Chapter 6

Statistical analysis and results

The signal and background models described in Chapter 5 are used to perform the statistical interpretation of the 12.9 fb⁻¹ of data, resulting in an independent observation of the Higgs boson in the diphoton decay channel and measurements of some of its properties.

The statistical analysis proceeds in several stages. The first step is to determine the best signal-plus-background fit of the models to the data. The floating parameters of the signal and background models are varied simultaneously in each analysis category to obtain the closest overall agreement with the observed invariant mass distributions. This procedure is described in Section 6.1, and establishes the existence of a Higgs boson signal in the data. The next step is to formally quantify the significance of this signal, and to reject the hypothesis that there exists no Higgs boson. The procedure for hypothesis testing and calculation of the significance is detailed in Section 6.2. Having established the existence of a SM-like Higgs boson which decays to photons, the final step is to make measurements of some of its properties, in particular those which relate to the rate at which it is produced or interacts with other SM particles. These measurements are described in Section 6.3 and Section 6.4.

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

6.1 Best fit of models to the data

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

$$\mathcal{L}_C(\mu, m_{\mathrm{H}}; \mathbf{n} | m_{\gamma\gamma}^{obs}) = \mu \cdot f_S(m_{\gamma\gamma}^{obs} | m_{\mathrm{H}}; \mathbf{n}_S) + f_B(m_{\gamma\gamma}^{obs} | \mathbf{n}_B), \tag{6.1}$$

where:

- \mathbf{n} is a set of floating nuisance parameters, composed of those which affect the signal and those which affect the background, labelled \mathbf{n}_S and \mathbf{n}_B respectively;
- f_S and f_B represent the normalised probability distribution functions of the signal and background components of the model respectively;
- $\mu = \sigma^H/\sigma_{SM}^H$ is a parameter of interest (POI) called the signal strength, where in this context σ^H is the observed Higgs boson cross section, and σ_{SM}^H is the SM cross-section.

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

The construction of the signal component of the model, f_S , was described in Section 5.1. When performing the signal-plus-background fit to the data, the values of the individual parameters of the signal model functional form are fixed. Specifically, since the SSF method was used, this refers to the coefficients of the seventeen polynomial functions which describe the dependence on $m_{\rm H}$ of the DCB+1G parameters for each of the RV and WV scenarios and their mixing fraction. The parameters which are allowed to vary are the nuisance parameters \mathbf{n}_S introduced into the signal modelling to account for systematic uncertainties (see Section 5.3) as well as the POIs μ and $m_{\rm H}$. The signal strength μ uniformly scales all the signal models for each process and for each category. This means that the contribution to the overall signal model from each process remains in proportion to what is predicted by the SM, but the normalisation of the overall signal model can be varied.

The handling of the background component of the model f_B was described in Section 5.2. The nuisance parameters \mathbf{n}_B in each category are composed of:

- the discrete nuisance parameter which corresponds to the choice of background function, as prescribed by the discrete profiling method (see Section 5.2.1);
- the individual parameters of all the candidate functions, which are allowed to float in the fit.

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

$$\{(\mu, m_{\rm H}, \mathbf{n}) : -2 \ln \mathcal{L}(\hat{\mu}, \, \hat{m}_{\rm H}; \hat{\mathbf{n}} | \, m_{\gamma\gamma}^{obs}) = \inf_{\mu, m_{\rm H}, \mathbf{n}} (-2 \ln \mathcal{L}(\mu, \, m_{\rm H}; \mathbf{n} | \, m_{\gamma\gamma}^{obs}))\}, \tag{6.2}$$

where $\inf_{\mu, m_{\rm H}, \mathbf{n}}$ represents the *infimum function* evaluated over all allowed values of μ , $m_{\rm H}$ and \mathbf{n} , taking into account their constraints. This technique can be generalised for an arbitrary number of POIs. In general, the 2NLL cannot be minimized analytically. Instead, this task is handled numerically with the Minuit2 package [71].

