Probing heavy charged Higgs boson at the LHC

Monoranjan Guchait[*] and Aravind H. Vijay[†]

*Department of High Energy Physics, Tata Institute of Fundamental Research,
Homi Bhabha Road, Mumbai-400005, India*

The signature of a heavier charged Higgs boson, much above the top quark mass, is investigated at the LHC Run 2 experiments, following its decay mode via the top and bottom quark, focusing on both the hadronic and semileptonic signal final states. The generic two Higgs doublet model framework is considered with a special emphasis on the supersymmetry motivated type II model. The signal is found to be heavily affected by huge irreducible backgrounds due to the top quark pair production and QCD events. The jet substructure technique is used to tag moderately boosted top jets in order to reconstruct the charged Higgs mass. A simple cut-based analysis is performed by optimizing various kinematic selections, and the signal sensitivity is found to be reasonable for only the lower range of charged Higgs masses, corresponding to 3000 fb^{-1} integrated luminosity. However, employing the multivariate analysis technique, a remarkable improvement in signal sensitivity is achieved. We find that the charged Higgs signal for the mass range of about 300–600 GeV is observable with 1000 fb^{-1} luminosity. However, for high luminosity, $\mathcal{L} = 3000$ fb^{-1}, the discovery potential can be extended to 700–800 GeV.

I. INTRODUCTION

The recent discovery of the 125 GeV Higgs boson [1,2] at the CERN Large Hadron Collider (LHC) has provided the last missing piece of the Standard Model (SM) and opened up a new window to explore physics beyond the standard model (BSM). Although the current precision measurements of various properties of the Higgs boson, in particular the couplings with fermions and gauge bosons, indicate that it is indeed the candidate for the SM Higgs [3], nonetheless, it does not rule out many BSM scenarios. Among the plethora of BSM candidates, the supersymmetry-based models, such as the minimal supersymmetric standard model (MSSM), which is the most popular and well-studied BSM scenario, provide elegant solutions to some of the shortcomings of the SM and predict a rich and diverse phenomenology that is testable directly in colliders.

Recall that the MSSM requires at least two Higgs doublets to make the theory anomaly free and also to generate the masses of up- and down-type fermions. The theories with an extended Higgs sector predict more Higgs boson neutral and charged states. In general, a two Higgs doublet model (2HDM) consisting of an extra SU(2)

Higgs doublet added with the SM Higgs doublet is well motivated and consistent with the Higgs discovery. In fact, the 2HDM can be interpreted as the effective theory at low energy of many BSM theories with UV completion. For example, the Higgs sector in the supersymmetric model may appear as a simple 2HDM (type II), if the masses of all sparticles decouple at a very high scale. Generally, 2HDM is classified into four categories, types I, II, III, and IV, depending on the nature of Yukawa couplings, subject to Z_2 symmetry in order to avoid flavor changing neutral currents (for more details about 2HDM, see Refs. [4,5]). In all classes of the 2HDM scenario, there exist five physical Higgs boson states, two CP even (h, H, with the assumption $m_h < m_H$); one CP odd (A); and two charged Higgs bosons ($H^{\pm}$). The lightest CP even Higgs h can be interpreted as the SM-like Higgs boson in the decoupling limit, where the other states turn out to be very heavy, much above the electroweak scale [6]. However, some other studies also show that CP even Higgs states may behave as SM-like with mass 125 GeV in the alignment limit even without decoupling [7–10]. The presence of extra physical Higgs boson states along with the SM-like Higgs is one of the characteristics of BSM. Needless to say, discovery of an extra Higgs boson certainly confirms the existence of BSM. Therefore, looking for these additional Higgs bosons in various decay channels over a wide range of masses is a top priority in the current LHC experiment.

In this context, searching for the charged Higgs boson signal is unique, since discovery of it clearly and unambiguously confirms the presence of BSM. Therefore, the

[*]guchait@tifr.res.in
[†]aravind.vijay@tifr.res.in

study of the charged Higgs boson has received special attention both phenomenologically and experimentally. For the lower mass range, less than the top quark mass, $m_{H^\pm} < m_t$, the phenomenology of the charged Higgs boson is well studied and also experimentally probed thoroughly in many of its decay channels. However, detection of the charged Higgs boson for the heavier mass range, greater than the top quark mass ($m_{H^\pm} \gg m_t$), is found to be very challenging due to huge contamination by the irreducible SM backgrounds. In this current study, we attempt to find the discovery potential of the charged Higgs boson for this heavier mass range ($m_{H^\pm} \gg m_t$). The study is carried out within the framework of the generic 2HDM with an emphasis on type II 2HDM motivated by supersymmetry. The charged Higgs boson couplings with fermions are strongly dependent on $\tan\beta$, and hence the production and subsequent decays of the charged Higgs boson are sensitive to $\tan\beta$. In hadron colliders, in the lower mass range ($m_{H^\pm} < m_t$), the charged Higgs boson is produced via a pair production of the top quark, $p\bar{p}/pp \rightarrow t\bar{t}$, following the decay $t \rightarrow H^+b$. For the intermediate and heavier mass range, it is mainly produced directly in association with a top quark (and also a b quark) [11]. Furthermore, the charged Higgs boson can be produced in SUSY cascade decays via heavier chargino and neutralino production in gluino and squark decays [12,13].

So far, nonobservation of any charged Higgs signal events in direct searches constrains its production and decay in a model independent way, which in turn can be translated to exclude the relevant parameter space, in particular $\tan\beta$ and $m_{H^\pm}$, for a given model framework. For example, in the past, direct searches at LEP [14] and Tevatron [15] experiments excluded the lower mass range of $m_{H^\pm}$ in terms of $\tan\beta$. At the LHC Run 1 experiments with $\sqrt{s} = 7$ and 8 TeV data, a lighter charged Higgs boson was probed in the decay channels $\tau\nu$ [16,17], cs [18,19], and also cb [20], while at Run 2 with $\sqrt{s} = 13$ TeV energy, mainly the decay modes $\tau\nu$ [21,22] and tb [23] were considered to probe it up to ~ 1 TeV mass. The absence of any signal event in $H^+ \rightarrow \tau^+\nu_\tau$ decay modes in CMS at 13 TeV energy with an integrated luminosity 12.9 fb^{-1} leads to an exclusion of the cross section times the respective branching ratio for the mass range 180 GeV $< m_{H^\pm} < 3$ TeV, whereas limits on the Br($t \rightarrow H^+b$) $\times$ Br($H^+ \rightarrow \tau^+\nu_\tau$) are set for the range 80 GeV $< m_{H^\pm} <$ 160 GeV [21]. Eventually, these exclusion limits rule out $m_{H^\pm} \sim 90$–160 GeV, corresponding to the entire range of $\tan\beta$ up to 60 in the context of MSSM with the $m_h^{\mathrm{mod}+}$ scenario [24], except for a hole around $m_{H^\pm} \sim 150$–160, and $\tan\beta \sim 10$. Similar results were also published from ATLAS [22] at $\sqrt{s} = 13$ TeV. The searches in the $H^+ \rightarrow t\bar{b}$ decay channel for a heavier mass range carried out by ATLAS at $\sqrt{s} = 13$ TeV and $\mathcal{L} = 13.2$ fb^{-1} excluded $m_{H^\pm} \sim 300$–900 GeV for a very low $\tan\beta(\sim 0.5$–1.7)

region [23], whereas for high values of $\tan\beta > 44(60)$, $m_{H^\pm} \sim 300(366)$ GeV were excluded. Note that this decay channel was also probed at $\sqrt{s} = 8$ TeV by ATLAS, including the s-channel charged Higgs production, and exclusion was presented for the cross section times Br($H^+ \rightarrow t\bar{b}$) [25]. However, these limits were found to be very weak in comparison to the predictions from $H^+ \rightarrow \tau^+\nu_\tau$ searches [16,22]. Remarkably, the most stringent constraints on the charged Higgs sector in the context of a SUSY motivated type II model are predicted indirectly by the neutral Higgs boson searches, $pp \rightarrow h, H, A \rightarrow \tau\bar{\tau}$ at the LHC [26]. It can be attributed to the fact that the neutral Higgs boson couplings with tau leptons very strongly depend on $\tan\beta$, in particular for higher values of it. The exclusion region predicted by these neutral Higgs boson searches implies a limit on $\tan\beta > 6$ for $m_A < 250$ GeV, whereas higher values of $\tan\beta$ (>20) are completely ruled out up to 60; for $m_A \sim m_{H^\pm} \sim 1000$ GeV, m_A is the mass of the pseudoscalar Higgs, related with the charged Higgs mass as $m_{H^\pm}^2 = m_W^2 + m_A^2$ in the SUSY model (like the type II model). In addition to these direct limits, the charged Higgs sector is also constrained by flavor physics data. Strong contribution via loops to the Br of rare decay modes of the B meson makes it very sensitive to flavor physics observables. Measurements of these Br by B-factories, and also at the LHC and LHCb, put a very strong limit on the charged Higgs sector. More details about these latest constraints in the framework of 2HDM can be found in the recent review of Ref. [27], and references therein.

In the phenomenological side, there have been numerous studies on exploring the $H^\pm$ signal in various decay channels in the context of the MSSM Higgs sector [28–38], as well as in 2HDM framework [30,39–41] using various interesting techniques. More details about charged Higgs phenomenology can be found in Ref. [42]. It is worth mentioning here the use of τ lepton polarization in its one- and three-prong decay for $H^+ \rightarrow \tau^+\nu_\tau$, which is found to be very useful in extracting the signal suppressing the $t\bar{t}$ and QCD background [43–45]. The signal of the charged Higgs boson is also probed in the subdominant production channels $H^\pm W^\mp$ [46] and H^+H^- [47,48]. The discovery potential of the charged Higgs for a heavier mass ($m_{H^\pm} > m_t$) range with its dominant decay mode $H^+ \rightarrow t\bar{b}$ has been investigated by many authors in the framework of SUSY models [49–52]. For instance, in Ref. [52], the authors used triple and quadruple b-tagging in order to suppress the SM background, which also costs the signal significantly as well. Consequently, for the heavier mass range, it is found to be very hard to achieve a reasonable signal sensitivity, due to the large $t\bar{t}$ and QCD backgrounds. A recent study [53] reported about the detection prospect of the charged Higgs signal for a heavier mass $\gtrsim 1$ TeV by applying a jet substructure technique to tag the top quarks from the charged

Higgs decay in the framework of 2HDM. The authors predicted reasonable sensitivities of the charged Higgs signal around the mass of 1 TeV and found it difficult to probe for the intermediate mass range. The jet substructure technique has also been used to look for the heavy charged Higgs boson signal in the decay channel $H^{\pm} \rightarrow W^{\pm}A$ for lighter A boson states [54,55]. In this current study, we explore the detection prospect of the charged Higgs boson for the intermediate to heavier mass range, 300–1000 GeV, considering the decay mode $H^{+} \rightarrow t\bar{b}$ with the hadronic and leptonic final state. For the heavier mass of $H^{\pm}$, the top quark from its decay is expected to be boosted. We try to exploit this feature by employing the technique of jet substructures to reconstruct the top quark and subsequently the charged Higgs boson. This method helps to avoid the combinatorial problem while reconstructing the top quark simply by combining the hard jets. In this study, first we attempt to obtain the signal sensitivity using a cut-based analysis. Then we try to improve the sensitivity employing the multivariate (MVA) analysis. Performing a detailed analysis in the MVA framework, we achieve a remarkable improvement in signal sensitivity, and the results are presented for three integrated luminosity options $\mathcal{L} = 300$, 1000, and 3000 fb^{-1}. Finally, for the sake of completeness, signal sensitivities are predicted for all classes of 2HDM corresponding to a few benchmark parameter spaces.

We present this study as follows. After briefly describing the 2HDM in Sec. II, we discuss the charged Higgs production in Sec. III. In Sec. IV, the signal and backgrounds are discussed, and subsequently, details of the simulation are presented in Sec. IV B, with a brief description about top tagging in Sec. IV A. The results based on cut and count analysis are discussed in Sec. IV C, while in Sec. V the results based on MVA analysis are presented. Finally we summarize in Sec. VI.

