

Precise probing of the inert Higgs-doublet model at the LHC

Anupam Ghosh^{ⓧ, 1,2,*} Partha Konar^{ⓧ, 1,†} and Satyajit Seth^{ⓧ, 1,‡}

¹Physical Research Laboratory (PRL), Ahmedabad, 380009, Gujarat, India

²Indian Institute of Technology, Gandhinagar, 382424, Gujarat, India



(Received 1 February 2022; accepted 3 June 2022; published 28 June 2022)

The inert Higgs-doublet model provides a simple framework to accommodate a viable Higgs portal scalar dark matter candidate, together with other heavier scalars of mass 100 GeV or more. We study the effect of next-to-leading order (NLO) QCD corrections in this scenario in the context of the Large Hadron Collider. $\mathcal{O}(\alpha_s)$ corrections to the gluon-gluon-Higgs effective coupling have been taken into account in this study wherever appropriate. We find such corrections have a significant impact on various kinematic distributions and reduce scale uncertainties substantially. Fixed order NLO results are matched to the PYTHIA8 parton shower (PS) and the di-fatjet signal associated with the missing transverse momentum is analyzed, as this channel has the ability to explore its entire parameter space during the next phase of the LHC run. A closer look at the NLO + PS computation indicates a sizable NLO effect together with a subdued contribution from associated production of the heavy scalar compared to the pair production, thereby leading to a refined analysis strategy during the multivariate analysis of this signal.

DOI: [10.1103/PhysRevD.105.115038](https://doi.org/10.1103/PhysRevD.105.115038)

I. INTRODUCTION

The Standard Model (SM) of particle physics elegantly describes three of the four fundamental interactions of the Universe. As a major triumph of this model, the last piece, i.e., the only scalar in the SM, the Higgs boson, was discovered by the ATLAS [1] and CMS [2] Collaborations in 2012. Certain drawbacks of the SM, either in terms of theoretical consistency or due to the lack of explanation for different fundamental observations, such as nonzero masses of the neutrinos, existence of the dark matter, matter antimatter asymmetry, etc., have prompted both theoretical as well as experimental communities to look further into the ideas and evidences beyond the Standard Model (BSM). Different cosmological observations, namely, galactic rotation curves, gravitational lensing, bullet cluster structure formation, etc., already establish the existence of some exotic nonluminous, weakly (to feebly) interacting matter, known as dark matter (DM) and has shown its existence almost at all length scales of the universe. Apart from the fact that DM has gravitational interaction, precise measurement of cosmic microwave background from the

WMAP [3] and PLANCK [4–6] data, has established the relic abundance corresponding to 26% of the present energy budget of our universe. However, any microscopic nature of the DM is yet unknown. Despite of considerable amount of effort by the astroparticle and high energy physics communities in order to detect and explain the microscopic nature of the DM, no experiment has so far been able to make any detection.

In the present study, we consider the inert Higgs doublet model (IDM) [7,8] as a prospective BSM scenario that gives a viable DM candidate. This model is renormalizable and is constructed by a simple particle extension of the SM containing an extra $SU(2)_L$ scalar doublet, which is odd under the discrete \mathbb{Z}_2 symmetry.¹ In contrast, all SM particles are even under the \mathbb{Z}_2 and therefore such a symmetry arrangement prevents any interaction between the SM fermions and BSM scalars—this stabilizes the lightest neutral scalar, which acts as a suitable DM candidate. The parameter space of this model is constrained [10–13] from the dark matter direct detection (DD) experiment, LHC data, and various other astronomical and cosmological observations. In its simplest form, the present model can satisfy the whole amount of observed relic density of the DM in some particular parameter space, the so-called *resonance region* and *degenerate region*. In the former case, the relic density of the DM is produced thermally through the resonant Higgs portal annihilation.

*anupam@prl.res.in

†konar@prl.res.in

‡seth@prl.res.in

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

¹Promoting such \mathbb{Z}_2 parity into a global U(1) symmetry is possible that provides a heavily constrained DM model with fewer model parameters [9].

Hence, the DM mass is required to be nearly half of the Higgs boson mass and other BSM scalars carry larger masses. This region is also known as the *hierarchical mass region* as DM is the lightest, while others are quite heavy. On the other hand, DM and all other BSM scalars are nearly of equal mass (~ 500 GeV or more) in the degenerate region [14–16]. As expected, this region is harder to probe at the LHC because of the kinematic suppression due to heavy final state production, narrow mass gap, and poor detection efficiency of the soft products coming from the decay of the BSM scalars. In spite of these complexities, this region is partially probed by the CMS Collaboration [17]. There are two additional regions: (i) heavy DM, $m_{\text{DM}} \sim 550$ GeV with the hierarchical mass spectrum where the mass difference, $\Delta M \sim \mathcal{O}(10 - 100)$ GeV between DM and other heavy BSM scalars, (ii) light DM, $m_{\text{DM}} < 80$ GeV with the degenerate mass spectrum, $\Delta M \sim \mathcal{O}(1 - 10)$ GeV. However, these two regions gather only a few percent of the observed relic density of the DM [18] and are hence less exciting. The light DM degenerate scenario is known to be satisfying only $\sim 10\%$ of the observed relic density [18] and that can be probed at the LHC with the monojet search [19]. Several other search strategies focused on different final states such as multilepton production with missing transverse energy [20–23], dilepton with dijet [22], and dijet with missing transverse momentum (\cancel{E}_T) [24].

In this study, we consider the hierarchical mass region that satisfies the total observed relic density of the DM and all other constraints. The significant mass difference between BSM scalars and the DM leads to a very interesting signal topology due to the boosted vector boson created through heavy scalar decay. We focus on the associated production and pair production of heavy scalars. Among heavy scalar pair production channels, pseudoscalar pair production is only possible through the Higgs boson mediator. Another channel (charged-scalar pair production) also gets contribution from the Higgs mediated diagrams. The major contribution to such production comes from the gluon fusion, which contains a loop at the leading order (LO). We work in the heavy top mass limit and that reduces the one loop diagram into an effective vertex. We consider $\mathcal{O}(\alpha_s)$ corrections to that effective term which is known to be as large as the LO alone. Therefore, the total Lagrangian is the sum of the IDM Lagrangian and the gluon-Higgs effective Lagrangian, and we consider $\mathcal{O}(\alpha_s)$ corrections to the total Lagrangian. The $pp \rightarrow HH$ process does not contribute to our phenomenological analysis, while loop induced $pp \rightarrow AA$ process has a minimal contribution compared to the other processes and we take this contribution into account in the heavy top mass limit. So even in the phase space regions where the heavy top mass limit is not a reasonable assumption, we restrict ourselves to this approximation as the effect coming from the above-mentioned channels are considerably smaller. This hierarchical mass spectrum of the IDM is hard to probe at the

LHC because of its tiny cross section over an immense SM background. We recognize in this study that the K factor due to NLO correction for the associated production of the heavy scalar ranges from 1.33 to 1.37. Vector boson mediated diproduction channels get corrections ranging from 1.35 to 1.56, while for the Higgs boson mediated ones it varies from 1.7 to 1.9. Correct understanding of signal distribution is important, where the signal is tiny compared to the background. These corrections give precise predictions improving the signal's statistical significance and lead us to change the analysis strategy based on modified signal output.

It was already demonstrated that the dijet + \cancel{E}_T signal could barely give 2σ statistical significance of the signal at high luminosity LHC [24]. The mono-fatjet signal was also studied in [25,26], which is not sufficient to achieve discovery potential. Hybrid topology (admixture of the mono-fatjet and di-fatjet+ \cancel{E}_T signal) was analyzed in [18]. Fatjets originate from boosted W^\pm/Z boson (denoted as J_V later on) after the decay of heavy BSM scalars and therefore possess a two-prong structure. Naturally, the pruned fatjet mass (M_J) and subjetteness (τ_{21}) become crucial variables to separate the tiny signal from the background. A sophisticated multivariate analysis (MVA) with jet-substructure variables is adopted in this analysis. It is demonstrated that di-fatjet detection from hybrid topology, together with the full potential of jet substructure variables, could effectively bring nearly the entire available parameter space well within reach of the 14 TeV HL-LHC. In MVA, we choose cuts on different variables optimally to increase the signal to background ratio and perform the analysis at NLO accuracy with jet substructure variables. Note that the di-fatjet signature relies upon a particular phase space region; hence the differential NLO K factor due to NLO correction plays a vital role to get more accurate signal efficiency. In this analysis, we find that NLO computation reduces the contribution in the di-fatjet final state arising from the associated production of the heavy scalar processes compared to the pair production channels, leading to rectified analysis strategies during MVA and that helps to reach a higher discovery potential of the IDM.

We organize the paper as follows: Sec. II briefly describes the IDM model and the Higgs-gluon effective Lagrangian that we adopt in this computation. Section III points out various constraints on the IDM model and lists benchmark points accordingly in the hierarchical mass region. In Sec. IV, we mainly discuss the computational setup and show numerical results including the differential NLO K factor and scale uncertainties. Section V presents the distributions of different high-level kinematical variables involving jets at LO and NLO for the associated and pair production channels, demonstrating the importance of the QCD corrections. Section VI explains the reason to consider $2J_V + \cancel{E}_T$ as the signal while dealing with a tiny IDM signal over an immense background. We also discuss here the

MVA, which uses a highly non-linear cut, and use the full potential of NLO computation and jet-substructure variables to separate this tiny signal from the large background. Finally, we conclude in Sec. VII.

II. THEORETICAL FRAMEWORK

IDM has a new $SU(2)_L$ doublet Φ_2 in addition to the SM Higgs doublet, Φ_1 , and a discrete \mathbb{Z}_2 symmetry is being imposed on it. All the fields of the SM are even under \mathbb{Z}_2 transformations. Φ_2 is odd under \mathbb{Z}_2 transformation and therefore the inert doublet cannot acquire vacuum expectation value (vev), as vev cannot change sign under any internal symmetry. As Φ_2 has no vev, we can write this doublet in terms of physical fields. \mathbb{Z}_2 symmetry also prevents the interaction between inert scalars and SM fermions at any order in the perturbation series, aiding the lightest inert neutral scalar to act as a dark matter. The doublet, Φ_2 , has hypercharge $Y = \frac{1}{2}$, which is equal to the hypercharge of Φ_1 . These two doublets can be written in the unitary gauge as

$$\Phi_1 = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}(v + h + iG^0) \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} H^+ \\ \frac{H+iA}{\sqrt{2}} \end{pmatrix}, \quad (1)$$

where G^+ and G^0 are the Goldstone bosons and the vev $v = 246$ GeV. H^+ is the charged BSM scalars. H and A are both neutral scalars; one is CP even, and the other is CP odd. Note that CP properties of the neutral scalars are basis dependent. The most general potential [11] can be written as

$$\begin{aligned} V_{\text{IDM}} = & \mu_1^2 \Phi_1^\dagger \Phi_1 + \mu_2^2 \Phi_2^\dagger \Phi_2 + \frac{\lambda_1}{2} (\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2} (\Phi_2^\dagger \Phi_2)^2 \\ & + \lambda_3 (\Phi_1^\dagger \Phi_1) (\Phi_2^\dagger \Phi_2) + \lambda_4 (\Phi_2^\dagger \Phi_1) (\Phi_1^\dagger \Phi_2) \\ & + \frac{\lambda_5}{2} [(\Phi_1^\dagger \Phi_2)^2 + (\Phi_2^\dagger \Phi_1)^2]. \end{aligned} \quad (2)$$

After electroweak symmetry breaking through SM Higgs doublet, Φ_1 , the masses of the BSM scalars at the tree level can be expressed as

$$\begin{aligned} m_h^2 &= \lambda_1 v^2, & m_{H^\pm}^2 &= \mu_2^2 + \frac{1}{2} \lambda_3 v^2, \\ m_A^2 &= \mu_2^2 + \frac{1}{2} \lambda_c v^2, & m_H^2 &= \mu_2^2 + \frac{1}{2} \lambda_L v^2. \end{aligned} \quad (3)$$

All free parameters are real, so the scalar sector does not contain any CP violations and $\lambda_{L/c} = (\lambda_3 + \lambda_4 \pm \lambda_5)$. Higgs portal coupling λ_L , which can be positive or negative, plays an important role in the DM sector as it determines the annihilation rate of the DM in the hierarchical mass region. m_h is the SM Higgs boson mass, and $m_{H^\pm, A, H}$ are the masses of the BSM scalars. The parameters

λ_1 and μ_1 can be written in terms of the mass of the Higgs boson and vev. So, IDM has five parameters—three masses of the inert scalars, self-coupling between inert scalars λ_2 and Higgs portal coupling λ_L . Self-coupling λ_2 does not affect the scalar masses and their phenomenology. In our study, we choose the inert scalar H as the dark matter candidate, but one can also choose the A as the dark matter without changing any phenomenology, just by flipping the sign of λ_5 preserving the CP properties of the DM candidate. The full IDM Lagrangian can be written as

$$\mathcal{L}_{\text{IDM}} = \mathcal{L}_{\text{SM}} + (\mathcal{D}_\mu \Phi_2)^\dagger (\mathcal{D}^\mu \Phi_2) + V_{\text{IDM}}, \quad (4)$$

where the covariant derivative, $\mathcal{D}_\mu = (\partial_\mu - ig_Y Y B_\mu - ig \frac{\sigma^i}{2} W_\mu^i)$, and σ^i are the Pauli matrices; g and g_Y are the coupling strength of the weak and hypercharge interactions, respectively. In addition, we consider the following five-dimensional effective term to take into account Higgs interactions with gluons in the heavy top mass limit,

$$\mathcal{L}_{\text{HEFT}} = -\frac{1}{4} C_{\text{eff}} h G_{\mu\nu}^a G^{a\mu\nu}. \quad (5)$$

Here, $G_{\mu\nu}^a$ represents QCD field strength tensor and $C_{\text{eff}} = \frac{\alpha_s}{3\pi v} (1 + \frac{11}{4} \frac{\alpha_s}{\pi}) = C_0 (1 + \frac{11}{4} \frac{\alpha_s}{\pi})$ contains terms up to $\mathcal{O}(\alpha_s^2)$, that basically takes part in the one loop corrected amplitude for Higgs boson mediated production channels.

III. CONSTRAINTS AND BENCHMARK POINTS

The parameter space of the IDM is very constrained from theoretical calculations, various experimental data and cosmological observations. We briefly demonstrate all these constraints and then set few benchmark points that will cover almost the entire hierarchical region of the IDM. Further details are provided in [11,18].

