Probing Trilinear Higgs Self–coupling at the HL-LHC via Multivariate Analysis

Jubin Park
IUEP and Chonnam National University

Collaborators: Prof. Kingman Cheung, Prof. Jae Sik Lee, Dr. Chih-Ting Lu, Dr. Jung Chang
Ref: 1. An exploratory study of Higgs-boson pair production (JHEP 1508(2015) 133)
   2. Higgs-boson-pair production $H(\to b\bar{b})H(\to \gamma\gamma)$ from gluon fusion
      at the HL-LHC and HL-100 TeV hadron collider (Arxiv:1804.07130)
   3. Probing Trilinear Higgs Self–coupling at the HL-LHC via Multivariate Analysis
      (Arxiv:1908.00753)

ML: (from SI 2019)
   1. Tagging boosted weak gauge bosons with deep learning – Prof. Cheng-Wei Chiang
   2. A New Architecture of Classification Model with the Abstraction of Physical Symmetry,
      - Kayoung Ban
   3. Portraying Double Higgs at the LHC – Minho Kim
      (bbWW dilepton channel)
   4. Complex-valued Neural Arithmetic Logic Units (CALU) for the Abstraction of
      Physical Symmetries – Dr. Won Sang Cho
Higgs In the SM

• Higgs field ($h$) : responsible for
  ① the spontaneous EW symmetry breaking
  ② the generation of masses of all the SM particle

• The potential is characterized by only two parameters :
  ① vacuum expectation value $v$
  ② the Higgs mass $m_H$

\[ v = \frac{1}{\sqrt{2} G_F} \approx 246 \text{ GeV} \]
\[ V_{SM}(h) = \frac{1}{2} m_H^2 h^2 + \lambda_3 v h^3 + \frac{1}{4} \lambda_4 h^4 \]
Trilinear and Quartic Higgs boson coupling in the SM

\[ \lambda_3^{SM} = \lambda_4^{SM} = \frac{m_H^2}{2\nu^2} \]

New Physics can affect the Higgs potential form

Sizeable departures from the SM form

\[ \lambda_3 = \lambda_3^{SM} + \delta \lambda_3^{SM}, \quad \lambda_4 = \lambda_4^{SM} + \delta \lambda_4^{SM} \]
Trilinear Higgs boson coupling

\[ \lambda_3 = \lambda_3^{SM} + \delta\lambda_3^{SM} \]
Therefore, it is really important to check (or directly measure) this trilinear Higgs coupling $\lambda_3$ to verify that SM is correct.
So, how to probe it?

Higgs pair production!
We now focus on Higgs pair production at the Hadron collider.
Why **Higgs pair production** so interesting?

- Allows accessing crucial components of the Higgs sector !!!
- Can probe the Higgs self-coupling
- Can help to reconstruct the electroweak symmetry breaking potential
- May reveal the doublet nature of the Higgs by means of the $hhVV$ coupling

A high priority goal on the physics program at the future collider
Higgs pair productions at the LHC

Production modes

Gluon Fusion

Vector Boson Fusion

Top associated productions

Higgs strahlung
Why **Higgs pair production** so difficult?

In the leading gluon fusion production mode, the cross section at 14 TeV is only **45 fb** (in the SM), further suppressed by each decay branching fractions.

45 fb ↔ NNLO accuracy including NNLL gluon resummation in the infinite top quark mass approximation.
Direct access to Higgs self-couplings

Multi-Higgs production !!!!

However,

$\mathcal{O}(\text{pb})$ $\gg$ $\mathcal{O}(\text{fb})$ $\gg$ $\mathcal{O}(\text{ab})$

Experimentally very challenging !!
So, at the HL-LHC

Higgs pair production
(Double Higgs production)
Challenging !!!

