Pramana – J. Phys. (2016) 87: 36
DOI 10.1007/s12043-016-1247-1

The model-independent analysis for Higgs boson

MD NAIMUDDIN and SHIVALI MALHOTRA*

Department of Physics and Astrophysics, University of Delhi, Delhi 110 007, India
*Corresponding author. E-mail: shivali.malhotra@cern.ch

Published online 23 August 2016

Abstract. The discovery of a 125 GeV particle, announced by the ATLAS and CMS Collaborations on July 04, 2012, is one of the most important events in the recent history of particle physics. This particle could be the last missing particle of the Standard Model of particle physics or it could be the beginning of the long list of particles predicted by the physics beyond the Standard Model. Before we jump to make the final conclusion about this particle, it is imperative to study all the properties of this newly discovered particle. Since the model-dependent analyses always have this danger of being biased, we can perform a model-independent search for the Higgs boson and also check if the 125 GeV particle is indeed the Standard Model Higgs boson or a particle belonging to the physics beyond the Standard Model.

Keywords. Higgs boson; model independent; HEP.

PACS Nos 14.80; 13.90.+i

1. Introduction

The discovery of a 125 GeV particle announced by ATLAS [1] and CMS [2] Collaborations in the summer of 2012 might turn out to be the last missing particle to complete the Standard Model of particle physics or it might just be the beginning of the unravelling of the physics beyond the Standard Model. Going by the importance this new particle carries, it is of utmost importance that we do not leave any stone unturned to find out each and every detail of this particle. The ATLAS and CMS Collaborations announced the discovery of the new particle in their search for the Higgs boson in its decay to $\gamma\gamma$ and four-leptons final state.

The Standard Model (SM) of particle physics has been extremely successful in describing the known phenomena so far with the only exception of the Higgs boson and the associated mass mechanism. The Higgs mechanism was proposed in order to overcome the question of mass generation. It was suggested that the spontaneous symmetry can be broken in the gauge theories [3–5] due to which gauge boson can acquire mass through the absorption of Nambu–Goldstone bosons. This leads to the introduction of a complex scalar doublet field in the theory. The masses of the W and Z bosons are generated when the spontaenous symmetry breaking is applied to the electroweak theory [6–8]. Nevertheless, SM in itself is not a complete theory and is unable to provide explanation for the many known facts. Therefore, there are strong reasons to expect new physics at the energies at or just above the electroweak scale.

The Higgs boson predicted by the SM interacts differently with different particles. The Higgs boson couples strongly to the heavy particles, like the W and Z bosons, the top quark and to a lesser extent to the b quark. The coupling of the Higgs boson with light particles are quite weak. The Higgs boson is dominantly produced at the high-energy colliders via gluon fusion, vector boson fusion, associated production with the weak boson and associated production with the heavy quarks. Figure 1a shows the production cross-section for Higgs boson from individual channels at 8 TeV centre-of-mass energy at the Large Hadron Collider (LHC), with the combined parametric and theoretical uncertainties as illustrated by the bands. The labels on the bands briefly indicate the type of radiative corrections that are included in the predictions. The Higgs boson decays through many channels. Figure 1b shows the decay branching ratio of Higgs boson and here the bands correspond to the theoretical uncertainties which are based on the error estimates for partial widths of the respective decay modes [9].

