Analysis of b quark pair production signal from neutral 2HDM Higgs bosons at future linear colliders

Majid Hashemi\textsuperscript{a}, Mostafa MahdaviKhorrami\textsuperscript{b}

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

Received: 26 April 2018 / Accepted: 31 May 2018 / Published online: 12 June 2018
© The Author(s) 2018

Abstract In this paper, the b quark pair production events are analyzed as a source of neutral Higgs bosons of the two Higgs doublet model type I at linear colliders. The production mechanism is $e^+e^- \rightarrow Z^{(*)} \rightarrow HA \rightarrow b\bar{b}bb$ assuming a fully hadronic final state. The analysis aim is to identify both CP-even and CP-odd Higgs bosons in different benchmark points accommodating moderate boson masses. Due to pair production of Higgs bosons, the analysis is most suitable for a linear collider operating at $\sqrt{s} = 1$ TeV. Results show that in selected benchmark points, signal peaks are observable in the $b$-jet pair invariant mass distributions at integrated luminosity of $500 \text{ fb}^{-1}$.

1 Introduction

One of the most significant accomplishments of standard model (SM) of particle physics is indubitably, observation of 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 Higgs boson) or a model accommodating a larger structure such as two Higgs doublet model (2HDM) [9–11] whose lightest Higgs boson respects the observed particle properties.

In the latter scenario, one would have a light Higgs boson (h) playing the role of the observed particle, plus additional Higgs bosons with different parities and electric charges. The extra Higgs bosons of the model are assumed to be heavier than the observed one. Therefore, a center of mass energy above the threshold of their masses is required to observe them. Moreover there may be needs for a cleaner collider with a dominant leptonic environment rather than LHC to provide reasonable signature of such particles.

Apart from different scenarios already introduced as beyond standard model (BSM), the 2HDM is considered as a framework for supersymmetry theory, in which each fermion (boson) particle has an associated boson (fermion) particle known as super partner. This theory offers an ingenious solution to the gauge coupling unification at high energies, dark matter candidate (lightest supersymmetric particle) and removal of quadratic divergence of the Higgs boson mass radiative corrections by a natural parameter tuning. In such a theory, the particle space is twice that in SM due to introducing super partners for SM particles and therefore, two Higgs doublets are required to give mass to the double space of particles [12–14]. Here we do not go through such supersymmetric theories, but instead work in the field of 2HDM without supersymmetry.

There are four types of 2HDM 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 the free parameter of the model and is considered as a measure of the Higgs-fermion coupling strength in all 2HDM types [15].

Our focus in this paper is 2HDM type I which allows for heavy quark pair production in Higgs boson decays while decays to light quarks and leptons are suppressed because the relevant Higgs-fermion couplings depend on the fermion mass. Therefore below the top quark pair production threshold, $H/A \rightarrow b\bar{b}$ is dominant as long as $A \rightarrow ZH$ is not kinematically allowed.

In other scenarios, such as type III, the Higgs boson coupling with down-type quarks can experience enhancement proportional to $\tan \beta$. However, the current analysis relies on type I in which the signal cross section and Higgs boson decays only depend on the Higgs boson and fermion masses. In recent studies, other types of 2HDM (types II and IV) were analyzed [16–18] leading to overall conclusion that linear colliders have a prominent potential for observation of 2HDM Higgs bosons with a supreme capability over LHC.

In total five physical Higgs bosons are predicted in 2HDM. The lightest Higgs boson, $h$, (sometimes denoted as $h_{SM}$) is the SM like Higgs boson and there are two heavier neutral
Higgs bosons, \( H \) (CP-even) and \( A \) (CP-odd), and two charged Higgs bosons, \( H^\pm \). Recently the theory and phenomenology of 2HDM has been extensively discussed in [19].

In addition to direct searches for the 2HDM Higgs bosons which look for direct signals of Higgs boson decays, there are indirect searches based on flavor Physics data. In such searches, deviations from SM observables are looked for when processes containing 2HDM Higgs bosons are added to their corresponding diagrams from SM [20]. Limits obtained from these type of studies have to be taken into account. However, the current analysis focuses on points in parameter space where there is no exclusion from flavor physics studies.

The Higgs boson mass range in this analysis is 150–250 GeV to be searched for at a future linear collider, operating at \( \sqrt{s} = 1 \) TeV. Signals of heavier Higgs bosons tend to become small when increasing the Higgs bosons masses. All Higgs bosons are assumed to be degenerate in mass, i.e., \( m_H = m_A = m_{H^0} \). This setting ensures that deviation from SM, in terms of \( \Delta \rho \), is small enough and consistent with experimental value [21].