The potential must be bounded from below for any realistic model, and the vacuum should be neutral, which leads to the following constraint [18]:

$$\begin{aligned} \lambda_1 &> 0, & \lambda_2 &> 0, & \lambda_3 + 2\sqrt{\lambda_1 \lambda_2} &> 0, \\ \lambda_3 + \lambda_4 + \lambda_5 + 2\sqrt{\lambda_1 \lambda_2} &> 0. \end{aligned} \quad (6)$$

The condition $\lambda_4 + \lambda_5 < 0$ ensures the inert vacuum to be charge neutral. Generically, depending on the nature of additionally imposed symmetry, the electroweak symmetry breaking pattern has the following possibilities:

$$\begin{aligned} v_1 &= v, & v_2 &= 0, & \text{inert vacuum,} \\ v_1 &= 0, & v_2 &= v, & \text{pseudo-inert vacuum,} \\ v_1 &\neq 0, & v_2 &\neq 0, & \text{mixed vacuum,} \end{aligned} \quad (7)$$

where v_1 denotes the vev of the doublet Φ_1 , and v_2 is the vev of the Φ_2 . v is the electroweak scale,

$(G_F\sqrt{2})^{-1/2} = 246$ GeV. We want the inert vacuum as the global minima, which leads to the following constraint [27,28]:

$$\frac{\mu_1^2}{\sqrt{\lambda_1}} - \frac{\mu_2^2}{\sqrt{\lambda_2}} > 0. \quad (8)$$

The eigenvalues of the $2 \rightarrow 2$ scalar scattering processes [29] are given in Eq. (9), and each eigenvalue ($|\Lambda_i|$) should be $\leq 8\pi$, coming from the perturbativity and unitarity constraints:

$$\begin{aligned} \Lambda_{1,2} &= \lambda_3 \pm \lambda_4, & \Lambda_{3,4} &= \lambda_3 \pm \lambda_5, & \Lambda_{5,6} &= \lambda_3 + 2\lambda_4 \pm 3\lambda_5, \\ \Lambda_{7,8} &= -\lambda_1 - \lambda_2 \pm \sqrt{(\lambda_1 - \lambda_2)^2 + \lambda_4^2}, \\ \Lambda_{9,10} &= -3\lambda_1 - 3\lambda_2 \pm \sqrt{9(\lambda_1 - \lambda_2)^2 + (2\lambda_3 + \lambda_4)^2}, \\ \Lambda_{11,12} &= -\lambda_1 - \lambda_2 \pm \sqrt{(\lambda_1 - \lambda_2)^2 + \lambda_5^2}. \end{aligned} \quad (9)$$

The contribution that navigates from the BSM physics to the electroweak radiative correction is parametrized by the S, T, U parameters [30], known as oblique parameters. The central values of the oblique parameters that we use in our analysis are [31]

$$S = 0.04 \pm 0.11, \quad T = 0.09 \pm 0.14, \quad U = -0.02 \pm 0.11. \quad (10)$$

The following parameter space of the IDM is ruled out from the neutralino search results at LEP-II [32,33]:

$$m_H < 80 \text{ GeV}, \quad m_A < 100 \text{ GeV}, \quad \text{and} \quad (m_A - m_H) > 8 \text{ GeV}. \quad (11)$$

The charged Higgs mass gets the following bound from the chargino search results at LEP-II [34]:

$$m_{H^\pm} > 70 \text{ GeV}. \quad (12)$$

More recently, analyzing a pair of boosted hadronically decaying bosons together with MET from 13 TeV LHC data, ATLAS gave constraints on the masses of the charginos and neutralinos of the minimal supersymmetric model [35]. Based on a similar production mechanism from IDM, Ref. [36] carried out a recasting analysis to show that the Higgs portal DM scenario and hierarchical heavy scalars of mass 123 GeV or above are allowed from this exclusion limit.

In the hierarchical region, the decay channels, $\Gamma(Z \rightarrow HA, H^+H^-)$ and $\Gamma(W^\pm \rightarrow H^\pm A, H^\pm H)$ are kinematically forbidden. The signal strength of the Higgs boson decay into the diphoton final state relative to SM prediction is [37–39]

$$\mu_{\gamma\gamma} = \frac{\sigma(pp \rightarrow h \rightarrow \gamma\gamma)}{\sigma(pp \rightarrow h \rightarrow \gamma\gamma)_{\text{SM}}} = 1.10_{-0.09}^{+0.10}. \quad (13)$$

The Higgs boson production rate is the same in both the SM and IDM models, dominated by the gluon gluon fusion channel, and so the signal strength turns out to be

$$\mu_{\gamma\gamma} = \frac{\text{BR}(h \rightarrow \gamma\gamma)_{\text{IDM}}}{\text{BR}(h \rightarrow \gamma\gamma)_{\text{SM}}}. \quad (14)$$

A sufficiently large value of the λ_3 coupling and lighter charged Higgs mass can lead to enhanced decay of $h \rightarrow \gamma\gamma$, thereby pushing the ratio beyond the experimental limit and hence excluded. The upper limit of the Higgs invisible decay branching ratio measured by the ATLAS Collaboration [40] is 0.11 at 95% confidence level (CL). This measurement puts stringent constraints on the Higgs portal coupling (λ_L) and DM mass (m_H) in the region $m_H < \frac{m_h}{2}$. The Higgs invisible decay width in the IDM model is given by

$$\Gamma_{h \rightarrow HH} = \frac{\lambda_L^2 v^2}{64\pi m_h} \sqrt{1 - \frac{4m_H^2}{m_h^2}}. \quad (15)$$

$\Gamma_{h \rightarrow HH}/(\Gamma_{\text{SM}} + \Gamma_{h \rightarrow HH}) \leq 0.11$ must be satisfied in the kinematically allowed region of the decay of the Higgs boson into pair of the DM. Moreover, extremely precise measurements from WMAP [3] and PLANCK [4–6] have established that the relic abundance of the DM is $\Omega_{\text{DM}} h^2 = 0.120 \pm 0.001$ [6] with $h = \frac{\text{Hubble Parameter}}{(100 \text{ km s}^{-1} \text{ Mpc}^{-1})}$. The dark matter annihilates into SM particles and the relic density of the DM is inversely proportional to this annihilation rate. The observed relic density of DM sets a rigid constraint on the parameter spaces of the IDM so as not to over-produce the relic in the IDM. The spin-independent cross section of the DM-nucleon scattering processes at leading order mediated by the Higgs boson is given by [7]

$$\sigma = \frac{\lambda_L^2 f^2}{4\pi} \frac{\mu^2 m_n^2}{m_h^4 m_{\text{DM}}^2}, \quad (16)$$

where m_n is the mass of the nucleon and $\mu = \frac{m_n m_{\text{DM}}}{m_n + m_{\text{DM}}}$. f is the Higgs-nucleon coupling strength and the allowed range of f is 0.26–0.63 [41]. However, the recent study suggests the value of f is 0.32 [42]. The upper bound of the DM-nucleon scattering cross section from the DM DD experiments like LUX [43] and XENON1T [44] poses a firm limit on the allowed values of λ_L . As already stated, we can divide the entire parameter space of the IDM into four distinct regions depending on the mass of the DM and the mass splitting between DM and other scalars—among these four, only the following two regions satisfy the observed relic density of the DM entirely.

TABLE I. Input parameters, masses of the BSM scalars (m_{H^\pm}, m_A), and the self-coupling constant (λ_2) between dark sector particles for several selected benchmark points that satisfy theoretical, DM relic density, DD data, and collider constraints listed in the text. Three other parameters are DM mass, $m_H = 53.71$ GeV, Higgs portal coupling, $\lambda_L = 5.4 \times 10^{-3}$, and Higgs boson mass $m_h = 125$ GeV.

Input parameters	BP1	BP2	BP3	BP4	BP5	BP6	BP7
m_{H^\pm} (GeV)	255.3	304.8	350.3	395.8	446.9	503.3	551.8
m_A (GeV)	253.9	302.9	347.4	395.1	442.4	500.7	549.63
λ_2	1.27	1.07	0.135	0.106	3.10	0.693	0.285

The hierarchical mass region consists of a Higgs portal mass region with $m_{\text{DM}} \equiv m_H < 80$ GeV, and the mass gap with other BSM scalars as, $\Delta M \equiv \Delta M_{\text{charged}} \simeq \Delta M_{\text{neutral}} \sim 100$ GeV or more, where $\Delta M_{\text{charged}} = (m_{H^\pm} - m_{\text{DM}})$ and $\Delta M_{\text{neutral}} = (m_A - m_{\text{DM}})$. In this region, no bound on the DM mass comes from the LEP Z-boson width measurements. Since the DM mass is less than 80 GeV, the annihilation of the DM into the pair of weak gauge bosons is significantly suppressed. In this region, relic density of DM is achieved only through the Higgs portal annihilation channel. Since the mass differences between DM and other BSM scalars are significant, the co-annihilation effects are absent. As the annihilation cross section is proportional to λ_L , any small value of λ_L leads to overproduction of relic density. We get the total observed relic density of the DM in the range where the DM mass varies between 53 and 70 GeV for substantial λ_L values, constrained from DD of DM.

The degenerate mass region consists of high mass region, $m_{\text{DM}} \geq 500$ GeV, with rather tiny mass gap $\Delta M \sim 1$ GeV. In this regime, the following annihilation and co-annihilation processes open up:

$$\begin{aligned}
 &\text{annihilation} \begin{cases} HH \rightarrow W^+ W^- \\ HH \rightarrow ZZ \end{cases} \quad \lambda_L \text{ sensitive,} \\
 &\text{co-annihilation} \begin{cases} H^+ H^- \rightarrow W^+ W^- \\ AA \rightarrow W^+ W^- \\ AA \rightarrow ZZ \end{cases} \quad \lambda_L \text{ sensitive,} \\
 &\text{co-annihilation} \begin{cases} H^\pm H \rightarrow W^\pm \gamma \\ H^\pm A \rightarrow W^\pm \gamma \end{cases} \quad \text{gauge couplings.} \quad (17)
 \end{aligned}$$

The quartic coupling between DM and the longitudinal gauge bosons in the annihilation processes $HH \rightarrow W_L^+ W_L^-$ and $HH \rightarrow Z_L Z_L$ is $(4m_{\text{DM}}\Delta M/v^2 + \lambda_L)$. In this degenerate mass spectrum $\Delta M \rightarrow 0$, and so this coupling remains sensitive to λ_L mostly. The relic density of DM increases with the DM mass and decreases with the annihilation cross section. Those combined effects set the correct relic density of DM in this region for $m_{\text{DM}} \geq 500$ GeV. Although this region is difficult to probe, with a charged long-lived Higgs boson, one can explore this region at LHC with the charged track signal [45].

The different benchmark points that we pick out for this study are given in Table I, and all of them satisfy the constraints discussed above.

IV. COMPUTATIONAL SETUP AND NUMERICAL RESULTS

We implement the Lagrangian given in Eq. (4) together with the leading term of Eq. (5) in FeynRules [46] and employ NLOCT [47] to generate UV and R_2 counterterms of the SM Lagrangian in order to have a NLO UFO model that we use under the MG5_aMC@NLO environment [48]. Inside this environment, real corrections are performed following the FKS subtraction method [49], whereas the OPP technique [50] is the one that is being used to take care of the virtual contributions. Nevertheless, for AA , HH , and $H^+ H^-$ pair production processes, gluon-gluon initiated processes mediated by the Higgs propagator play a significant role and we insert the corresponding analytic form of the one loop amplitude in MadGraph5 virtual routine and that in the $d = (4 - 2\epsilon)$ dimension reads as

$$\begin{aligned}
 \overline{\mathcal{R}(M_0 M_v^\dagger)} &= \left(\frac{\alpha_s}{2\pi}\right) \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)} \left(\frac{\mu^2}{s_{12}}\right)^\epsilon |\overline{M_0}|^2 \\
 &\times \left[-\frac{6}{\epsilon^2} - \frac{2b_0}{\epsilon} + 11 + 3\pi^2 \right], \quad (18)
 \end{aligned}$$

while setting the renormalization scale $\mu^2 = s_{12}$, the partonic center-of-mass (c.m.) energy. M_0 and M_v represent tree-level and one-loop amplitudes, respectively. The leading term of the QCD β function $b_0 = \frac{11}{6} C_A - \frac{2}{3} n_f T_F$, where n_f represents the number of active quark flavors and $C_A = 3$, $T_F = 1/2$. Note that the strong coupling is renormalized following the $\overline{\text{MS}}$ scheme and the $\mathcal{O}(\alpha_s^2)$ term of the Lagrangian given in Eq. (5) is taken into account in the above expression. The color and spin averaged tree level squared amplitude in the $d = (4 - 2\epsilon)$ dimension can be written as

$$|\overline{M_0}|^2 = \frac{1}{128} (1 + \epsilon + \epsilon^2) \frac{C_0 \Lambda^2 v^2 s_{12}^2}{(s_{12} - m_h^2)^2 + \Gamma_h^2}. \quad (19)$$

Here $C_0 = \frac{\alpha_s}{3\pi v}$, Γ_h is the Higgs boson width, and Λ corresponds to $\Lambda_{L/c/3}$ as given in the Feynman rules

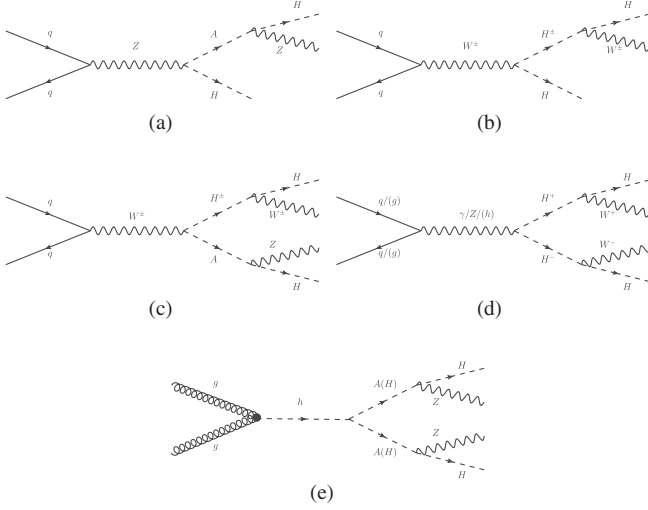


FIG. 1. Parton level representative diagrams at LO of (a),(b) associate production of heavy scalar, and (c),(d),(e) pair production of heavy scalars. In our study, we consider one loop correction in α_s of all these diagrams.

furnished in the Appendix. Final state heavy scalar particles are decayed via MadSpin [51], which retains spin information at the tree level accuracy. NLO events thus obtained are then matched to the PYTHIA8 [52,53] parton shower following the MC@NLO formalism [54] to avoid any double counting. For the signal, we use in-built NN23LO1 and NN23NLO PDF sets [55] for LO and NLO, respectively. We use DELPHES3 [56] to include the detector effects in our simulation, where we use the default card of the CMS. Jets are formed by clustering the particle-flow tower objects and particle-flow tracks. We employ an anti- k_T [57] clustering algorithm to form jets, where we have set radius parameter $R = 0.5$. Using FastJet 3.2.2 [58] package, we reconstruct fatjets, utilizing DELPHES tower objects as input for clustering. The Cambridge-Aachen (CA) [59] algorithm is hired for

fatjets clustering where the radius parameter is set to $R = 0.8$. Fatjets are characterized by the radius parameter, $R \sim 2m_V/P_T$ ($V \equiv \{W^\pm, Z\}$), where P_T is the transverse momentum and m_V is the mass of the weak boson. We apply minimum $P_T = 180$ GeV for each FatJet formation. MVA analysis is done in the TMVA framework [60]. We implement the boosted decision tree (BDT) algorithm in our MVA analysis. A decision tree splits the high-level input data recursively depending on a set of input features. The method that combines many trees (weak learners) into a strong classifier is called boosting. Figure 1 displays representative LO Feynman diagrams of the associated production of the heavy scalar and pair production of the heavy scalars, which we ultimately decay hadronically.

Production cross sections for these channels before hadronic decay of the heavy scalars are given in Table II, and Table III at 14 TeV LHC. We choose the renormalization scale and the factorization scale as $\mu_R = \zeta_R \sqrt{s_{12}}$ and $\mu_F = \zeta_F \sqrt{s_{12}}$, respectively, where $\zeta_R = \zeta_F = 1$ represents the central scale choice. We vary $\zeta_R, \zeta_F = \{1/2, 1, 2\}$, which has a total of nine datasets. All the cross sections are given corresponding to the central scale where superscripts and subscripts denote the envelope of those nine scale choices. The Monte Carlo uncertainties are also given in Table II, and Table III. We get reduced scale uncertainty in the total cross section at NLO than LO for both the associated and pair production processes, except in a few benchmark points for the associated production processes and the reason could be the crossover of the envelopes around the maximum differential LO cross section, unlike NLO [see Figs. 3(a), bottom; 3(b), bottom]. Fractional scale uncertainty is defined as the envelope of the ratios of the differential cross sections at eight additional (ζ_R, ζ_F) choices to the central one. Dashed and solid lines in the fractional scale uncertainty subplot correspond to the lower and upper envelope, respectively. Our study includes one

TABLE II. Cross sections for the associated production of heavy scalar at LO and NLO with integrated K factor are given in this table at 14 TeV LHC before the decay of heavy scalars into DM and SM particles. The superscript and subscript denote the scale uncertainties in the total cross section (the percentages are given in bracket), while the last entry is the Monte Carlo uncertainty. AH and $H^\pm H$ channels are produced in five and four massless quark flavors, respectively.