II. TWO HIGGS DOUBLET MODEL

In the context of our present study, it is instructive to discuss the 2HDM very briefly. In this model, an extra SU(2) Higgs doublet is added with the SM Higgs doublet. The most general 2HDM potential consisting of two doublets ϕ_1 and ϕ_2 with hypercharge Y $= +1$ is given by [4,5]

$$V = m_{11}^2 \phi_1^2 + m_{22}^2 \phi_2^2 - m_{12}^2 (\phi_1^{\dagger}\phi_2 + \phi_1\phi_2^{\dagger})$$
$$+ \frac{\lambda_1}{2}(\phi_1^{\dagger}\phi_1)^2 + \frac{\lambda_2}{2}(\phi_2^{\dagger}\phi_2)^2 + \lambda_3(\phi_1^{\dagger}\phi_1)(\phi_2^{\dagger}\phi_2)$$
$$+ \lambda_4(\phi_1^{\dagger}\phi_2)(\phi_2^{\dagger}\phi_1) + \frac{\lambda_5}{2}[(\phi_1^{\dagger}\phi_2)^2 + (\phi_2^{\dagger}\phi_1)^2)]. \quad (2.1)$$

For simplification, all the free parameters are assumed to be real to conserve CP property, and the discrete Z_2 symmetry, $\phi_1 \rightarrow -\phi_1$ and $\phi_2 \rightarrow +\phi_2$, is imposed to suppress flavor-changing neutral current at the tree level. The Z_2 symmetry

is softly broken by the terms proportional to m_{12}. The minimum of the potential V is ensured by two vacuum expectation values (VEVs), which break the symmetry down to U(1)$_{\text{em}}$ symmetry,

$$\langle \phi_1 \rangle = \begin{bmatrix} 0 \\ \frac{v_1}{\sqrt{2}} \end{bmatrix}, \qquad \langle \phi_2 \rangle = \begin{bmatrix} 0 \\ \frac{v_2}{\sqrt{2}} \end{bmatrix}, \qquad (2.2)$$

where v_1 and v_2 are two VEVs corresponding to neutral components of ϕ_1 and ϕ_2, respectively, with $v = \sqrt{v_1^2 + v_2^2}$. The ratio of two VEVs defined to be $\tan\beta = \frac{v_2}{v_1}$ is considered as one of the free parameters of the model. Expanding the doublets around the minimum of the potential, the Higgs fields can be given by [4,5]

$$\phi_1 = \begin{pmatrix} \phi_1^+ \\ \frac{1}{\sqrt{2}}(v\cos\beta + \phi_1^0) \end{pmatrix},$$
$$\phi_2 = \begin{pmatrix} \phi_2^+ \\ \frac{1}{\sqrt{2}}(v\sin\beta + \phi_2^0) \end{pmatrix}. \quad (2.3)$$

As already mentioned in the previous section, after symmetry breaking, the potential predicts five physical Higgs boson states: two neutral CP even states, h and H ($m_h < m_H$); one neutral CP odd state A; and two charged states $H^{\pm}$. The physical charged state and CP odd neutral states are expressed as

$$H^{\pm} = -\phi_1^{\pm}\sin\beta + \phi_2^{\pm}\cos\beta, \quad (2.4)$$
$$A = \sqrt{2}(-\text{Im}\phi_1^0\sin\beta + \text{Im}\phi_2^0\cos\beta). \quad (2.5)$$

The two CP even neutral weak states mix through an angle α providing two mass eigenstates, h and H. The input parameters present in the potential V can be reexpressed in terms of physical masses and other parameters, such as

$$m_h, \quad m_H, \quad m_A, \quad m_{H^{\pm}}, \quad \tan\beta, \quad \sin(\beta - \alpha), \quad v, \quad m_{12}^2. \quad (2.6)$$

Note that v is set to be at the electroweak scale ($= 246$ GeV), and one of the CP even Higgs bosons can be interpreted as the recently discovered Higgs boson of mass 125 GeV under certain scenarios of the model, which are already mentioned in the earlier section [6–10]. The topic of our interest in this current study is to look for the charged Higgs signal; hence, we focus only this sector of 2HDM. In the generic 2HDM model, the Yukawa couplings of the charged Higgs with fermions are given by [4,5]

TABLE I. λ's in charged Higgs couplings with fermions [Eq. (2.7)] for all four types of 2HDM.

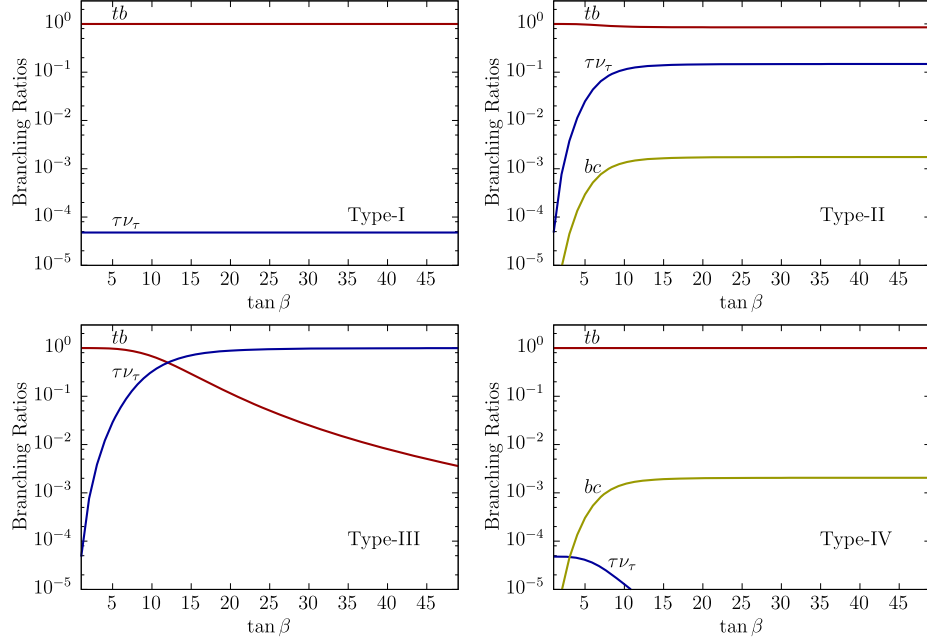|	Type I	Type II	Type III	Type IV
λ_u	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
λ_d	$\cot\beta$	$-\tan\beta$	$\cot\beta$	$-\tan\beta$
λ_{ℓ}	$\cot\beta$	$-\tan\beta$	$-\tan\beta$	$\cot\beta$

FIG. 1. Charged Higgs branching ratios for four classes of 2HDM, setting $m_H = m_A = m_{H^\pm} = 500$ GeV and $\sin(\beta - \alpha) = 1$.

$$\mathcal{L}_{H^\pm} = -H^+ \left[\frac{\sqrt{2}V_{ud}}{v} \bar{u}(m_u \lambda_u P_L + m_d \lambda_d P_R)d \right.$$

$$\left. + \frac{\sqrt{2}m_\ell}{v} \lambda_\ell \bar{\nu}_\ell \ell_R \right] + \text{H.c.,} \qquad (2.7)$$

where V_{ud} is the Cabibbo–Kobayashi–Maskawa matrix elements, and the couplings λ represent either $\tan\beta$ or $\cot\beta$, depending on the assignments of Z_2 charges to right-handed fermions, which finally define the four types of 2HDM. Table I presents λ's corresponding to the four types of the 2HDM model. As shown in the type I model, the couplings of the charged Higgs with fermions are heavily suppressed for $\tan\beta \gg 1$, in the same way as the type III model, except the coupling with leptons, which is enhanced, making it lepton specific. In the type II model, which is same as the supersymmetric Higgs sector, couplings are favored with u-type quarks for the low $\tan\beta$ case, whereas for d-type quarks and leptons, high values of $\tan\beta$ are preferred. The type IV model is found to be leptophobic for the high $\tan\beta$ scenario, but the couplings with quarks are the same for both the type II and type IV models. Consequently, the charged Higgs decay Br to fermions is very much $\tan\beta$ dependent. The decay channels of the charged Higgs to $\tau\bar{\nu}_\tau$ or $\bar{t}b$ channels are very much sensitive to $\tan\beta$ once they are kinematically allowed. The charged Higgs Br computed by HDECAY [56,57] is demonstrated for various values of $\tan\beta$, setting $m_{H^\pm} = 500$ GeV, corresponding to the four types of 2HDM in Fig. 1. The input parameters are set as $m_h = 125$ GeV, $m_H = m_A = m_{H^\pm}$, and $\sin(\beta - \alpha) = 1$, like the MSSM scenario [27] with a decoupling limit. In the type I model, due to the $\cot\beta$ dependence of the coupling, the $\text{Br}(H^+ \to \tau^+ \nu_\tau)$ is

suppressed by m_τ^2 / m_t^2 over $\text{Br}(H^+ \to t\bar{b})$, leading almost 100% Br to $\bar{t}b$ mode. The dominant decay mode of the charged Higgs in the type II model, as expected, is in the $\bar{t}b$ channel, following the subdominant $\bar{\tau}\nu_\tau$ channel with Br $\sim 10\%–15\%$, followed by other suppressed modes, such as $H^+ \to \bar{b}c, c\bar{s}$. However, in the case of the type III model, which is lepton specific, the charged Higgs decays to the $\bar{\tau}\nu_\tau$ mode dominantly, except in the lower region of $\tan\beta \sim 1$–12, where the $t\bar{b}$ mode becomes important. On the contrary, the $\bar{\tau}\nu_\tau$ mode gets suppressed in the type IV model, because of $\cot\beta$ dependence, and the $\bar{t}b$ channel takes over. It is to be noted that the pattern of these Br's is expected to be different in the presence of the $H^\pm \to W^\pm \phi$ $(\phi = h, H, A)$ mode, whose decay width is proportional to $\cos(\beta - \alpha)$, leading it to be dominant ($\sim 100\%$) for the choice of $\sin(\beta - \alpha) = 0$.[1] Interestingly, in the case of the SUSY motivated Higgs sector, i.e., in the type II model, if kinematically allowed, the charged Higgs can decay also to a chargino and neutralino pair, $H^\pm \to \tilde{\chi}_i^\pm \tilde{\chi}_j^0$; (i:1-2, j:1-4), which may be dominant for a Higgsino-like scenario [58]. As pointed out earlier, the charged Higgs sector is severely constrained by flavor physics data in addition to the direct searches; details can be found in reviews [27,42].

III. CHARGED HIGGS PRODUCTION

In the intermediate to heavier mass range ($m_{H^\pm} \gtrsim m_t$), the charged Higgs is produced directly in proton-proton collisions via the process

---

[1]This scenario is equivalent to the MSSM inverted scenario, where H is SM-like and $m_H = 125$ [27].

TABLE II.   Charged Higgs boson production cross sections (in fb) in the 4FS and 5FS schemes, at $\sqrt{s} = 13$ TeV and $\tan\beta = 30$ in the type II model. The last row presents cross sections for type I. μ and $m_b(\mu)$ represent the QCD scales and mass of the bottom quark, respectively, at the scale μ.

$m_{H^\pm}$ (GeV) →	300	500	600	800	1000
μ (GeV)	236.5	336.5	386.5	486.5	586.5
$m_b(\mu)$ (GeV)	2.64	2.58	2.56	2.51	2.48
$\sigma(pp \to tbH^\pm)$ (4FS) LO	290.2	60.4	30.6	9.0	3.1
$\sigma(pp \to tbH^\pm)$ (4FS) NLO	359.4	73.3	39.9	11.4	4.1
$\sigma(pp \to tH^\pm)$ (5FS) LO	581.3	126.0	64.8	19.7	6.9
$\sigma(pp \to tH^\pm)$ (5FS) NLO	748.6	166.2	86.1	26.5	9.3
Matched (NLO)	625.3	140.9	74.1	22.8	8.1
Matched (NLO) (type I)	2.9	0.66	0.35	0.11	3.9×10^{-2}

$$pp \to tH^- + X. \qquad (3.1)$$

At the parton level, the production mechanism is initiated via two subprocesses,

$$gg, q\bar{q} \;\; \to t\bar{b}H^-(4\text{FS}) \qquad gb \to tH^-(5\text{FS}) \qquad (3.2)$$

in the four-flavor scheme (4FS) and five-flavor scheme (5FS) at the leading order (LO), respectively. In fact, the process in the 4FS is part of the next to leading order (NLO) QCD correction to the 5FS scheme mechanism. The total NLO QCD effects on the inclusive $H^\pm$ production are essentially the NLO correction to the process $gb \to tH^-$ plus the total contribution due to the tree-level processes [59]. In the 5FS, the NLO QCD corrections have been known for sometime in the literature [60–63], and also very recently approximate next to next to leading order (NNLO) calculations were also published [64]. The total theoretical uncertainty in $H^\pm$ production in association with the top quark (5FS) is found to be of the range 15%–20% [65]. In the 4FS, the final state bottom quark which originates due to the hard scattering is assumed to have a nonzero mass, whereas in the 5FS, the b quark is treated as massless, being part of the parton flux. In the 4FS, the corresponding NLO correction is estimated to be around 20% for the lower range of $m_{H^\pm}$, and it goes up a little for higher masses [65].

