Bayesian Model comparison of Higgs couplings

Johannes Bergström\textsuperscript{a} Stella Riad\textsuperscript{b}

\textsuperscript{a}Departament d’Estructura i Constituents de la Matèria and Institut de Ciencies del Cosmos, Universitat de Barcelona, Diagonal 647, E-08028 Barcelona, Spain
\textsuperscript{b}Department of Theoretical Physics, School of Engineering Sciences, Royal Institute of Technology (KTH) – AlbaNova University Center, Roslagstullsbacken 21, S-106 91 Stockholm, Sweden

E-mail: bergstrom@ecm.ub.edu, sriad@kth.se

Abstract: We investigate the possibility of contributions from physics beyond the Standard Model (SM) to the Higgs couplings, in the light of the LHC data. The work is performed within an interim framework where the magnitude of the Higgs production and decay rates are rescaled though Higgs coupling scale factors. We perform Bayesian parameter inference on these scale factors, concluding that there is good compatibility with the SM. Furthermore, we carry out Bayesian model comparison on all models where any combination of scale factors can differ from their SM values and find that typically models with fewer free couplings are strongly favoured. We consider the evidence that each coupling individually equals the SM value, making the minimal assumptions on the other couplings. Finally, we make a comparison of the SM against a single “not-SM” model, and find that there is moderate to strong evidence for the SM.

Keywords: Statistical methods, Higgs physics
1 Introduction

The discovery of a boson with a mass of approximately 125.5 GeV was announced in July 2012 by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) at CERN [1, 2]. This discovery is compatible with the previous one from proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV at the Tevatron [3]. Using all of the available data, with a total luminosity of 25 fb$^{-1}$ from the proton-proton collisions with energies of $\sqrt{s} = 7$ and 8 TeV runs at the LHC, properties of the Higgs boson properties, such as spin, parity, mass, and the couplings to other Standard Model (SM) particles, has been further investigated [4–6]. So far, however, there are no indications of major deviations from the properties of the SM Higgs boson, and the boson does in fact seem to be a CP even scalar [7, 8]. The discovery of the Higgs boson marks an important milestone in the history of particle physics, especially for our understanding of electroweak symmetry breaking and the generation of particle masses [9–12].

The Higgs boson was discovered in the decays into gauge bosons, i.e., the Higgs decays to $WW$, $ZZ$, and $\gamma\gamma$. However, using the full set of data from the LHC, there is now also evidence for decays to fermions, $b\bar{b}$ and $\tau^+\tau^-$ [13, 14].
Even though the properties of the new boson so far are compatible with those of the SM Higgs, the possibility for new physics in the Higgs sector should be investigated. New physics can either be manifested as effective interactions at tree-level or in the form of additional particles contributing to the loop-induced production and decay modes or the invisible decay modes of the boson. These new contributions could then be detected as they would effectively give rise to scaling of the magnitude or change in the structure of the Higgs boson couplings. Hence, the natural step forward in experimental Higgs physics is precision measurements of the Higgs boson couplings to fermions and gauge bosons. There are rather severe bounds from the existing data on the couplings of the boson, however, hadron colliders are in general not ideal for Higgs precision measurements since they cannot measure the narrow total width and thus not the values of individual couplings and new contributions to the total decay width at the same time. Hence, in order to determine these quantities with even greater precision the upgrade of the LHC to 14 TeV won’t suffice, experiments have to be performed for example at a Higgs factory [15].

The status of the Higgs couplings as measured by the LHC can be studied by means of so-called Higgs coupling scale factors, introduced by the LHC Higgs cross section working group (LHC HXSWG) as an interim treatment of the Higgs couplings [16]. In this framework, the magnitude of the Higgs production and decay rates is simply rescaled, which is of special use with respect to the presentation of the experimental data from the collaborations in terms of so-called global signal strengths. In this treatment, we assume a single underlying Higgs boson at the mass of about 125 GeV, i.e., an excitation of a field whose vacuum expectation value breaks electroweak symmetry. Furthermore, it is compatible with data, i.e., it is a CP-even scalar. Furthermore, a zero-width approximation is used, which significantly simplifies the treatment. This so-called interim framework has been studied by the ATLAS and CMS collaborations as well as in several phenomenological studies [17–25]. In this work, we use the software HIGGS SIGNALS for the statistical analysis of the LHC data [26–28].

In the present work we apply Bayesian inference within the framework of coupling scale factors. We shall use Bayesian parameter inference to get a rough idea of how the parameters are constrained. However, since the most important question is rather that of which model best describes the data, we will focus on model comparison – in particular of different models in which any combination of couplings differ from their SM values. This framework makes it possible to compare many models to each other at once, but the main advantage is that it is possible to obtain evidence in favour of simpler models – in the present case for models where the couplings are given by their SM values.

The paper is organized as follows. In Sec. 2, we give an introduction to Higgs physics and the concept of coupling scale factors. In Sec. 3, we discuss the Bayesian method used in the present work, especially model comparison in the context of Higgs couplings. In addition, we discuss the models used in the present work as well as the priors used. In Sec. 4, we discuss the results of parameter estimation, and the results concerning the different questions addressed using model comparison. Finally, in Sec. 5, we give a short summary and give our conclusions.
2 Higgs physics

Whether the discovered particle at 125.5 GeV actually is the SM Higgs boson, or only a part of some bigger picture, needs to be determined. In general additional degrees of freedom in the Higgs sector will influence the Higgs couplings to the SM particles as well as the loop-induced production and decay modes.

One common way to investigate the possibility of new physics in the Higgs sector is to study specific renormalizable, physical models for BSM physics, such as two Higgs Doublet Models [29–31], composite Higgs models [32, 33], a dilaton model [19], and supersymmetry [34, 35]. However, this case would be specific since comparisons would only be made between the SM and this specific model, and obviously lack in generality. Another way is to consider the SM extended with effective operators, which result from new physics above the TeV scale. Since this new physics is heavy it will give rise to modifications of the couplings, which necessarily are small in magnitude. Both these investigation methods will alter the Higgs boson couplings, some or all of them, in general both magnitude and tensor structure. Another approach is simply to not consider a physical and realistic model, but instead make a statistical analysis based on the “naive” rescaling of the magnitude of the Higgs couplings. In such a framework it is possible to investigate whether there are any significant deviations of the couplings from their SM value but impossible to determine which underlying model gave rise to the deviations. Thus the relevant result of the analysis is whether the couplings deviate from their SM value or not, rather than the exact value of the couplings. This treatment, with coupling scale factors shall be considered here.