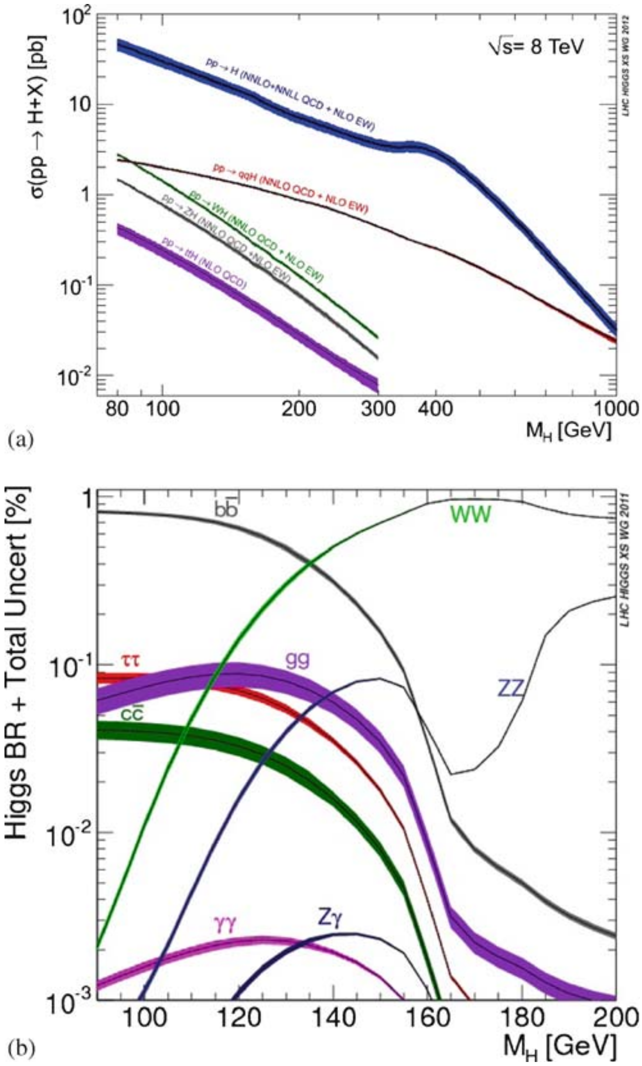

Figure 1. Higgs production cross-section at 8 TeV (**a**) and Higgs branching ratio as a function of Higss mass (**b**).

2. Model-independent approach to search for Higgs boson

The current search for Higgs boson is mainly focussed on analysing its SM decay modes and look for an excess of events compared to what is predicted by the SM, where Higgs boson signals are excluded. The LHC experiments ATLAS and CMS are mainly performing the search for the Higgs boson in the final states of $\gamma\gamma$, four leptons and $ll\nu\nu$ [10–15]. Recently, the channels having fermionic decays of Higgs boson, like two τ leptons [16,17] or two b-quarks in association with the weak bosons and production of Higgs with the top quarks have also been added [18–20]. Prior to this, the Tevatron experiments CDF and DØ performed searches for Higgs boson in the final states of $l\nu bb$, $llbb$ and $\nu\nu bb$ [21–26]. While performing all these searches based on a particular final state, one fine-tunes the analysis by optimizing the cuts for the maximum signal selection and maximum background rejection. In all these analyses many statistical tools were used for discriminating the signal from background.

The ultimate goal in all the statistical approach is to enhance the signal selection probability and increase the background rejection probability. In principle, there is nothing wrong with this approach if things are done in the right way. But being always eager to enhance the signal causes a potential danger of being biased. For a discovery as important as that of the Higgs boson, we cannot afford to leave an iota of doubt in our methodology and experimental approach. Hence, we also need to look for alternative approach which is not biased by any such consideration as the signal or the background. A global approach to search for the new particles and phenomena in a model-independent way completely removes this kind of bias. This approach has already been applied to the data collected by the DØ [27–30], CDF [31,32] and CMS [33,34] experiments.

We can use the approach of model-independent analysis to search for the Higgs boson as well. In such an analysis, we do not make data selection specific to a particular model or final state. This way we can analyse many channels without tuning our cuts to a specific channel or model. By doing this we do lose a little bit of sensitivity but the breadth of the search increases considerably and the search also remains unbiased towards any final state. The search for the Higgs boson can be performed by dividing the data in as many exclusive final state as possible. Then we take all the SM Monte Carlo (MC), except the Higgs, to estimate the background contributions from known physics. SM MC samples are also divided into as many exclusive final states as possible. Then different kinematic properties like invariant mass, transverse momenta, transverse missing energy, etc. are made both for the data and MC. Finally, a comparison is made between the data and the MC distributions across all the final states and any discrepancy between them will point towards a possible discovery of the new particle or the new physics. The observed discrepancy can then be quantified in terms of the statistical variables and the strength of the discrepancy can be inferred as follows.

