



Sensitivities of new resonance couplings to W -bosons at the LHC*

Ying-nan Mao (毛英男)[†]  Kechen Wang (王科臣)[‡]  Yiheng Xiong (熊以恒)[§]

Department of Physics, School of Physics and Mechanics, Wuhan University of Technology, Wuhan 430070, China

Abstract: We propose a search strategy at the HL-LHC for a new neutral particle X that couples to W -bosons, using the process $pp \rightarrow W^\pm X (\rightarrow W^+ W^-)$ with a tri- W -boson final state. Focusing on events with two same-sign leptonic W -boson decays into muons and a hadronically decaying W -boson, our method leverages the enhanced signal-to-background discrimination achieved through a machine-learning-based multivariate analysis. Using the heavy photophobic axion-like particle (ALP) as a benchmark, we evaluate the discovery sensitivities on both production cross section times branching ratio $\sigma(pp \rightarrow W^\pm X) \times \text{Br}(X \rightarrow W^+ W^-)$ and the coupling g_{aWW} for particle mass over a wide range of 170–3000 GeV at the HL-LHC with center-of-mass energy $\sqrt{s} = 14$ TeV and integrated luminosity $\mathcal{L} = 3 \text{ ab}^{-1}$. Our results show significant improvements in discovery sensitivity, particularly for masses above 300 GeV, compared to existing limits derived from CMS analyses of Standard Model (SM) tri- W -boson production at $\sqrt{s} = 13$ TeV. This study demonstrates the potential of advanced selection techniques in probing the coupling of new particles to W -bosons and highlights the HL-LHC's capability to explore physics beyond the SM.

Keywords: Axion ; Axion-like particles, LHC phenomenology, Tri- W process at LHC

DOI: 10.1088/1674-1137/add115

CSTR: 32044.14.ChinesePhysicsC.49083106

I. INTRODUCTION

The search for new particles beyond the Standard Model (SM) is a cornerstone of high-energy physics, driven by unresolved questions such as the nature of dark matter, origin of neutrino mass, baryogenesis, and dark energy. Among the various extensions to the SM, new particles coupling to W -bosons represent a particularly intriguing avenue for exploration. In physics beyond the SM (BSM), there exist various types of neutral particles X that couple to W -bosons. Examples include (i) extended gauge models with additional gauge bosons (Z') [1], which couple to $W^+ W^-$ via mixing with SM gauge bosons γ or Z ; (ii) extended scalar sectors like the Two-Higgs-Doublet Model (2HDM) [2] and supersymmetric models [3], where the new CP-even scalar couples to $W^+ W^-$ via the mixing with the SM Higgs boson, which may also be related to some recent findings at the Large Hadron Collider (LHC) [4]; and (iii) pseudo-scalars, such as axions [5–9] or axion-like particles (ALPs) a , which

couple to $W^+ W^-$ via a dimension-five operator $a W_{\mu\nu}^+ \widetilde{W}^{-\mu\nu}$. Besides, new neutral particles coupling to W -bosons have also emerged in theoretical frameworks such as composite Higgs models [10], extra-dimensional scenarios [11], and dark matter models [12].

As the new neutral particle X couples to di- W -bosons, it can be produced and decay via the X - W - W vertex at the LHC. In this study, we consider X produced in association with a W -boson through the s -channel exchange of a $W^{(*)}$ boson, followed by its decay into a $W^+ W^-$ pair, resulting in a tri- W -boson final state. The mass of X is assumed to be greater than 170 GeV, ensuring that both its di- W -boson decay products are on-shell. We choose the heavy photophobic ALP [13] as a benchmark model in this work.

Axion models were originally proposed to address the strong CP problem in QCD [9, 14]. In these models, the axion arises as a pseudo-Nambu-Goldstone boson from a spontaneously broken $U(1)$ symmetry at a high energy scale, a mechanism known as the Peccei-Quinn mechanism.

Received 4 March 2025; Accepted 28 April 2025; Published online 29 April 2025

* Ying-nan Mao and Yiheng Xiong are supported by the National Natural Science Foundation of China (12205227). Kechen Wang is supported by the National Natural Science Foundation of China (11905162), the Excellent Young Talents Program of the Wuhan University of Technology (40122102), and the research program of the Wuhan University of Technology (2020IB024). The simulation and analysis work of this article was completed with the computational cluster provided by the Theoretical Physics Group at the Department of Physics, School of Physics and Mechanics, Wuhan University of Technology

[†] E-mail: E-mail: ynmao@whut.edu.cn

[‡] E-mail: kechen.wang@whut.edu.cn

[§] E-mail: yiheng@whut.edu.cn



Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Article funded by SCOAP³ and published under licence by Chinese Physical Society and the Institute of High Energy Physics of the Chinese Academy of Sciences and the Institute of Modern Physics of the Chinese Academy of Sciences and IOP Publishing Ltd

ism [5, 6]. In such original models [5–8, 15–21], coupling of the axions is strictly related to their mass m_a . Under the extended concept of ALPs, the mass and couplings of ALPs are treated as independent parameters. This flexibility significantly broadens the ALP parameter space, making them promising candidates for astrophysical and collider-based searches [22–24]. At colliders, ALPs are typically studied through their interactions with SM particles [25–35]. Among these interactions, previous experiments have primarily focused on ALPs' coupling to the diphoton $g_{a\gamma\gamma}$. The constraints on $g_{a\gamma\gamma}$ are stringent across most of the ALP mass range, with limits coming from astrophysical phenomena for $m_a \lesssim \mathcal{O}(\text{MeV})$ [36, 37], low-energy e^-e^+ collisions for $0.1 \text{ GeV} \lesssim m_a \lesssim 10 \text{ GeV}$ [38–40], Pb-Pb collisions at the LHC for $10 \text{ GeV} \lesssim m_a \lesssim 100 \text{ GeV}$ [41, 42], and p - p collisions at the LHC for $100 \text{ GeV} \lesssim m_a \lesssim 2 \text{ TeV}$ [29]. Consequently, the ALP-photon coupling is expected to be small, motivating the study of photophobic ALPs [13]. Photophobic ALPs are characterized by suppressed couplings to diphotons, with their primary interactions occurring with other SM electroweak bosons.

Previous studies [13, 43, 44] on heavy photophobic ALPs at the LHC have derived limits based on the tri- W -boson final state by reinterpreting experimental analyses at center-of-mass energy $\sqrt{s} = 8$ or 13 TeV, with limited luminosities and mass ranges. Ref. [13] derived constraints on ALP masses in the range of 40–500 GeV by reinterpreting triboson searches in the SM framework at $\sqrt{s} = 8$ TeV, where ATLAS analyses of $pp \rightarrow W^\pm(\rightarrow \mu^\pm \nu) W^\pm(\rightarrow \mu^\pm \nu) W^\pm(\rightarrow jj)$ with $\mathcal{L} = 20.3 \text{ fb}^{-1}$ [45] were utilized as an ALP-mediated signal process $pp \rightarrow W^\pm a(\rightarrow W^\pm W^\mp)$. Ref. [43] reinterpreted the CMS analyses of the SM process $pp \rightarrow jj W^\pm(\rightarrow \ell^\pm \nu) W^\pm(\rightarrow \ell^\pm \nu)$ at $\sqrt{s} = 13$ TeV [46] as a non-resonant ALP-mediated vector boson scattering process, with the ALP serving as an off-shell mediator; the results revealed sensitivities for a photophobic ALP with a mass below 100 GeV at the High-Luminosity LHC (HL-LHC) with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3 \text{ ab}^{-1}$. Ref. [44] reinterpreted the CMS analyses for the SM process $pp \rightarrow W^\pm W^\pm W^\pm$ at $\sqrt{s} = 13$ TeV with $\mathcal{L} = 35.9 \text{ fb}^{-1}$ [47] as an ALP-mediated signal process $pp \rightarrow W^\pm a(\rightarrow W^\pm W^\mp)$, setting limits on ALP mass up to 1 TeV.

More recently, some of our authors conducted detailed analyses of the γZ decay mode for heavy photophobic ALPs at the HL-LHC with $\sqrt{s} = 14$ TeV [48]. In that study, ALPs were produced with two jets through s -channel vector boson exchange and fusion, resulting in the process $pp \rightarrow jj a(\rightarrow \gamma Z(\rightarrow \ell^+ \ell^-))$, where $\ell = e, \mu$. Machine-learning-based multivariate analyses were used to optimize background rejection. Discovery sensitivities for the ALP coupling to di- W -bosons, g_{aWW} , were evaluated over masses from 100 to 4000 GeV with $\mathcal{L} = 3 \text{ ab}^{-1}$ and 140 fb^{-1} . Other studies on heavy photophobic ALPs at the LHC were also reviewed in Ref. [48]. Besides, heavy

photophobic ALPs have also been studied through a global fit analysis of electroweak precision observables [49].

We note that because the abovementioned experimental analyses were designed to target SM production processes rather than heavy photophobic ALPs, discrepancies in signal kinematics and background characteristics frequently result in conservative or potentially unreliable discovery sensitivities in existing reinterpretation studies. In this study, we conducted detailed analyses and derived discovery sensitivities at the HL-LHC for a new heavy neutral particle X that couples to W -bosons, using the process $pp \rightarrow W^\pm X(\rightarrow W^\pm W^\mp)$ with a tri- W -boson final state. Compared with previous reinterpreting studies based on tri- W -boson final state, key advancements include (i) performing simulations for the signal process of $pp \rightarrow W^\pm X(\rightarrow W^\pm W^\mp)$ and related background processes to fully capture signal and background kinematics at the detector-level; (ii) using machine-learning-based multivariate analysis (MVA) for optimal signal-to-background discrimination; (iii) presenting results over a wide mass range of 170–3000 GeV at $\sqrt{s} = 14$ TeV with $\mathcal{L} = 3 \text{ ab}^{-1}$; and (iv) providing the model-independent discovery sensitivities for the production cross-section time branching ratio, $\sigma(pp \rightarrow W^\pm X) \times \text{Br}(X \rightarrow W^\pm W^\mp)$, which can be applied to general models involving a new heavy neutral particle X that couples to W -bosons.

