ATLAS Note

ANA-HDBS-2021-31-INT1

20th January 2022

Draft version 0.1

HEFT interpretation of Di-Higgs searches in $b\bar{b}\gamma\gamma$, $b\bar{b}\tau\tau$ final states and their combination in ATLAS

The ATLAS Collaboration

This is a bare bones ATLAS document. Put the abstract for the document here.

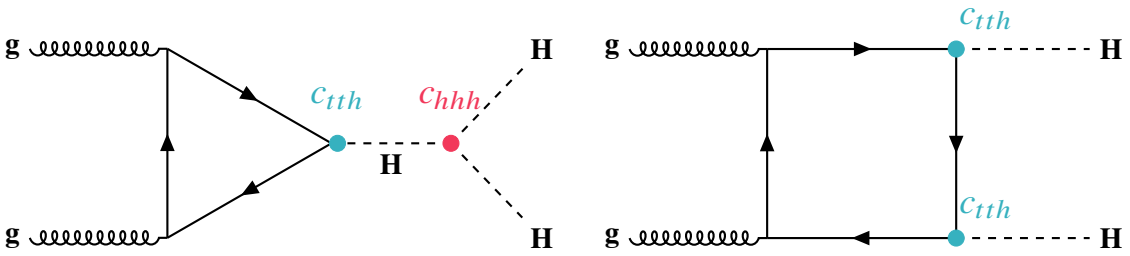
Contents

1 Introduction

The Higgs Boson (H), a cornerstone of the SM, was discovered by the ATLAS and CMS collaborations in 2012 **REF**. Since its discovery, numerous measurements of the spin and couplings have been performed **REF**, showing consistency with Standard Model (SM) prediction. The Higgs boson self-coupling, (λ_{HHH}), provides direct access to the shape of the Higgs potential, and its measurement is a primary physics goal of the LHC & its forthcoming upgrade, the High Luminosity LHC (HL-LHC). Modifications to the Higgs self-coupling can occur through physics beyond the SM (BSM), which can lead to an enhancement of the rate of Higgs pair-production. BSM enhancements can occur through either resonant processes ($X \rightarrow HH$) or through non-resonant processes, with this document focusing on the latter.

1.1 Non-resonant HH production

In the SM, the dominant HH production process is gluon-gluon fusion (ggF), accounting for over 90% of the total production rate. Two Feynman diagrams describe ggF HH production in the SM, shown in Figure 1.



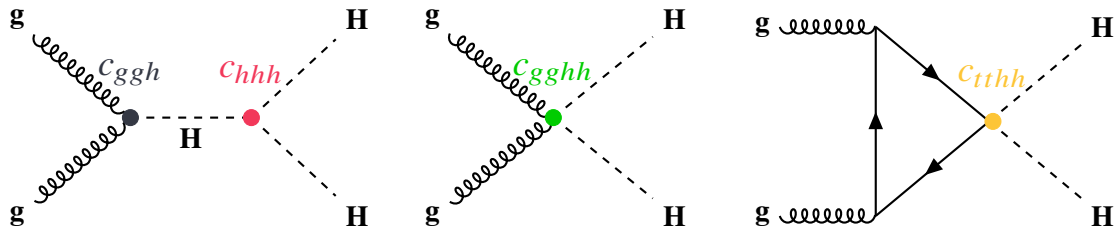
(a) ttH and HHH (b) ttH

Figure 1: HEFT Feynman diagrams for Higgs pair production through gluon-gluon fusion (orders 4). In SM, $c_{tth} = c_{hhh} = 1$ and $c_{tthh} = c_{ghh} = c_{gghh} = 0$.

The triangle diagram, shown in Figure 1(a), is sensitive to λ_{HHH}. The two diagrams shown in Figure 1 interfere destructively, resulting in a HH production cross-section of $\sigma^{SM}_{ggF} = 31.05$ fb which is around three orders of magnitude smaller than the dominant single Higgs production process. However, anomalous couplings in BSM scenarios can result in an enhanced HH production rate, and any observation of such an enhancement could be evidence for BSM physics. Anomalous couplings in HH production can be probed by parameterising the Lagrangian using an Effective Field Theory (EFT) approach, which will be described in the subsequent section.

1.2 EFT parameterisation

This document studies the potential modifications to the Higgs self-coupling by probing
The production of Higgs pairs is described in HEFT through 5 couplings: $c_{hhh}, c_{tth}, c_{tthh}, c_{ggh}$ and c_{gghh}. Where $c_{hhh}, c_t = \kappa_\lambda, c_{tth}$ represent the coupling strengh modifiers for the Higgs self coupling and the top yukawa coupling and therefore in the SM $c_{hhh} = c_{tth} = 1$ whereas c_{tthh}, c_{ggh} and c_{gghh} represent the beyond the SM processes and therefore $c_{tt} = c_{ggh} = c_{gghh} = 0$ in the SM. The BSM Feynman diagrams for Higgs pair production in HEFT are shown in Figure 2.