At finite order, the cross section in the 4FS does not match with the 5FS, as expected, due to different ways of treating perturbative calculation. However, it is expected that the results will match within the respective uncertainties once all orders are taken into account in perturbation. In order to obtain a precise estimation of the charged Higgs production cross section, one needs to combine the four- and five-flavor scheme predictions appropriately. This combination is performed following the prescription, the so-called Santander matching [66]. In the IR limit ($\frac{m_{H^\pm}}{m_b} \to 1$), the cross sections obtained from the 4FS and 5FS match nicely. The main difference between the 4FS and 5FS occurs because of the presence of a large logarithm, which arises due to the splitting of the incoming gluon into two nearly collinear b quarks [67]. Thus, the calculated cross sections using two schemes should be combined in such a manner that

such logarithmic effects are taken into account appropriately. The prescription to match these cross sections computed in two schemes is given by [11,65]

$$\sigma = \frac{\sigma^{4\text{FS}} + w\sigma^{5\text{FS}}}{1 + w} \quad \text{with} \quad w = \ln\frac{m_{H^\pm}}{m_b} - 2. \qquad (3.3)$$

Similarly, the theoretical uncertainties are combined as

$$\Delta\sigma = \frac{\Delta\sigma^{4\text{FS}} + w\Delta\sigma^{5\text{FS}}}{1 + w} \quad \text{with} \quad w = \ln\frac{m_{H^\pm}}{m_b} - 2. \qquad (3.4)$$

With this matching methodology, the overall theoretical uncertainty of the combined NLO cross section is found to be around 10%, whereas the individual 4FS and 5FS cross sections at NLO are in reasonable agreement within $\sim$20% from the central value [11,65]. The production cross section and the corresponding uncertainty are very sensitive to $\tan\beta$, owing to the dependence of the Yukawa coupling on it. The scale of uncertainty decreases with the decrease of $\tan\beta$ through the correction of the bottom Yukawa coupling, which is proportional to $\tan\beta$. We first estimate the charged Higgs boson production in type II 2HDM motivated by the SUSY providing inputs $\tan\beta$ and $m_{H^\pm}$, and then we predict the corresponding cross sections for other classes of 2HDM (types I, III, and IV) simply by appropriately rescaling the couplings. It is to be noted that in the MSSM, the NLO QCD corrections may involve additional loop contributions from gluinos and squarks, which also depend on $\tan\beta$. This extra contribution can be absorbed through the rescaling of the NLO QCD prediction of the bottom Yukawa coupling [68]. The total cross section primarily governed by the $tbH^\pm$ coupling is found to be minimum in strength for $\tan\beta \approx 7$–8. In Table II, the charged Higgs boson production cross sections for both schemes and the final matched values are presented for a few representative choices for $m_{H^\pm}$ and $\tan\beta = 30$ in the type II model. The cross sections are computed both at LO and NLO, using MadGraph5-2.6.1 [69], with the FeynRules [70] model file uploaded by the authors of [71]. We notice that for $\tan\beta = 3$, the cross sections go down by a factor of $\sim\frac{1}{2}$ in the type II model. In calculating these cross sections, factorization and
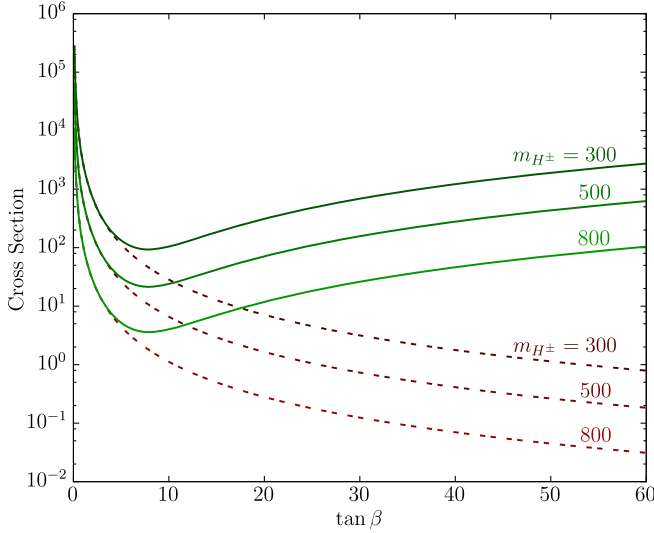
FIG. 2.   Matched charged Higgs production cross section (fb) at $\sqrt{s} = 13$ TeV for three different masses (in GeV) in type II (green solid lines) and type I (red dashed lines) 2HDM.

renormalization scales are set as $\mu^2 = (\frac{m_{H^\pm}+m_t}{2})^2$, as shown in the first row, along with the value of the running b-quark mass [72]. The variation of the cross sections is found to be within a range from $\mathcal{O}(100)$ to $\mathcal{O}(1)$ fb, corresponding to the range $m_{H^\pm} \sim 300$–1000 GeV.

In the type I model (see Table I), the charged Higgs boson couplings with the top and bottom quark go by $\sim (m_b + m_t)\cot\beta$. The cross sections in the type I model can be obtained from the values corresponding to the type II model by simply rescaling the Yukawa couplings [11,65,71]. The total cross section can be parametrized by $\sigma_{H^\pm}^{\text{Type II}} \propto g_t^2 \sigma_t \cot^2\beta + g_b \sigma_b \tan^2\beta + g_t g_b \sigma_{tb}$, where g_t and g_b are the part of the Yukawa couplings proportional to the top and bottom quark masses, respectively. Evaluating the contributions by setting $m_t = 0$ (i.e., $g_t = 0$) and $m_b = 0$ (i.e., $g_b = 0$), σ_b, σ_t, and σ_{bt} can be obtained. Thus, the cross sections in the type I model can be estimated by rescaling each contribution by $\cot\beta$. This prescription works to all orders in QCD, but not appropriate to all orders in the electroweak corrections [11]. The cross sections for both type I and II 2HDM are presented in Fig. 2 for various values of $\tan\beta$ and three choices of $m_{H^\pm} = 300$, 500, and 800 GeV. Clearly, as expected, the cross sections in the type I model are suppressed over the type II model by approximately $\sim \tan^2\beta$, for $\tan\beta \gg 1$. The cross sections in the type III (type IV) model are the same as the type I (type II) model due to the identical Yukawa coupling structures with quarks. A dip is observed for type II model around $\tan\beta \sim 7$–8, unlike type I, which can be understood from the respective coupling's dependence on $\tan\beta$ or $\cot\beta$.

IV. SIGNAL AND BACKGROUND

As mentioned before, in this current study, the signature of the charged Higgs boson is explored with its decay

mode, $H^+ \to t\bar{b}$. The Br($H^+ \to t\bar{b}$) is almost dominant, more than 70% for a wide range of $\tan\beta$, and for all classes of 2HDM as shown in Fig. 1, except for the type III model which is lepton specific. The signal is simulated considering $H^\pm$ production mechanisms [Eq. (3.2)], and eventually the final results are obtained by combining them following the recipe, given in Eq. (3.3).

The resulting signal final state consists of multiple top quarks via the following processes:

$$\text{5FS}: \quad gb \to tH^- \to t\bar{t}b$$
$$\text{4FS}: \quad gg \to t\bar{b}H^- \to t\bar{b}\,\bar{t}\,b. \qquad (4.1)$$

Both leptonic and hadronic final states are considered following the semileptonic and hadronic decays of the top quarks, respectively. Note that the final states consist of multiple b quarks, a characteristic of the heavier charged Higgs signal for the $H^+ \to t\bar{b}$ decay channel [51,52]. The top quark originating from $H^\pm$ decay is tagged in its hadronic mode and, combined with the appropriately identified b-jet, the charged Higgs mass is reconstructed. Tagging of the top quark is performed by implementing the powerful jet substructure analysis [73], which is postponed for discussion in the next section. In the case of a pure hadronic signal final state, the associated top quark is also identified through kinematic reconstruction in order to make the signal more robust. In addition to the reconstruction of top quarks, we exploit the presence of extra hard b-jets in the final state in order to separate out backgrounds. Therefore, we collect the charged Higgs signal final states into two categories:

$$(a): H^\pm_{\text{reco}} + t_{\text{reco}} + n_b(\geq 1)b - \text{jet}$$
$$(b): H^\pm_{\text{reco}} + n_\ell(\geq 1) + n_b(\geq 1)b - \text{jet}, \qquad (4.2)$$

where $H^\pm_{\text{reco}}$ and t_{reco} represent the reconstructed charged Higgs and top quark, and n_ℓ and n_b are the number of leptons and b-jets, respectively, and required to be at least one. The main dominant source of irreducible SM backgrounds is due to $t\bar{t}$ and inclusive hard QCD jet production. However, in both cases, extra b-jets may arise via gluon splitting in the initial state radiation. The QCD jet production becomes the dominant source of irreducible background, in particular corresponding to the hadronic signal final state, due to the non-negligible mistagging probability of a hard jet as a top jet. Moreover, the process $t\bar{t}g$ which predominantly produces the final state $t\bar{t}b\bar{b}$ is also taken into account in our background estimation. Before discussing the signal and background simulation strategy, we discuss briefly the top tagging methodology used in our simulation.

A. Top tagging

The top tagged jets are the essential components of our considered signal events. It has been pointed out earlier that
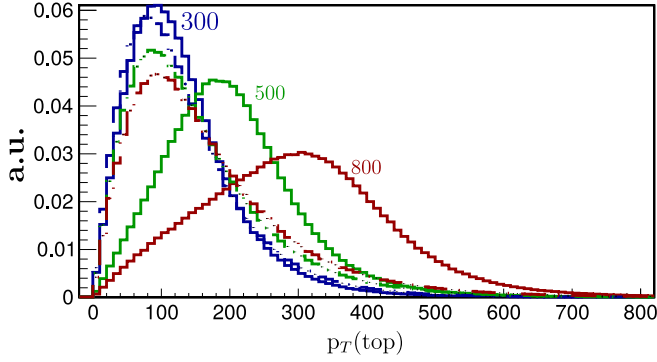
FIG. 3. Parton level transverse momentum of the top quark from charged Higgs decay (solid lines) and associated with it (dashed lines) for $m_{H^{\pm}} = 300$, 500, and 800 GeV, normalized to arbitrary units (a.u.). The distribution of p_T of the top quark corresponds to the combined processes [Eq. (3.2)] and then is appropriately matched using Eq. (3.3).