The remainder of this paper is organized as follows. In Sec. II, we introduce the signal process under study. Sec. III outlines the main SM backgrounds relevant to our signal. In Secs. IV and V, we describe the simulation setup and search strategy, respectively. Our results are presented in Sec. VI, and we conclude in Sec. VII. Additional details supporting the main text are provided in the appendices.

II. SIGNAL PRODUCTION

As shown in Fig. 1, we consider the production of X in association with a W -boson at the HL-LHC with cen-

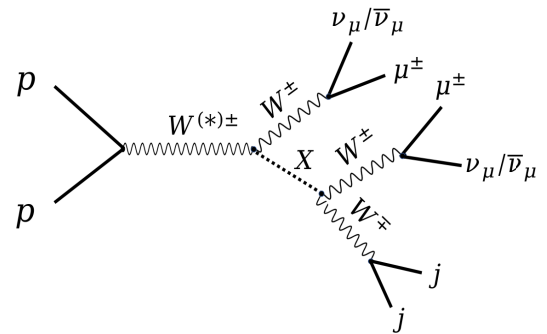


Fig. 1. Production and decay of new particle X coupling to di- W boson at pp colliders, leading to the tri- W final state. The same-charged W^\pm bosons are considered to decay into μ^\pm and ν_μ (or $\bar{\nu}_\mu$), while the remaining W^\mp boson decays into di-jet.

ter-of-mass energy $\sqrt{s} = 14$ TeV. X then decays to the di- W boson. The corresponding signal process is $pp \rightarrow W^\pm X (\rightarrow W^+ W^-)$, leading to the tri- W -boson final state. To eliminate the background effectively, two same-charged W -bosons are required to decay leptonically into muons, leading to the existence of same-sign muons in the final state. The remanent W -boson decays hadronically to di-jet, which has larger branching ratio and can increase the signal production cross-section.

For a concrete study, we take the ALP theory model as an example. The neutral particle X is assumed to be a heavy photophobic ALP that couples to electroweak gauge bosons only. The effective linear ALP Lagrangian is [50]

$$\mathcal{L}_{\text{eff}} \supset \frac{1}{2} (\partial_\mu a) (\partial^\mu a) - \frac{1}{2} m_a^2 a^2 - c_{\tilde{W}} \frac{a}{f_a} W_{\mu\nu}^b \tilde{W}^{b,\mu\nu} - c_{\tilde{B}} \frac{a}{f_a} B_{\mu\nu} \tilde{B}^{\mu\nu}, \quad (1)$$

where $W_{\mu\nu}^b$ and $B_{\mu\nu}$ represent the field strength tensors of the $SU(2)_L$ and $U(1)_Y$ gauge groups, respectively, and $\tilde{O}^{\mu\nu} \equiv \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} O_{\alpha\beta}$ ($O = W, B$) is the dual-field strength tensor. m_a and f_a denote the mass of ALP and its decay constant¹⁾, respectively, and they are assumed to be independent parameters in this work.

After electroweak symmetry breaking, the interactions between ALP and gauge bosons are generally expressed as

$$\mathcal{L}_{\text{eff}} \supset -\frac{1}{4} g_{a\gamma\gamma} a F_{\mu\nu} \tilde{F}^{\mu\nu} - \frac{1}{2} g_{a\gamma Z} a Z_{\mu\nu} \tilde{F}^{\mu\nu} - \frac{1}{4} g_{aZZ} a Z_{\mu\nu} \tilde{Z}^{\mu\nu} - \frac{1}{2} g_{aWW} a W_{\mu\nu}^+ \tilde{W}^{-,\mu\nu}, \quad (2)$$

where $F_{\mu\nu}$, $Z_{\mu\nu}$, and $W_{\mu\nu}^+$ are the field strength tensors of the electromagnetic, Z -, and W -fields. In addition, the coupling constants are expressed as

$$\begin{aligned} g_{a\gamma\gamma} &= \frac{4}{f_a} (s_\theta^2 c_{\tilde{W}} + c_\theta^2 c_{\tilde{B}}), \\ g_{aZZ} &= \frac{4}{f_a} (c_\theta^2 c_{\tilde{W}} + s_\theta^2 c_{\tilde{B}}), \\ g_{a\gamma Z} &= \frac{2s_{2\theta}}{f_a} (c_{\tilde{W}} - c_{\tilde{B}}), \\ g_{aWW} &= \frac{4}{f_a} c_{\tilde{W}}, \end{aligned} \quad (3)$$

where $c_{\tilde{W}}$ and $c_{\tilde{B}}$ are the same constants as in Eq. (1), and $c_\theta = \cos \theta_W$, $s_\theta = \sin \theta_W$, and $s_{2\theta} = \sin 2\theta_W$ are all trigono-

metric functions of the Weinberg angle θ_W . In this work, focusing on the photophobic ALP scenario, we assume the coupling between ALP and di-photon $g_{a\gamma\gamma} = \frac{4}{f_a} (s_\theta^2 c_{\tilde{W}} + c_\theta^2 c_{\tilde{B}}) = 0$. As a consequence, $s_\theta^2 c_{\tilde{W}} + c_\theta^2 c_{\tilde{B}} = 0$, making a constant proportional relationship between $c_{\tilde{W}}$ and $c_{\tilde{B}}$. Thus, $g_{a\gamma Z}$ and g_{aZZ} can be expressed as functions of g_{aWW} :

$$g_{a\gamma Z} = t_\theta g_{aWW}, \quad \text{and} \quad g_{aZZ} = (1 - t_\theta^2) g_{aWW}, \quad (4)$$

where $t_\theta = \tan \theta_W$.

III. SM BACKGROUND

Because the signal final state contains two same-sign muons plus di-jet and moderate missing energy, relevant SM processes that can mimic the signal are listed as follows:

- (i) one-boson production with two jets: $W^\pm jj$ and Zjj ;
- (ii) two-boson production with two jets: $W^\pm W^\pm jj$, $W^+ W^- jj$, $W^\pm Zjj$, and $ZZjj$;
- (iii) three-boson production: $W^\pm W^\pm W^\mp$;
- (iv) top quark pair production: $t\bar{t}$.

After W -boson decays, the processes $W^\pm W^\pm jj$ and $W^\pm W^\pm W^\mp$ can result in the same final state as the signal and serve as irreducible backgrounds, despite their relatively low production cross-sections. The process $W^\pm jj$, which has a significantly higher production cross-section, also contributes to the background due to the potential misidentification of muons. Additionally, charge mis-measurement of final-state muons can lead to backgrounds from processes like Zjj , $W^+ W^- jj$, and $t\bar{t}$. Furthermore, the $W^\pm Zjj$ process becomes relevant when one of the muons in the final state remains undetected, and the $ZZjj$ process can contribute both through missed muons and mis-measured muon charges.

IV. EVENT SIMULATION

We first use the MadGraph5_aMC@NLO program (version 2.6.7) [51] to simulate the proton-proton collision events at the parton level, where the NNPDF23 parton distribution function (PDF) [52] of the proton is utilized. At the parton level, to produce data as close to the experimental results as possible, the following relatively loose thresholds are applied for both the signal and back-

¹⁾ In ALP models, f_a always affects on physical observables together with the coefficients c_i [such as the coefficients $c_{\tilde{W},\tilde{B}}$ in Eq. (1)], thus people can only set constraints on the combinations c_i/f_a . As a new physics scale, f_a is expected to be much higher than the electro-weak scale, thus we expect $f_a \gtrsim O(\text{TeV})$.

ground processes: (a) the minimal transverse momentum of jets, photons, and charged leptons is set to 0.5 GeV, *i.e.*, $p_T(j/\gamma/l^\pm) > 0.5$ GeV; (b) the maximal pseudorapidity of jets is set to 10, *i.e.*, $|\eta(j)| < 10$, and it is set to 5 for the photons and charged leptons, *i.e.*, $|\eta(\gamma/l^\pm)| < 5$; and (c) the solid angular difference $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between objects is set to 0.1. These threshold criteria are set in the "run_card.dat" file of the MadGraph program.

All signal and background events are then passed through the PYTHIA program (version 8.2) [53] for parton showering, hadronization, and decay of unstable particles. We use the Delphes program (version 3.4.2) [54] to perform the detector simulation, where the ATLAS detector configuration is adopted and the minimal transverse momenta of leptons and photons accepted by the detector is set to 2 GeV. Due to limited computing resources, when simulating background processes of $W^\pm jj$, Zjj , and $W^\pm W^\mp jj$, the decay mode of one gauge boson is fixed to decay to a muon, *i.e.*, $W^\pm(\rightarrow \mu^\pm \nu_\mu)jj$, $Z(\rightarrow \mu^+ \mu^-)jj$, and $W^\pm(\rightarrow \mu^\pm \nu_\mu)W^\mp jj$, respectively, so that these processes can still have sufficient event numbers after all selection criteria.

For the signal, we apply the ALP model file with the linear Lagrangian [50] in the Universal FeynRules Output (UFO) format [55] in the MadGraph5 program and generate the $pp \rightarrow W^\pm a(\rightarrow W^+ W^-)$ events. A scan of the ALP mass, m_a , is performed in the following manner: individual mass 170; 25 GeV increments in the mass range 200–300; 100 GeV increments in the mass range 300–600 GeV; individual masses 750, 800, and 900; and 400 GeV increments in the mass range 1000–3000 GeV. For each ALP mass, the coupling g_{aWW} is set to 1 TeV^{-1} . For each iteration of ALP mass m_a , at least 10^6 events are generated to ensure that statistical uncertainties are minimized as much as possible within the bounds of computational resources.