BP	$\sigma(pp \rightarrow AH)$ (fb)			K-fac	$\sigma(pp \rightarrow H^\pm H)$ (fb)		
	NLO	LO			NLO	LO	
BP1	$46.55^{+0.79(1.7\%)}_{-0.65(1.4\%)} \pm 8.0 \times 10^{-2}$	$35.12^{+0.04(0.1\%)}_{-0.14(0.4\%)} \pm 1.3 \times 10^{-1}$	1.33	$82.42^{+1.26(1.5\%)}_{-0.98(1.2\%)} \pm 2.0 \times 10^{-2}$	$61.67^{+0.12(0.2\%)}_{-0.37(0.6\%)} \pm 2.2 \times 10^{-1}$	1.34	
BP2	$25.34^{+0.40(1.6\%)}_{-0.29(1.1\%)} \pm 4.0 \times 10^{-2}$	$19.01^{+0.17(0.9\%)}_{-0.23(1.2\%)} \pm 7.0 \times 10^{-2}$	1.33	$45.27^{+0.87(1.9\%)}_{-0.40(0.9\%)} \pm 8.0 \times 10^{-2}$	$33.77^{+0.37(1.1\%)}_{-0.47(1.4\%)} \pm 1.0 \times 10^{-2}$	1.34	
BP3	$15.50^{+0.21(1.4\%)}_{-0.21(1.4\%)} \pm 3.0 \times 10^{-2}$	$11.60^{+0.17(1.5\%)}_{-0.21(1.8\%)} \pm 4.0 \times 10^{-2}$	1.34	$27.78^{+0.44(1.6\%)}_{-0.3(1.1\%)} \pm 5.0 \times 10^{-2}$	$20.69^{+0.37(1.8\%)}_{-0.41(2\%)} \pm 8.0 \times 10^{-2}$	1.34	
BP4	$9.68^{+0.15(1.6\%)}_{-0.14(1.4\%)} \pm 2.0 \times 10^{-2}$	$7.19^{+0.16(2.2\%)}_{-0.17(2.3\%)} \pm 3.0 \times 10^{-2}$	1.35	$17.86^{+0.27(1.5\%)}_{-0.29(1.6\%)} \pm 3.0 \times 10^{-2}$	$13.23^{+0.32(2.4\%)}_{-0.33(2.5\%)} \pm 4.0 \times 10^{-2}$	1.35	
BP5	$6.32^{+0.11(1.7\%)}_{-0.10(1.6\%)} \pm 1.0 \times 10^{-2}$	$4.67^{+0.13(2.8\%)}_{-0.13(2.8\%)} \pm 2.0 \times 10^{-2}$	1.35	$11.34^{+0.2(1.8\%)}_{-0.19(1.7\%)} \pm 2.0 \times 10^{-2}$	$8.39^{+0.25(3\%)}_{-0.25(3\%)} \pm 3.0 \times 10^{-2}$	1.35	
BP6	$3.90^{+0.07(1.8\%)}_{-0.07(1.8\%)} \pm 6.5 \times 10^{-3}$	$2.87^{+0.10(3.5\%)}_{-0.10(3.5\%)} \pm 1.0 \times 10^{-2}$	1.36	$7.19^{+0.14(1.9\%)}_{-0.13(1.8\%)} \pm 1.0 \times 10^{-2}$	$5.33^{+0.19(3.6\%)}_{-0.19(3.6\%)} \pm 2.0 \times 10^{-2}$	1.35	
BP7	$2.69^{+0.05(1.9\%)}_{-0.05(1.9\%)} \pm 4.6 \times 10^{-3}$	$1.97^{+0.08(4\%)}_{-0.07(3.6\%)} \pm 7.3 \times 10^{-3}$	1.37	$5.01^{+0.1(2\%)}_{-0.1(2\%)} \pm 8.9 \times 10^{-3}$	$3.68^{+0.15(4\%)}_{-0.15(4\%)} \pm 1.0 \times 10^{-2}$	1.36	

TABLE III. Same as Table II but for the pair production of the heavy scalars. $H^\pm A$ is produced in four massless quark flavors, while H^+H^- and the AA channels are produced in five massless quark flavors.

BP	$\sigma(pp \rightarrow H^\pm A)$ (fb)			$\sigma(pp \rightarrow H^+H^-)$ (fb)		
	NLO	LO	K-fac	NLO	LO	K-fac
BP1	$16.93^{+0.28(1.7\%)}_{-0.25(1.5\%)} \pm 3.0 \times 10^{-2}$	$12.53^{+0.34(2.7\%)}_{-0.34(2.7\%)} \pm 5.0 \times 10^{-2}$	1.35	$11.01^{+0.43(3.9\%)}_{-0.42(3.8\%)} \pm 2.0 \times 10^{-2}$	$7.98^{+0.48(6\%)}_{-0.4(5\%)} \pm 2.0 \times 10^{-2}$	1.38
BP2	$8.41^{+0.14(1.7\%)}_{-0.18(2.1\%)} \pm 2.0 \times 10^{-2}$	$6.21^{+0.23(3.7\%)}_{-0.22(3.5\%)} \pm 3.0 \times 10^{-2}$	1.35	$6.01^{+0.37(6.2\%)}_{-0.32(5.3\%)} \pm 8.5 \times 10^{-3}$	$4.17^{+0.37(8.9\%)}_{-0.29(7\%)} \pm 1.0 \times 10^{-2}$	1.44
BP3	$4.78^{+0.1(2.1\%)}_{-0.1(2.1\%)} \pm 8.3 \times 10^{-3}$	$3.48^{+0.15(4.3\%)}_{-0.14(4\%)} \pm 1.0 \times 10^{-2}$	1.37	$3.76^{+0.29(7.7\%)}_{-0.24(6.4\%)} \pm 5.8 \times 10^{-3}$	$2.57^{+0.29(11.3\%)}_{-0.23(8.9\%)} \pm 7.6 \times 10^{-3}$	1.46
BP4	$2.81^{+0.06(2.1\%)}_{-0.06(2.1\%)} \pm 5.0 \times 10^{-3}$	$2.04^{+0.1(4.9\%)}_{-0.1(4.9\%)} \pm 7.8 \times 10^{-3}$	1.38	$2.50^{+0.23(9.2\%)}_{-0.19(7.6\%)} \pm 4.1 \times 10^{-3}$	$1.68^{+0.24(14.3\%)}_{-0.18(10.7\%)} \pm 5.3 \times 10^{-3}$	1.49
BP5	$1.69^{+0.04(2.4\%)}_{-0.04(2.4\%)} \pm 3.0 \times 10^{-3}$	$1.22^{+0.07(5.7\%)}_{-0.06(4.9\%)} \pm 4.7 \times 10^{-3}$	1.38	$1.68^{+0.18(10.7\%)}_{-0.15(8.9\%)} \pm 3.0 \times 10^{-3}$	$1.10^{+0.19(17.3\%)}_{-0.14(12.7\%)} \pm 3.1 \times 10^{-3}$	1.52
BP6	$0.97^{+0.02(2.1\%)}_{-0.02(2.1\%)} \pm 1.7 \times 10^{-3}$	$0.70^{+0.05(7.1\%)}_{-0.04(5.7\%)} \pm 2.4 \times 10^{-3}$	1.38	$1.14^{+0.14(12.3\%)}_{-0.12(10.5\%)} \pm 2.4 \times 10^{-3}$	$0.74^{+0.15(20.3\%)}_{-0.11(14.9\%)} \pm 2.2 \times 10^{-3}$	1.54
BP7	$0.63^{+0.018(2.8\%)}_{-0.018(2.8\%)} \pm 1.1 \times 10^{-3}$	$0.45^{+0.03(6.7\%)}_{-0.03(6.7\%)} \pm 1.6 \times 10^{-3}$	1.39	$0.85^{+0.11(12.9\%)}_{-0.09(10.6\%)} \pm 1.9 \times 10^{-3}$	$0.55^{+0.13(23.6\%)}_{-0.09(16.4\%)} \pm 1.7 \times 10^{-3}$	1.56

BP	$\sigma(pp \rightarrow AA)$ (fb)		
	NLO	LO	K-fac
BP1	$0.88^{+0.18(20.4\%)}_{-0.14(15.9\%)} \pm 3.3 \times 10^{-3}$	$0.46^{+0.15(32.6\%)}_{-0.11(23.9\%)} \pm 1.6 \times 10^{-3}$	1.92
BP2	$0.72^{+0.14(19.4\%)}_{-0.12(16.7\%)} \pm 2.6 \times 10^{-3}$	$0.38^{+0.13(34.2\%)}_{-0.09(23.7\%)} \pm 1.4 \times 10^{-3}$	1.87
BP3	$0.59^{+0.12(20.3\%)}_{-0.10(16.9\%)} \pm 2.2 \times 10^{-3}$	$0.32^{+0.11(34.3\%)}_{-0.08(25\%)} \pm 1.1 \times 10^{-3}$	1.86
BP4	$0.48^{+0.1(20.8\%)}_{-0.08(16.7\%)} \pm 1.7 \times 10^{-3}$	$0.27^{+0.09(33.3\%)}_{-0.07(25.9\%)} \pm 9.6 \times 10^{-4}$	1.80
BP5	$0.40^{+0.08(20\%)}_{-0.07(17.5\%)} \pm 1.7 \times 10^{-3}$	$0.22^{+0.08(36.4\%)}_{-0.05(22.7\%)} \pm 8 \times 10^{-4}$	1.78
BP6	$0.31^{+0.06(19.4\%)}_{-0.05(16.1\%)} \pm 1.1 \times 10^{-3}$	$0.18^{+0.06(33.3\%)}_{-0.05(27.8\%)} \pm 6.5 \times 10^{-4}$	1.75
BP7	$0.26^{+0.05(19.2\%)}_{-0.05(19.2\%)} \pm 9.3 \times 10^{-4}$	$0.15^{+0.05(33.3\%)}_{-0.04(26.7\%)} \pm 5.4 \times 10^{-4}$	1.70

order in α_S corrections to all these channels. The cross section of the $pp \rightarrow HH$ channel at LO is 0.332 pb and at NLO it is 0.617 pb (i.e., K factor = 1.858) at 14 TeV LHC, independent of benchmark points since the cross section depends only on m_H and λ_L , and both remain the same for chosen benchmark points. This channel has a larger cross section than any other pair or associated production channels because of being less s -channel suppressed due to the presence of an on-shell Higgs boson mediator. Total transverse momentum distribution of the DM pair for BP2 of the channel, $pp \rightarrow HH$ is shown in Fig. 2 for fixed order NLO (dashed blue) and NLO matched with a parton shower (solid red). It is clear from this figure that NLO + PS describes the low P_T region more vividly compared to a fixed order estimation. Note that, although such calculation is essential for a traditional monojet search, possible contributions of $pp \rightarrow HH$ can only come in our di-fatjet study at the NNLO level. Characteristically, this process is background like and we find that much of the events will not pass the event selection criteria even while starting from a reasonably significant contribution. This channel is shown here for completeness, but we would not add such a contribution to our conservative estimate.

In the subsequent figures, on the left panel, we show the improvement in NLO + PS results over the LO + PS ones

on the invariant mass distribution (top) along with the differential K factor (middle) and fractional scale uncertainties (bottom) for all remaining production channels. The differential K factor is vital in extracting the correct signal

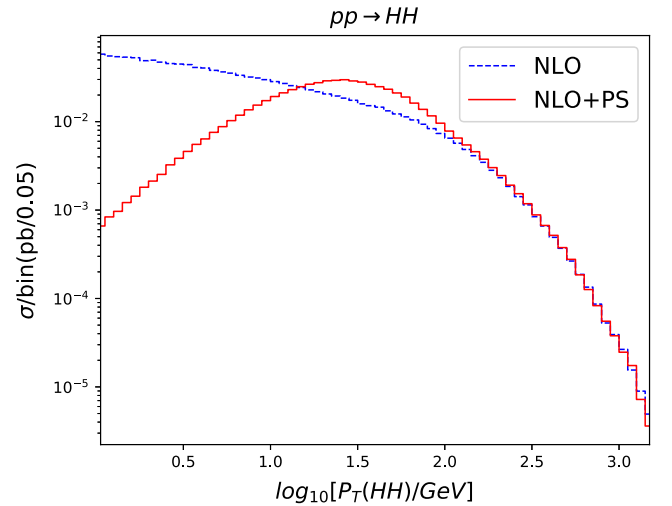


FIG. 2. Differential distribution of the total transverse momentum of the DM pair for the channel $pp \rightarrow HH$ at fixed order NLO (dashed blue) and NLO + PS (solid red) accuracy.

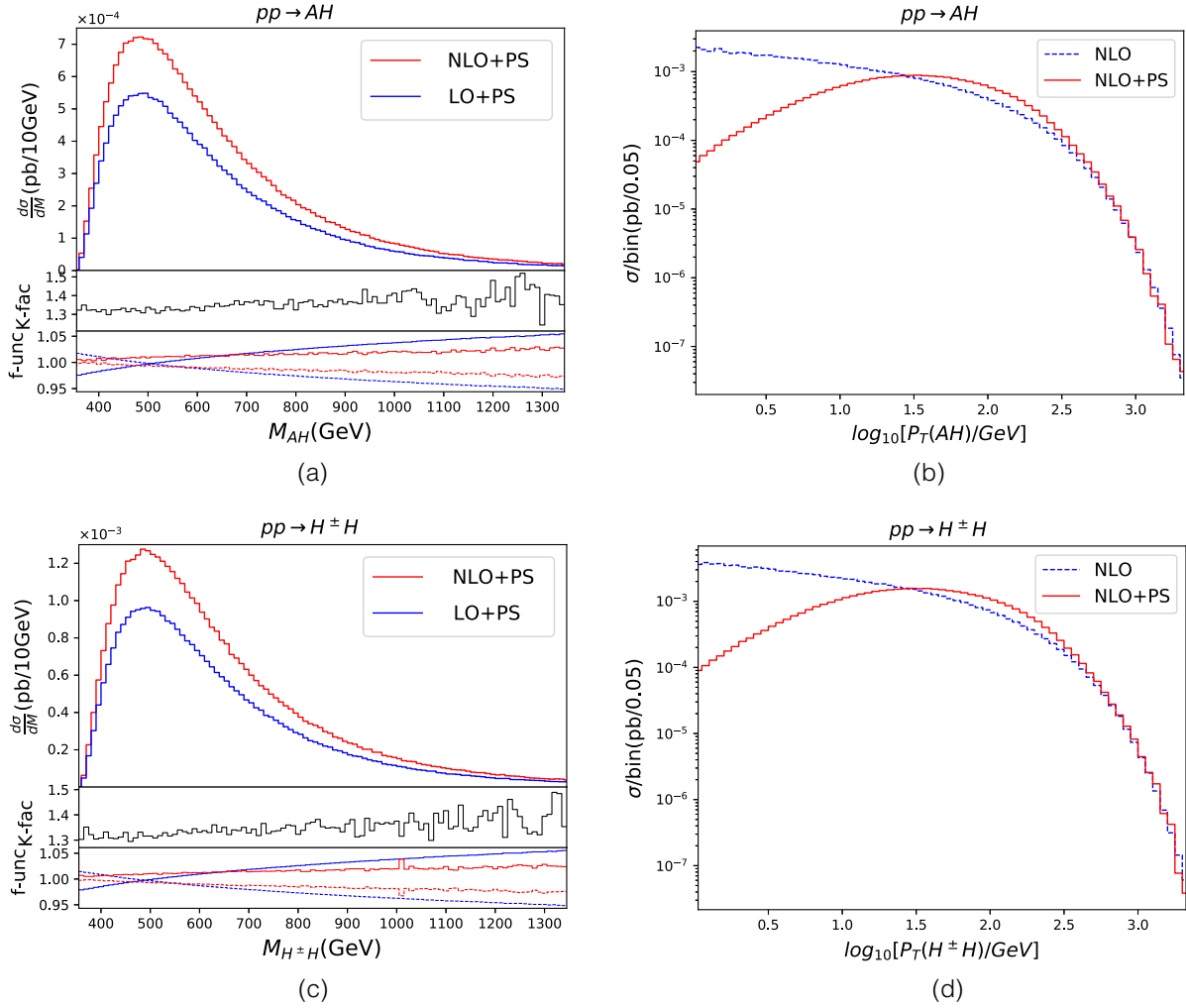


FIG. 3. NLO effects on the associated production of heavy scalar channels, such as $pp \rightarrow AH$ [(a),(b)], $pp \rightarrow H^\pm H$ [(c),(d)]. In each plot of the left panel, the top subplot shows the invariant mass distribution of the heavy scalar and DM pair at LO + PS (dashed blue) and NLO + PS (solid red) accuracy. The middle subplot displays the differential NLO K factor, the ratio of the NLO + PS cross section to the LO + PS one in each bin, while the bottom subplot presents the scale uncertainties for LO + PS (blue) and NLO + PS (red). The right panel shows the differential distribution of the total transverse momentum of the heavy scalar and DM pair for the respective channel at fixed order NLO (dashed blue) and NLO + PS (solid red) accuracy. All distributions are given for sample benchmark point BP2.