The resulting parametrisations of the observed data are shown for each analysis category separately in Figures 6.1 and 6.2. The combined parametrisations and data resulting from a direct sum of each category is shown in Figure 6.3a, while the sum weighted by the S/(S+B) in $\pm \sigma_{\rm eff}$ around the best-fit $m_{\rm H}$ in each category is shown in Figure 6.3b. The best-fit values of the POIs are $\hat{\mu}=0.95$ and $\hat{m}_{\rm H}=126.0\,{\rm GeV}$, indicating a strong SM-like Higgs boson signal.

6.2 Significance of observation

Given the best-fit value of the signal strength $\hat{\mu} = 0.95$ determined in Section 6.1, a frequentist approach is used to determine the degree of certainty with which the hypothesis that there is no Higgs boson can be rejected in favour of an alternative hypothesis which is that a SM-like Higgs boson exists. The hypotheses can be formulated in terms of the signal strength: the null hypothesis H_0 corresponds to the case where $\mu = 0$, while the alternative hypothesis H_μ corresponds to $\mu > 0$.

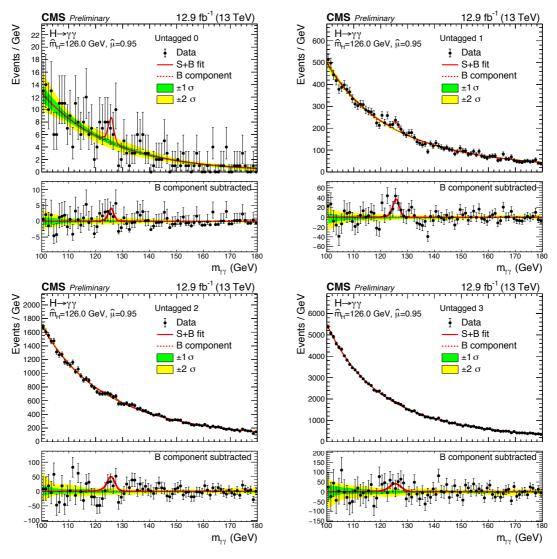


Figure 6.1: The signal-plus-background fit (solid red line) of the $m_{\gamma\gamma}$ distribution in data (black points) for the Untagged analysis categories. The background-only fit is shown as a dashed red line, while the green and yellow bands denote the 1σ and 2σ uncertainties on the background shape respectively.

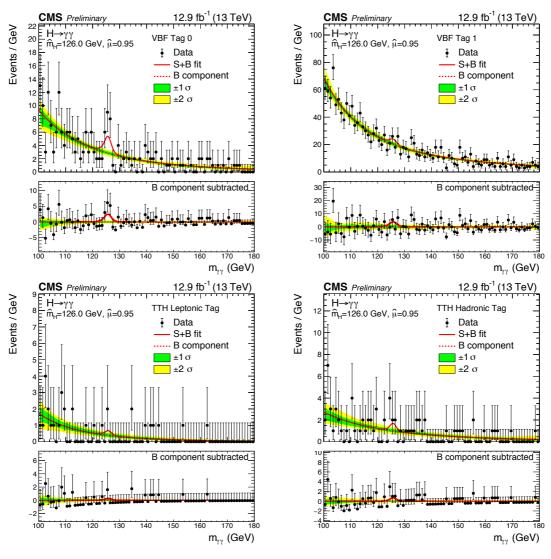


Figure 6.2: The signal-plus-background fit (solid red line) of the $m_{\gamma\gamma}$ distribution in data (black points) for the VBF-tagged and ttH-tagged analysis categories. The background-only fit is shown as a dashed red line, while the green and yellow bands denote the 1σ and 2σ uncertainties on the background shape respectively.