2.1 Production modes

Four production modes of the Higgs boson in the SM are significant at the LHC. The predominant production mode is the loop-induced gluon fusion $gg \to H$, with heavy quarks running in a triangular loop, with the main contribution coming from the top quark. Since this process is loop-induced it is of particular interest in searches for new physics. The subdominant processes are vector boson fusion $q\bar{q}' \to q\bar{q}' H$, associated production with a vector boson $W/ZH$, and the associated production with a top-quark pair $q\bar{q}/gg \to t\bar{t}H$. Here we use the notation where $l$ stands for lepton and $q$ stands for quark.

2.2 Decay modes

The Higgs can decay to either a fermion-antifermion pair or two gauge bosons. In the SM, the Higgs coupling to fermions is proportional to the fermion mass, hence the heaviest kinematically accessible fermion mode will have the largest partial decay. Currently five decay modes of the Higgs boson have been detected at the LHC, namely the $\gamma\gamma$, $ZZ^{(*)}$ (which in turn is followed by a decay to $4l, 2l\nu, 2l2q, 2l2\tau$), $WW^{(*)}$ (followed by decays to $l\nu l\nu, l\nu q\bar{q}$), $bb$, and $\tau^+\tau^-$, which then decay leptonically and hadronically. The decay mode $H \to \mu^+\mu^-$ have been investigated at the LHC, with the result of a rather loose upper bound [36].

Due to the excellent mass resolution in the diphoton and four-lepton channel, they are the two most important decay channels for the mass determination of a light Higgs
are $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4l$. The most common decay mode for a light Higgs, $bb$, has a lot of background and the Higgs decay is therefore difficult to detect. The SM branching ratios are given by the LHC Higgs Cross Section Working group [37]. There are additional decay channels, such as $H \rightarrow Z\gamma$ and $H \rightarrow gg$, which have not yet been detected and therefore will not be considered in the present analysis. These two processes would however be interesting for physics beyond the SM, since they are loop-induced and therefore new particles, in principle, could participate in the loop.

2.3 Definition of coupling scale factors

The LHC Higgs results are commonly presented in terms of global signal strengths, defined as

$$
\mu = \frac{\sigma(X) \cdot BR(H \rightarrow Y)}{\sigma(X)^{SM} \cdot BR(H \rightarrow Y)^{SM}}.
$$

where $\sigma(X)$ is the cross section for the production mode $X$ and $BR(H \rightarrow Y)^{SM}$ the branching ratio of the decay mode $Y$. In the case of a SM process the value of $\mu$ is naturally 1. In the SM the Higgs boson couples to the other particles with couplings $y_i^{SM}$, where $i \in \{t, b, \tau, \mu, W, Z\}$. The couplings to the fermions are the Yukawa couplings where

$$
y_f^{SM} = \frac{m_f}{v},
$$

where $m_f$ is the mass of the fermion, $f \in \{b, t, \tau, \mu\}$, and $v$ is the Higgs vev. The upper perturbative limit for these couplings is $4\pi$. For the gauge couplings we have

$$
y_W^{SM} = \frac{2m_W^2}{v}, \quad y_Z^{SM} = \frac{m_Z^2}{v},
$$

where $m_W, m_Z$ are the $W$ and $Z$ masses respectively. Note that these couplings are dimensionful.

A simple extension to the SM can be made by rescaling the magnitude of the SM decay and production rates, which effectively will lead to a rescaling of the Higgs, couplings by introducing so called coupling scale factors, $\kappa_i$. For the processes which exist at tree-level in the SM the couplings are rescaled as

$$
y_i = \kappa_i \cdot y_i^{SM}.
$$

Naturally, the SM is recovered for $\kappa_i = 1$. In addition coupling scale factors can be introduced for the loop-induced processes. We introduce $\kappa_g$ and $\kappa_\gamma$ for the $gg \rightarrow H$ and $H \rightarrow \gamma\gamma$ respectively. In principle, a scale factor, $\kappa_{Z\gamma}$, should be introduced for a third loop-induced process $H \rightarrow Z\gamma$. However, there are no significant bounds on this process in the LHC data. Hence the cross section of the process $ii \rightarrow H \rightarrow ff$ is given by

$$
(\sigma \cdot BR)(ii \rightarrow H \rightarrow ff) = \sigma^{SM}(ii \rightarrow H) \cdot BR^{SM}(H \rightarrow ff) \cdot \frac{\kappa_i^2 \kappa_f^2}{\kappa_H^2},
$$

where $\kappa_i$ and $\kappa_f$ corresponds to the initial and final states respectively and $\kappa_H$ is the scale factor for the total Higgs decay width.
The coupling scale factors $\kappa_g$ and $\kappa_\gamma$ can either be functions of the other coupling scale factors or free parameters of the fit if new physics is allowed to participate in the loops. In the SM these scale factors have the values $\kappa_g = \kappa_\gamma = 1$. However, in the case when only the tree-level scale factors are varied, the scale factors of the loop-induced processes will vary depending on the other scale factors. The effects of the rescaled tree-level couplings would have to be cancelled by some new physics, if these parameters were fixed to their SM values. If the scale couplings are free, new physics is allowed to propagate in the loop.

Furthermore the factor $\kappa_g$ can be defined in two different ways, either in terms of partial cross-sections or decay widths. In the present case we define the coupling scale factor, $\kappa_g(\kappa_t, \kappa_b)$ using the cross sections, since gluon fusion is the more important process. Thus the scale factor is given by

$$\kappa^2_g(\kappa_b, \kappa_t) = \frac{\kappa^2_t \cdot \sigma_{tt}^{ggH} + \kappa^2_b \cdot \sigma_{bb}^{ggH} + \kappa_t \kappa_b \cdot \sigma_{tb}^{ggH}}{\sigma_{tt}^{ggH} + \sigma_{bb}^{ggH} + \sigma_{tb}^{ggH}}$$

(2.6)

In terms of the other $\kappa$s, $\kappa_\gamma$ is given by

$$\kappa^2_\gamma(\kappa_b, \kappa_t, \kappa_\tau, \kappa_W) = \frac{\sum_{ij} \kappa_i \kappa_j \cdot \Gamma_{ij}^{\gamma\gamma}}{\sum_{ij} \Gamma_{ij}^{\gamma\gamma}},$$

(2.7)

where $\Gamma_{ij}^{\gamma\gamma}$ are the partial decay widths and the pairs $(i, j)$ are given by $bb, tt, \tau\tau, WW, bt, b\tau, bW, t\tau, tW, \tau W$ [16].