Triple Higgs production
Impossible !!!
Search channels for Higgs pair production at Collider

Our Channel reconstruct $\tau / W$

<table>
<thead>
<tr>
<th>Decay channels</th>
<th>$HH \rightarrow bb\gamma\gamma$</th>
<th>$HH \rightarrow bb\tau\tau$</th>
<th>$HH \rightarrow bbWW$</th>
<th>$HH \rightarrow bbbb$</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branching ratios</td>
<td>0.263%</td>
<td>7.29%</td>
<td>24.8%</td>
<td>33.3%</td>
<td></td>
</tr>
</tbody>
</table>

Huge $t\bar{t}$ BG
Huge hadronic BG

Expected events with $3 ~ ab^{-1}$

<table>
<thead>
<tr>
<th>Decay channels</th>
<th>Expected events</th>
<th>Expected events</th>
<th>Expected events</th>
<th>Expected events</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$HH \rightarrow bb\gamma\gamma$</td>
<td>290</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$HH \rightarrow bb\tau\tau$</td>
<td>8000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$HH \rightarrow bbWW$</td>
<td>27000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$HH \rightarrow bbbb$</td>
<td>37000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Recent updates

- HH NLO generator – POWHEG BOX V2
  Use of NLO kinematic distributions (or variables)

- Improved yields and significance($Z$)

- Improved Likelihood fit using the $M_{HH}$ kinematic distribution

- Use of the latest multivariate analysis using Toolkit for Multivariate Data Analysis (TMVA) with ROOTv6.18

<table>
<thead>
<tr>
<th>Signal</th>
<th>Generator/Parton Shower</th>
<th>$\sigma \cdot BR$ [fb]</th>
<th>Order</th>
<th>PDF used</th>
</tr>
</thead>
<tbody>
<tr>
<td>$gg \rightarrow HH \rightarrow bb\gamma\gamma$</td>
<td>POWHEG-BOX-V2/PYTHIA8</td>
<td>0.096</td>
<td>NNLO</td>
<td>PDF4LHC15_nlo</td>
</tr>
<tr>
<td>Backgrounds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background(BG)</td>
<td>Process</td>
<td>Generator/Parton Shower</td>
<td>$\sigma \cdot BR$ [fb]</td>
<td>Order</td>
</tr>
<tr>
<td>Single-Higgs</td>
<td>$ggH(\rightarrow \gamma\gamma)$</td>
<td>POWHEG-BOX/PYTHIA8</td>
<td>1.20 $\times$ 10$^3$</td>
<td>NNNLO</td>
</tr>
<tr>
<td>associated BG</td>
<td>$t\bar{t}H(\rightarrow \gamma\gamma)$</td>
<td>PYTHIA8/PYTHIA8</td>
<td>1.37</td>
<td>NLO</td>
</tr>
<tr>
<td></td>
<td>$ZH(\rightarrow \gamma\gamma)$</td>
<td>PYTHIA8/PYTHIA8</td>
<td>2.24</td>
<td>NLO</td>
</tr>
<tr>
<td></td>
<td>$t\bar{t}H(\rightarrow \gamma\gamma)$</td>
<td>PYTHIA8/PYTHIA8</td>
<td>1.26</td>
<td>NLO</td>
</tr>
<tr>
<td>Non-resonant BG</td>
<td>$b\bar{b}\gamma\gamma$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>1.12 $\times$ 10$^3$</td>
<td>LO</td>
</tr>
<tr>
<td></td>
<td>$c\bar{c}\gamma\gamma$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>1.08 $\times$ 10$^3$</td>
<td>LO</td>
</tr>
<tr>
<td></td>
<td>$j\bar{j}\gamma\gamma$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>1.40 $\times$ 10$^4$</td>
<td>LO</td>
</tr>
<tr>
<td></td>
<td>$b\bar{b}j\gamma$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>2.72 $\times$ 10$^5$</td>
<td>LO</td>
</tr>
<tr>
<td></td>
<td>$c\bar{c}j\gamma$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>9.17 $\times$ 10$^5$</td>
<td>LO</td>
</tr>
<tr>
<td></td>
<td>$b\bar{b}j\bar{j}$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>3.00 $\times$ 10$^8$</td>
<td>LO</td>
</tr>
<tr>
<td></td>
<td>$Z(\rightarrow b\bar{b})\gamma\gamma$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>5.03</td>
<td>LO</td>
</tr>
<tr>
<td>$t\bar{t}$ and $t\bar{t}\gamma$ BG</td>
<td>$t\bar{t}$</td>
<td>POWHEG BOX/PYTHIA8</td>
<td>5.30 $\times$ 10$^5$</td>
<td>NNLO</td>
</tr>
<tr>
<td></td>
<td>+NLL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(\geq 1$ lepton)</td>
<td>$t\bar{t}\gamma$</td>
<td>MG5_aMC@NLO/PYTHIA8</td>
<td>1.60 $\times$ 10$^3$</td>
<td>NLO</td>
</tr>
</tbody>
</table>
Analysis Methods