efficiency, as most collider analyses usually do not cover the entire phase space and apply various kinematical cuts to distinguish signal from the background. Fractional scale uncertainty denotes how stable the NLO result is as compared to the LO under scale variation. On the right panel, Sudakov suppression due to NLO + PS computation is explicitly shown for each corresponding channel and that ensures resummation of large logarithm terms in the low P_T region because of incorporating the parton shower effect on top of the fixed order calculation. Note that, in these sets of representative figures, hadronic decays of final state heavy scalars are not considered for the time being. Figure 3 collects all the associated production channels of heavy scalars, whereas Fig. 4 contains various pair production channels of heavy scalars. In all these figures, BP2 is

considered as the representative benchmark point. The invariant mass distributions for the associated production channels peak around the same region, close to 485 GeV for both $pp \rightarrow AH$ [Fig. 3(a), top] and $pp \rightarrow H^\pm H$ [Fig. 3(c), top]. However, among the pair production channels, vector boson mediated processes viz. $pp \rightarrow H^\pm A$ [Fig. 4(a), top] and $pp \rightarrow H^+ H^-$ [Fig. 4(c), top] peak around 785 and 730 GeV, respectively, but the peak for the other one, i.e., $pp \rightarrow AA$ [Fig. 4(e), top] occurs near to 650 GeV, which is solely scalar mediated. This indicates that the final state particles coming from the associated production processes would be softer compared to the pair production processes. The K factor varies substantially, and in some kinematic regions, it indicates correction up to 90%. The nature of scale uncertainties for associated production processes is

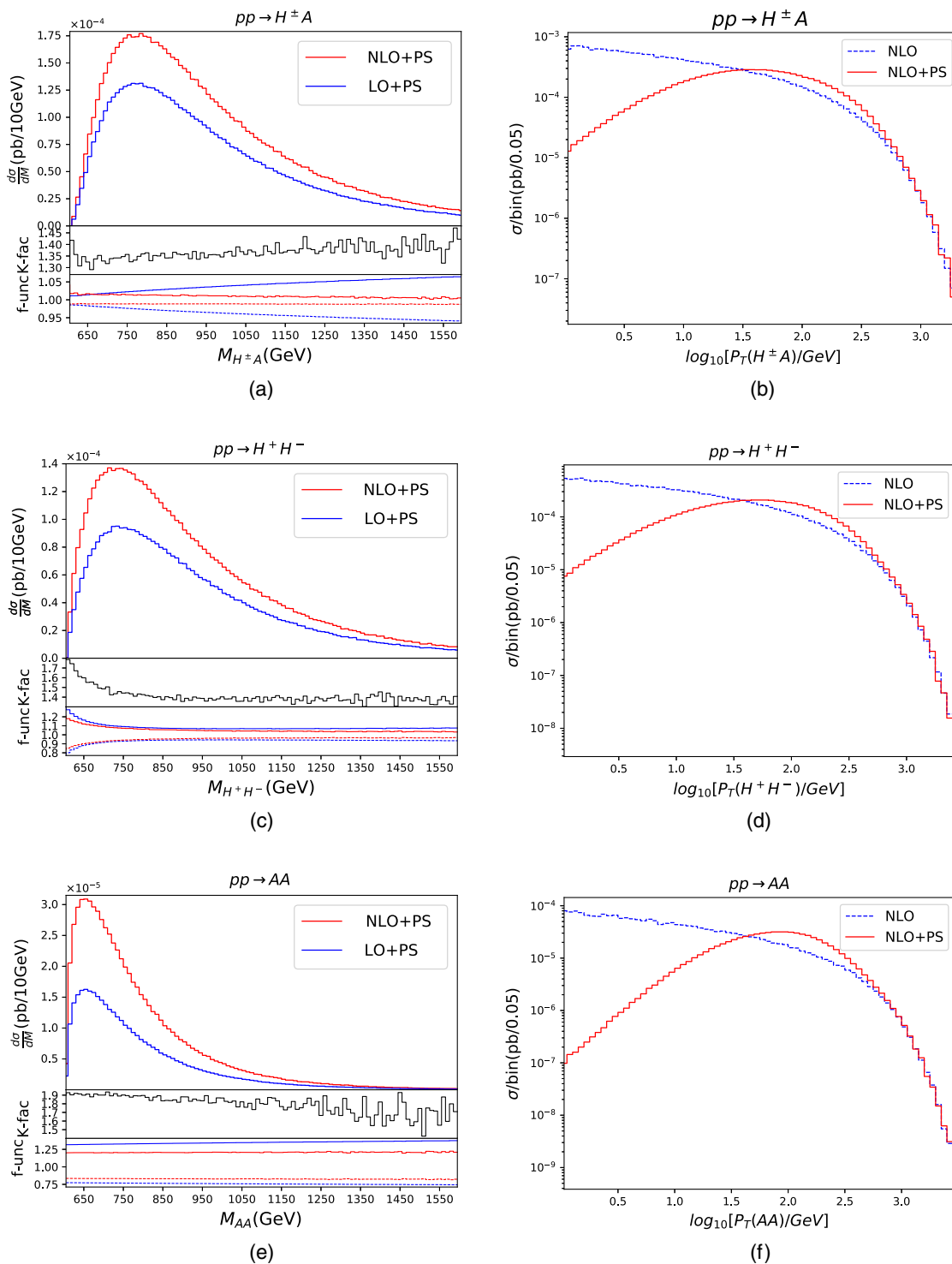


FIG. 4. Kinematic variables in the left and right panels are same as in Fig. 3, but these are for the pair production of the heavy scalar channels, such as, $pp \rightarrow H^\pm A$ [(a),(b)], $pp \rightarrow H^+ H^-$ [(c),(d)], and $pp \rightarrow AA$ [(e),(f)].

quite similar. Among pair production processes, fractional scale uncertainties for $pp \rightarrow H^+ H^-$ [Fig. 4(c), bottom] and $pp \rightarrow AA$ [Fig. 4(e), bottom] are mostly stable in the high invariant mass region, whereas for $pp \rightarrow H^\pm A$ [Fig. 4(a), bottom] such uncertainties are monotonically increasing.

Although these results are metaphorical as hadronic decay of the final state heavy scalars are not being considered here, they show the importance of doing $\mathcal{O}(\alpha_s)$ corrections to all the production channels to have a better estimation of the production rate and reduced scale uncertainty.

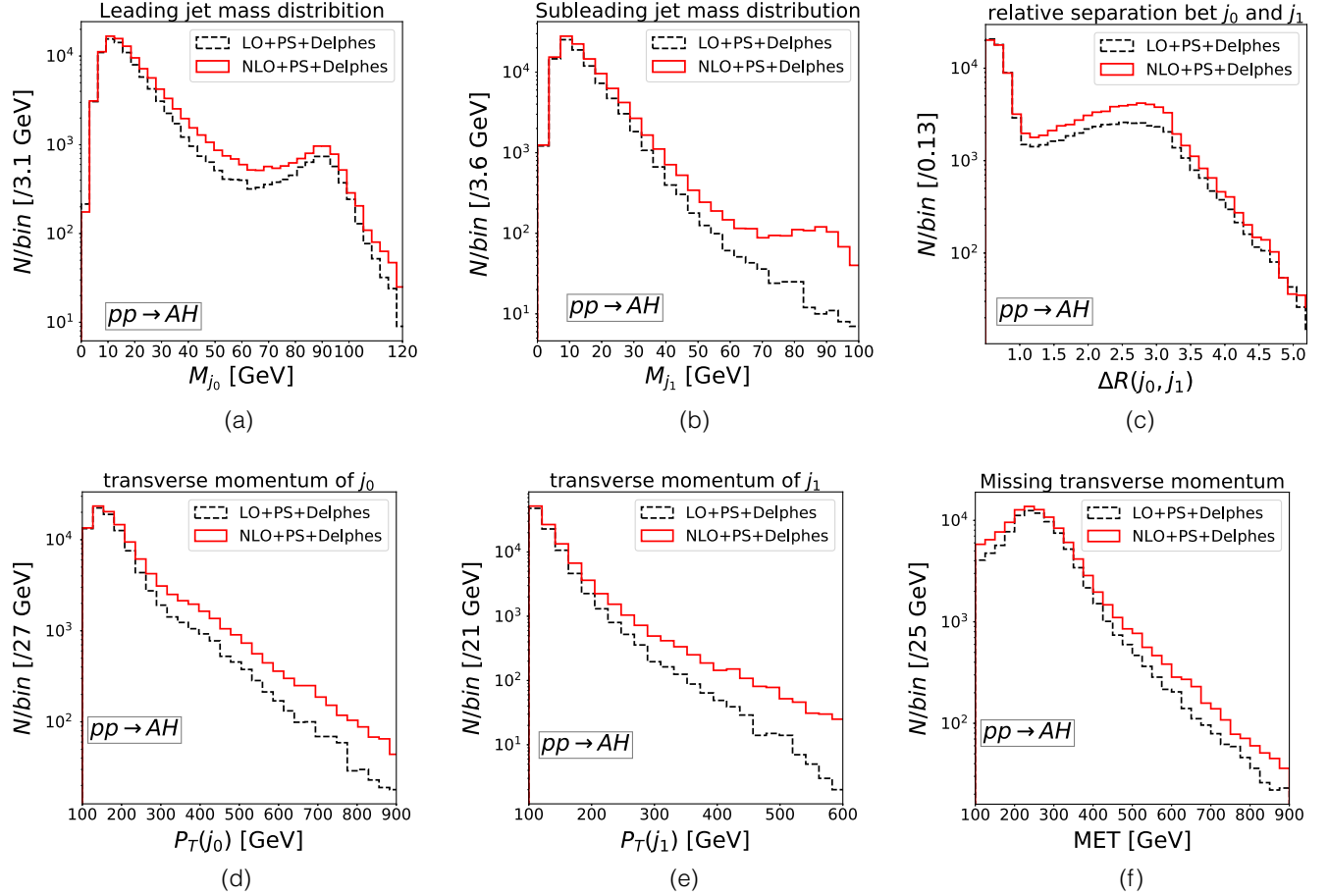


FIG. 5. Distributions of the various kinematic observables at LO (dashed black) and NLO (solid red) for the selected events with $\cancel{E}_T, P_T(j_0), P_T(j_1) > 100$ GeV from the channel $pp \rightarrow AH$, where A decay hadronically. This demonstration is for the benchmark point BP2. Plots (a) and (b) show distributions of the leading (j_0) and subleading (j_1) jet mass (M_{j_0}, M_{j_1}), respectively, (c) is the distribution of the relative separation between these two leading jets $\Delta R(j_0, j_1)$, while (d) and (e) are transverse momentum distribution of j_0 and j_1 , respectively. Plot (f) shows the distribution of the total missing transverse energy; here the label MET represents \cancel{E}_T .

V. QCD JETS FROM HEAVY SCALAR DECAY

Heavy scalars, after their creation through the associated channel along with DM candidate H , or from a pair production, primarily decay into H and a gauge boson, which is further decayed hadronically. It is imperative to look into their dominant hadronic decay channels as a possible probe for IDM. We select the simulated events including the parton shower and detector effect with a minimum missing transverse energy, $\cancel{E}_T > 100$ GeV, and the minimum transverse momentum of the two leading jets $P_T(j_i) > 100$ GeV (for $i = 0, 1$). Particle-flow towers and particle-flow tracks are used as input to cluster the jets of radius parameter 0.5, where we use the anti- K_T algorithm for clustering. The jet mass is defined by $M_j = (\sum_{i \in j} P_i)^2$, where P_i is the four-momentum of the i th constituent within the jet. The missing transverse energy is defined as the negative sum of the transverse momentum of all the reconstructed constituents, $\cancel{E}_T = -\sum_i \vec{P}_{T,i}$. The angular distance between two jets in the transverse plane is denoted

as $\Delta R(j_i, j_j)$. This section aims to examine relevant distributions of the jets from the signal to motivate the significance of NLO QCD calculation over the LO. In addition to upward shift, NLO corrections can change the shape of the distribution for a variety of kinematical variables. This has a profound effect in constructing the phenomenological study. These distributions also make a case for large-radius jets (fatjets) originated from boosted Z/W^\pm boson decay, which comes naturally in probing the hierarchical mass region of the IDM.

The distribution of the different high-level observables for one of the associated production channels of heavy scalar,² $pp \rightarrow AH$, and vector boson mediated pair production of the heavy scalars, $pp \rightarrow H^\pm A$ and scalar mediated pair production, $pp \rightarrow AA$ are shown in Figs. 5, 6, and 7, respectively, displaying the LO (dashed

²Both $pp \rightarrow AH$ and $pp \rightarrow H^\pm H$ channels follow similar distributions, as both A and H^\pm masses are nearly degenerate and produced through a vector mediator.

