Analysis of top quark pair production signal from neutral 2HDM Higgs bosons at LHC

Majid Hashemi a, Mahbobeh Jafarpour b

1 Physics Department, College of Sciences, Shiraz University, Shiraz, 71946-84795, Iran

Abstract In this paper, the top quark pair production events are analyzed as a source of neutral Higgs bosons of two Higgs doublet model type I at LHC. The production mechanism is $pp \rightarrow H/A \rightarrow t\bar{t}$ assuming a fully hadronic final state through $t \rightarrow Wb \rightarrow jjb$. In order to distinguish the signal from the main background which is the standard model $t\bar{t}$, we benefit from the fact that the top quarks in signal events acquire a large Lorentz boost due to the heavy neutral Higgs boson. This feature leads to three collinear jets (a fat jet) which is a discriminating tool for identification of the top quarks from the Higgs boson resonances. Events with two identified top jets are selected and the invariant mass of the top pair is calculated for both signal and background. It is shown that the low $\tan\beta$ region has still some parts which can be covered by this analysis and has not been excluded yet by flavour physics data.

1 Introduction

The standard model (SM) of particle physics has taken a major step forward by observing the Higgs boson at LHC [1, 2] based on a theoretical framework known as the Higgs mechanism [3–8]. The observed particle may belong to a single SU(2) doublet (SM) or a two Higgs doublet model (2HDM) [9–11] whose lightest Higgs boson respects the observed particle properties.

One of the motivations for the two Higgs doublet model is supersymmetry where each particle has a superpartner. The supersymmetry provides an elegant solution to the gauge coupling unification, dark matter candidate and the Higgs boson mass radiative correction by a natural parameters tuning. In such a model two Higgs doublets are required to give mass to the double space of the particles [12–14].

There are four types of 2HDMs with different scenarios of Higgs-fermion couplings. The ratio of vacuum expectation values of the two Higgs doublets (tan $\beta = v_2/v_1$) is a measure of the Higgs-fermion coupling in all 2HDM types [15].

In general, 2HDM involves five physical Higgs bosons due to the extended degrees of freedom added to the model by introducing the second Higgs doublet. The lightest Higgs boson, $h$, is like the SM Higgs boson. The rest are two neutral Higgs bosons, $H, A$ (subjects of this study), and two charged bosons, $H^\pm$. A review of the theory and phenomenology of 2HDM can be found in [16].

In addition to direct searches for the 2HDM Higgs bosons at colliders, there are indirect searches based on flavor Physics data by investigating sources of deviations from SM when processes which involve 2HDM Higgs bosons are introduced [17]. Limits obtained from these type of studies are one of the strongest limits on the mass of the charged and neutral Higgs bosons and $\tan\beta$ and will be referred to when presenting the final results.

The adopted scenario in this analysis is a search for heavy neutral Higgs boson with mass in the range 0.5-1 TeV at LHC operating at $\sqrt{s} = 14$ TeV. All heavy Higgs bosons (CP-even, CP-odd and the charged Higgs) are assumed to be degenerate, i.e., $m_H = m_A = m_{H^\pm}$. The region of interest is low $\tan\beta$ and the final results will be limited to $\tan\beta < 2$. The signal process is $pp \rightarrow H/A \rightarrow tt \rightarrow W^+bW^-\bar{b} \rightarrow jjbjjb$. The fully hadronic final state is expected to result in two fat jets (each consisting of three sub-jets associated with the top quark) which are examined using the updated HEPTopTagger [18,19]. Events which contain two identified (tagged)
top jets are used to fill the top pair invariant mass distribution histogram. The same approach is applied on background events and a final shape discrimination is performed to evaluate the signal significance. Before going to the details of the analysis, a brief review of the theoretical framework is presented in the next section.