- **Cut Based Analysis**

- **Machine Learning**
  
  Multi-Variate Analysis : BDT (Boosted Decision Tree)  
  (work in progress)

**TMVA (Toolkit for Multivariate Data Analysis with ROOT)**

A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss,

*TMVA - Toolkit for Multivariate Data Analysis,*

Various ML methods in TMVA

Ideal line of ROC curve

ROC = Receiver Operating Characteristic curve

A better direction:

A good way to illustrate the performance of given classifier

AUC = Area under the ROC curve

AUC = 1
BDT-related methods give the best results.
BDT machine optimized with $\lambda_{3H} = 1$
Impact of NLO kinematic distributions
Results: Improved Expected Yields and Significance (Z)
<table>
<thead>
<tr>
<th>Signal and Backgrounds</th>
<th>Expected yields (3000 fb(^{-1}))</th>
<th>Pre-Selection</th>
<th>BDT(_{SM})</th>
<th>Cut-and-Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H(b\bar{b})H(\gamma\gamma), \lambda_{3H} = -5)</td>
<td>223.22</td>
<td>49.36</td>
<td>90.19</td>
<td></td>
</tr>
<tr>
<td>(H(b\bar{b})H(\gamma\gamma), \lambda_{3H} = 0)</td>
<td>33.69</td>
<td>10.22</td>
<td>16.70</td>
<td></td>
</tr>
<tr>
<td>(H(b\bar{b})H(\gamma\gamma), \lambda_{3H} = 1)</td>
<td><strong>17.77</strong></td>
<td><strong>6.19</strong></td>
<td><strong>9.63</strong></td>
<td></td>
</tr>
<tr>
<td>(H(b\bar{b})H(\gamma\gamma), \lambda_{3H} = 5)</td>
<td>26.37</td>
<td>2.76</td>
<td>6.77</td>
<td></td>
</tr>
<tr>
<td>(ggH(\gamma\gamma))</td>
<td>68.76</td>
<td>1.92</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>(t\bar{t}H(\gamma\gamma))</td>
<td>158.14</td>
<td>3.26</td>
<td>13.21</td>
<td></td>
</tr>
<tr>
<td>(ZH(\gamma\gamma))</td>
<td>23.89</td>
<td>1.28</td>
<td>3.62</td>
<td></td>
</tr>
<tr>
<td>(b\bar{b}H(\gamma\gamma))</td>
<td>2.52</td>
<td>0.08</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>(b\bar{b}\gamma\gamma)</td>
<td>6968.37</td>
<td>2.69</td>
<td>15.09</td>
<td></td>
</tr>
<tr>
<td>(c\bar{c}\gamma\gamma)</td>
<td>7051.90</td>
<td>1.30</td>
<td>7.13</td>
<td></td>
</tr>
<tr>
<td>(jj\gamma\gamma)</td>
<td>1015.48</td>
<td>0.66</td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>(b\bar{b}jj)</td>
<td>10018.32</td>
<td>0.82</td>
<td>13.91</td>
<td></td>
</tr>
<tr>
<td>(c\bar{c}jj)</td>
<td>4679.49</td>
<td>0.82</td>
<td>4.78</td>
<td></td>
</tr>
<tr>
<td>(b\bar{b}jj)</td>
<td>2525.67</td>
<td>0.36</td>
<td>3.42</td>
<td></td>
</tr>
<tr>
<td>(Z(b\bar{b})\gamma\gamma)</td>
<td>184.09</td>
<td>0.15</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>(t\bar{t} (\geq 1 \text{ leptons}))</td>
<td>5433.74</td>
<td><strong>0.22</strong></td>
<td><strong>4.98</strong></td>
<td></td>
</tr>
<tr>
<td>(t\bar{t}\gamma (\geq 1 \text{ leptons}))</td>
<td>1916.50</td>
<td>0.49</td>
<td>3.61</td>
<td></td>
</tr>
<tr>
<td>Total Background</td>
<td>40046.87</td>
<td><strong>14.04</strong></td>
<td><strong>80.26</strong></td>
<td></td>
</tr>
<tr>
<td>Significance (Z, \lambda_{3H} = 1)</td>
<td></td>
<td><strong>1.55</strong></td>
<td><strong>1.05</strong></td>
<td></td>
</tr>
</tbody>
</table>