The probability that the excess observed in the data is due to a statistical fluctuation of the SM sample in the channel f_s is determined from $p = 1 - (1 - p_{\mathrm{fs}})^{N_{\mathrm{fs}}}$, where p_{fs} is the probability of observing a discrepancy in an individual final state before the trial factor and N_{fs} is the number of unique exclusive final states observed in the data. In the limiting case of $p_{\mathrm{fs}} \ll 1$, the probability p becomes $p = N_{\mathrm{fs}} \times p_{\mathrm{fs}}$. The p_{fs} can be estimated as

$$p_{\mathrm{fs}} = \int_0^{\infty} \exp\left[-\frac{(N - N_B)^2}{2\sigma_B^2}\right] \mathrm{d}N \sum_{N_{\mathrm{data}}}^{\infty} \frac{N^i}{i!} \mathrm{e}^{-N},$$

where N_B and σ_B are the SM event yield expected from the background and its uncertainty, respectively, and N_{data} is the number of events observed in any channel. This is then converted into units of standard deviation using the equation

$$\int_{\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \mathrm{d}x = p.$$

3. Model-independent approach to confirm the Higgs boson

Once the mass of the new boson is measured, we now have to check whether the experimentally observed particle at a given mass is really the Higgs boson as expected from the SM predictions or does it imply some physics beyond the Standard Model. To evaluate the consistency of the data observed in different channels with the expectations for a SM Higgs boson, various properties of this boson have been measured using signal strength parameter, spin-parity measurements, couplings to vector bosons and fermions, etc. The mass of the new boson has been measured with two high-resolution channels, i.e. $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow 4l$ (where $l = e, \mu$) while the properties of the boson have been studied in five decay modes: $\gamma\gamma$, ZZ, WW, $\tau\tau$ and bb. For simplicity, $H \rightarrow b\bar{b}$ is denoted as $H \rightarrow bb$, $H \rightarrow \tau^+\tau^-$ as $H \rightarrow \tau\tau$, etc. further. There are four main Higgs boson production modes in proton–proton (pp) collisions. The gluon–gluon fusion production mode ($gg \rightarrow H$) has the largest cross-section, followed by vector boson fusion (VBF), associated WH and ZH production (VH) and production in association with top quarks ($t\bar{t}H$).

To incorporate all statistical uncertainties, systematic uncertainties and their correlations from the data selected by all individual analyses, a statistical methodology is used which was developed by the ATLAS and CMS Collaborations in the context of the LHC Higgs Combination Group [35]. The profile likelihood ratio [36] may then be used to determine how likely the test statistic q, is signal-like or background-like. According to the frequentist approach, systematic uncertainties can be incorporated in the analysis via nuisance parameters. The excess of events over the expected background can be estimated using a test statistic where the likelihood appearing in the numerator is for the background-only hypothesis:

$$q_0 = -2 \ln \frac{\mathcal{L}(\mathrm{obs}|b, \hat{\theta}_0)}{\mathcal{L}(\mathrm{obs}|\hat{\mu} \cdot s + b, \hat{\theta})},$$

where s stands for the signal expected under the SM Higgs hypothesis, μ is a signal strength modifier which accommodates deviations from the SM Higgs predictions, b are for backgrounds which are estimated from SM contributions and q are the nuisance parameters describing systematic uncertainties. The point at which the likelihood reaches its global maximum is defined by $\hat{\mu}$ and $\hat{\theta}$ whereas the value $\hat{\theta}_0$ maximizes the likelihood in the numerator under the background-only hypothesis (i.e. $\mu = 0$). A scan of the profile likelihood ratio, $q(a)$, is performed to evaluate the signal model parameter, a, using

$$q(a) = -2 \ln \frac{\mathcal{L}(\mathrm{obs}|s(a) + b, \hat{\theta}_a)}{\mathcal{L}(\mathrm{obs}|s(\hat{a}) + b, \hat{\theta})}.$$

The parameters $\hat{a}$ and $\hat{\theta}$ that maximize the likelihood, $\mathcal{L}(\mathrm{obs}|s(\hat{a}) + b, \hat{\theta}) = \mathcal{L}_{\max}$ are called the best-fit values. For $q(a_i) = 1$ and $q(a_i) = 3.84$, the 68% and 95% confidence limit (CL) on a given parameter of interest a_i can be evaluated, having the other unconstrained model parameters treated as the nuisance parameters.

