Kinematic corrections and reconstruction methods for neutral Higgs boson decay to $b\bar{b}$ in 2HDM type I at future lepton colliders

Majid Hashemi[*] and Elnaz Ebrahimi[†]

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

In this paper, an approach for neutral Higgs boson searches is described based on 2HDM type I at electron-positron linear colliders operating at $\sqrt{s} = 1$ TeV. The beam is assumed to be unpolarized and fast detector simulation is included. The signal process produces a fully hadronic final state through $e^+e^- \to AH \to b\bar{b}b\bar{b}$ where both CP-even and CP-odd Higgs bosons (H and A) are assumed to decay to a pair of b-jets. Several benchmark scenarios are introduced as the baseline for the analysis taking $m_{H/A}$ in the range 150–300 GeV. In order to avoid Higgs boson conversion $A \to ZH$, Higgs boson masses are chosen with $m_A - m_H < m_Z$. It is shown that with a proper kinematic correction applied on final state b-jet four momenta, true combinations of b-jets can be found for simultaneous reconstruction of both Higgs bosons through $b\bar{b}$ invariant mass calculation. Results show that observable signals can be achieved with statistical significance exceeding 5σ well before the target integrated luminosity of 8 ab^{-1}.

I. INTRODUCTION

The Standard Model of particle physics (SM) is one of the most precisely tested theories which has been verified in various experiments. After the discovery of the electroweak gauge bosons, there has been extensive search for the missing key element of the standard model, i.e., the Higgs boson, h_{SM} [1–6]. The result of these searches is the observation of a new boson at the Large Hadron Collider (LHC) by the two collaborations ATLAS and CMS [7,8].

The properties of the observed boson show reasonable compatibility with SM predictions as verified at LHC [9–16]. However, these measurements still allow possible extensions to the Higgs sector such as the two Higgs doublet model [17–19] which is the basis for several beyond SM scenarios such as supersymmetry [20–22].

The two Higgs doublet model (2HDM) has attracted attention even as a stand-alone model without necessarily embedding it in a supersymmetric theory. Since there are two Higgs boson doublets with complex fields, a total number of five Higgs bosons are predicted including the lightest Higgs boson h (the SM-like Higgs boson), the heavy neutral CP-even(CP-odd) Higgs bosons $H(A)$ and the two charged Higgs bosons $H^{\pm}$.

After the discovery of the new boson, there has been extensive search for the extra Higgs bosons at LHC. The ATLAS collaboration has reported an analysis of $pp \to A + X \to ZH + X$ followed by $Z \to \ell\ell$ and $H \to b\bar{b}$ at $\sqrt{s} = 13$ TeV. The pseudoscalar Higgs decay is kinematically allowed if $m_A - m_H \geq m_Z$. There are two subsequent reports for this analysis based on integrated luminosity of 139 fb^{-1} [23] and 36.1 fb^{-1} [24]. Four types of 2HDM based on Higgs-fermion couplings have been analyzed and exclusion contours have been presented in parameter space of m_A vs m_H for various model parameter values. As mentioned before, the main characteristics of such analyses is the limitation in the Higgs boson mass parameter space which leaves the region of $m_A - m_H < m_Z$ untouched.

The CMS collaboration has also reported similar analysis of $pp \to A(H) \to ZH(A)$ with $Z \to \ell\ell$ and $H(A) \to b\bar{b}$ using integrated luminosity of 35.9 fb^{-1} [25]. There has been also analysis of heavy pseudoscalar A boson decaying to Z and SM-like h boson with $m_h = 125$ GeV followed by $h \to b\bar{b}$ and $Z \to \ell^+\ell^-$ [26].

The overall conclusion from these searches is that the upper left region of the parameter space in m_A vs m_H plane defined by $m_A - m_H > m_Z$ is excluded up to $m_H \simeq 300$ GeV. The area of the excluded region, however, depends on the type of 2HDM.

Before going to the analysis details, the theoretical framework of the analysis is described and the working points in the parameter space are introduced. These points represent example analysis scenarios which target the unexplored region of the parameter space by LHC.

[*]majid.hashemi@cern.ch
[†]ebrahimielnaz69@gmail.com

II. TWO HIGGS DOUBLET MODEL

The two $SU(2)$ doublets introduced in 2HDM contain complex fields resulting in eight degrees of freedom, three of which are eaten by the electroweak gauge bosons to receive their masses. Therefore five degrees of freedom remain leading to five physical Higgs bosons which are denoted as h, H, A, and $H^{\pm}$. The neutral Higgs masses are assumed to be in the same order as listed above, i.e., h is the lightest Higgs boson which is considered to be SM-like and the other two neutral Higgs bosons are heavier.

Each doublet has its own vacuum expectation value or "vev" (v_1 and v_2). They are related to the SM-like $v = 246$ GeV through $v_1 = v \cos \beta$ and $v_2 = v \sin \beta$ resulting in the ratio $\tan \beta = v_2/v_1$ which is the free parameter of the model.

In addition to β parameter, there is also rotation angle, denoted by α, which diagonalizes the mass-squared matrix of the neutral Higgs bosons. The two free parameters α and β determine the Higgs-fermion couplings in various 2HDM types.

The neutral scalar Higgs couplings with gauge bosons also depend on these parameters through $\sin(\beta - \alpha)$ (h-gauge) or $\cos(\beta - \alpha)$ (H-gauge) and therefore the lightest Higgs boson of the model acquires the same couplings with gauge bosons as those of h_{SM} if $\sin(\beta - \alpha) = 1$ which is the alignment limit [27]. With this requirement, H/A-fermion couplings will solely depend on $\tan \beta$ or $\cot \beta$ and h-fermion couplings coincide their SM values which are m_f/v with m_f being the fermion mass [28–31]. It should be noted that the alignment limit is naturally realized in the decoupling regime where the other Higgs bosons are decoupled by assuming that they are much heavier than the electroweak scale v [32]. However, the Higgs boson masses under study in this work, are of $\mathcal{O}(v)$. Therefore the chosen scenario is the alignment limit without decoupling.

The Yukawa Lagrangian for Higgs-fermion interactions can be written in this form:

$$\mathcal{L} = v^{-1}(m_d d\bar{d} + m_u u\bar{u} + m_\ell \ell\bar{\ell})h$$
$$+ v^{-1}(\rho^d m_d d\bar{d} + \rho^u m_u u\bar{u} + \rho^\ell m_\ell \ell\bar{\ell})H$$
$$+ iv^{-1}(-\rho^d m_d \bar{d}\gamma_5 d + \rho^u m_u \bar{u}\gamma_5 u - \rho^\ell m_\ell \bar{\ell}\gamma_5 \ell)A \quad (1)$$

According to Table I, the heavy Higgs couplings acquire additional type dependent factor ρ^f which can be used to distinguish the model type as well as the Higgs boson decay properties [33,34].

In Table I, $u(d)$ and ℓ denote the up(down)-type quarks and leptons. Type III is also called "*Flipped*" and type IV is called "*lepton-specific*".

We have recently performed various studies of different 2HDM types at future lepton colliders. The main focus has been on Higgs boson pair production, through $e^+e^- \rightarrow HA$.

In type I, $H \rightarrow b\bar{b}$ has been shown to be the most promising decay channel with $A \rightarrow b\bar{b}$ [35] or $A \rightarrow ZH$

TABLE I. The Higgs-fermion couplings in different 2HDM types.