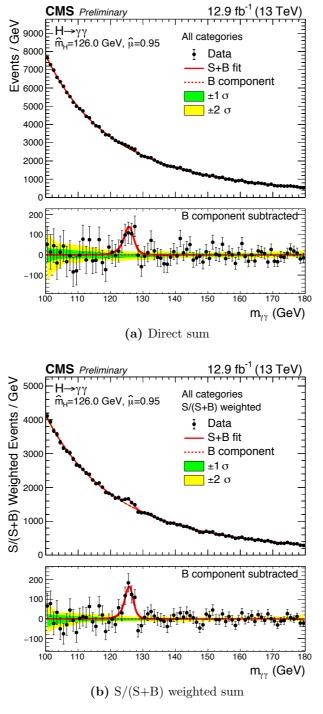


Figure 6.3: The signal-plus-background fit (solid red line) of the $m_{\gamma\gamma}$ distribution in data (black points) for all categories combined, either using a direct sum (a) or a sum weighted by the S/(S+B) in $\pm \sigma_{\rm eff}$ around the best-fit value of $m_{\rm H}$ (b). The background-only fit is shown as a dashed red line, while the green and yellow bands denote the 1σ and 2σ uncertainties on the background shape respectively.

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

$$P(m_{\gamma\gamma}^{obs} \in w|H_0) \le \alpha, \tag{6.3}$$

where $P(m_{\gamma\gamma}^{obs} \in w|H_0)$ is the probability (assuming that H_0 is correct), of observing the data inside the critical region w, and α is a small predetermined threshold [72].

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

$$P(m_{\gamma\gamma}^{obs} \in w | H_{\mu}) = 1 - \beta. \tag{6.4}$$

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

A common choice, which is found to maximise β [72], is to define the critical region in terms of a test statistic q_{μ} , corresponding to the difference between the best-fit 2NLL and the 2NLL evaluated for a particular μ . This difference is abbreviated as 2Δ NLL, and can be expressed as:

$$q_{\mu} = \begin{cases} -2 \ln \mathcal{L}(\mu, \ \hat{m}_{\mathrm{H}}^{\mu}; \hat{\mathbf{n}}^{\mu} | \ m_{\gamma\gamma}^{obs}) - \ln \mathcal{L}(\hat{\mu}, \ \hat{m}_{\mathrm{H}}; \hat{\mathbf{n}} | \ m_{\gamma\gamma}^{obs}) & \text{when } \hat{\mu} \ge 0, \\ 0 & \text{when } \hat{\mu} < 0, \end{cases}$$
(6.5)

where $\hat{m}_{\rm H}^{\mu}$ and $\hat{\bf n}^{\mu}$ denote the best-fit values $m_{\rm H}$ and $\bf n$ for a fixed value of μ . In this case, $m_{\rm H}$ is said to be profiled, although the test statistic can be modified to evaluate the $2\Delta {\rm NLL}$ for a given $m_{\rm H}$ hypothesis by fixing it to a particular value instead. When trying to exclude hypothesis H_0 , the test statistic q_0 in particular should be used to define a critical region. In the limit of a large sample of data, the probability distribution function of the test statistic (f_q) , is Gaussian. The fact that $q_0 = 0$ for $\hat{\mu} < 0$ reflects the fact that only excesses in the data are regarded as significant. This simplifies the definition of the critical region, since increasingly large values of q_0 indicate increasing incompatibility with H_0 , and therefore only the right-hand tail of f_q is considered when assessing probabilities. Assuming H_0 , the probability of obtaining a value of q_0^{obs} (corresponding to observed

data $m_{\gamma\gamma}^{obs}$) or higher is given by the integral of f_q from q_0^{obs} to infinity. This probability is commonly referred to as the *p*-value. We can therefore define the critical region as:

$$w = \{ m_{\gamma\gamma}^{obs} : \int_{q_0^{obs}}^{+\infty} f_q(q_0) dq_0 \le \alpha \},$$
 (6.6)