The signal strength defined as $\mu = \hat{\sigma}/\sigma_{\mathrm{SM}}$, is calculated for each decay mode separately. The likelihood depends on the signal strength parameter such that $\mu = 0$ corresponds to the background-only hypothesis and $\mu = 1$ corresponds to the predicted Higgs boson signal in addition to the background. The likelihood is calculated as the product of the probabilities of observing each event, where the individual event probabilities depend on the measured masses (or m_{H}) of the Higgs boson candidates. The evaluation accounts for systematic uncertainties. The signal strength and the parameters that describe the systematic uncertainties are varied to maximize the likelihood. The ratio of the likelihood with the best-fit signal to that with a specified signal, $\mu = 1$ or 0, is calculated. These likelihood ratios are then used to quantify the exclusion of the signal hypothesis ($\mu = 1$) or the rejection of the background hypothesis ($\mu = 0$). The statistical tests were repeated at various values of m_{H} and μ. A SM Higgs boson with mass m_{H} was considered excluded when $\mu = 1$ is excluded at 95% CL at that mass. This is equivalent to an upper limit on μ at 95% CL being less than 1. On the other hand, a significant rejection of the background hypothesis was interpreted as evidence for the SM Higgs boson because this is the alternate hypothesis. As the SM does not predict the value of m_{H}, and because background fluctuations can occur anywhere in the search region of m_{H}, the local significance is an overestimate of the true significance. So, the global significance takes into account the 'look-elsewhere' effect. In our study, the above-mentioned likelihood is used to find an excess over the background which corresponds to 125 GeV signal with $\mu = 1$.

It is very crucial to measure the quantum numbers of the new boson-like spin and parity to determine its identity. Study of hypothesis tests has been performed between the SM Higgs boson with floating signal strength (0^+) and a pseudoscalar (0^-) or a spin-2 resonance with minimal coupling which are produced in gluon–gluon fusion ($2^+_m(gg)$). It has been presented by CMS in the $ZZ \to 4l$ channel [37] where the data disfavour the pure pseudoscalar hypothesis (0^-) and in the $H \to WW \to l\nu l\nu$ channel [38] where it also disfavours the hypothesis of a graviton-like boson ($2^+_m(gg)$). The observed value from the data is consistent with the expected one for the $J^P = 0^+$ hypothesis.

It is also a good test to check the compatibility of the observed data with the SM Higgs boson couplings. Given the production cross-section σ_x for an initial state x, the partial decay width Γ_{ff} into the final state ff and the total Higgs boson decay width Γ_{tot}, the event yield in any (production)×(decay) mode is assumed to be related as

$$(\sigma \cdot \mathrm{BR})(x \to H \to ff) = \frac{\sigma_x \cdot \Gamma_{ff}}{\Gamma_{\mathrm{tot}}},$$

where the initial state x includes gluon–gluon fusion, VBF, WH and ZH and $t\bar{t}H$ and ff include WW, ZZ, bb, $\tau\tau$, $\gamma\gamma$, and $Z\gamma$ final states. To accommodate the possibility of Higgs boson decaying into beyond the Standard Model (BSM) particles with a partial width Γ_{BSM}, the total width is considered as a dependent parameter such that $\Gamma_{\mathrm{tot}} = \sum \Gamma_{ii} + \Gamma_{\mathrm{BSM}}$, where Γ_{ii} stands for the partial width of the Higgs boson decay to all SM particles. These partial widths are proportional to the square of the effective Higgs boson couplings to the corresponding particles. To check for possible deviations in the rates expected in the different channels for the SM Higgs boson from the data, the modified couplings denoted by scale factors κ_i are introduced to fit the data according to these new parameters [39]. Here i can stand for: V (vector boson), W (W boson), Z (Z boson), f (fermions), l (leptons), q (quarks), u (up-type quarks), d (down-type quarks), b (b quark), t (top quark), τ (tau lepton), g (gluons), γ (photons). If any anomaly is observed in the measurement of these κ's, the new physics beyond the SM Higgs boson hypothesis may become true, although the measurements of the couplings do not show any statistically significant discrepancy from the observations.

