

# Analysis of the $gg \rightarrow H \rightarrow hh \rightarrow 4\tau$ process in the 2HDM lepton specific model at the LHC

Yan Ma,<sup>1,\*</sup> A. Arhrib,<sup>2,†</sup> S. Moretti,<sup>3,‡</sup> S. Semlali,<sup>4,§</sup> Y. Wang,<sup>5,¶</sup> and Q.S. Yan<sup>6,\*\*</sup>

<sup>1</sup>*College of Physics and Electronic Information,*

*Inner Mongolia Normal University, Hohhot 010022, PR China*

<sup>2</sup>*Abdelmalek Essaadi University, Faculty of Sciences and Techniques, B.P. 2117 Tétouan, Tanger, Morocco &*

*Department of Physics and Center for Theory and Computation,*

*National Tsing Hua University, Hsinchu, Taiwan 300*

<sup>3</sup>*School of Physics & Astronomy, University of Southampton, Southampton SO17 1BJ, UK &*

*Department of Physics & Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden*

<sup>4</sup>*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK &*

*Particle Physics Department, Rutherford Appleton Laboratory,*

*Chilton, Didcot, Oxon OX11 0QX, United Kingdom*

<sup>5</sup>*College of Physics and Electronic Information,*

*Inner Mongolia Normal University, Hohhot, 010022, China &*

*Key Laboratory for Physics and Chemistry of Functional Materials,*

*Inner Mongolia Normal University, Hohhot, 010022, China*

<sup>6</sup>*Center for Future High Energy Physics, Chinese Academy of Sciences, Beijing 100049, PR China &*

*School of Physics Sciences, University of Chinese Academy of Sciences, Beijing 100039, PR China*

## Abstract

We analyse the signature of a light Higgs boson pair in the 2-Higgs Doublet Model(2HDM) Type-X (or lepton specific) over the parameter spaces allowed by theoretical self-consistency requirements as well as the latest experimental constraints from the Large Hadron Collider (LHC), precision test data and  $B$  physics. Over the viable regions of the latter, wherein the Standard Model (SM)-like Higgs boson discovered at the LHC in 2012 is the heavier CP-even state of the 2HDM,  $H$ , it is found that the SM-like Higgs boson can decay into a pair of the lighter CP-even Higgs boson,  $h$ , via the process  $H \rightarrow hh$  with a Branching Ratio (BR) of 5% – 10% or so, (with  $2m_h < m_H = 125$  GeV). Furthermore, in the Type-X scenario, the lighter Higgs bosons  $h$  can dominantly decay into two  $\tau$ 's due to a large  $\tan\beta$ . Therefore, the pair of lighter Higgs bosons can altogether decay into a  $4\tau$  final state. In order to suppress the huge SM background events, we confine ourself to consider the fraction of signal events with two Same-Sign (SS)  $\tau$ 's further decaying into same sign leptons while the other two  $\tau$ 's decay hadronically. By using Monte Carlo (MC) and Machine Learning (ML) tools, we thus focus on the analysis of the signal process  $pp \rightarrow H \rightarrow hh \rightarrow \tau^+\tau^-\tau^+\tau^- \rightarrow \ell\nu_\ell\ell\nu_\ell\tau_h\tau_h$  (where  $\ell = e, \mu$  and  $\tau_h$  means a hadronic decay of the  $\tau$ ) and explore the feasibility of such a search at the LHC for a collision energy  $\sqrt{s} = 14$  TeV and a luminosity  $300 \text{ fb}^{-1}$ .

KEYWORDS: 2-Higgs Doublet Model, Type-X Scenario, Higgs Physics

---

\* 20214016005@mails.imnu.edu.cn

† a.arhrib@gmail.com

‡ s.moretti@soton.ac.uk; stefano.moretti@physics.uu.se

§ s.semlali@soton.ac.uk

¶ wangyan@imnu.edu.cn

\*\* yanqishu@ucas.ac.cn

## CONTENTS

I. Introduction	2
II. The 2HDM	4
III. Bounds and constraints in the parameter space scans	5
IV. Collider phenomenology	8
A. Event generation and selection	9
B. Event reconstruction	10
C. Multi-variate analysis	13
D. Significances	15
V. Conclusions	17
Acknowledgments	17
References	17

## I. INTRODUCTION

The LHC has played an irreplaceable role in our understanding of the microcosmic subatomic structure, particularly with the discovery of a 125 GeV scalar particle by the ATLAS and CMS experimental teams in 2012 [1, 2]. This was the last particle predicted by the Standard Model (SM) and its properties measured in the above experiments are currently consistent with the particle predicted by the SM at the  $5\sigma$  level [3, 4]. To investigate the Higgs sector at the LHC, wherein the presence of additional (pseudo)scalar particles might exist, remains an intriguing task, particularly in new physics models with additional doublet fields. Therefore, to discover extra Higgs bosons at the LHC is one of the prime targets of Beyond the SM (BSM) signal searches.

One of the simplest extensions of the SM is the 2HDM, which contains two complex Higgs doublets [5–7]. After Electro-Weak Symmetry Breaking (EWSB), there are three Goldstone bosons, which are ‘eaten’ by the  $W^\pm$  and Z bosons while the remaining five degrees of freedom incarnate into physical Higgs bosons: two neutral CP-even scalars ( $h$  and  $H$ , with

$M_h < M_H$ ), one CP-odd pseudoscalar ( $A$ ) and two charged Higgs states  $H^\pm$  (with mixed CP status).

In order to be consistent with the stringent experimental constraints from Flavour Changing Neutral Currents (FCNCs) at the tree level, typically, a  $Z_2$  symmetry is introduced into the Yukawa sector such that each type of fermion only couples to one of the doublets of the 2HDM [8]. Depending on the  $Z_2$  charge assignments of the Higgs doublets, we can define four basic scenarios, known as (Yukawa) Types. In the lepton specific scenario (or Type-X), the extra Higgs doublet  $\phi_2$  only couples to the lepton sector and not to the quark sector, where the light Higgs boson mainly decays to  $\tau\tau$  for the wide range of parameter spaces [9].

There have been several experiments for searching the SM Higgs exotic decays through four fermion final states  $pp \rightarrow H \rightarrow XX \rightarrow f_1 f_2 f_3 f_4$ . Such process benefits from the large cross-section of  $\sigma(pp \rightarrow H)$ , and the branching ratio of the SM-like Higgs exotic decay  $Br(H \rightarrow hh)$  could reach up to 10% in some BSM models. Motivated by the recent searches for  $\tau\tau\tau\tau$  by CMS experiments [10–12], we will study the process  $pp \rightarrow H \rightarrow hh\tau\tau\tau\tau$  in the 2HDM type-X model. In such a scenario,  $h \rightarrow \tau^+\tau^-$  decay could be enhanced, which can lead to a sizable  $4\tau$  final state. Therefore, it is worth to examine whether this statement is robust enough after taking into account the effects of parton shower, hadronization, heavy flavor decays and detector resolution.