2 The Higgs sector of 2HDM

The 2HDM Lagrangian for neutral Higgs-fermion couplings as introduced in [20] takes the form:

\[
\mathcal{L}_Y = \frac{1}{\sqrt{2}} \sum_f \bar{f} \left[ \kappa^f \sigma_{\beta - \alpha} + \rho^f c_{\beta - \alpha} \right] f h \\
+ \frac{1}{\sqrt{2}} \sum_f \bar{f} \left[ \kappa^f c_{\beta - \alpha} - \rho^f s_{\beta - \alpha} \right] f H \\
+ \frac{i}{\sqrt{2}} \bar{f} \gamma_5 \rho^f f A
\]

with \(U(D)\) being the up(down)-type quarks, \(L\) the lepton fields, \(h, H, A\) the neutral Higgs boson fields, \(\kappa^f = \sqrt{2} \frac{m_f}{v}\) for any fermion type \(f\) and \(s_{\beta - \alpha} = \sin(\beta - \alpha)\) and \(c_{\beta - \alpha} = \cos(\beta - \alpha)\). The \(\rho^f\) parameters define the model type and are proportional to \(\kappa^f\) as in Tab. 1 [21]. Therefore the four types of interactions (2HDM types) depend on the values of \(\rho^f\) [22].

In this study, we require \(s_{\beta - \alpha} = 1\) which has two advantages. The first one is that the \(s_{\beta - \alpha}\) factor in the lightest Higgs-gauge coupling is set to unity while the heavier Higgs, \(H\), decouples from gauge bosons [16]. On the other hand, the SM-like Higgs-fermion interactions are \(\tan \beta\) independent.

According to Tab. 1, the type I is interesting for low \(\tan \beta\) as all couplings in the neutral Higgs sector are proportional to \(\cot \beta\). This feature leads to cancellation of this factor as long as Higgs boson branching ratio of decay to leptons and quarks is concerned. The mass of the fermion thus plays an important role in the decay rate and as seen from Figs. 1 and 2 the Higgs boson decay to \(t\bar{t}\) dominates for all relevant Higgs boson masses and \(\tan \beta\) values. The decay to a pair of gluons proceeds through a preferably top quark loop and stands as the second channel. The third channel is \(H/A \to b\bar{b}\) which has been shown to be visible at LHC [23]. The current study focuses on \(H/A \to t\bar{t}\) with branching ratio being near unity and independent of the Higgs boson mass (Fig. 1) and \(\tan \beta\) (Fig. 2).

3 Signal and background cross sections

The signal process under study is a Higgs boson production with the Higgs boson masses in the range 500–1000 GeV. The three Higgs bosons masses are set to be equal for minimizing \(\Delta \rho\) [24]. All selected points are checked to be consistent with the potential stability, perturbativity and unitarity requirements and the current experimental limits on Higgs boson masses using 2HDMC 1.6.3 [25,26].

There has been phenomenological searches for leptonophilic Higgs boson within type IV 2HDM at LHC [27] and linear colliders [28, 29]. These searches are based on leptonic decay of the Higgs boson. On the other hand, the type I 2HDM can be considered as a leptonophobic model where the Higgs boson decay to quarks
plays an important role. At the first glance, decays to all fermions are relevant at low \( \tan \beta \) values. However, the fermion mass in the Higgs-fermion vertex enhances the top quark coupling dramatically compared to other channels. This is due to the fact that the common \( \cot \beta \) factors cancel out when calculating branching ratio of Higgs decays to fermions. Therefore in this analysis, the Higgs boson decay to \( tt \) is considered as the signal.

While the neutral Higgs boson searches at LEP \cite{30,31} leads to \( m_A \geq 93.4 \text{ GeV} \), the LHC results \cite{32,33} indicate that the neutral Higgs boson mass in the range \( m_{H/A} = 200 - 400 \text{ GeV} \) is excluded for \( \tan \beta \geq 5 \). This result is based on minimal supersymmetric standard model (MSSM) which has a different Higgs boson spectrum from 2HDM due to supersymmetry constraints. Since our region of interest is Higgs boson masses above 500 GeV, no constraints from LEP or LHC limits the current analysis and the Higgs boson masses under study.