	2HDM Types			
	I	II	III	IV
ρ^d	$\cot \beta$	$-\tan \beta$	$-\tan \beta$	$\cot \beta$
ρ^u	$\cot \beta$	$\cot \beta$	$\cot \beta$	$\cot \beta$
ρ^ℓ	$\cot \beta$	$-\tan \beta$	$\cot \beta$	$-\tan \beta$

with possible leptonic or hadronic decay of the Z boson [36–38]. The four b-jet final state through $H/A \rightarrow b\bar{b}$ has also shown discovery potential in the flipped type or type III [39]. In the lepton-specific type or type IV, the leptonic decay channels, i.e., $H/A \rightarrow \tau\tau$ or $\mu\mu$, result in observable signals in parts of the parameter space as reported in [40–42].

In this work, the same Higgs boson pair production in the four b-jet final state is considered as the signal. If $m_{H/A} < 2m_t$, the Higgs boson decay to top quark pair is kinematically forbidden. Given all Higgs-fermion couplings proportional to $\cot \beta$, the branching ratio of Higgs decay to fermions becomes independent of $\cot \beta$ as the common factors from partial decay rates and the total width cancel out. The remaining key parameter is thus the fermion mass resulting in dominant Higgs decay to b-jet pair.

It will be shown that using kinematic correction applied on final state four-momenta, a dramatic improvement of the results is obtained compared to those reported in [35]. Moreover two possible approaches for the simultaneous reconstruction of the Higgs bosons are introduced and compared.

III. SIGNAL AND BACKGROUND PROCESSES

The signal process is assumed to be the Higgs boson pair production producing four b-jet final state in the framework of type I 2HDM, i.e., $e^+e^- \rightarrow AH \rightarrow b\bar{b}b\bar{b}$. Since Higgs-fermion couplings are proportional to the fermion mass and the common $\cot \beta$ factor cancels out in branching ratio calculations at tree level, both Higgs bosons predominantly decay to $b\bar{b}$ which is the heaviest accessible fermion pair, provided that the Higgs boson mass is below the top quark pair production threshold.

The linear collider is assumed to be e^+e^- collider operating at center-of-mass energy of $\sqrt{s} = 1$ TeV which is realized at the upgrade phase of ILC with target integrated luminosity of 8 ab^{-1} [43,44].

For illustrative purposes, several benchmark scenarios are introduced and the analysis focuses on the selected points in the Higgs boson mass parameter space as shown in Fig. 1. The masses of the CP-even (H) and CP-odd (A) Higgs bosons are chosen to be in the region between the two dashed lines shown in Fig. 1.
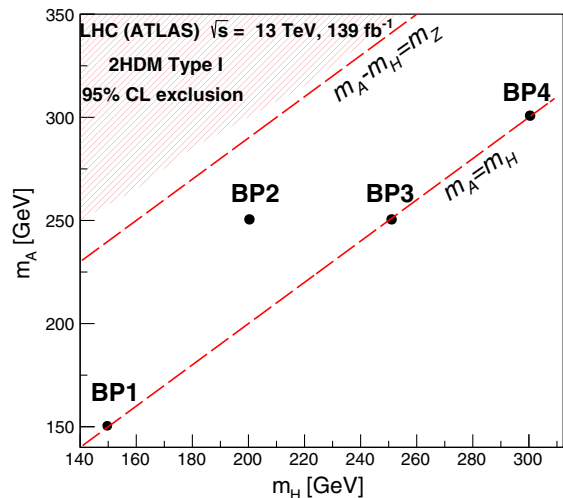
FIG. 1.   The selected benchmark points in the parameter space of m_A vs m_H. The current excluded region of LHC has also been shown based on [23].

The analysis strategy is different from what is adopted by LHC experiments (ATLAS and CMS). They take $pp \rightarrow A + X$ as the signal followed by $A \rightarrow ZH$ decay which requires $m_A - m_H \geq m_Z$ [23]. This requirement limits their explorable region in m_A vs m_H plane which is shown with the upper dashed line in Fig. 1. In the current analysis, the Higgs boson fermionic decay, i.e., $H/A \rightarrow b\bar{b}$ is adopted for analysis. Therefore it is possible to explore regions in parameter space which are inaccessible by the current LHC analyses, i.e., those with $m_A - m_H < m_Z$. These points are well outside the excluded region of type I 2HDM reported by LHC [23].

Table II shows parameter values for the four benchmark points BP1–BP4. The $\tan\beta$ parameter is set to 10 and $\sin(\beta - \alpha)$ is required to be 1 for all scenarios. The Higgs potential mass parameter m_{12}^2 is determined by searching for values which satisfy the theoretical requirements of potential stability [45], unitarity [46–48] and perturbativity. Therefore, for each benchmark point, a range of allowed m_{12}^2 values is obtained as shown in Table II.

TABLE II.   The selected benchmark points of the analysis and parameter values.

Higgs bosons mass	BP1	BP2	BP3	BP4
m_h			125	
m_H	150	200	250	300
m_A	150	250	250	300
$m_{H^\pm}$	150	250	250	300
m_{12}^2	1987–2243	3720–3975	5948–6203	8671–8926
$\tan\beta$			10	
$\sin(\beta - \alpha)$			1	

All calculations related to the parameter values and theoretical requirements as well as the Higgs boson branching ratio of decays are performed with the help of 2HDMC 1.8.0 [49–51]. Agreement with experimental results is confirmed by embedding HiggsBounds 5.9.0 [52–56] and HiggsSignal 2.6.0 [57–59] in 2HDMC 1.8.0 where the selected benchmark points are checked to be not in the excluded regions reported by LHC and TeVatron experiments.

Since, the signal final state consists of four b-jets, any SM process with the same final state should be regarded as the background. Moreover, detector effects, b-tagging fake rate and the final state radiation can also be a source of background. Therefore, Drell-Yan Z/γ^* (single neutral gauge boson), ZZ (gauge boson pair) and $t\bar{t}$ (top quark pair) are the main background production processes. The WW background was also analyzed and passed through the full chain of the analysis. This background is a source of light jets and was very well suppressed by the reasonable performance of the b-tagging requirement.

The so-called overlay hadronic background from photon interactions, i.e., $\gamma\gamma \rightarrow$ hadrons is not simulated in the analysis. However we follow the same approach as adopted by the CLIC collaboration reported in [60] by adding the jet momentum smearing to account for the effect of the hadronic overlay on the jet reconstruction. The jet smearing at 1.5 TeV collisions proposed in [60] assumes 1% and 5% relative smearing applied to the jet momentum with $|\eta| < 0.76$ and $|\eta| \geq 0.76$ respectively. We perform a rough tuning of the above values to 0.7% and 3% in the corresponding pseudorapidity bins for 1 TeV collisions.

IV. ANALYSIS STRATEGY FOR HIGGS BOSON RECONSTRUCTION

In this section, the software setup for event generation, cross section calculation, detector response simulation and analysis is presented in detail including package versions which are all the current latest versions.

The b-tagging algorithm and kinematic corrections applied on b-jets based on full four momentum conservation are described in the next sub-sections. The analysis details are then presented where two approaches for finding the true combinations of final state objects are described and the best approach is adopted.

