\,C_{HB}C_{HWB}$
Total width scaling	$1 + 0.117\,C_{H\Box} - 0.0291\,C_{HDD} + 1.37\,C_{HG} - 0.0444\,C_{HW} - 0.0705\,C_{HB} + 0.0571\,C_{HWB} - 0.0058\,\lvert C_{eH}\rvert - 0.0087\,\lvert C_{uH}\rvert - 0.0974\,\lvert C_{dH}\rvert - 0.123\,C_{H\ell}^{(3)} + 0.0026\,C_{Hq}^{(3)} + 0.0631\,C_{\ell\ell}^{(1)} + 0.0035\,C_{H\Box}^2 + 13.4\,C_{HG}^2 + 0.146\,C_{HW}^2 + 1.12\,C_{HB}^2 + 0.329\,C_{HWB}^2 + 0.003\,\lvert C_{dH}\rvert^2 + 0.0043\,C_{H\ell}^{(3)2} - 0.0018\,C_{H\Box}C_{HDD} - 0.0059\,C_{H\Box}\lvert C_{dH}\rvert - 0.0075\,C_{H\Box}C_{H\ell}^{(3)} + 0.0038\,C_{H\Box}C_{\ell\ell}^{(1)} + 0.0015\,C_{HDD}\lvert C_{dH}\rvert + 0.0019\,C_{HDD}C_{H\ell}^{(3)} + 0.487\,C_{HW}C_{HB} - 0.404\,C_{HW}C_{HWB} - 1.05\,C_{HB}C_{HWB} + 0.0059\,\lvert C_{dH}\rvert C_{H\ell}^{(3)} - 0.003\,\lvert C_{dH}\rvert C_{\ell\ell}^{(1)} - 0.0044\,C_{H\ell}^{(3)}C_{\ell\ell}^{(1)}$

Appendix J

Input feature distributions for the e/γ identification BDT in the HGCAL L1T

This Appendix shows the distributions of the nine input features ($\vec{x}$) for the e/γ identification BDT algorithm introduced in Section 8.3. The distributions are shown for both signal (red) and background (black) trigger primitive clusters, for the low $|\eta|$ ($1.5 < |\eta| < 2.7$, solid) and the high $|\eta|$ ($2.7 < |\eta| < 3.0$, dashed) regions separately.

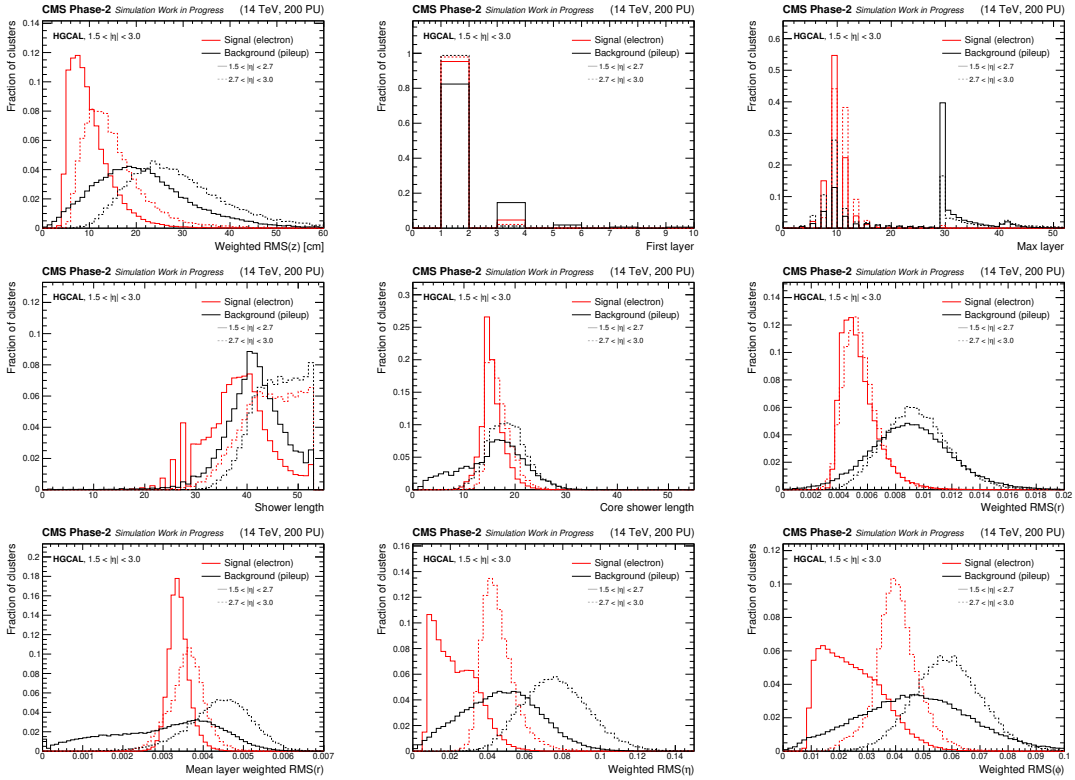


Figure J.1: The distributions of the input features to the HGCAL L1T e/γ identification BDT algorithm. The first five plots represent the five longitudinal shower shape variables, whilst the final four plots shows the four lateral shower shape variables. The signal clusters (generator-level matched electron) and background clusters (pileup) are shown in red and black respectively, and are separated into the two pseudorapidity regions in which the BDTs are trained.