In addition, the total Higgs width scales with a coupling scale factor, which is defined in terms of the other coupling scale factors as

$$\kappa^2_H = \sum_X \kappa^2_X \cdot BR_{SM}(H \to X),$$

(2.8)

where the summation runs over all possible decay modes in the SM. This parametrization requires that the resonance width is small and therefore the zero-width approximation is assumed. In principle new physics could contribute to the total Higgs width, in which case $\kappa_H$ should be a free parameter. For an extensive description of the concept of coupling scale factors, see Ref. [16].

In the present work we shall focus on the coupling scale factors in two settings. First, the scale couplings corresponding to the SM tree-level couplings, which are currently constrained by the LHC, i.e., the ones rescaling the Higgs couplings to $bb, tt, \tau\tau, \mu^+\mu^-$, $WW$, and $ZZ$. In the second case, the loop-induced processes are scaled as well. Thus, in addition, the coupling scale factors $\kappa_g$ and $\kappa_\gamma$ are introduced. Naturally, when the scale of the coupling is allowed to vary, the coupling is no longer completely determined by the corresponding particle mass. We shall not consider the total decay width to be a free parameter in the present case. Throughout this work, the information on effective scale couplings from LHC data were implemented using the HiggsSignals software [26–28].

The new particle is assumed to resemble the SM Higgs boson. In principle, however, new physics will not only change the magnitude of the couplings but also their tensor structure. These new couplings usually are referred to as anomalous couplings. The method of analysis applied here is however applicable to the extended problem concerning anomalous couplings.
3 Statistical approach

In this work, we will be using Bayesian probability theory, where each proposition is associated with a probability or **plausibility**, defined to lie between 0 and 1. This is the only consistent extension of boolean logic, which incorporates uncertainty [38–40].

In order to calculate the probabilities of different hypotheses or models, the laws of probability are used when conditioned on some known (or presupposed) information. Of interest in the present case is Bayes’ theorem, which can be used in order to reverse the conditioning,

\[
\Pr(A|B) = \frac{\Pr(B|A) \Pr(A)}{\Pr(B)}. \tag{3.1}
\]

A given set of hypotheses or models \( H_j \), can be compared using some set of collected data, \( D \), through calculation of the **posterior odds**, which is given by

\[
\frac{\Pr(M_i|D)}{\Pr(M_j|D)} = \frac{\Pr(D|M_i) \Pr(M_i)}{\Pr(D|M_j) \Pr(M_j)}. \tag{3.2}
\]

The **prior odds** \( \Pr(M_i)/\Pr(M_j) \) quantifies how much more plausible one model is than the other **a priori**. This ratio is typically taken equal to unity, which however must be discussed more carefully in some cases. The **evidence**, \( Z_i = \Pr(D|M_i) \), is the likelihood for the model quantifying how well the model describes the data. The **Bayes factor**, \( B_{ij} = Z_i/Z_j \), which is the ratio of the evidences, quantifies how much better the model \( M_i \) describes the data than \( M_j \).

Given that the model \( M \) contains the free parameters \( \Theta \), the evidence is given by

\[
Z = \Pr(D|M) = \int \Pr(D, \Theta|M)d^N \Theta
\]

\[
= \int \Pr(D|\Theta, M) \Pr(\Theta|M)d^N \Theta
\]

\[
= \int L(\Theta) \pi(\Theta)d^N \Theta, \tag{3.3}
\]

where \( L(\Theta) \equiv \Pr(D|\Theta, M) \) is the **likelihood function**. The prior probability density of the parameters is given by \( \pi(\Theta) \equiv \Pr(\Theta|M) \), and should always be normalized, i.e., it should integrate to unity. The assignment of priors are probably the most discussed and controversial part of Bayesian inference. This is often far from trivial, but nevertheless this assignment is an important, even essential, part of any Bayesian analysis.

The Bayes factors, or rather the posterior odds, are interpreted or “translated” into ordinary language using the so-called **Jeffreys scale**, given in Tab. 1 as used in, e.g., Refs. [41, 42] (“log” denotes the natural logarithm). Even though the Bayes factor in general will favour the correct model once “enough” data have been obtained, the evidence is often highly dependent on the choice of prior.

Under the assumption that a model \( M \) is true, complete inference of its parameters is given by the posterior distribution,

\[
\Pr(\Theta|D, M) = \frac{\Pr(D|\Theta, M) \Pr(\Theta|M)}{\Pr(D|M)} = \frac{L(\Theta)\pi(\Theta)}{Z}. \tag{3.4}
\]
In this case, the evidence is only a normalization factor, since it is independent of the values of the parameters \( \Theta \) and it is therefore often disregarded in parameter estimation. However, the actual values of the parameter within a pre-specified model are often not of great interest. Instead the primary question is which values are preferred from data, given a set of models.

After model comparison, there might still be a significant amount of uncertainty regarding which model actually is the best, which should not be ignored. Therefore instead distributions of parameters can be considered, without assuming that the model with maximum probability is correct. Model uncertainty can be taken into account by calculating the model-averaged posterior distribution \([43, 44] \)

\[
\Pr(\eta|\mathbf{D}) = \sum_{i} \Pr(\eta|H_i, \mathbf{D}) \Pr(H_i|\mathbf{D}).
\]

This gives the average of the individual distributions over the full space of the models considered, weighted by the posterior model probabilities. Averaging over models can be done for both prior and posterior distributions, however the parameters \( \eta \), which could be derived, obviously need to be well-defined in all of the models. The posterior in Eq. (3.4) is obtained by setting all prior model probabilities except one, equal to zero. For applications in physics and cosmology see Refs. [44–46].

The main result of Bayesian parameter inference is the posterior and its marginalised versions (usually in one or two dimensions). Commonly, point estimates such as the posterior mean or median, are given together with credible intervals (regions), which are defined as intervals (regions) containing a certain amount of posterior probability. These regions are not unique, without further restrictions, similarly to classical confidence intervals, and in general they do not describe all the information contained in the posterior. We use MultiNest [47–49] for the evaluation of all evidences and posterior distributions in this work.