(a) ggH and HHH (b) ggHH (c) ttHH

Figure 2: HEFT Feynman diagrams for Higgs pair production through gluon-gluon fusion (orders 6). In SM, $c_{tth} = c_{hhh} = 1$ and $c_{tthh} = c_{ghh} = c_{gghh} = 0$.

2 Data and simulation samples

- Data and MC (same as combination)

Both the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau\tau$ analyses use a consistent set of Monte Carlo (MC) samples, as shown in this section. A consistent value of 125 GeV is used across the two analyses for the Higgs boson mass in the MC simulation for both single Higgs production and pair-production.

SM non-resonant HH via ggF production was simulated using full next-to-leading order (NLO) corrections and finite top mass effects using **POWHEG-BOX v2 generator**, using the **PDF4LHC15_nlo_30_pdfas** parton distribution function (PDF) set. **Pythia 8.244** was used for the parton shower and hadronization simulation, utilising the **NNPDF23LO PDF set** and **A14 tune**. The simulation of the HEFT benchmark signals is described in the next section.

In the $b\bar{b}\gamma\gamma$ channel, the background contribution from single Higgs production via the ggF, VBF, associated VH, $t\bar{t}H$, tH and $b\bar{b}H$ production processes were considered. The tH process is separated into the $s-/t-$channel (tHq) and the W-channel (tWH). The continuum processes $\gamma\gamma$+jets and $t\bar{t}\gamma\gamma$ are also considered.

In the $b\bar{b}\tau\tau$ channel, background processes including the $t\bar{t}$, $W\bar{t}$, V+jets, diboson processes, $t\bar{t}V$ and the single Higgs production via the ggF, VBF, associated VH and $t\bar{t}H$ production processes were considered. Fake τ background from multijet processes is estimated using data.

A summary of the MC simulation setups used for the SM backgrounds is shown in Table 1

Process	Generator	PDF	Showering	Tune

Table 1: Summary of the simulation used for the MC backgrounds in the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau\tau$ analyses.

2.1 HEFT re-weighting

It is very computationally expensive to simulate HH processes with the multiple possible variations of the Wilson coefficients described by EFTs and we re-weight the SM sample to obtain the different BSM samples described by HEFT.

The cross section for HH production via gluon-gluon fusion, $\sigma(HH)$, normalized to its SM value $\sigma^{SM}(HH)$, can be parametrized at next-to-leading order (NLO) by the expression given in Equation 1.

$$\begin{aligned}\frac{\sigma(HH)}{\sigma^{SM}(HH)} = Poly(\mathbf{A}) = A_1 c_{tth}^4 + A_2 c_{tthh}^2 + (A_3 c_{tth}^2 + A_4 c_g^2)\cdot c_{hhh}^2 + A_5 c_{gghh}^2 \\ +(A_6 c_{tthh} + A7 c_{tth} c_{hhh})\cdot c_{hhh}^2 + (A_8 c_{tth} c_{hhh} + A_9 c_{ggh} chhh)\cdot c_{tthh} + A_{10} c_{tthh} c_{gghh} \\ +(A_{11} c_{ggh} c_{hhh} + A_{12} c_{gghh})\cdot c_{tth}^2 + (A_{13} c_{hhh} c_{ggh} + A_{14} c_{gghh})\cdot c_{tth} c_{hhh} \\ +A_{15} c_{ggh} c_{gghh} c_{hhh} + A_{16} c_{tth}^3 c_{ggh} + A_{17} c_{tth} c_{tthh} c_{ggh} + A_{18} c_{tth} c^2 c_{ggh} c_{hhh} \\ +A_{19} c_{tth} c_{ggh} c_{gghh} + A_{20} c_{tth}^2 c_{ggh}^2 + A_{21} c_{tthh} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} + A_{23} c_{ggh}^2 c_{gghh} \end{aligned} \quad (1)$$

To predict the difference in the m_{HH} shape of the BSM process from the **A** coefficients events are weighted with w_{EFT}, defined in Equation 2. The HEFT cross section, $\sigma_{HEFT}(HH)$, can be obtained through Equation 3.