the top quark originating from $H^{\pm}$ decay is expected to be boosted (with a boost factor $\gamma_t \sim m_{H^{\pm}}/m_t$), in particular for heavier charged Higgs boson masses. The p_T's of those top quarks are demonstrated in Fig. 3 for three masses of $H^{\pm}$, along with the same for associated top quarks. Clearly, this figure indicates that the top quark from heavier $H^{\pm}$ decay is moderately boosted; however, the p_T distribution of associated top quarks is found to be not so sensitive to $m_{H^{\pm}}$. A top quark decays to a b quark and a W, which subsequently decays to a pair of light quarks, leading to jets in the calorimeter. However, for a fast top quark, these decay products may not appear well separated to resolve as separate jets. In such cases, the boosted top quarks may look like a single jet, called a fat jet, with three or more subjets as constituents, corresponding to its decay products. These subjets are well separated within an angular cone of the order $\sim 2m_t/p_T$. Following these kinematic features, we attempt to tag top jets, surrounded by a busy hadronic environment using the top tagger HepTopTagger [74–77]. Our procedure of tagging tops begins by first clustering particles with $p_T \geq 0.5$ GeV and $|\eta| < 5$, using the Cambridge/Aachen [78] jet algorithm implemented in `Fastjet-3.3.0` [79] for jet radius $R = 1.5$ to form fat jets. Then require at least one hard fat jet in the event with $p_T \geq 200$ GeV. In our searches, top tagged jets are likely to be contaminated by QCD radiation, since a wider radius $R = 1.5$ is considered to contain all subjets from the moderately boosted top quark decay. Therefore, it is suggested to take extra measures to eliminate QCD effects due to soft radiation in reconstructing subjets. The substructures of fat jets are obtained following the mass drop method using some recursive steps which are built in to HepTopTagger [74–77]. In this process, the last step of the clustering process is declustered to obtain two subjets j_1 and j_2, such that $m_{j_1} > m_{j_2}$. If $m_{j_1} + m_{j_2} \sim m_j$, and $m_{j_1} > 0.8 m_j$, then

it is expected that j_2 originates from QCD emission or underlying events, and we discard j_2; otherwise we keep both j_1 and j_2. If the mass of the subjet is 30 GeV or less, then we keep it or decompose it further (both j_1 and j_2 or just j_1, depending on how symmetrically the mass splits). The subjets which are obtained at the end of the recursive declustering procedure are also cleaned further through filtering [73] to eliminate the contamination from the QCD radiation. Two subjets are supposed to originate from W decay, and it is ensured by requiring the invariant mass of two subjets $m_{jj} = m_W \pm 15$ GeV. Finally, the top is tagged by adding the third subjet, which is a b-like jet, examined by matching with the parton level b quark in the event. The invariant mass of three subjets after filtering is required to be $m_{jjb} = m_t \pm 30$ GeV. If there is more than one top tagged jet, we choose the one which is the closest to the pole mass of the top quark. Using the default conditions in HepTopTagger, we find that the single top tagging efficiency is about 10% for these kinds of moderately boosted tops in signal events. Note that in calculating this efficiency no pile-up effects are taken into account. The mistagging efficiencies are obtained using the QCD events and they are found to be around 2%–3%.

We attempt to recover this top tagging efficiency to a better level by employing multivariate analysis. The multivariate analysis is implemented within TMVA [80] combining the HepTopTagger mass drop method, and instead of using the full chain of HepTopTagger, some other additional kinematic variables, including N-subjettiness and energy correlation, are used as listed below:

(1) N-subjettiness [81]: Variables are defined as

$$\frac{\tau_2}{\tau_1}, \qquad \frac{\tau_3}{\tau_2}, \qquad \frac{\tau_4}{\tau_3}, \qquad (4.3)$$

where τ_N is the Nth subjettiness variable [82] as defined by

$$\tau_N = \frac{1}{R_0 \sum_k P_{T,k}} \sum_k P_{T,k} \min\left(\Delta R_{1,k}, \Delta R_{2,k} \ldots \Delta R_{N,k}\right). \qquad (4.4)$$

ΔR_{ik} is defined to be the geometrical separation between the ith subjet and the kth reference axis, and R_0 is the jet cone size parameter. Clearly, a smaller τ_N implies more radiation around the given axes, i.e., a better description of jets with N or less subjets, whereas large τ_N means a better description of jets with more than N subjets. It is found that τ_N/τ_{N-1} is an efficient discriminating variable to distinguish boosted objects [81–83].

(2) Mass difference: It is defined as $\Delta m_t = |m_{j_t} - m_t|$, where m_{j_t} is the mass of the tagged top jet. This mass difference is also very crucial in tagging tops.

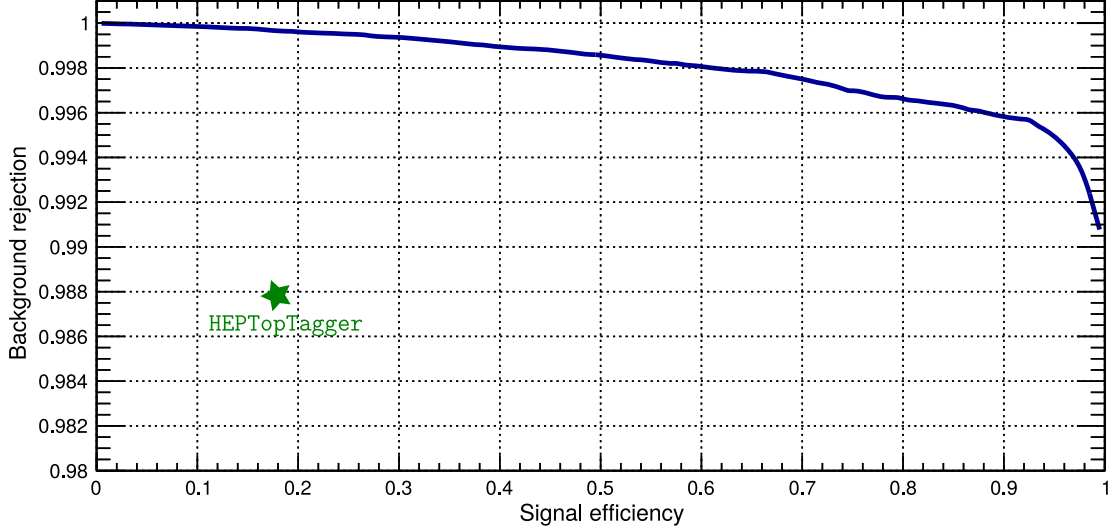(3) Invariant mass of 2 and 3 subjets: Invariant mass of the 3 subjets m_{123} and 2 subjets m_{ij}, where

FIG. 4.   ROC curve for the MVA TopTagger obtained from signal events in the hadronic final state and QCD for mistagging. Efficieny obtained using HepTopTagger are also shown.

($i, j \in \{1, 2, 3\}$) is computed for each possible combination of subjets.

(4) Number of b-like subjets: The number of b-like subjets n_b^j is counted by matching subjets with b-partons within $|\eta| < 2.5$, and $p_T > 5$ GeV using the matching cone $\Delta R < 0.3$ around the subjet.

(5) Variable related with reconstructed masses: It is defined to be

$$f_{\text{rec}} \equiv \min_{i,j} \left| \frac{\left(\frac{m_{ij}}{m_{123}}\right)}{\left(\frac{m_W}{m_t}\right)} - 1 \right|. \qquad (4.5)$$

This ratio determines the quality of reconstructed W with respect to the overall quality of reconstructed top mass.

(6) Energy correlations: The energy correlators among the subjets or particles inside a jet distinguish the various properties of jets [84]. The correlation function uses the information about the energies and pairwise angles between particles within a jet. This generalized energy correlation function interestingly can be made infrared and collinear safe. This energy correlation function is found to be very effective to classify jets. For details, see Ref. [84].

With this set of variables, 1–6, we train boosted decision trees to tag top jets in $t\bar{t}$ and mistags in QCD process. In Fig. 4, we show the results as receiver operator response (ROC) curves for both signal acceptance and background (QCD) rejection efficiencies. This figure clearly demonstrates an improvement in top tagging efficiencies, along with the suppressed background mistag rates. The efficiency obtained using HepTopTagger is also shown by a star. Undoubtedly, the top tagging efficiency through MVA method is improved significantly. We use

this improved efficiency in the simulation of signal and background.

B. Signal and background simulation

The PYTHIA8-8.2.26 (PYTHIA8) [85] is used to generate events via the process $gb \to tH^-$, whereas MadGraph_aMC@NLO-2.6.1 (MG5) [69] is used for $gg \to t\bar{b}H^-$ and then showering through PYTHIA8. The dominant SM background processes $t\bar{t}$ and QCD events are generated using PYTHIA8, while MG5 interfacing with PYTHIA8 is used for the $t\bar{t}b\bar{b}$ process. Events are generated by dividing the phase space into $\hat{p}_T$ bins; $\hat{p}_T$ is the transverse momentum of the final state partons in the center of mass frame. For instance, in the case of signal events, the bins are chosen as 0–200, 200–400, and 400–∞ GeV, whereas for backgrounds ($t\bar{t}$ and QCD), the bins are set to 0–200 (20–200 for QCD), 200–500, 500–800, and 800–∞ GeV. Various event selections imposed on the simulation for both signal and backgrounds are described below:

(1) Lepton selection: Leptons, both electrons and muons, are selected with cuts on the transverse momentum (p_T^ℓ) and rapidity (η_ℓ),

$$p_T^\ell \geq 20 \text{ GeV}, \qquad |\eta_\ell| \leq 2.5. \qquad (4.6)$$

Isolation of a lepton is ensured by requiring $E_T^{AC} \leq$ 30% of p_T^ℓ, where E_T^{AC} is the sum of the transverse momenta of the particles which are within the cone $\Delta R (= \sqrt{\Delta\eta^2 + \Delta\phi^2}) < 0.3$ along the direction of the lepton. It is to be noted that the lepton isolation criteria are not imposed while selecting events applying a lepton veto; otherwise genuine leptonic events would contribute to the hadronic events.

(2) *b*-jet identification: In the simulation, jets are reconstructed using Fastjet [79] with an anti-k_T algorithm [86] and jet size parameter $R = 0.5$. Reconstructed jets are subject to $p_T^j > 20$ GeV, $|\eta_j| < 4$. A given reconstructed jet is identified as a *b*-like jet if there is a match with a parton level *b* quark with a matching cone $\Delta R < 0.3$. In addition, the matched jets are required to have $|\eta| < 2.5$. We found that in about 70% of cases the *b* quarks are identified as *b*-like jets. Finally, in the simulation to select b-like jets, we apply a hard cut $p_T > 30$ GeV. It is to be noted that in our simulation the mistags are not taken into account, which is out of the scope of the present analysis. However, from the studies in [87,88], we found that mistags of the order of few percent are not expected to affect our results significantly.

(3) Top reconstruction: The details of the top tagging are already discussed in the previous section. However, among the tops tagged using this technique, we found that 60%–70% are from the decay of $H^\pm$, while the remaining are the associated tops for the case of $m_{H^\pm} = 500$ GeV and it goes up with the increase of $m_{H^\pm}$. In addition, after the reconstruction of a charged Higgs using tagged top jets, an additional top quark is also reconstructed through kinematic fitting out of the remaining jets for hadronic signal events. This extra kinematically reconstructed top quark is likely to correspond to the associated top quark. For leptonic signal events no such top quark is reconstructed.

(4) Charged Higgs mass reconstruction: We observed via matching that the leading identified *b*-jet with $p_T > 50$ GeV corresponds ($\sim$70%–80%) to a *b* quark originating from $H^\pm$ decay (for $m_{H^\pm} \gtrsim 500$ GeV). Hence, the charged Higgs mass is reconstructed combining the leading top tagged jet with the leading *b*-like jet. In Fig. 5, we show the reconstructed

mass ($m_{t_j b_1}$) of a charged Higgs for three input values, $m_{H^\pm} = 300$, 500, and 800 GeV, along with the dominant background from $t\bar{t}$ corresponding to hadronic final states [Eq. (4.2a))] subject to selection cuts on *b*-jets. The distribution due to $t\bar{t}b\bar{b}$ appears to be almost the same as that due to $t\bar{t}$, whereas for QCD it comes out as flat without any visible peak. The distributions from both of these sources are not shown in this Fig. 5; otherwise it would be very crowded. Notice that the peaks are not appearing exactly at the input mass of charged Higgs because of the smearing of the momenta of a tagged top and *b* jet. The wide spread of $m_{t_j b_1}$ distribution around the peak is due to the incorrect combination of the reconstructed top and *b*-like jet. The events are selected requiring the reconstructed mass $m_{t_j b_1}$ within the range

$$m_{t_j b_1} = m_{H^\pm} \pm 0.3 m_{H^\pm}, \tag{4.7}$$

which is 30% around the peak.