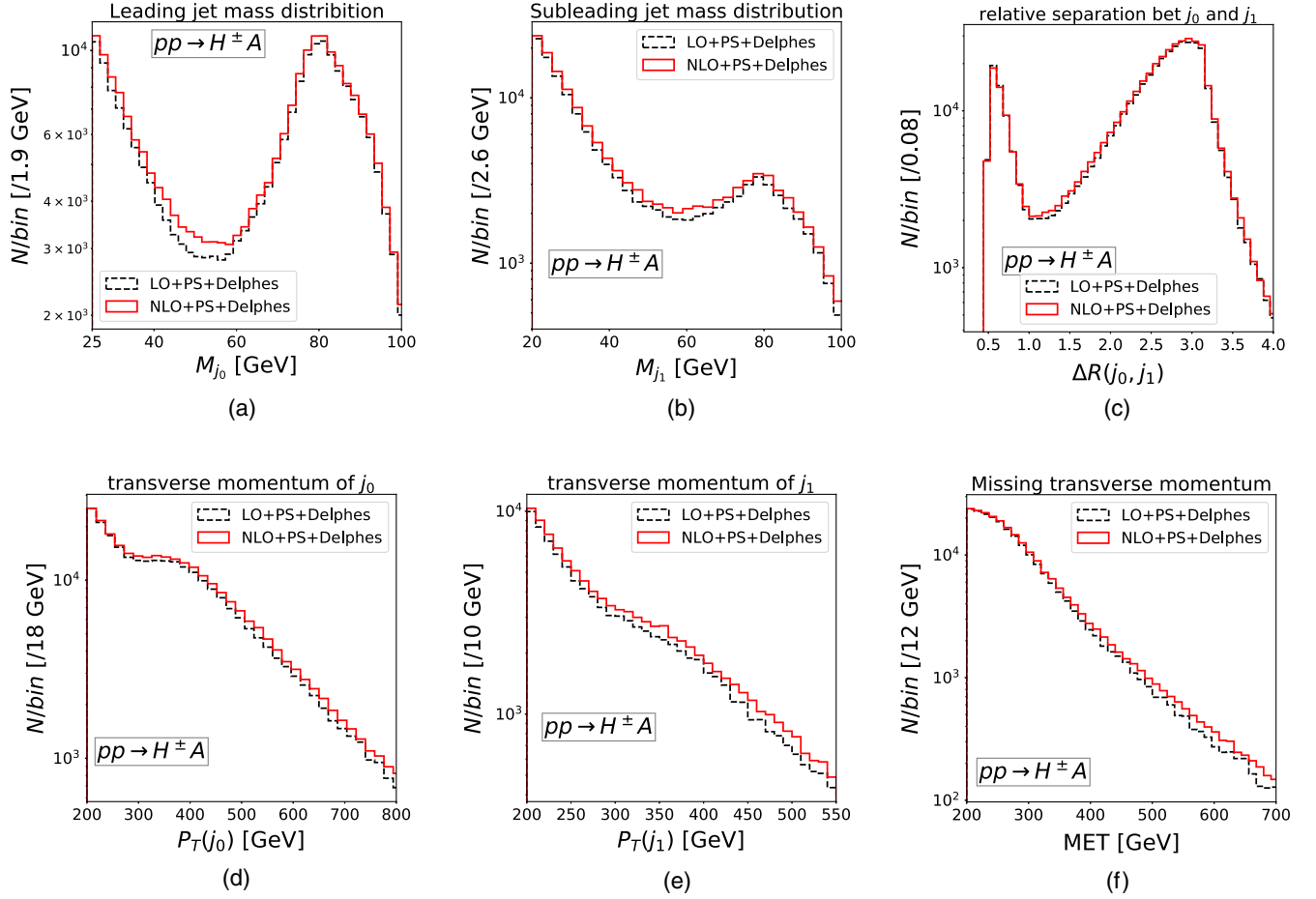


FIG. 6. Panels are the same as in Fig. 5, but for the pair production of the heavy scalars channel, such as $pp \rightarrow H^\pm A$, where both A , H^\pm decay hadronically.

black) and NLO (solid red) contributions considering a sample benchmark point BP2. In each figure, the first two plots (a) and (b) present distributions of the leading (j_0) and subleading (j_1) jet mass, respectively. Plot (c) presents the distribution of the relative separation between these leading and subleading jets, whereas plots (d) and (e) exhibit their transverse momentum distributions, respectively. Finally, plot (f) shows the distribution of the total missing transverse energy from such production.

The channel $pp \rightarrow AH$ at the partonic level produces three hard jets, two from Z -boson decay, and the other is the NLO radiation, while at LO, it has only two hard jets from Z -boson decay. The first peak in the leading jet mass distribution [Fig. 5(a)] is generated when a QCD hard parton forms a jet after PS and detector simulation. Interestingly, this same distribution points to a second peak both for LO and NLO results. This occurs when the Z boson is produced with enough boost to form a merged jet out of its full decay products, resulting into a peak at Z -boson mass. The second hard parton from the Z boson forms the subleading jet, causing a peak near $M_{j_1} = 10$ GeV [Fig. 5(b)] but no other peak in the LO j_1 mass distribution.

However, the NLO distribution can have extra hard radiation. Occasionally when that carries enough transverse momentum to form a leading jet, Z -boson decay still forms a merged subleading jet resulting in a second peak near Z -boson mass [Fig. 5(b)] deviating from a leading order estimate. Hence the NLO estimate predicts an upward trend in the number of boosted dijet events even from such associated production channels. One can also expect such abundance in boosted jets for other benchmark points with heavier scalars. Our previous argument is even more evident in the next distribution plot of the relative separation between two leading jets [Fig. 5(c)] for the same channel $pp \rightarrow AH$. The number of events with smaller jet separation $\Delta R_{j_0, j_1} < 1.0$ is 1 order larger than in the other region. For a significant event sample, both leading and subleading jets come from the Z boson's decay and are closely separated. Naturally, the construction of large-radius jets embeds them together to form a single fatjet carrying properties of the originating gauge boson. It is even more pronounced in larger masses of scalar. The distribution of the transverse momentum of the leading [Fig. 5(d)] and subleading [Fig. 5(e)] jets and the total missing transverse energy

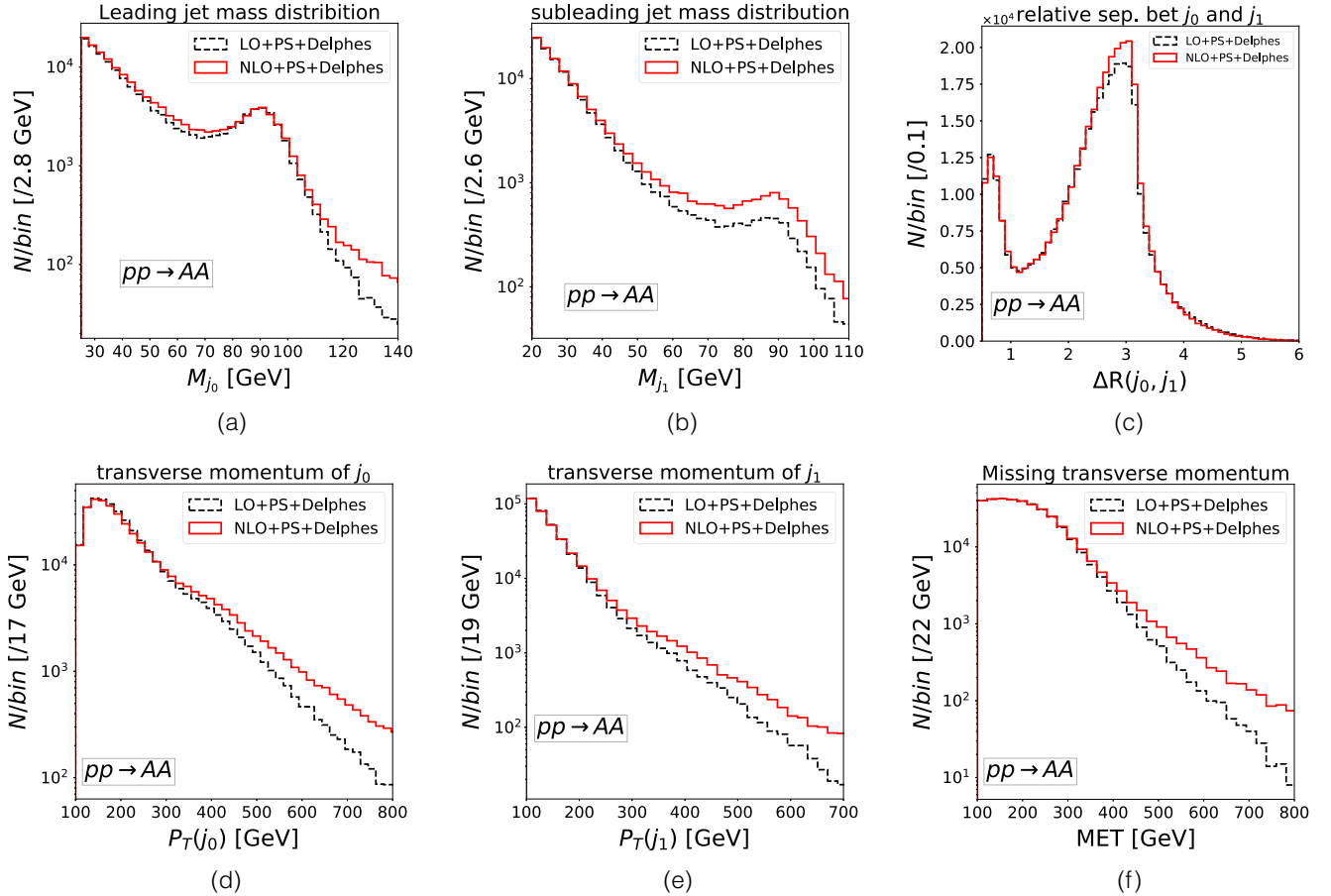


FIG. 7. Panels are the same as in Fig. 5, but for the pair production of the heavy scalar channel, such as $pp \rightarrow AA$, where both A decay hadronically.

[Fig. 5(f)] shows an upswing in NLO at larger PT. This is significant in view of the final selection of events (or, during multivariate analysis at the boosted decision tree) comes with higher weightage from these distribution tails to deal with a tiny signal over an overwhelmingly large background.

Similarly, one requires to follow distributions from pair production channels of the heavy scalars. The leading and subleading jet mass distributions for vector boson mediated [Figs. 6(a), 6(b)] and scalar boson mediated [Figs. 7(a), 7(b)] channels in pair production of heavy scalars demonstrates two clear mass peaks both at LO and NLO. In this case, pairs of heavy scalars produce two boosted vector bosons, and as they have enough boost, it results into the second peak in both cases. Again, with the increase of scalar mass, the second peak rises, ensuring enhancement of di-fatjet events. The distributions of the relative separation for pair production of heavy scalars shown in Figs. 6(c) and 7(c) contain two peaks. The second peak at $\Delta R \sim \pi$ appears when two jets originate from two different vector bosons. The first peak is when both the jets arrive from the same vector boson, which gradually diminishes for heavier mass. Pair production channel $pp \rightarrow AA$ has a significant shift between NLO and LO distributions in

comparison to the $pp \rightarrow H^\pm A$ channel, as the former is Higgs mediated and has a larger K factor. It is evident from this discussion that the tagging of large-radius jets originating from boosted vector bosons can significantly improve the efficiency of probing the hierarchical mass spectrum of the IDM. In the next section, we will describe the selection and properties of such boosted fatjets.

VI. BOOSTED FATJET AS A PROXY FOR HEAVY SCALAR PRODUCTION

Our discussion in the previous section demonstrates that the multijet + \cancel{E}_T search is not sufficient to explore the hierarchical mass region of IDM. Jet pair originated from the vector bosons, which comes out as a boosted decay product of heavy scalar, is already collimated as a merged hadronic object. This process of getting a fatjet becomes more and more evident while probing a heavier scalar mass. A large radius fatjet can effectively identify this combined hadronic yield from the boosted vector boson. Moreover, it can carry a significant amount of information hidden inside the internal structure of jet formation through the orientation of fragmented hadrons and their energy

deposits, revealing the properties and identity of the originating particle.

A. Signal and background processes

Representative LO Feynman diagrams both for associate production and pair production of heavy scalars are already depicted in Fig. 1. Our primary focus is to analyze the NLO accurate di-fatjet signal arising from heavy IDM scalar decay using jet substructure variables. We do not discriminate W -jet or Z -jet and dub them as V -fatjet (J_V) since we consider a suitable mass window to accommodate both in our analysis. We will discuss the usefulness of the sophisticated multivariate analysis that can make the signature of $2J_V + \cancel{E}_T$ into the better discriminator in order to separate out tiny signal from an overwhelmingly large SM background. However, a multivariate description creates a highly performant nonlinear cut at the cost of blurring the exact physical description of how different high-level variables affect our analysis. Hence, to better understand the kinematic variables that may affect LO and NLO computations, we would analyze them first with the usual cut-based method before moving on to the MVA analysis. In passing, it is to be noted that the cross section of the di-Higgs production while one Higgs boson decay into a pair of bottom quarks ($h \rightarrow b\bar{b}$) and the other decays into pair of dark matter ($h \rightarrow HH$) is 1.05 fb. Although this channel has a sizable effect on the di-fatjet final state, we do not include this in our analysis since this process drops sharply after applying the b -veto.

All the significant backgrounds that contribute to the $2J_V + \cancel{E}_T$ signal are included in our analysis. We do two to four additional jets merging using the MLM matching [61,62] scheme for different background processes, and normalize the cross section according to the available higher-order QCD corrections. Inclusive Z -boson production is the principal background where the Z boson decays invisibly ($pp \rightarrow Z + \text{jets} \rightarrow \nu\nu + \text{jets}$) and gives a large \cancel{E}_T together with fatjets arising from QCD jets. This process is matched to four extra partons using the MLM scheme. Second, inclusive W^\pm boson production has a significant contribution when the lepton from the leptonic decay of the W boson remains undetected ($pp \rightarrow W + \text{jets} \rightarrow l_{e,\mu}\nu + \text{jets}$). The neutrino from W decay gives a substantial amount of \cancel{E}_T and fatjets arise from QCD jets. This process is generated up to four extra partons with MLM matching. Note that the contribution from the above two background processes counts only when the missing transverse momentum is sufficiently large. We apply the generation level hard cut $\cancel{E}_T > 100$ GeV, as the region with the lower missing transverse energy is of no interest for this present analysis. Additionally, di-boson production can offer a considerable amount of contribution in the background. The three different di-boson processes $pp \rightarrow WZ, WW,$ and ZZ , are possible, where the WZ process gives the most significant

contribution among these three. All three processes are generated and merged up to two extra partons. One of the vector bosons in all these processes decays hadronically, giving rise to a J_V . Other vector boson decaying invisibly ($Z \rightarrow \nu\nu$) or leptonically ($W \rightarrow l_{e,\mu}\nu$) with lepton being undetected, gives a large \cancel{E}_T . Another fatjet in all these di-boson processes arises from the QCD jets. Single top production is possible in SM through three different types of process, S-channel ($pp \rightarrow tb$), t-channel ($pp \rightarrow tj$), and associated production ($pp \rightarrow tW$), where associated production gives a considerable amount of contribution to the background of our signal. This process is merged up to two extra partons using the MLM scheme. Finally, top pair production contributes to the background when one top decays leptonically and the lepton is escaping detection. Whereas the other top decays hadronically and that essentially gives rise to a vectorlike fatjet J_V . Since such an event comes with a couple of b -jets, b -veto can effectively reduce this background. This process is generated to two extra partons with MLM matching. The other fatjet arises from the QCD jets or untagged b -jets. We found negligible contributions to the background from the QCD multijet and tri-boson processes compared to the processes mentioned above. Therefore we do not include these processes into our analysis. For our simulated backgrounds at 14 TeV LHC, we normalize their cross section according to the available higher-order QCD corrections, as tabulated in Table IV of Ref. [18].

The associated production of heavy scalar with two jets merging and pair production of the heavy scalars are analyzed at LO [18] where it was found that the former processes contribute dominantly in the di-fatjet final state than the latter. A further estimate at NLO accuracy modifies the contribution in two vital directions. First, both for the associated production and pair production of heavy scalar processes, the differential NLO K factor plays an important role, as already described in the previous section. Second, two jets merged associated production channels can mimic the Higgs mediated pair production of heavy scalar processes, and therefore may contribute to double counting in a particular phase space region. NLO estimate eliminates such possibility giving nonoverlapping contributions from all processes.

Now, along with both these effects, our estimate at NLO predicts reduced contribution from associated production, thereby enhancing the part from the pair production. This has a profound significance in setting up the phenomenological analysis. On contrary to a more complex mixed-signal region analysis by taking into account the admixture of $1J_V$ and $2J_V$, that has been carried out in Ref. [18], it is tempting to concentrate only on the $2J_V$ identification for a significant gain. Demanding that both the fatjets have V -jet like characteristics, one finds a more effective background control and, as a result, a higher statistical significance.

