Eur. Phys. J. C (2023) 83:148
https://doi.org/10.1140/epjc/s10052-023-11303-6

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Feasibility of the observation of a heavy scalar through the fully hadronic final state at the LHeC

Elias Malwa[1,a], **Mukesh Kumar**[1,b], **Bruce Mellado**[1,2,c], **Xifeng Ruan**[1,d]

[1] School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand (Wits), Johannesburg 2050, South Africa
[2] iThemba LABS, National Research Foundation, PO Box 722, Somerset West 7129, South Africa

Received: 3 September 2022 / Accepted: 3 February 2023 / Published online: 15 February 2023

Abstract The proposed future Large Hadron electron collider provides sufficient center of mass energies, $\sqrt{s}$, to probe heavy particles decaying into $W^{\pm}(Z)-$boson of mass $> 2m_W$ $(2m_Z)$. In this work we present a study to produce one such heavy CP even scalar H of mass $2m_h < m_H < 2m_t$ through charged-current production mode where $H \to W^+W^-$, where hadronic decay of $W^{\pm}-$boson is considered to reconstruct m_H. Due to the presence of missing energy and forward jet in this channel, it is challenging to reconstruct m_H with above final state and thus we employed three different reconstruction methods and discuss the significance of each one. For this analysis we consider a benchmark value of $m_H = 270$ GeV and $\sqrt{s} \approx 1.3$ TeV with an assumed luminosity of 1 ab^{-1}.

1 Introduction

To date many existing models beyond the Standard Model (BSM) like the two-Higgs doublet models [1] and its extensions incorporates scalars of mass lower or higher than the SM Higgs-boson ($m_h = 125$ GeV) [2] with models parameters heavily constrained by existing experimental data and theoretical limits. The multi-lepton anomalies seen in Run 1 data at ATLAS and CMS are explained in a two-Higgs doublet model with additional real singlet scalar (2HDM+S) [3–8].[1] In this model the mass of the heaviest CP-even scalar H is considered in the interval $2m_h \leq m_H < 2m_t$, where m_t is the mass of top-quark. The 2HDM+S model with different mass ranges of scalars are also well motivated from theories BSM [10–14], possibilities of existence of BSM scalars at the large Hadron collider data [15,16] and future e^+e^- collider [17,18], to explain dark matter abundance [19–21], di-Higgs production [22], excess seen at 96 GeV [23,24] and to explain recent CDF [25] W-mass measurements [26]. Heavy scalars searches in WW/ZZ channels are considered at CMS and ATLAS [27–29]. The discovery potential of heavy Higgs-boson through the resonant di-Higgs production in HL-LHC and FCC-hh has been studied with 4τ and $bb\gamma\gamma$ channels in "xSM" model [30].[2] Even the physics of dark matter and axions or axions like particles can be connected with CP-even or odd scalars [31–33].

In this work we investigate the possibility of probing H at the proposed future electron-proton colliders *via* the deep-inelastic scattering charged-current (CC) process. The proposed large Hadron electron collider (LHeC) facility at CERN provides sufficient center of mass energy $\sqrt{s} \approx 1.3$ TeV following electron (proton) energy of $E_{e(p)} = 60$ GeV (7 TeV) to explore the allowed mass range of H. Interestingly with this mass range one can explore the resonance H via its decay to $W^{\pm}$ and $Z-$bosons. In this work we consider $H \to W^+W^-$, where $W^{\pm}$ decay to hadronic final states. However, the mass reconstruction of H through this final state is challenging due to (a) the $W^{\pm}-$boson emanating from heavy H is boosted with respect to the laboratory system, and hence the jets coming from $W^{\pm}$ are collimated, and (b) in the e^-p production process, the scattered jet from the proton-line is not easily distinguishable from jets coming from $W^{\pm}$ (Fig. 1). However, the high rapidity (η_j) region of the scattered jets can be exploited to reconstruct the signal. We also employ a machine learning approach to distinguish the signal and potential backgrounds in this work.

[1] For a recent review of anomalies see Ref. [9].

[a] e-mail: elias.malwa@cern.ch
[b] e-mail: mukesh.kumar@cern.ch (corresponding author)
[c] e-mail: bmellado@mail.cern.ch
[d] e-mail: xifeng.ruan@cern.ch

[2] The "xSM" model is the extension of the SM scalar sector with a single real singlet scalar.

In Sect. 2 we discuss the framework needed to perform this analysis. A description of event simulation and tools needed are discussed in Sect. 3. The mass reconstruction methods are described in Sect. 4. Summary and discussion of this work is presented in Sect. 5.

2 Model

To investigate the discovery potential of heavy Higgs boson of mass $2m_h \leq m_H < 2m_t$ in $e^- p$ environment, we consider a model where H corresponds to a real singlet scalar field Φ_H which mixes with the SM $SU(2)$ doublet Higgs field Φ. Then the Higgs-boson Lagrangian will be modified and can be written as [34–36]:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^2 + (\partial_\mu \Phi_H)^2 + \mu_h^2 |\Phi|^2 - \lambda_h |\Phi|^4 + \mu_H^2 |\Phi_H|^2 - \lambda_H |\Phi_H|^4 + \xi |\Phi|^2 |\Phi_H|^2 . \quad (1)$$

In general the, parameters μ_h, μ_H, λ_h and λ_H are all positive in order to have stable potential but ξ may not require any particular sign. We assume that in the above Lagrangian the scalar fields acquire a vacuum expectation values and hence the component fields can be written as:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} G^\pm \\ \phi + \mathrm{v} + iG^0 \end{pmatrix}, \Phi_H = \frac{1}{\sqrt{2}} \left(\phi_H + \mathrm{v}_H + iG'\right). \quad (2)$$

Here the fields G are Goldstone bosons absorbed by the vector bosons, and so no physical pseudoscalar states are left in the spectrum. But the scalar spectrum has two physical states h and H rather than just one of the SM. Also since the singlet do not couple to the $SU(2)_L \times U(1)_Y$ gauge bosons, they do not contribute to m_W and m_Z and hence v must take the SM value $\mathrm{v} = 246$ GeV. We can also redefine the coefficient of Eq. (1) such that $\mathrm{v}_H = 0$. Note that we are not imposing any extra possible symmetries like $\mathbb{Z}_2$ in the scalar sector, and in general ϕ will mix with the ϕ_H to form the mass eigenstates. We assume the masses of h and H as in previous case, $m_h < m_H$, where $m_h = 125$ GeV is taken as the SM Higgs boson and m_H as mass of the heavy scalar singlet. The mass eigenstates h and H are related to the gauge eigenstates ϕ and ϕ_H by a 2×2 unitary matrix [3] V:

$$\begin{pmatrix} \phi \\ \phi_H \end{pmatrix} = V \begin{pmatrix} h \\ H \end{pmatrix}. \quad (3)$$