A. Analysis software setup

The signal and background generation is performed with the use of WHIZARD 3.0.0-β [61,62] including the beam spectrum and initial state radiation (ISR). The beam spectrum file is taken from the official package repository [63]. The generated event files are stored in LHEF format [64] and passed to PYTHIA 8.3.03 [65] for the multi-particle interaction, showering and final state radiation (FSR). The PYTHIA 8.3.03 output is used by DELPHES 3.4.2 [66] for

detector response simulation using `ILCgen` detector card proposed for ILC.

For the detector coordinate system we use the azimuthal angle ϕ and pseudorapidity defined as $\eta = -\ln\tan(\theta/2)$ where θ is the polar angle with respect to the beam axis.

The detector acceptance implemented in `ILCgen` card allows charged track reconstruction with p_T, η dependent momentum smearing and tracking efficiencies up to $|\eta| < 3$.

The jet reconstruction is performed using `FASTJET` 3.3.4 [67,68] with anti-k_t algorithm [69] and the jet cone size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.5$.

The event reconstruction in `DELPHES` is performed using the particle flow approach [70] which is based on the particle composition of the jets [71]. The charged tracks mainly from the charged hadrons are propagated to the calorimeter segments to make the energy flow [72] which, together with the contribution from neutral hadrons and photons, provide the input for the jet reconstruction by `FASTJET`. Therefore the jet energy resolution is mainly due to the tracking resolution and calorimeter segmentation.

The `ILCgen` card contains detailed b-tagging scenarios with the b-tagging efficiencies implemented as a function of the jet energy and pseudorapidity with average b-jet identification efficiencies of 80%, 70%, and 50%. The fake rate has also been included for all b-tagging scenarios.

The detector simulation output is stored in `ROOT` files which are created using `ROOT` 6.22/08 [73] and serve as the datasets for the final numerical and graphical analysis.

B. Cross sections

The signal and background cross sections are obtained using both `PYTHIA` and `WHIZARD`. The signal process is defined in `WHIZARD` using built-in models THDM or MSSM (both models give identical cross sections) and the Higgs boson masses are set in the `SINDARIN` command files. Using `PYTHIA` for signal production requires mass spectrum files in `LHA` format [74] which are generated using `2HDMC`. Tables III and IV show the signal and background cross sections respectively with values from the two generators. Final results are normalized to `WHIZARD` cross sections. The $Z^{(*)}/\gamma^*$ has been generated

TABLE III. Signal cross sections assuming different benchmark points. The letters "W" and "P" denote the results from `WHIZARD` and `PYTHIA`.

Signal process				
Benchmark point	BP1	BP2	BP3	BP4
W: Total $\sigma[fb]$	12.3	9.1	8.0	5.9
P: Total $\sigma[fb]$	12.1	9.4	8.5	6.2
BR($H \to b\bar{b}$)	0.71	0.62	0.51	0.38
BR($A \to b\bar{b}$)	0.54	0.13	0.29	0.16
$\sigma \times BR[fb]$	4.7	0.73	1.18	0.36

TABLE IV. Background cross sections from `WHIZARD` (denoted by "W") and `PYTHIA` ("P").

Background processes			
	ZZ	Z/γ^*	$t\bar{t}$
W: $\sigma[fb]$	181	3473	197
P: $\sigma[fb]$	176	3015	211

in the fully hadronic final state and is slightly higher than the corresponding result from `PYTHIA`. The ZZ background includes only Z boson pair production. The `WHIZARD` cross sections include beam spectrum and ISR.

C. b-tagging

Since there are four b-jets in the signal final state, every event is required to contain exactly four b-jets. Therefore, the event selection starts with selecting events with exactly four jets requiring all of them to pass the b-tagging.

The kinematic requirement for the jet selection is $E_T > 10$ GeV (soft jet veto) and $|\eta| < 2$ (central jet selection).

Figure 2 shows the jet multiplicity in signal and background events and provides the reason for excluding events with more than four jets which are dominated by $t\bar{t}$. Figure 3 shows the b-jet multiplicities in three b-tagging scenarios. The contribution of $t\bar{t}$ background is suppressed very well by choosing the third b-tagging scenario (with average efficiency of 50%).

D. Kinematic correction

When an event containing four b-jets is selected, a kinematic correction is applied on the b-jets four momenta according to what is expected from momentum and energy conservation.

The energy conservation relies on the fact that at lepton colliders, the beam energy is known within the uncertainty arisen from the ISR and beamstrahlung. At hadron colliders, the situation is more complicated due to the fact that
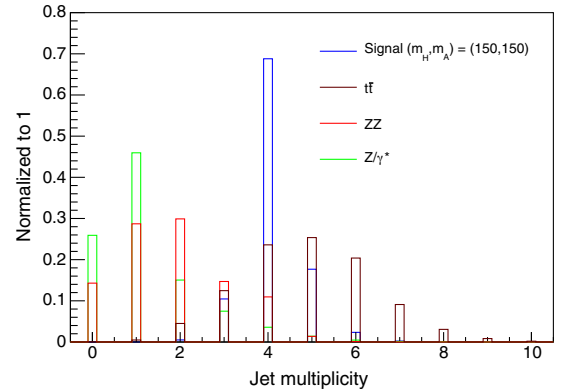
FIG. 2. The reconstructed jet multiplicity in signal and background events. Only BP1 has been shown representing the signal.
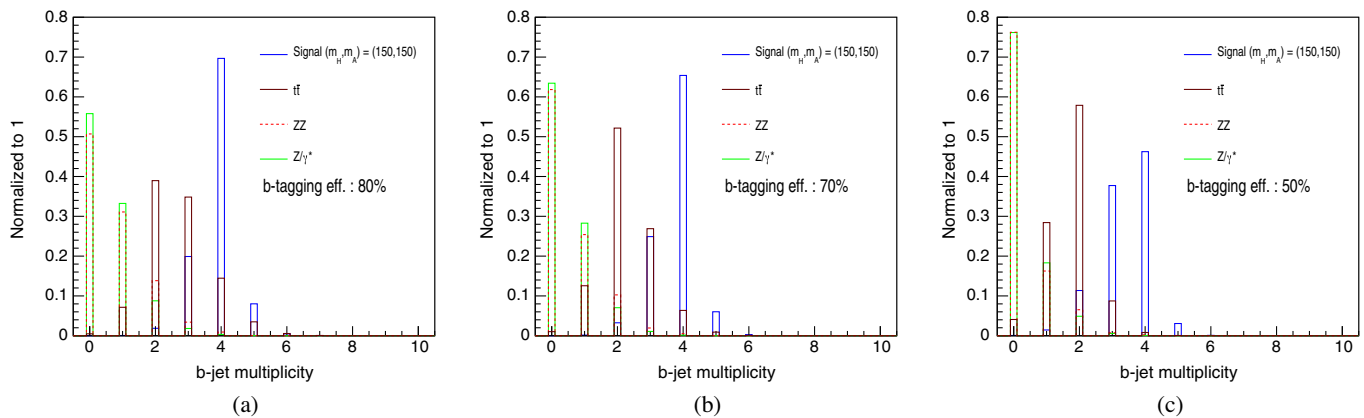
FIG. 3. The b-jet multiplicity in signal and background events with different b-tagging efficiencies.

the effective center of mass energy varies event by event due to the parton distribution functions. The beam spectrum and ISR certainly affect the correction procedure in this analysis, however, as can be seen from the final results, a reasonable performance is observed even including such effects.

It should be noted that the beam crossing angle can also affect the kinematic correction performance which implies that there is no total momentum component in any direction. The effect of beam crossing angle can easily be activated in WHIZARD. However, since it is not yet implemented in DELPHES, we did not apply it for the current analysis.

