Probing the CP nature of the Higgs couplings in $b\bar{b}h$ and $t\bar{t}h$ at the LHC

Braga meeting – Universidade do Minho, Escola de Ciências

Rodrigo Capucha

Collaboration: A. Onofre, D. Azevedo, R. Santos

30-31 January 2020
Introduction

• In 2012, a new scalar particle with a mass close to 125 GeV, later identified as the Higgs boson, was discovered at the Large Hadron Collider (LHC). Electroweak symmetry breaking confirmed and SM completed. ATLAS, Phys. Lett. B716 (2012), and CMS, Phys. Lett. B 716 (2012).

• Tested in a wide range of precise experimental measurements at colliders. No experimental result strongly deviates from its predictions (few exceptions: B meson decays and muon’s anomalous magnetic moment).

• But it is incomplete. Some unanswered question: dark matter, matter-antimatter asymmetry, gravity.
CP-Violation (CPV)

• **CP**: symmetry that changes particle into antiparticle and inverts the spatial coordinates of a particle.

• Why does it matter? According to Sakharov’s conditions for baryogenesis, matter-antimatter discrepancy may be explained if CP is violated.

• Is there CPV in the SM? Yes, in the Yukawa sector, but not enough. **New sources of CPV are needed.** A possibility is to add CPV in an extended scalar sector (like the C2HDM hep-ph/0211371).

• How to look for CPV at the LHC?
The SM Higgs is CP-even. What about the discovered scalar? Pure CP-odd Higgs already ruled out at 99.98% in VH production \cite{PhysLettB759,672}. Is it CP-even? Not necessarily. It may have a CP-even and a CP-odd component. This would imply CPV in the scalar sector.

- The Higgs CP can be probed at the LHC by looking for CP-sensitive observables in Higgs decays into boson pairs, fermions, or in Higgs production channels. $hVV$ ($V=Z,W$) couplings: only the CP-even component is projected out. Fermions: both components can contribute equally. Precise measurement of the Yukawa couplings still lacking.

- Two processes were considered: $b\bar{b}h$ and $t\bar{t}h$. Complicated backgrounds and relatively low rates, but we are directly probing the vertices. Difficult channel, but $t\bar{t}h$ has already been observed. CMS, Phys. Rev. Lett. 120 (2018) and ATLAS, Phys. Lett. B 784 (2018)
The Model

- **SM + generic CP-violating Yukawa coupling** for $f = b, t$.

\[ \mathcal{L}_{h \bar{f} f} = - \sum_f g_{h \bar{f} f} \kappa_{h \bar{f} f} \bar{f} (\cos \alpha + i \sin \alpha \gamma_5) f h, \]

- $g_{h \bar{f} f}$ are the SM couplings. We set $k_{h \bar{f} f} = 1$. Two limits considered: CP-even ($\alpha = 0$, $h = H$) and CP-odd ($\alpha = \pi/2$, $h = A$).

- For $b\bar{b}h$: $h \rightarrow \tau^+\tau^-$. $m_h = 10 – 125$ GeV.

- For $t\bar{t}h$: $h \rightarrow b\bar{b}$, $t (\bar{t}) \rightarrow W^+ b$ ($W^- \bar{b}$) and $W^+ (W^-) \rightarrow l^+ \nu_l (l^- \bar{\nu}_l)$ (dileptonic state), with $l = e, \mu$. $m_h = 40 – 500$ GeV.
Generation of signal and background events

- Generation of **pp collisions at 13 TeV** done with the Monte Carlo event generator MadGraph5_aMC@NLO. JHEP 1407, 079 (2014)

- h = H, A implemented with the **HC_NLO_X0 model**, with NLO corrections. JHEP 1311 (2013)
  Backgrounds simulated with the **SM implementation**.

- **MadSpin** used to decay heavy resonances (t, h, and V) preserving full spin correlations. JHEP 1303, 015 (2013)

- **Pythia 6** used for showering and hadronization. JHEP 0605, 026 (2006)

- **DELPHES 3** used for a fast detector simulation of a LHC-like experiment. JHEP 1402, 057 (2014)