Hence the couplings of the gauge bosons and fermions with h will be same as in the SM if $|V_{11}| = 1$ which implies $|V_{12}| = \sqrt{1 - |V_{11}|^2} = 0$. However in this work we considered $|V_{11}| \neq 1$ and $|V_{12}| \neq 0$. Then the production rates of the h and H are suppressed by a factor $|V_{1i}|^2$ relative to the SM h production rates. The branching ratios (BRs) of h to the SM particles are identical to the SM BRs, while the BRs of heavy H depend on whether the channel $H \to hh$ are kinematically accessible. For our analysis we scale the HW^+W^- coupling with respect to the SM Higgs boson hW^+W^- coupling.

3 Event simulation and tools

The simulation of CC process (signal) for the heavy scalar H production follows through $pe^- \to \nu_e Hj$, where ν_e is electron-neutrino (and is the source of missing energy) and j represents jets emanating from proton-line (we refer to this j as scattered or forward jet in the text). Further the decay of $H \to W^+W^-$ and $W^\pm \to jj$ is taken at the matrix element level for this signal process (see Fig. 1). Note that H can also be produced in neutral current process through the fusion of Z-bosons at tree-level as $pe^- \to e^- Hj$, but the cross-section is sub-dominant and approximately 5.5 times smaller than the CC process which follows through $W^\pm$-fusion for unpolarized e^- beam.

To generate event samples for signal and potential backgrounds we use a Monte Carlo generator `MadGraph5` [37], interfaced with a customised `Pythia-PGS` [38] for parton showers and hadronization (for details see Ref. [39]). The detector simulation is performed using `Delphes` [40] with parameters optimised for the detector in LHeC. The jets are clustered using `FastJet` [41] with the anti-k_T algorithm [42] and distance parameter $R = 0.4$. The factorisation

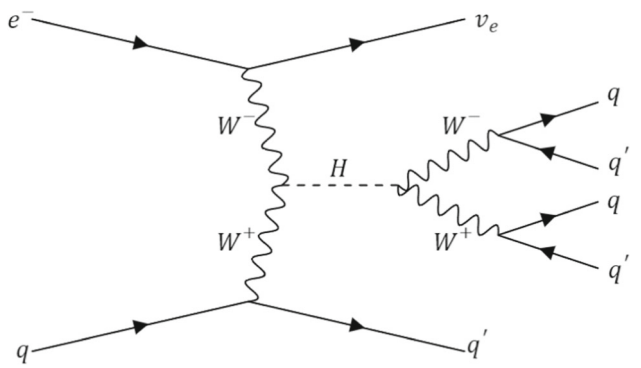

Fig. 1 Leading order diagram for signal process $pe^- \to \nu_e Hj$, $H \to W^+W^-$, $W^\pm \to jj$. Here, $q \equiv u, c, \bar{d}, \bar{s}$ and $q' \equiv d, s, \bar{u}, \bar{c}$

[3] In general, a 2×2 unitary matrix V can be formed with one parameter θ as:

$$V = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \equiv \begin{pmatrix} V_{11} & V_{12} \\ -V_{12} & V_{11} \end{pmatrix},$$

where $|V_{11}|^2 + |V_{12}|^2 = 1$.

Table 1 Total cross-sections (in fb) for signal production (see text) and potential backgrounds with $E_e = 60$ GeV and $E_p = 7$ TeV. The polarisation of e^- is taken to be -80%. The first row represents the signal process and the other four rows are for the dominant background processes

Process	Cross section (fb)
Signal	0.49
$e^- W^+ W^- j$	26.7
$e^- ZZj$	0.13
$\nu_e W^+ W^- j$	7.66
$\nu_e ZZj$	2.54

and renormalisation scales for the signal simulation are fixed to the heavy Higgs boson mass m_H. The background simulations are done with the default `MadGraph5` dynamic scales. The polarization of the charged electron is assumed to be -80%. This enhances the polarized cross-sections by ~ 1.8 times with respect to the unpolarized e^- beam for both signal and background.

An estimation of cross-section for the signal[4] and potential background processes are calculated at leading order using `MadGraph5` with applied minimal cuts on transverse momentum of jets $p_{T_j} > 20$ GeV, jet pseudo-rapidity $-1 < \eta_j < 5$ and there is no requirements for transverse missing energy E_T^{miss}, and presented in Table 1 for a benchmark value of $m_H = 270$ GeV. Before going for mass reconstruction of H with appropriate methodologies we made preliminary selection criteria to estimate the significance, and those are as follows: (a) since the final state of signal (Fig. 1) contains five jets at matrix element level (four from decay of $W^{\pm}$-boson and one scattered jet), we chose at least five leading p_T-ordered jets in simulated events and (b) $E_T^{miss} > 20$ GeV. In Table 2 we presented the number of weighted events of signal (S) and backgrounds (B) at luminosity $\mathcal{L} = 1$ ab^{-1} after these selection criteria where in the last column significance of signal over background is calculated with formula $\sigma = S/\sqrt{B}$. It is interesting to note the there is slight increase ($\approx 2.3\%$) in σ after the selection of five leading jets, though $E_T^{miss} > 20$ GeV reduces the σ by $\approx 7\%$ in comparison with initial weighted events. In order to estimate the systematic errors in the shape of signal and background distributions due to detector resolution, E_T^{miss} measurement, reconstruction efficiency etc., as well as on the expected number of events we calculate significance as function of systematic factor δ_{sys}: $\sigma(\delta_{sys}) = S/\sqrt{B + (\delta_{sys} \cdot B)^2}$ and added the estimation in Table 2.

[4] We scaled the HW^+W^- coupling such that the cross-section for signal should be ~ 20 times less than the corresponding cross-section of h with $m_h = 270$ GeV. This factor is very optimistic in order to not evade any theoretical and experimental limits for m_H cross-section in the considered signal.

It is important to investigate and account for these observations during the mass reconstruction procedure of H and further discuss in next section.