To ensure consistency in our analysis, we use the production cross-sections calculated by the MadGraph5 program to estimate the event yields for both signal and background processes. Figure 2 illustrates the signal production cross-section $\sigma(pp \rightarrow W^\pm a)$, multiplied by the branching ratio $\text{Br}(a \rightarrow W^+ W^-)$, as a function of the ALP mass m_a in the range of 170 GeV to 3000 GeV at the HL-LHC with $\sqrt{s} = 14 \text{ TeV}$, where the coupling g_{aWW} is fixed to 1 TeV^{-1} . As m_a increases beyond 160 GeV, the phase space for the decay mode $a \rightarrow W^+ W^-$ expands rapidly, leading to a sharp rise in the branching ratio, which reaches approximately 50% at $\sim 190 \text{ GeV}$ [48]. Beyond this point, the branching ratio increases more gradually, with only minor changes at higher masses. At the same time, the production cross-section for process $pp \rightarrow W^\pm a$ decreases steadily with increasing m_a . Consequently, the production cross-section times branching ratio peaks when $m_a \sim 190 \text{ GeV}$, just above the WW decay threshold of 160 GeV.

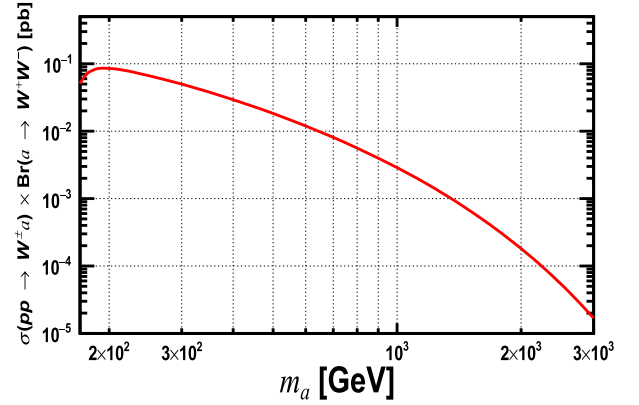


Fig. 2. (color online) Signal production cross-section $\sigma(pp \rightarrow W^\pm a)$ multiplied by branching ratio $\text{Br}(a \rightarrow W^+ W^-)$ as a function of ALP mass m_a from 170 to 3000 GeV at the HL-LHC with $\sqrt{s} = 14 \text{ TeV}$, assuming the coupling $g_{aWW} = 1 \text{ TeV}^{-1}$.

V. SEARCH STRATEGY

The following preselection criteria were employed before a multivariate analysis was performed, where final-state muons and jets are ordered based on their transverse momenta and labeled as O_i ($i = 1, 2, \dots$), with $O = \mu, j$, respectively:

- (1) Events are required to have exactly two muons, *i.e.*, $N(\mu) = 2$.
- (2) The minimal p_T of the two muons is 10 GeV, *i.e.*, $p_T(\mu) > 10 \text{ GeV}$.
- (3) The two muons have the same charge, *i.e.*, $\mu^\pm \mu^\pm$.
- (4) The minimal p_T of all jets is 30 GeV, *i.e.*, $p_T(j) > 30 \text{ GeV}$.
- (5) Event are required to have at least two jets, *i.e.*, $N(j) \geq 2$, and the number of b -tagged jets must be zero, *i.e.*, $N(j_b) = 0$.

Table 1 presents the expected numbers of events, N_{exp} , for the signal, with a benchmark mass of $m_a = 500 \text{ GeV}$ and coupling $g_{aWW} = 1 \text{ TeV}^{-1}$, as well as for the background processes. Here, N_{exp} is calculated by the formula

$$N_{\text{exp}} = \sigma_{\text{pro}} \times \mathcal{L} \times \epsilon_{\text{pre}}, \quad (5)$$

where σ_{pro} is the production cross-section of the signal or background process; \mathcal{L} is the integrated luminosity; and ϵ_{pre} is the preselection efficiency, which is evaluated based on our analyses. The numbers are shown after applying preselection criteria (1)–(5) sequentially at the

Table 1. Expected numbers of events for the signal with benchmark mass $m_a = 500$ GeV and coupling $g_{aWW} = 1 \text{ TeV}^{-1}$ and background processes after applying preselection criteria (1)–(5) sequentially at the HL-LHC with $\sqrt{s} = 14 \text{ TeV}$ and $\mathcal{L} = 3 \text{ ab}^{-1}$.

	initial	(1)-(2)	(3)	(4)-(5)
$W^\pm a(\rightarrow W^\pm W^\mp)$	5.51×10^4	1.27×10^3	4.26×10^2	2.54×10^2
$W^\pm(\rightarrow \mu^\pm \nu_\mu)jj$	3.62×10^{11}	7.24×10^6	2.69×10^6	4.89×10^5
$Z(\rightarrow \mu^\pm \mu^\mp)jj$	3.31×10^{10}	1.13×10^{10}	3.33×10^5	6.67×10^4
$W^\pm W^\pm jj$	1.10×10^6	5.22×10^3	5.19×10^3	3.88×10^3
$W^\pm(\rightarrow \mu^\pm \nu_\mu)W^\mp jj$	5.05×10^8	2.43×10^7	8.81×10^3	4.53×10^3
$W^\pm Zjj$	1.03×10^9	1.42×10^7	6.17×10^5	1.99×10^5
$ZZjj$	3.10×10^8	8.42×10^6	4.32×10^4	1.55×10^4
$W^\pm W^\pm W^\mp$	3.79×10^5	5.92×10^3	1.73×10^3	8.55×10^2
$t\bar{t}$	1.79×10^9	1.22×10^7	8.71×10^4	2.91×10^4

HL-LHC, with $\sqrt{s} = 14 \text{ TeV}$ and $\mathcal{L} = 3 \text{ ab}^{-1}$. As shown in Table 1, after applying all preselection criteria, the number of background events for most processes is reduced by at least 10^4 -fold. Nevertheless, the total background rate remains much larger than the expected number of signal events.

After applying the preselection criteria, the Toolkit for Multivariate Analysis (TMVA) package [56] is employed to perform a multivariate analysis (MVA), which enhances the ability to distinguish between signal and background events. The events that pass the preselection steps are then subjected to analysis using the Boosted Decision Trees (BDT) algorithm within the TMVA framework. In this process, the discrimination between signal and background is achieved by utilizing the following set of kinematic observables:

(a) Missing transverse energy and its azimuth angle: $\cancel{E}_T, \phi(\cancel{E}_T)$.

(b) The x , y , and z -components of momentum (p_x, p_y, p_z) and energy (E) of the first two leading jets and muons: $p_x(j_1), p_y(j_1), p_z(j_1), E(j_1); p_x(j_2), p_y(j_2), p_z(j_2), E(j_2); p_x(\mu_1), p_y(\mu_1), p_z(\mu_1), E(\mu_1); p_x(\mu_2), p_y(\mu_2), p_z(\mu_2), E(\mu_2)$.

(c) The number of charged tracks (N_{track}) and ratio (R_E) of the hadronic versus electromagnetic energy deposited in the calorimeter cells for the first two leading jets: $N_{\text{track}}(j_1), N_{\text{track}}(j_2), R_E(j_1)$, and $R_E(j_2)$. R_E is typically greater than one for a jet.

(d) The solid angular separation ΔR between j_1 and j_2 , = invariant mass (m) of system of $(j_1 + j_2)$, and number (N) of jets: $\Delta R(j_1, j_2), m(j_1 + j_2)$, and $N(j)$.

(e) Observables related to the isolation quality of muons. (i) The summed p_T of other objects excluding the muon in an $R = 0.4$ cone around the muon: $p_T^{\text{iso}}(\mu_1)$,

$p_T^{\text{iso}}(\mu_2)$, and the larger of the two $p_{T,\text{max}}^{\text{iso}}(\mu)$. (ii) The ratio of the transverse energy in a 3×3 grid surrounding the muon to the p_T of the muon ("etrat" in Ref. [57]), which is a percentage between 0 and 0.99: $R_{\text{grid}}(\mu_1), R_{\text{grid}}(\mu_2)$. (iii) We make all possible combinations between μ_1/μ_2 and every jet, compare the ΔR , and find the minimal value: $\Delta R_{\text{min}}(\mu, j)$. For well-isolated muons, p_T^{iso} and R_{grid} should be small, while $\Delta R_{\text{min}}(\mu, j)$ is large.

(f) To reconstruct the hadronically decaying W -boson (jj_w), we make all di-jet combinations and select the pair with the invariant mass closest to 80 GeV. The corresponding two jets are labeled as j_{w1} and j_{w2} sorted by their p_T . We input the following observables of these two jets: $p_x(j_{w1}), p_y(j_{w1}), p_z(j_{w1}), E(j_{w1}); p_x(j_{w2}), p_y(j_{w2}), p_z(j_{w2}), E(j_{w2}); N_{\text{track}}(j_{w1}), N_{\text{track}}(j_{w2}); R_E(j_{w1}), R_E(j_{w2}); \Delta R(j_{w1}, j_{w2}), m(j_{w1} + j_{w2})$.

(g) The transverse mass m_T [58] of the system, which includes one muon and the missing transverse energy: $m_T(\mu_1 + \cancel{E}_T), m_T(\mu_2 + \cancel{E}_T)$. Here, the transverse mass $m_T \equiv \sqrt{(E_T^{\text{vis.}} + \cancel{E}_T)^2 - (\vec{p}_T^{\text{vis.}} + \vec{p}_T)^2}$, where $E_T^{\text{vis.}}$ ($\vec{p}_T^{\text{vis.}}$) is the transverse energy (momentum) of the visible object or system, while \cancel{E}_T (\vec{p}_T) is the missing transverse energy (momentum). $E_T^{\text{vis.}} = \sqrt{(\vec{p}_T^{\text{vis.}})^2 + (m^{\text{vis.}})^2}$, where $m^{\text{vis.}}$ is the invariant mass of the visible object or system. $\cancel{E}_T = |\vec{p}_T|$, assuming the invariant mass of the invisible object is zero.

(h) We calculate and compare ΔR values of two combinations between jj_w and μ_1/μ_2 . The muon with smaller (larger) ΔR value is labeled as μ_a (μ_b). We input the minimal value $\Delta R(j_{w1} + j_{w2}, \mu_a)$ and maximal value $\Delta R(j_{w1} + j_{w2}, \mu_b)$.