B. Construction of high level variables

The total energy of the fatjet originated from the hadronic decay of boosted W, Z is distributed around two subject axes. N -subjettiness ratio (τ_{21}) and the jet mass (M_J) are two potent variables to classify such fatjets J_V from those that arise from the fragmentation of QCD parton. The jet mass is defined by $M_J = (\sum_{i \in J} P_i)^2$, where P_i is the four-momentum of the i th hit in the calorimeter. Large-radius jets are prone to attract additional soft contributions from underlying QCD radiation, which needs to be eliminated to get reliable estimates from the different high-level variables. Pruning, filtering, and trimming [63–66] are different grooming techniques prescribed to remove those soft and wide-angle radiations. We consider pruned jet in our analysis as discussed in Refs. [65,66].

We run the pruning algorithm repeatedly to remove the soft and wide-angle emission and veto such recombinations. One has to estimate two variables, the angular separation of the two proto jets, ΔR_{ij} and softness parameter $Z = \min(P_{Ti}, P_{Tj})/P_{T(i+j)}$, at every recombination step. The recombination between i th and j th proto jets is not performed dropping the softer one, if $\Delta R_{ij} > R_{\text{fact}}$ and $Z < Z_{\text{cut}}$. We choose standard default values of $R_{\text{fact}} = 0.5$ and $Z_{\text{cut}} = 0.1$ [65]. The N subjettiness determines the jet shape of hadronically decaying boosted V bosons. Considering that N number of subjets exist within the jet, N subjettiness (τ_N) is defined by the angular separation between constituents of the jet from their nearest subjet axis as given below [67,68]:

$$\tau_N = \frac{1}{N_0} \sum_i P_{T,i} \min\{\Delta R_{i,1}, \Delta R_{i,2}, \dots, \Delta R_{i,N}\}. \quad (20)$$

The summation runs over all the constituents of the jet, and $P_{T,i}$ is the transverse momentum of the i th constituent. $N_0 = \sum_i P_{T,i} R$ is the normalization factor, and R is the jet radius. τ_{21} denotes the ratio of τ_2 and τ_1 , which is an

excellent variable to tag a hadronically decaying boosted V boson as it tends to zero (far from zero) for a correctly identified two-prong (one-prong) jet.

To proceed further, we define the following *preselection criteria* based on which signal and background event samples are prepared: (i) each event has to have at least two fatjets constructed by the Cambridge-Aachen (CA) jet clustering algorithm with radius parameter $R = 0.8$, and the minimum transverse momentum of each fatjet $P_T(J_i) > 180$ GeV, (ii) since a pair of DM particles are produced in the signal, a minimum missing transverse energy $\cancel{E}_T > 100$ GeV is applied to select the events, (iii) we also impose a minimum azimuthal angle separation between the identified fatjet and missing transverse momentum direction, so that, $|\Delta\phi(J_i, \cancel{E}_T)| > 0.2$. This minimizes any jet mismeasurement effect contributing to \cancel{E}_T , (iv) since no leptons are expected in the signal region, backgrounds can be further suppressed by vetoing a lepton tag. So, events are vetoed if they contain leptons that have pseudorapidity $|\eta(l)| < 2.4$ and transverse momentum $P_T(l) > 10$ GeV.

It is clear from our previous discussion on boosted fatjet that several interesting variables can contribute to strengthen the signal efficiency. We would demonstrate the distribution of all such variables, but before that we point out some of the significant changes that appeared due to NLO computation in the signal region. In Table IV, we show the expected number of $2J_V + \cancel{E}_T$ final state events corresponding to the central scale, originated from different pair production of heavy scalar processes. Such numbers for $pp \rightarrow H^\pm A$, and $pp \rightarrow H^+ H^-$ at NLO (N_S^{NLO}) level are given for three sample benchmark points, together with LO level numbers multiplied by overall NLO K factor ($N_S^{\text{LO} \times K}$) for 3000 fb^{-1} integrated luminosity at 14 TeV LHC. Superscripts and subscripts are the change in the corresponding number of events due to the envelope of eight different (μ_R, μ_F) scale choices. In both cases, that makes the overall cross section normalized to the NLO value.

TABLE IV. N_S^{NLO} and $N_S^{\text{LO} \times K}$ are shown to represent the expected number of the $2J_V + \cancel{E}_T$ final state events for different pair production of heavy scalars channels, $pp \rightarrow H^\pm A$, and $pp \rightarrow H^+ H^-$. These two numbers coming from NLO computation and LO with integrated NLO K -factor multiplication, respectively, at 14 TeV HL-LHC. Superscripts and subscripts are the change in the corresponding number of events due to the envelope of eight different (μ_R, μ_F) scale choices. Additional cuts over existing preselection (see text) are $\cancel{E}_T > 200$ GeV, b -veto, $65 \text{ GeV} < M(J_0), M(J_1) < 105 \text{ GeV}$, and $\tau_{21}(J_0), \tau_{21}(J_1) < 0.35$. Relative change, $(N_S^{\text{NLO}} - N_S^{\text{LO} \times K})/N_S^{\text{LO} \times K}$ reflects the importance of the differential K factor in the present computation. Relative changes are given corresponding to the central scale.

Preselection cuts + $\cancel{E}_T > 200$ GeV, b -veto, $65 \text{ GeV} < M(J_0), M(J_1) < 105 \text{ GeV}$, $\tau_{21}(J_0), \tau_{21}(J_1) < 0.35$						
BP	$H^\pm A$			$H^+ H^-$		
	N_S^{NLO}	$N_S^{\text{LO} \times K}$	Relative change %	N_S^{NLO}	$N_S^{\text{LO} \times K}$	Relative change %
BP1	$168.2^{+2.8}_{-2.5}$	$119.5^{+3.3}_{-3.2}$	40.75%	$121.2^{+4.6}_{-4.6}$	$82.7^{+5.0}_{-4.1}$	46.55%
BP2	$190.7^{+3.1}_{-4.1}$	$155.6^{+5.7}_{-5.5}$	22.56%	$150.4^{+9.2}_{-8.0}$	$111.1^{+9.8}_{-7.7}$	35.37%
BP3	$202.8^{+4.2}_{-4.2}$	$162.8^{+7.0}_{-6.5}$	24.57%	$153.8^{+11.8}_{-9.8}$	$122.5^{+13.8}_{-10.9}$	25.55%

TABLE V. Expected number of events from different signal and background processes at 14 TeV HL-LHC corresponding to the central scale after applying the preselection cuts with leading and subleading fatjet mass $M_{J_0}, M_{J_1} > 40$ GeV and b -veto.

Topology	BP1	BP2	BP3	BP4	BP5	BP6	BP7
Associated production	452.29	377.73	327.56	266.9	217.53	176.9	138.11
Pair production	1677.13	1432.67	1184.16	969.0	785.63	622.99	516.62
Z + jets	W + jets	tW + jets	tt + jets	WZ + jets	ZZ + jets	WW + jets	Total
652519	527312	46011.8	54635	36126.5	3689.51	12002.4	1.3323×10^6

Signal region criteria in conjunction with preselection cuts are described in Table IV.³ Relative change, defined as $(N_S^{\text{NLO}} - N_S^{\text{LO} \times K})/N_S^{\text{LO} \times K}$, is given for the central scale. Relative change is independent of the luminosity and ascertains the necessity of considering actual NLO events instead of using LO events multiplied by a flat K factor. It is evident that NLO and LO computations have different efficiencies for the given kinematic cuts. Relative change between these two estimations exhibits the role of the differential NLO K factor by changing the LO estimation up to 40% for the process $pp \rightarrow H^\pm A$, and 46% for $pp \rightarrow H^+ H^-$ for the given kinematic cuts mentioned at the top of the Table IV.

In addition to the preselection cuts described above, final event selection criteria for multivariate analysis includes a very relaxed cut on pruned jet mass. All other variables are kept free to provide the multivariate analysis with enough scope to optimize the nonlinear cut based on suitable variables. We select the signal and background events after applying the following cuts: (i) both leading and subleading fatjets have to have a minimum pruned jet mass of 40 GeV to reduce the contribution of fatjets originated from QCD, (ii) b -veto is applied on the jets that are formed using anti- k_t algorithm with radius parameter $R = 0.5$ and this significantly reduces the $t\bar{t}$ background.

C. Multivariate analysis (MVA)

In Table V, we present the expected number of signal events coming from the associated production and pair production of the heavy scalar channels together with all background processes at 14 TeV LHC with integrated luminosity 3000 fb^{-1} . From Table V, we see a contribution in the $2J_V + \cancel{E}_T$ final state coming from the pair production of the heavy scalars is always more prominent than the associated production after these cuts. We construct two independent event samples for our multivariate analysis, one for the signal and another for the background.

³One can, in principle, use such stiff event selection criteria for a realistic cut based analysis. Our purpose is purely for demonstration, as we would finally employ multivariate analysis to construct the suitable optimization based on rather loosely set criteria.

The entire dataset is split randomly—50% for the training and the remaining for testing purposes for both samples. We employ an adaptive BDT algorithm for MVA. We generate different signal processes separately at NLO and combine them according to their weights to get the kinematic distributions of the combined signal. Similarly, the different background processes are generated separately at LO with two to four extra jet MLM matching and combined thereafter according to their weights to get the kinematic distributions of the combined background. A set of kinematic variables is chosen from a bigger group of variables employed in the MVA analysis depending on their relative importance while discriminating the signal class from the background class. We present in Fig. 8 the normalized kinematic distributions of all nine input variables that are used in MVA. We obtain the signal distributions using sample benchmark point BP2, including all the associated production and pair production of the heavy scalars at NLO. We do not include the process $pp \rightarrow HH$ in our analysis although it has a larger cross section than any other associated or pair-production channels, as b -veto and cuts on the fatjet mass and N -subjettiness ratio τ_{21} weaken its effect and the remaining events reside well away from the maximum BDT response region. The background comprises of all the processes discussed in Sec. VIA after applying the cuts $M_{J_0}, M_{J_1} > 40$ GeV and b -veto along with the preselection cuts mentioned in Sec. VIB at 14 TeV LHC. The distributions of the pruned jet mass $M_{J_{0,1}}$ of the leading [Fig. 8(a)] and subleading [Fig. 8(c)] fatjets, have a peak near 80–90 GeV for the signal close to the vector boson mass, however, no such peak for the background reflects that fatjets are predominantly formed from QCD jets. The distributions of the N -subjettiness ratio, $\tau_{21}(J_{0,1})$ of the leading [Fig. 8(c)] and subleading [Fig. 8(d)] fatjets establish that both the fatjets of the signal have a two-prong structure as they peak at a smaller value of τ_{21} . In contrast, both the fatjets in the background has a characteristic one-prong structure producing a larger value for this variable. Hence these four jet substructure variables are crucial in discriminating the signal from the background. The relative separation between the leading (J_0) and subleading (J_1) fatjets $\Delta R(J_0, J_1)$ [Fig. 8(e)], azimuthal separation between

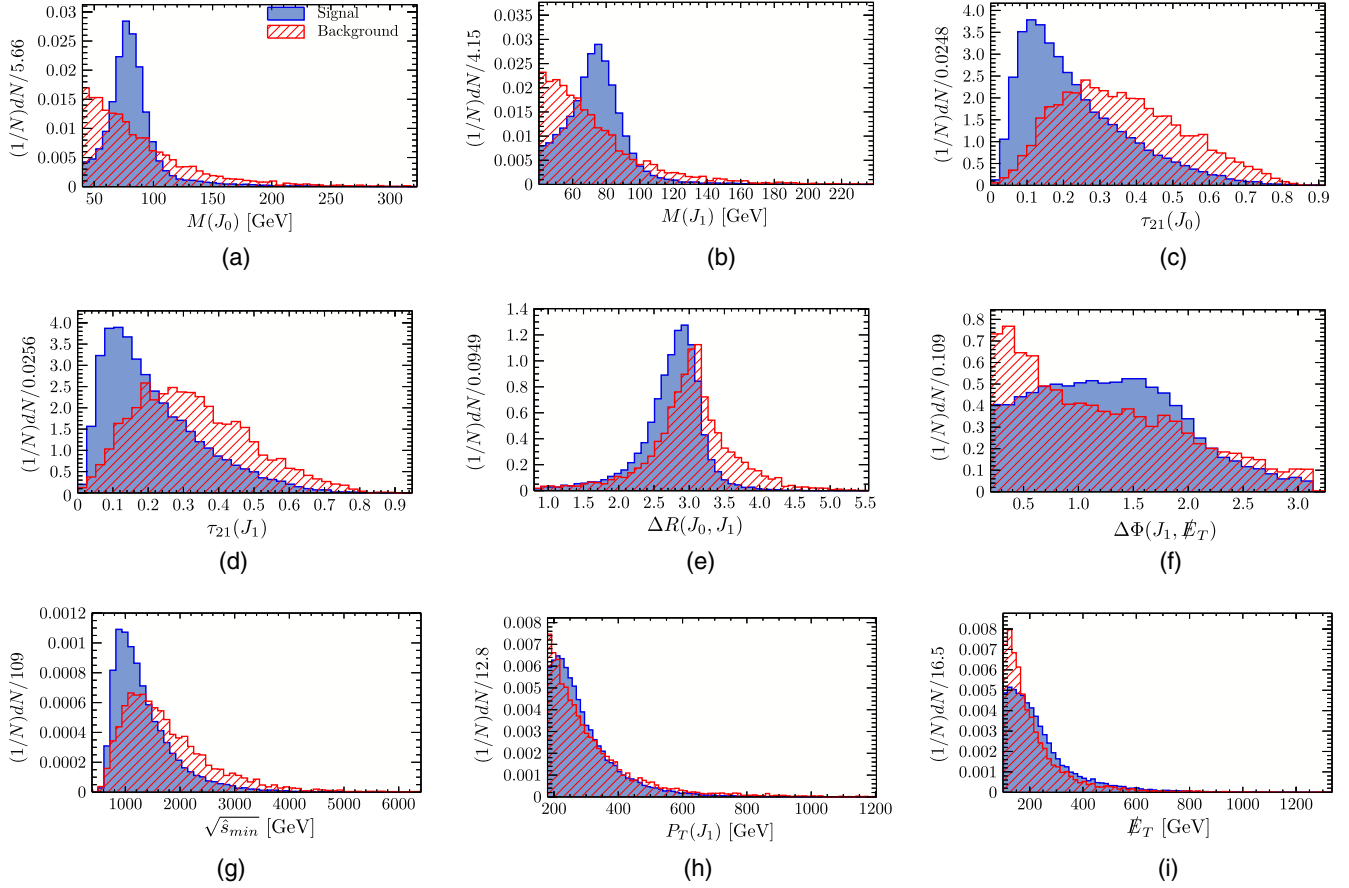


FIG. 8. Normalized kinematic distributions of the different input variables used in MVA for the background (red) and the signal (blue). Plot (a) and (b) represent the distribution of pruned jet mass of the leading and subleading fatjets, whereas plot (c) and (d) are the distributions of the N -subjettiness ratio of the leading and subleading fatjets, respectively. Plot (e) shows distribution of the relative separation between the two leading fatjets. Azimuthal separation distribution of the subleading fatjet from the missing energy direction is depicted in plot (f). Plot (g) shows distribution of the global inclusive variable $\sqrt{\hat{S}_{\min}}$ and distribution of the transverse momentum of the subleading fatjet and total missing transverse momentum are presented in plot (h) and (i), respectively. We display the signal distributions for BP2, including all contributions from the associated production of the heavy scalar and pair production of the heavy scalars at NLO. The background comprises all the processes discussed in Sec. VI A after applying the cuts $M_{J_{0,1}} > 40$ GeV and b -veto together with the preselection criteria mentioned in Sec. VI B.