In this paper, we assume that the heavier CP-even Higgs boson  $H$  is the observed SM-like Higgs boson whose properties agree with the LHC measurements. In such a parameter space, a heavier Higgs boson can decay into two light Higgs bosons  $h$  when kinematically allowed, i.e.,  $2M_h < 125$  GeV. We perform a detailed MC study on the signal process  $pp \rightarrow H \rightarrow hh \rightarrow 4\tau \rightarrow lv_l lv_l \tau_h \tau_h$  (with  $l = e, \mu$  and  $\tau_h$ ) and relevant background processes, in order to examine its feasibility at the LHC. We do so after taking into account relevant theoretical and experimental conditions. In order to improve the experimental significance, we will require two leptons to be the same sign (SS) leptons. Our analysis demonstrates that, by using this signature, the relevant parameter space can be either discovered or ruled out by using the current integrated luminosity at the LHC.

The paper is organised as follows. In the next section, we briefly describe the 2HDM and its Yukawa scenarios, then introduce a few Benchmark Points (BPs) for our MC analysis which pass all present theoretical and experimental constraints. In the following section, we perform a detailed collider analysis of these BPs and examine the potential to discover

the aforementioned signature within the 2HDM Type-X scenario. Finally, we present some conclusions.

## II. THE 2HDM

The scalar sector of the 2HDM consists of two weak isospin doublets with hypercharge  $Y = 1$ . The most general Higgs potential for the 2HDM that complies with the  $SU(2)_L \times U(1)_Y$  gauge structure of the EW sector of the SM has the following form [7]:

$$\begin{aligned}
V(\phi_1, \phi_2) = & m_{11}^2(\phi_1^\dagger\phi_1) + m_{22}^2(\phi_2^\dagger\phi_2) - [m_{12}^2(\phi_1^\dagger\phi_2) + \text{h.c.}] \\
& + \frac{1}{2}\lambda_1(\phi_1^\dagger\phi_1)^2 + \frac{1}{2}\lambda_2(\phi_2^\dagger\phi_2)^2 + \lambda_3(\phi_1^\dagger\phi_1)(\phi_2^\dagger\phi_2) + \lambda_4(\phi_1^\dagger\phi_2)(\phi_2^\dagger\phi_1) \\
& + \frac{1}{2} [\lambda_5(\phi_1^\dagger\phi_2)^2 + \text{h.c.}] + \left\{ [\lambda_6(\phi_1^\dagger\phi_1) + \lambda_7(\phi_2^\dagger\phi_2)] (\phi_1^\dagger\phi_2) + \text{h.c.} \right\}, \quad (1)
\end{aligned}$$

where  $\phi_1$  and  $\phi_2$  are the two Higgs doublet fields. By hermiticity of such a potential,  $\lambda_{1,2,3,4}$  as well as  $m_{11,22}^2$  are real parameters while  $\lambda_{5,6,7}$  and  $m_{12}^2$  can be complex, in turn enabling possible Charge and Parity (CP) violation effects in the Higgs sector. Upon enforcing two minimization conditions of the potential,  $m_{11}^2$  and  $m_{22}^2$  can be replaced by  $v_{1,2}$ , which are the Vacuum Expectation Values (VEVs) of the Higgs doublets  $\phi_{1,2}$ , respectively. Moreover, the coupling  $\lambda_{1,2,3,4,5}$  can be substituted by the four physical Higgs masses (i.e.,  $M_h, M_H, M_A$  and  $M_{H^\pm}$ ) and the parameter  $\sin(\beta - \alpha)$ , where  $\alpha$  and  $\beta$  are, respectively, the mixing angles between CP-even and the angle related to the VEV values, i.e.,  $\tan\beta = v_1/v_2$ . Thus, the independent input parameters can be taken as  $M_h, M_H, M_A$  and  $M_{H^\pm}$ ,  $\lambda_6, \lambda_7, \sin(\beta - \alpha)$ ,  $\tan(\beta)$  and  $m_{12}^2$ .

If both Higgs doublet fields of the general 2HDM couple to all fermions, the ensuing scenario can induce FCNCs in the Yukawa sector at tree level. As intimated, to remedy this, a  $Z_2$  symmetry is imposed on the Lagrangian such that each fermion type interacts with only one of the Higgs doublets [8]. As a consequence, there are four possible types of 2HDM, namely Type-I, Type-II, Type-X (or lepton specific) and Type-Y (or flipped). However, such a symmetry is explicitly broken by the quartic couplings  $\lambda_{6,7}$  and softly broken by the (squared) mass term  $m_{12}^2$ . In what follows, we shall consider a CP-conserving (i.e.,  $m_{12}^2$  and  $\lambda_5$  are real) 2HDM Type-X and assume that  $\lambda_6 = \lambda_7 = 0$  to forbid the explicit breaking of  $Z_2$ , while also taking  $m_{12}^2$  to be generally small, thereby preventing large FCNCs at tree level, which are incompatible with experiment.

In general, the couplings of the neutral and charged Higgs bosons to fermions can be described by the Yukawa Lagrangian given by [7]

$$-\mathcal{L}_{\text{Yukawa}} = \sum_{f=u,d,l} \left( \frac{m_f}{v} \kappa_f^h \bar{f} f h + \frac{m_f}{v} \kappa_f^H \bar{f} f H - i \frac{m_f}{v} \kappa_f^A \bar{f} \gamma_5 f A \right) + \left( \frac{V_{ud}}{\sqrt{2}v} \bar{u} (m_u \kappa_u^A P_L + m_d \kappa_d^A P_R) d H^+ + \frac{m_l \kappa_l^A}{\sqrt{2}v} \bar{\nu}_L l_R H^+ + \text{h.c.} \right), \quad (2)$$

where  $\kappa_f^S$  ( $S = h, H$  and  $A$ ) are the Yukawa couplings in the 2HDM, which are illustrated in Tab. I for the Type-X under consideration. Here,  $V_{ud}$  refers to a Cabibbo-Kobayashi-Maskawa (CKM) matrix element and  $P_{L,R}$  denote the left- and right-handed projection operators. The coupling of the two CP-even states  $h$  and  $H$  to gauge bosons  $VV$  ( $V = W^\pm, Z$ ) are proportional to  $\sin(\beta - \alpha)$  and  $\cos(\beta - \alpha)$ , respectively. If we assume that either  $h$  or  $H$  can be the observed SM-like Higgs boson, the coupling to gauge bosons is obtained for  $h$  when  $\cos(\beta - \alpha) \rightarrow 0$  and for  $H$  when  $\sin(\beta - \alpha) \rightarrow 0$ . Therefore, each scenario can explain the 125 GeV Higgs signal at the LHC. However, following our works [13–16], though, we shall focus in the present paper on the scenario where  $H$  mimics the observed signal with mass  $\sim 125$  GeV (as previously intimated).

	$\kappa_u^S$	$\kappa_d^S$	$\kappa_\ell^S$
$h$	$\cos \alpha / \sin \beta$	$\cos \alpha / \sin \beta$	$-\sin \alpha / \cos \beta$
$H$	$\sin \alpha / \sin \beta$	$\sin \alpha / \sin \beta$	$\cos \alpha / \cos \beta$
$A$	$\cot \beta$	$-\cot \beta$	$\tan \beta$

TABLE I. Yukawa couplings of the fermions  $f = u, d$  and  $\ell$  to the neutral Higgs bosons  $S = h, H$  and  $A$  in the 2HDM Type-X.