**C&C**

- # of signal = 9.63
- # of bg = 80.26
- ∴ significance \(Z = 1.05\)

**BDT\(_{SM}\)**

- # of signal = 6.19
- # of bg = 14.04
- ∴ BDT-improved \(Z = 1.55\)

48% enhancement on \(Z\)
Significance of the signal over the background versus $\lambda_{3H}$ at the HL-LHC

95% CL: $0.53 \leq \lambda_{3H} \leq 6.80$ for the central line
Improved Likelihood fit using the (NLO) $M_{HH}$ kinematic distribution

<table>
<thead>
<tr>
<th>Scenario</th>
<th>1σ CI</th>
<th>2σ CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDTλ SM, $\lambda_3^H = 1$ nominal set</td>
<td>$-0.1 &lt; \lambda_3^H &lt; 2.8$ and $4.9 &lt; \lambda_3^H &lt; 7.0$</td>
<td>$-1.1 &lt; \lambda_3^H &lt; 8.1$</td>
</tr>
<tr>
<td>BDTλ SM, $\lambda_3^H = 0$ nominal set</td>
<td>$-1.0 &lt; \lambda_3^H &lt; 1.0$ and $6.5 &lt; \lambda_3^H &lt; 7.6$</td>
<td>$-1.9 &lt; \lambda_3^H &lt; 2.9$ and $4.6 &lt; \lambda_3^H &lt; 8.7$</td>
</tr>
</tbody>
</table>
We find a bulk region of $0.5 \leq \lambda_{3H} \leq 4.5$, where it is hard to pin down the trilinear coupling.
Conclusion
• **HL-LHC**: constraint the $\lambda_{3H}$

  With 3000 fb$^{-1}$

1. Cut-Based Analysis: $-1.1 < \lambda_{3H} < 7.0$ at 95% CL,

   $$Z = 1.05 \ (\lambda_{3H}=1)$$

2. BDT Analysis + NLO$\_\text{dist}$: $0.5 < \lambda_{3H} < 6.8$ at 95% CL,

   $$Z = 1.55 \ (\lambda_{3H}=1)$$

3. We find a bulk region of $0.5 \leq \lambda_{3H} \leq 4.5$,

   where it is hard to pin down the trilinear coupling.
Last, but not least

• Combined analysis : $b\bar{b}b\bar{b} + b\bar{b}\gamma\gamma + b\bar{b}\tau\tau \rightarrow Z \ (\geq 3\sigma)$

• Advanced technology in the future ...
  Increased luminosity, improved tagging efficiency, improved resolution and so on....

• More precise simulation, higher order QCD correction ..
  Improved MC Event generators (at the NLO, NNLO QCD order),
  QCD NLO, NNLO, NNNLO corrections ...
  ....
It may be possible to see this trilinear Higgs coupling at the HL-LHC.