Bibliography

[1] CMS Collaboration. "Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s} = 13$ TeV". *Submitted to JHEP* (2021). arXiv: 2103.06956 [hep-ex].

[2] CMS Collaboration. "Combined Higgs boson production and decay measurements with up to 137 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-HIG-19-005. (2020).

[3] CMS Collaboration. "The Phase-2 Upgrade of the CMS Level-1 Trigger". Tech. rep. CERN-LHCC-2020-004. CMS-TDR-021. (2020).

[4] M. Cepeda et al. "Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC". *CERN Yellow Rep. Monogr.* 7 (2019). Ed. by Andrea Dainese et al., pp. 221–584. DOI: 10.23731/CYRM-2019-007.221. arXiv: 1902.00134 [hep-ph].

[5] S. L. Glashow. "Partial Symmetries of Weak Interactions". *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[6] Steven Weinberg. "A Model of Leptons". *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[7] Abdus Salam. "Weak and Electromagnetic Interactions". *Conf. Proc.* C680519 (1968), pp. 367–377.

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

[9] P.W. Higgs. "Broken symmetries, massless particles and gauge fields". *Physics Letters* 12.2 (1964), pp. 132–133. DOI: https://doi.org/10.1016/0031-9163(64)91136-9.

[10] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". *Phys. Rev. Lett.* 13 (1964), pp. 508–509. 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), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[12] Peter W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". *Phys. Rev.* 145.4 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.

[13] T. W. B. Kibble. "Symmetry Breaking in Non-Abelian Gauge Theories". *Phys. Rev.* 155.5 (1967), pp. 1554–1561. DOI: 10.1103/PhysRev.155.1554.

[14] 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.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

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

[16] CMS Collaboration. "Observation of a New Boson with Mass Near 125 GeV in pp Collisions at $\sqrt{s} = 7$ and 8 TeV". *JHEP* 06 (2013), p. 081. DOI: 10.1007/JHEP06(2013)081. arXiv: 1303.4571 [hep-ex].

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

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

[19] Lyndon Evans and Philip Bryant. "LHC Machine". *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[20] N. Aghanim et al. "Planck 2018 results. VI. Cosmological parameters". *Astron. Astrophys.* 641 (2020), A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

[21] Steven Weinberg. "Implications of dynamical symmetry breaking". *Phys. Rev. D* 13.4 (1976), pp. 974–996. DOI: 10.1103/PhysRevD.13.974.

[22] Leonard Susskind. "Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory". *Phys. Rev. D* 20.10 (1979), pp. 2619–2625. DOI: 10.1103/PhysRevD.20.2619.

[23] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

[24] P.A. Zyla et al. "Review of Particle Physics". *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[25] K. Kajantie et al. "The Electroweak phase transition: A Nonperturbative analysis". *Nucl. Phys. B* 466 (1996), pp. 189–258. DOI: 10.1016/0550-3213(96)00052-1. arXiv: hep-lat/9510020.

[26] F. Csikor, Z. Fodor, and J. Heitger. "Endpoint of the hot electroweak phase transition". *Phys. Rev. Lett.* 82 (1999), pp. 21–24. DOI: 10.1103/PhysRevLett.82.21. arXiv: hep-ph/9809291.

[27] Stephen P. Martin. "A Supersymmetry primer". *Adv. Ser. Direct. High Energy Phys.* 21 (2010). Ed. by Gordon L. Kane, pp. 1–153. DOI: 10.1142/9789812839657_0001. arXiv: hep-ph/9709356.

[28] Oliver Witzel. "Review on Composite Higgs Models". *PoS* LATTICE2018 (2019), p. 006. DOI: 10.22323/1.334.0006. arXiv: 1901.08216 [hep-lat].

[29] Mariano Quiros. "Higgs Bosons in Extra Dimensions". *Mod. Phys. Lett. A* 30.15 (2015), p. 1540012. DOI: 10.1142/S021773231540012X. arXiv: 1311.2824 [hep-ph].

[30] ATLAS Collaboration. "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector". *Phys. Lett. B* 784 (2018), pp. 173–191. DOI: 10.1016/j.physletb.2018.07.035. arXiv: 1806.00425 [hep-ex].

[31] ATLAS Collaboration. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". *Phys. Lett. B* 786 (2018), pp. 59–86. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[32] ATLAS Collaboration. "Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". *Phys. Rev. D* 99 (2019), p. 072001. DOI: 10.1103/PhysRevD.99.072001. arXiv: 1811.08856 [hep-ex].

[33] CMS Collaboration. "Observation of $t\bar{t}H$ production". *Phys. Rev. Lett.* 120.23 (2018), p. 231801. DOI: 10.1103/PhysRevLett.120.231801. arXiv: 1804.02610 [hep-ex].

[34] CMS Collaboration. "Observation of Higgs boson decay to bottom quarks". *Phys. Rev. Lett.* 121.12 (2018), p. 121801. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex].

[35] CMS Collaboration. "Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector". *Phys. Lett. B* 779 (2018), pp. 283–316. DOI: 10.1016/j.physletb.2018.02.004. arXiv: 1708.00373 [hep-ex].

[36] CMS Collaboration. "Evidence for Higgs boson decay to a pair of muons" (2020). arXiv: 2009.04363 [hep-ex].

[37] ATLAS Collaboration. "Measurements of the Higgs boson inclusive and differential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV". *Eur. Phys. J. C* 80.10 (2020), p. 942. DOI: 10.1140/epjc/s10052-020-8223-0. arXiv: 2004.03969 [hep-ex].