### III. BOUNDS AND CONSTRAINTS IN THE PARAMETER SPACE SCANS

In order to identify regions that satisfy both theoretical requirements and experimental observations, the following numerical explorations of the possible parameter spaces have been studied:

$$M_h \in [15, 60] \text{ (GeV)}, \quad M_H = 125, \quad M_A = 245 \text{ (GeV)}, \quad M_{H^\pm} = 258 \text{ (GeV)} \\ \sin(\beta - \alpha) \in [-0.25, -0.05], \quad \tan(\beta) \in [2, 13], \quad m_{12}^2 \in [0, 500] \text{ (GeV)}^2 \quad (3)$$

The program `2HDMC-1.8.0` [17] was used for calculating the cross sections ( $\sigma$ 's) and BRs as well as testing the following theoretical and experimental constraints.

- Vacuum stability [6]: the scalar potential should be bounded by  $\lambda_1 > 0$ ,  $\lambda_2 > 0$ ,  $\lambda_3 > -\sqrt{\lambda_1\lambda_2}$ ,  $\lambda_3 + \lambda_4 - \lambda_5 > -\sqrt{\lambda_1\lambda_2}$ .
- Perturbativity constraints [21, 22], which imply the condition  $|\lambda_i| < 8\pi$  ( $i = \dots, 5$ ).
- Tree-level perturbative unitarity [20–22]: requiring unitary when  $2 \rightarrow 2$  scattering processes involving Higgs and gauge bosons at high energy.
- Oblique EW parameters ( $S$ ,  $T$  and  $U$ ) [23, 24], which control the mass splitting between the Higgs states, with the following measured values [25]:

$$S = -0.02 \pm 0.10, \quad T = 0.03 \pm 0.12. \quad (4)$$

The correlation factor between  $S$  and  $T$  is set as 0.92 for consistency at 95% Confidence Level (CL).

To account for potential additional Higgs bosons, exclusion bounds at 95% CL are enforced using the `HiggsBounds-5.9.0` program [26], which systematically checks each parameter point against the 95% CL exclusion limits derived from Higgs boson searches conducted by LEP, Tevatron and LHC experiments.

To ensure agreement with the measurements of the SM-like Higgs state, constraints are enforced using the `HiggsSignals-2.6.0` program [27], which incorporates the combined measurements of the SM-like Higgs boson from LHC Run-1 and Run-2 data.

Constraints from flavor physics are incorporated using `SuperIso v4.1` [28] by utilizing the following observables:

- $\text{BR}(B \rightarrow X_s \gamma) = (3.32 \pm 0.15) \times 10^{-4}$  [29],
- $\text{BR}(B^0 \rightarrow \mu^+ \mu^-)_{(\text{LHCb})} = (1.2^{+0.8}_{-0.7}) \times 10^{-10}$  [30, 31],
- $\text{BR}(B^0 \rightarrow \mu^+ \mu^-)_{(\text{CMS})} = (0.37^{+0.75}_{-0.67}) \times 10^{-10}$  [32],
- $\text{BR}(B_s \rightarrow \mu^+ \mu^-)_{(\text{LHCb})} = (3.09^{+0.46}_{-0.43}) \times 10^{-9}$  [30, 31],
- $\text{BR}(B_s \rightarrow \mu^+ \mu^-)_{(\text{CMS})} = (3.83^{+0.38}_{-0.36}) \times 10^{-9}$  [32],

- $\text{BR}(B \rightarrow \tau\nu) = (1.09 \pm 0.20) \times 10^{-4}$  [29].

In the left panel of Fig. 1, we show the branching ratio of the light Higgs decaying into pairs of  $\tau$  leptons within the allowed parameter space of the 2HDM Type-X. The total width of the  $h$  state ( $\Gamma_h$ ) is obviously dominated by  $h \rightarrow \tau\tau$ . Such a significant branching ratio is attributed to the enhancement of the light Higgs couplings to leptons at large  $\tan\beta$ . The branching ratio of the SM-like Higgs boson decaying into a pair of light Higgs bosons is displayed on the right panel. The precise measurements of the Higgs boson couplings [3, 4] tighten the constraints on Higgs decay to non-SM particles and force the  $\text{BR}(H \rightarrow hh)$  to be below 12%. Both ATLAS and CMS collaborations have reported an upper limit of 12% [3] and 16 % [4], respectively, on  $\text{BR}(H \rightarrow \text{BSM})$  at 95%CL.

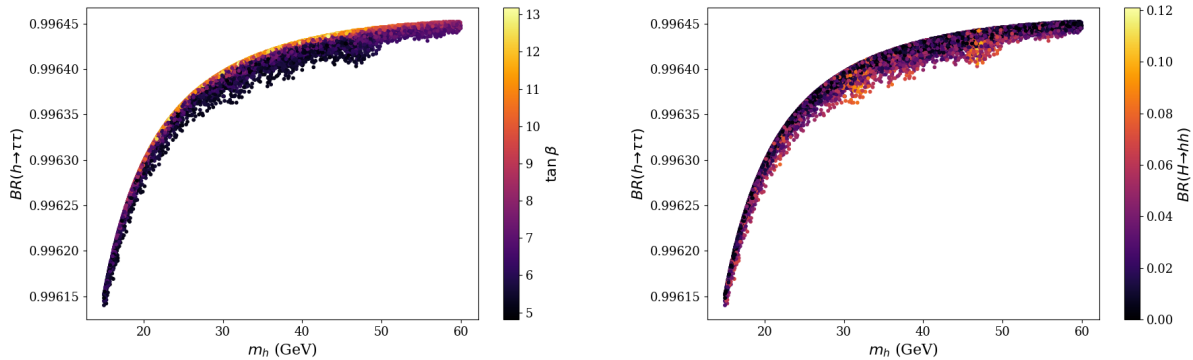


FIG. 1.  $\text{BR}(h \rightarrow \tau\tau)$  as a function of  $m_h$  vs.  $\tan\beta$  (left panel) and  $\text{BR}(H \rightarrow hh)$  (right panel) in the 2HDM type-X.

To test the allowed parts of the parameter space, twelve BPs are given in Tab. II, in terms of the 2HDM Type-X parameters entering our production and decay process, including the corresponding  $pp \rightarrow H \rightarrow 2h \rightarrow 4\tau$  cross sections at the LHC with  $\sqrt{s} = 14$  TeV. As one can see from such a table, the light Higgs boson mass varies from about 15 GeV to 60 GeV, which can be paired produced from a 125 GeV Higgs boson. The cross sections of the BPs are around 2 – 5 pb, which are large enough to be detected at the collider.

In this paper, we mainly focus on two SS taus further leptonically decay, while the other two tau hadronically decay (tau-jets). Thus, the signal will be SS leptons and two ( $\tau$ -)jets. The leptons are used for triggering purposes, so its kinematics is bound to comply with the trigger requirements.

	$m_h$	$\sin(\beta - \alpha)$	$\tan \beta$	$\sigma_{4\tau, 14 \text{ TeV}}$ (pb)
BP1	19.04	-0.12	8.06	2.53
BP2	21.8	-0.162	6.35	3.29
BP3	25.11	-0.16	5.82	2.18
BP4	28.00	-0.17	5.61	2.72
BP5	32.14	-0.15	6.36	4.01
BP6	36.02	-0.107	9.72	3.31
BP7	44.00	-0.188	5.10	2.43
BP8	46.88	-0.094	10.29	4.74
BP9	47.30	-0.10	9.92	3.88
BP10	50.98	-0.13	7.57	2.81
BP11	55.10	-0.149	6.45	2.65
BP12	58.16	-0.11	8.91	1.80

TABLE II. 2HDM Type-X input parameters and Leading Order (NLO) cross sections (at parton level with  $\sqrt{s} = 14$  TeV) for each BP are presented. The unit of all masses is GeV and we fix the SM-like Higgs boson mass as  $M_H = 125$  GeV.