There are also results from flavor physics data which impose lower limits on the charged Higgs mass in type II and III at 480 GeV \cite{34}. An update to this work is reported in \cite{35} where low \( \tan \beta \) values are excluded to some extent. The idea in such analyses is based on the contribution from additional Feynman diagrams which involve charged Higgs bosons and their effect depends on the type of the 2HDM. The type I and IV behave different from type II and III as far as the charged Higgs coupling to quarks is concerned. In the former, the charged Higgs coupling to all quark types is suppressed at low \( \tan \beta \), while in the latter, coupling with at least one type of the quarks (up type or down type) is enhanced with \( \tan \beta \). Therefore charged Higgs limits from flavor physics in type I and IV are very soft and basically relevant at \( \tan \beta \) values as low as 2. This is the region of search in this analysis. Although we are dealing with neutral Higgs bosons, since the scenario under study is a degenerate scenario based on \( m_H = m_A = m_{H^\pm} \), limits on the charged Higgs are propagated into the final results.

The signal cross sections times branching ratio of Higgs (\( H/A \)) decay to \( tt \) are shown in figs. 3 and 4. The cross section decreases with increasing the Higgs boson mass as well as \( \tan \beta \). Therefore the most suitable area for search is where the mass is as low as possible and \( \tan \beta \) is also very small.

The main SM background processes are \( tt \), gauge boson pair production \( WW, WZ, ZZ \), \( s \)-channel and \( t \)-channel single top, single \( W \) and single \( Z/\gamma^* \). The signal and background cross sections are listed in tab. 2.

**Fig. 3** The signal cross section times BR(\( H/A \to tt \)) at \( \sqrt{s} = 14 \text{ TeV} \) as a function of \( \tan \beta \).

**Fig. 4** The signal cross section times BR(\( H/A \to tt \)) at \( \sqrt{s} = 14 \text{ TeV} \) as a function of \( \tan \beta \).

### 4 Signal selection and analysis

The generation of signal and background events starts with PYTHIA 8 \cite{36} followed by jet reconstruction using FASTJET 2.8 \cite{37,38}.

The jet reconstruction algorithms are classified according to their different subjet distance measures which can be written as \( d_{1,2,j} = \Delta R_{1,2,j}^2 / D2 \times \min(p_{T,j1}, p_{T,j2}) \) with \( n = -1, 0, 1 \) for anti-\( k_T \), Cambridge/Aachen (CA) and \( k_T \) algorithms respectively. The \( k_T \) algorithm first combines the soft and collinear subjets and is suitable for reconstructing the QCD splitting history in top tagging algorithm. The anti-\( k_T \) algorithm, first combines the hardest subjets to obtain a stable jet with clean jet boundary. The CA algorithm always combines the most collinear subjets while not being sensitive to soft splittings and therefore is suitable for top tagging reconstruction. The algorithm adopted by HEPTopTagger is thus CA with a cone size of \( \Delta R = 1.5 \).

The HEPTopTagger is one of recent algorithms introduced for boosted top quark reconstruction \cite{39}. It is based on a CA jet reconstruction with \( \Delta R = 1.5 \) and...
the top jet candidate $p_T$ above 200 GeV. The threshold can be lowered down to 150 GeV without significant loss of efficiency. Having the collection of fat jets in the first step, the top tagging algorithm starts with undoing the last clustering of the top jet candidate $j$ and requiring the mass drop criterion as $m_{j_i} < 0.8m_j$ where $j_i$ is the $i$th subjet from the jet $j$. Subjets with $m_j < 30$ GeV are not considered to end the unclustering iteration.

In the second step a filtering is applied to find a three-subjet combination with a jet mass within $m_{t \pm 25}$ GeV.

In the last step, having sorted jets in $p_T$, several requirements are applied to find the best combination of subjets with two subjets giving the best $W$ boson invariant mass and the whole three subjets to be consistent with the top quark invariant mass. Details of these criteria are expressed in [40].

Performing the algorithm, a selection efficiency for each signal sample is obtained. The same procedure is applied on background samples. An event is required to have two top jets identified. The invariant mass of the two top jets are calculated as the Higgs boson candidate mass. Both signal and background distributions of top quark pair invariant masses are normalized according to the corresponding cross sections. The signal on top of this criterion is then plotted for each benchmark point as seen in Figs. [5, 10].