[38] ATLAS Collaboration. "Measurements and interpretations of Higgs-boson fiducial cross sections in the diphoton decay channel using 139 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector". Tech. rep. ATLAS-CONF-2019-029. (2019).

[39] ATLAS Collaboration. "Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector" (2020). arXiv: 2007.02873 [hep-ex].

[40] CMS Collaboration. "Inclusive search for highly boosted Higgs bosons decaying to bottom quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV". *JHEP* 12 (2020), p. 085. DOI: 10.1007/JHEP12(2020)085. arXiv: 2006.13251 [hep-ex].

[41] CMS Collaboration. "Measurement of the inclusive and differential Higgs boson production cross sections in the leptonic WW decay mode at $\sqrt{s} = 13$ TeV". *JHEP* 03 (2021), p. 003. DOI: 10.1007/JHEP03(2021)003. arXiv: 2007.01984 [hep-ex].

[42] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". *CERN* (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[43] W. Buchmüller and D. Wyler. "Effective lagrangian analysis of new interactions and flavour conservation". *Nuclear Physics B* 268.3 (1986), pp. 621–653. DOI: https://doi.org/10.1016/0550-3213(86)90262-2.

[44] Kaoru Hagiwara, R. Szalapski, and D. Zeppenfeld. "Anomalous Higgs boson production and decay". *Phys. Lett.* B318 (1993), pp. 155–162. DOI: 10.1016/0370-2693(93)91799-S. arXiv: hep-ph/9308347 [hep-ph].

[45] Gian Francesco Giudice et al. "The strongly-interacting light Higgs". *JHEP* 2007.06 (2007), pp. 045–045. DOI: 10.1088/1126-6708/2007/06/045.

[46] B. Grzadkowski et al. "Dimension-Six Terms in the Standard Model Lagrangian". *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[47] Roberto Contino et al. "Effective Lagrangian for a light Higgs-like scalar". *JHEP* 07 (2013), p. 035. DOI: 10.1007/JHEP07(2013)035. arXiv: 1303.3876 [hep-ph].

[48] CMS Collaboration. "Constraints on the Higgs boson self-coupling from ttH + tH, H $\rightarrow \gamma\gamma$ differential measurements at the HL-LHC". Tech. rep. CMS-PAS-FTR-18-020. (2018).

[49] CMS Collaboration. "A measurement of the Higgs boson mass in the diphoton decay channel". *Phys. Lett. B* 805 (2020), p. 135425. DOI: 10.1016/j.physletb.2020.135425. arXiv: 2002.06398 [hep-ex].

[50] Emmy Noether. "Invariant variation problems". *Transport Theory and Statistical Physics* 1.3 (1971), pp. 186–207. DOI: 10.1080/00411457108231446.

[51] I. J. R. Aitchison and A. J. G. Hey. "Gauge theories in particle physics: A practical introduction. Vol. 1: From relativistic quantum mechanics to QED". CRC Press, (2012).

[52] Paul Adrien Maurice Dirac and Ralph Howard Fowler. "The quantum theory of the electron". *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.778 (1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023.

[53] I. J. R. Aitchison and A. J. G. Hey. "Gauge theories in particle physics: A practical introduction. Vol. 2: Non-Abelian gauge theories: QCD and the electroweak theory". CRC Press, (2012).

[54] F. Halzen and Alan D. Martin. "Quarks and Leptons: An Introductory Course in Modern Particle Physics". Wiley, (1984).

[55] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". *Phys. Rev. Lett.* 30.26 (1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[56] Gavin P. Salam. "Towards Jetography". *Eur. Phys. J. C* 67 (2010), pp. 637–686. DOI: 10.1140/epjc/s10052-010-1314-6. arXiv: 0906.1833 [hep-ph].

[57] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". *Phys. Rev.* 105.4 (1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413.

[58] Mark Thomson. "Modern Particle Physics". Cambridge University Press, (2013).

[59] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. "Broken Symmetries". *Phys. Rev.* 127 (1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965.

[60] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". *Phys. Rev. Lett.* 10 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531.

[61] Makoto Kobayashi and Toshihide Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[62] B. Pontecorvo. "Inverse beta processes and nonconservation of lepton charge". *Zh. Eksp. Teor. Fiz.* 34 (1957), p. 247.

[63] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the unified model of elementary particles". *Prog. Theor. Phys.* 28 (1962), pp. 870–880. DOI: 10.1143/PTP.28.870.

[64] R.D. Peccei and Helen Quinn. "CP Conservation in the Presence of Pseudoparticles". *Phys. Rev. Lett.* 38 (1977), pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440.

[65] ATLAS Collaboration. "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment". *Phys. Rev. D* 101.1 (2020), p. 012002. DOI: 10.1103/PhysRevD.101.012002. arXiv: 1909.02845 [hep-ex].

[66] CMS Collaboration. "Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV". *JHEP* 11 (2018), p. 185. DOI: 10.1007/JHEP11(2018)185. arXiv: 1804.02716 [hep-ex].

[67] Nicolas Berger et al. "Simplified Template Cross Sections - Stage 1.1". *LHC Higgs Cross Section Working Group Report* LHCHXSWG-2019-003, DESY-19-070 (2019). arXiv: 1906.02754 [hep-ph].

[68] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet User Manual". *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[69] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[70] Enrico Fermi. "Tentativo di una teoria dell'emissione dei raggi beta". *Ric. Sci.* 4 (1933), pp. 491–495.