The set of four equations representing four-momentum conservation includes four correction factors which are named c_i with $i \in 1$–4 assigned to the four b-jets in the event:

$$c_1 p_x^{b_1} + c_2 p_x^{b_2} + c_3 p_x^{b_3} + c_4 p_x^{b_4} = 0$$
$$c_1 p_y^{b_1} + c_2 p_y^{b_2} + c_3 p_y^{b_3} + c_4 p_y^{b_4} = 0$$
$$c_1 p_z^{b_1} + c_2 p_z^{b_2} + c_3 p_z^{b_3} + c_4 p_z^{b_4} = 0$$
$$c_1 E^{b_1} + c_2 E^{b_2} + c_3 E^{b_3} + c_4 E^{b_4} = \sqrt{s}. \quad (2)$$

Therefore all four momentum components of each b-jet receive the same correction factor without changing the b-jet flight direction. In order to avoid negative energy values, all correction factors are required to be positive.

The set of Eq. (2) consists of four linear equations with four unknowns. The momentum components of the four b-jets make the 4×4 coefficient matrix which has a nonzero determinant due to the random nature of the momentum components in events. Therefore, a unique nontrivial solution for the four coefficient factors is expected.

The energy correction may change the order of the b-jets in the list as they are sorted according to descending energies by default. Therefore, the energy sorting algorithm is applied again after the correction and the four b-jets are selected for the rest of the analysis.

Tables V and VI present the preselection efficiencies of the jet reconstruction and four jet selection, b-tagging (having four b-tagged jets in the event) and the positive correction factor requirement in signal and background processes.

According to Tab. V, the correction efficiencies increase with increasing the Higgs boson masses. In other words, the correction performance is better for harder jets from heavier Higgs bosons and reaches 86% for BP4.

The single Z/γ^* production is expected to have a low four jet selection efficiency. However, as seen from Tab. VI, The ZZ background is also suppressed very well by the four jet selection which is due to the performance of the kinematic cuts applied on the jets. The main kinematic difference between the signal and ZZ events is due to the pseudorapidity distributions which are shown in Fig. 4.

As seen from Fig. 4, the jets from ZZ events tend to proceed through the forward/backward region resulting in central jet selection efficiency ($|\eta| < 2$) of about 66%. Therefore, all four jets appear in $|\eta| < 2$ region with probability of $\sim 18\%$. The cut on the jet transverse energy

TABLE V. Preselection efficiencies of the signal events.

	$BP1$	$BP2$	$BP3$	$BP4$
Four jet eff.	0.69	0.47	0.41	0.34
Four b-jet eff.	0.54	0.53	0.52	0.51
Correction eff.	0.71	0.80	0.82	0.86
Total pre-sel. eff.	0.26	0.20	0.18	0.15

TABLE VI. Preselection efficiencies of the background events.

	ZZ	Z/γ^*	$t\bar{t}$
Four jet eff.	0.11	0.035	0.24
Four b-jet eff.	0.18	0.004	0.002
Correction eff.	0.54	0.54	0.31
Total pre-sel. eff.	1×10^{-3}	7×10^{-5}	2×10^{-4}
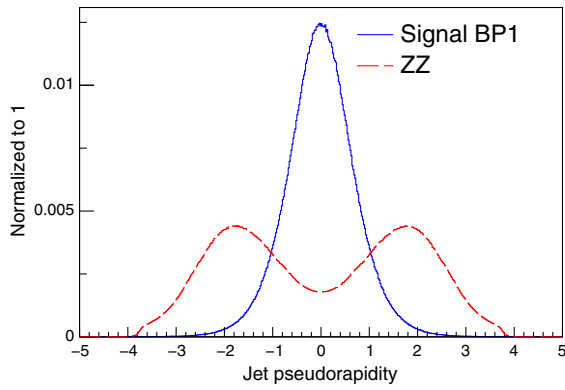
FIG. 4. Reconstructed jet pseudorapidity in signal (BP1) and ZZ background.

FIG. 6. The corrected and uncorrected distributions of the b-jet pair invariant mass (BP2).

$(E_{\text{jet}} \times \sin\theta)$ further reduces the selection efficiency to 11% which appears in Table VI. The b-tagging requirement suppresses this background further, which, followed by the correction efficiency, results in preselection efficiency of 10^{-3}.

Figure 5 shows correction factor distributions in signal events (BP1) with average values of 1.16, 1.25, 1.21, and 1.26. While the mean values of the distributions are close to unity, their widths (RMS values) are 0.6, 1.1, 1.45, 1.98 for c_1 to c_4 respectively and again shows better performance of the correction for harder jets.

In order to verify the correction efficiency, using Fig. 5 as the example, the four correction factors c_1 to c_4 are found to be positive with efficiencies of 98, 96, 92, and 85% respectively. Since all factors are required to be positive, the quoted efficiencies are multiplied to yield a total efficiency of $\sim 70\%$ which is what we obtain in the event analysis.

The effect of the kinematic correction on the b-jet pair invariant mass is shown in Fig. 6 using BP2 as the example. Details of the b-jet pair selection are presented in the next sections.

The correction factor sensitivity to the reconstructed jet energies is verified by estimating the uncertainty of the correction factors due to the jet energy smearing.
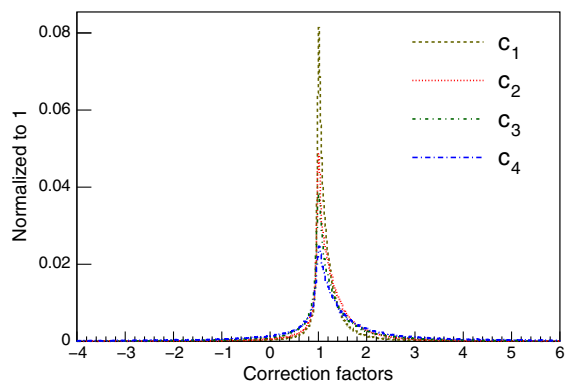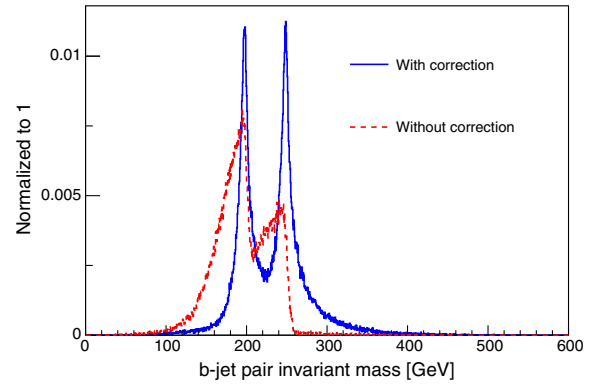
In order to do so, 1% additional smearing is applied on the jet four-momentum components on event-by-event basis and the distributions of relative errors of the correction factors are obtained. Results are shown in Fig. 7 and can be regarded as the correction factor smearing due to 1% uncertainty in the jet energies. The average uncertainties are 9, 11, 14, and 18% for c_1 to c_4 respectively.

The above estimates are one of the sources of the total uncertainty in the final distributions. However, a detailed study of different sources of uncertainties and their influence on the final distributions is beyond the scope of the current analysis and can be performed in a more detailed analysis based on full simulation of the detector.