4 Reconstruction of the invariant mass

In order to reconstruct m_H it is important to select appropriate hadronic jets in our signal and observe the features with respect to the dominant backgrounds. To begin the procedure we must isolate and identify the hadronic jets after detector simulations. In Fig. 2a, number of hadronic jets are shown which are constructed with requirement on $\Delta R = 0.4$.[5] It is clear that the number of hadronic jets from ZZ backgrounds are competitive in comparison to the signal. Also a similar feature can be observed in the pseudo-rapidity of forward jets, η_j, as shown in Fig. 2b. And therefore, the ZZ backgrounds needs to be optimize with the help of missing transverse energy cut $E_T^{miss} > 20$ GeV (see Fig. 3) and corresponding significant reduction in weighted events can be seen in Table 2.

To compare the reconstructed invariant m_H with the truth-level mass, the hadronic jets originated from W^+ and W^- bosons are selected using the truth-level information (note that $W^{\pm}$ are decaying from H in signal). An illustration of invariant mass of two-jets, m_{jj}, from W^+ (W^-) is shown in Fig. 4a (Fig. 4b). Note that along with signal we only showed backgrounds with $W^{\pm}$ final states as there is no information stored for Z-bosons in truth-level.

After analysing these observable, we apply three different methodologies to reconstruct m_H in the mentioned channel and compare the significance. In Method 1, selection of four p_T-ordered leading jets are considered. Method 2 is to select four hadronic jets excluding the most forward jet (which corresponds to largest η_j), while a high-level machine learning (ML) techniques used in Method 3.

4.1 Method 1: selection of four p_T-ordered leading jets

In this method, all jets are sorted according to the corresponding p_T and the four out-of five leading (p_T-ordered) jets are selected from the weighted signal and background events. We expect an inherent uncertainty in this method from the forward jet (which may not originate from either of W^+ or W^-) and this may contaminate the reconstruction of m_H in the signal. The invariant mass distribution of four selected jets, $m_{4j} \equiv m_H$, using this method is shown in Fig. 5a. The corresponding significance σ are shown in Table 3 (second column). Here $\sigma_{m_{4j}}$ represents the significance in full available

[5] The distance parameter ΔR between any two particles is defined as: $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, where ϕ and η are the azimuthal angle and rapidity, respectively, of particles into consideration.

Table 2 A summary table of event selections. In the first column the selection criteria are given. The second column contains the weight of the signal process $pe^- \to \nu_e Hj$, $H \to W^+W^-$, $W^\pm \to jj$ for $m_H = 270$ GeV. From column third to sixth dominant weights for backgrounds are given. Seventh column is weighted total number of backgrounds. All weights are calculated with $\mathcal{L} = 1$ ab^{-1}. The significance of signal over total background is given in the eight column. In the last column significance with $\delta_{sys} = 2\%$ is estimated

Cuts	Signal (S)	e^-WW+j	e^-ZZ+j	$\nu_e WW+j$	$\nu_e ZZ+j$	Total background (B)	$S/\sqrt{B}$	$\sigma(\delta_{sys})$
Initial	499	2680	128	7660	2540	13008	4.4	1.8
At least $5j$	211	264	20	1390	568	2242	4.5	3.2
$E_T^{miss} > 20$ GeV	182	52	4	1330	542	1928	4.1	3.1

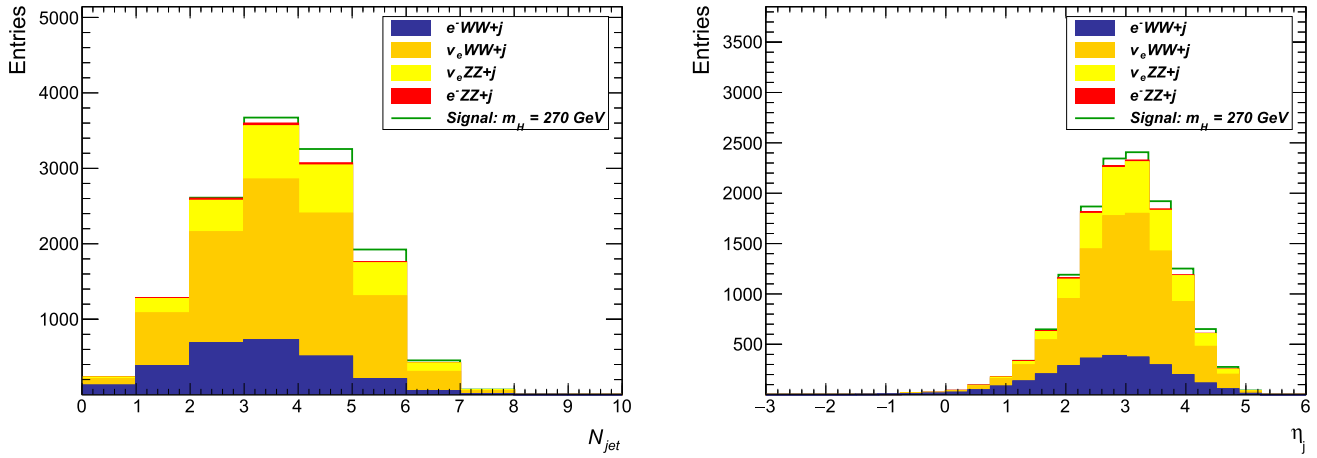

Fig. 2 **a** Multiplicity of jets in signal and backgrounds. **b** The pseudo-rapidity distribution of the forward jet after five jet selection in signal and backgrounds

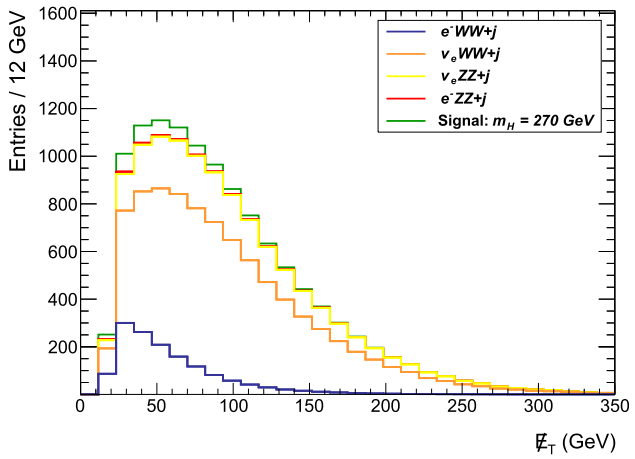

Fig. 3 The missing transverse energy distribution after applying the > 20 GeV requirement

range in m_{4j}, and σ_{max} is the range where maximum σ can be achieved. This method results maximum of 4.0σ within the invariant mass-range of $m_{4j} \in [190, 540]$ GeV and the improvement from full range of m_{4j} is by 2.5% with initial events. However after selecting $E_T^{miss} > 20$ GeV, accuracy of measurement improves with 4.9σ in $m_{4j} \in [190, 540]$ GeV (4.1% improvement from full range). And an improvement of $\sim$ 16% in comparison with significance shown in Table 2.