(i) The observables related to the reconstruction of the ALP mass are $m(j_{w1} + j_{w2} + \mu_a), m(j_{w1} + j_{w2} + \mu_b), m_T(j_{w1} + j_{w2} + \mu_a + \cancel{E}_T)$, and $m_T(j_{w1} + j_{w2} + \mu_b + \cancel{E}_T)$.

(j) The observables related to the reconstruction of the off-shell $W^{(*)\pm}$ boson: $m_T(j_{w1} + j_{w2} + \mu_a + \mu_b + \cancel{E}_T)$.

Details of the observables $N_{\text{track}}(j_i)$, $R_E(j_i)$, $R_{\text{grid}}(\mu_i)$, and $p_T^{\text{iso}}(\mu)$ can be found in [57]. The hadronic W -boson (jj_w) is reconstructed using the method described above. Appendix A presents distributions of kinematical observables after applying preselection criteria for signal and background processes at the HL-LHC with $\sqrt{s} = 14$ TeV. Among them, for m_a values below 500 GeV, the endpoint positions of the distributions for both $m(j_{w1} + j_{w2} + \mu_a)$ and $m_T(j_{w1} + j_{w2} + \mu_a + \cancel{E}_T)$ align closely with m_a ; when m_a exceeds 500 GeV, the endpoint positions of the $m_T(j_{w1} + j_{w2} + \mu_a + \cancel{E}_T)$ distribution are still related to m_a , but this phenomenon becomes less obvious. Appendix B provides detailed information about the MVA analyses performed using the TMVA package [56].

Figure 3 presents the BDT response distributions for the total background and signal at the HL-LHC, with benchmark masses of $m_a = 400$ GeV and $m_a = 900$ GeV. The left and right plots show similar trends in the distributions for both mass points. The clear separation between the signal and background distributions indicates that the BDT criteria are effective for background rejection at the HL-LHC. Additionally, the separation is more pronounced for $m_a = 900$ GeV compared to $m_a = 400$ GeV, suggesting improved discriminating power at higher masses. Appendix C presents distributions of BDT responses after applying preselection criteria for the signal and background processes at the HL-LHC with $\sqrt{s} = 14$ TeV, assuming various m_a cases.

After the preselection, the BDT cut is optimized according to the signal statistical significance calculated by Eq. (6) for each mass case:

$$\sigma_{\text{stat}} = \sqrt{2 \left[(N_s + N_b) \ln \left(1 + \frac{N_s}{N_b} \right) - N_s \right]}, \quad (6)$$

where N_s and N_b are the expected numbers of events for signal and total background, respectively, after applying

both the preselection and BDT criteria. Appendix D shows selection efficiencies of preselection and BDT criteria for signal and background processes at the HL-LHC with $\sqrt{s} = 14$ TeV, assuming different ALP masses, where "-" means the number of events can be reduced to negligible with $\mathcal{L} = 3 \text{ ab}^{-1}$.

VI. RESULTS

Using our search strategy, Fig. 4 presents the discovery sensitivities for the production cross-section $\sigma(pp \rightarrow W^\pm X)$ multiplied by the branching ratio $\text{Br}(X \rightarrow W^+ W^-)$ as a function of m_X in the mass range of 170–3000 GeV at the HL-LHC, with $\sqrt{s} = 14$ TeV and an integrated luminosity of $\mathcal{L} = 3 \text{ ab}^{-1}$. The red and green curves correspond to the 2- σ and 5- σ significances, respectively. As shown in the BDT distributions in Appendix C, the separation between the signal and background improves significantly for larger masses, leading to more effective background rejection. Consequently, the sensitivities for heavier masses do not decrease very rapidly, demonstrating strong discovery potential across a wide mass range.

In Fig. 5, for the concrete case of heavy photophobic ALPs, we show the discovery sensitivities on the coupling g_{aWW} with 2- σ and 5- σ significances as a function of the ALP mass m_a in the range of 170–3000 GeV at the HL-LHC, with $\sqrt{s} = 14$ TeV and integrated luminosities of 3 ab^{-1} and 35.9 fb^{-1} . For comparison, we also display the 95% C.L. limit (blue curve) from Ref. [44], which is derived from reinterpreting the CMS analyses of the SM production $pp \rightarrow W^\pm W^\pm W^\mp$ at $\sqrt{s} = 13$ TeV with $\mathcal{L} = 35.9 \text{ fb}^{-1}$. Our results demonstrate that at the HL-LHC with $\mathcal{L} = 3 \text{ ab}^{-1}$, the 2- σ sensitivity for g_{aWW} decreases from 1.06 TeV^{-1} to 0.75 TeV^{-1} as m_a increases from 170 GeV to 225 GeV, remains relatively stable for m_a in the range of 225–600 GeV, and then rises sharply to 9.42 TeV^{-1} at $m_a = 3000$ GeV. This pattern reflects the varying sensitivity of the search strategy across different mass ranges.

One observes that with the same luminosity of 35.9 fb^{-1} , our strategy achieves better sensitivities for m_a

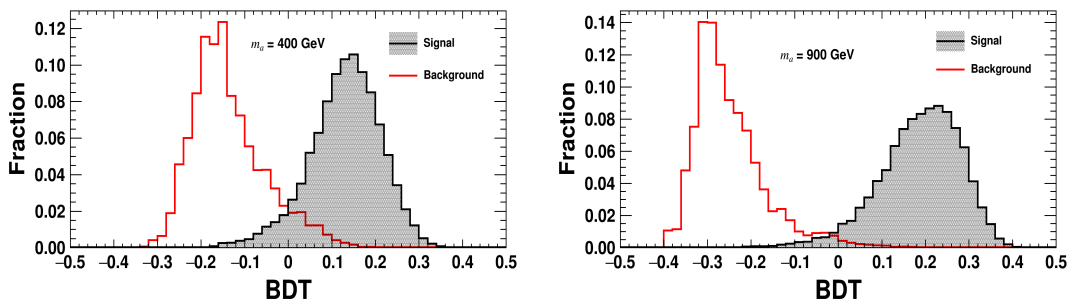


Fig. 3. (color online) BDT response distributions for total SM background and signal when ALP mass $m_a = 400$ GeV (left) and 900 GeV (right) at the HL-LHC with $\sqrt{s} = 14$ TeV.

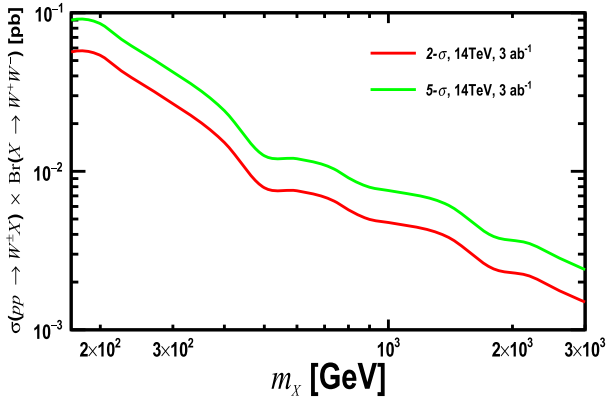


Fig. 4. (color online) Discovery sensitivities on the production cross-section $\sigma(pp \rightarrow W^\pm X) \times \text{Br}(X \rightarrow W^+ W^-)$ in the mass range of 170–3000 GeV at the HL-LHC, with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3 \text{ ab}^{-1}$. Red and green curves correspond to 2- σ and 5- σ significances, respectively.

above 300 GeV compared to the referenced approach [44]. This improvement is primarily due to differences in the optimization of the analysis strategies. The referenced studies focused on the SM process $pp \rightarrow W^\pm W^\pm W^\mp$, while our strategy is specifically tailored to the tri- W final state produced by a heavy resonance. For larger masses, the kinematics of these two processes differ significantly, allowing our approach to reject background more effectively, as demonstrated by the BDT distributions in Appendix C.

VII. CONCLUSION

In this paper, we proposed a search strategy at the HL-LHC for a new neutral particle X that couples to a pair of W -bosons. The particle X is produced with a W -boson and decays into two W -bosons, resulting in the process $pp \rightarrow W^\pm X(\rightarrow W^+ W^-)$ and a tri- W -boson final state. To suppress background, we focused on events where two same-charge W -bosons decay leptonically into muons, producing same-sign muons in the final state. The third W -boson decays hadronically into jets, leveraging its higher branching ratio to enhance the signal. As a case study, we used the heavy photophobic ALP as an example.

Signal and background events were simulated at the detector level. The signal production cross-section times branching ratio, $\sigma(pp \rightarrow W^\pm a) \times \text{Br}(a \rightarrow W^+ W^-)$, was evaluated as a function of the ALP mass m_a at the HL-LHC with $\sqrt{s} = 14$ TeV. Events were selected with ex-

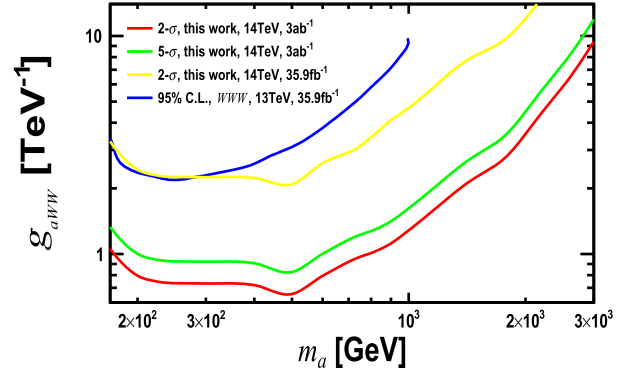


Fig. 5. (color online) Discovery sensitivities with 2- σ and 5- σ significances on the coupling g_{aWW} in the mass range of 170–3000 GeV at the HL-LHC, with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3 \text{ ab}^{-1}$ and 35.9 fb^{-1} . The 95% C.L. limit (blue curve) [44] derived from reinterpreting CMS analyses of SM production $pp \rightarrow W^\pm W^\pm W^\mp$ at $\sqrt{s} = 13$ TeV and $\mathcal{L} = 35.9 \text{ fb}^{-1}$ is displayed for comparison.

actly two same-sign di-muons and at least two non- b -tagged jets. A machine-learning-based MVA was applied to enhance signal-background discrimination. Distributions of key input variables and the corresponding BDT responses were presented, and selection efficiencies for both preselection and BDT criteria were provided for various ALP masses.