(v) Multiplicity of *b* jets: In signal events the multiplicity of *b*-jets is higher than the $t\bar{t}$ and QCD backgrounds. A hard *b*-jet remains in the signal final state, even after the reconstruction of two (one) tops, and subsequently a charged Higgs in the hadronic (leptonic) final state. The additional *b*-jets appearing in background events are due to the gluon splitting and are not expected to be hard. Therefore, the requirement of at least one hard *b* jet in the final state is expected to be useful in rejecting backgrounds. Hence, the selection

$$n_b \geq 1 \quad \text{with} \quad p_T^b \geq 30 \text{ GeV} \tag{4.8}$$

is imposed in the simulation.

C. Results

We simulate both the production processes in 4FS and 5FS, and then we obtain the final yield by appropriately weighting both the resulting cross sections, as per prescription given by Eqs. (3.3) and (3.4). For illustration purposes, in Table III, the event yields in terms of cross sections are presented after each set of cuts as described above, for signal and backgrounds corresponding to the hadronic final state; see Eq. (4.2a)). The second row shows the total production cross sections of the respective processes at a 13 TeV center of mass energy. The results for the signal events are shown only for a representative choice of a single mass of charged Higgs, $m_{H^\pm} = 500$ GeV, although simulations are performed for a wide range of masses, up to 1 TeV. Also note that the results are presented for $\tan\beta = 30$ and within the framework of a supersymmetric-based model (type II). Fat jets are reconstructed by selecting events with a lepton veto and at least one *b*-identified jet. In order to access the boosted region,
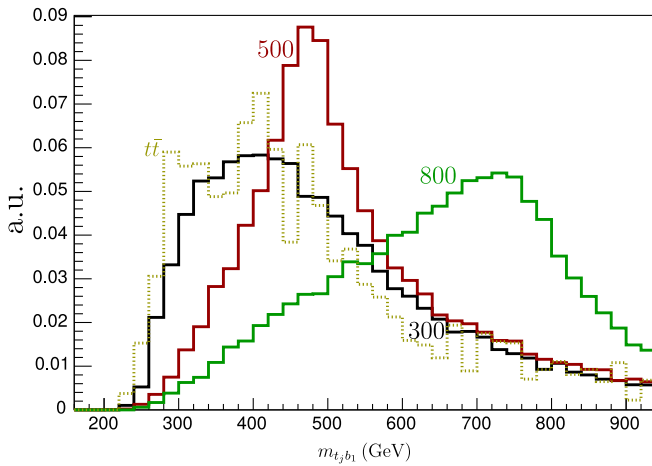

FIG. 5. Charged Higgs mass (= $m_{t_j b_1}$, matched with 4FS and 5FS) reconstruction for $m_{H^\pm} = 300$, 500, and 800 GeV and $\tan\beta = 30$ along with the background from $t\bar{t}$ (dashed lines).

TABLE III. Cross section yields for signal and backgrounds in the hadronic signal final state [Eq. (4.2)a)] for $m_{H^\pm} = 500$ GeV, $\tan\beta = 30$ in type II 2HDM. The first row presents the production cross sections. For signal, Br($H^+ \to t\bar{b}$) is multiplied with the signal cross sections.

Selection	5FS × Br	4FS × Br	$t\bar{t}$	$t\bar{t}b\bar{b}$	QCD
σ (fb)	124.9	64.4	8.3×10^5	1.4×10^4	7.2×10^{11}
$N_b \geq 1$ and lepton veto	55.8	28.8	4.6×10^5	7.4×10^3	1.2×10^{10}
$N_{FJ} \geq 1$	44.8	24.2	1.1×10^5	2.6×10^3	9.4×10^6
$N_{t_j} \geq 1$	13.3	7.4	3.2×10^4	790.4	2.5×10^5
$p_T^{b_1} \geq 50$ GeV	12.3	6.9	2.0×10^4	633.5	9.2×10^4
$m_{t_j b_1} \in [0.7 m_{H^\pm}, 1.3 m_{H^\pm}]$	8.8	5.2	1.3×10^4	387.6	4.2×10^4
$N_{t_{Associated}^{Hadronic}} = 1$	2.2	1.6	364.1	78.8	1.2×10^3
Extra b, $p_T \geq 30$ GeV	0.5	0.5	20.3	15.9	50.2

TABLE IV. Same as in Table III, but for the leptonic signal final state.

Selection	5FS × Br	4FS × Br	$t\bar{t}$	$t\bar{t}b\bar{b}$	QCD
σ (fb)	124.9	64.4	8.3×10^5	1.4×10^4	7.2×10^{11}
$N_b \geq 1$ and $N_\ell \geq 1$	39.3	20.3	2.5×10^5	4.2×10^3	5.0×10^6
$N_{FJ} \geq 1$	29.3	15.8	5.0×10^4	1.3×10^3	3.6×10^3
$N_{t_j} \geq 1$	5.4	3.0	1.0×10^4	276.1	103.1
$p_T^{b_1} \geq 50$ GeV	4.9	2.8	6.7×10^3	221.8	71.7
$m_{t_j b_1} \in [0.7 m_{H^\pm}, 1.3 m_{H^\pm}]$	3.4	2.0	4.4×10^3	138.3	20.5
$p_T^{b_2} \geq 30$ GeV	2.1	1.3	301.0	66.1	$\lesssim 1.0$

events are selected with a $p_T > 200$ GeV on fat jets. These high p_T fat jets are used as an input to HepTopTagger to tag them as top jets. We employ the MVA method as described above to tag top jets, and we have found that about 30% of events are tagged as a top jet. Subsequently, after top tagging, we look for the hardest leading b-jet with a cut $p_T > 50$ GeV, which is found to be originating from $H^\pm$ decay for about 70%–80% of events. Combining the top tagged jets and the hardest b jet, we reconstruct the charged Higgs mass, and we select events within a mass window that is ±30% of the input charged Higgs mass. Notice that a good fraction of background events remain within this reconstructed charged Higgs mass window. With the remaining untagged jets and identified b-jets, the associated top quark is reconstructed. The requirement of a second reconstructed top quark suppresses the background, in particular QCD, more than the signal. Finally, demanding a hard b-jet with $p_T > 30$ GeV rejects backgrounds substantially.

Similarly cross section yields for leptonic final states [Eq. (4.2b)] are presented in Table IV. The events are selected with at least one identified b-jet and one isolated lepton. A top jet is tagged, and it is observed that the efficiency of top tagging is lower, due to the lack of availability of many hadronic top quarks. As before, requiring a hard identified b-jet, with $p_T > 50$ GeV and combining it with tagged top jet, the charged Higgs mass is reconstructed. Finally, the requirement of a hard b-jet

suppresses the background more than the signal. Use of an additional cut on missing transverse momentum due to the presence of neutrinos in the leptonic decay of a top quark is found to be not so helpful.

In Table V, we summarize the signal and background cross sections normalized by the kinematic acceptance efficiencies for both the hadronic and leptonic final state,

TABLE V. Cross sections (fb) normalized by acceptance efficiencies (ϵ_{ac}) for signal and background. Signal significances for three integrated luminosity options for a hadronic (leptonic) final state are performed for the type II model and $\tan\beta = 30$.

$m_{H^\pm} \to$	$\sigma \times \epsilon_{ac}$ (fb)		
	300 GeV	500 GeV	800 GeV
5FS	0.4 (1.4)	0.5 (2.1)	0.1 (0.51)
4FS	0.3 (0.95)	0.5 (1.3)	0.1 (0.34)
$t\bar{t}$	5.9 (140.0)	20.3 (301.0)	15.5 (142.3)
$t\bar{t}b\bar{b}$	5.4 (22.0)	15.9 (66.1)	8.8 (37.6)
QCD	$\lesssim 1.0$ ($\lesssim 1.0$)	50.2 ($\lesssim 1.0$)	21.4 ($\lesssim 1.0$)
Matched signal cross section (S)	0.4 (1.3)	0.5 (1.9)	0.1 (0.47)
Total background cross section (B)	11.3 (161.9)	86.4 (367.1)	45.7 (179.9)
$\mathcal{L}$ (fb^{-1})		$S/\sqrt{B}$	
300	1.9 (1.73)	0.92 (1.71)	0.28 (0.61)
1000	3.4 (3.16)	1.7 (3.13)	0.51 (1.11)
3000	5.9 (5.48)	2.9 (5.42)	0.88 (1.92)

respectively. For illustration, we show the results for three choices of the charged Higgs mass, $m_{H^\pm} = 300, 500,$ and 800 GeV, corresponding to the signal cross sections in both 4FS and 5FS mechanisms. The signal cross sections are found to be $\mathcal{O}(\text{fb})$, whereas the total background contribution is huge, in particular for the hadronic final state. But for the leptonic final state, the level of background contamination is comparatively less. In this case, the presence of leptons and a hard b jet requirement in the final state help to get rid of a good fraction of the QCD background.

The signal significances are presented for three integrated luminosity options $\mathcal{L} = 300, 1000,$ and 3000 fb^{-1}. Table V reveals that the charged Higgs boson of mass 300 GeV can be discovered for high luminosity options $(3000$ fb$^{-1})$ with a reasonable significance, but for higher masses ~ 500 GeV or more, the signal is merely observable. Clearly, it is hard to achieve discoverable signal sensitivity for a heavier charged Higgs mass in this channel. However, the discovery potential of a charged Higgs in the leptonic final state is comparatively better. For instance, Table V shows that the charged Higgs signal is observable with a moderate significance for the mass range around 500 GeV even for the 1000 fb^{-1} integrated luminosity option.

In summary, undoubtedly, this cut-based analysis indicates how difficult it is to achieve discoverable sensitivity of the charged Higgs signal in the $t\bar{b}$ decay mode owing to the huge background cross section with identical event topology. The present set of cuts is not very efficient at suppressing backgrounds at the required level in order to make signal sensitivity better. One may think of a better construction of kinematic observables and devise a set of cuts providing efficient optimization to reduce the background effect. It is a very challenging task to find the feasibility of the charged Higgs signal for heavier masses at the LHC. It motivates us further to develop a search strategy using the technique of multivariate analysis, which is discussed in the next section.

V. MULTIVARIATE ANALYSIS

In the previous section, we observed that there is no single or a combination of kinematic variables which has the potential to isolate a tiny signal out of huge backgrounds. In this section, we discuss MVA in order to improve a signal to background ratio, aiming to achieve a better significance for a given luminosity option. The basic idea of this method is to combine many kinematic variables which are the characteristics of signal events into a single discriminator, and eventually this single discriminator is used to separate out the signal-suppressing backgrounds. The MVA framework is a powerful tool used very widely in high energy physics, to extract the tiny signal events out of huge background events, including single top discovery [89] and recently the Higgs boson at the LHC [90]. Here we

carry out MVA through the boosted decision tree (BDT) method within the framework of TMVA [80].

In the BDT method, events are classified by applying sequentially a set of cuts making subsets of events with a different signal purity. Several disjoint decision trees consisting of two branches are constructed using the best selection of cuts out of the listed input variables of the given process, and it is repeated using subsequent sets of cuts until all the events are classified. While training the sample events, if an event is misclassified, i.e., a signal event labeled as background or a background event as a signal event, then it is boosted by increasing the weight of that event. Subsequently, a second tree is made using the new weights, which may not be same as the previous tree. This process is repeated and we constructed about 1000 trees. There are a few methods of boosting [91], and we use the gradient boosting technique [92]. In the BDT algorithm, these trees are made by training half of the signal and background events. The remaining half of the signal and background events is used to check the performance of the trained BDT.

Following the production and decay mechanism [Eq. (4.1)], events are selected for the final state consisting of one top tagged jet, more than one identified b jet, and untagged jets corresponding to the hadronic signal final state. For a leptonic signal, in addition, at least one isolated lepton is required. A large number of kinematic variables are constructed out of the momenta of these objects to train event samples, and eventually 10 input variables are used in BDT to train the signal and background sample. In Tables VI and VII, the set of input variables is shown, ranking them according to their importance in the BDT analysis for $m_{H^\pm} = 500$ GeV, corresponding to hadronic and leptonic final states, respectively. In the third row of this table, a brief description is provided for each of the variables. The importance here means the effectiveness of

TABLE VI. Kinematic variables used to train the signal and background sample for the hadronic final state and $m_{H^\pm} = 500$ GeV.