From distribution of m_{4j} in Fig. 5a it is noticed that the width of invariant mass is wide and reason for this could be the contamination of forward jets as discussed. Thus a method to narrower the width suppose to result better mass reconstruction by removing the forward jet and discussed in next subsection.

4.2 Method 2: elimination of forward jet

As Method 1 slightly improved the accuracy in the measurement of m_H through four p_T-ordered leading jets using m_{4j} (comparing the significance obtained in Table 2), we employ a second approach where forward jet corresponding to largest η_j are eliminated and remaining four p_T-ordered jets are selected. In addition we also verified that the selected jets originate from $W^\pm$-bosons using the truth-level information. The corresponding invariant mass distribution is shown in Fig. 5b. Clearly the m_{4j} distribution has narrower width comparing with Method 1 (Fig. 5a) and this approach should improve the accuracy of measuring m_H. This approach also

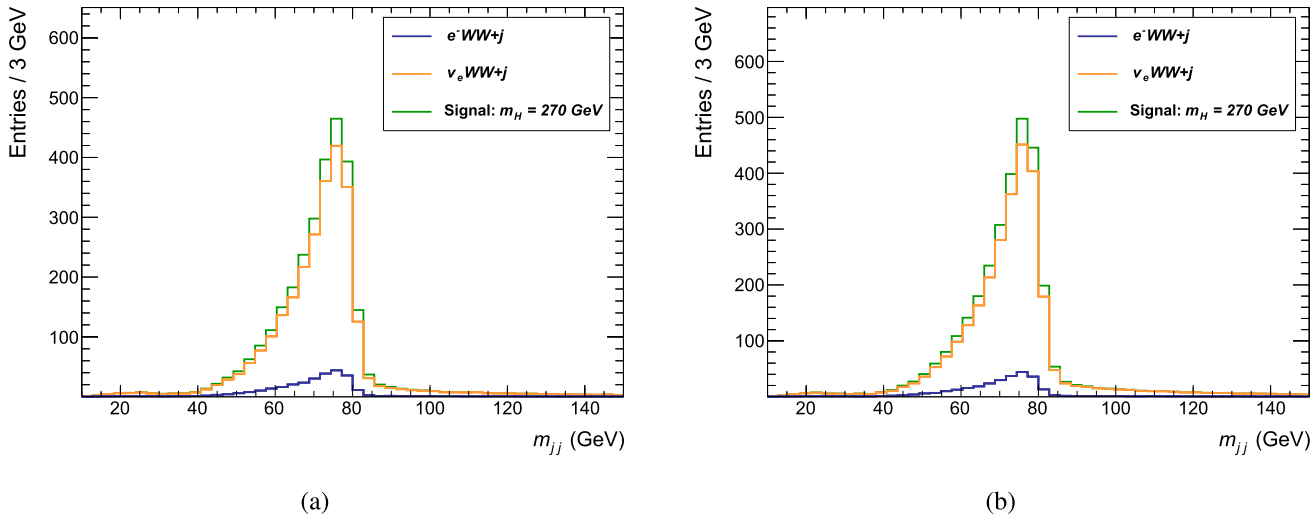

Fig. 4 Invariant di-jet mass distribution m_{jj} from truth-level information of **a** W^+ and **b** W^-, where $H \to W^+W^-$ with $m_H = 270$ GeV

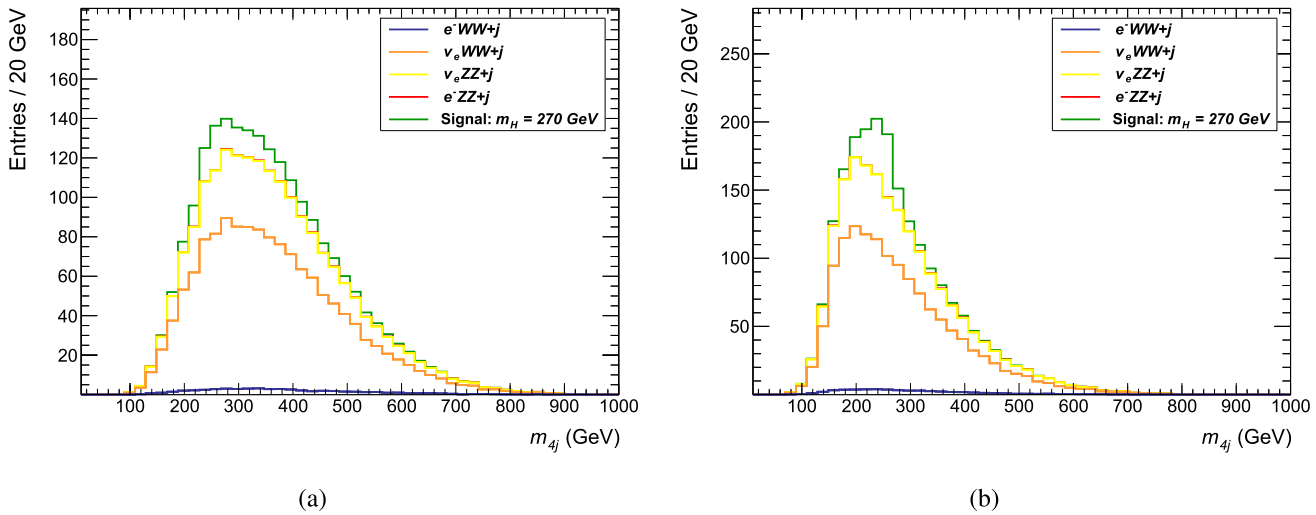

Fig. 5 **a** Invariant mass distribution of four p_T-ordered leading jets [Method 1 (Sect. 4.1)]. **b** Invariant mass distribution of four p_T-ordered jets by removing the forward jet [Method 2 (Sect. 4.2)]

uses the same number of initial weighted events as the above method. When reconstructing the invariant mass of H, this method achieved a maximum significance of 5.0σ before applying the missing energy cut. A maximum significance of 6.1σ can be attained with 24% improvement after selecting jets with $E_T^{miss} > 20$ GeV. In Table 3 (third column) significance obtained for Method 2 is shown. Overall applying this method shows improvement in significance of about 33% in comparison with significance obtained selecting at least $5j$ with $E_T^{miss} > 20$ GeV as in Table 2.