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

The p-value can be evaluated separately for different assumption about the value of $m_{\rm H}$. In this case, the value of $m_{\rm H}$ is fixed to the given value rather than being profiled in the $2\Delta {\rm NLL}$ minimisation. The result of a scan of the p-values as a function of $m_{\rm H}$ in 0.1 GeV steps is shown in Figure 6.4. The black solid line represents the p-value scan for the observed data. The dashed lines represent the expected p-values for a SM Higgs boson. These are obtained by generating an Asimov dataset [62] from the best-fit background-only model and a signal of strength $\mu = 1$, and then performing a signal-plus-background fit as for oberved data. For the blue dashed line, the signal was injected at $m_{\rm H} = 125.09 \, {\rm GeV}$ (the best fit value from the previous combined measurement of the Higgs boson mass by CMS and ATLAS [23]), while for the red dashed line the signal was injected at the corresponding $m_{\rm H}$ for each step.

The observed significance at $m_{\rm H}=125.09\,{\rm GeV}$ is 5.6σ , where 6.2σ was expected for the SM Higgs boson. The maximum observed significance is found at $m_{\rm H}=126.0\,{\rm GeV}$, corresponding to 6.1σ . Since the observed data fall in the critical region where the signifiance is above 5σ (i.e p-value is less than 2.87×10^{-7}), the null hypothesis that there is no Higgs boson is rejected. Therefore, the data correspond to a new observation of the Higgs boson decaying to photons.

6.3 Measurements of the signal strength

6.3.1 Global signal strength

One of the advantages of using $2\Delta NLL$ as a test statistic is that to a very good approximation, the $\pm 1\sigma$ uncertainty on the measured value of the POI can be obtained

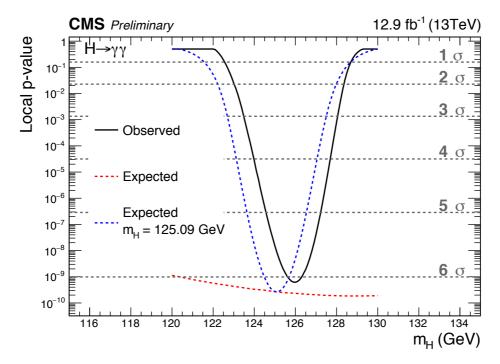


Figure 6.4: The p-value for the observation as a function of the Higgs boson mass (black), shown with the expected p-values for a SM Higgs boson, across the range 120-130 GeV. The expected p-values are obtained using Asimov datasets [62]. The blue dashed line shows the expected p-value when the mass of the injected signal is $m_{\rm H} = 125.09$ GeV, while the red line shows the maximum significance for any injected signal in the range of 120 to 130 GeV.

by finding the values for which $2\Delta NLL = 1$ [72]. This fact is used to produce a measurement of the global signal strength μ .

Figure 6.5 shows the value of $2\Delta \text{NLL}$ evaluated for fixed values of μ in small steps in the range of 0.5 to 1.5. This type of plot is referred to as a $2\Delta \text{NLL}$ scan. In this case, m_{H} is profiled in the minimisation at each step. By definition, the best-fit point $\hat{\mu}$ has a $2\Delta \text{NLL}$ value of 0. This gives the central value for the measurement. The upper and lower uncertainties are obtained by graphically finding the intercepts of the curve with $2\Delta \text{NLL} = 1$. The contribution to the total uncertainty on the signal strength arising from the statistical, theory systematic and experimental systematic components are assessed by freezing the corresponding nuisance parameters, and calculating the difference in quadrature with respect to the total uncertainty. The measured value of the signal strength is:

$$\hat{\mu} = 0.95^{+0.21}_{-0.19} = 0.95 \pm 0.17 \text{ (stat.)} ^{+0.09}_{-0.06} \text{ (theo. syst.)} ^{+0.10}_{-0.07} \text{ (exp. syst.)}.$$

This measurement indicates that the observed global signal strength is compatible with the SM expectation within one standard deviation.