E. b-jet pair selection based on their energies

Finding the true b-jet combination for Higgs boson reconstruction relies on two approaches. In the first approach, we note that decay products which move closer to the beam axis in the Higgs boson rest frame, acquire the highest and lowest energies when the Lorentz boost is applied to transform them from the Higgs boson frame to the laboratory frame. The other decay products belong to

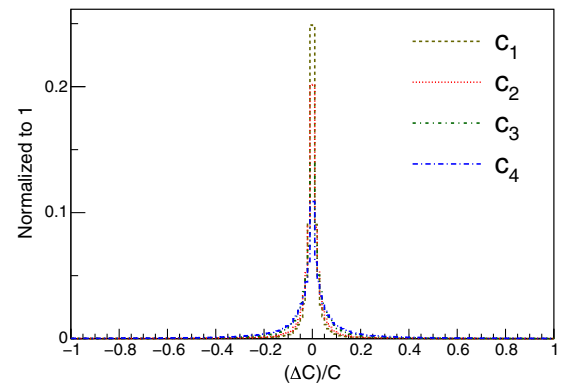FIG. 5. Correction factor distributions (BP1).

FIG. 7. The correction factor uncertainty due to 1% smearing on the jet energies.

the latter Higgs boson whose decay occurs at a larger angle with respect to the beam axis. Therefore having sorted the four b-jets in terms of their energies, b_1 and b_4 are expected to be the decay products of one Higgs boson and b_2 and b_3 from the other.

Since the decay products of the Higgs bosons fly at random angles with respect to the beam in each event, two possible scenarios may occur: $H \rightarrow b_1 b_4$, $A \rightarrow b_2 b_3$ or $H \rightarrow b_2 b_3$, $A \rightarrow b_1 b_4$. Some events choose the former scenario and the others choose the latter. Therefore, the b-jet pair invariant mass distribution dramatically shows both Higgs boson signals even if the distribution is obtained using only $b_1 b_4$ or $b_2 b_3$.

The above approach works best for equal Higgs boson masses. However, still a reasonable performance is observed for different Higgs boson masses in BP2 and two separated peaks are observed while for equal Higgs boson masses only one peak is observed. If the Higgs boson masses are significantly different, the efficiency of the algorithm is expected to be reduced. Figure 8 shows an example of the signal distributions (BP2) with two pairings, i.e., $b_1 b_2$ and $b_2 b_3$. The latter pairing results in slightly better distribution due to using more central b-jets.

F. b-jet pair selection based on their spatial distance

In the alternative approach, the two b-jet pairs are selected requiring minimum ΔR between each pair. In order to do so, sum of the two ΔR values are required to be minimum for the selected pairs.

The idea is based on the fact that, in general, b-jets from the decay of a particle, are expected to proceed at smaller ΔR values compared to two randomly selected b-jets, each one belonging to a different particle. Figure 9 shows an example of a signal event (BP1) in four b-jet final state in the detector using DELPHES display module [66].

The above requirement turns out to perform the b-jet pairing very similar to the previous scenario. As an example, in 97% of the BP1 events, the selected pairs
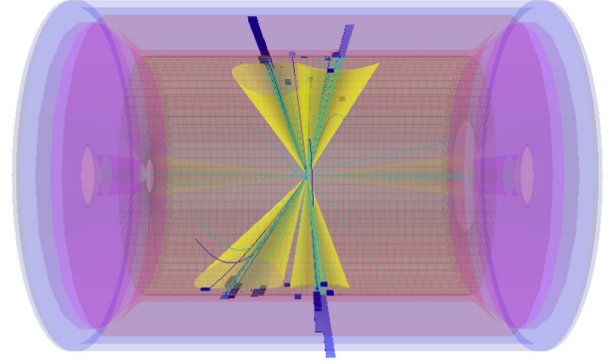

FIG. 9. 3D view of a signal event in the four b-jet final state in the detector. Jets from each Higgs boson tend to be collinear. The visible non-zero total momentum along the beam axis is due to the ISR + beamstrahlung as verified by inspecting the WHIZARD event file which was used for the simulation.

are $b_1 b_4$ and $b_2 b_3$ and $\Delta R(b_1 b_4) + \Delta R(b_2 b_3)$ is minimum among other possible combinations.

Therefore, for the final event selection, the two approaches described in sub-sections IV E and IV F are combined by requiring $\min(\Delta R(b_i b_j) + \Delta R(b_k b_l))$ and then demanding $i = 1$, $j = 4$, $k = 2$, and $l = 3$.

The performance of this requirement depends on the Higgs boson masses and their momenta and decreases when $m_H + m_A$ reaches the kinematic threshold $\sqrt{s}$. In such cases, the two Higgs bosons are almost created at rest with their decay products flying back-to-back at the maximum ΔR in the laboratory frame. However, we keep this requirement to suppress the large Z/γ^* background, which otherwise, extends to the signal region.

Figure 10 compares the two pairing scenarios with $\min(\Delta R)$ applied. The two distributions are in general better than those shown in Fig. 8 and the high energy tail is well suppressed. Again the $b_2 b_3$ pairing shows a better distribution compared to $b_1 b_4$. Therefore the final event distributions are obtained using $b_2 b_3$ with $\min(\Delta R)$ requirement applied which is the best scenario among the four possible choices shown in Figs. 8 and 10.
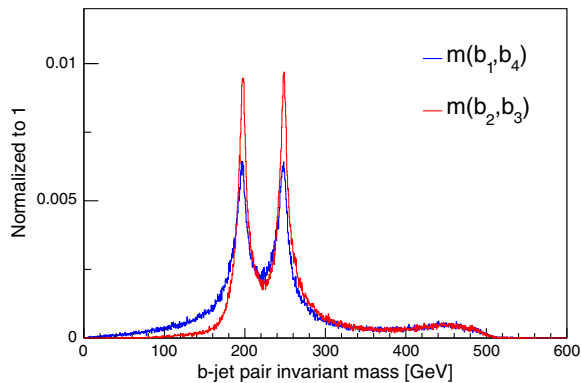

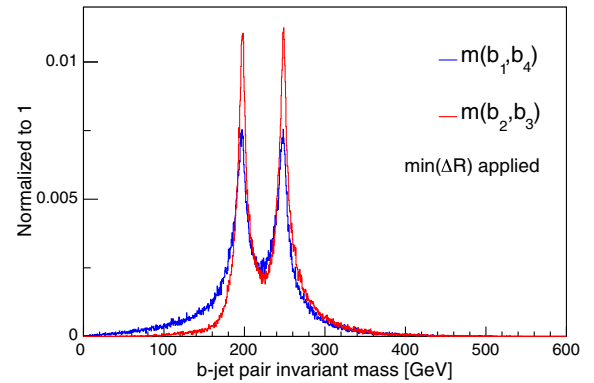FIG. 8. Comparison between the two selection scenarios: $b_1 b_4$ vs $b_2 b_3$

FIG. 10. Comparison between the two selection scenarios: $b_1 b_4$ vs $b_2 b_3$ with $\min(\Delta R)$ requirement applied.