We presented the model-independent discovery sensitivities for the production cross-section times branching ratio $\sigma(pp \rightarrow W^\pm X) \times \text{Br}(X \rightarrow W^+ W^-)$ as a function of the particle mass m_X in the range of 170–3000 GeV at the HL-LHC, with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3 \text{ ab}^{-1}$, which can be applied to general models involving a new neutral particle X that couples to W -bosons. The sensitivities for the coupling g_{aWW} of heavy photophobic ALPs were also shown for 2- σ and 5- σ significances in the same mass range under luminosities of 3 ab^{-1} and 35.9 fb^{-1} . At $\mathcal{L} = 3 \text{ ab}^{-1}$, the 2- σ sensitivity for g_{aWW} decreased from 1.06 TeV^{-1} to 0.75 TeV^{-1} as m_a increased from 170 GeV to 225 GeV, remained almost constant for m_a between 225 and 600 GeV, and then sharply rose to 9.42 TeV^{-1} at $m_a = 3000$ GeV. Compared to previous limits derived from reinterpreting CMS analyses of SM production $pp \rightarrow W^\pm W^\pm W^\mp$ at $\sqrt{s} = 13$ TeV, our strategy provides improved sensitivity for m_a above 300 GeV.

APPENDIX A: DISTRIBUTIONS OF REPRESENTATIVE OBSERVABLES

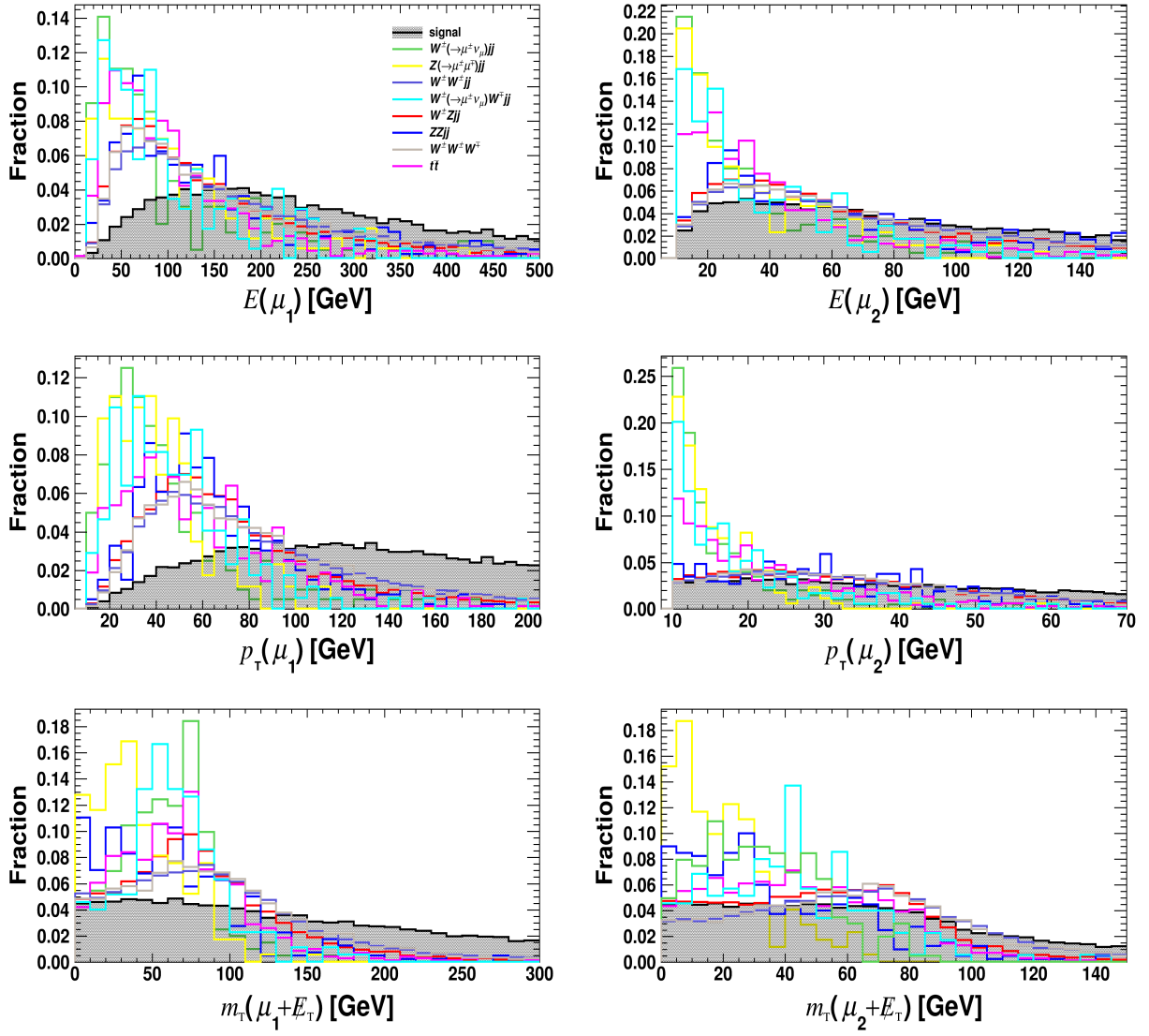
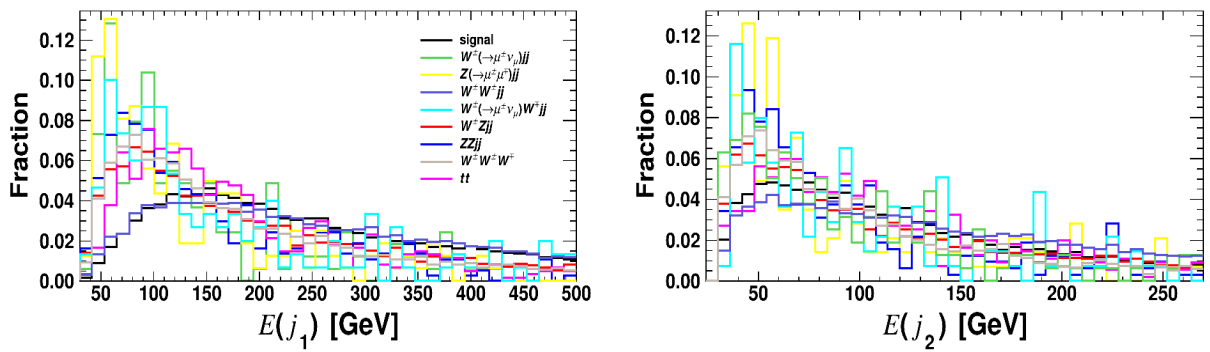


Fig. A1. (color online) Distributions of kinematical observables for μ_1 and μ_2 after applying preselection criteria for the signal (black, shadow) and background processes at the HL-LHC with $\sqrt{s} = 14$ TeV, assuming the benchmark $m_a = 400$ GeV.



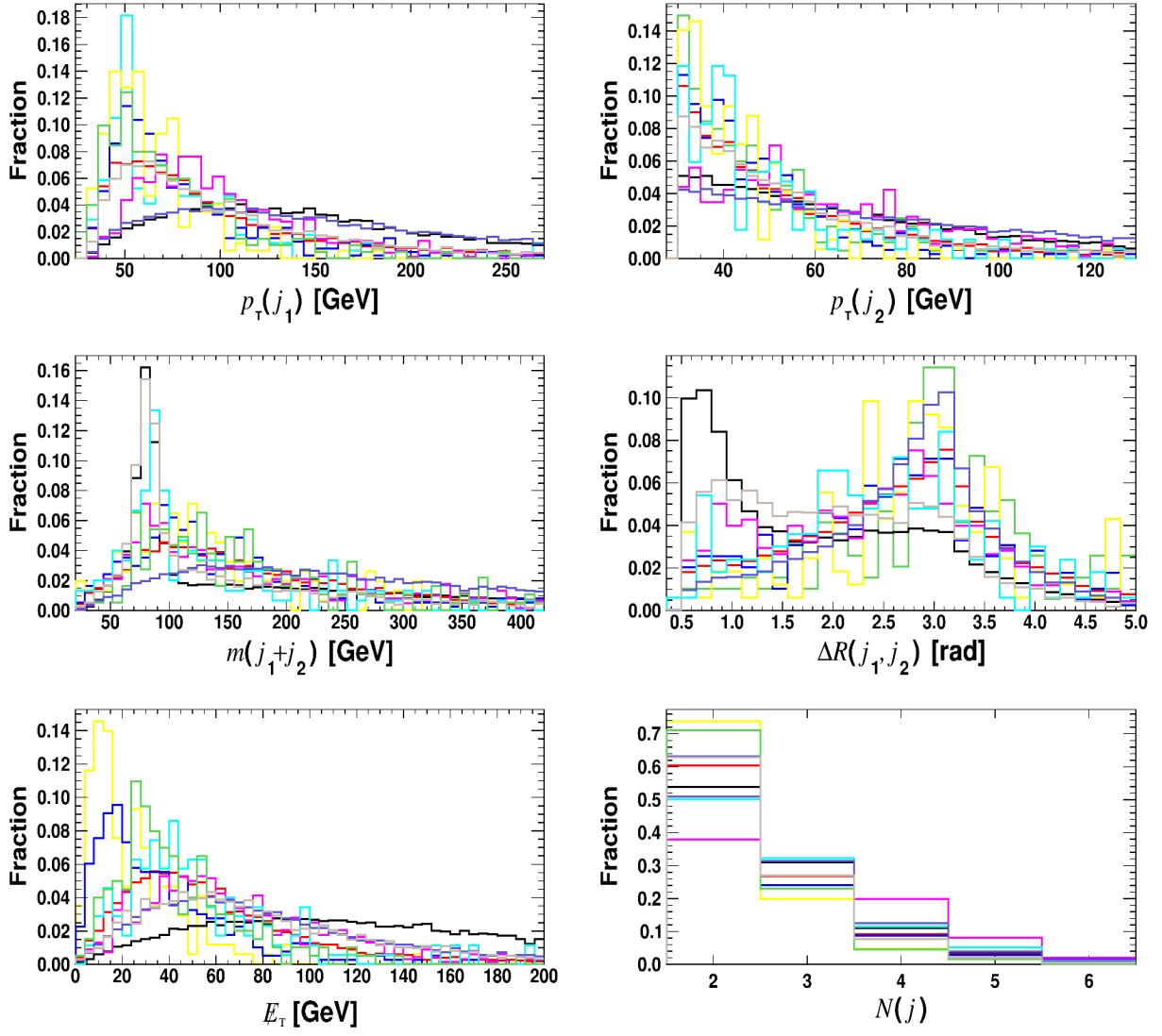
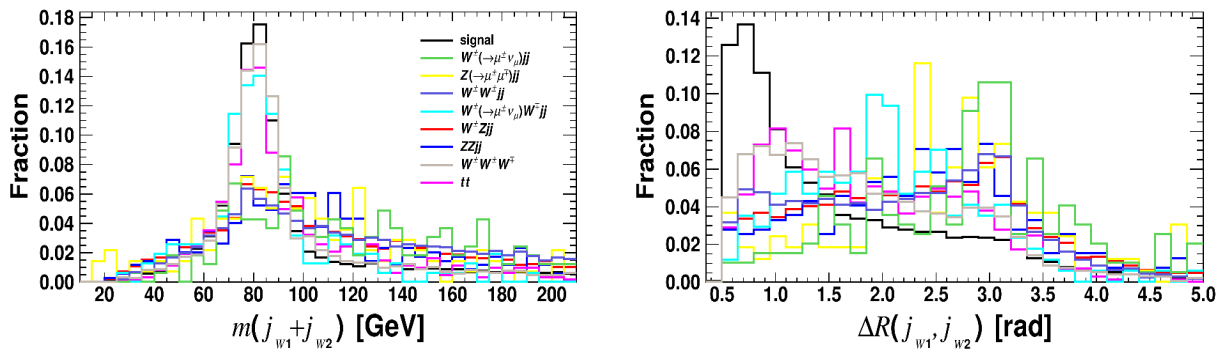