### 3.1 Model comparison and Higgs couplings

We want to determine whether there is any evidence in the LHC data for deviations from the SM values of the couplings, \( \text{i.e., } \kappa_i \neq \kappa_i^{\text{SM}} \), or if \( \kappa_i = \kappa_i^{\text{SM}} \) is sufficient to describe the data. From a statistical viewpoint, a model with \( \kappa_i = \kappa_i^{\text{SM}} \) can also be interpreted as a model where there is some undetectable deviation from the SM value, for a discussion on

| \[\log(\text{odds})\] | odds  | \(\Pr(M_1|\mathbf{D})\) | Strength of evidence |
|----------------|-------|----------------|---------------------|
| < 1.0          | \(\lesssim 3 : 1\) | \(\lesssim 0.75\) | Inconclusive |
| 1.0            | \(\approx 3 : 1\) | \(\approx 0.75\) | Weak evidence |
| 2.5            | \(\approx 12 : 1\) | \(\approx 0.92\) | Moderate evidence |
| 5.0            | \(\approx 150 : 1\) | \(\approx 0.993\) | Strong evidence |

Table 1. The Jeffreys scale, which is used for interpretation of Bayes factors, odds, and model probabilities. The posterior model probabilities for the preferred model are calculated assuming only two competing hypotheses and equal prior probabilities. Note that \(\log\) denotes natural logarithm.
this subject see Ref. [43]. In the present case, we are interested in if there is a deviation from the SM couplings, and not exactly what it is. For each coupling this gives two distinct cases and in order to differentiate between them, we want to perform Bayesian model comparison. Beforehand it is not specified whether the other couplings, i.e., the couplings with indices \( j \neq i \), should be fixed to their SM value or not, which gives rise to a complication. In principle, there is an important distinction since, without making the assumption of a particular model, different combinations of the couplings can deviate significantly from each other.

Thus, we can consider the models \(^1\) \( H_\alpha \), with \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \), where each \( \alpha_i = 0 \) if \( \kappa_i = \kappa_i^{\text{SM}} \) and \( \alpha_i = 1 \) if \( \kappa_i \neq \kappa_i^{\text{SM}} \). In total there are \( 2^n \) models, where \( n \) is the number of free parameters. In fact we can consider \( \alpha \) as a discrete parameter, for which the posterior odds is given by

\[
\frac{\Pr(\alpha|D)}{\Pr(\beta|D)} = \frac{\Pr(D|\alpha)\Pr(\alpha)}{\Pr(D|\beta)\Pr(\beta)} = \frac{Z_\alpha \pi_\alpha}{Z_\beta \pi_\beta}
\]  

(3.6)

where the calculable Bayes factor \( B^\alpha_\beta = Z_\alpha / Z_\beta \) quantifies how much better \( \alpha \) describes the data than \( \beta \). The natural baseline model is \( \text{SM} = \bar{0} = (0, 0, \ldots, 0) \), and all the \( B^\alpha_\beta \) can be obtained from the \( B^\alpha_{\text{SM}} \) as \( B^\alpha_\beta = B^\alpha_{\text{SM}} / B^\beta_{\text{SM}} \). If also finite prior odds \( \pi_\alpha / \pi_\beta \) are assigned to the full set of models it is assumed that complete, finite posteriors \( \Pr(\alpha|D) \) can be calculated, even though we will typically refrain from doing this. Calculating the Bayes factor does, however, require assignment of priors on the couplings in all the models, which is non-trivial and will be discussed in detail in Sec. 3.4.

A different, but equivalent, approach is to instead consider a single model with a prior which is a mixture of the continuous prior and a point mass at the SM value,

\[
\pi(\kappa_i) = (1 - p_i)f_i(\kappa_i) + p_i\delta(\kappa_i - \kappa_i^{\text{SM}}),
\]  

(3.7)

for each coupling.\(^2\) Here the continuous part of the prior given by \( f_i \) (which is normalized to unity) corresponds to the prior assuming \( \alpha_i = 1 \), is assigned a total probability \( 1 - p_i \) and the SM value of the coupling is assigned a probability \( p_i \) using the \( \delta \)-function.

Note that \( \alpha \) is a function of \( y \) and hence that the induced priors and posteriors of \( \alpha \) always can be extracted from the distributions obtained using (3.7). In addition, the Bayes factors (which are independent of the prior on \( \alpha \)) can be calculated using (3.6) by factoring out the prior odds.

In fact, this can be used in order to improve the numerical accuracy since a, possibly very, non-uniform prior on \( \alpha \) can be chosen in such a way that the corresponding posterior becomes relatively uniform. Then all values of \( \alpha \) will be sampled adequately, which in turn will make the posterior more accurate over all the values. Using this method, the prior odds can be factored out, enabling the evaluation of all the \( 2^n \) Bayes factors, which span many orders of magnitude, simultaneously. In the present case models with many couplings at their SM values will be preferred to the models with many couplings free. The larger the

---

\(^1\)Remember we refer to a statistical model and not necessarily a full, realistic physical model.\(^2\)In the general non-separable case all the quantities in the equation can depend on couplings \( y_j \) for \( j < i \).
number of free parameter the better the fit has to be in order for the more complex model to be preferred. Since the data do not indicate any major deviations these more complex models will not be preferred. Therefore in the model comparison \( p_i = p_0 \) has been chosen relatively small in order to obtain adequate sampling.

### 3.2 Inclusion of individual couplings

In the previous section, we discussed the comparison of \( 2^n \) models, with different numbers of Higgs scale couplings kept free to vary. However, when \( n \) grows in size, comparing this large number of models to each other rapidly becomes less transparent.

One can test if a particular variable should be included by comparing the cases \( \kappa_i = \kappa_i^{\text{SM}} \) and \( \kappa_i \neq \kappa_i^{\text{SM}} \), and hence calculating the Bayes factors

\[
B_i = \frac{\Pr(D|\alpha_i = 0)}{\Pr(D|\alpha_i = 1)}.
\]

However, except the usual specification of the prior on \( \kappa_i \), we, one also needs to decide which assumptions, \( i.e., \) priors, should be imposed on the other couplings. Possibilities could be

- fixed to the SM value (\( S \)),
- free and different from the SM value (\( F \)), or
- either of the above, \( i.e., \) an average (\( A \)).