$$w_{EFT} = \frac{Poly(\mathbf{A}|m_{HH})}{Poly(\mathbf{A})} \quad (2)$$

$$\sigma_{EFT}(HH) = Poly(\mathbf{A}) \times \sigma_{SM}(HH) \quad (3)$$

3 Individual final states

3.1 $HH \to b\bar{b}\gamma\gamma$

3.2 $HH \to b\bar{b}\tau\tau$

The $HH \to b\bar{b}\tau^+\tau^-$ benefits from a relatively large branching fraction (7.3%)while also having a smaller SM background than some of the HH decay channels. This channel splits events by the decay of the tau leptons, with lephad events containing one leptonic tau and one hadronic decaying tau, while hadhad events contain two hadronic decaying taus. Lephad events use a combination of single-lepton triggers (SLT) and lepton + tau triggers (LTT). Hadhad events utilise the single-tau triggers (STT) and di-tau triggers (DTT). For both lephad and hadhad events, priority is given to events which trigger the single-lepton/tau trigger.

Lephad events are selected by first requiring exactly one electron or muon and exactly one hadronic tau of opposite charge to the lepton. Hadhad events require exactly two hadronic taus with opposite charge and veto events with electrons or muons. Both lephad and hadhad require two $b-$tagged jets (DL1r 77% working point) and the invariant mass of the $\tau\tau$ system (calculated using the Missing Mass Calculator (MMC) [1]) to be > 60 GeV. Events passing this preliminary selection form the signal regions (SRs) for the lephad and hadhad categories. Events in the lephad category are separated into SLT and LTT SRs by the trigger used to capture the event, while the hadhad category has a single SR.

The dominant SM backgrounds for the $b\bar{b}\tau\tau$ channel arise from top-quark, W/Z+jets, diboson, single Higgs and multi-jet production processes. The MC estimates for the $t\bar{t}$ and Z plus heavy flavour processes

are normalised to data in a dedicated control region (CR). For both the lephad and hadhad categories, an estimate of the rate of quark- or gluon-initiated jets faking a hadronic tau (fake tau rate) is made. For the lephad category, the method closely follows that described in [2]. Fake factors are derived for the $t\bar{t}$ and multi-jet processes, which are then combined to give an overall estimation of the fake tau rate. For the hadhad category, a fake factor method is used for the multi-jet process, while $t\bar{t}$ events with fake taus are estimated using a scale factor correction to the MC estimate.

In both the lephad and hadhad categories, a multivariate classifer is trained to separate the SM HH signal from the backgrounds. For the lephad category, a deep neural network (DNN) is trained in each of the SRs using 11 (8) kinematic and angular observables for SLT (LTT) events. A boosted decision tree (BDT) classifier is trained for hadhad events using five kinematic and angular observables.

4 HEFT interpretation

5 Systematic uncertainties

6 Statistical Interpretation

A combined likelihood function is constructed to take into account the data, the models and the systematic uncertainties (correlated where needed) from each of the input channels. The form of this function is shown in Equation 4.

$$\mathcal{L}\left(\mathcal{D}, \mathcal{G}|\mu, \alpha\right) = \prod_{c \in C} \text{Pois}\left(n_c|\nu_c\left(\mu, \alpha\right)\right) \prod_{e=1}^{n_c} f_c\left(x_{ce}|\mu, \alpha\right) \prod_{p \in \mathcal{S}} f_p\left(a_p|\alpha_p\right) \tag{4}$$

The individual terms of Equation 4 are outlined below:

- $\mathcal{D}$, the set of observed or Asimov data
- $\mathcal{G} = \{a_p\}$, the set of global observables, each of which corresponds to a measurement of a single experiment, theoretical or statistical parameter.
- μ, the parameter of interest (POI) to be measured in the analysis. In general it is either the signal strength (relative or absolute) of the model being tested. It can also be a parameter of a model which varies the signal prediction.
- α, the nuisance parameters (NPs) with respect to μ, such as the experimental or systematic uncertainties.
- Index $c \in C$, indicating one channel (c) in the set (C) of channels under consideration.
- Index e, indicating one event in n_c number of events in the observed or Asimov data in the channel c.
- Index $p \in \mathcal{S}$, numbering the constrained NPs in the set ($\mathcal{S}$).
- Pois, the probability function of a Poisson distribution to model the total number of events.
- n_c, the total number of events in the data $\mathcal{D}$ of the channel c.

- $\nu_c(\mu, \alpha)$, the total number of events as a function of the POI and the NPs.
- $f_c\,(x_{ce}|\mu, \alpha)$, the expected distribution of the variable x_{ce} in the channel c, as a funcion of the POI and the NPs.
- x_{ce}, the value of the variable x of an event e, used for the fit in the search for signal in the channel c.
- $f_p\,(a_p|\alpha_p)$, the constraint function for a given NP with the prediction α_p and the measurement a_p.