[71] Ilaria Brivio and Michael Trott. "The Standard Model as an Effective Field Theory". *Phys. Rept.* 793 (2019), pp. 1–98. DOI: 10.1016/j.physrep.2018.11.002. arXiv: 1706.08945 [hep-ph].

[72] Rodrigo Alonso et al. "Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology". *JHEP* 04 (2014), p. 159. DOI: 10.1007/JHEP04(2014)159. arXiv: 1312.2014 [hep-ph].

[73] Christopher W. Murphy. "Dimension-8 operators in the Standard Model Effective Field Theory". *JHEP* 10 (2020), p. 174. DOI: 10.1007/JHEP10(2020)174. arXiv: 2005.00059 [hep-ph].

[74] Steven Weinberg. "Baryon and Lepton Nonconserving Processes". *Phys. Rev. Lett.* 43 (1979), pp. 1566–1570. DOI: 10.1103/PhysRevLett.43.1566.

[75] Andre de Gouvea, Juan Herrero-Garcia, and Andrew Kobach. "Neutrino Masses, Grand Unification, and Baryon Number Violation". *Phys. Rev. D* 90.1 (2014), p. 016011. DOI: 10.1103/PhysRevD.90.016011. arXiv: 1404.4057 [hep-ph].

[76] D. Marzocca et al. "BSM Benchmarks for Effective Field Theories in Higgs and Electroweak Physics". *LHC Higgs Cross Section Working Group Report* LHC-HXSWG-2019-006 (2020). arXiv: 2009.01249 [hep-ph].

[77] John Ellis et al. "Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data". *JHEP* 06 (2018), p. 146. DOI: 10.1007/JHEP06(2018)146. arXiv: 1803.03252 [hep-ph].

[78] G. F. Giudice et al. "The Strongly-Interacting Light Higgs". *JHEP* 06 (2007), p. 045. DOI: 10.1088/1126-6708/2007/06/045. arXiv: hep-ph/0703164.

[79] Ilaria Brivio, Yun Jiang, and Michael Trott. "The SMEFTsim package, theory and tools". *JHEP* 12 (2017), p. 070. DOI: 10.1007/JHEP12(2017)070. arXiv: 1709.06492 [hep-ph].

[80] K. Aamodt et al. "The ALICE experiment at the CERN LHC". *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[81] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[82] CERN. "LEP design report". CERN-LEP-84-01, (1984).

[83] CMS Collaboration. "CMS Luminosity - Public Results". http://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. [Online; accessed 14-Apr-2021].

[84] Tai Sakuma and Thomas McCauley. "Detector and Event Visualization with SketchUp at the CMS Experiment". *Journal of Physics: Conference Series* 513.2 (2014), p. 022032. DOI: 10.1088/1742-6596/513/2/022032.

[85] CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". *JINST* 9.10 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569 [physics.ins-det].

[86] CMS Collaboration. "The Phase-2 Upgrade of the CMS Tracker". Tech. rep. CERN-LHCC-2017-009. CMS-TDR-014. (2017).

[87] CMS Collaboration. "CMS Technical Design Report for the Pixel Detector Upgrade". Tech. rep. CERN-LHCC-2012-016, CMS-TDR-011. (2012).

[88] Andrea Benaglia. "The CMS ECAL performance with examples". Tech. rep. CMS-CR-2013-430. Geneva: CERN, 2013. DOI: 10.1088/1748-0221/9/02/C02008.

[89] CMS Collaboration. "The CMS electromagnetic calorimeter project: Technical Design Report". Tech. rep. CERN-LHCC-97-033, CMS-TDR-4. (1997).

[90] CMS Collaboration. "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV". *JINST* 10.08 (2015), P08010. DOI: 10.1088/1748-0221/10/08/P08010. arXiv: 1502.02702 [physics.ins-det].

[91] CMS Collaboration. "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV". *JINST* 10.06 (2015), P06005. DOI: 10.1088/1748-0221/10/06/P06005. arXiv: 1502.02701 [physics.ins-det].

[92] CMS Collaboration. "2015 ECAL detector performance plots". CMS-DP-2015-057 (2015).

[93] CMS Collaboration. "ECAL 2016 refined calibration and Run 2 summary plots". CMS-DP-2020-021 (2020).

[94] CMS Collaboration. "The CMS hadron calorimeter project: Technical Design Report". Tech. rep. CERN-LHCC-97-031, CMS-TDR-2. (1997).

[95] CMS Collaboration. "The Performance of the CMS Muon Detector in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV at the LHC". *JINST* 8 (2013), P11002. DOI: 10.1088/1748-0221/8/11/P11002. arXiv: 1306.6905 [physics.ins-det].

[96] 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 [physics.ins-det].

[97] CMS Collaboration. "The CMS trigger system". *JINST* 12.01 (2017), P01020. DOI: 10.1088/1748-0221/12/01/P01020. arXiv: 1609.02366 [physics.ins-det].

[98] CMS Collaboration. "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger". Tech. rep. CERN-LHCC-2002-026. (2002).

[99] CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". *JINST* 12.10 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965 [physics.ins-det].

[100] P. Billoir and S. Qian. "Simultaneous pattern recognition and track fitting by the Kalman filtering method". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 294.1 (1990), pp. 219–228. DOI: https://doi.org/10.1016/0168-9002(90)91835-Y.

[101] Johann Brehmer and Kyle Cranmer. "Simulation-based inference methods for particle physics" (2020). To appear in *Artificial Intelligence for Particle Physics*, World Scientific Publising Co. arXiv: 2010.06439 [hep-ph].