Rank	Variables	Description
1	$m_{b\bar{b}}^{12}$	Invariant mass of two b jets
2	p_T^j	p_T of the leading jet
3	Tagged TopMVA	MVA discriminator for top tagging
4	$m_{b_1 b_2 b_3}$	Invariant mass of three b-jets
5	$p_{T_2}^{b_{\text{jet}}}$	p_T of second b-jet after top tagging
6	$m_{t_j b_1}^{\text{reco}}$	Reconstructed Higgs mass
7	H_T	Scalar sum of p_T of all final detectable particles
8	n_j	Number of untagged
9	$p_{T_3}^{b_{\text{jet}}}$	p_T of third b-jet
10	H_T^b	Scalar sum of p_T of all b-jets

TABLE VII.   Same as Table VI, but for the leptonic final state.

Rank	Variables	Description
1	H_T^b	Scalar sum of p_T of all b-jets
2	m_{t_j}	Mass of the tagged top jet
3	$m_{b\bar{b}}^{12}$	Invariant mass of two b-jets
4	m_{tb_2}	Invariant mass of tagged top jet and second b-jet
5	H_T/MHT	Ratio over H_T and MHT
6	$p_T^{j_1}$	p_T of leading untagged jet
7	H_T	Scalar sum of p_T of all jets
8	$p_T^{b_{\text{jet}}}$	p_T of leading b-jet
9	$p_{T_2}^{b_{\text{jet}}}$	p_T of second b-jet after top tagging
10	MHT	Vector sum of p_T of all jets and leptons

those variables in suppressing backgrounds while maintaining a better signal purity.

We have observed that for $m_{H^\pm} = 300$ GeV, the importance or ranking of some of the variables is altered. For example, for 300 GeV, $p_{T_3}^{b_{\text{jet}}}$ and H_T^b are found to be more important than m_{tb}^{reco}. Similarly, for a much heavier charged Higgs mass, the H_T is expected to be more important in suppressing background; hence it is ranked second. Interestingly, the invariant mass of the first two leading b-jets seems to be a very strong discriminant variable in separating the signal and background. Moreover, the MVA discriminator for top tagging using HepTopTagger, the multiplicity of untagged jets, and the p_T of the second b-jet all appear to be useful variables in eliminating the background events.

In Table VII, the set of kinematic variables is presented for a leptonic final state and for $m_{H^\pm} = 500$ GeV. However, as before, this set remains the same for $m_{H^\pm} = 300$ and 1000 GeV, but the ranking becomes different for obvious reasons. For instance, for the lower mass of $m_{H^\pm} = 500$ GeV, the variable 3 becomes more important than the variable 1. Due to the presence of neutrinos, the variable related with missing transverse energy, MHT plays a role in discriminating the background, in particular from QCD. Like the hadronic case, the number of b-jets and their corresponding transverse momentum are very effective in increasing the signal to background ratio.

In this type of analysis based on machine learning models, one of the issues often encountered is the problem of overtraining the sample. The training of the sample can be checked using a test data sample. Ideally, for a sufficiently large and random set of Monte Carlo data, the performance of the training and testing data should be similar. If significant deviations between these two are found, that would be an indication of overtraining of the sample. These overtraining tests are performed for all $m_{H^\pm} = 300$–1000 GeV masses.
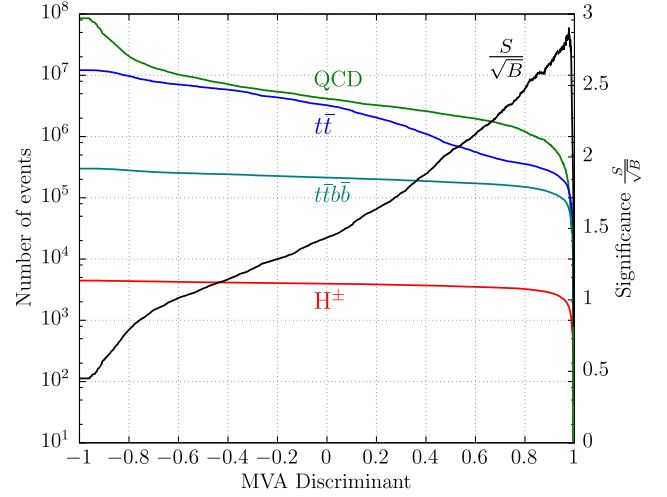

FIG. 6.   MVA output (D) distribution for signal and backgrounds corresponding to the hadronic signal final state and $m_{H^\pm} = 500$ GeV, $\tan\beta = 30$ for the type II 2HDM.

In Fig. 6, the distribution of the MVA output discriminator (D) with the number of events is presented, for signal events with $m_{H^\pm} = 500$ GeV and backgrounds from QCD, $t\bar{t}$, and $t\bar{t}b\bar{b}$, along with the significance $S/\sqrt{B}$ for an integrated luminosity of 300 fb^{-1}. A significance close to 3σ can be achieved with a selection of the discriminator, D > 0.9. With this cut on D, and for an integrated luminosity of 300 fb^{-1}, the number of events turn out to be 2830 for signal, and 1140000 for total backgrounds, where 70% of the contribution comes from QCD. The selection of D > 0.9 leads to a significance ~2.65σ, which goes up more for higher luminosity options.

Unlike the hadronic case, in the leptonic signal final state (see Fig. 7), the dominant background appears to be due to $t\bar{t}$ production. A cut on BDT output D > 0.9 leads to a significance of about 3σ for $\mathcal{L} = 300$ fb^{-1}. The study is extended up to the 1000 GeV mass of the charged Higgs.
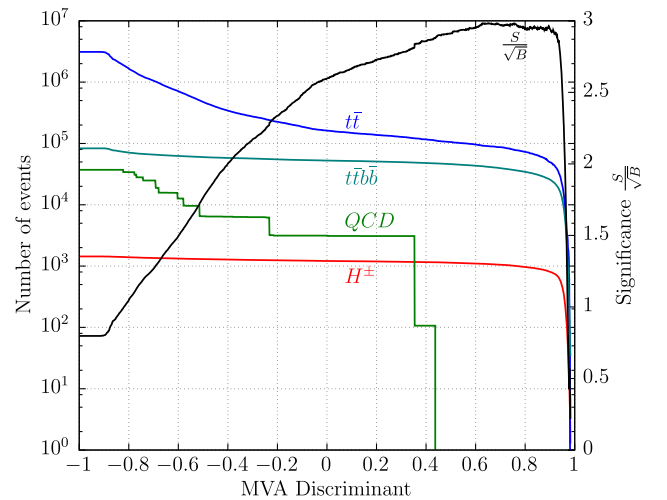

FIG. 7.   Same as for Fig. 6, but for the leptonic signal.

TABLE VIII. Significances for the hadronic (leptonic) final states and for three luminosity options, in the SUSY motivated type II model with $\tan\beta = 30$.

$m_{H^\pm}$(GeV) $\rightarrow$	300	$S/\sqrt{B}$ 500	800	1000
$\mathcal{L} = 300$ fb^{-1}	6.1 (5.2)	2.7 (2.94)	0.61 (0.96)	0.22 (0.39)
$\mathcal{L} = 1000$ fb^{-1}	11.0 (9.5)	4.8 (5.4)	1.1 (1.7)	0.40 (0.71)
$\mathcal{L} = 3000$ fb^{-1}	19.1 (16.5)	8.4 (9.3)	1.9 (3.0)	0.70 (1.2)

TABLE X. Same as Table IX, but with $\tan\beta = 3$.

$m_{H^\pm}$ (GeV)	$\mathcal{L}$ (fb^{-1})	Type I	Type II	Type III	Type IV
	300	4.3 (3.7)	4.3 (3.7)	4.3 (3.7)	4.3 (3.7)
300	1000	7.8 (6.7)	7.9 (6.8)	7.9 (6.8)	7.8 (6.7)
	3000	13.5 (11.7)	13.7 (11.8)	13.7 (11.8)	13.5 (11.7)
	300	1.7 (1.9)	1.7 (1.9)	1.7 (1.9)	1.7 (1.9)
500	1000	3.1 (3.4)	3.1 (3.5)	3.1 (3.5)	3.1 (3.4)
	3000	5.3 (5.9)	5.4 (6.0)	5.4 (6.0)	5.3 (5.9)
	300	0.39 (0.62)	0.40 (0.63)	0.40 (0.63)	0.39 (0.62)
800	1000	0.72 (1.1)	0.73 (1.1)	0.73 (1.1)	0.72 (1.1)
	3000	1.2 (2.0)	1.3 (2.0)	1.3 (2.0)	1.2 (2.0)

Signal significances are presented for both the hadronic and leptonic final states (in parentheses) in Table VIII for three masses of the charged Higgs and for three integrated luminosity options. Remarkably, using the MVA technique, a significant improvement in sensitivity for both the hadronic and leptonic signal is achieved. This table suggests that in the hadronic channel the charged Higgs boson of mass up to $\sim$500 GeV can be probed with a reasonable sensitivity, much better than that obtained using a simple cut-based analysis, as shown in Table V.

Table VIII shows that the signature of a charged Higgs of mass around 800 GeV is observable in the leptonic channel for the 3000 fb^{-1} luminosity option unlike the hadronic final state. For a lower range of masses ($\sim$500 GeV), the signal is feasible even for the 1000 fb^{-1} luminosity option.

The results presented in Table VIII correspond to the SUSY motivated type II model. However, the signal cross sections for other classes of 2HDM can be obtained out of these estimated values simply by rescaling the couplings and appropriately multiplying Br$(H^+ \rightarrow t\bar{b})$. The significances for all four types of models are presented for both the hadronic and leptonic (decay modes) in Tables IX and X, corresponding to $\tan\beta = 30$ and 3, respectively. Table IX suggests that for the high $\tan\beta$ scenario, the discovery potential of a charged Higgs in the context of the type II and type IV models is quite promising for masses up to around 600–700 GeV; however, due to a little increase ($\sim$20%) of Br$(H^+ \rightarrow t\bar{b})$, the sensitivity is better for the type IV model. Because of the suppressed coupling of a charged Higgs with top and bottom quarks, for the high $\tan\beta$ scenario, the signal sensitivity is very poor for both

the type I and III models. However, for the low $\tan\beta$ scenario (10), results suggest that the discovery potential is quite promising for this kind of model parameter space. Interestingly significances corresponding to all types of models are found to be almost the same for a given mass and luminosity option. It can be attributed to the fact that, in all the classes of 2HDM, the dominant part of charged Higgs couplings is proportional to $m_t \cot\beta$, which results in the same significances.

The discovery potential of a charged Higgs of mass 300 GeV in the $\bar{t}b$ decay channel is quite promising even for the 300 fb^{-1} luminosity option at 13 TeV energy. However, for higher masses, e.g., for $m_{H^\pm} = 500$ GeV, one needs high luminosity options such as 1000 fb^{-1} or more. This study shows that for higher masses $\sim$1000 GeV, it is very hard to achieve better signal sensitivity even for the high luminosity option.