J_1 and \cancel{E}_T is represented as $\Delta\phi(J_1, \cancel{E}_T)$ [Fig. 8(f)], and the inclusive global variable $\sqrt{\hat{S}_{\min}}$ [Fig. 8(g)] are effective observables to separate the signal from the background. The inclusive variable $\sqrt{\hat{S}_{\min}}$, defined as the minimum c.m. energy required to satisfy all observed objects and \cancel{E}_T was proposed in [69–71] to find the new physics mass scale for the signals containing invisible particles like ours. All the reconstructed objects of the detectors are used to

construct the reconstructed object-level $\sqrt{\hat{S}_{\min}}$ that demonstrate better efficiency than the other inclusive variables H_T , \cancel{H}_T , etc.

A variable is considered to be more powerful discriminator, if it possesses a larger separation between the signal and background. For different kinematic variables, the method unspecific relative importance is shown in Table VI, where we principally keep the variables that have

TABLE VI. Method unspecific relative importance (or separation power) of the different variables according to their rank before using at MVA. We obtain the numbers for BP2 from the TMVA package during MVA. Those numbers can change modestly for different benchmark points and different algorithms.

Variable	$\tau_{21}(J_0)$	$M(J_0)$	$\tau_{21}(J_1)$	$M(J_1)$	$\Delta R(J_0, J_1)$	$\sqrt{\hat{S}_{\min}}$	$\Delta\phi(J_1, \cancel{E}_T)$	\cancel{E}_T	$P_T(J_1)$
Separation	16.58	15.71	13.71	11.57	11.27	9.039	3.011	2.451	1.324

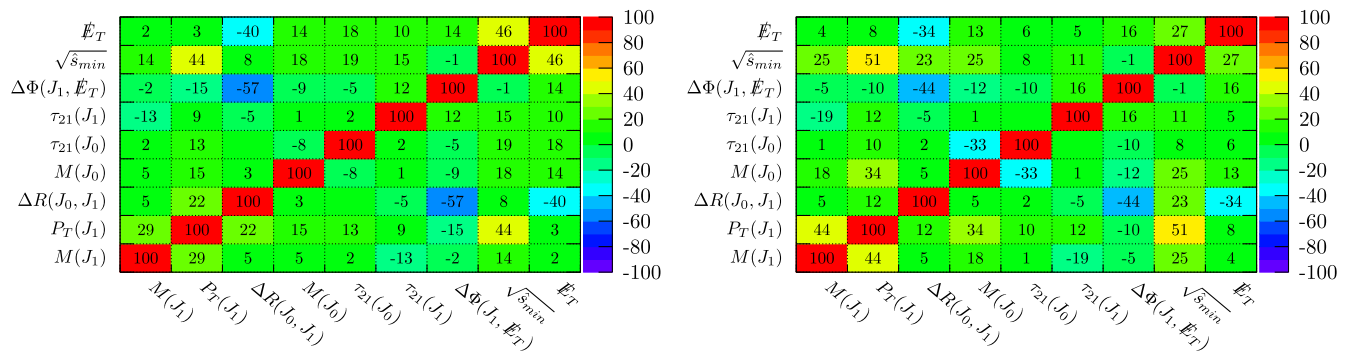


FIG. 9. The linear correlation coefficients among different kinematic variables used in MVA (in percentage) for the signal (left panel, BP2) and background (right panel). The positive and negative signs signify the positive and negative correlations (anticorrelated) among the two variables.

less (anti-)correlation among themselves both for the signal and background. We notice four jet substructure variables $M_{J_{0,1}}$ and $\tau_{21}(J_{0,1})$ are very good discriminators. The relative importance of the different kinematic variables can change modestly for different benchmark points. Although very high P_T for both fatjets and large \cancel{E}_T are considered during event selection, transverse momentum of the subleading fatjet, $P_T(J_1)$ [Fig. 8(h)] and \cancel{E}_T [Fig. 8(i)] still can take a role in discriminating the signal from the background in MVA. Note that $P_T(J_0)$ and $P_T(J_1)$ are highly correlated (positively) both in signal and background classes, so we keep only $P_T(J_1)$ in the analysis as it has more relative importance than $P_T(J_0)$. Similarly, $\Delta\phi(J_0, \cancel{E}_T)$ and $\Delta\phi(J_1, \cancel{E}_T)$ are highly anticorrelated, but we keep $\Delta\phi(J_1, \cancel{E}_T)$ because of its larger relative importance. The linear correlation coefficients among different kinematic variables used in MVA (in %) for the signal and background are shown in Fig. 9. The positive (negative) signs signify the positive (negative) correlation (anti-correlation) among the two variables. Modestly large anticorrelation between $\Delta\phi(J_1, \cancel{E}_T)$ and $\Delta R(J_0, J_1)$ is present, although we kept them both as they have large relative importance.

Finally, we present the normalized BDT response for the training and testing samples for both signal and background classes in the left panel of Fig. 10. The signal distribution is presented for BP2. The distributions of the BDT response get well separated for the signal and background. Cut efficiencies can be estimated by applying a cut $\text{BDT}_{\text{res}} > \text{BDT}_{\text{cut}}$ on the BDT response. In the right panel of Fig. 10, such cut efficiencies are demonstrated for the background (red) and signal (blue), along with the statistical significance of the signal over the background (green) as a function of the cut value applied on the BDT response. We use the prescription $\sigma = \frac{N_S}{\sqrt{N_S + N_B}}$ for computing the statistical significance. N_S and N_B are, respectively, the expected number of signal and background events after using the optimal cut BDT_{opt} at 3000 fb^{-1} luminosity at 14 TeV LHC. N_S , N_B , and σ are shown in the right panel of Fig. 11 for different benchmark points. We find more than a 5σ discovery potential for four different benchmark points. In the left panel of Fig. 11, we summarize the result in terms of statistical significance of the signal as a function of the masses of the heavy BSM scalars (solid red) at 14 TeV LHC with integrated luminosity 3000 fb^{-1} . At the same time, the

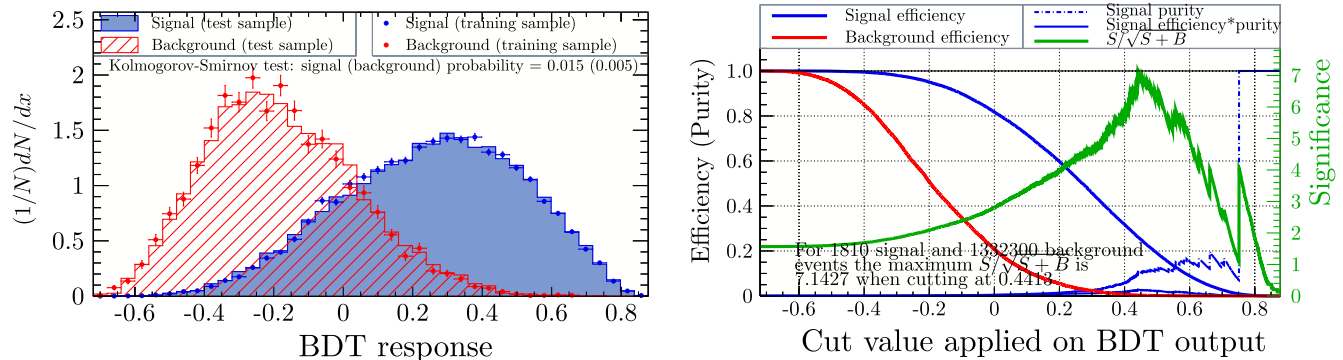
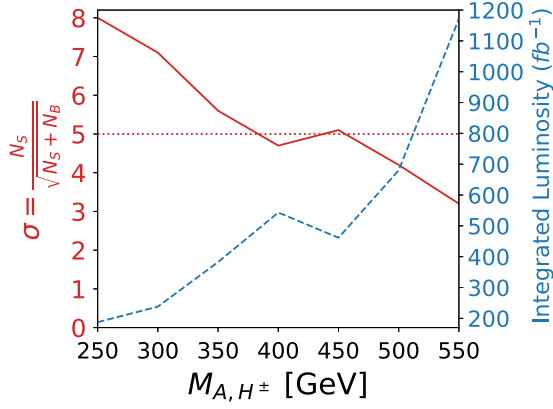


FIG. 10. The left panel shows the normalized BDT response for the training and testing samples for both signal (BP2) and background classes. The right panel contains the cut efficiencies for the background (red) and the signal (blue) and the statistical significance of the signal over the background (green) as a function of the cut value applied on the BDT response.



Model	N_S^{bc}	BDT_{opt}	N_S	N_B	σ
BP1	2129	0.6927	357	1650	8.0
BP2	1810	0.4413	474	3930	7.1
BP3	1511	0.6003	390	4516	5.6
BP4	1236	0.5270	207	1738	4.7
BP5	1003	0.6499	178	1041	5.1
BP6	800	0.6098	102	485	4.2
BP7	654	0.6462	97	820	3.2
N_{SM}	1332300				

FIG. 11. The upper panel shows the statistical significance of the signal over the background as a function of masses of the heavy BSM scalars (solid red line) at 14 TeV HL-LHC. The dashed blue curve on the same plot exhibits the required luminosity for two sigma (2σ) exclusion for different benchmark points—the horizontal dotted red line to mark 5σ discovery potential. The lower panel demonstrates the corresponding expected number of signal events (N_S^{bc}) at NLO and background events (N_{SM}) before applying the BDT cut, where N_S and N_B are the expected number of signal and background events that survive after applying the optimum BDT_{opt} cut, respectively.

dashed blue line exhibits the required luminosity for 2σ exclusion for different benchmark points.

VII. CONCLUSIONS

IDM is a simple extension of the SM where a new $SU(2)_L$ scalar doublet owning a discrete \mathbb{Z}_2 symmetry provides a viable DM candidate together with additional heavy BSM scalars. This model offers two distinct parameter spaces, consisting of hierarchical mass spectrum and degenerate mass spectrum of these scalars, that satisfy the observed relic density of the dark matter and other theoretical and experimental constraints.

Despite of several studies being performed in exploring this viable dark matter model at the LHC, in this paper we initiate the effort of looking into a promising channel with NLO QCD precision. This study focuses on the hierarchical mass region and considers NLO QCD corrections on the associated and pair production channels of heavy scalars. We find that the effect of QCD correction is significant for

encrypting the correct search strategy at the LHC. Table II, and Table III encapsulate the correction factors for different benchmark points. We get an overall correction of about 33%–39% for the associated production processes and for a gauge boson mediated pair production channel, $pp \rightarrow H^\pm A$. Similarly, the $pp \rightarrow H^+ H^-$ process, which encompasses both gauge boson and Higgs mediator, has the correction factor in between 38% and 56%. In contrast, $pp \rightarrow AA$ being scalar mediated, receives a correction factor in the range of 70%–92%. Nevertheless, notable improvement on scale uncertainties is achieved due to the inclusion of NLO corrections. We also take into account the parton shower effect and demonstrate its practicality at the low transverse momentum region.

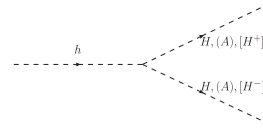
After jet clustering and detector simulation, we compare distributions of various crucial kinematic observables at LO and NLO. Noted shifts in the shape of these distributions over the LO computation can significantly influence the construction of phenomenological analysis. We notice a substantial relative change in the number of survived signal events as an effect of the differential NLO K factor. For example, this change is up to 46% for the gauge mediated pair production of heavy scalar processes. We also emphasize that gauge boson mediated decay products of hadronically decayed heavy scalars are highly collimated in this signal region and therefore large-radius fatjets come naturally in probing the hierarchical mass region. The internal structure and properties of the fatjet are key ingredients to know about their genesis. Fatjets originated from the QCD radiation of partons pose different characteristics compared to the fatjets generated from boosted vector boson. The jet substructure is a powerful tool to get control over the colossal SM background and identify the signal correctly. We find jet-substructure observables $M_{J_{0,1}}$ and $\tau_{21}(J_{0,1})$ are excellent discriminators in discriminating fatjets originated from the boosted vector boson and the QCD jets. We work with the parton shower matched NLO QCD corrected signal and employ sophisticated multivariate analysis to distinguish the signal using these powerful jet-substructure variables. We discuss the set of nine variables that are used in the MVA analysis and their linear correlation coefficients are presented for the signal at a sample benchmark point and for the background.

We observe that the discovery potential for different benchmark points nearly up to 350 GeV of heavy scalar mass in the hierarchical mass region has a statistical significance above 5σ at the HL-LHC. Hence, this parameter space of the hierarchical mass spectrum which is well motivated having a dark matter candidate of mass $m_{DM} \sim m_h/2$, would be quite interesting to look into. We also notice through this study that the heavy BSM scalar mass falling in the range of 250–550 GeV can be excluded with 1200 fb^{-1} integrated luminosity at the 14 TeV LHC.

ACKNOWLEDGMENTS

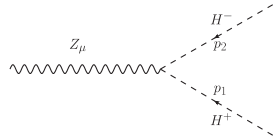
We thank Akanksha Bhardwaj for her continuous help and fruitful discussions regarding the phenomenological part of this analysis. This work is supported by the Physical Research Laboratory (PRL), Department of Space, Government of India. Computational work was performed using the HPC resources (Vikram-100 HPC) and TDP project at PRL.

APPENDIX: FEYNMAN RULES



$$= -i\Lambda v = \begin{cases} -i\lambda_c v & \text{for } AA \\ -i\lambda_L v & \text{for } HH \\ -i\lambda_3 v & \text{for } H^+H^- \end{cases}, \quad (\text{A1})$$

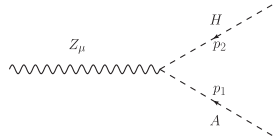
where $\lambda_{L/c} = (\lambda_3 + \lambda_4 \pm \lambda_5)$.