The evidences are given by the likelihoods integrated not only over the prior on \( \kappa_i \), but also over the prior on all other couplings. In particular,

\[
\Pr(D|\alpha_i) = \sum_{\alpha_i^*} \Pr(D|\alpha_i, \alpha_i^*) \pi(\alpha_i^*),
\]

which depends on the prior on \( \alpha_i^* = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n) \). The evidences in Eq. (3.9) are simply the evidences discussed in Sec. 3.1. The three cases then corresponds to \( p_k = \pi(\alpha_k) \) being equal to either \( p_k = 1, 0 \), or some intermediate values, most naturally 0.5 (see [43] for detailed discussion). The results are only expected to be independent of this choice in the case where the constraints one one parameter is independent of the values of the others and the likelihood factorizes into one factor for each coupling.

### 3.3 Single comparison with SM

In physics there is often a theoretically \( a \ priori \) motivated “baseline” model which all extended models should be compared to, unlike in applications in other fields as discussed for example in Ref. [43]. In the present case the obvious choice for such a reference model is the SM. Furthermore, Bayesian model comparison treats all models on equal footing, which enables quantification of how much the SM is favoured with respect to extended models. Again, this could be done in the context of specific renormalizable high-energy models, but here we will focus on the effective case only considering the rescaled couplings.
We want to compare the SM with a model “not-SM”, or $\overline{\text{SM}}$. The question is how this model for comparison should be defined. For example, one could compare the SM with a model with only a single coupling free, which is just one of the cases discussed in the previous chapter. However, this is obviously not satisfactory since there are several such models\textsuperscript{3}, and at the same time we are completely neglecting models with two or more couplings free\textsuperscript{4}. Alternatively, one could compare with the most general model in which all couplings are free. However, the issue is the same, still neglecting the possibility that there could be significant deviations in more than one coupling, but not in all at once. The most general model could be punished for the inclusion of the couplings for which the SM value is preferred. Therefore the most general comparison appears to be the one between the SM and a model in which each coupling \emph{either} takes the SM values, or differs from it (with the possibility that all couplings having the SM value simultaneously excluded).

Indeed, probability theory again yields

$$\Pr(D|\overline{\text{SM}}) = \sum_\alpha \Pr(D|\alpha)\pi(\alpha|\overline{\text{SM}}),$$

(3.10)

and all the above cases are just cases for a specific choice of prior $\pi(\alpha|\overline{\text{SM}})$. Due to lack of further information, we take $\pi(\alpha_i|\overline{\text{SM}}) = 1/2$, which means that in the $\overline{\text{SM}}$ model it is equally probable that each coupling deviates (significantly) from the SM value, as it is that there is no (or negligible) deviation. In this case, however, the couplings of the model $\overline{\text{SM}}$ equal the SM couplings with prior probability $1/2^n$. This part of $\overline{\text{SM}}$, \emph{i.e.}, the part that is statistically equivalent to the SM, can of course then just be excluded in the analysis, and this will be our default choice. In principle, however, taking the SM model to mean the exact SM, one could also motivate its inclusion by saying there could still be a deviation from the SM values, however negligible. Note that adding additional couplings, which are unconstrained by data, does not affect the comparison of SM and $\overline{\text{SM}}$.

### 3.4 Choice of prior

As discussed in Sec. 2, we will consider the two cases: (i) all tree-level couplings are allowed to vary, with the loop-induced couplings calculated assuming no additional contribution from new physics; (ii) all couplings, including the loop-induced ones, are allowed to vary, which implies that new physics is allowed to participate in the loop processes. In the first case there are 6 free parameters, whereas there are 8 free parameters in the second case. Note that the default “SM” values of these couplings are those calculated in Eqs. (2.6) and (2.7) assuming no new particles, which do not necessarily correspond to the actual exact SM value (equal to 1). In addition, the value of the scale factor for the total Higgs width, $\kappa_H$, will depend on the other ones according to Eq. (2.8). We shall however not consider this as a free parameter in either case.

\textsuperscript{3}This could be remedied by comparing with a model in which any of the couplings are free, but only one at a time.

\textsuperscript{4}Again, we remind the reader that the statistical model for which $\kappa_i = \kappa_i^{\text{SM}}$ does not necessarily mean that this holds “exactly”, but rather to a very good approximation, \emph{i.e.}, that the deviation is negligible.
In order to calculate the evidence of the models in which the couplings differ from the SM value, a prior for each coupling $\kappa_i$ is needed. The assignment of prior is important in the Bayesian analysis since not only the posteriors within these models depend on it, but perhaps more importantly, so does the evidence. It is therefore important to take care to include as much known information into the prior without making any assumptions based on the data under consideration.

- **Default: uniform.** A common choice is to take a uniform prior on each of the couplings in order to implement a priori “ignorance”, usually unbounded or with “wide enough” limits. However, there are a number of problems with such a prior. Firstly, contrary to the intuition of many physicists a uniform prior does not actually quantify ignorance in a parameter. This can easily be realised since a uniform prior in one parameter will not imply a uniform in another parameter, given by a non-linear transformation of the first one [50]; Secondly, an unbounded (improper) prior typically gives meaningless answers for the evidence, but so does any prior in the limit when the width goes to infinity. However, this does not necessarily imply that the uniform prior as such is useless or in general should be avoided. On the contrary, as any other prior it can be used in the cases where it is motivated for some given parameter. We shall still use this form of the priors on the $\kappa_i$’s in Sec. 4.1 to get a rough idea of what the parameter constraints are.

- **Couplings: uniform.** In the case where the tree-level couplings are free, one can consider the actual couplings appearing in the Lagrangian as the free parameters. In the Higgs sector there are Yukawa couplings for the Higgs coupling to the fermions as well as the Higgs coupling to the gauge bosons. In principle, one could argue that a priori all couplings should be of order one. However, the masses of the SM particles differ by so many order of magnitudes, which is known as the “flavour puzzle” [51]. Hence, a roughly uniform prior on each of the couplings, with an effective upper limit of some constant of order one (in principle there is also a perturbative upper limit which could be considered), would seem appropriate.\(^5\) However, if the measured couplings have a small (absolute) errors compared to one, this will lead to a very strong “Occam effect” (with a factor of “error” appearing in the evidence) which will strongly disfavour modifications of the couplings and give strong preference to the SM values. This is in fact the case for all of the couplings, with the possible exception of the top Yukawa and therefore all models with additional couplings will be severely disfavoured. Hence, in the present work we will not perform a detailed analysis of this case, even though these conclusions are worth to bear in mind.