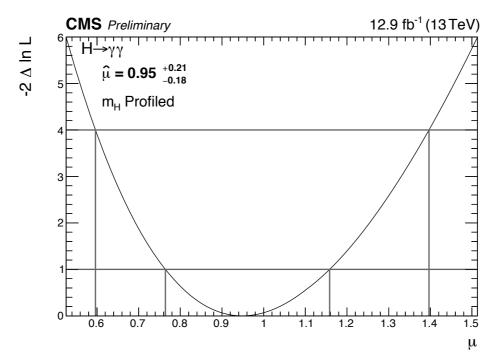


Figure 6.5: The 2Δ NLL scan of the overall signal strength for a Higgs boson decaying to two photons. The mass of the Higgs boson is profiled in the fit. The 1σ and 2σ uncertainties correspond to the crossings with 2Δ NLL = 1 and 2Δ NLL = 4.

Instead of allowing the $m_{\rm H}$ parameter the be profiled, it can instead be fixed to particular values. A similar likelihood scan can be repeated for given values of $m_{\rm H}$ in the 120-130 GeV range, in small steps. The result is shown in Figure 6.6, where the best-fit signal strength is plotted as a function of the fixed value of $m_{\rm H}$, where the green bands represent the $\pm 1\sigma$ uncertainty obtained by finding the crossing with $2\Delta {\rm NLL} = 1$ as described above. This figure illustrates that no other excesses other than the one at the best-fit exist in the region of interest.

6.3.2 Fermionic and bosonic components of the signal strength

When making the measurement of the global signal strength as in Section 6.3.1, a single POI which uniformly scales all production processes in all categories is defined. However, this measurement makes the assumption that the contribution of each production process to the total Higgs boson cross-section is in proportion to the SM prediction. In order to

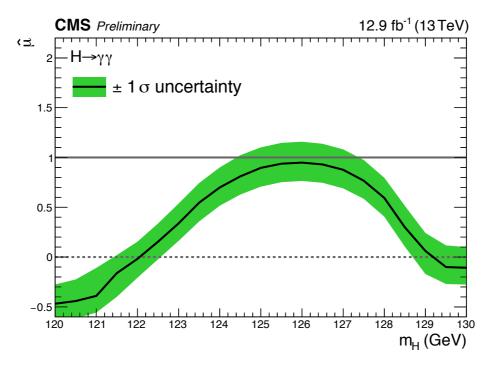


Figure 6.6: The best-fit signal strength for fixed values of $m_{\rm H}$ in the 120-1230 GeV range, where the $m_{\rm H}$ parameter is fixed in the fitting procedure. The green bands show the $\pm 1\sigma$ uncertainty obtained by finding the crossing with $2\Delta {\rm NLL} = 1$.

test this assumption, the single POI representing to global signal strength can be split up into components. For example, the contributions from the production modes where the Higgs boson is produced from fermions (ggH and ttH) and vector bosons (VBF and VH) are separated, to test if they individually agree with the SM expectation.

The measurement is made by producing a two-dimensional $2\Delta NLL$ scan with a slightly modified test statistic, where two POIs, $\mu_{ggH,ttH}$ and $\mu_{VBF,VH}$, are considered in the place of μ . The likelihood function is modified such that the $\mu_{ggH,ttH}$ parameter scales the yield of the signal models for the ggH and ttH processes in all categories uniformly, but does not affect the yields of the models for the VBF or VH processes, and vice-versa for the $\mu_{VBF,VH}$ parameter. The m_H parameter is profiled in the scan.

The result of the two-dimensional scan can be seen in Figure 6.7. The z-axis, representing the value of $2\Delta NLL$, has been omitted for clarity. The black cross shows the location of the best-fit point, with the red diamond indicating the SM expectation. The 1σ and 2σ contours, which represent the intersection with $2\Delta NLL = 2.30$ and $2\Delta NLL = 6.18$ respectively, are shown as solid and dashed lines. The plot shows that the observed

best-fit of ($\mu_{ggH,ttH}$, $\mu_{VBF,VH}$) = (0.82, 1.60) is consistent with the SM hypothesis within 1σ .