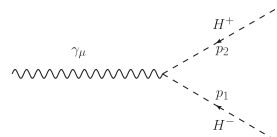
$$= i e (p_1 - p_2)_\mu \frac{\cos(2\theta_W)}{2 C_W S_W}, \quad (\text{A2})$$

where θ_W is the Weinberg angle. S_W and C_W correspond to the sine and cosine of the Weinberg angle, respectively:

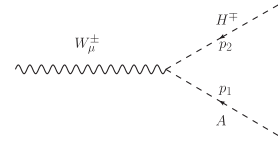
$$S_W = \sqrt{1 - (M_W/M_Z)^2} \quad \text{and} \quad C_W = \sqrt{1 - S_W^2}. \quad (\text{A3})$$



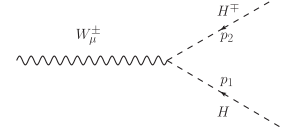
$$= e \frac{(p_1 - p_2)_\mu}{2 C_W S_W}, \quad (\text{A4})$$



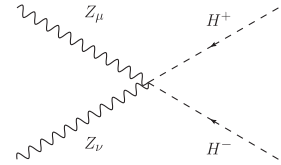
$$= i e (p_2 - p_1)_\mu, \quad (\text{A5})$$



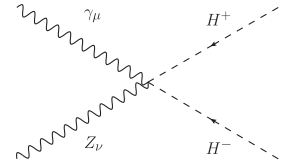
$$= -e \frac{(p_1 - p_2)_\mu}{2 S_W}, \quad (\text{A6})$$



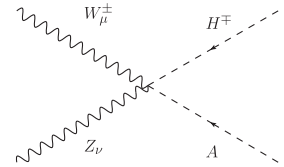
$$= \pm i e \frac{(p_1 - p_2)_\mu}{2 S_W}, \quad (\text{A7})$$



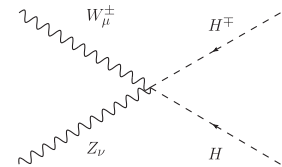
$$= i e^2 \frac{\cos^2(2\theta_W) g_{\mu\nu}}{2 C_W^2 S_W^2}, \quad (\text{A8})$$



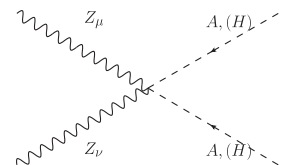
$$= i e^2 \frac{\cos(2\theta_W) g_{\mu\nu}}{C_W S_W}, \quad (\text{A9})$$



$$= \pm e^2 \frac{g_{\mu\nu}}{2 C_W}, \quad (\text{A10})$$



$$= -i e^2 \frac{g_{\mu\nu}}{2 C_W}, \quad (\text{A11})$$



$$= i e^2 \frac{g_{\mu\nu}}{2 C_W^2 S_W^2}, \quad (\text{A12})$$

$$= i 2e^2 g_{\mu\nu}, \quad (\text{A13})$$

$$= \mp e^2 \frac{g_{\mu\nu}}{2 S_W}, \quad (\text{A14})$$

$$= i e^2 \frac{g_{\mu\nu}}{2 S_W}, \quad (\text{A15})$$

$$= i e^2 \frac{g_{\mu\nu}}{2 S_W^2}, \quad (\text{A16})$$

$$= -i 6\lambda_2, \quad (\text{A17})$$

$$= -i 2\lambda_2 (-i 4\lambda_2), \quad (\text{A18})$$

$$= -i 2\lambda_2, \quad (\text{A19})$$

$$= -i \lambda_L, (-i \lambda_c), [-i \lambda_3], \quad (\text{A20})$$

$$= -i C_{eff} \delta^{ab} (p_1^\nu p_2^\mu - g^{\mu\nu} \frac{\hat{s}_{12}}{2}), \quad (\text{A21})$$

where $\hat{s}_{12} = (p_1 + p_2)^2$ and the black blob indicates to the effective vertex arising from the Lagrangian $\mathcal{L}_{\text{HEFT}}$, given in Eq. (5).

$$= -g_s C_{eff} f^{abc} [(p_1 - p_2)^\rho g^{\mu\nu} + (p_2 - p_3)^\mu g^{\nu\rho} + (p_3 - p_1)^\nu g^{\rho\mu}], \quad (\text{A22})$$

$$= i g_s^2 C_{eff} [f^{abe} f^{cde} (g^{\mu\sigma} g^{\nu\rho} - g^{\mu\rho} g^{\nu\sigma}) + f^{ace} f^{bde} (g^{\mu\sigma} g^{\nu\rho} - g^{\mu\nu} g^{\rho\sigma}) + f^{ade} f^{bce} (g^{\mu\rho} g^{\nu\sigma} - g^{\mu\nu} g^{\rho\sigma})]. \quad (\text{A23})$$

- [1] Georges Aad *et al.*, 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] Serguei Chatrchyan *et al.*, 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. Hinshaw *et al.*, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Cosmological parameter results, *Astrophys. J. Suppl. Ser.* **208**, 19 (2013).
- [4] P. A. R. Ade *et al.*, Planck 2013 results. XVI. Cosmological parameters, *Astron. Astrophys.* **571**, A16 (2014).
- [5] P. A. R. Ade *et al.*, Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, A13 (2016).
- [6] N. Aghanim *et al.*, Planck 2018 results. VI. Cosmological parameters, *Astron. Astrophys.* **641**, A6 (2020); **652**, C4(E) (2021).
- [7] Riccardo Barbieri, Lawrence J. Hall, and Vyacheslav S. Rychkov, Improved naturalness with a heavy Higgs: An alternative road to LHC physics, *Phys. Rev. D* **74**, 015007 (2006).
- [8] Marco Cirelli, Nicolao Fornengo, and Alessandro Strumia, Minimal dark matter, *Nucl. Phys.* **B753**, 178 (2006).
- [9] Adil Jueid, Jinheung Kim, Soojin Lee, So Young Shim, and Jeonghyeon Song, Phenomenology of the inert doublet model with a global $U(1)$ symmetry, *Phys. Rev. D* **102**, 075011 (2020).
- [10] Agnieszka Ilnicka, Maria Krawczyk, and Tania Robens, Inert doublet model in light of LHC run I and astrophysical data, *Phys. Rev. D* **93**, 055026 (2016).
- [11] Alexander Belyaev, Giacomo Cacciapaglia, Igor P. Ivanov, Felipe Rojas-Abatte, and Marc Thomas, Anatomy of the inert two Higgs doublet model in the light of the LHC and non-LHC dark matter searches, *Phys. Rev. D* **97**, 035011 (2018).
- [12] Abdesslam Arhrib, Yue-Lin Sming Tsai, Qiang Yuan, and Tzu-Chiang Yuan, An updated analysis of inert Higgs doublet model in light of the recent results from LUX, PLANCK, AMS-02 and LHC, *J. Cosmol. Astropart. Phys.* **06** (2014) 030.
- [13] Daniel Dercks and Tania Robens, Constraining the inert doublet model using vector boson fusion, *Eur. Phys. J. C* **79**, 924 (2019).
- [14] A. Goudelis, B. Herrmann, and O. Stål, Dark matter in the inert doublet model after the discovery of a Higgs-like boson at the LHC, *J. High Energy Phys.* **09** (2013) 106.
- [15] Nikita Blinov, Jonathan Kozaczuk, David E. Morrissey, and Alejandro de la Puente, Compressing the inert doublet model, *Phys. Rev. D* **93**, 035020 (2016).
- [16] Marco Aurelio Díaz, Benjamin Koch, and Sebastián Urrutia-Quiroga, Constraints to dark matter from inert Higgs doublet model, *Adv. High Energy Phys.* **2016**, 8278375 (2016).
- [17] Albert M Sirunyan *et al.*, Search for disappearing tracks in proton-proton collisions at $\sqrt{s} = 13$ TeV, *Phys. Lett. B* **806**, 135502 (2020).
- [18] Akanksha Bhardwaj, Partha Konar, Tanumoy Mandal, and Soumya Sadhukhan, Probing the inert doublet model using jet substructure with a multivariate analysis, *Phys. Rev. D* **100**, 055040 (2019).
- [19] A. Belyaev, T. R. Fernandez Perez Tomei, P. G. Mercadante, C. S. Moon, S. Moretti, S. F. Novaes, L. Panizzi, F. Rojas, and M. Thomas, Advancing LHC probes of dark matter from the inert two-Higgs-doublet model with the monojet signal, *Phys. Rev. D* **99**, 015011 (2019).
- [20] Xinyu Miao, Shufang Su, and Brooks Thomas, Trilepton signals in the inert doublet model, *Phys. Rev. D* **82**, 035009 (2010).
- [21] Michael Gustafsson, Sara Rydbeck, Laura Lopez-Honorez, and Erik Lundstrom, Status of the inert doublet model and the role of multileptons at the LHC, *Phys. Rev. D* **86**, 075019 (2012).
- [22] Majid Hashemi and Saereh Najjari, Observability of inert scalars at the LHC, *Eur. Phys. J. C* **77**, 592 (2017).
- [23] Amitava Datta, Nabanita Ganguly, Najimuddin Khan, and Subhendu Rakshit, Exploring collider signatures of the inert Higgs doublet model, *Phys. Rev. D* **95**, 015017 (2017).
- [24] P. Poulou, Shibananda Sahoo, and K. Sridhar, Exploring the inert doublet model through the dijet plus missing transverse energy channel at the LHC, *Phys. Lett.* **765**, 300 (2017).
- [25] Vardan Khachatryan *et al.*, Search for dark matter in proton-proton collisions at 8 TeV with missing transverse momentum and vector boson tagged jets, *J. High Energy Phys.* **12** (2016) 083; **08** (2017) 035(E).
- [26] M. Aaboud *et al.*, Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, *J. High Energy Phys.* **10** (2018) 180.
- [27] I. F. Ginzburg, K. A. Kanishev, M. Krawczyk, and D. Sokolowska, Evolution of Universe to the present inert phase, *Phys. Rev. D* **82**, 123533 (2010).
- [28] Bogumiła Świeżewska, Yukawa independent constraints for two-Higgs-doublet models with a 125 GeV Higgs boson, *Phys. Rev. D* **88**, 055027 (2013); **88**, 119903(E) (2013).
- [29] Abdesslam Arhrib, Rachid Benbrik, and Naveen Gaur, $H \rightarrow \gamma\gamma$ in inert Higgs doublet model, *Phys. Rev. D* **85**, 095021 (2012).
- [30] Michael E. Peskin and Tatsu Takeuchi, Estimation of oblique electroweak corrections, *Phys. Rev. D* **46**, 381 (1992).
- [31] Johannes Haller, Andreas Hoecker, Roman Kogler, Klaus Mönig, Thomas Peiffer, and Jörg Stelzer, Update of the global electroweak fit and constraints on two-Higgs-doublet models, *Eur. Phys. J. C* **78**, 675 (2018).
- [32] Erik Lundstrom, Michael Gustafsson, and Joakim Edsjo, The inert doublet model and LEP II limits, *Phys. Rev. D* **79**, 035013 (2009).
- [33] Genevieve Belanger, Beranger Dumont, Andreas Goudelis, Bjorn Herrmann, Sabine Kraml, and Dipan Sengupta, Dilepton constraints in the inert doublet model from run 1 of the LHC, *Phys. Rev. D* **91**, 115011 (2015).
- [34] Aaron Pierce and Jesse Thaler, Natural dark matter from an unnatural Higgs boson and new colored particles at the TeV scale, *J. High Energy Phys.* **08** (2007) 026.
- [35] Georges Aad *et al.*, Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, *Phys. Rev. D* **104**, 112010 (2021).
- [36] Shankha Banerjee, Fawzi Boudjema, Nabaran Chakrabarty, and Hao Sun, Relic density of dark matter in the inert doublet model beyond leading order for the low mass

- region: 1. Renormalisation and constraints, *Phys. Rev. D* **104**, 075002 (2021).
- [37] M. Tanabashi *et al.*, Review of particle physics, *Phys. Rev. D* **98**, 030001 (2018).
- [38] Georges Aad *et al.*, Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment, *Eur. Phys. J. C* **76**, 6 (2016).
- [39] A. M. Sirunyan *et al.*, Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV, *J. High Energy Phys.* **11** (2018) 185.
- [40] ATLAS Collaboration, Combination of searches for invisible Higgs boson decays with the ATLAS experiment, CERN Report No. ATLAS-CONF-2020-052, 2020, <https://cds.cern.ch/record/2743055>.
- [41] Y. Mambrini, Higgs searches and singlet scalar dark matter: Combined constraints from XENON100 and the LHC, *Phys. Rev. D* **84**, 115017 (2011).
- [42] Joel Giedt, Anthony W. Thomas, and Ross D. Young, Dark Matter, the CMSSM and Lattice QCD, *Phys. Rev. Lett.* **103**, 201802 (2009).
- [43] D. S. Akerib *et al.*, Results from a Search for Dark Matter in the Complete LUX Exposure, *Phys. Rev. Lett.* **118**, 021303 (2017).
- [44] E. Aprile *et al.*, Dark Matter Results from 225 Live Days of XENON100 Data, *Phys. Rev. Lett.* **109**, 181301 (2012).
- [45] V. Khachatryan *et al.*, Search for disappearing tracks in proton-proton collisions at $\sqrt{s} = 8$ TeV, *J. High Energy Phys.* **01** (2015) 096.
- [46] Adam Alloul, Neil D. Christensen, Céline Degrande, Claude Duhr, and Benjamin Fuks, FeynRules 2.0—A complete toolbox for tree-level phenomenology, *Comput. Phys. Commun.* **185**, 2250 (2014).
- [47] Celine Degrande, Automatic evaluation of UV and R2 terms for beyond the Standard Model Lagrangians: A proof-of-principle, *Comput. Phys. Commun.* **197**, 239 (2015).
- [48] 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.
- [49] S. Frixione, Z. Kunszt, and A. Signer, Three jet cross-sections to next-to-leading order, *Nucl. Phys.* **B467**, 399 (1996).
- [50] Giovanni Ossola, Costas G. Papadopoulos, and Roberto Pittau, Reducing full one-loop amplitudes to scalar integrals at the integrand level, *Nucl. Phys.* **B763**, 147 (2007).
- [51] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robert Rietkerk, Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations, *J. High Energy Phys.* **03** (2013) 015.
- [52] Torbjorn Sjostrand, Leif Lonnblad, and Stephen Mrenna, PYTHIA6.2: Physics and Manual, [arXiv:hep-ph/0108264](https://arxiv.org/abs/hep-ph/0108264).
- [53] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands, An introduction to PYTHIA8.2, *Comput. Phys. Commun.* **191**, 159 (2015).
- [54] Stefano Frixione and Bryan R. Webber, Matching NLO QCD computations and parton shower simulations, *J. High Energy Phys.* **06** (2002) 029.
- [55] Richard D. Ball *et al.*, Parton distributions with LHC data, *Nucl. Phys.* **B867**, 244 (2013).
- [56] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi, DELPHES3, A modular framework for fast simulation of a generic collider experiment, *J. High Energy Phys.* **02** (2014) 057.
- [57] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, The anti- k_t jet clustering algorithm, *J. High Energy Phys.* **04** (2008) 063.
- [58] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, FastJet user manual, *Eur. Phys. J. C* **72**, 1896 (2012).
- [59] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, Better jet clustering algorithms, *J. High Energy Phys.* **08** (1997) 001.
- [60] Andreas Hocker *et al.*, TMVA—Toolkit for multivariate data analysis, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [61] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, and Michele Treccani, Matching matrix elements and shower evolution for top-quark production in hadronic collisions, *J. High Energy Phys.* **01** (2007) 013.
- [62] Stefan Hoeche, Frank Krauss, Nils Lavesson, Leif Lonnblad, Michelangelo Mangano, Andreas Schalick, and Steffen Schumann, Matching parton showers and matrix elements, [arXiv:hep-ph/0602031](https://arxiv.org/abs/hep-ph/0602031).
- [63] David Krohn, Jesse Thaler, and Lian-Tao Wang, Jet trimming, *J. High Energy Phys.* **02** (2010) 084.
- [64] Jonathan M. Butterworth, Adam R. Davison, Mathieu Rubin, and Gavin P. Salam, Jet Substructure as a New Higgs Search Channel at the LHC, *Phys. Rev. Lett.* **100**, 242001 (2008).
- [65] Stephen D. Ellis, Christopher K. Vermilion, and Jonathan R. Walsh, Techniques for improved heavy particle searches with jet substructure, *Phys. Rev. D* **80**, 051501 (2009).
- [66] Stephen D. Ellis, Christopher K. Vermilion, and Jonathan R. Walsh, Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches, *Phys. Rev. D* **81**, 094023 (2010).
- [67] Jesse Thaler and Ken Van Tilburg, Identifying boosted objects with N-subjettiness, *J. High Energy Phys.* **03** (2011) 015.
- [68] Jesse Thaler and Ken Van Tilburg, Maximizing boosted top identification by minimizing N-subjettiness, *J. High Energy Phys.* **02** (2012) 093.
- [69] Partha Konar, Kyoungchul Kong, and Konstantin T. Matchev, $\sqrt{\hat{s}}_{\min}$: A global inclusive variable for determining the mass scale of new physics in events with missing energy at hadron colliders, *J. High Energy Phys.* **03** (2009) 085.
- [70] Partha Konar, Kyoungchul Kong, Konstantin T. Matchev, and Myeonghun Park, RECO level \sqrt{s}_{\min} and subsystem \sqrt{s}_{\min} : Improved global inclusive variables for measuring the new physics mass scale in E_T events at hadron colliders, *J. High Energy Phys.* **06** (2011) 041.
- [71] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev, and M. Park, Guide to transverse projections and mass-constraining variables, *Phys. Rev. D* **84**, 095031 (2011).