4.3 Method 3: machine learning technique

Though the use of Method 2 results a higher significance of about 6σ shows the efficacy of this approach to reconstruct m_H, we also analyse the event samples using high-level machine learning technique as Method 3 and compare the significance. For our analysis we employed the Toolkit for Multivariate Data Analysis (TMVA) package [43] in which all multivariate methods respond to supervised learning only, i.e., the input information is mapped in feature space to the desired outputs.

To start with, the four-momentum information of jets from the signal and backgrounds' event samples are used to construct the low-level observables like jet's transverse momenta p_{T_j}, pseudo-rapidity η_j, azimuthal angle ϕ_j, energy E_j and mass m_j. The signal samples with these observables are passed in two equal proportions for training and testing, respectively, to reconstruct m_{4j}. Here we include three different analysis routines known as: Boosted Decision Trees with

Table 3 The significance is calculated at each stage of the optimised selection criteria using $\sigma = S/\sqrt{B}$ and $\sigma(\delta_{sys} = 2\%) = S/\sqrt{B + (\delta_{sys} \cdot B)^2}$ where S and B are the expected signal and background yields at a luminosity of 1 ab^{-1} respectively. Here $\sigma_{m_{4j}}$ represents the significance in full available range in m_{4j}. And $\sigma_{max}(m_{4j})$ is the range where maximum σ can be achieved, corresponding minimum to maximum range $m_{4j} \in [m_{4j}^{min}, m_{4j}^{max}]$ are specified for each approach (corresponding S and B are given in the next row)

	Method 1	Method 2	Method 3		
			BDTG	DNN	LD
Initial					
$\sigma_{m_{4j}}$	3.9σ	3.9σ	3.9σ	3.9σ	3.9σ
$\sigma_{max}(m_{4j} \in [m_{4j}^{min}, m_{4j}^{max}])$	$4.0\sigma \in [190, 540]$	$5.0\sigma \in [210, 280]$	$4.0\sigma \in [210, 270]$	$4.2\sigma \in [215, 270]$	$3.9\sigma \in [225, 270]$
$S\ (B)$	257 (4145)	187 (1380)	243 (3712)	237 (3258)	237 (3689)
$\sigma(\delta_{sys} = 2\%)$	2.4σ	4.0σ	2.5σ	2.7σ	2.5σ
$E_T^{miss} > 20$ GeV					
$\sigma_{m_{4j}}$	4.7σ	4.7σ	4.7σ	4.7σ	4.7σ
$\sigma_{max}(m_{4j} \in [m_{4j}^{min}, m_{4j}^{max}])$	$4.9\sigma \in [190, 540]$	$6.1\sigma \in [210, 280]$	$4.8\sigma \in [205, 270]$	$4.9\sigma \in [210, 270]$	$4.8\sigma \in [220, 270]$
$S\ (B)$	222 (2088)	161 (691)	211 (1941)	206 (1755)	214 (1955)
$\sigma(\delta_{sys} = 2\%)$	3.6σ	5.4σ	3.6σ	3.8σ	3.6σ

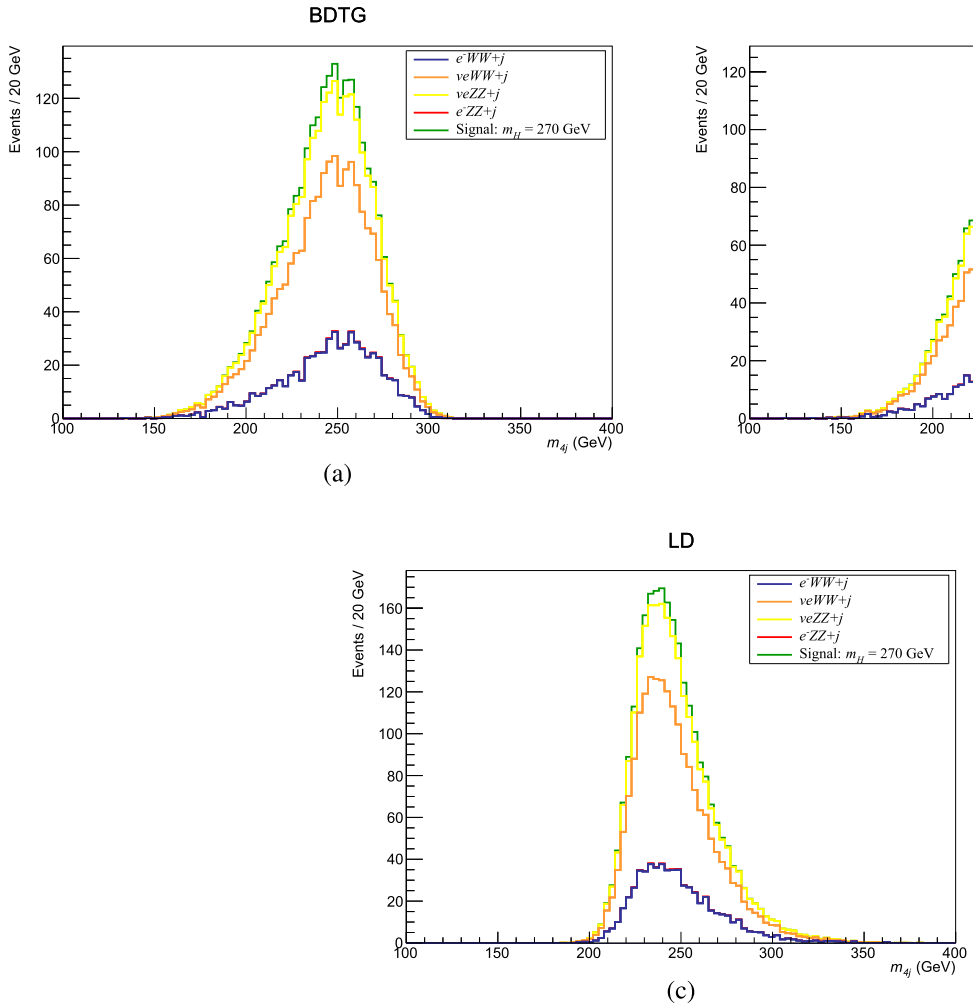

Fig. 6 Invariant mass distribution m_{4j} of the trained signal and evaluated background sample using the BDTG, DNN and LD method