To measure the Higgs boson properties further, one need some more quantities in addition to the already existing measurements like signal strength, as the best-fit value for the global signal strength factor μ does not give any direct information on the relative contributions from different production modes. After the measurement of m_{H}, the production cross-sections in the SM becomes completely fixed. But we cannot fix the ratios of the production cross-sections to the ratios predicted by the SM, as it will induce bias in our measurements. For a model-independent analysis we only assume a common signal strength scale factor $\mu_{ggF+ttH}$ which is assigned to both gluon fusion production (ggF) and the very small $t\bar{t}H$ production modes, as they both scale dominantly with the $t\bar{t}H$ coupling in the SM. Similarly, the VBF and VH production modes scale with the WH/ZH gauge coupling. Hence, a common signal strength scale factor $\mu_{\mathrm{VBF+VH}}$ has been assigned.

To extend our perspective and to consolidate our results, many more quantities are being measured; one of them being the double ratios [40], i.e. the ratio of the production cross-section times decay branching fraction between two different decay modes (say, $H \to XX$ and $H \to YY$) and check if it matches with the ratio as predicted by SM (i.e. 1). The double ratios for a given production mode is measured as

$$r\_XX\_YY = \frac{\mathrm{BR}(XX)}{\mathrm{BR}(YY)} \times \frac{\mathrm{BR}_{\mathrm{SM}}(YY)}{\mathrm{BR}_{\mathrm{SM}}(XX)}$$

such that the expected uncertainties related to the production and decay of the Higgs boson may cancel out. These double ratios only include the statistical uncertainties from the two decay modes involved, and may also contain, to some extent, the correlated experimental uncertainties (apart from the uncertainties from luminosity measurements and Higgs branching ratios). The right-hand term of this equation depends only on the partial decay widths rather than the total decay width of Higgs which may include the contribution from the invisible Higgs decay channels. Thus, the double ratios remove the ambiguities from the new physics scenarios and are only left with the statistical and some of the systematical uncertainties. While scanning the likelihood for the pseudodata as a function of a given r_XX_YY ratio, the production cross-section modifiers $\mu_{ggF+ttH}$ and $\mu_{\mathrm{VBF+VH}}$, as well as the other double ratios having $H \to YY$ decay mode in the denominator, are profiled. For example, consider $H \to WW$ decay mode as our denominator. Here we get four double ratios namely: $r\_\gamma\gamma\_WW$, r_ZZ_WW, $r\_\tau\tau\_WW$ and r_bb_WW.

The likelihood as a function of the double ratio, while profiling over all parameters can be measured for all possible permutations between the $H \to \gamma\gamma$, $H \to ZZ$, $H \to WW$, $H \to \tau\tau$ and $H \to bb$ channels. For this measurement, it is only necessary to assume that the same boson H is responsible for all the observed Higgs-like signals and that the separation of gluon fusion-like events and VBF-like events within the individual analysis based on the event kinematic properties is valid. Figure 2 shows the likelihood as a function of pairwise ratios of branching ratios corresponding to some pseudodata and signal values. The best-fit values are $r\_hgg\_hww = 1.139^{+0.641}_{-0.439}$ and $r\_hgg\_htt = 0.895^{+0.745}_{-0.395}$ which are in agreement with the SM expectation of one within uncertainty. The rather large uncertainty on the double ratios is dominated by lack of statistics in the $\gamma\gamma$ and *ZZ* channels while its systematic dominated in the other three channels.