The region of interest is \( \tan \beta < 50 \). The signal process, i.e., \( e^+e^- \rightarrow Z(\pm) \rightarrow HA \) is independent of \( \tan \beta \) as the \( Z-H-A \) vertex does not depend on Higgs-fermion couplings and the 2HDM type. As will be seen in the next section, the branching ratio of Higgs boson decay to \( b\bar{b} \) is also independent of \( \tan \beta \) for \( \tan \beta < 50 \).

The fully hadronic final state is expected to result in two pairs of \( b\)-jets (totally four \( b\)-jets) coming from neutral Higgs boson \( H/A \) decays to two \( b\)-jets. Events which contain four identified (tagged) \( b\)-jets, are used to produce the \( H/A \) invariant mass distribution. 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 Theoretical framework

Couplings of heavy neutral Higgs bosons (\( H \) and \( A \)) with quarks in Yukawa Lagrangian of the 2HDM, as introduced in [22], takes the form:

\[
\mathcal{L}_Y = \overline{D} \rho^D_{\beta - \alpha} U + \overline{U} \rho^U D + U \rho^U U \tag{1}
\]

in which \( U(D) \) are the up(down)-type quarks fields, \( H \) and \( A \) the neutral Higgs boson fields, \( \kappa^q = \frac{m^q}{v} \), for any up(down)-type quark \( U(D) \) and \( s_{\beta - \alpha} = \sin(\beta - \alpha) \) and \( c_{\beta - \alpha} = \cos(\beta - \alpha) \). The \( \rho^D \) parameters depend on the 2HDM type and are proportional to \( \kappa^q \) as shown in Table 1 [23]. Therefore the four types of interactions (2HDM types) depend on the values of \( \rho^D \) which is \( \kappa^q \) (SM coupling) times a \( \tan \beta \) or \( \cot \beta \) factor which makes possible deviations from SM [24].

Table 1 Different types of 2HDM in terms of the Higgs boson couplings with \( U \) (up-type quarks), \( D \) (down-type quarks) and \( L \) (leptons)

| 2HDM Type | I | II | III | IV |
|-----------|---|----|-----|----|
| \( \rho^D \) | \( \kappa^D \cot \beta \) | \(-\kappa^D \tan \beta \) | \(-\kappa^D \tan \beta \) | \( \kappa^D \cot \beta \) |
| \( \rho^U \) | \( \kappa^U \cot \beta \) | \( \kappa^U \cot \beta \) | \( \kappa^U \cot \beta \) | \( \kappa^U \cot \beta \) |
| \( \rho^L \) | \( \kappa^L \cot \beta \) | \(-\kappa^L \tan \beta \) | \( \kappa^L \cot \beta \) | \(-\kappa^L \tan \beta \) |

In Yukawa Lagrangian of the 2HDM type I, the light neutral Higgs (\( h \)) coupling to fermions and gauge-bosons takes the SM value by setting \( s_{\beta - \alpha} = 1 \). This setting suppresses the heavy neutral Higgs (\( H \)) coupling with gauge bosons which is proportional to \( c_{\beta - \alpha} \) [19]. Therefore in our study, we set \( s_{\beta - \alpha} = 1 \), in respect of the correspondence principle so that 2HDM SM-like Higgs boson behaves the same as SM Higgs. This leads to the brief form of the Lagrangian as shown in Eq. 1:

\[
\mathcal{L}_Y = \overline{D} \rho^D D + \overline{U} \rho^U U \tag{2}
\]

According to Table 1, the type I appears interesting for low \( \tan \beta \) as all couplings in the neutral Higgs sector are proportional to \( \cot \beta \). However, the \( \cot \beta \) factor cancels out when calculating branching ratio of Higgs boson decays because all Higgs-fermion couplings are proportional to the same factor. This has two subcases: firstly the Higgs boson decay to fermions is independent of \( \tan \beta \) at tree level, and secondly the Higgs boson decays to heavy quarks become dominant due to their larger coupling with the Higgs boson. As a result, as long as the Higgs boson mass is below the top quark pair production threshold, decay to \( b\bar{b} \) is dominant while above the threshold (Higgs boson mass above twice the top quark mass) decay to \( t\bar{t} \) starts to grow as seen from Fig. 1a, b.

It should be noted that loop diagrams such as \( H \rightarrow gg \) proceed through preferably virtual top quarks in the loop, as seen in Fig. 2. Such decays grow when the Higgs boson mass increases and result in the reduction of decay to \( b \) or \( c \) quark pairs. The above conclusion is of course valid as long as Higgs boson decay to \( t\bar{t} \) is kinematically impossible, i.e., \( m_{H/A} \lesssim 350 \) GeV.