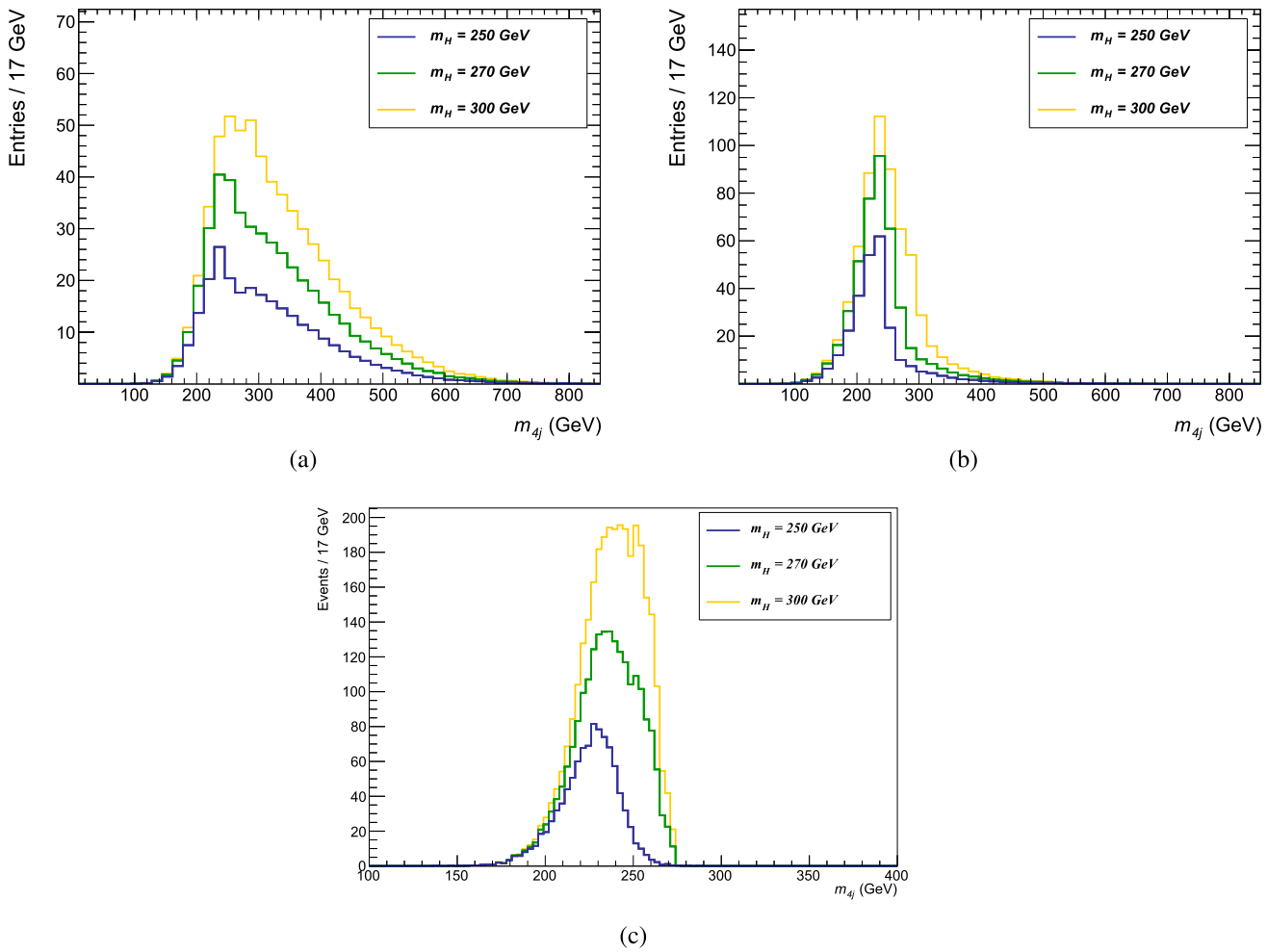

Fig. 7 Comparison of m_{4j} (signal only) for three different masses: $m_H = 250$, 270 and 300 GeV following **a** Method 1 (Sect. 4.1), **b** Method 2 (Sect. 4.2) and **c** DNN method (Sect. 4.3)

gradient boosting (BDTG), Deep Neural Network (DNN) and Linear Discriminator (LD). The details of all three analysis procedure and mechanism are documented in Ref. [43]. All background samples are passed through evaluation with default parameters in TMVA regression application with Boosted Decision Trees (BDTG), Deep Neural Networks (DNN) and Linear Discriminants (LD). The combination of outputs are shown in Fig. 6. The default parameters are later tested and tuned to give maximum significance with target mass as 270 GeV.[6] In Table 3, the significance obtained through all three analysis techniques are presented. All three analysis routines provides the maximum significance of mass measurement $\sim 5\sigma$, which is a little less in comparison with Method 2 while is similar to Method 1. Though the improvements after $E_T^{miss} > 20$ GeV requirement are high in comparison with Method 1. However among the three analysis routines the DNN performance is better with maximum significance of 4.9σ in $m_{4j} \in [210, 270]$.

By analysing the m_{4j} distributions shown in Fig. 6 the ML algorithms used here seems to accumulate the signal as well as the backgrounds region towards the target mass. Though the significance are consistent with other two methods and even better than Method 1 by using DNN as shown in Table 3.

4.4 Scanning m_H

Among the three methods, the Method 2 - elimination of forward jet corresponding to the largest η_j is the most efficient to reconstruct the m_H. So we will use this technique for two different $m_H = 250$ and 300 GeV, and compare the significance with the benchmark $m_H = 270$ GeV taken in this study to understand how other masses affect the sensitivity of measurement method(s). This will allow us to investigate such masses at LHeC with considered $\sqrt{s} \approx 1.3$ TeV. For completeness we also analyse and compare the significance

[6] This mass is set as data-loader and defined in testing and training sample dataset as truth mass. The target mass is the reconstructed m_{4j}, where the selected four jets originate from $W^{\pm}$ as in Method 2.