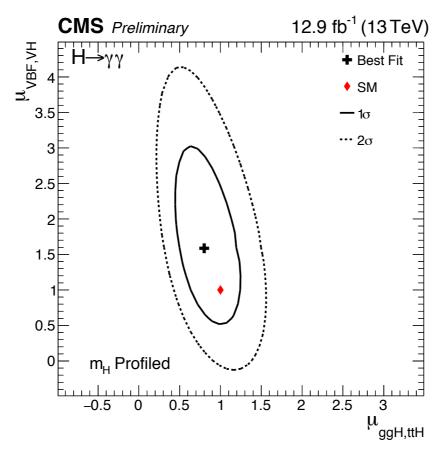
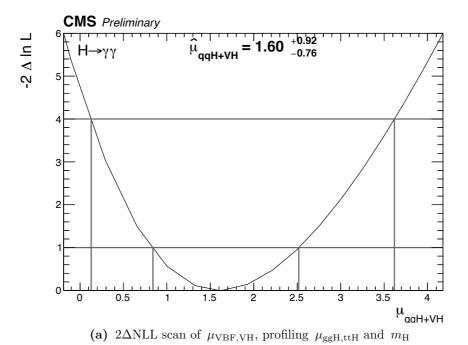


Figure 6.7: The result of a two-dimensional $2\Delta NLL$ scan of the $\mu_{ggH,ttH}$ and $\mu_{VBF,VH}$ components of the signal strength. The red diamond indicates the SM expectation, while the black cross shows the location of the best-fit point. The measurement is consistent with the SM within the uncertainty contours, which are shown in dashed lines. The value of m_H was profiled in the scan.

To correctly extract the uncertainties on $\mu_{ggH,ttH}$ and $\mu_{VBF,VH}$ individually, a 2 Δ NLL scan of each parameter is performed while profiling the other. The resulting scans can be seen in Figures 6.8a and 6.8b. The result of the measurement is found to be:

$$\hat{\mu}_{\mathrm{ggH,ttH}} = 0.82^{+0.27}_{-0.24} \text{ and } \hat{\mu}_{\mathrm{VBF,VH}} = 1.60^{+0.90}_{-0.76}.$$

The fermionic and bosonic components of the signal strength are therefore found to be compatible with the SM expectations.



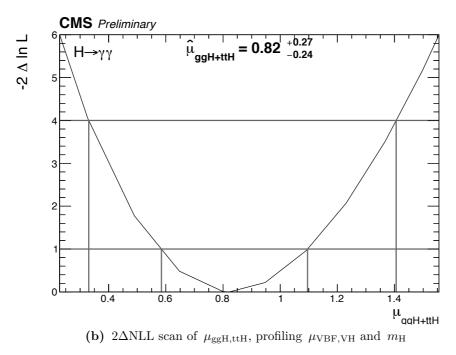


Figure 6.8: The result of performing $2\Delta NLL$ scans of $\mu_{ggH,ttH}$ while profiling $\mu_{VBF,VH}$ (a) and vice-versa (b). In both cases the mass of the Higgs boson is profiled in the fit.

6.3.3 Per-process signal strengths

Using a similar procedure to what was described in Section 6.3.2, measurements of the signal strengths of the individual Higgs boson production modes can be made. In this case, the likelihood function and test statistic are modified to contain one POI for each production process (μ_{ggH} , μ_{VBF} , μ_{VH} , and μ_{ttH}). Each per-process signal strength independently scales the yield for the signal models for the corresponding process in all categories, leaving the signal models for the other processes unchanged.