At this step, since a large number of background is still filling the signal region, a mass window is applied to select the signal and increase the signal to background ratio. The position of the mass window (both left and right sides) is determined in an automatic search based on requiring the maximum signal significance. This is performed in a loop over bins of the histogram and finding the left and right bins inside which the signal significance is maximum.

Table 2 shows mass window position, total efficiencies for signal and background events, final number of signal and background events passed the mass window cut, their ratio and the signal significance as $S/\sqrt{B}$ at two values of $\tan \beta = 0.5$ and 1. The integrated lumi-

| $m(H/A)$ [GeV] | 500 | 600 | 700 | 800 | 900 | 1000 |
|-----------------|-----|-----|-----|-----|-----|-----|
| $\sigma \times BR$ [fb] | 61.1 | 36.0 | 25.4 | 15.0 | 9.1 | 5.6 |

| Background | tt | WW | WZ | ZZ | sts | stt | W | Z |
|------------|----|----|----|----|-----|-----|---|---|
| $\sigma \times BR$ [fb] | 390 | 32.6 | 12.1 | 5.33 | 5.6 | 117 | 1.02×10^5 | 5.7×10^4 |

Figure 11 shows the signal significance as a function of the Higgs boson mass for different $\tan \beta$ values. The dashed horizontal line indicates the 5$\sigma$ significance. Using the analysis results for Higgs boson masses from 500 GeV to 1000 GeV, one can obtain the 95% C.L. exclusion region and the 5$\sigma$ discovery contours. Figure 12 shows the exclusion region at 95% C.L. including the recent result from [35] (the result reported in [35] is based on charged Higgs mass as a function of $\tan \beta$, however, it is included in the current work as a limit for all Higgs bosons since the Higgs boson masses are equal in the scenario adopted in this analysis). The 5$\sigma$ contour is also shown in Fig. 13.

As seen from Figs. 12 and 13, both exclusion and discovery are possible at regions not yet excluded by any experimental or phenomenological analysis. Therefore any sign of extra top pair signals on top of SM background could be regarded as a signal for new physics especially 2HDM. It should be noted that in this analysis, a full set of background processes was studied. However, all background processes led to very small number of events which were negligible compared to the SM $t\bar{t}$. Therefore final plots are based on signal on top of the $t\bar{t}$ distribution without any sizable error.

### 5 Conclusions

Extra sources of $t\bar{t}$ events from what we expect from standard model can appear from theories beyond standard model such as two Higgs doublet models. In 2HDM type I, the heavy neutral (CP-even or odd) Higgs decay to $t\bar{t}$ dominates the other channels. In such a scenario a proton-proton collision may create a neutral Higgs decaying to $t\bar{t}$. The signal from such a process, can be observed as an excess of top pair events over what is expected from SM. The discriminating tool can be a top pair invariant mass distribution filled with events containing two top jets from both signal and background.
processes. The analysis performed in this work, shows that such a signal is observable at integrated luminosity of 300 $fb^{-1}$ for $\tan \beta$ values which depend on the Higgs boson mass. The exclusion at 95% C.L. is also possible at the same integrated luminosity for $\tan \beta < 2$ with $m_{(H/A)} = 600$ GeV as the best point.

Acknowledgements

This work was performed using the computing cluster at Shiraz University, college of sciences. We would like to thank the personnel involved in the operation and maintenance of the cluster.
Table 3  Signal and background analysis results at $\tan\beta = 0.5$ and 1

| Mass window [GeV] | $m_{H/A}$ [GeV] |
|-------------------|-----------------|
| 450 – 635         | 500             |
| 525 – 630         | 600             |
| 610 – 730         | 700             |
| 695 – 825         | 800             |
| 780 – 925         | 900             |
| 865 – 1020        | 1000            |