#### IV. COLLIDER PHENOMENOLOGY

In this section, we present a detailed MC analysis for both signal and background events at the detector level. In the  $pp \rightarrow H \rightarrow hh \rightarrow \tau^+\tau^-\tau^+\tau^- \rightarrow lv_l lv_l \tau_h \tau_h$  process, a light Higgs boson decays into two  $\tau$  leptons and one of the  $\tau$  decays into a charged lepton and neutrino, while the other  $\tau$  decays hadronically (thus labelled as  $\tau_h$ ). As a result, the final states consist of two SS leptons and two jets. The SM backgrounds are:  $t\bar{t} \rightarrow lv_l lv_l b\bar{b}$ ,  $W^\pm tb \rightarrow lv_l lv_l b\bar{b}$  (where the  $W^\pm$  boson and  $b$ -jet are not from a top quark resonance),  $WWjj \rightarrow lv_l lv_l jj$ ,  $Zjj \rightarrow lv_l lv_l jj$ ,  $ZZ \rightarrow lv_l lv_l \tau_h \tau_h$ ,  $t\bar{t}Z \rightarrow lv_l lv_l b\bar{b} \tau \tau$ ,  $t\bar{t}ZZ \rightarrow lv_l lv_l lv_l b\bar{b} \tau \tau$  and  $t\bar{t}WW \rightarrow lv_l lv_l lv_l jj$ . In order to suppress the huge SM background process  $Zjj$ , we deliberately choose two SS leptons plus two hadronic  $\tau$ 's in the signal events, which is about 10% of the total number of the  $4\tau$  ones.



### A. Event generation and selection

We use MadGraph5\_aMC@NLO v3.4.0 [33] to compute the cross sections and generate both signal and background events at parton level. The following kinematic cuts are adopted for the signal and background generations:

$$|\eta(l, j)| < 2.5, p_T(l, j) > 10 \text{ GeV}, E_T^{\text{miss}} > 5 \text{ GeV}, \Delta R(ll, lj, jj) > 0.4. \quad (5)$$

Then parton level events are then passed to Pythia [34, 35] to simulate initial and final state radiation (i.e., the QED and QCD emissions), parton shower, hadronization and heavy flavor decays. We further use Delphes v3.4.2 [36] to simulate the detector effects. For each event, the anti- $k_t$  jet algorithm [37] is used to cluster jets with jet parameter  $\Delta R = 0.4$  in the FastJet package [38]. The following kinematic cuts are adopted in order to emulate the detector acceptance:

$$|\eta(l, j)| < 2.5, p_T(l, j) > 10 \text{ GeV}. \quad (6)$$

The leptons are used for triggering purposes, so its kinematics is bound to comply with the trigger requirements.<sup>1</sup>

Then, we select two leptons and two jets final states from detector simulated events. In order to improve the signal event rate, the tau tagging is not applied. In order to further suppress the huge SM backgrounds, the SS leptons are selected. The cross sections for the signal and background processes with detector acceptance cuts, with  $2l2j$  selection and with two SS lepton selection are shown in the Tab. III and in Tab. IV, respectively.

$\sigma$ (fb)	BP1	BP2	BP3	BP4	BP5	BP6	BP7	BP8	BP9	BP10	BP11	BP12
parton level generation	8.95	11.37	7.38	9.05	13.32	11.33	10.43	23.16	19.26	15.81	17.00	12.75
selection of $2l2j$	0.67	0.88	0.59	0.76	1.10	0.96	0.88	1.97	1.62	1.46	1.70	1.38
selection of SS $2l$	0.337	0.438	0.31	0.41	0.59	0.52	0.50	1.14	0.92	0.83	0.93	0.75

TABLE III. Cross sections for signals in the parton level, detector level, and after selection of SS leptons at  $\sqrt{s} = 14 \text{ TeV}$  for  $300 \text{ fb}^{-1}$ .

<sup>1</sup> The CMS developed a double muon trigger to search for low pt muons from B meson decays that are close together. The same approach has been adopted recently to search for a pairs of electron from B meson decays, by tightening the topological selection on the two electrons L1 objects. In our recent paper [39], we discussed the possibility of developing an electron-muon trigger to search for low pt leptons (pt 10 GeV)

$\sigma$ (fb)	$t\bar{t}$	$W^\pm tb$	$W^+W^-jj$	$Zjj$	$ZZ$	$t\bar{t}Z$	$t\bar{t}ZZ$	$t\bar{t}W^+W^-$
parton level generation	16060	518.3	1053	317600	18.89	0.49	$1.14 \times 10^{-4}$	0.02
selection of 2l2j	8787.7	289.9	530.1	151086	10.0	0.33	$1.1 \times 10^{-4}$	0.018
selection of SS 2l	19.43	0.62	1.99	0	2.51	0.079	$3.3 \times 10^{-5}$	$7.6 \times 10^{-3}$

TABLE IV. Background rates after the acceptance cuts in Eq. (6) at  $\sqrt{s} = 14$  TeV for  $300 \text{ fb}^{-1}$ .

In Fig. 2, the transverse momentum distributions of the leading and the second-leading jets and leptons as well as the transverse missing energy (MET) are presented for, e.g., BP2 and background events. It is clear that the transverse momenta of the signal final state are always much softer than for background events, because they emerge from a rather light Higgs bosons, which energies are capped at a mass of 125 GeV.

## B. Event reconstruction

To further improve the signal-to-background discrimination, we reconstruct some kinematic features of signal events. Specifically, because the light Higgs boson decays into two  $\tau$  and one of the  $\tau$ 's decays to a lepton and two neutrinos plus two  $\tau$  neutrinos emerge from hadronic  $\tau$  decays, there will be six neutrinos in a signal event. It is thus extremely hard to completely reconstruct the momentum of each light Higgs.

However, since the lepton momenta and MET are all very small, we can approximately reconstruct the light Higgs boson with only one lepton and one jet. Thus, there are two possible combinations with two leptons and two  $\tau$ -jets in each event. By defining a suitable  $\chi^2$ ,

$$\chi^2 = (M_{lj}^1 - M_h)^2 + (M_{lj}^2 - M_h)^2, \quad (7)$$

where  $M_h$  is the light Higgs mass, we pair the two leptons and two jets to find a combination which minimizes it and then assign these to  $M_{lj}^1$  and  $M_{lj}^2$ , respectively. Taking again BP2 as an example, the reconstructed light Higgs mass distributions are shown in Fig. 3.