A profile likehood ratio is constructed using the asymptotic formulae shown in Equation 5:

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{\mathcal{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathcal{L}(0,\hat{\hat{\theta}}(0))} & \hat{\mu} < 0, \\ \frac{\mathcal{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\theta})} & \hat{\mu} > 0 \end{cases} \tag{5}$$

Limit setting is performed by constructing a test statistic, $\tilde{q}_\mu$, using the profile likehood ratio as described in Equation 6:

$$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{6}$$

7 Results

7.1 $HH \to b\bar{b}\gamma\gamma$

7.2 $HH \to b\bar{b}\tau\tau$

7.2.1 $b\bar{b}\tau\tau$ lephad channel

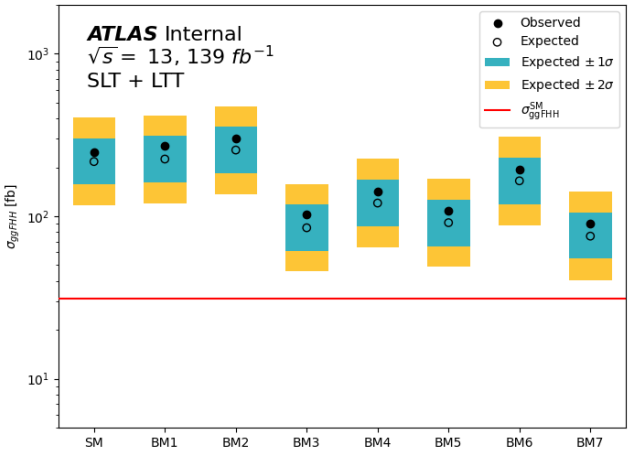
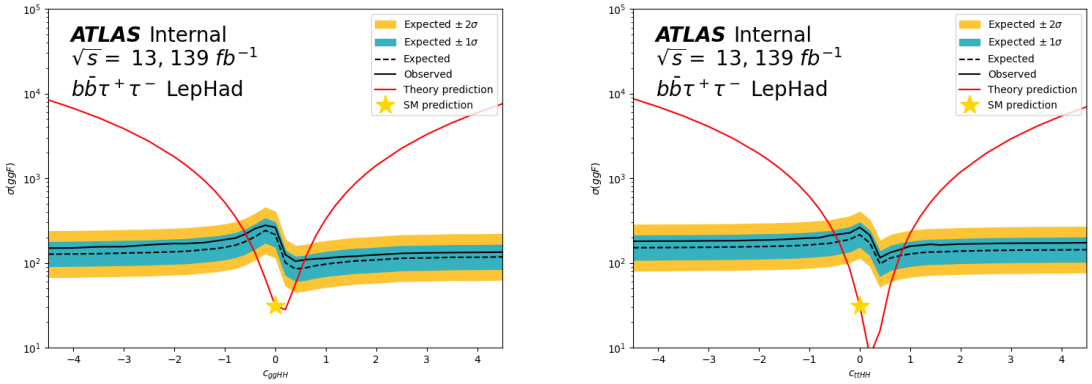


Figure 3: Cross-section limits for SM HH production and the 7 HEFT benchmarks in the $b\bar{b}\tau\tau$ lephad channel.



(a) (b)

Figure 4: Cross-section limit scans for a range of c_{gghh} (Figure 4(a)) and c_{tthh} (Figure 4(b)) couplings in the $b\bar{b}\tau\tau$ lephad channel.

7.3 Combination

8 Conclusions

References

[1] A. Elagin, P. Murat, A. Pranko and A. Safonov,
A new mass reconstruction technique for resonances decaying to $\tau\tau$,
Nucl. Instrum. Meth. A **654** (2011) 481, arXiv: 1012.4686 [hep-ex] (cit. on p. 5).

[2] D. Alvarez Piqueras et al., *Search for neutral MSSM Higgs bosons $H/A \to \tau_{lep}\tau_{had}$ and $Z' \to \tau_{lep}\tau_{had}$ produced in 13 TeV collisions with the ATLAS detector*, tech. rep., CERN, 2016,
URL: https://cds.cern.ch/record/2131232 (cit. on p. 6).

The supporting notes for the analysis should also contain a list of contributors. This information should usually be included in `mydocument-metadata.tex`. The list should be printed either here or before the Table of Contents.

List of contributions

Appendices

In an ATLAS note, use the appendices to include all the technical details of your work that are relevant for the ATLAS Collaboration only (e.g. dataset details, software release used). This information should be printed after the Bibliography.