Table 4 Same as Table 3 for $m_H = 250$ and 300 GeV in comparison with $m_H = 270$ GeV

$E_T^{miss} > 20$ GeV	Method 1	Method 2	DNN
$m_H = 250$ GeV			
$\sigma_{max}(m_{4j} \in [m_{4j}^{min}, m_{4j}^{max}])$	$5.5\sigma \in [160, 470]$	$7.0\sigma \in [190, 250]$	$5.8\sigma \in [170, 240]$
$\sigma(\delta_{sys} = 2\%)$	4.2σ	6.2σ	5.0σ
$m_H = 300$ GeV			
$\sigma_{max}(m_{4j} \in [m_{4j}^{min}, m_{4j}^{max}])$	$3.9\sigma \in [220, 580]$	$5.0\sigma \in [240, 310]$	$4.1\sigma \in [237, 310]$
$\sigma(\delta_{sys} = 2\%)$	2.9σ	4.4σ	3.1σ
$m_H = 270$ GeV			
$\sigma_{max}(m_{4j} \in [m_{4j}^{min}, m_{4j}^{max}])$	$4.9\sigma \in [190, 540]$	$6.1\sigma \in [210, 280]$	$4.9\sigma \in [210, 270]$
$\sigma(\delta_{sys} = 2\%)$	3.6σ	5.4σ	3.8σ

with Method 1 and DNN routines (as this method gives highest significance in comparison to BDTG and LD).

In Fig. 7a–c we compare m_{4j} (signal only) using Method 1, Method 2 and DNN routines for $m_H = 250$, 270 and 300 GeV, respectively. In Table 4, the maximum significance obtained using Method 1, Method 2 and DNN are shown as in Table 3. A comparison with $m_H = 270$ GeV shows $\sim 1\sigma$ difference in significance for both masses. Since the cross-section of $m_H = 250$ (300) GeV is higher (lower) than the corresponding cross-section of $m_H = 270$ GeV, the enhancement (suppression) in significance is expected.

5 Discussion and summary

The existence of heavy particles are usually known in physics BSM and strategies to search such particles in colliders are very important. Specially in the scalar-sector it is most important since these particles are responsible for mass generation of several bosons and fermions in SM as well in BSM. In this article we attempted to prescribe mass reconstruction methods for a heavy scalar boson in a mass range of $m_H \in (2m_h, 2m_t)$, where H particularly decays to hadronic jets through $W^{\pm}$ and the production is followed through charged-current in the LHeC environment.

As a benchmark, a heavy scalar of mass $m_H = 270$ GeV produced in CC channel in LHeC with $E_e = 60$ GeV and $E_p = 7$ TeV. Further we considered $H \to W^+W^-$ and $W^{\pm} \to jj$ channel to develop a prescription for mass reconstruction. In doing so we explained the possible methods of selecting final state hadronic jets as the scattered jets in this channel are the source of contamination. Overall Method 2 gives a significance of about 6σ using m_{4j}, which is better compared to the other two methodologies discussed. It is also noted that $E_T^{miss} > 20$ GeV plays a significant role to improve the significance only when a proper selection of four hadronic jets are taken out of at least five jets. Similarly, a significant results for mass reconstruction of $m_H = 250$ and 300 GeV with 7σ and 5σ, respectively, indicates the efficiency to discover such heavy masses at future LHeC. By accounting for the systematics effect of 2% mentioned, the significance reduces from 6.1σ to 5.4σ for $m_H = 270$ GeV in Method 2.

Future opportunities: A similar analysis can be performed with $H \to ZZ$, $Z \to \ell^+\ell^-$, jj in addition with the neutral current channel $pe^- \to e^-Hj$. Also, these studies can be carried forward in the HL-LHC and proposed FCC facilities.

Acknowledgements The Institute for Collider Particle Physics is grateful for the support from the National Research Foundation, the National Institute of Theoretical Physics and the Department of Science and Technology through the SA-CERN consortium and other forms of support.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: For this work authors used publicly available data where necessary, and are referenced properly.]

Funded by SCOAP[3]. SCOAP[3] supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

1. G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, M. Sher, J.P. Silva, Theory and phenomenology of two-Higgs-doublet models. Phys. Rep. **516**, 1–102 (2012)