In order to further separate signal from background events, we can also combine the two leptons and two jets together, denoting their invariant masses as  $M_{ll}$  and  $M_{jj}$ , respectively, as shown in Fig. 4. Since, for signal events, all particles come from the SM-like Higgs boson decaying, any such combination only yield a smaller invariant mass, typically, about half

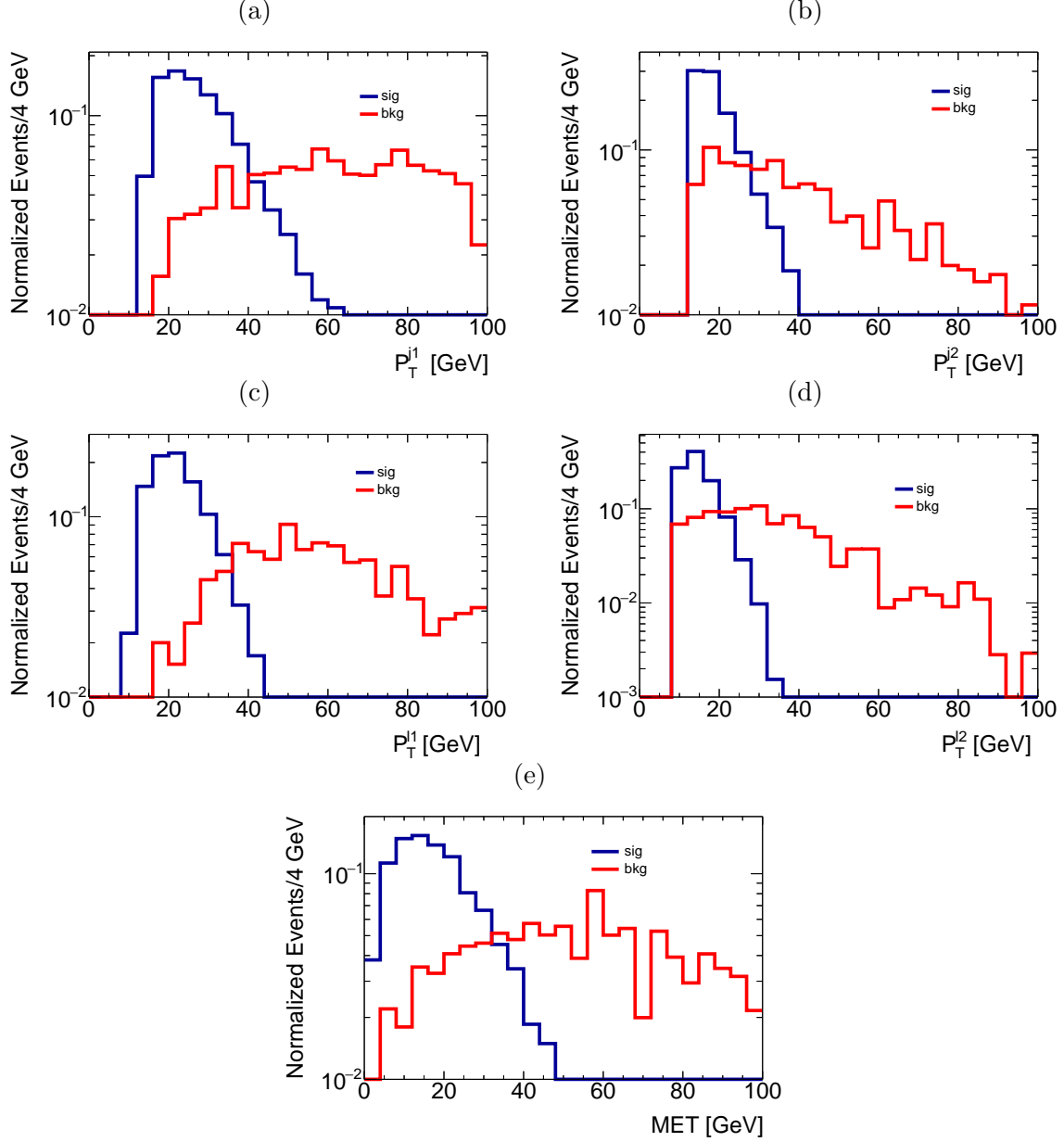


FIG. 2. Distributions in (a)-(d) transverse momenta of the leading/second-leading jets/leptons and (e) MET for BP2 and backgrounds when  $\sqrt{s} = 14$  TeV and  $L = 300 \text{ fb}^{-1}$ .

of the SM-like Higgs boson mass. In contrast, for background events, these lepton and jet pairs always come from a heavy resonance, a top (anti)quark and a  $W^\pm$  or  $Z$  boson, so the background spectra are somewhat harder in general.

We can further reconstruct the invariant mass of the SM-like Higgs boson  $M_H$  with two leptons and two jets but without any MET. It is expected that the sampled values of it will be lower than the real SM-like Higgs boson mass. Nevertheless, it is helpful enough to

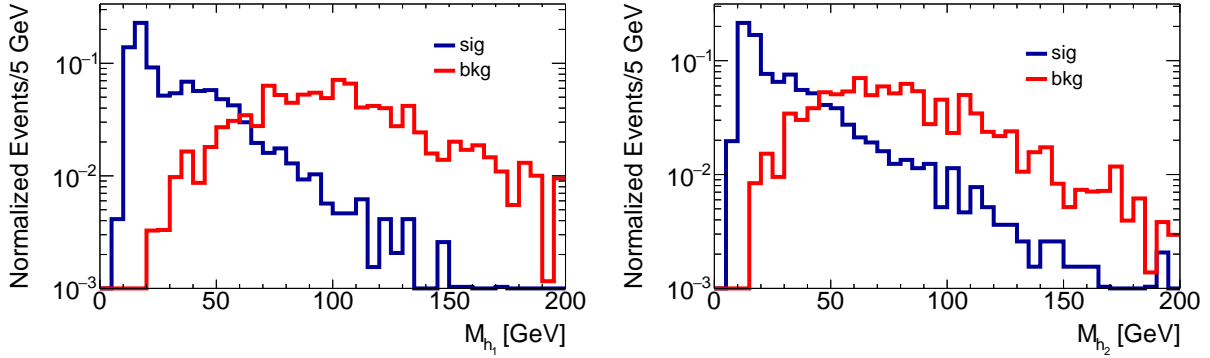


FIG. 3. Distributions in invariant mass of a lepton and a jet for BP2 and backgrounds. The reconstructed  $M_h$  values, corresponding to the best  $\chi^2$  fit, is labelled as  $M_{lj}^1$  (left) and  $M_{lj}^2$  (right), respectively, when  $\sqrt{s} = 14$  TeV and  $L = 300$  fb $^{-1}$ .

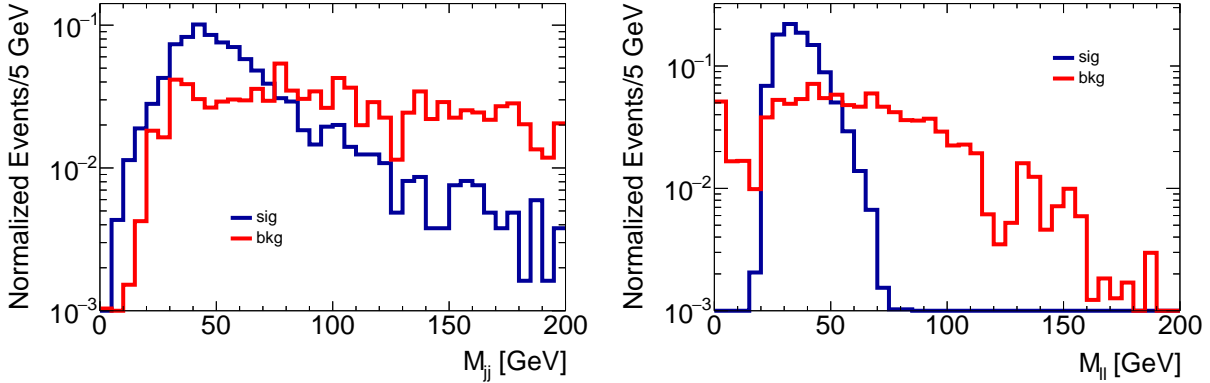


FIG. 4. Distributions in invariant mass of lepton (left) and jet (right) pairs for BP2 and backgrounds when  $\sqrt{s} = 14$  TeV and  $L = 300$  fb $^{-1}$ .

distinguish the signal and background events further, see Fig. 5 (left).