[102] CMS Collaboration. "Measurements of Higgs boson properties in the diphoton decay channel at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-HIG-19-015. (2020).

[103] Stefan Höche. "Introduction to parton-shower event generators". *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*. (2015), pp. 235–295. DOI: 10.1142/9789814678766_0005. arXiv: 1411.4085 [hep-ph].

[104] Michelangelo L. Mangano and Timothy J. Stelzer. "Tools for the Simulation of Hard Hadronic Collisions". *Annual Review of Nuclear and Particle Science* 55.1 (2005), pp. 555–588. DOI: 10.1146/annurev.nucl.55.090704.151505.

[105] Raymond Brock et al. "Handbook of perturbative QCD: Version 1.0". *Rev. Mod. Phys.* 67 (1995), pp. 157–248. DOI: 10.1103/RevModPhys.67.157.

[106] Andy Buckley et al. "General-purpose event generators for LHC physics". *Phys. Rept.* 504 (2011), pp. 145–233. DOI: 10.1016/j.physrep.2011.03.005. arXiv: 1101.2599 [hep-ph].

[107] Charalampos Anastasiou et al. "High precision determination of the gluon fusion Higgs boson cross-section at the LHC". *JHEP* 05 (2016), p. 058. DOI: 10.1007/JHEP05(2016)058. arXiv: 1602.00695 [hep-ph].

[108] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[109] V. N. Gribov and L. N. Lipatov. "Deep inelastic e p scattering in perturbation theory". *Sov. J. Nucl. Phys.* 15 (1972), pp. 438–450.

[110] Yuri L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." *Sov. Phys. JETP* 46 (1977), pp. 641–653.

[111] Guido Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language". *Nucl. Phys. B* 126 (1977), pp. 298–318. DOI: 10.1016/0550-3213(77)90384-4.

[112] CMS Collaboration. "Event generator tunes obtained from underlying event and multiparton scattering measurements". *Eur. Phys. J. C* 76.3 (2016), p. 155. DOI: 10.1140/epjc/s10052-016-3988-x. arXiv: 1512.00815 [hep-ex].

[113] CMS Collaboration. "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements". *Eur. Phys. J. C* 80.1 (2020), p. 4. DOI: 10.1140/epjc/s10052-019-7499-4. arXiv: 1903.12179 [hep-ex].

[114] Bo Andersson et al. "Parton Fragmentation and String Dynamics". *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[115] Bo Andersson. "The Lund model". Vol. 7. Cambridge University Press, (2005).

[116] D. Amati and G. Veneziano. "Preconfinement as a Property of Perturbative QCD". *Phys. Lett. B* 83 (1979), pp. 87–92. DOI: 10.1016/0370-2693(79)90896-7.

[117] Johan Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". *Eur. Phys. J. C* 53 (2008), pp. 473–500. DOI: 10.1140/epjc/s10052-007-0490-5. arXiv: 0706.2569 [hep-ph].

[118] S. Agostinelli et al. "GEANT4–a simulation toolkit". *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[119] J. Alwall 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), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[120] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[121] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[122] Simone Alioli et al. "NLO Higgs boson production via gluon fusion matched with shower in POWHEG". *JHEP* 04 (2009), p. 002. DOI: 10.1088/1126-6708/2009/04/002. arXiv: 0812.0578 [hep-ph].

[123] Paolo Nason and Carlo Oleari. "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG". *JHEP* 02 (2010), p. 037. DOI: 10.1007/JHEP02(2010)037. arXiv: 0911.5299 [hep-ph].

[124] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[125] Heribertus B. Hartanto et al. "Higgs boson production in association with top quarks in the POWHEG BOX". *Phys. Rev. D* 91.9 (2015), p. 094003. DOI: 10.1103/PhysRevD.91.094003. arXiv: 1501.04498 [hep-ph].

[126] T. Gleisberg et al. "Event generation with SHERPA 1.1". *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[127] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[128] CMS Collaboration. "Measurements of Higgs boson production via gluon fusion and vector boson fusion in the diphoton decay channel at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-HIG-18-029. (2019).

[129] CMS Collaboration. "Measurements of $t\bar{t}H$ Production and the CP Structure of the Yukawa Interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel". *Phys. Rev. Lett.* 125.6 (2020), p. 061801. DOI: 10.1103/PhysRevLett.125.061801. arXiv: 2003.10866 [hep-ex].

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

[131] Keith Hamilton et al. "NNLOPS simulation of Higgs boson production". *JHEP* 10 (2013), p. 222. DOI: 10.1007/JHEP10(2013)222. arXiv: 1309.0017 [hep-ph].

[132] Richard D. Ball et al. "Parton distributions for the LHC Run II". *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[133] Richard D. Ball et al. "Parton distributions from high-precision collider data". *Eur. Phys. J. C* 77.10 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[134] E. Spyromitros-Xioufis et al. "Multi-target regression via input space expansion: treating targets as inputs". *Machine Learning* 104 (2016), p. 55. DOI: 10.1007/s10994-016-5546-z. arXiv: 1211.6581 [cs.LG].

[135] CMS Collaboration. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". *JINST* 12.02 (2017), P02014. DOI: 10.1088/1748-0221/12/02/P02014. arXiv: 1607.03663 [hep-ex].