| $\text{Tot-eff}(S)$ | $\sqrt{B}$ | $\tan\beta = 0.5$ | $\sqrt{B}$ | $\tan\beta = 1$ |
|---------------------|-----------|-------------------|-----------|----------------|
17. F. Mahmoudi, O. Stal, Phys. Rev. D81, 035016 (2010). DOI 10.1103/PhysRevD.81.035016
18. T. Plehn, M. Spannowsky, M. Takeuchi, D. Zerwas, JHEP 10, 078 (2010). DOI 10.1007/JHEP10(2010)078
19. T. Plehn, M. Spannowsky, M. Takeuchi, Phys. Rev. D85, 034029 (2012). DOI 10.1103/PhysRevD.85.034029
20. S. Davidson, H.E. Haber, Phys. Rev. D72, 035004 (2005). DOI 10.1103/PhysRevD.72.035004. [Erratum: Phys. Rev.D72,099902(2005)]
21. V.D. Barger, J.L. Hewett, R.J.N. Phillips, Phys. Rev. D41, 3421 (1990). DOI 10.1103/PhysRevD.41.3421
22. M. Aoki, S. Kanemura, K. Tsumura, K. Yagyu, Phys. Rev. D80, 015017 (2009). DOI 10.1103/PhysRevD.80.015017
23. G.C. Dorsch, S.J. Huber, K. Mimasu, J.M. No, Phys. Rev. Lett. 113(21), 211802 (2014). DOI 10.1103/PhysRevLett.113.211802
24. W. Grimus, L. Lavoura, O.M. Ogreid, P. Osland, J. Phys. G35, 075001 (2008). DOI 10.1088/0954-3899/35/7/075001
25. D. Eriksson, J. Rathmann, O. Stal, Comput. Phys. Commun. 181, 189 (2010). DOI 10.1016/j.cpc.2009.09.011
26. D. Eriksson, J. Rathmann, O. Stal, Comput. Phys. Commun. 181, 833 (2010). DOI 10.1016/j.cpc.2009.12.016
27. S. Su, B. Thomas, Phys. Rev. D79, 095014 (2009). DOI 10.1103/PhysRevD.79.095014
28. M. Hashemi, The European Physical Journal C 77, 302 (2017). DOI 10.1140/epjc/s10052-017-4863-0
29. M. Hashemi, G. Haghighat, Physics Letters B 772, 426 (2017). DOI 10.1016/j.physletb.2017.06.068
30. R. Barate, et al., Phys. Lett. B487, 253 (2000). DOI 10.1016/S0370-2693(00)00822-4
31. M. Acciarri, et al., Phys. Lett. B496, 34 (2000). DOI 10.1016/S0370-2693(00)01272-7
32. CMS Collaboration, CMS-PAS-HIG-16-037
33. The ATLAS Collaboration, ATLAS-CONF-2016-085
34. M. Misiak, et al., Phys. Rev. Lett. 114(22), 221801 (2015). DOI 10.1103/PhysRevLett.114.221801
35. M. Misiak, M. Steinhauser, The European Physical Journal C 77, 201 (2017). DOI 10.1140/epjc/s10052-017-4776-y
36. T. Sjostrand, S. Mrenna, P.Z. Skands, Comput. Phys. Commun. 178, 852 (2008). DOI 10.1016/j.cpc.2008.01.036
37. M. Cacciari, in Deep inelastic scattering. Proceedings, 14th International Workshop, DIS 2006, Tsukuba, Japan, April 29-24, 2006 (2006), pp. 487–490. [,125(2006)]
38. M. Cacciari, G.P. Salam, G. Soyez, Eur. Phys. J. C72, 1896 (2012). DOI 10.1140/epjc/s10052-012-1896-2
39. T. Plehn, M. Spannowsky, J. Phys. G39, 083001 (2012). DOI 10.1088/0954-3899/39/8/083001
40. S. Yang, Q.S. Yan, JHEP 02, 074 (2012). DOI 10.1007/JHEP02(2012)074
41. M. Hashemi, G. Haghighat, JHEP 02, 040 (2016). DOI 10.1007/JHEP02(2016)040