- **Logarithmic.** Dropping the assumption that the couplings should be of order one, given that we know this is not the case in the SM, it might seem more appropriate

\(^5\)In the Bayesian analysis of the lepton sector of Ref. [52] this was used, although there the prior was derived using the symmetries present.
that instead orders of magnitude of the couplings have equal \textit{a priori} weight. Thus, the choice is instead a logarithmically uniform prior on $y_i$ between some lower limit and the perturbative upper limit, which is $4\pi$. The lower limit must be chosen by hand; we will use $10^{-7}$ as the default choice. However, it turns out that the results are insensitive to changing this lower limit by at least a few orders of magnitude and this can therefore be considered a reasonable choice. Furthermore, we will always assume positive couplings. However, due to the very small sensitivity of the sign of the couplings, moving some of the prior to negative couplings will in fact have a negligible effect on the evidences.

- \textit{Gaussian.} Instead of making the same assumption as in the previous cases, \textit{i.e.}, that the couplings are \textit{a priori} unrelated to the SM ones, one can consider that many SM extensions, such as the ones mentioned in Sec. 2, will all lead to modifications of the same size of the SM couplings. Without considering a specific model, we cannot determine whether this contribution should be positive or negative, hence a priori, $\langle \kappa_i \rangle = \kappa_i^{SM}$, and as discussed above we then would expect a typical deviation of $\sigma(\kappa_i) = s_i$. Out of all the (prior) distributions on the real number, there is a unique one which has maximal entropy (or equivalently “minimal information”), namely the Gaussian distribution [38, 50]. We will consider values $s_i = s$ in the range $1 - 4$ as appropriate, with a default value of $s = 2$.

Finally, we mention that one in principle could consider the SM in the context of effective operators, with additional higher-dimensional effective operators. These modifies the SM couplings by an amount proportional to $v^2/\Lambda^2$, where $\Lambda$ is the scale of new physics [53–55]. These operators could be implemented in a Bayesian analysis such as in Ref. [55], but they could also be implemented in the present analysis by simply using a prior on $\Lambda$ to obtain expected sizes of the additional contributions and from this construct priors on the $\kappa_i$'s.

If one expects that $\Lambda$ could be of any order of magnitude, much of the prior would be piled up close to the SM values, which would imply that it would be possible to obtain significant evidence against the SM, but not in favour of it. However, if the scale of new physics is assumed to be close to the electroweak scale as in Ref. [55], the typical modification would be of order one, in which case one will get a result to the one for the Gaussian prior above.

In the present work we shall consider the following models and priors. In the case with only the tree-level couplings free we shall make the analysis both using a logarithmic prior, which is placed directly on the actual couplings, and a Gaussian prior, which instead is placed on the coupling scale factors, $\kappa_i$. In the second case, where both tree and loop-level couplings are free, we shall only make an analysis using the Gaussian prior placed on the scale factors. In this case it should be noted that the expectation value of the now free parameters $\kappa_g$ and $\kappa_\gamma$ are the values given from the other scale factors, \textit{i.e.}, the values given by Eqs. (2.6) and (2.7), and not the SM value of these scale factors (which is 1).
4 Results

4.1 Default parameter constraints

In this section, we obtain the “default” parameter constraints on the coupling scale factors by calculating the likelihood using HiggsSignals and imposing a uniform prior on the $\kappa$’s with zero as the lower limit and a “large enough” upper limit. Although this prior does not impose a priori ignorance, and it cannot be used for model comparison, the derived parameter constraints will be valid as long as the uniform prior is reasonable in the the region of parameter space which are not completely ruled out by the data. A fixed Higgs boson mass of $m_H = 125.5$ GeV was used, and will be used throughout this work.

Similar to the model comparison performed later, we first simultaneously estimate only the scale factors present at tree-level, and then additionally also the loop-induced scale factors. In addition to these two cases, we shall consider the special case where new physics only contribute to the loop-induced processes and thus only the scale factors corresponding to these processes, i.e., $\kappa_g$ and $\kappa_\gamma$, are free.

In Fig. 1 we present the results in terms of one- and two-dimensional posterior distributions. In the two-dimensional plots the blue shading denotes the natural logarithm of the posteriors and the black contours the 1σ and 3σ credible regions, while the one-dimensional posteriors are also black in the plots on the diagonal. Superimposed on these, in red, are the 1σ and 3σ contours as well as the one-dimensional posteriors for the case when only the tree-level scale factors are free. As previously discussed, $\kappa_g$ and $\kappa_\gamma$ are given as functions of the free scale factors. Finally, the same quantities are presented in green (in the bottom right) for the case when the tree-level couplings remain fixed at their SM values but new physics is allowed to participate in the loop-induced processes. The SM values are marked with stars and vertical lines.

In all three fits, the SM values are inside (or extremely close to) the 1σ regions, which is in fact rather surprising. As expected, adding $\kappa_g$ and $\kappa_\gamma$ to the set of free parameters will relax constraints on the six free tree-level couplings. The main effects should be seen in the scale factors corresponding to the particles which give the main contribution to the loop processes. Thus, the largest effect will be for the top quark which gives the absolutely dominating effect to the loop in the gluon fusion process, while a smaller effect should also be seen in the bottom quark coupling. Apart from the top quark this is also the only particle that participates in both the gluon fusion and $H \rightarrow \gamma\gamma$ processes. The modifications to the other couplings are marginal. In a similar manner, the constraints on the loop-induced couplings are weaker in the eight-dimensional fit than in the two-dimensional one.

From the plots it is evident that there is quite strong evidence (in an informal way) that all couplings differ from zero, with the exception of $\kappa_\mu$, $\kappa_t$ is the 8-parameter fit, and to some extent $\kappa_b$.

---

6Defining the contours by the usual $\chi^2$-thresholds on $-2\log(L(\theta)/L_{\text{max}})$, with $L(\theta)$ the Bayesian marginal likelihood, yields virtually identical contours in all cases.
Figure 1. Results of (default) parameter estimation. Two-dimensional log-posterior distribution (blue shading), 1σ and 3σ Bayesian credible regions and one-dimensional posteriors (both black) of the eight-parameter fit. 1σ and 3σ credible regions and one-dimensional posteriors of the tree-level six-parameter fit (red) and the two-parameter fit (green). The SM values of unity are marked with vertical lines and stars, respectively.