[136] CMS Collaboration. "Pileup mitigation at CMS in 13 TeV data". *JINST* 15.09 (2020), P09018. DOI: 10.1088/1748-0221/15/09/P09018. arXiv: 2003.00503 [hep-ex].

[137] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. "The Elements of Statistical Learning". Springer New York Inc., (2001).

[138] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. "Deep Learning". The MIT Press, (2016).

[139] Christopher M. Bishop. "Pattern Recognition and Machine Learning". Springer, (2006).

[140] CMS Collaboration. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV". *JINST* 13.05 (2018), P05011. DOI: 10.1088/1748-0221/13/05/P05011. arXiv: 1712.07158 [physics.ins-det].

[141] CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with pp collisions at $\sqrt{s} = 13$ TeV". *JINST* 13.06 (2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528 [physics.ins-det].

[142] CMS Collaboration. "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector". *JINST* 14.07 (2019), P07004. DOI: 10.1088/1748-0221/14/07/P07004. arXiv: 1903.06078 [hep-ex].

[143] CMS Collaboration. "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV". *Eur. Phys. J. C* 75.5 (2015), p. 212. DOI: 10.1140/epjc/s10052-015-3351-7. arXiv: 1412.8662 [hep-ex].

[144] J R Andersen et al. "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties". *CERN* (2013). DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph].

[145] Wouter Verkerke and David P. Kirkby. "The RooFit toolkit for data modeling". *eConf* C0303241 (2003), MOLT007. arXiv: physics/0306116.

[146] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". *Eur. Phys. J. C* 71 (2011). [Erratum: *Eur. Phys. J. C* 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[147] P.D. Dauncey et al. "Handling uncertainties in background shapes: the discrete profiling method". *JINST* 10.04 (2015), P04015. DOI: 10.1088/1748-0221/10/04/P04015. arXiv: 1408.6865 [physics.data-an].

[148] R. A. Fisher. "On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P". *Journal of the Royal Statistical Society* 85.1 (1922), pp. 87–94.

[149] CMS Collaboration. "Observation of the Diphoton Decay of the Higgs Boson and Measurement of Its Properties". *Eur. Phys. J. C* 74.10 (2014), p. 3076. DOI: 10.1140/epjc/s10052-014-3076-z. arXiv: 1407.0558 [hep-ex].

[150] CMS Collaboration. "CMS Luminosity Measurements for the 2016 Data Taking Period". Tech. rep. CMS-PAS-LUM-17-001. (2017).

[151] CMS Collaboration. "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-LUM-17-004. (2018).

[152] CMS Collaboration. "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-LUM-18-002. (2019).

[153] CMS Collaboration. "Jet algorithms performance in 13 TeV data". Tech. rep. CMS-PAS-JME-16-003. (2017).

[154] K. Becker et al. "Precise predictions for boosted Higgs production". *CERN* CERN-TH-2020-074 (2020). arXiv: 2005.07762 [hep-ph].

[155] ATLAS Collaboration. "Evaluation of theoretical uncertainties for simplified template cross section measurements of V-associated production of the Higgs boson". Tech. rep. ATL-PHYS-PUB-2018-035. (2018).

[156] Sayipjamal Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". *Phys. Rev. D* 93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph].

[157] L. A. Harland-Lang et al. "Parton distributions in the LHC era: MMHT 2014 PDFs". *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph].

[158] Stefano Carrazza et al. "An Unbiased Hessian Representation for Monte Carlo PDFs". *Eur. Phys. J. C* 75.8 (2015), p. 369. DOI: 10.1140/epjc/s10052-015-3590-7. arXiv: 1505.06736 [hep-ph].

[159] Jun Gao and Pavel Nadolsky. "A meta-analysis of parton distribution functions". *JHEP* 07 (2014), p. 035. DOI: 10.1007/JHEP07(2014)035. arXiv: 1401.0013 [hep-ph].

[160] CMS Collaboration. "Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons". *JHEP* 02 (2019), p. 149. DOI: 10.1007/JHEP02(2019)149. arXiv: 1811.06625 [hep-ex].

[161] CMS Collaboration. "Measurement of the cross section ratio ttbar+bbbar/ttbar+jj using dilepton final states in pp collisions at 13 TeV". Tech. rep. CMS-PAS-TOP-16-010. (2016).

[162] ATLAS and CMS Collaborations, and the LHC Higgs Combination Group. "Procedure for the LHC Higgs boson search combination in Summer 2011". Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. (2011).

[163] Adam Alloul, Benjamin Fuks, and Verónica Sanz. "Phenomenology of the Higgs Effective Lagrangian via FEYNRULES". *JHEP* 04 (2014), p. 110. DOI: 10.1007/JHEP04(2014)110. arXiv: 1310.5150 [hep-ph].

[164] Céline Degrande et al. "Automated one-loop computations in the SMEFT". *CERN* CERN-TH-2020-140, CP3-20-42 (2020). arXiv: 2008.11743 [hep-ph].

[165] John Ellis, Veronica Sanz, and Tevong You. "The Effective Standard Model after LHC Run I". *JHEP* 03 (2015), p. 157. DOI: 10.1007/JHEP03(2015)157. arXiv: 1410.7703 [hep-ph].

[166] Chris Hays, Veronica Sanz Gonzalez, and Gabija Zemaityte. "Constraining EFT parameters using simplified template cross sections". *LHC Higgs Cross Section Working Group Report* LHC-HXSWG-2019-004 (2019).

[167] CMS Collaboration. "Measurement of the associated production of a Higgs boson and a pair of top-antitop quarks with the Higgs boson decaying to two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-HIG-18-018. (2018).