Another useful kinematic variable is  $H_T$ , which is the scalar sum of all visible final states transverse momenta, which are shown in Fig. 5 (right).

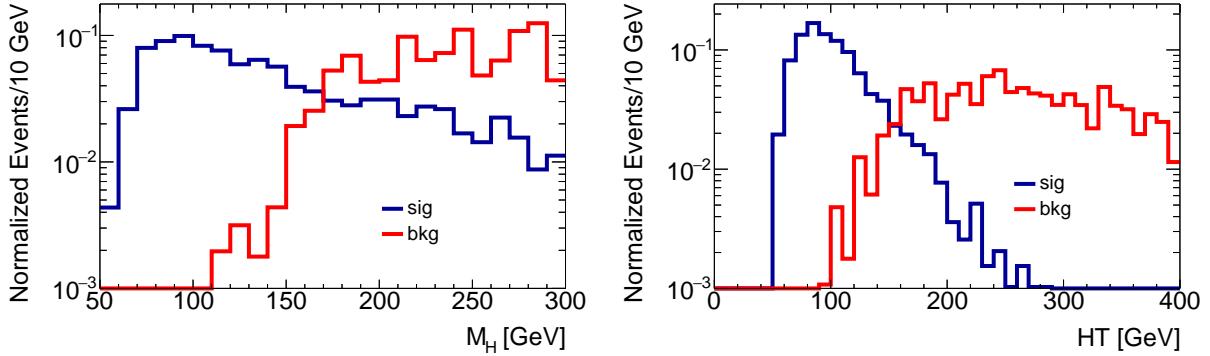


FIG. 5. Distributions in invariant mass of two leptons and two jets (left) and in scalar sum of visible momenta (right) for BP2 and backgrounds when  $\sqrt{s} = 14$  TeV and  $L = 300 \text{ fb}^{-1}$ .

### C. Multi-variate analysis

To optimize the signal and background distinction, a Gradient-Boosted Decision Tree (GBDT) approach is further applied, which is implemented in the Toolkit for Multi-Variate Analysis (TMVA) within ROOT [40].

Ten input variables in total are used for the GBDT/TMVA analysis, which are shown in Tab. V. In addition to the above invariant masses and  $H_T$ , we also calculate five angles between pairs of objects in the final state,  $\cos(\theta_{l_1 j_1})$ ,  $\cos(\theta_{l_1 j_2})$ ,  $\cos(\theta_{l_2 j_1})$ ,  $\cos(\theta_{l_2 j_2})$  and  $\cos(\theta_{h_1 h_2})$ , which is defined as the cosine of the angle between the two reconstructed final states. The distributions are shown in Fig. 6. As we expect these to be useful, since in the signal they all tend to be collinear to each other (because of the boost structure herein) unlike the background case.

Energy variables	$M_{l\tau_j}^1$	$M_{l\tau_j}^2$	$M_{ll}$	$M_{jj}$	$H_T$
Angular variables	$\cos(\theta_{h_1 h_2})$	$\cos(\theta_{l_1 j_1})$	$\cos(\theta_{l_1 j_2})$	$\cos(\theta_{l_2 j_1})$	$\cos(\theta_{l_2 j_2})$

TABLE V. The input observables used in the GBDT analysis.

In the training stage, it is found that the invariant mass variables are powerful, but angular variables are still useful. Ultimately, for the GBDT output, a very good separation between the signal and backgrounds can be achieved, which is shown in Fig. 7.

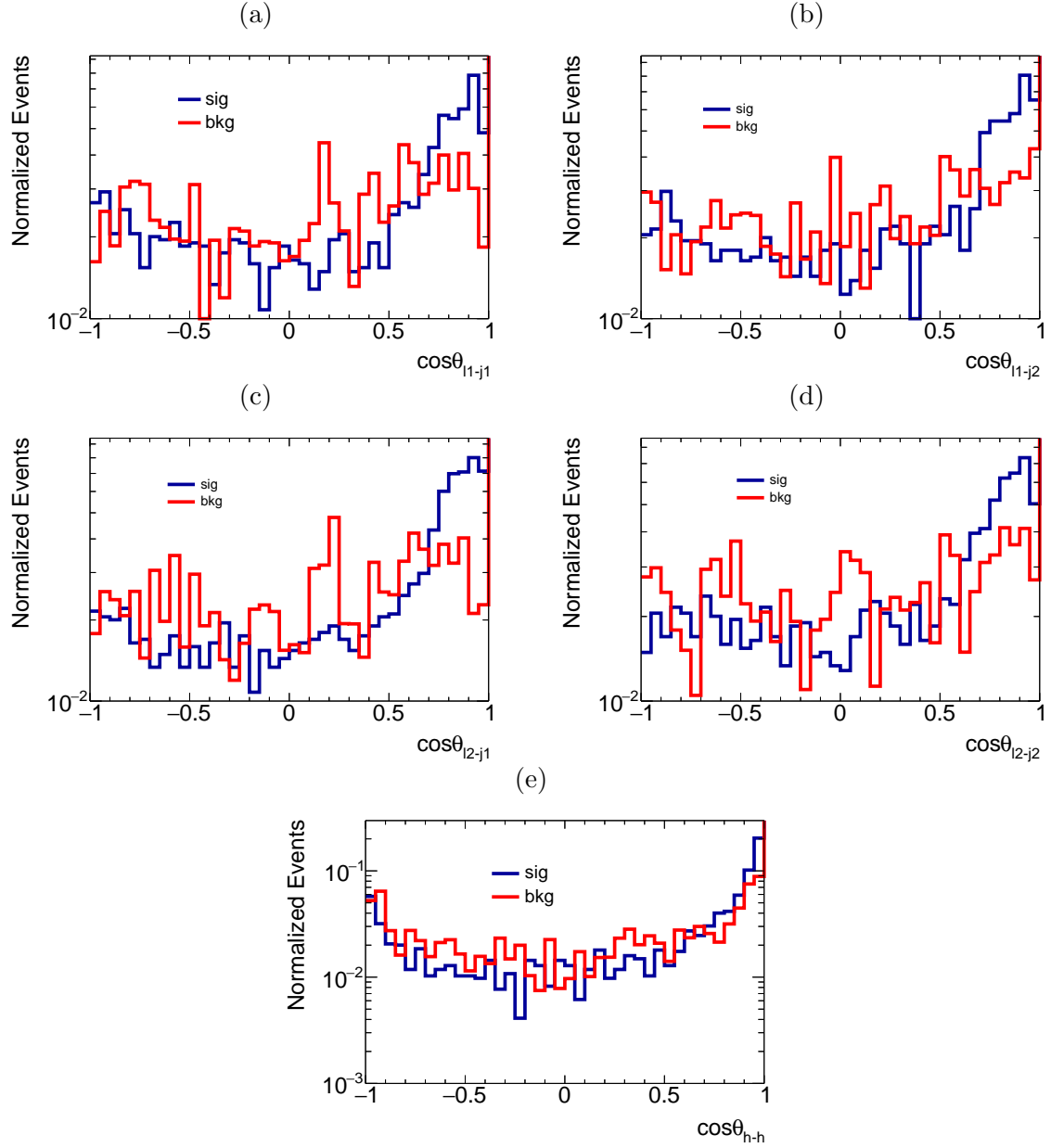


FIG. 6. Distributions in the (cosine of the) angle between (a)-(d) pairs of jets/leptons and (e) the two reconstructed light Higgs states for BP2 and backgrounds when  $\sqrt{s} = 14$  TeV and  $L = 300 \text{ fb}^{-1}$ .