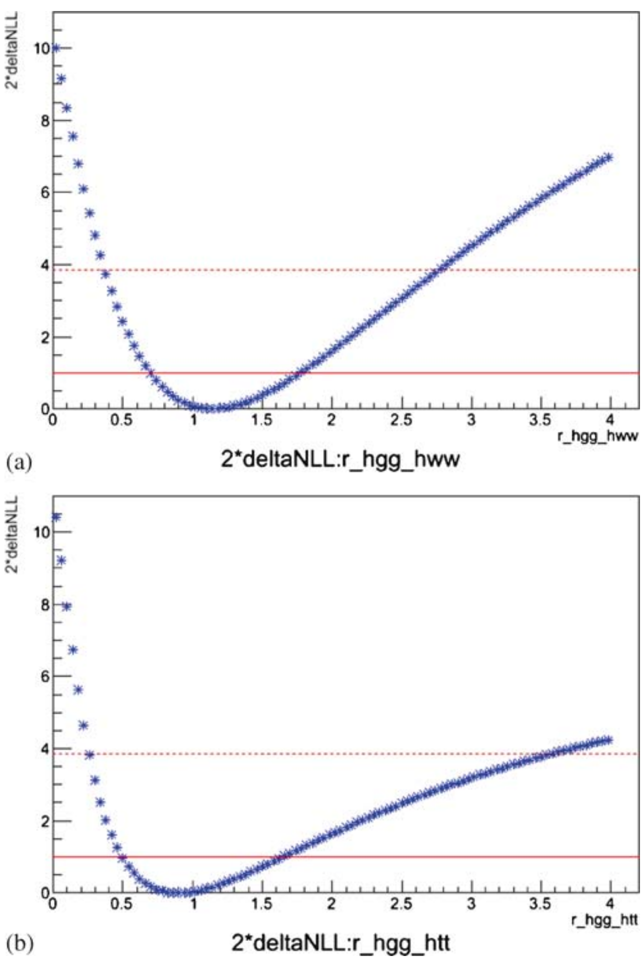

Figure 2. Likelihood curves for the ratio of branching ratios normalized to their SM expectations; (**a**) $H \to \gamma\gamma$ and $H \to WW$, (**b**) $H \to \gamma\gamma$ and $H \to \tau\tau$. Red horizontal line at 1 shows 68% CL and red dotted line at 3.84 shows 95% CL.

The LHC is gearing up for the higher centre of mass energy and high luminosity runs. The 13/14 TeV LHC run would provide approximately 100 fb^{-1} of data by the end of run 2, and as a result the uncertainties on the double ratios are expected to reduce drastically which would provide a more robust and precise measurement of the tests beyond the SM. Considering $H \to \gamma\gamma$ as the denominator in the double ratios, four double ratios can be obtained with *ZZ*, *WW*, $\tau\tau$ and *bb* each in the numerator. The statistical uncertainty on the $ZZ/\gamma\gamma$ ratio is ~21% with 25 fb^{-1} of 7 and 8 TeV data. This uncertainty is expected to reduce to about 6.5% with 100 fb^{-1} of 13 TeV data. Similarly, the statistical uncertainty with 25 fb^{-1} of 7 and 8 TeV data for $WW/\gamma\gamma$ double ratio is ~7% which is expected to reduce to 2.3% with 100 fb^{-1} of 13 TeV data, though this is a systematic dominated channel as is the case with the other two ratios involving $\tau\tau$ and *bb*. It is to be noted however, that the systematic and theoretical uncertainties are also expected to improve over this period which will further improve the accuracy of these measurements.

4. Conclusions

We performed a model-independent search for the SM Higgs boson at LHC. The 125 GeV particle that was discovered in 2012 could either be the long-sought SM Higgs boson or it might just turn out to be a particle predicted by the models of physics beyond SM. Therefore, it is very important to independently verify the nature of this new particle. Our approach of model-independent analysis can independently rediscover this new particle in a completely unbiased way. Though the recent measurement of properties of this particle hints that this particle is a SM Higgs boson but in order to remove any bias from these measurements, we propose to measure the ratios and double ratios of signal strength in a model-independent way. We are already working on these two measurements using the data collected by the CMS experiment. We hope that our approach will for ever settle the question of this new particle being a SM Higgs boson or something else.

Acknowledgements

The authors would like to thank the Department of Science and Technology (DST), Government of India for providing the financial support to work in the CMS experiment. They would also like to thank University of Delhi for providing R&D grants to partially support this work. Shivali Malhotra would also like to thank University Grants Commission (UGC), Government of India for providing financial support.