Finally, in Fig. 8 the discovery region is presented in the $m_{H^\pm} - \tan\beta$ plane in the context of the SUSY motivated type II model, requiring a 5σ significance for the given luminosity options, as shown in the figure. Contours show the minimum value of $\tan\beta$ required to discover a charged Higgs of given mass at the 5σ level for a given luminosity option. The parameter space above the contours is discoverable corresponding to that luminosity option at $\sqrt{s} = 13$ TeV energy. In the hadronic channel, even for the high luminosity option, it is very hard to find a charged Higgs of

TABLE IX. Sensitivity ($S/\sqrt{B}$) for hadronic (leptonic) signal events corresponding to all four types of 2HDM, with $\tan\beta = 30$ and $\sin(\beta - \alpha) = 1$.

| $m_{H^\pm}$ (GeV)	$\mathcal{L}$ (fb^{-1})	Type I	Type II	Type III	Type IV
	300	0.043 (0.037)	7.4 (6.4)	0.0007 (0.0006)	9.270 (7.999)
300	1000	0.08 (0.07)	13.5 (11.6)	0.001 (0.001)	16.925 (14.603)
	3000	0.14 (0.12)	23.3 (20.1)	0.002 (0.002)	29.316 (25.294)
	300	0.017 (0.019)	3.1 (3.4)	0.0004 (0.0004)	3.621 (4.017)
500	1000	0.031 (0.034)	5.6 (6.2)	0.0007 (0.0009)	6.611 (7.335)
	3000	0.053 (0.059)	9.8 (10.8)	0.001 (0.001)	11.451 (12.704)
	300	0.004 (0.006)	0.71 (1.1)	0.0001 (0.0002)	0.823 (1.294)
800	1000	0.007 (0.011)	1.3 (2.0)	0.0002 (0.0003)	1.502 (2.363)
	3000	0.01 (0.02)	2.2 (3.5)	0.0003 (0.0005)	2.601 (4.093)
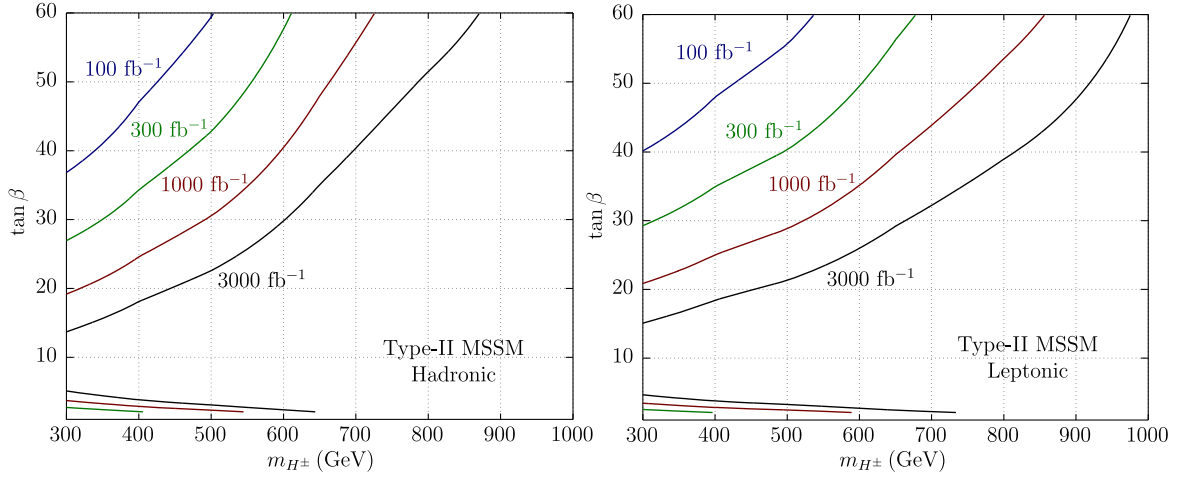
FIG. 8.   Discovery region for the given luminosity (fb^{-1}) options in the context of the SUSY (type II) model for the hadronic (left) and leptonic (right) final state at $\sqrt{s} = 13$ TeV.

mass beyond 850 GeV. On the other hand, for the leptonic final state, the charged Higgs can be explored almost up to $m_{H^\pm} \sim 1$ TeV with high luminosity. Discovery regions below the contours for much lower $\tan\beta$ are also shown with three luminosity options. For a given $m_{H^\pm}$, the lowest $\tan\beta$ corresponds to 3000 fb^{-1} and then decreases to 1000 and 300 fb^{-1} for the other two lines, respectively.

It is to be noted that while calculating the signal significance, the uncertainties of the background are not taken into account. The estimation of systematics in the background evaluation is currently out of the scope of the present analysis. However, due to the tiny signal size in comparison to background events, i.e., with less purity, the impact of systematic uncertainties is expected to be severe. It can be understood by evaluating the significance as $S/\sqrt{B + (\delta B)^2}$, where δ stands for the level of uncertainties. For instance, corresponding to the moderate range of charged Higgs masses, and for about 20% uncertainties in background estimation, the significances go down drastically, for both the hadronic and leptonic case. For a heavier mass range, $m_{H^\pm} \sim 800$ GeV, the impact of systematics to significance is not that severe due to the smaller number of background events. Clearly, in order to achieve a reasonable significance to discover the charged Higgs for the intermediate mass range, one needs to perform the background estimation as precisely as possible.

VI. SUMMARY

In this study, we explore the detection prospect of the charged Higgs boson for the heavier mass range at the LHC in Run 2 experiments with the center of mass energy $\sqrt{s} = 13$ TeV, within the framework of generic 2HDM. A very brief discussion of 2HDM is presented in order to set up the model framework to carry out the analysis. It is observed that, in all classes of 2HDM, the Br($H^+ \rightarrow t\bar{b}$) is

always the dominant one, except in the type III model, where it holds only for the lower range of $\tan\beta (< 10)$. It is to be noted that the other decay modes, such as $H^\pm \rightarrow W^\pm \phi$ ($\phi \in \{h, H, A\}$) also open up with a large Br once the condition $\sin(\beta - \alpha) = 1$ is relaxed, as discussed in Sec. II. The charged Higgs boson production cross section at $\sqrt{s} = 13$ TeV is computed in both the 4FS and 5FS mechanisms, and finally the matched values are presented for a few representative choices of $m_{H^\pm}$. In the context of the SUSY motivated type II model, the matched cross sections vary from $\mathcal{O}(100$ fb$)$ to $\mathcal{O}(10$ fb$)$ for the range of $m_{H^\pm} \sim$ 300–1000 GeV, corresponding to large values of $\tan\beta$, and are found to be less than those for other classes of 2HDM. The signature of the charged Higgs is analyzed for the final state consisting of a reconstructed charged Higgs mass and extra b-jets, plus an additional reconstructed top quark for hadronic events, while in leptonic events, a lepton is required without reconstruction of a second top. The jet substructure technique is used to tag a moderately boosted top quark from the heavier charged Higgs decay in order to avoid the recombinatorial problem while reconstructing the charged Higgs mass. The MVA method is employed including inputs from HepTopTagger to tag topjets. A better top tagging efficiency with a lower mistagging rate is achieved in comparison to the result obtained using only the default HepTopTagger. A detailed simulation is performed for the signal and the main dominant irreducible SM backgrounds from the top quark pair production and QCD. The cut-based analysis predicts a very poor signal sensitivity even for the high luminosity options. However, for the lower mass range of the charged Higgs, $m_{H^\pm} \sim 300$ GeV, one can expect a modest sensitivity for the 3000 fb^{-1} luminosity option. In order to improve the signal significance, the analysis is carried out using the techniques of BDT within the framework of TMVA. Several kinematic

variables are constructed to train BDT. Remarkably, MVA analysis yields a substantial improvement in signal significance. For example, this MVA-based analysis shows that with $\mathcal{L} = 1000$ fb^{-1}, the signature of the charged Higgs boson for the mass range ∼300–700 GeV can be probed for both the hadronic and leptonic channels. For an even higher luminosity option, such as 3000 fb^{-1}, the discovery reach of $m_{H^{\pm}}$ can be extended up to ∼800 GeV for the hadronic final state, whereas for the leptonic case, it can be extended further, up to almost 1 TeV for high values of $\tan\beta$. In Fig. 8, the discovery potential of the charged Higgs boson is presented in the $m_{H^{\pm}} - \tan\beta$ plane for a few integrated luminosity options. This figure indicates that the discovery reach corresponding to the leptonic final state is better than the hadronic signal case. By simply scaling the charged Higgs couplings with fermions, and then the production cross sections, we present signal significances for all classes of 2HDM for three representative choices of $m_{H^{\pm}}$ and two values of $\tan\beta = 30$ and 3. The results show that for the high $\tan\beta = 30$ scenario, it is difficult to achieve any

detectable signal sensitivity, except for the type II and type IV models. However, for the low $\tan\beta (= 3)$ case, the signal of a charged Higgs for the mass range ∼300–600 GeV seems to be detectable with a ∼3σ sensitivity for the leptonic final state with $\mathcal{L} = 1000$ fb^{-1}. Indeed, it is hard to discover the signal of the charged Higgs boson of mass beyond 800 GeV for the low $\tan\beta$ scenario, even for the higher luminosity options. Definitively, to probe the charged Higgs boson of large mass, more than 800 GeV, one needs a very high energy option, such as a 100 TeV hadron collider[93].

ACKNOWLEDGMENTS

[1] 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).

[2] 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).

[3] 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.

[4] J. F. Gunion, S. Dawson, H. E. Haber, and G. L. Kane, *The Higgs Hunter's Guide* (Brookhaven Nat. Lab., Upton, NY, 1989), Vol. 80.

[5] 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).

[6] 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).

[7] M. Carena, I. Low, N. R. Shah, and C. E. M. Wagner, Impersonating the Standard Model Higgs boson: Alignment without decoupling, J. High Energy Phys. 04 (2014) 015.

[8] P. S. Bhupal Dev and A. Pilaftsis, Maximally symmetric two Higgs doublet model with natural Standard Model alignment, J. High Energy Phys. 12 (2014) 024; Erratum, J. High Energy Phys. 11 (2015) 147(E).

[9] M. Carena, H. E. Haber, I. Low, N. R. Shah, and C. E. M. Wagner, Complementarity between nonstandard Higgs

boson searches and precision Higgs boson measurements in the MSSM, Phys. Rev. D **91**, 035003 (2015).

[10] S. Profumo and T. Stefaniak, Alignment without decoupling: The portal to light dark matter in the MSSM, Phys. Rev. D **94**, 095020 (2016).

[11] D. de Florian *et al.* (LHC Higgs Cross Section Working Group), Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector, arXiv:1610.07922.

[12] A. Datta, A. Djouadi, M. Guchait, and Y. Mambrini, Charged Higgs boson production from supersymmetric particle cascade decays at the CERN LHC, Phys. Rev. D **65**, 015007 (2001).

[13] A. Datta, A. Djouadi, M. Guchait, and F. Moortgat, Detection of MSSM Higgs bosons from supersymmetric particle cascade decays at the LHC, Nucl. Phys. **B681**, 31 (2004).

[14] G. Abbiendi *et al.* (LEP, DELPHI, OPAL, ALEPH, and L3 Collaborations), Search for charged Higgs bosons: Combined results using LEP data, Eur. Phys. J. C **73**, 2463 (2013).

[15] T. Aaltonen *et al.* (CDF Collaboration), Search for Charged Higgs Bosons in Decays of Top Quarks in p Anti-p Collisions at s ∗ ∗(1/2) = 1.96 TeV, Phys. Rev. Lett. **103**, 101803 (2009).

[16] G. Aad *et al.* (ATLAS Collaboration), Search for charged Higgs bosons decaying via $H^{\pm} \to \tau^{\pm}\nu$ in fully hadronic final states using pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector, J. High Energy Phys. 03 (2015) 088.

[17] V. Khachatryan *et al.* (CMS Collaboration), Search for a charged Higgs boson in pp collisions at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 11 (2015) 018.

[18] G. Aad *et al.* (ATLAS Collaboration), Search for a light charged Higgs boson in the decay channel $H^+ \to c\bar{s}$ in $t\bar{t}$ events using pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, Eur. Phys. J. C **73**, 2465 (2013).

[19] V. Khachatryan *et al.* (CMS Collaboration), Search for a light charged Higgs boson decaying to c$\bar{s}$ in pp collisions at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 12 (2015) 1.

[20] CMS Collaboration, Search for charged Higgs boson to c$\bar{b}$ in lepton + jets channel using top quark pair events, Tech. Rep. CMS-PAS-HIG-16-030, CERN, Geneva, 2016.

[21] CMS Collaboration, Search for charged Higgs bosons with the $H^\pm \to \tau^\pm \nu_\tau$ decay channel in the fully hadronic final state at $\sqrt{s} = 13$ TeV, Technical Report CMS-PAS-HIG-16-031, CERN, Geneva, 2016.

[22] M. Aaboud *et al.* (ATLAS Collaboration), Search for charged Higgs bosons produced in association with a top quark and decaying via $H^\pm \to \tau\nu$ using pp collision data recorded at $\sqrt{s} = 13$ TeV by the ATLAS detector, Phys. Lett. B **759**, 555 (2016).