The technique described above exploits the categorisation scheme described in Chapter 4: in particular, the fact that the relative contribution from each process to the overall signal model differs from category to category. This means that varying each per-process signal strength has a different effect on the overall likelihood function. However, since no VH-tagged categories are included in this analysis, it is not possible to resolve the effect of varying μ_{VH} from other POIs, in particular μ_{ggH} , since most VH events are included in the Untagged categories along with the ggH events. To break the degeneracy, when making the measurements of the other POIs, the parameter μ_{VH} is fixed to a value of 1.

The measurements of the μ_{ggH} , μ_{VBF} , and μ_{ttH} are performed by producing a $2\Delta \text{NLL}$ scan of each parameter, while profiling the others. The m_{H} parameter is also profiled. The best-fit values and their uncertainties are shown on Figure 6.9. The measurement of the global signal strength obtained in Section 6.5 is shown as the vertical black line with green bands showing the 1σ uncertainties. The SM expectation is shown as the vertical dashed red line. The per-process signal strength measurements are all compatible with the SM expectation within 1σ .

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

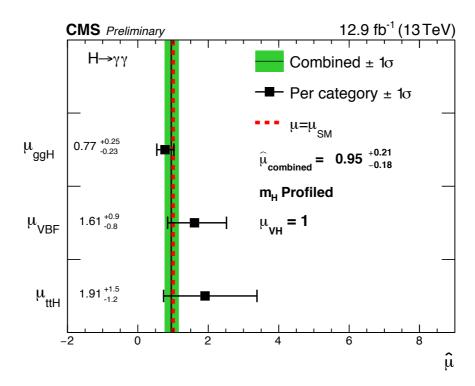


Figure 6.9: The measurements of the per-process signal strengths μ_{ggH} , μ_{VBF} , μ_{ttH} , obtained by performing $2\Delta \text{NLL}$ scans of each one while profiling the others. In each case m_{H} is also profiled in the fit, and $\mu_{VH}=1$ is imposed since this analysis does not include any categories specifically targeting the VH process. The vertical black line and green bands represent the measurement of the overall signal strength μ , and the SM expectation is shown in the vertical red dashed line.

6.3.4 Compatibility of result with SM in each category

Using an analogous method to the one described in Section 6.3.3, it is possible to make a measurement of the signal strength for each category separately. In this case, one POI per analysis category is defined, which scales the yield of the signal models of all processes uniformly, but independently within each category. Although these signal strengths do not have any physical meaning, they can be used to check that each category gives a result consistent with the overall measurement, and that no bias is introduced by any particular category. The result of the check is shown in Figure 6.10, which determines that all the per-category signal strengths are compatible with the SM expectation and the overall result.

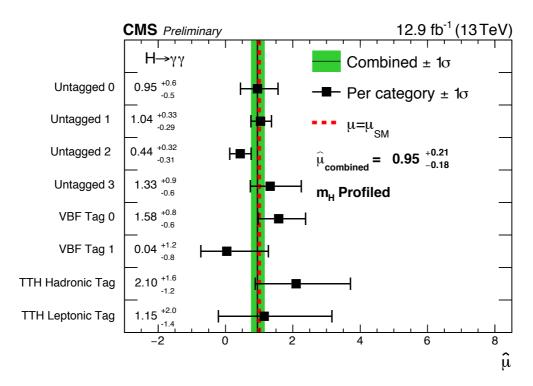


Figure 6.10: The measurements of the per-category signal strengths, obtained by performing $2\Delta \text{NLL}$ scans of each one while profiling the others. In each case m_{H} is also profiled in the fit, The vertical black line and green bands represent the measurement of the overall signal strength μ , and the SM expectation is shown in the vertical red dashed line. The per-category signal strengths are not meaningful physical observables, and this result is a check that no particular category is introducing a large bias into the overall measurement.