- Analysis done with **MadAnalysis 5**. EPJC 74, no 10, 3103 (2014)
<table>
<thead>
<tr>
<th>Topology</th>
<th>Order</th>
<th>Generated cross-section (pb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t \bar{t} H^{SM}$</td>
<td>NLO</td>
<td>0.025</td>
</tr>
<tr>
<td>$t \bar{t} b \bar{b}$</td>
<td>NLO</td>
<td>0.79</td>
</tr>
<tr>
<td>$t \bar{t} + 3$ jets</td>
<td>LO</td>
<td>37.89</td>
</tr>
<tr>
<td>$t \bar{t} V +$ jets</td>
<td>LO</td>
<td>0.0618</td>
</tr>
<tr>
<td>Single top s-channel</td>
<td>LO</td>
<td>2.1916</td>
</tr>
<tr>
<td>Single top t-channel + jets</td>
<td>LO</td>
<td>46.863</td>
</tr>
<tr>
<td>Single top Wt-channel</td>
<td>LO</td>
<td>15.1827</td>
</tr>
<tr>
<td>$W + 4$ jets</td>
<td>LO</td>
<td>34500</td>
</tr>
<tr>
<td>$W b \bar{b} + 2$ jets</td>
<td>LO</td>
<td>289</td>
</tr>
<tr>
<td>$Z + 4$ jets</td>
<td>LO</td>
<td>3120</td>
</tr>
<tr>
<td>$Z b \bar{b} + 2$ jets</td>
<td>LO</td>
<td>123</td>
</tr>
<tr>
<td>$W W + 3$ jets</td>
<td>LO</td>
<td>84.2</td>
</tr>
<tr>
<td>$W Z + 3$ jets</td>
<td>LO</td>
<td>37.9</td>
</tr>
<tr>
<td>$Z Z + 3$ jets</td>
<td>LO</td>
<td>11</td>
</tr>
</tbody>
</table>

- **Backgrounds (only for $t \bar{t}h$):** all processes that can lead to 4 jets and two leptons of opposite charge. Jets: gluons or non-b quarks.
Event Selection and Reconstruction (for $t\bar{t}h$ and backgrounds)

• Results for $b\bar{b}h$ only at parton level.

• Selection cuts:
  ➢ $N_{\text{jets}} \geq 4$ and $N_{\text{lep}} \geq 2$ (one-to-one correspondence)
  ➢ $p_T > 20$ GeV and $|\eta| < 2.5$ (experimental limitations)
  ➢ $|m_{l^+l^-} - m_Z| > 10$ GeV and at least 3 b-tagged jets (after reconstruction).

• Reconstruction: matching and actual reconstruction done separately.
  ➢ Boosted decision tree (BDT) algorithm to assign jets to the b-quarks of the top-quarks and Higgs.
    Constraints: only 6 jets with highest $p_T$. $m_{l^+b_t} < 150$, $m_{l^-\bar{b}_t} < 150$ and $20 < m_{b\bar{h}\bar{h}} < 300$ GeV.
  ➢ After matching, 6 objects are identified. These are used to find the momentum of the remaining particles.
Event Selection and Reconstruction (for $t\bar{t}h$)

- All missing energy due to neutrinos. With $m_\nu = 0$, we are left with six unknowns:

$$\begin{align*}
(p_{l+} + p_\nu)^2 &= m_{W+}^2 \\
(p_{l-} + p_{\bar{\nu}})^2 &= m_{W-}^2 \\
(p_{W+} + p_{b}\bar{t})^2 &= m_t^2 \\
(p_{W-} + p_{\bar{b}t})^2 &= m_{\tilde{t}}^2 \\
p_\nu^x + p_{\bar{\nu}}^x &= E^x \\
p_\nu^y + p_{\bar{\nu}}^y &= E^y.
\end{align*}$$

- $m_{W\pm}$ and $m_t(m_{\tilde{t}})$ randomly generated from 2D p.d.f.s. If no solution is found up to 500 trials, event is discarded. If several solutions are found, a likelihood function is built using parton level information and the solution with the highest likelihood is picked.

$$L_{t\bar{t}h} \propto \frac{1}{p_{T\nu}p_{T\bar{\nu}}} P(p_{T\nu}) P(p_{T\bar{\nu}}) P(p_{Tt}) P(p_{T\tilde{t}}) P(p_{T\bar{t}}) P(m_t, m_{\tilde{t}}) P(m_h),$$
CP observables

- **Two types of observables**, based on previous works. Designed to increase the sensitivity in discriminating CP signals from irreducible backgrounds, but also to probe the CP nature of the Yukawa coupling.

  - angle between the 4-momenta of the Y system, measured in the rest frame of X, with respect to the direction of X in the rest frame of its mother.
  - Two ways to compute: direct or sequential boost.
  - Set of functions \( f(\theta^X_Y)g(\theta'^X_Y) \) were considered, with \( f, g \), simple trigonometric functions.

- **Gunion-He variables** (both in the LAB and \( \vec{t}\vec{h} \) CM system): Phys. Rev. Lett. 76 (1996)

\[
b_2 = (\vec{p}_t \times \hat{k}_Z) \cdot (\vec{p}_{\tilde{t}} \times \hat{k}_Z) / |\vec{p}_t| |\vec{p}_{\tilde{t}}|, \quad b_4 = (p_{\tilde{t}}^z \cdot \vec{p}_{\tilde{t}}^z) / (|\vec{p}_t| \cdot |\vec{p}_{\tilde{t}}|)
\]
Results for $b\bar{b}h$

• **No visible asymmetry for $b\bar{b}h$.** Explicit evaluation of the $f\bar{f}h$ production shows that the asymmetry term is proportional to $m_f^2$. Only significant when $m_f$ is big enough relative to $m_h$. *Phys. Rev. Lett. 76* (1996)

• What if we lower $m_h$ (down to 10 GeV)? Still nothing.