[23] ATLAS Collaboration, Search for charged Higgs bosons in the $H^\pm \to tb$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector, Technical Report ATLAS-CONF-2016-089, CERN, Geneva, 2016.

[24] CMS Collaboration, Search for a high-mass resonance decaying into a dilepton final state in 13 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV, Technical Report CMS-PAS-EXO-16-031, CERN, Geneva, 2016.

[25] G. Aad *et al.* (ATLAS Collaboration), Search for charged Higgs bosons in the $H^\pm \to tb$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector, J. High Energy Phys. 03 (2016) 127.

[26] A. M. Sirunyan *et al.* (CMS Collaboration), Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 09 (2018) 007.

[27] A. Arbey, F. Mahmoudi, O. Stal, and T. Stefaniak, Status of the charged Higgs boson in two Higgs doublet models, Eur. Phys. J. C **78**, 182 (2018).

[28] M. Krawczyk, S. Moretti, P. Osland, G. Pruna, and R. Santos, Prospects for 2HDM charged Higgs searches, J. Phys. Conf. Ser. **873**, 012048 (2017).

[29] A. Arhrib, R. Benbrik, R. Enberg, W. Klemm, S. Moretti, and S. Munir, Identifying a light charged Higgs boson at the LHC run II, Phys. Lett. B **774**, 591 (2017).

[30] L. Basso, A. Lipniacka, F. Mahmoudi, S. Moretti, P. Osland, G. M. Pruna, and M. Purmohammadi, The *CP*-violating type-II 2HDM and charged Higgs boson benchmarks, Proc. Sci., Corfu2012 (**2013**) 029 [arXiv:1305.3219].

[31] A. Arhrib, R. Benbrik, and S. Moretti, Bosonic decays of charged Higgs bosons in a 2HDM type-I, Eur. Phys. J. C **77**, 621 (2017).

[32] K. A. Assamagan, M. Guchait, and S. Moretti, Charged Higgs bosons in the transition region M(H+-) m(t) at the LHC, in *Physics at TeV Colliders. Proceedings, Workshop, Les Houches, France, 2003* (CNRS, Marseille, France, 2004), p. 459.

[33] D. P. Roy, The hadronic τ decay signature of a heavy charged Higgs boson at LHC, Phys. Lett. B **459**, 607 (1999).

[34] S. Moretti and K. Odagiri, Production of charged Higgs bosons of the minimal supersymmetric Standard Model in b quark initiated processes at the Large Hadron Collider, Phys. Rev. D **55**, 5627 (1997).

[35] S. Moretti, R. Santos, and P. Sharma, Optimising charged Higgs boson searches at the Large Hadron Collider across $b\bar{b}W^\pm$ final states, Phys. Lett. B **760**, 697 (2016).

[36] R. Enberg, W. Klemm, S. Moretti, S. Munir, and G. Wouda, Charged Higgs boson in the $W^\pm$ Higgs channel at the Large Hadron Collider, Nucl. Phys. **B893**, 420 (2015).

[37] R. Guedes, S. Moretti, and R. Santos, Charged Higgs bosons in single top production at the LHC, J. High Energy Phys. 10 (2012) 119.

[38] J. L. Diaz-Cruz, J. Hernandez-Sanchez, S. Moretti, and A. Rosado, Charged Higgs boson phenomenology in supersymmetric models with Higgs triplets, Phys. Rev. D **77**, 035007 (2008).

[39] S. Moretti, 2HDM charged Higgs boson searches at the LHC: Status and prospects, Proc. Sci., CHARGED2016 (**2016**) 014 [arXiv:1612.02063].

[40] M. Aoki, R. Guedes, S. Kanemura, S. Moretti, R. Santos, and K. Yagyu, Light charged Higgs bosons at the LHC in 2HDMs, Phys. Rev. D **84**, 055028 (2011).

[41] L. Basso, A. Lipniacka, F. Mahmoudi, S. Moretti, P. Osland, G. M. Pruna, and M. Purmohammadi, Probing the charged Higgs boson at the LHC in the *CP*-violating type-II 2HDM, J. High Energy Phys. 11 (2012) 011.

[42] A. G. Akeroyd *et al.*, Prospects for charged Higgs searches at the LHC, Eur. Phys. J. C **77**, 276 (2017).

[43] S. Raychaudhuri and D. P. Roy, Sharpening up the charged Higgs boson signature using τ polarization at LHC, Phys. Rev. D **53**, 4902 (1996).

[44] M. Guchait and D. P. Roy, Using tau polarisation for charged Higgs boson and SUSY searches at the LHC, in *Physics at the Large Hadron Collider*, (2009), p. 205, DOI: 10.1007/978-81-8489-295-6_13

[45] M. Guchait, R. Kinnunen, and D. P. Roy, Signature of heavy charged Higgs boson at LHC in the 1 and 3 prong hadronic tau decay channels, Eur. Phys. J. C **52**, 665 (2007).

[46] S. Moretti and K. Odagiri, The phenomenology of $W^\pm H^\pm$ production at the large hadron collider, Phys. Rev. D **59**, 055008 (1999).

[47] A. A. Barrientos Bendezu and B. A. Kniehl, H+ H- pair production at the large hadron collider, Nucl. Phys. **B568**, 305 (2000).

[48] A. Krause, T. Plehn, M. Spira, and P. M. Zerwas, Production of charged Higgs boson pairs in gluon-gluon collisions, Nucl. Phys. **B519**, 85 (1998).

[49] J. F. Gunion, Detecting the tb decays of a charged Higgs boson at a hadron supercollider, Phys. Lett. B **322**, 125 (1994).

[50] V. D. Barger, R. J. N. Phillips, and D. P. Roy, Heavy charged Higgs signals at the LHC, Phys. Lett. B **324**, 236 (1994).

[51] D. J. Miller, S. Moretti, D. P. Roy, and W. J. Stirling, Detecting heavy charged Higgs bosons at the CERN LHC with four b quark tags, Phys. Rev. D **61**, 055011 (2000).

[52] S. Moretti and D. P. Roy, Detecting heavy charged Higgs bosons at the LHC with triple b tagging, Phys. Lett. B **470**, 209 (1999).

[53] S. Yang and Q.-S. Yan, Searching for heavy charged Higgs boson with jet substructure at the LHC, J. High Energy Phys. 02 (2012) 074.

[54] R. Patrick, P. Sharma, and A. G. Williams, Exploring a heavy charged Higgs using jet substructure in a fully hadronic channel, Nucl. Phys. **B917**, 19 (2017).

[55] K. Pedersen and Z. Sullivan, Probing the two Higgs doublet wedge region with charged Higgs boson decays to boosted jets, Phys. Rev. D **95**, 035037 (2017).

[56] A. Djouadi, J. Kalinowski, M. Muehlleitner, and M. Spira, HDECAY: Twenty$_{++}$ years after, arXiv:1801.09506.

[57] A. Djouadi, J. Kalinowski, and M. Spira, HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension, Comput. Phys. Commun. **108**, 56 (1998).

[58] M. Bisset, M. Guchait, and S. Moretti, Signatures of MSSM charged Higgs bosons via chargino neutralino decay channels at the LHC, Eur. Phys. J. C **19**, 143 (2001).

[59] S. Dittmaier, M. Kramer, M. Spira, and M. Walser, Charged-Higgs-boson production at the LHC: NLO supersymmetric QCD corrections, Phys. Rev. D **83**, 055005 (2011).

[60] T. Plehn, Charged Higgs boson production in bottom gluon fusion, Phys. Rev. D **67**, 014018 (2003).

[61] E. L. Berger, T. Han, J. Jiang, and T. Plehn, Associated production of a top quark and a charged Higgs boson, Phys. Rev. D **71**, 115012 (2005).

[62] N. Kidonakis, Charged Higgs production via bg → tH- at the LHC, J. High Energy Phys. 05 (2005) 011.

[63] N. Kidonakis, Charged Higgs production: Higher-order corrections, Proc. Sci., HEP2005 (**2006**) 336 [arXiv:hep-ph/0511235].

[64] N. Kidonakis, Charged Higgs production in association with a top quark at approximate NNLO, Phys. Rev. D **94**, 014010 (2016).

[65] M. Flechl, R. Klees, M. Kramer, M. Spira, and M. Ubiali, Improved cross-section predictions for heavy charged Higgs boson production at the LHC, Phys. Rev. D **91**, 075015 (2015).

[66] R. Harlander, M. Kramer, and M. Schumacher, Bottom-quark associated Higgs-boson production: Reconciling the four- and five-flavour scheme approach, arXiv:1112.3478.

[67] F. Maltoni, G. Ridolfi, and M. Ubiali, b-initiated processes at the LHC: A reappraisal, J. High Energy Phys. 07 (2012) 022; Erratum 04 (2013) 095.

[68] F. Borzumati, J.-L. Kneur, and N. Polonsky, Higgs-strahlung and R-parity violating slepton-strahlung at hadron colliders, Phys. Rev. D **60**, 115011 (1999).

[69] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, J. High Energy Phys. 07 (2014) 079.

[70] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, FeynRules 2.0—A complete toolbox for tree-level phenomenology, Comput. Phys. Commun. **185**, 2250 (2014).

[71] C. Degrande, M. Ubiali, M. Wiesemann, and M. Zaro, Heavy charged Higgs boson production at the LHC, J. High Energy Phys. 10 (2015) 145.

[72] A. V. Bednyakov, B. A. Kniehl, A. F. Pikelner, and O. L. Veretin, On the b-quark running mass in QCD and the SM, Nucl. Phys. **B916**, 463 (2017).

[73] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, Jet Substructure as a New Higgs Search Channel at the LHC, Phys. Rev. Lett. **100**, 242001 (2008).

[74] T. Plehn and M. Spannowsky, Top tagging, J. Phys. G **39**, 083001 (2012).

[75] G. Kasieczka, T. Plehn, T. Schell, T. Strebler, and G. P. Salam, Resonance searches with an updated top tagger, J. High Energy Phys. 06 (2015) 203.

[76] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, Stop reconstruction with tagged tops, J. High Energy Phys. 10 (2010) 078.

[77] T. Plehn, G. P. Salam, and M. Spannowsky, Fat Jets for a Light Higgs, Phys. Rev. Lett. **104**, 111801 (2010).

[78] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, Better jet clustering algorithms, J. High Energy Phys. 08 (1997) 001.

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

[80] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, TMVA: Toolkit for multivariate data analysis, Proc. Sci., ACAT2007 (**2007**) 040 [arXiv:physics/0703039].

[81] J. Thaler and K. Van Tilburg, Maximizing boosted top identification by minimizing N-subjettiness, J. High Energy Phys. 02 (2012) 093.

[82] J. Thaler and K. Van Tilburg, Identifying boosted objects with N-subjettiness, J. High Energy Phys. 03 (2011) 015.

[83] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, N-Jettiness: An Inclusive Event Shape to Veto Jets, Phys. Rev. Lett. **105**, 092002 (2010).

[84] A. J. Larkoski, G. P. Salam, and J. Thaler, Energy correlation functions for jet substructure, J. High Energy Phys. 06 (2013) 108.

[85] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 physics and manual, J. High Energy Phys. 05 (2006) 026.

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

[87] Expected performance of the ATLAS b-tagging algorithms in Run-2, Technical Report ATL-PHYS-PUB-2015-022, CERN, Geneva, 2015. http://cds.cern.ch/record/2037697.

[88] CMS Collaboration, Identification of b quark jets at the CMS experiment in the LHC run 2, Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[89] V. M. Abazov et al. (D0 Collaboration), Evidence for Production of Single Top Quarks and First Direct Measurement of |Vtb|, Phys. Rev. Lett. **98**, 181802 (2007).

[90] V. Khachatryan et al. (CMS Collaboration), Observation of the diphoton decay of the Higgs boson and measurement of its properties, Eur. Phys. J. C **74**, 3076 (2014).

[91] R. E. Schapire, The strength of weak learnability, Mach. Learn. **5**, 197 (1990).

[92] J. H. Friedman, Greedy function approximation: A gradient boosting machine, Ann. Stat. **29**, 1189 (2001).

[93] M. Mangano, Physics at the FCC-hh, a 100 TeV pp collider, arXiv:1710.06353.