Fig. A2. (color online) Same as Fig. A1 but for j_1 , j_2 , E_T , and $N(j)$, assuming the benchmark $m_a = 400$ GeV.



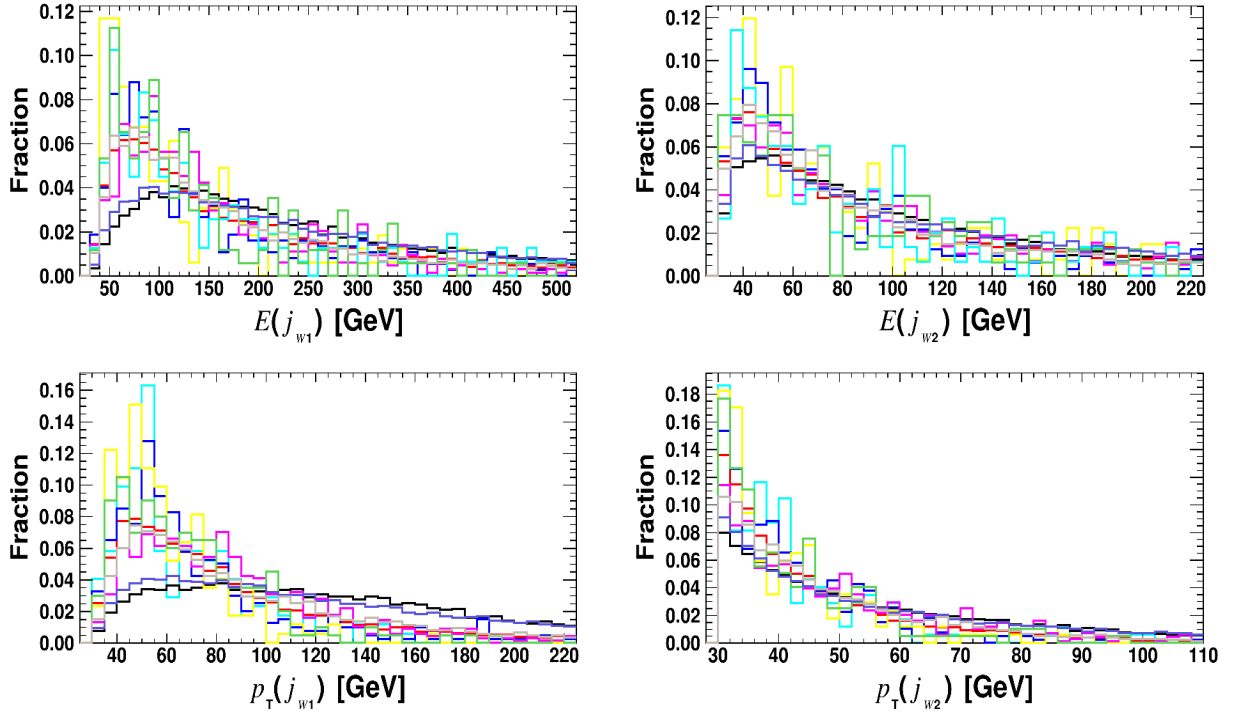
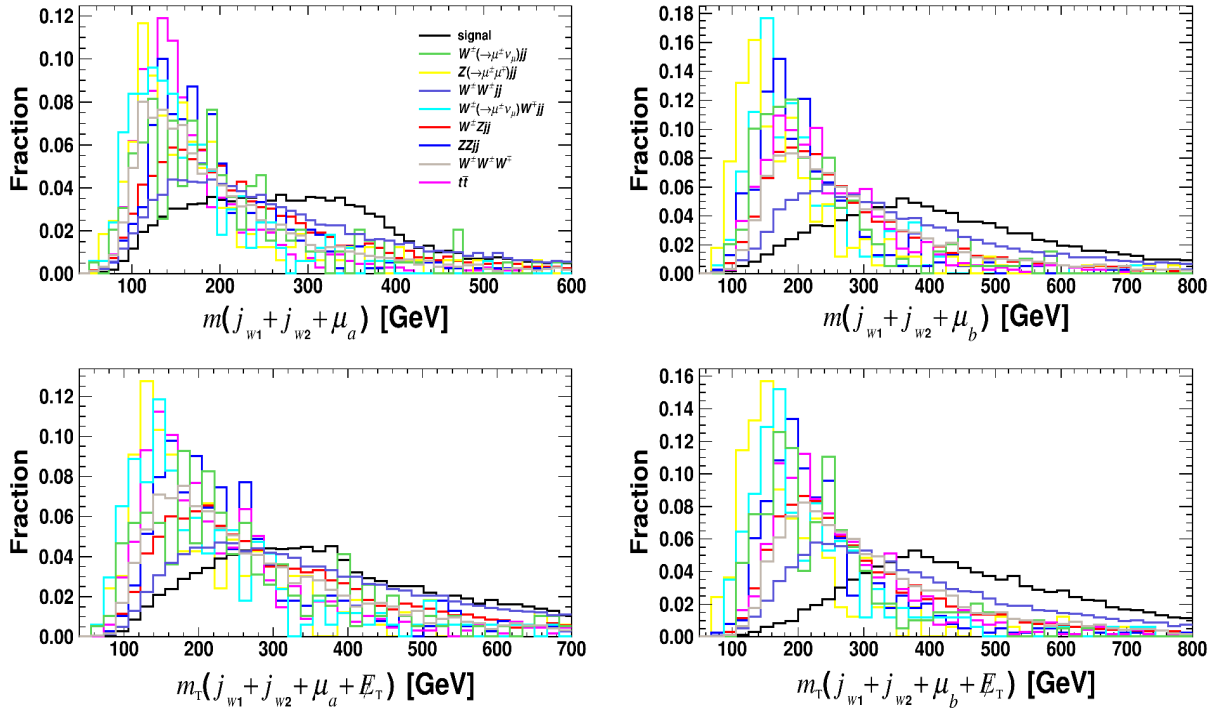


Fig. A3. (color online) Same as Fig. A1 but for j_{w1} and j_{w2} , assuming the benchmark $m_a = 400$ GeV.



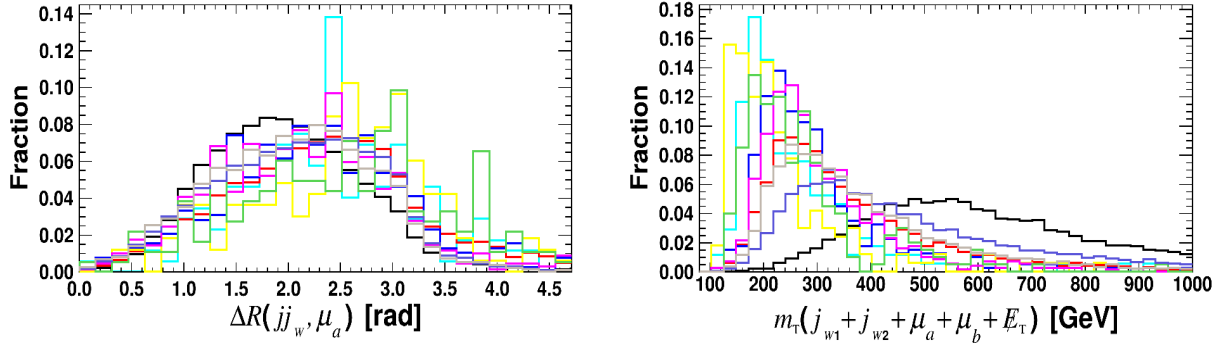


Fig. A4. (color online) Same as Fig. A1 but for observables related to the reconstruction of the ALP mass and off-shell $W^{(*)\pm}$ boson, assuming the benchmark $m_a = 400$ GeV.