6.4 Measurements of Higgs boson coupling modifiers

The measurements presented so far have all dealt with Higgs boson signal strengths. Such observables are sensitive the variations in the rate at which the Higgs boson is produced, but do not take into account the possible variations in the partial width of the subsequent decay. An alternative set of measurements addresses this shortcoming by being sensitive to variations in the coupling strength of the Higgs boson with individual particles, relative to the SM expectation. The so-called kappa framework assigns a modifier to the coupling strength of the Higgs boson to a particle or group of particles X directly in the amplitude of the process. The corresponding modifier is labelled as κ_X . A detailed description of the scheme is available in [73]. The assumptions which underly this framework are as follows:

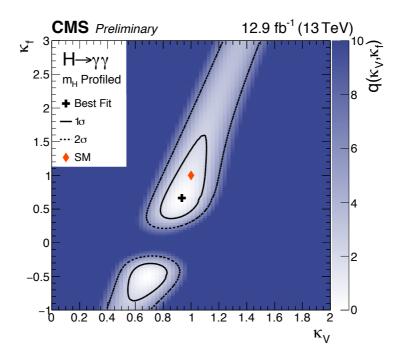
- any calculated deviations from the SM are due to only one Higgs-boson-like state with mass around 125 GeV;
- the natural width of this state is sufficiently small that it can be neglected, allowing the cross-section and branching fraction for a process $ii \to H \to ff$ to be decomposed as $(\sigma_{ii}^H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

The coupling modifier κ_X for a particle X interacting with the Higgs boson is applied directly as a factor to the corresponding σ_{XX}^H if the interaction produces a Higgs boson or Γ_{XX}^H if the interaction involves the decay of a Higgs boson. For processes which only occur via loops of particles, the effective coupling modifier is defined as a function of the coupling modifiers for particles in which play a large role in the loop, e.g. $\kappa_{\gamma} = \kappa_{\gamma}(\kappa_b, \kappa_t, \kappa_{\tau}, \kappa_W)$ (for the decay $H \to \gamma \gamma$) and $\kappa_g = \kappa_{\gamma}(\kappa_b, \kappa_t)$ (for ggH production).

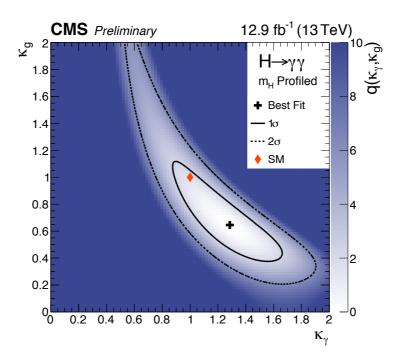
In this measurement, coupling modifiers $\kappa_{\rm f}$ and $\kappa_{\rm V}$ are defined which uniformly scale the Higgs boson's interactions with all fermions and vector bosons respectively. The result of a two-dimensional $2\Delta {\rm NLL}$ scan of $\kappa_{\rm f}$ and $\kappa_{\rm V}$ is shown in Figure 6.11a, where the $m_{\rm H}$ parameter was profiled in the fit. The black cross indicates the best-fit while the red diamond indicates the SM expectation. The 1σ and 2σ contours are indicated by solid and dashed lines. The best-fit indicates compatibility with the SM.

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

Alternatively, the measurement can be made in terms of the Higgs boson's effective coupling to gluons and photons using a two-dimensional $2\Delta NLL$ scan of κ_g and κ_{γ} , where the m_H parameter was profiled. The results are presented in Figure 6.11b. The best-fit point agrees with the SM within the uncertainties.



(a) $2\Delta NLL$ scan of κ_f versus κ_V



(b) $2\Delta NLL$ scan of κ_{γ} versus κ_{g}

Figure 6.11: The result of a two-dimensional $2\Delta NLL$ scan of the effective coupling strength modifiers for: fermions and bosons (b), and gluons and photons (b). The best-fit values are denoted with black crosses and the SM expected values with red diamonds. The best-fit points agree with the SM within the 1σ and 2σ uncertainty contours denoted by the solid and dashed lines respectively.