2. G. 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**, 1–29 (2012)
3. S. von Buddenbrock, N. Chakrabarty, A.S. Cornell, D. Kar, M. Kumar, T. Mandal, B. Mellado, B. Mukhopadhyaya, R.G. Reed, The compatibility of LHC Run 1 data with a heavy scalar of mass around 270 GeV. arXiv:1506.00612 [hep-ph]
4. S. von Buddenbrock, A.S. Cornell, A. Fadol, M. Kumar, B. Mellado, X. Ruan, Multi-lepton signatures of additional scalar bosons beyond the Standard Model at the LHC. J. Phys. G **45**(11), 115003 (2018)
5. S. von Buddenbrock, N. Chakrabarty, A.S. Cornell, D. Kar, M. Kumar, T. Mandal, B. Mellado, B. Mukhopadhyaya, R.G. Reed, X. Ruan, Phenomenological signatures of additional scalar bosons at the LHC. Eur. Phys. J. C **76**(10), 580 (2016)
6. S. Buddenbrock, A.S. Cornell, Y. Fang, A. Fadol Mohammed, M. Kumar, B. Mellado, K.G. Tomiwa, The emergence of multi-lepton anomalies at the LHC and their compatibility with new physics at the EW scale. JHEP **10**, 157 (2019)
7. S. von Buddenbrock, R. Ruiz, B. Mellado, Anatomy of inclusive $t\bar{t}W$ production at Hadron colliders. Phys. Lett. B **811**, 135964 (2020)
8. Y. Hernandez, M. Kumar, A.S. Cornell, S.-E. Dahbi, Y. Fang, B. Lieberman, B. Mellado, K. Monnakgotla, X. Ruan, S. Xin, The anomalous production of multi-lepton and its impact on the measurement of Wh production at the LHC. Eur. Phys. J. C **81**(4), 365 (2021)
9. O. Fischer et al., Unveiling hidden physics at the LHC. Eur. Phys. J. C **82**(8), 665 (2022)
10. M. Muhlleitner, M.O.P. Sampaio, R. Santos, J. Wittbrodt, The N2HDM under theoretical and experimental scrutiny. JHEP **03**, 094 (2017)
11. M. Mühlleitner, M.O.P. Sampaio, R. Santos, J. Wittbrodt, Phenomenological comparison of models with extended Higgs sectors. JHEP **08**, 132 (2017)
12. M. Krause, D. Lopez-Val, M. Muhlleitner, R. Santos, Gauge-independent renormalization of the N2HDM. JHEP **12**, 077 (2017)
13. I. Engeln, M. Mühlleitner, J. Wittbrodt, N2HDECAY: Higgs Boson decays in the different phases of the N2HDM. Comput. Phys. Commun. **234**, 256–262 (2019)
14. P.M. Ferreira, M. Mühlleitner, R. Santos, G. Weiglein, J. Wittbrodt, Vacuum instabilities in the N2HDM. JHEP **09**, 006 (2019)
15. A. Arhrib, R. Benbrik, M. El Kacimi, L. Rahili, S. Semlali, Extended Higgs sector of 2HDM with real singlet facing LHC data. Eur. Phys. J. C **80**(1), 13 (2020)
16. T. Biekötter, A. Grohsjean, S. Heinemeyer, C. Schwanenberger, G. Weiglein, Possible indications for new Higgs bosons in the reach of the LHC: N2HDM and NMSSM interpretations. Eur. Phys. J. C **82**(2), 178 (2022)
17. D. Azevedo, P. Ferreira, M. Mühlleitner, R. Santos, J. Wittbrodt, Extended Higgs sectors at future e^+e^- colliders (2018)
18. D. Azevedo, P. Ferreira, M.M. Mühlleitner, R. Santos, J. Wittbrodt, Models with extended Higgs sectors at future e^+e^- colliders. Phys. Rev. D **99**(5), 055013 (2019)
19. S. Glaus, M. Mühlleitner, J. Müller, S. Patel, R. Santos, Electroweak corrections to dark matter direct detection in the dark singlet phase of the N2HDM. Phys. Lett. B **833**, 137342 (2022)
20. D. Azevedo, P. Gabriel, M. Muhlleitner, K. Sakurai, R. Santos, One-loop corrections to the Higgs boson invisible decay in the dark doublet phase of the N2HDM. JHEP **10**, 044 (2021)
21. I. Engeln, P. Ferreira, M.M. Mühlleitner, R. Santos, J. Wittbrodt, The Dark phases of the N2HDM. JHEP **08**, 085 (2020)
22. H. Abouabid, A. Arhrib, D. Azevedo, J.E. Falaki, P.M. Ferreira, M. Mühlleitner, R. Santos, Benchmarking di-Higgs production in various extended Higgs sector models. JHEP **09**, 011 (2022)
23. T. Biekötter, M. Chakraborti, S. Heinemeyer, A 96 GeV Higgs boson in the N2HDM. Eur. Phys. J. C **80**(1), 2 (2020)
24. S. Heinemeyer, C. Li, F. Lika, G. Moortgat-Pick, S. Paasch, Phenomenology of a 96 GeV Higgs boson in the 2HDM with an additional singlet. Phys. Rev. D **106**(7), 075003 (2022)
25. T. Aaltonen et al., High-precision measurement of the W boson mass with the CDF II detector. Science **376**(6589), 170–176 (2022)
26. T. Biekötter, S. Heinemeyer, G. Weiglein, Excesses in the low-mass Higgs-boson search and the W-boson mass measurement. **4** (2022)
27. Search for high mass resonances decaying into W^+W^- in the dileptonic final state with 138 fb^{-1} of proton–proton collisions at $\sqrt{s} = 13$ TeV (2022)
28. M. Aaboud et al., Searches for heavy ZZ and ZW resonances in the $\ell\ell qq$ and $\nu\nu qq$ final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. JHEP **03**, 009 (2018)
29. G. Aad et al., Search for heavy resonances decaying into a pair of Z bosons in the $\ell^+\ell^-\ell'^+\ell'^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states using 139 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Eur. Phys. J. C **81**(4), 332 (2021)
30. A. Abada et al., FCC-hh: the Hadron collider: future circular collider conceptual design report volume 3. Eur. Phys. J. ST **228**(4), 755–1107 (2019)
31. N.F. Bell, G. Busoni, I.W. Sanderson, Self-consistent dark matter simplified models with an s-channel scalar mediator. JCAP **03**, 015 (2017)
32. N.F. Bell, G. Busoni, I.W. Sanderson, Two Higgs doublet dark matter portal. JCAP **01**, 015 (2018)
33. F.A. Ghebretinsaea, Z.S. Wang, K. Wang, Probing axion-like particles coupling to gluons at the LHC. JHEP **07**, 070 (2022)
34. R.M. Schabinger, J.D. Wells, A Minimal spontaneously broken hidden sector and its impact on Higgs boson physics at the large Hadron collider. Phys. Rev. D **72**, 093007 (2005)
35. S. Dawson, W. Yan, Hiding the Higgs boson with multiple scalars. Phys. Rev. D **79**, 095002 (2009)
36. S. Dawson, M. Sullivan, Enhanced di-Higgs boson production in the complex Higgs singlet model. Phys. Rev. D **97**(1), 015022 (2018)
37. J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, T. Stelzer, MadGraph 5: going beyond. JHEP **06**, 128 (2011)
38. T. Sjostrand, S. Mrenna, P.Z. Skands, PYTHIA 6.4 physics and manual. JHEP **05**, 026 (2006)
39. M. Kumar, X. Ruan, R. Islam, A.S. Cornell, M. Klein, U. Klein, B. Mellado, Probing anomalous couplings using di-Higgs production in electron–proton collisions. Phys. Lett. B **764**, 247–253 (2017)
40. J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, M. Selvaggi, DELPHES 3, a modular framework for fast simulation of a generic collider experiment. JHEP **02**, 057 (2014)
41. M. Cacciari, G.P. Salam, G. Soyez, FastJet user manual. Eur. Phys. J. C **72**, 1896 (2012)
42. M. Cacciari, G.P. Salam, G. Soyez, The anti-ktjet clustering algorithm. J. High Energy Phys. **2008**, 063–063 (2008)
43. A. Hocker, et al., TMVA—toolkit for multivariate data analysis. **3** (2007)