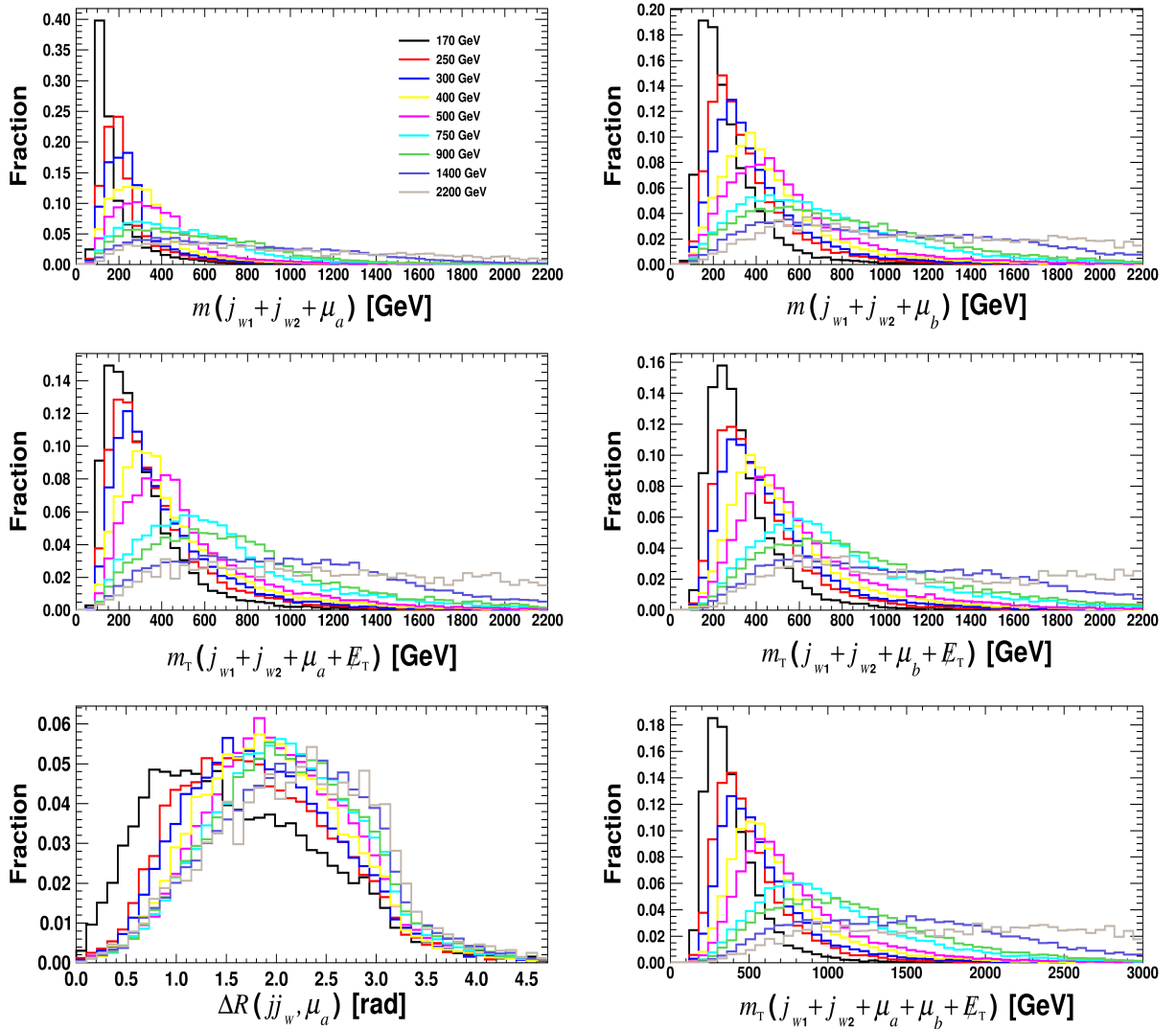


Fig. A5. (color online) Distributions of same observables as Fig. A4 but for the signal only assuming different m_a values.

APPENDIX B: DETAILS OF THE MVA ANALYSES

In this appendix, we provide detailed information about the MVA analyses performed using the TMVA package [56]. Figure B1 illustrates the importance values of input observables as a function of m_a , evaluated during the MVA training. The importance values for BDT input variables are quantified by counting how often the variables are used to split decision tree nodes and by

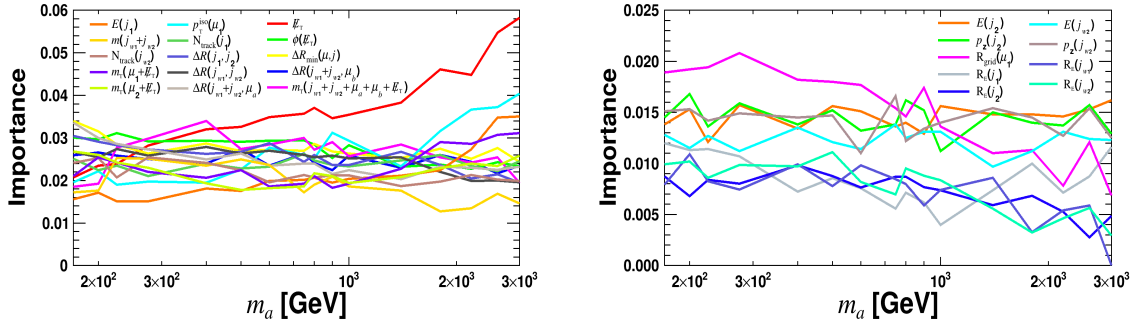


Fig. B1. (color online) Importance values of input observables as a function of m_a , evaluated during the MVA training.

Table B1. Correlations between two observables for the total background and signal with benchmark mass $m_a = 400$ GeV.

observables	signal	background
$p_T^{\text{iso}}(\mu_1), p_{T,\text{max}}^{\text{iso}}(\mu)$	1.000	0.998
$m_T(j_{w1} + j_{w2} + \mu_b + \cancel{E}_T), m_T(j_{w1} + j_{w2} + \mu_a + \mu_b + \cancel{E}_T)$	0.919	0.974
$m(j_{w1} + j_{w2} + \mu_b), m_T(j_{w1} + j_{w2} + \mu_a + \mu_b + \cancel{E}_T)$	0.797	0.955
$m(j_{w1} + j_{w2} + \mu_a), m(j_{w1} + j_{w2})$	0.845	0.865
$R_E(j_1), R_E(j_{w1})$	0.762	0.896
$E(j_1), E(j_{w1})$	0.747	0.850
$m(j_{w1} + j_{w2} + \mu_b), m_T(j_{w1} + j_{w2} + \mu_b + \cancel{E}_T)$	0.724	0.926
$m(j_{w1} + j_{w2} + \mu_b), m(j_{w1} + j_{w2})$	0.710	0.873
$m_T(j_{w1} + j_{w2} + \mu_b + \cancel{E}_T), m(j_{w1} + j_{w2})$	0.715	0.861
$p_z(j_1), p_z(j_{w1})$	0.791	0.838
$m(j_{w1} + j_{w2} + \mu_b), m(j_{w1} + j_{w2} + \mu_a)$	0.740	0.861
$N_{\text{track}}(j_1), N_{\text{track}}(j_{w1})$	0.761	0.772

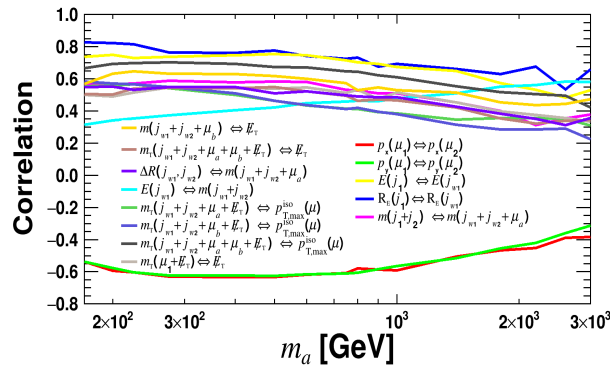


Fig. B2. (color online) Correlations between pairs of observables for the signal as a function of m_a .

weighting each split occurrence by the separation gain-squared it has achieved and by the number of events in the node [56, 59]. The left plot highlights observables with the highest importance values, while those with the lowest importance values are displayed in the right panel.

Table B1 presents the correlations between two observables for the total background and signal with benchmark mass $m_a = 400$ GeV. The listed 12 pairs have strong correlations and remain stable under variations of m_a . Figure B2 displays the correlations between pairs of ob-

servables for the signal as a function of m_a . The correlations for the selected pairs exhibit significant dependence on m_a .