4.2 Model comparison: all models

Although the previous results were interesting, they were all derived under the assumption that the scale factors actually differed from those of the SM. Following Sec. 3 we would instead like to perform model comparison. We will use the priors discussed in Sec. 3.4 and aim to evaluate how much the results depend on these different prior choices.

In this section we follow Sec. 3.1 and compare models with any combination of free
parameters. In particular, we use MultiNest with the priors in (3.7) and $p_i = p_0$ chosen so that the posterior over the space of models becomes as uniform as possible.\footnote{We check that the statistical error on each Bayes factor is reasonably small by considering the effective sample size $(\sum w_i)^2 / \sum w_i^2$, with $w_i$ the weight of each sample belonging to a certain model. Typically it is of the order of $10^2 - 10^3$, but at the very least 10.} There are in total $2^n$ models, with $n = 6$ when the tree-level couplings are free and $n = 8$ when also loop-processes are included.

In the left panel of Fig. 2 we present the logarithms of Bayes factors for all of the $2^6 = 64$ models, compared to the SM and using the logarithmic prior (on $[10^{-7}, 4\pi]$) for the tree-level couplings, i.e., the Yukawa couplings and gauge boson couplings. The models are divided into unicoloured groups depending on the number of couplings which are free. In the model to the far left in the figure all couplings are free, the models in the next group have 5 parameter free, etc., until the model to the far right, which is the SM (and has no visible bar since $\log B = 0$). The blue stars are the values calculated by extrapolating the comparison of the SM with the models with a single coupling free, and then assuming that adding an additional parameter has the same effect on $\log B$ regardless of the assumptions on the other parameters. This would be exact if the shape of the likelihood as a function of each parameter did not depend on the values of the other parameters; although not exact this seems to be a reasonable approximation.

As expected, there is a clear trend. The larger the number of free couplings, the smaller the values of $\log B$, i.e., the stronger the evidence against that model. Hence the evidence against the model with all couplings free is very strong. Adding any of the parameters makes the model worse with about the same amount, with the exception of $\kappa_\mu$, which only decreases the evidence with a small amount (roughly one log-unit). Letting $\kappa_W$ free, corresponding to most heavily constrained coupling, will have the largest effect on the evidence of the model. Furthermore, one should remember that the log-odds only equals $\log B$ when the priors are equal. In this case, one might argue that the SM should have the same prior probability as all the other models together, which (assuming that prior is uniformly distributed) would lead to the log-odds being $\log 2^6 \simeq 4$ smaller than the $\log B$’s in the plot. Again, we note that the dependence on the prior limits is very weak. For example, decreasing the lower limit to $10^{-15}$ would lead to a decrease of $\log B$ smaller than 0.7 for the addition of each coupling.

In the right panel of Fig. 2 we present $\log B$ for the same models, but with Gaussian priors on the coupling scale factors. The bars are obtained using a standard deviation of $s = 4$, and the solid black line using $s = 1$. Naturally, the choice of priors affects the exact values of the evidences, but the general trend is the same in all cases. Adding a parameter with a Gaussian prior is not as influential as adding one with a logarithmic prior, and the difference between the two Gaussian priors is only about one log-unit per parameter.

Next, we consider the case when also the loop-induced couplings are allowed to differ from the SM values, or rather those calculated in Eqs. (2.6) and (2.7), giving a total of $2^8 = 256$ combinations of free couplings. The same Gaussian priors as in the right panel of Fig. 2 has been used, but with the expectation values of $\kappa_g$ and $\kappa_\gamma$ given by Eqs. (2.6) and (2.7), since this is the expectation without any contribution from new physics. The trend is...
similar to the previous case with tree-level couplings in that models with few free couplings are preferred to models with more free couplings. However, when approaching the models with most free parameters, there seems to be a “leveling off” in the sense that adding more parameters is less damaging. This makes sense since if adding more free parameters deteriorates the constraints on the parameters already varied, the evidence will tend to be larger. Finally, in a similar way to the previous case one might consider the SM not on equal footing with each of the other models, making the posterior odds smaller than the Bayes factor (now with $\log 2^8 \approx 5.5$ log-units).

4.3 Inclusion of individual couplings

In the previous section we studied how all the different combinations of free couplings compared to each other. Although some conclusions could be drawn, the result was not completely transparent. In this section we instead follow Sec. 3.2 and evaluate the evidence for or against the inclusion of each individual coupling.

In Fig. 4 we give the logarithms of the Bayes factors in Eq. (3.8), i.e., against the inclusion of each of the couplings, both for the case of the six tree-level couplings with a logarithmic priors, and for the Gaussian priors on the scale factors. Here we use the value $s = 2$ for the standard deviation, although the difference from $s = 4$ and $s = 1$ as used previously is expected to be quite small. As in Eq. (3.9), the other (nuisance) couplings are either fixed to their SM values ($S$), allowed to vary with the same priors as the coupling of interest ($F$), or averaged over these two cases ($A$). However, in Eq. (3.9), the size of each contribution is proportional to the evidence of that particular model, and since typically
Figure 3. Logarithms of Bayes factors (with respect to the SM) for models with up to 8 free parameters, with a Gaussian prior with standard deviations $s = 4$ (bars) and $s = 1$ (black line).

the evidences are much larger when the other couplings equal their SM values\textsuperscript{8}, the average is dominated by these components. Hence the result for $\mathcal{A}$ equals that of $\mathcal{S}$ to a very good approximation. Note that these Bayes factors are evaluated separately using dedicated MultiNest runs. Hence, these number might differ somewhat from those which can be read from Figs. 2 and 3. The Bayes factors in the table have significantly smaller numerical errors of about 0.1.

Some general conclusions which can be drawn are that the logarithmic prior yields a stronger preference for the SM couplings than the Gaussian (as in previous chapter), and $\mathcal{S}$ stronger than $\mathcal{F}$ (which is reasonable since the constraints are relaxed).