Although the charged Higgs bosons acquire a strong limit from flavor physics in 2HDM types II and III as reported in [25–27], types I and IV receive a small excluded region at low \( \tan \beta \). Since the scenario assumed in this paper assumes degenerate Higgs boson states, one may assume the same excluded region for neutral Higgs bosons. However, these limits are at low \( \tan \beta \) values and do not affect results of this analysis.
3 Signal identification and the search scenario

The signal process, i.e., $e^+e^- \rightarrow Z^{(*)} \rightarrow HA$ followed by the Higgs bosons decays ($H/A \rightarrow b\bar{b}$) depends on the Higgs bosons masses and the cross section decreases when the sum of the masses of the two Higgs bosons reaches the threshold of center of mass energy provided by the collider. As it is shown in Fig. 1a, b, the proper range of the Higgs boson mass is basically 150–350 GeV. However at masses above 250 GeV, the cross section and decay rates decrease so that a reasonable signal is not observed. The Higgs boson masses are assumed to be equal while a scenario with $m_A = m_H \pm 50$ GeV is also studied.

The total cross section of the signal is $\sim 12 \text{ fb}$ corresponding to $m_{H/A} = 150$ GeV and decreases to $\sim 9 \text{ fb}$ with $m_H = 200$ GeV and $m_A = 250$ GeV. Taking into account the branching ratio of Higgs bosons decays, these values decrease down to $\sim 5 \text{ fb}$ and $\sim 1 \text{ fb}$ respectively.

All benchmark 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 1.7.0 [28,29]. A discussion on this subject will be presented before conclusion.

In order to identify the right pairs of $b$-jets, two methods are tried. The first method is based on the special relativity kinematics; In this method, if the Higgs bosons masses are equal ($m_H = m_A$) which is valid in the scenario presented in this analysis, one can distinguish the two pairs of the $b$-jets.
As the example, according to Fig. 3b, almost equal thus the conclusion holds for particle energies.

Due to the energy conservation, \( E_1 + E_4 > E_2 + E_3 \) and therefore the \( b \)-jet with the maximum energy has to come along with the \( b \)-jet with the lowest energy to keep the sum of the two energies conserved. Without any loss of generality, one may assume \( E_1 \) to be the highest energy. Then the four \( b \)-jets energies are sorted like \( E_1 > E_2 > E_3 > E_4 \). In order to apply this algorithm, the tagged \( b \)-jets are sorted in terms of their energies and labeled as \( b_1, b_2, b_3 \) and \( b_4 \). Then two independent pairs, i.e., \((b_1, b_4)\) and \((b_2, b_3)\) are used to calculate the invariant masses of the parent Higgs bosons.

The second method is based on minimizing the angular separation between the two \( b \)-jets, \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \), in which \( \eta = -\ln(\tan(\theta/2)) \) is the pseudo-rapidity and \( \phi \) is the azimuthal angle. This algorithm finds the correct \( b \)-jet pair with the minimum angular separation. One of the advantages of this method is that it is independent of whether the Higgs bosons masses are equal or not.

The two methods discussed above were tested separately, however, the second method leads to better results (narrower invariant mass peak, smaller tails and better signal to background ratio due to the less wrong pairing). Therefore results presented in this analysis are based on the second method.

### 4 Software setup and cross sections

In this work, the event generation and cross section calculation of signal processes are handled by PYTHIA 8.2.15 [30] which uses 2HDM spectrum files in LHA format [31] generated using 2HDMC 1.7.0 for each benchmark point separately. The LHA files contain information about the parameters of the model which include Higgs bosons masses, widths and the branching ratio of decay to different channels for each benchmark point. After multi-particle interaction and shower production by PYTHIA, the jet reconstruction is performed by FASTJET 3.1 [32,33] using anti-\( k_t \) algorithm with a jet cone size of 0.4.

The main SM background processes are \( t \bar{t} \), gauge boson pair production \( WW, ZZ \) and single \( Z/\gamma^* \). These background processes are all generated using PYTHIA and their cross sections are also obtained using the same package besides an additional NLO scaling which is applied on the \( t \bar{t} \) cross section according to MCFM 6.1 [34–37]. Tables 2 and 3 indicate the signal benchmark points and background processes and their cross sections.
5 Signal selection and analysis

The event selection starts from $b$-tagging using reconstructed jets from FASTJET. A kinematic requirement is applied on jets as $E_j > 5$ GeV, $|\eta| < 5$ to reject soft jets or jets close to the beam pipe. The $b$-tagging algorithm is based on a simplified matching between the reconstructed jet and the $b$-quark in the event. For each reconstructed jet, a search for $b$-quarks in the jet cone is performed. If the jet accommodates a $b$ or $c$ quark in the reconstruction cone, it is tagged as a $b$-jet with a $b$-tagging probability of 70% and fake rate of 10% (from c-jets).