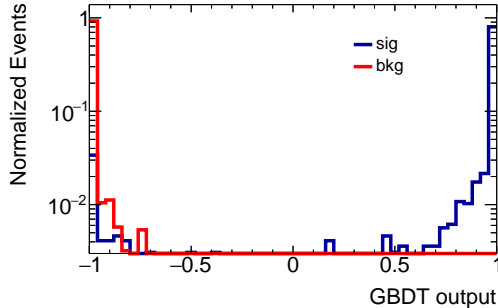


FIG. 7. GBDT output for the BP2 signal and backgrounds.

#### D. Significances

The process of searching for signals at high energy colliders using ML techniques is ultimately a statistical process. Therefore, to look for signals, a large number of events need to be collected and the kinematic variables of the final state particles need to be analyzed to find differences from the noise, but then a sizeable statistical sample needs to be collected, so has to have enough significance.

The latter is defined as  $\Sigma = \frac{N_S}{\sqrt{(N_S+N_B)}}$ , where  $N_S$  and  $N_B$  represent the number of events for the signal ( $S$ ) and (total) background ( $B$ ), respectively, at a given luminosity  $L$ .

To distinguish the signal from the background and to compute the ensuing significance, we apply six kinematic cuts on reconstructed observables and one MVA cut on GBDT outputs, the details of which are listed in Tab. VI, for BP2 and where  $L = 300 \text{ fb}^{-1}$  (with a CM energy  $\sqrt{s} = 14 \text{ TeV}$ , as previously mentioned). Following the acceptance cuts and upon applying the SS lepton criterion (the combination of which we refer to as pre-selection), we see a strong decrease of the initially huge SM backgrounds to the same order as the signal. For example, the initially troublesome  $Zjj$  is almost eliminated after this cut (thus, it is no longer included in this table). As Figs. 3–5 show, the kinematic selection is already quite efficient in separating signal from background, as – depending on the BP adopted –  $\Sigma$  values can reach values around 4, even with rather loose cuts. However, exploitation of the GBDT output more than double the final significance.

$L = 300 \text{ fb}^{-1}$	$S$	$t\bar{t}$	$t\bar{t}W^+W^-$	$t\bar{t}Z$	$t\bar{t}ZZ$	$Wtb$	$W^+W^-jj$	$ZZ$	$B$	$\Sigma$
Acceptance	262.6	$2.17 \times 10^6$	2.3	63	0.01	53751.6	119699	2790.8	$2.35 \times 10^6$	0.17
SS Leptons	131.5	5829.8	1.5	23.6	0.01	186.6	595.5	754.1	7391.1	1.52
$H_T \in [40, 200]$	129.9	1040.7	0.002	0.5	1e-05	40.4	127.9	379.1	1588.7	3.13
$M_{ll} \in [10, 80]$	129.8	905.8	0.001	0.5	1e-05	32.7	58.4	283.8	1281.1	3.45
$M_{h_1} \in [0, 150]$	128.7	809.4	0.001	0.4	1e-05	31.1	56.9	271.4	1169.2	3.57
$M_{h_2} \in [0, 150]$	126.9	732.3	0.001	0.4	8e-06	28	55.3	260.5	1076.5	3.66
$M_H \in [20, 230]$	96.9	269.8	0	0.2	5e-06	10.9	26.9	124.3	432	4.21
GBDT $\in [0.5, 1]$	92.6	0	0	0.02	0	0	3.2	7.1	10.2	9.13

TABLE VI. Response to our selection cuts for the signal (e.g., BP2) and background (separately and total) rates computed at  $\sqrt{s} = 14 \text{ TeV}$  for  $300 \text{ fb}^{-1}$ .

The results for all twelve BPs are shown in Tab. VII. With only the acceptance and SS lepton criteria, the significances can reach 3 for BP8, BP9 and BP11. After applying the kinematic cuts, the significances can be all larger than 3, except for BP3. Upon applying the MVA selection, significances can reach values as large as 14 for BP8. In short, it is clear that a judicious combination of pre-selection and kinematic cuts combined with ML analysis can result in large sensitivity to a significant expanse of 2HDM Type-X parameter space through our chosen signal.

$\Sigma$	BP1	BP2	BP3	BP4	BP5	BP6	BP7	BP8	BP9	BP10	BP11	BP12
after selecting SS leptons	1.17	1.52	1.06	1.40	2.04	1.79	1.71	3.89	3.15	2.86	3.18	2.58
after kinematic cuts, w/o GBDT	3.09	4.21	2.66	3.79	5.20	4.50	4.50	8.84	7.58	6.65	7.41	6.24
after kinematic cuts, w/ GBDT	7.65	9.13	6.27	8.61	10.41	9.45	9.34	14.32	12.91	11.25	12.14	10.74

TABLE VII. Significances following our different selections for all signals (BP1–BP12) at  $\sqrt{s} = 14 \text{ TeV}$  for  $300 \text{ fb}^{-1}$ .



## V. CONCLUSIONS

In this paper, we have investigated the feasibility of producing the SM-like Higgs boson ( $H$ ) in single mode at the LHC decaying via a resonance into two lighter neutral CP-even Higgs bosons ( $hh$ ) in the 2HDM Type-X. We have focused on  $4\tau$  final states, where two  $\tau$ 's decay leptonically to form a pair of SS leptons (electrons and/or muons) and two  $\tau$ 's decay hadronically, thus producing two jets. In order to motivate future studies of this signature, we have also proposed 12 BPs and processed these through a realistic MC analysis for a CM (proton-proton) energy  $\sqrt{s} = 14$  TeV and (integrated) luminosity  $L = 300 \text{ fb}^{-1}$ . Eventually, we have proven that this kind of signal could be found at the end of Run 3 of the LHC, following a dedicated selection based on kinematic and TMVA analysis.

## ACKNOWLEDGMENTS

SM is supported in part through the NExT Institute and the STFC Consolidated Grant ST/L000296/1. SS is supported in full through the NExT Institute and acknowledges the use of the IRIDIS High Performance Computing Facility, and associated support services, at the University of Southampton, in the completion of this work. YW's work is supported by the Natural Science Foundation of China Grant No. 12275143, the Inner Mongolia Science Foundation Grant No. 2020BS01013 and the Fundamental Research Funds for the Inner Mongolia Normal University Grant No. 2022JBQN080. QSY is supported by the Natural Science Foundation of China under the Grants No. 11875260 and No. 12275143.