[168] CMS Collaboration. "Measurements of properties of the Higgs boson in the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-HIG-19-001. (2019).

[169] CMS Collaboration. "Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV". *Phys. Lett. B* 791 (2019), p. 96. DOI: 10.1016/j.physletb.2018.12.073. arXiv: 1806.05246 [hep-ex].

[170] CMS Collaboration. "Measurement of Higgs boson production and decay to the $\tau\tau$ final state". Tech. rep. CMS-PAS-HIG-18-032. (2019).

[171] CMS Collaboration. "Search for the associated production of the Higgs boson and a vector boson in proton-proton collisions at $\sqrt{s} = 13$ TeV via Higgs boson decays to τ leptons". *JHEP* 06 (2019), p. 093. DOI: 10.1007/JHEP06(2019)093. arXiv: 1809.03590 [hep-ex].

[172] CMS Collaboration. "Evidence for the Higgs boson decay to a bottom quark-antiquark pair". *Phys. Lett. B* 780 (2018), pp. 501–532. DOI: 10.1016/j.physletb.2018.02.050. arXiv: 1709.07497 [hep-ex].

[173] CMS Collaboration. "Measurement of $\mathrm{t\bar{t}H}$ production in the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay channel in 41.5 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13\,\mathrm{TeV}$". Tech. rep. CMS-PAS-HIG-18-030. (2019).

[174] CMS Collaboration. "Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair". *Phys. Rev. Lett.* 120.7 (2018), p. 071802. DOI: 10.1103/PhysRevLett.120.071802. arXiv: 1709.05543 [hep-ex].

[175] CMS Collaboration. "Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV". *JHEP* 08 (2018), p. 066. DOI: 10.1007/JHEP08(2018)066. arXiv: 1803.05485 [hep-ex].

[176] CMS Collaboration. "Measurement of the associated production of a Higgs boson with a top quark pair in final states with electrons, muons and hadronically decaying τ leptons in data recorded in 2017 at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-HIG-18-019. (2018).

[177] CMS Collaboration. "Search for the Higgs boson decaying to two muons in proton-proton collisions at $\sqrt{s} = 13$ TeV". *Phys. Rev. Lett.* 122.2 (2019), p. 021801. DOI: 10.1103/PhysRevLett.122.021801. arXiv: 1807.06325 [hep-ex].

[178] ATLAS 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 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[179] Celine Degrande et al. "Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD". *Eur. Phys. J. C* 77.4 (2017), p. 262. DOI: 10.1140/epjc/s10052-017-4793-x. arXiv: 1609.04833 [hep-ph].

[180] Olivier Mattelaer. "On the maximal use of Monte Carlo samples: re-weighting events at NLO accuracy". *Eur. Phys. J. C* 76.12 (2016), p. 674. DOI: 10.1140/epjc/s10052-016-4533-7. arXiv: 1607.00763 [hep-ph].

[181] Andy Buckley et al. "Rivet user manual". *Comput. Phys. Commun.* 184 (2013), pp. 2803–2819. DOI: 10.1016/j.cpc.2013.05.021. arXiv: 1003.0694 [hep-ph].

[182] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[183] ATLAS Collaboration. "Constraints on an effective Lagrangian from the combined $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data collected with the ATLAS detector". Tech. rep. ATL-PHYS-PUB-2017-018. (2017).

[184] ATLAS Collaboration. "Interpretations of the combined measurement of Higgs boson production and decay". Tech. rep. ATLAS-CONF-2020-053. (2020).

[185] Sabine Kraml et al. "Constraining new physics from Higgs measurements with Lilith: update to LHC Run 2 results". *SciPost Phys.* 7.4 (2019), p. 052. DOI: 10.21468/SciPostPhys.7.4.052. arXiv: 1908.03952 [hep-ph].

[186] Waleed Abdallah et al. "Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2". *SciPost Phys.* 9.2 (2020), p. 22. DOI: 10.21468/SciPostPhys.9.2.022.

[187] Adam Falkowski et al. "Rosetta: an operator basis translator for Standard Model effective field theory". *Eur. Phys. J. C* 75.12 (2015), p. 583. DOI: 10.1140/epjc/s10052-015-3806-x. arXiv: 1508.05895 [hep-ph].

[188] CMS Collaboration. "Search for new physics in top quark production with additional leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV using effective field theory". *JHEP* 03 (2021), p. 095. DOI: 10.1007/JHEP03(2021)095. arXiv: 2012.04120 [hep-ex].

[189] CMS Collaboration. "W^+W^- boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV". *Phys. Rev. D* 102.9 (2020), p. 092001. DOI: 10.1103/PhysRevD.102.092001. arXiv: 2009.00119 [hep-ex].

[190] CMS Collaboration. "Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV". *Submitted to EPJC* (2021). arXiv: 2103.04956 [hep-ex].

[191] G. Apollinari et al. "High Luminosity Large Hadron Collider HL-LHC". *CERN Yellow Report* 5 (2015). DOI: 10.5170/CERN-2015-005.1. arXiv: 1705.08830 [physics.acc-ph].

[192] D Contardo et al. "Technical Proposal for the Phase-II Upgrade of the CMS Detector". Tech. rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02. (2015).

[193] CMS Collaboration. "The Phase-2 Upgrade of the CMS Endcap Calorimeter". Tech. rep. CERN-LHCC-2017-023. CMS-TDR-019. (2017).