G. Mass window optimization and final statistics

Figure 11 shows the $b\bar{b}$ invariant mass distributions in signal (BP2) and background events with b-tagging efficiencies of 80% and 70%. The $t\bar{t}$ contribution is sizable in these two b-tagging scenarios, however, this background is suppressed well by the tight b-tagging

scenario with average efficiency of 50%. Therefore the final results are shown in Fig. 12 based on the tight b-tagging scenario and normalized to integrated luminosity of 1 ab^{-1} which is enough for signal observation and can be easily used to rescale the results to any other integrated luminosity.
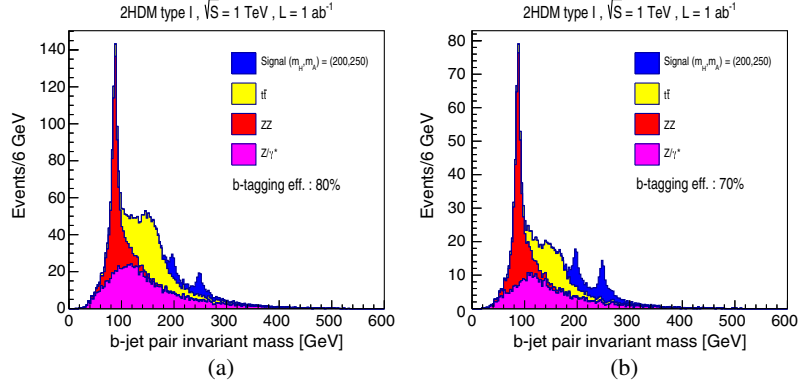

FIG. 11. The b-jet pair invariant mass distributions in signal and background events with b-tagging efficiencies of 80% and 70% respectively.
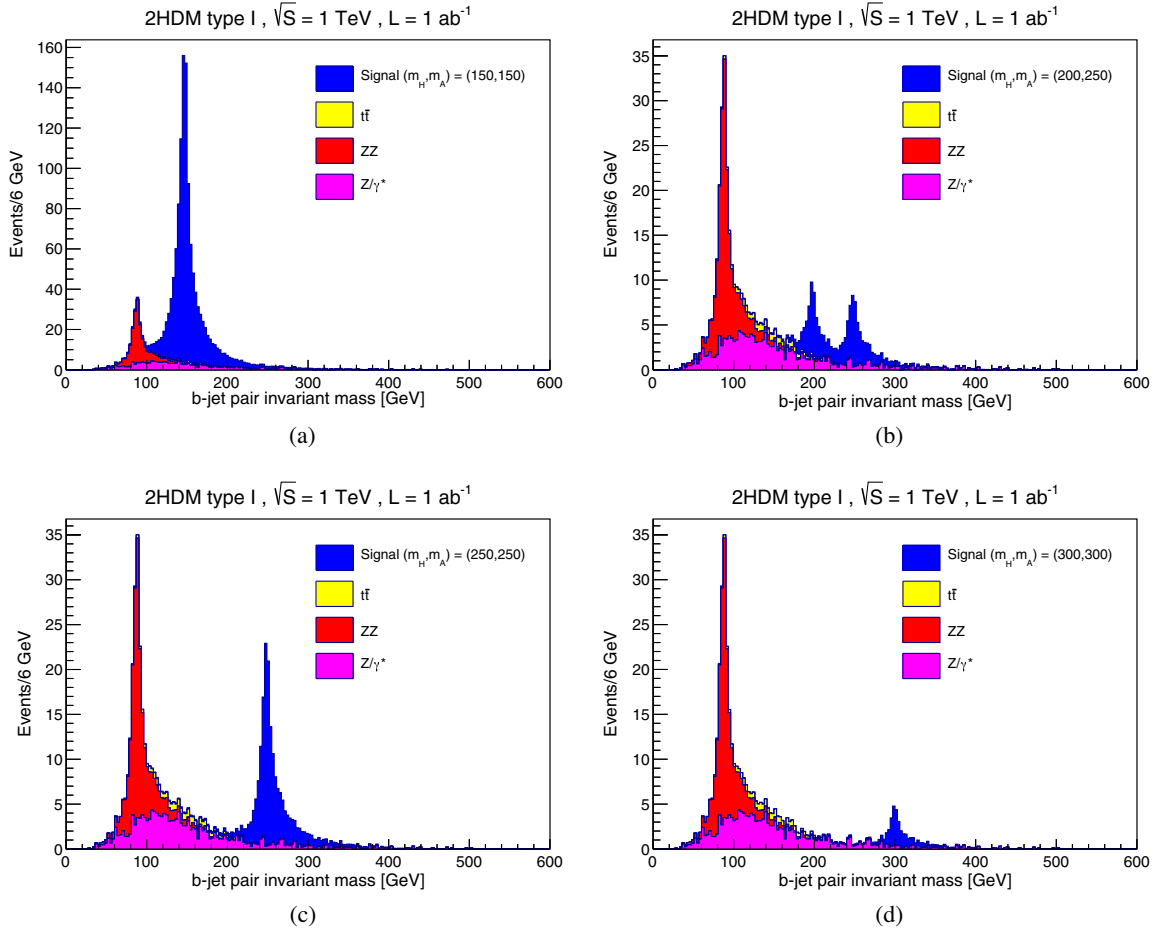

FIG. 12. The $b_2 b_3$ pair invariant mass distributions in signal and background events at $\sqrt{s} = 1$ TeV. The $\min(\Delta R)$ requirement has been applied as described in Sec. IV F.

TABLE VII. Selection efficiencies and the signal significance in different benchmark points based on the scenario described in Sec. IV F. The min(ΔR) cut means that indices which minimize $\Delta R(b_i b_j) + \Delta R(b_k b_l)$ are $i = 1$, $j = 4$, $k = 2$, and $l = 3$.

	BP1	BP2	BP3	BP4
min (ΔR) eff.	0.97	0.84	0.84	0.67
Eff. before mass win. cut	0.25	0.17	0.15	0.10
Mass window [GeV]	135–165	189–279	246–255	291–312
Mass window eff.	0.68	0.80	0.30	0.52
Total signal eff.	0.17	0.13	0.04	0.05
S	798	101	56	19
B	40	30	1.4	2.1
S/B	20	3.4	39	9.2
$S/\sqrt{B}$	125	18	47	13

The final statistics can be obtained by applying a mass window cut which is optimized to achieve the maximum signal significance. Results are presented in terms of an optimized interval (window) in the distribution and the number of signal and background events in the mass window are counted for signal significance evaluation.

Table VII shows the results of the adopted scenario presented in Sec. IV F, including the min ΔR cut efficiency, the mass window, total signal selection efficiency, number of signal (S) and background (B) events, signal to background ratio and the signal significance (defined as $S/\sqrt{B}$), for each benchmark point, at $\sqrt{s} = 1$ TeV and integrated luminosity of 1 ab^{-1}. The corresponding signal significance values using the first scenario (using $b_2 b_3$ without min(ΔR) cut) are 109, 14, 31, 9.3.

The quoted results presented in Table VII show that all selected benchmark points are observable, however, these results can easily be normalized to any other integrated luminosity.

The signal cross section is independent of $\tan \beta$ as the $Z.H.A$ vertex coupling is $\sin(\beta - \alpha)$ (which was set to unity in the current analysis). The Higgs boson decays with the above assumption depend on $\cot \beta$ and below the kinematic threshold of Higgs boson decay to $t\bar{t}$, BR($H/A \to b\bar{b}$) is dominant and independent of $\tan \beta$ at tree level. Therefore, results obtained in this analysis hold for other $\tan \beta$ values which are not yet excluded by experiments.