- 
- [1] G. Aad *et al.* [ATLAS], Phys. Lett. B **716** (2012) 1, [arXiv:1207.7214 [hep-ex]].
  - [2] S. Chatrchyan *et al.* [CMS], Phys. Lett. B **716** (2012) 30, [arXiv:1207.7235 [hep-ex]].
  - [3] G. Aad *et al.* [ATLAS], Nature **607**, no.7917, 52-59 (2022) [erratum: Nature **612**, no.7941, E24 (2022)] doi:10.1038/s41586-022-04893-w [arXiv:2207.00092 [hep-ex]].
  - [4] A. Tumasyan *et al.* [CMS], Nature **607**, no.7917, 60-68 (2022) doi:10.1038/s41586-022-04892-x [arXiv:2207.00043 [hep-ex]].
  - [5] T. D. Lee, Phys. Rev. D **8** (1973), 1226-1239.
  - [6] N. G. Deshpande and E. Ma, Phys. Rev. D **18** (1978), 2574.

- [7] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher and J. P. Silva, Phys. Rept. **516** (2012), 1-102 [arXiv:1106.0034 [hep-ph]].
- [8] S. L. Glashow and S. Weinberg, Phys. Rev. D **15** (1977), 1958
- [9] M. Aoki, S. Kanemura, K. Tsumura and K. Yagyu, Phys. Rev. D **80**, 015017 (2009) doi:10.1103/PhysRevD.80.015017 [arXiv:0902.4665 [hep-ph]].
- [10] V. Khachatryan *et al.* [CMS], JHEP **01**, 079 (2016) doi:10.1007/JHEP01(2016)079 [arXiv:1510.06534 [hep-ex]].
- [11] V. Khachatryan *et al.* [CMS], JHEP **10**, 076 (2017) doi:10.1007/JHEP10(2017)076 [arXiv:1701.02032 [hep-ex]].
- [12] A. M. Sirunyan *et al.* [CMS], Phys. Lett. B **800**, 135087 (2020) doi:10.1016/j.physletb.2019.135087 [arXiv:1907.07235 [hep-ex]].
- [13] A. Arhrib, R. Benbrik, M. Krab, B. Manaut, S. Moretti, Y. Wang and Q. S. Yan, JHEP **10**, 073 (2021) doi:10.1007/JHEP10(2021)073 [arXiv:2106.13656 [hep-ph]].
- [14] Y. Wang, A. Arhrib, R. Benbrik, M. Krab, B. Manaut, S. Moretti and Q. S. Yan, JHEP **12**, 021 (2021) doi:10.1007/JHEP12(2021)021 [arXiv:2107.01451 [hep-ph]].
- [15] A. Arhrib, R. Benbrik, M. Krab, B. Manaut, S. Moretti, Y. Wang and Q. S. Yan, Symmetry **13**, no.12, 2319 (2021) doi:10.3390/sym13122319 [arXiv:2110.04823 [hep-ph]].
- [16] Z. Li, A. Arhrib, R. Benbrik, M. Krab, B. Manaut, S. Moretti, Y. Wang and Q. S. Yan, [arXiv:2305.05788 [hep-ph]].
- [17] D. Eriksson, J. Rathsman and O. Stål, Comput. Phys. Commun. **181** (2010), 189-205 [arXiv:0902.0851 [hep-ph]].
- [18] S. Kanemura, T. Kubota and E. Takasugi, Phys. Lett. B **313**, 155-160 (1993) doi:10.1016/0370-2693(93)91205-2 [arXiv:hep-ph/9303263 [hep-ph]].
- [19] A. G. Akeroyd, A. Arhrib and E. M. Naimi, Phys. Lett. B **490**, 119-124 (2000) doi:10.1016/S0370-2693(00)00962-X [arXiv:hep-ph/0006035 [hep-ph]].
- [20] A. Arhrib, [arXiv:hep-ph/0012353 [hep-ph]].
- [21] S. Kanemura, T. Kubota and E. Takasugi, Phys. Lett. B **313** (1993), 155-160 [arXiv:hep-ph/9303263 [hep-ph]].
- [22] A. G. Akeroyd, A. Arhrib and E. M. Naimi, Phys. Lett. B **490** (2000), 119-124 [arXiv:hep-ph/0006035 [hep-ph]]. A. Arhrib, [arXiv:hep-ph/0012353 [hep-ph]].
- [23] M. E. Peskin and T. Takeuchi, Phys. Rev. Lett. **65** (1990), 964-967.

- [24] M. E. Peskin and T. Takeuchi, *Phys. Rev. D* **46** (1992), 381-409.
- [25] R. L. Workman *et al.* [Particle Data Group], *PTEP* **2022**, 083C01 (2022)  
doi:10.1093/ptep/ptac097
- [26] P. Bechtle, D. Dercks, S. Heinemeyer, T. Klingl, T. Stefaniak, G. Weiglein and J. Wittbrodt, *Eur. Phys. J. C* **80** (2020) no.12, 1211 [arXiv:2006.06007 [hep-ph]].
- [27] P. Bechtle, S. Heinemeyer, T. Klingl, T. Stefaniak, G. Weiglein and J. Wittbrodt, *Eur. Phys. J. C* **81** (2021) no.2, 145 [arXiv:2012.09197 [hep-ph]].
- [28] F. Mahmoudi, *Comput. Phys. Commun.* **180** (2009), 1579-1613 [arXiv:0808.3144 [hep-ph]].
- [29] Y. S. Amhis *et al.* [HFLAV], *Phys. Rev. D* **107**, no.5, 052008 (2023)  
doi:10.1103/PhysRevD.107.052008 [arXiv:2206.07501 [hep-ex]].
- [30] R. Aaij *et al.* [LHCb], *Phys. Rev. D* **105**, no.1, 012010 (2022)  
doi:10.1103/PhysRevD.105.012010 [arXiv:2108.09283 [hep-ex]].
- [31] R. Aaij *et al.* [LHCb], *Phys. Rev. Lett.* **128**, no.4, 041801 (2022)  
doi:10.1103/PhysRevLett.128.041801 [arXiv:2108.09284 [hep-ex]].
- [32] A. Tumasyan *et al.* [CMS], *Phys. Lett. B* **842**, 137955 (2023)  
doi:10.1016/j.physletb.2023.137955 [arXiv:2212.10311 [hep-ex]].
- [33] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli and M. Zaro, *JHEP* **07** (2014), 079 [arXiv:1405.0301 [hep-ph]].
- [34] T. Sjostrand, S. Mrenna and P. Z. Skands, *JHEP* **05** (2006), 026 [arXiv:hep-ph/0603175 [hep-ph]].
- [35] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen and P. Z. Skands, *Comput. Phys. Commun.* **191** (2015), 159-177 [arXiv:1410.3012 [hep-ph]].
- [36] J. de Favereau *et al.* [DELPHES 3], *JHEP* **02** (2014), 057 [arXiv:1307.6346 [hep-ex]].
- [37] M. Cacciari, G. P. Salam and G. Soyez, *JHEP* **04**, 063 (2008) doi:10.1088/1126-6708/2008/04/063 [arXiv:0802.1189 [hep-ph]].
- [38] M. Cacciari, G. P. Salam and G. Soyez, *Eur. Phys. J. C* **72**, 1896 (2012)  
doi:10.1140/epjc/s10052-012-1896-2 [arXiv:1111.6097 [hep-ph]].
- [39] A. Arhrib, S. Moretti, S. Semmlali, C. H. Shepherd-Themistocleous, Y. Wang and Q. S. Yan, [arXiv:2310.02736 [hep-ph]].
- [40] J. Therhaag, *PoS ICHEP2010* (2010), 510