Each event has to have four $b$-jets to be analyzed. A search among the four $b$-jets in the event is performed to find the correct $b$-jet pair with minimum $\Delta R$ and calculates their invariant mass. The same algorithm is applied on background processes. The distributions of $b$-jet pair invariant masses are stored in histograms for both signal and background processes.

In order to select events in the signal region (region of the signal peak in the $b$-jet pair invariant mass distribution), a mass window is applied on the signal plus background distributions. The mass window enhances the signal to background ratio leading to a higher signal purity. Figure 4a–c show the distributions of $b\bar{b}$ invariant mass from signal events on top the background.

A polynomial fit describes the background distribution and is used as the input probability distribution function (PDF) for the background (the red curves shown in Fig. 4a–c). A Gaussian fit is then added to the background PDF (the poly-
nominal function) and is applied on signal plus background distribution using the input parameter values taken from the previous step. The result of the fits are shown as green curves in Fig. 4a–c. The error bars are taken into account in the fit which is based on \( \chi^2 \) minimization. They are however of statistical origin and systematic uncertainties are not taken into account.

Table 4 shows mass windows and the total signal and background efficiencies and the final number of signal and background. The signal to background ratio and signal statistical significance in different points are also included. The mass window is set by maximizing the signal significance in a search for the best position of the left and right sides of the window. As seen from Table 4, the signal significance is reasonable and above 5\( \sigma \) for all selected benchmark points.

It should be noted that we obtain reasonable results only for BP1, BP2 and BP3. The fourth benchmark (BP4) which involves different Higgs boson masses leads to smaller branching ratio of Higgs boson decays compared to other points. The reason is mainly due to the CP-odd Higgs decay branching ratio of Higgs boson decays compared to other which involves different Higgs boson masses. The theoretical framework was set to 2HDM type I where the Higgs (H or A) boson decay to \( b \bar{b} \) is dominant as long as the Higgs boson mass is below the threshold of on-shell decay to \( t \bar{t} \). The leptonic decays are also suppressed as cot\( \beta \). The signal cross section is almost independent of tan\( \beta \) and therefore all results were quoted based on a typical tan\( \beta \) value of 10. The collider integrated luminosity was set to 500 fb\(^{-1}\). Results show that the Higgs boson signal in \( b \bar{b} \) invariant mass distribution is well observable for all benchmark points. There is also a fit possibility for those points with equal Higgs boson masses.

### 6 Discussion

It should be noted that the theoretical space available to this study is limited in \( m_{T2}^2 \) vs tan\( \beta \) parameter space. In order to respect theoretical requirements, a scan over possible \( m_{T2}^2 \) values was performed for each tan\( \beta \) value resulting in a narrow range of \( m_{T2}^2 \) which satisfies all requirements of potential stability, perturbativity and unitarity. Results are plotted in Fig. 5. As seen from Fig. 5, the available range of \( m_{T2}^2 \) becomes smaller with increasing tan\( \beta \) for each benchmark point. Any value of \( m_{T2}^2 \) in the allowed range leads to the same branching ratio of Higgs boson decay to \( b \bar{b} \) for tan\( \beta < 50 \). Therefore the signal significance values obtained for tan\( \beta = 10 \) are valid up to tan\( \beta \approx 50 \) provided that for each value of tan\( \beta \), the allowed value of \( m_{T2}^2 \) is taken from the range provided in Fig. 5.

### 7 Conclusions

The observability of 2HDM Higgs bosons was studied at an \( e^+e^- \) linear collider operating at \( \sqrt{s} = 1 \) TeV. Different benchmark points were studied focusing on moderate values of the Higgs bosons masses. The theoretical framework was set to 2HDM type I where the Higgs (H or A) boson decay to \( b \bar{b} \) is dominant as long as the Higgs boson mass is below the threshold of on-shell decay to \( t \bar{t} \). The leptonic decays are also suppressed as cot\( \beta \). The signal cross section is almost independent of tan\( \beta \) and therefore all results were quoted based on a typical tan\( \beta \) value of 10. The collider integrated luminosity was set to 500 fb\(^{-1}\). Results show that the Higgs boson signal in \( b \bar{b} \) invariant mass distribution is well observable for all benchmark points. There is also a fit possibility for those points with equal Higgs boson masses.

Acknowledgements We would like to thank the college of sciences at Shiraz university for providing computational facilities and maintaining the computing cluster during the research program.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Funded by SCOAP³.