V. CONCLUSIONS

Possibility of observing CP-even and CP-odd neutral Higgs bosons, H and A, was studied at a lepton collider operating at $\sqrt{s} = 1$ TeV. The theoretical framework was chosen to be the type I 2HDM with $\sin(\beta - \alpha) = 1$ and $\tan \beta = 10$. The Higgs boson decay to b-jet pair, $H/A \to b\bar{b}$, was analyzed using a fast detector simulation and two different approaches for the signal observation were presented including a kinematic correction based on the four momentum conservation. It was illustrated that a clear signal can be observed on top of the background in the $b\bar{b}$ invariant mass distribution and the signal significance exceeds 5σ in all benchmark points at integrated luminosity of 1 ab^{-1}. The current analysis contains improvements in several aspects, i.e., the use of beam spectrum in event generation, more dedicated ILC detector card (`ILCgen`), application of several b-tagging scenarios and using updated software related to theoretical and experimental considerations. Compared to the previous results reported in [35] the observed signal distributions are much sharper and well located above the assumed Higgs boson masses, the total background is more suppressed and the signal to background ratio and the signal significance also show reasonable enhancements and more points proved to be explorable with the current analysis setup.

[1] P. W. Higgs, Spontaneous symmetry breakdown without massless bosons, Phys. Rev. **145,** 1156 (1966).

[2] T. W. Kibble, Symmetry breaking in Non-Abelian gauge theories, Phys. Rev. **155,** 1554 (1967).

[3] P. W. Higgs, Broken symmetries, massless particles and gauge fields, Phys. Lett. **12,** 132 (1964).

[4] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, Phys. Rev. Lett. **13,** 508 (1964).

[5] F. Englert and R. Brout, Broken Symmetry and the Mass of Gauge Vector Mesons, Phys. Rev. Lett. **13,** 321 (1964).

[6] G. S. Guralnik, C. R. Hagen, and T. W. Kibble, Global Conservation Laws and Massless Particles, Phys. Rev. Lett. **13,** 585 (1964).

[7] G. Aad *et al.* (ATLAS Collaboration), Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. B **716,** 1 (2012).

[8] S. Chatrchyan *et al.* (CMS Collaboration), Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys. Lett. B **716,** 30 (2012).

[9] G. Aad *et al.* (ATLAS Collaboration), Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment, Phys. Rev. D **101,** 012002 (2020).

[10] G. Aad *et al.* (ATLAS Collaboration), *CP* Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and

tH Processes Using $H \to \gamma\gamma$ with the ATLAS Detector, Phys. Rev. Lett. **125**, 061802 (2020).

[11] G. Aad *et al.* (ATLAS Collaboration), Test of *CP* invariance in vector-boson fusion production of the Higgs boson in the $H \to \tau\tau$ channel in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Lett. B **805**, 135426 (2020).

[12] G. Aad *et al.* (ATLAS Collaboration), Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector, Eur. Phys. J. C **75**, 476 (2015); Erratum, Eur. Phys. J. C **76**, 152 (2016).

[13] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement and interpretation of differential cross sections for Higgs boson production at $\sqrt{s} = 13$ TeV, Phys. Lett. B **792**, 369 (2019).

[14] A. M. Sirunyan *et al.* (CMS Collaboration), Combined measurements of Higgs boson couplings in proton–proton collisions at $\sqrt{s} = 13$TeV, Eur. Phys. J. C **79**, 421 (2019).

[15] A. M. Sirunyan *et al.* (CMS Collaboration), Measurements of t̄t*H* Production and the *CP* Structure of the Yukawa Interaction Between the Higgs Boson and Top Quark in the Diphoton Decay Channel, Phys. Rev. Lett. **125**, 061801 (2020).

[16] G. Aad *et al.* (ATLAS and CMS Collaborations), Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV, J. High Energy Phys. 08 (2016) 045.

[17] G. C. Branco, Spontaneous *CP* nonconservation and natural flavor conservation: A minimal model, Phys. Rev. D **22**, 2901 (1980).

[18] T. D. Lee, A theory of spontaneous T violation, Phys. Rev. D **8**, 1226 (1973).

[19] S. L. Glashow and S. Weinberg, Natural conservation laws for neutral currents, Phys. Rev. D **15**, 1958 (1977).

[20] I. J. Aitchison, Supersymmetry and the MSSM: An elementary introduction, arXiv:hep-ph/0505105.

[21] E. Ma and D. Ng, New supersymmetric option for two Higgs doublets, Phys. Rev. D **49**, 6164 (1994).

[22] A. Djouadi, The anatomy of electroweak symmetry breaking Tome II: The Higgs bosons in the minimal supersymmetric model, Phys. Rep. **459**, 1 (2008).

[23] G. Aad *et al.* (ATLAS Collaboration), Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the $\ell\ell bb$ and $\ell\ell WW$ final states in *PP* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Eur. Phys. J. C **81**, 396 (2021).

[24] M. Aaboud *et al.* (ATLAS Collaboration), Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the $\ell\ell bb$ final state in *PP* collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B **783**, 392 (2018).

[25] A. M. Sirunyan *et al.* (CMS Collaboration), Search for new neutral Higgs bosons through the $H \to ZA \to \ell^+\ell^- b\bar{b}$ process in *pp* collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 03 (2020) 055.

[26] A. M. Sirunyan *et al.* (CMS Collaboration), Search for a heavy pseudoscalar boson decaying to a Z and a Higgs boson at $\sqrt{s} = 13$ TeV, Eur. Phys. J. C **79**, 564 (2019).

[27] B. Grzadkowski, H. E. Haber, O. M. Ogreid, and P. Osland, Heavy Higgs boson decays in the alignment limit of the 2HDM, J. High Energy Phys. 12 (2018) 056.

[28] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, Theory and phenomenology of two-Higgs-doublet models, Phys. Rep. **516**, 1 (2012).

[29] S. Davidson and H. E. Haber, Basis-independent methods for the two-Higgs-doublet model, Phys. Rev. D **72**, 035004 (2005); Erratum, Phys. Rev. D **72**, 099902 (2005).

[30] J. Mrazek, A. Pomarol, R. Rattazzi, M. Redi, J. Serra, and A. Wulzer, The other natural two Higgs doublet model, Nucl. Phys. **B853**, 1 (2011).

[31] F. Mahmoudi and O. Stl, Flavor constraints on two-Higgs-doublet models with general diagonal Yukawa couplings, Phys. Rev. D **81**, 035016 (2010).

[32] J. F. Gunion and H. E. Haber, The *CP* conserving two Higgs doublet model: The approach to the decoupling limit, Phys. Rev. D **67**, 075019 (2003).

[33] M. Aoki, S. Kanemura, K. Tsumura, and K. Yagyu, Models of Yukawa interaction in the two Higgs doublet model, and their collider phenomenology, Phys. Rev. D **80**, 015017 (2009).

[34] V. Barger, J. L. Hewett, and R. J. N. Phillips, New constraints on the charged Higgs sector in two-Higgs-doublet models, Phys. Rev. D **41**, 3421 (1990).

[35] M. Hashemi and M. Mahdavikhorrami, Analysis of b quark pair production signal from neutral 2HDM Higgs bosons at future linear colliders, Eur. Phys. J. C **78**, 485 (2018).

[36] M. Hashemi and G. Haghighat, Observability of 2HDM neutral Higgs bosons with different masses at future e^+e^- linear colliders, Nucl. Phys. **B951**, 114903 (2020).

[37] M. Hashemi and G. Haghighat, Capability of future linear colliders to discover heavy neutral *CP*-even and *CP*-odd Higgs bosons within Type-I 2HDM, J. Phys. G **45**, 095005 (2018).