[194] J. Duarte et al. "Fast inference of deep neural networks in FPGAs for particle physics". *Journal of Instrumentation* 13.07 (2018), P07027–P07027. DOI: 10.1088/1748-0221/13/07/p07027.

[195] CMS Collaboration. "The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report". Tech. rep. CERN-LHCC-2017-014. CMS-TDR-018. (2017).

[196] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". Proceedings for the Association for Computing Machinery, (2016). DOI: 10.1145/2939672.2939785.

[197] Andrew P. Bradley. "The Use of the Area under the ROC Curve in the Evaluation of Machine Learning Algorithms". *Elsevier Science Inc.* 30.7 (1997). DOI: 10.1016/S0031-3203(96)00142-2.

[198] "Addendum to the report on the physics at the HL-LHC, and perspectives for the HE-LHC: Collection of notes from ATLAS and CMS". *CERN Yellow Rep. Monogr.* 7 (2019). DOI: 10.23731/CYRM-2019-007.Addendum. arXiv: 1902.10229 [hep-ex].

[199] Roel Aaij et al. "Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era". LHCB-PUB-2018-009, CERN-LHCC-2018-027 (2016). arXiv: 1808.08865 [hep-ex].

[200] J. de Favereau et al. "DELPHES 3, A modular framework for fast simulation of a generic collider experiment". *JHEP* 02 (2014), p. 057. DOI: 10.1007/JHEP02(2014)057. arXiv: 1307.6346 [hep-ex].

[201] CMS Collaboration. "Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s}$ = 13 TeV". *Phys. Rev. Lett.* 122.12 (2019), p. 121803. DOI: 10.1103/PhysRevLett.122.121803. arXiv: 1811.09689 [hep-ex].

[202] ATLAS Collaboration. "Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector". *Phys. Lett. B* 800 (2020), p. 135103. DOI: 10.1016/j.physletb.2019.135103. arXiv: 1906.02025 [hep-ex].

[203] CMS Collaboration. "Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s}$ = 13 TeV". *JHEP* 03 (2021), p. 257. DOI: 10.1007/JHEP03(2021)257. arXiv: 2011.12373 [hep-ex].

[204] Massimiliano Grazzini et al. "Higgs boson pair production at NNLO with top quark mass effects". *JHEP* 05 (2018), p. 059. DOI: 10.1007/JHEP05(2018)059. arXiv: 1803.02463 [hep-ph].

[205] Giuseppe Degrassi et al. "Probing the Higgs self coupling via single Higgs production at the LHC". *JHEP* 12 (2016), p. 080. DOI: 10.1007/JHEP12(2016)080. arXiv: 1607.04251 [hep-ph].

[206] Fabio Maltoni et al. "Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC". *Eur. Phys. J. C* 77.12 (2017), p. 887. DOI: 10.1140/epjc/s10052-017-5410-8. arXiv: 1709.08649 [hep-ph].

[207] Martin Gorbahn and Ulrich Haisch. "Indirect probes of the trilinear Higgs coupling: $gg \to h$ and $h \to \gamma\gamma$". *JHEP* 10 (2016), p. 094. DOI: 10.1007/JHEP10(2016)094. arXiv: 1607.03773 [hep-ph].

[208] Wojciech Bizon et al. "Constraints on the trilinear Higgs coupling from vector boson fusion and associated Higgs production at the LHC". *JHEP* 07 (2017), p. 083. DOI: 10.1007/JHEP07(2017)083. arXiv: 1610.05771 [hep-ph].

[209] Stefano Di Vita et al. "A global view on the Higgs self-coupling". *JHEP* 09 (2017), p. 069. DOI: 10.1007/JHEP09(2017)069. arXiv: 1704.01953 [hep-ph].

[210] Ambresh Shivaji and Xiaoran Zhao. "Higgs Trilinear self-coupling determination through one-loop effects". https://cp3.irmp.ucl.ac.be/projects/madgraph/wiki/HiggsSelfCoupling. [Online; accessed 12-Sep-2018].

[211] Ilya Sutskever et al. "On the importance of initialization and momentum in deep learning". *Proceedings of the 30th International Conference on Machine Learning*. Vol. 28. 3. PMLR, (2013), pp. 1139–1147.

[212] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". *3rd International Conference for Learning Representations* (2017). arXiv: 1412.6980 [cs.LG].

[213] Hai-Jun Yang, Byron P. Roe, and Ji Zhu. "Studies of boosted decision trees for MiniBooNE particle identification". *Nuclear Instruments and Methods in Physics Research Section A* 555.1-2 (2005), pp. 370–385. DOI: 10.1016/j.nima.2005.09.022.

[214] J. R. Quinlan. "Induction of Decision Trees". *MACH. LEARN* 1 (1986), pp. 81–106.

[215] Leo Breiman. "Arcing classifier (with discussion and a rejoinder by the author)". *The Annals of Statistics* 26.3 (1998), pp. 801–849. DOI: 10.1214/aos/1024691079.

[216] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine." *The Annals of Statistics* 29.5 (2001), pp. 1189–1232. DOI: 10.1214/aos/1013203451.

[217] CMS Collaboration. "Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV". *JHEP* 10 (2017), p. 005. DOI: 10.1007/JHEP10(2017)005. arXiv: 1707.03316 [hep-ex].

[218] ATLAS Collaboration. "Methodology for EFT interpretation of Higgs boson Simplified Template Cross-section results in ATLAS". Tech. rep. ATL-PHYS-PUB-2019-042. (2019).