• What about single bottom? Also nothing.
Results for $t\bar{t}h$

- In previous works $m_h = 125$ GeV. Here $m_h = 40$-300 GeV.

- For the tops significant differences are found.
Asymmetries

\[
A_{FB}^{Y} = \frac{\sigma(x_Y > x_Y') - \sigma_Y(x_Y < x_Y')}{\sigma_Y(x_Y > x_Y') + \sigma_Y(x_Y \leq x_Y')},
\]

- Asymmetry vanishes for very large Higgs masses. Exact value depends on the variable.
Expected confidence levels (CL) for $t\bar{t}h$

• Computed from **likelihood ratios** obtained from binned distributions of the observables, as a function of the integrated luminosity. Both **signal and backgrounds** are described by **Poisson distributions**.

• **Only statistical uncertainties** are considered. **Dileptonic channel alone**.

• Shown up to the **High Luminosity LHC (HL-LHC)**, maximum expected is $3000 \, fb^{-1}$, for $m_h = 40, 80, 120, 160$ and $200$ GeV.

• **Four different scenarios:**
  
  1: **CP-even exclusion**. $H_0 = \text{SM}$ and $H_1 = \text{SM+CP-even signal}$.
  
  2: **CP-odd exclusion**. $H_0 = \text{SM}$ and $H_1 = \text{SM+CP-odd signal}$.
  
  3: **CP-odd exclusion (vs CP-even)**. $H_0 = \text{SM +CP-even signal and } H_1 = \text{SM+CP-odd signal}$.
  
  4: **SM exclusion (vs CP-even)**. $H_0 = \text{SM+CP-even signal and } H_1 = \text{SM}$.
• CP-even with $m_h = 40, 80$ GeV can already be excluded at the LHC using the dilepton channel alone.

• Heavier masses harder to exclude, since $\sigma_{t\bar{t}h}$ decreases with $m_h$.

• Scenarios 1 and 4 give similar results.
• CP-odd harder to exclude for lighter masses, but better for heavier masses.

• $\sigma_{t\bar{t}A}$ starts lower than $\sigma_{t\bar{t}H}$. Decreases more smoothly. Around $m_h = 160-180$ GeV, it becomes larger.
• $m_h = 40$ and $m_h = 200$ similar to CP-even and CP-odd exclusion, respectively.

• If CP-even is found, a pure CP-odd can be excluded from a pure CP-even scalar.
Conclusions

• Asymmetry term proportional to $m_f^2$ for $f \bar{f} h$ production.

• Very difficult to determine the nature of the bottom quark Yukawa couplings at the LHC, even for very light Higgses.

• For the tops, significant differences found for several Higgs masses. Those vanish for large scalar masses.

• Expected confidence levels computed in four different scenarios. Lighter CP-even can already be excluded at the LHC using the dilepton channel alone. Heavier CP-even scalars not excluded. CP-odd exclusion requires higher luminosity, but improves for heavier masses. If new CP-even Higgs is found, CP-odd exclusion is possible.
$\sigma_{t\bar{t}h} \text{ vs } m_h$

- Computed in MadGraph5_aMC@NLO at NLO. No decays.
Signal vs Background that go into CL
Confidence levels computation

\[
\ln Q = - (\lambda_{1\text{tot}} - \lambda_{0\text{tot}}) + \sum_{i=1}^{N_{\text{chan}}} n_i \ln \left( \frac{\lambda_{1i}}{\lambda_{0i}} \right).
\]

- \(n_i\) - number of observed events.
- \(\lambda_{1\text{tot}}\) and \(\lambda_{0\text{tot}}\) - total number of predicted events assuming \(H_1\) or \(H_0\).

\[
CL_s = P(Q \geq Q_{\text{obs}}| H_1),
\]

- 1 000 000 toy experiments
- \(n_i\) random, using \(\lambda_{1i}\) or \(\lambda_{0i}\) for mean value.
- \(\ln Q\) in each experiment is computed, for \(H_1\) and \(H_0\).
- \(\ln Q_{\text{obs}} = \text{Median of } dP(Q|H_0)/dQ\)
Reconstruction plots
Reconstruction plots

<table>
<thead>
<tr>
<th>$N_{jets} \geq 4$ &amp; $N_{lep} \geq 2$</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching with TM</td>
<td>5-20</td>
</tr>
<tr>
<td>Reconstruction with TM</td>
<td>18-61</td>
</tr>
<tr>
<td>Reconstruction without TM</td>
<td>66-73</td>
</tr>
<tr>
<td></td>
<td>49-63</td>
</tr>
</tbody>
</table>