[38] M. Hashemi and G. Haghighat, Search for neutral Higgs bosons within Type-I 2HDM at future linear colliders, Eur. Phys. J. C **79**, 419 (2019).

[39] M. Hashemi and G. Haghighat, Search for neutral Higgs bosons decaying to $b\bar{b}$ in the flipped 2HDM at future e^-e^+ linear colliders, Phys. Rev. D **100**, 015047 (2019).

[40] M. Hashemi, Leptophilic neutral Higgs bosons in two Higgs doublet model at a linear collider, Eur. Phys. J. C **77**, 302 (2017).

[41] M. Hashemi and G. Haghighat, Search for heavy neutral *CP*-even Higgs within lepton-specific 2HDM at a future linear collider, Phys. Lett. B **772**, 426 (2017).

[42] M. Hashemi, Possibility of observing Higgs bosons at the ILC in the lepton-specific 2HDM, Phys. Rev. D **98**, 115004 (2018).

[43] T. Barklow, J. Brau, K. Fujii, J. Gao, J. List, N. Walker, and K. Yokoya, ILC operating scenarios, arXiv:1506.07830.

[44] P. Bambade, T. Barklow, T. Behnke, M. Berggren, J. Brau, P. Burrows, D. Denisov, A. Faus-Golfe, B. Foster and K. Fujii *et al.*, The international linear collider: A global project, arXiv:1903.01629.

[45] N. G. Deshpande and E. Ma, Pattern of symmetry breaking with two Higgs doublets, Phys. Rev. D **18**, 2574 (1978).

[46] H. Hffel and G. Pcsik, Unitarity bounds on Higgs boson masses in the Weinberg-Salam model with two Higgs doublets, Z. Phys. C **8**, 13 (1981).

[47] J. Maalampi, J. Sirkka, and I. Vilja, Tree level unitarity and triviality bounds for two-Higgs models, Phys. Lett. B **265,** 371 (1991).

[48] A. G. Akeroyd, A. Arhrib, and E. Naimi, Note on tree-level unitarity in the general two Higgs doublet model, Phys. Lett. B **490,** 119 (2000).

[49] D. Eriksson, J. Rathsman, and O. Stl, 2HDMCtwo-Higgs-doublet model calculator, Comput. Phys. Commun. **181,** 189 (2010).

[50] D. Eriksson, J. Rathsman, and O. Stl, 2HDMCtwo-Higgs-doublet model calculator, Comput. Phys. Commun. **181,** 833 (2010).

[51] R. Harlander, M. Mühlleitner, J. Rathsman, M. Spira, and O. Stål, Interim recommendations for the evaluation of Higgs production cross sections and branching ratios at the LHC in the two-Higgs-doublet model, arXiv:1312.5571.

[52] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein, and K. E. Williams, HiggsBounds: Confronting arbitrary Higgs sectors with exclusion bounds from LEP and the Tevatron, Comput. Phys. Commun. **181,** 138 (2010).

[53] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein, and K. E. Williams, HiggsBounds 2.0.0: Confronting neutral and charged Higgs sector predictions with exclusion bounds from LEP and the Tevatron, Comput. Phys. Commun. **182,** 2605 (2011).

[54] P. Bechtle, O. Brein, S. Heinemeyer, O. Stal, T. Stefaniak, G. Weiglein, and K. Williams, Recent developments in HiggsBounds and a preview of HiggsSignals, Proc. Sci., CHARGED2012 (**2012**) 024.

[55] P. Bechtle, O. Brein, S. Heinemeyer, O. Stl, T. Stefaniak, G. Weiglein, and K. E. Williams, HiggsBounds-4: Improved tests of extended Higgs sectors against exclusion bounds from LEP, the Tevatron and the LHC, Eur. Phys. J. C **74,** 2693 (2014).

[56] P. Bechtle, S. Heinemeyer, O. Stal, T. Stefaniak, and G. Weiglein, Applying exclusion likelihoods from LHC searches to extended Higgs sectors, Eur. Phys. J. C **75,** 421 (2015).

[57] P. Bechtle, S. Heinemeyer, O. Stl, T. Stefaniak, and G. Weiglein, HiggsSignals: Confronting arbitrary Higgs sectors with measurements at the Tevatron and the LHC, Eur. Phys. J. C **74,** 2711 (2014).

[58] O. Stål and T. Stefaniak, Constraining extended Higgs sectors with HiggsSignals, Proc. Sci., EPS-HEP2013 (**2013**) 314.

[59] P. Bechtle, S. Heinemeyer, O. Stål, T. Stefaniak, and G. Weiglein, Probing the Standard Model with Higgs signal rates from the Tevatron, the LHC and a future ILC, J. High Energy Phys. 11 (2014) 039.

[60] E. Leogrande, P. Roloff, U. Schnoor, and M. Weber, A DELPHES card for the CLIC detector, arXiv:1909.12728.

[61] W. Kilian, T. Ohl, and J. Reuter, WHIZARD: Simulating multi-particle processes at LHC and ILC, Eur. Phys. J. C **71,** 1742 (2011).

[62] M. Moretti, T. Ohl, and J. Reuter, O'Mega: An optimizing matrix element generator, arXiv:hep-ph/0102195.

[63] The CIRCE2 files are available at the URL https://whizard .hepforge.org/circe_files/.

[64] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos, A. Buckley, J. M. Butterworth, L. Dudko, S. Frixione and L. Garren *et al.*, A Standard format for Les Houches event files, Comput. Phys. Commun. **176,** 300 (2007).

[65] T. Sjstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA 8.2, Comput. Phys. Commun. **191,** 159 (2015).

[66] J. De Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaitre, A. Mertens, M. Selvaggi, and (Delphes 3 Collaboration), DELPHES 3: A modular framework for fast simulation of a generic collider experiment, J. High Energy Phys. 02 (2014) 057.

[67] M. Cacciari, FASTJET: A code for fast k_t clustering, and more, in *Deep Inelastic Scattering. Proceedings, 14th International Workshop, DIS 2006, Tsukuba, Japan* (2006), pp. 487–490 [arXiv:hep-ph/0607071].

[68] M. Cacciari, G. P. Salam, and G. Soyez, FASTJET user manual, Eur. Phys. J. C **72,** 1896 (2012).

[69] M. Cacciari, G. P. Salam, and G. Soyez, The anti-k_t clustering algorithm, J. High Energy Phys. 04 (2008) 063.

[70] M. A. Thomson, Particle flow calorimetry and the Pandor-aPFA algorithm, Nucl. Instrum. Methods Phys. Res., Sect. A **611,** 25 (2009).

[71] I. G. Knowles and G. D. Lafferty, Hadronization in Z^0 decay, J. Phys. G **23,** 731 (1997).

[72] D. Buskulic *et al.* (ALEPH Collaboration), Performance of the ALEPH detector at LEP, Nucl. Instrum. Methods Phys. Res., Sect. A **360,** 481 (1995).

[73] R. Brun and F. Rademakers, ROOTan object oriented data analysis framework, Nucl. Instrum. Methods Phys. Res., Sect. A **389,** 81 (1997).

[74] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos, A. Buckley, J. M. Butterworth, L. Dudko, S. Frixione, L. Garren, and S. Gieseke, A standard format for Les Houches event files, Comput. Phys. Commun. **176,** 300 (2007).