References

[1] ATLAS Collaboration, *Phys. Lett. B* **716**, 1 (2012)
[2] CMS Collaboration, *Phys. Lett. B* **716**, 30 (2012)
[3] P W Higgs, *Phys. Lett.* **12**, 132 (1964)
[4] F Englert and R Brout, *Phys. Rev. Lett.* **13**, 321 (1964)
[5] G S Guralnik, C R Hagen and T W B Kibble, *Phys. Rev. Lett.* **13**, 585 (1964)
[6] S L Glashow, *Nucl. Phys.* **22**, 579 (1961)
[7] S Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967)
[8] A Salam, Weak and electromagnetic interactions, in: *Elementary particle physics: Relativistic groups and analyticity* edited by N Svartholm (Almqvist & Wiskell, 1968) p. 367
[9] LHC Higgs Cross Section Working Group Collaboration, *TWiki*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections, (2012)
[10] ATLAS Collaboration, *Phys. Rev. Lett.* **108**, 111803 (2012)
[11] ATLAS Collaboration, *Phys. Lett. B* **710**, 383 (2012)
[12] ATLAS Collaboration, *Phys. Lett. B* **716**, 62 (2012)
[13] CMS Collaboration, *Phys. Lett. B* **710**, 403 (2012), arXiv:1202.1487
[14] CMS Collaboration, *Phys. Rev. Lett.* **108**, 111804 (2012), arXiv:1202.1997
[15] CMS Collaboration, *Phys. Lett. B* **710**, 91 (2012), arXiv:1202.1489
[16] ATLAS Collaboration, *J. High Energy Phys.* **09**, 070 (2012), arXiv:1206.5971
[17] CMS Collaboration, *J. High Energy Phys.* **05**, 104 (2014), arXiv:1401.5041
[18] ATLAS Collaboration, arXiv:1207.0210
[19] CMS Collaboration, *Phys. Lett. B* **710**, 284 (2012), arXiv:1202.4195
[20] CMS Collaboration, *J. High Energy Phys.* **05**, 145 (2013), arXiv:1303.0763
[21] CDF Collaboration, arXiv:1207.1704
[22] CDF Collaboration, *Phys. Rev. Lett.* **109**, 111804 (2012)
[23] CDF Collaboration, arXiv:1207.1711
[24] DØ Collaboration, arXiv:1207.5819
[25] DØ Collaboration, arXiv:1208.0653
[26] DØ Collaboration, arXiv:1207.5689
[27] DØ Collaboration, *Phys. Rev. D* **62**, 092004 (2000)
[28] DØ Collaboration, *Phys. Rev. Lett.* **86**, 3712 (2001)
[29] DØ Collaboration, *Phys. Rev. D* **64**, 012004 (2001)
[30] DØ Collaboration, *Phys. Rev. D* **85**, 092015 (2012)
[31] CDF Collaboration, *Phys. Rev. D* **78**, 012002 (2007)
[32] CDF Collaboration, *Phys. Rev. D* **79(1)**, 011101 (2009)
[33] CMS Collaboration, *MUSiC – An automated scan for deviations between data and Monte Carlo simulation*, CMS Physics Analysis Summary, CMS PAS EXO-08-005
[34] CMS Collaboration, *Model unspecific search for new Physics in pp collisions at* $\sqrt{s} = 7$ TeV, CMS Physics Analysis Summary, CMS PAS EXO-10-021
[35] ATLAS and CMS Collaborations, LHC Higgs Combination Group, Procedure for the LHC Higgs boson search combination in Summer 2011, Technical Report ATL-PHYS-PUB 2011-11, CMS NOTE 2011/005, 2011
[36] G Cowan, K Cranmer, E Gross and O Vitells, *Eur. Phys. J. C* **71**, 1554 (2011)
[37] CMS Collaboration, *Phys. Rev. Lett.* **110**, 081803 (2013), arXiv:1212.6639
[38] CMS Collaboration, *Search for the standard model Higgs boson in the* $H \to WW \to l\nu l\nu$ *channel in pp collisions at* $\sqrt{s} = 7$ *and* 8 TeV, CMS Physics Analysis Summary, CMS-PAS-HIG-13-003 (2013)
[39] LHC Higgs Cross Section Working Group: A David, A Denner, M Duehrssen, M Grazzini and others, arXiv:1209.0040
[40] CMS Collaboration, *Eur. Phys. J. C* **75**, 212 (2015)