The Higgs decay to $\mu^+\mu^-$ is rather weakly constrained and the results for this coupling is quantitatively different to the other tree-level couplings. For the logarithmic priors there is barely weak evidence in favour of the SM, while for the Gaussian case there is not even that. Moving on to the other tree-level couplings, for the log prior there is weak to moderate evidence for all the couplings, with $\mathcal{F}$ giving about $1 - 2$ log-units weaker preference than $\mathcal{S}$ and $\mathcal{A}$. For the Gaussian prior, the evidence is also weak to moderate, but typically weaker than the logarithmic case. For the Gaussian prior for the tree-level couplings, there

\textsuperscript{8}The muon coupling is an exception, but since this is due to a lack of constraints rather than the existence of a tension with the SM value, this has no effect.
Figure 4. Logarithms of Bayes factors against inclusion of couplings for the eight coupling scale factors. Values larger than 0 means the SM value of the coupling is prefered. The other couplings are either fixed to their SM values (S), or allowed to vary with the same prior as the coupling of interest (F), or averaged over these two cases (A). Since typically the evidences are much larger when the other couplings equal their SM values, the average is dominated by these components, and hence A yields essentially the same result as S.

is no significant difference between the cases where the loop-induced couplings are free or not.

The loop-induced couplings enter only in two cases, both with a Gaussian prior. When the tree-level coupling scale factors are fixed, there is just moderate evidence in favour of the SM values for both $\kappa_g$ and $\kappa_\gamma$, while in the case when the other couplings are free, this preference essentially disappears completely.

However, as discussed in Sec. 3.2, making the weakest assumption on the tree-level couplings, Bayesian probability theory tells us that one really ought to use the model-averaged results. Hence, one can make the following conclusions: the couplings moderately prefer the SM values for $b\bar{b}$, $t\bar{t}$, $WW$, for $\tau^+\tau^-$ and $ZZ$ too for logarithmic priors, but only barely so for Gaussian priors, for $gg$ and $\gamma\gamma$ the preference is barely moderate, and for the coupling to $\mu^+\mu^-$ the evidence is barely weak or none at all.

4.4 SM vs $\overline{\text{SM}}$

We consider the $\overline{\text{SM}}$ model as discussed in Sec. 3.3, with the most appropriate assumption is that all the couplings can either take their SM value, or differ from it, with a prior

\[
\frac{d}{d\theta} \log B_{\text{Gaussian, } F} = \frac{d}{d\theta} \log B_{\text{Gaussian, } F (\text{tree})} = \frac{d}{d\theta} \log B_{\text{Gaussian, } S} = \frac{d}{d\theta} \log B_{\text{log, } F} = \frac{d}{d\theta} \log B_{\text{log, } S}
\]
Table 2. log \( B \) between SM and SM for three different priors, in the two cases when the SM either is included in SM or not. In addition, we give the evidences for the most general model.

<table>
<thead>
<tr>
<th>Prior</th>
<th>SM</th>
<th>General</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log (with SM)</td>
<td>-3.8</td>
<td>-</td>
</tr>
<tr>
<td>Log (without SM)</td>
<td>-4.9</td>
<td>-20.2</td>
</tr>
<tr>
<td>Gauss (tree, with SM)</td>
<td>-3.1</td>
<td>-</td>
</tr>
<tr>
<td>Gauss (tree, without SM)</td>
<td>-3.5</td>
<td>-12.7</td>
</tr>
<tr>
<td>Gauss (all, with SM)</td>
<td>-4.4</td>
<td>-</td>
</tr>
<tr>
<td>Gauss (all, without SM)</td>
<td>-4.8</td>
<td>-14.7</td>
</tr>
</tbody>
</table>

probability of 0.5 for each. The special case where all couplings simultaneously take on their SM values would typically be excluded from, but could also be included in, the SM.

In Tab. 2, we present the comparison of SM with the above model for these two cases and for the different continuous priors. For reference we also give the Bayes factors for comparing with the most general model. When the SM part is excluded, the evidence for the SM is actually just about strong for the logarithmic and Gaussian (with \( s = 2 \)) priors on all couplings, and moderate for the case of tree couplings. In the present case, the evidence of the SM is dominated by the contribution from models with a single coupling free, weighted by their priors within the SM.

In the second case, there is also a contribution from the part equivalent with the SM, which can be relatively large (and even dominating in the logarithmic case). Still, the conclusions do not change significantly, although the 1.1 log-units difference for the logarithmic prior takes the evidence for the SM from just about strong to moderate. Of course, the most general model is very much disfavoured in all cases.

5 Summary and Conclusions

We have performed a Bayesian analysis of the LHC Higgs data and used an interim framework where the magnitude of the Higgs couplings are rescaled by coupling scale factors, whereas the tensor structure of the couplings is unaltered with respect to the SM. In the present work, we have limited our discussion to the couplings which are constrained by the LHC, in total six tree-level couplings and two loop-induced couplings.

We have performed Bayesian parameter inference on these coupling scale factors in the following three cases: either the tree-level couplings, the loop-level couplings, or both simultaneously, were free. In each case the SM values were well within the 1\( \sigma \)-region. However, when all couplings were free, neither \( \kappa_t \) nor \( \kappa_\mu \) were well-constrained and could in principle be zero.

Since the most important question is rather that of which model best describes the data, we have instead focused on Bayesian model comparison, considering models with either only the tree-level couplings in the Lagrangian, or all couplings, allowed to vary. In the first case, we used two types of prior distributions. A logarithmic prior, which was
imposed directly on the tree-level couplings, and a Gaussian prior, imposed on the coupling scale factors. In the second case, when the loop-induced couplings were also treated as free parameters, the analysis was made with a Gaussian prior imposed on the coupling scale factors. In each case we performed model comparison between models with one, several, or all of the couplings free. The larger the number of free parameters, the more disfavoured the model was.

We have considered a single coupling at a time in the cases where the other couplings could either be fixed to the SM values or allowed to vary with the same prior as the coupling of interest. The favoured models are those with the couplings fixed to the SM value, although the evidence is virtually non-existent for the coupling to $\mu^+\mu^-$. All this was performed with the combinations of free parameters and priors discussed above. Finally, we discussed the definition of the model $\overline{\text{SM}}$, and compared this single model to the SM, finding that the SM is moderately to strongly favoured.

Acknowledgments

J.B. acknowledges support from FP7 ITN INVISIBLES (Marie Curie Actions PITN-GA-2011-289442).
References