APPENDIX C: DISTRIBUTIONS OF BDT RESPONSES

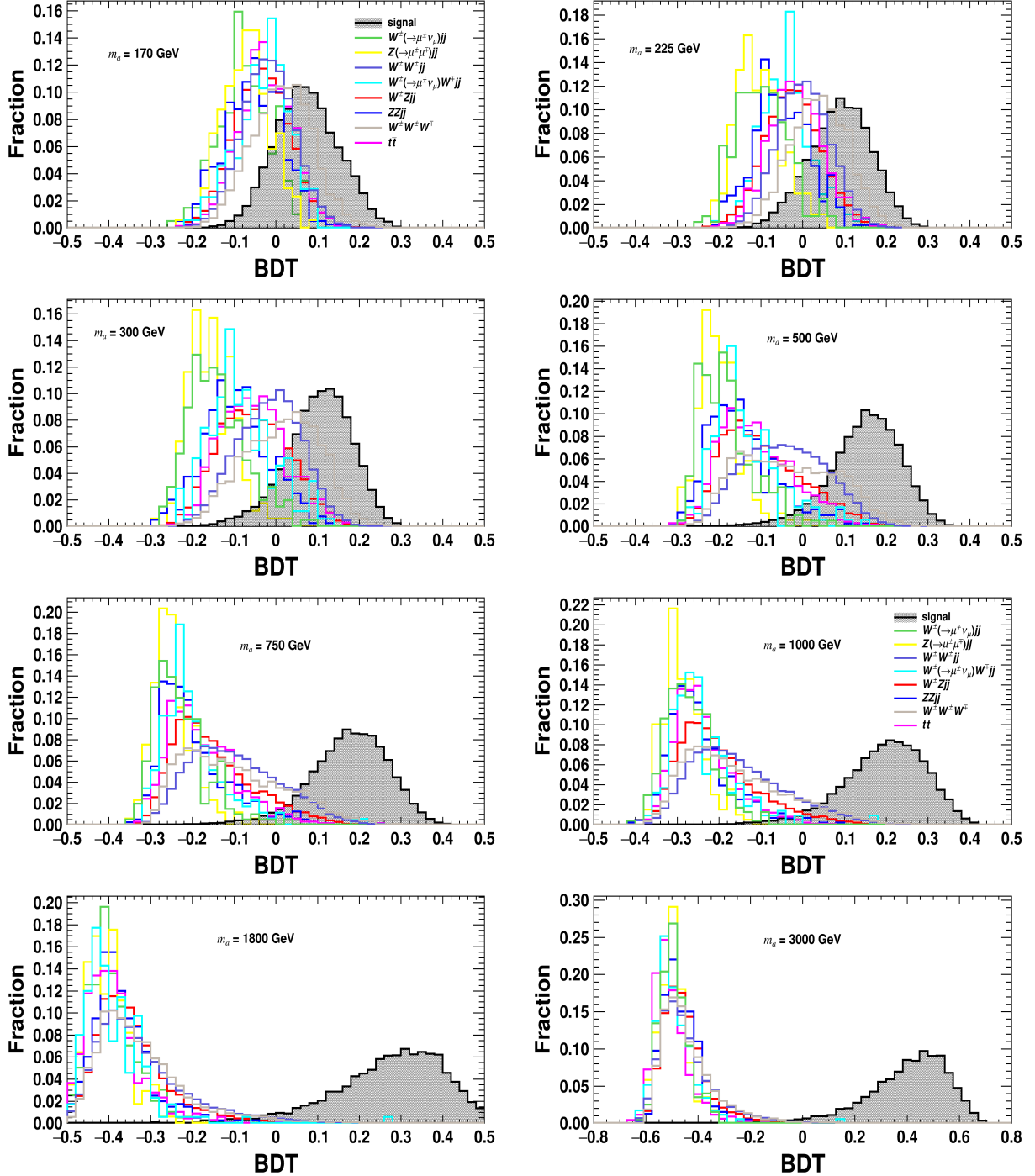


Fig. C1. (color online) Distributions of BDT responses after applying preselection criteria for the signal (black, shadow) and background processes at the HL-LHC with $\sqrt{s} = 14$ TeV in representative m_a cases.

APPENDIX D: SELECTION EFFICIENCY TABLE

Table D1. Selection efficiencies of preselection and BDT criteria for signal and background processes at the HL-LHC with $\sqrt{s} = 14$ TeV, assuming different ALP masses, where "-" means the number of events can be reduced to negligible with $\mathcal{L} = 3 \text{ ab}^{-1}$.

m_a	selection	signal	$W^\pm Z jj$	$ZZ jj$	$Z(\rightarrow \mu^\pm \mu^\mp) jj$	$t\bar{t}$	$W^\pm(\rightarrow \mu^\pm \nu_\mu) W^\mp jj$	$W^\pm(\rightarrow \mu^\pm \nu_\mu) jj$	$W^\pm W^\pm jj$	$W^\pm W^\pm W^\mp$
170 GeV	preselection	2.18×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.127	2.46×10^{-1}	8.06×10^{-3}	7.50×10^{-3}	—	5.76×10^{-3}	—	—	1.08×10^{-2}	5.61×10^{-2}
200 GeV	preselection	2.79×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.108	4.15×10^{-1}	3.93×10^{-2}	1.50×10^{-2}	—	2.74×10^{-2}	1.14×10^{-2}	—	5.77×10^{-2}	1.86×10^{-1}
225 GeV	preselection	3.24×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.101	4.54×10^{-1}	3.70×10^{-2}	7.50×10^{-3}	—	4.61×10^{-2}	1.71×10^{-2}	—	5.63×10^{-2}	1.91×10^{-1}
275 GeV	preselection	3.76×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.115	4.79×10^{-1}	3.04×10^{-2}	1.00×10^{-2}	—	1.59×10^{-2}	5.71×10^{-3}	—	4.95×10^{-2}	1.66×10^{-1}
400 GeV	preselection	4.40×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.136	4.99×10^{-1}	1.12×10^{-2}	—	—	5.76×10^{-3}	—	—	2.49×10^{-2}	8.76×10^{-2}
500 GeV	preselection	4.61×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.165	5.38×10^{-1}	3.60×10^{-3}	—	—	1.44×10^{-3}	—	—	1.20×10^{-2}	3.49×10^{-2}
600 GeV	preselection	4.71×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.173	4.72×10^{-1}	2.45×10^{-3}	—	—	1.44×10^{-3}	1.14×10^{-2}	—	7.13×10^{-3}	2.54×10^{-2}
750 GeV	preselection	4.64×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.167	5.68×10^{-1}	2.02×10^{-3}	—	—	1.44×10^{-3}	1.14×10^{-2}	—	9.64×10^{-3}	2.04×10^{-2}
800 GeV	preselection	4.77×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.165	7.17×10^{-1}	3.46×10^{-3}	—	—	1.44×10^{-3}	1.14×10^{-2}	—	1.00×10^{-2}	2.02×10^{-2}
900 GeV	preselection	4.70×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.180	5.68×10^{-1}	1.44×10^{-3}	—	—	1.44×10^{-3}	5.71×10^{-3}	—	5.01×10^{-3}	1.01×10^{-2}
1000 GeV	preselection	4.72×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.140	8.69×10^{-1}	3.17×10^{-3}	—	—	2.88×10^{-3}	1.14×10^{-2}	—	1.35×10^{-2}	1.80×10^{-2}
1400 GeV	preselection	4.77×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.150	8.36×10^{-1}	2.02×10^{-3}	—	—	1.44×10^{-3}	5.71×10^{-3}	—	8.17×10^{-3}	9.11×10^{-3}
1800 GeV	preselection	4.78×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.150	8.31×10^{-1}	5.76×10^{-4}	—	—	1.44×10^{-3}	5.71×10^{-3}	—	3.50×10^{-3}	3.20×10^{-3}
2200 GeV	preselection	4.67×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.080	9.27×10^{-1}	7.20×10^{-4}	—	—	—	5.71×10^{-3}	—	5.01×10^{-3}	3.20×10^{-3}
2600 GeV	preselection	4.68×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.004	9.34×10^{-1}	4.32×10^{-4}	—	—	—	5.71×10^{-3}	—	3.07×10^{-3}	2.22×10^{-3}
3000 GeV	preselection	4.61×10^{-3}	1.93×10^{-4}	5.00×10^{-5}	2.01×10^{-6}	1.63×10^{-5}	8.97×10^{-6}	1.35×10^{-6}	3.53×10^{-3}	2.26×10^{-3}
	BDT>0.003	9.69×10^{-1}	8.64×10^{-4}	—	—	—	5.71×10^{-3}	—	4.82×10^{-3}	5.17×10^{-3}

ACKNOWLEDGMENTS

We thank Bin Diao and Zilong Ding for helpful discussions.

References

- [1] P. Langacker, *Rev. Mod. Phys.* **81**, 1199 (2009), arXiv: [0801.1345](#)
- [2] G. C. Branco, P. M. Ferreira, L. Lavoura *et al.*, *Phys. Rept.* **516**, 1 (2012), arXiv: [1106.0034](#)

- [3] A. Djouadi, *Phys. Rept.* **459**, 1 (2008), arXiv: [hep-ph/0503173](#)
- [4] S. Bhattacharya, G. Coloretti, A. Crivellin *et al.*, *Growing Excesses of New Scalars at the Electroweak Scale*, arXiv: [2306.17209](#)
- [5] R. D. Peccei and H. R. Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977)
- [6] R. D. Peccei and H. R. Quinn, *Phys. Rev. D* **16**, 1791 (1977)
- [7] S. Weinberg, *Phys. Rev. Lett.* **40**, 223 (1978)
- [8] F. Wilczek, *Phys. Rev. Lett.* **40**, 279 (1978)
- [9] J. E. Kim and G. Carosi, *Rev. Mod. Phys.* **82**, 557 (2010), arXiv: [0807.3125](#)
- [10] G. Panico and A. Wulzer, *The Composite Nambu-Goldstone Higgs*, vol. 913, Springer (2016), [10.1007/978-3-319-22617-0](#), arXiv: [1506.01961](#)
- [11] C. Csaki, M. L. Graesser, and G. D. Kribs, *Phys. Rev. D* **63**, 065002 (2001), arXiv: [hep-th/0008151](#)
- [12] F. Kahlhoefer, *Int. J. Mod. Phys. A* **32**, 1730006 (2017), arXiv: [1702.02430](#)
- [13] N. Craig, A. Hook, and S. Kasko, *JHEP* **09**, 028 (2018), arXiv: [1805.06538](#)
- [14] M. Dine, *TASI lectures on the strong CP problem*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2000): Flavor Physics for the Millennium*, pp. 349–369, (2000), arXiv: [hep-ph/0011376](#)
- [15] J. E. Kim, *Phys. Rev. Lett.* **43**, 103 (1979)
- [16] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Nucl. Phys. B* **166**, 493 (1980)
- [17] A.R. Zhitnitsky, *Sov. J. Nucl. Phys.* **31**, 260 (1980)
- [18] M. Dine, W. Fischler, and M. Srednicki, *Phys. Lett. B* **104**, 199 (1981)
- [19] A. Hook, S. Kumar, Z. Liu *et al.*, *Phys. Rev. Lett.* **124**, 221801 (2020), arXiv: [1911.12364](#)
- [20] J. Quevillon and C. Smith, *Eur. Phys. J. C* **79**, 822 (2019), arXiv: [1903.12559](#)
- [21] L. Di Luzio, M. Giannotti E. Nardi *et al.*, *Phys. Rept.* **870**, 1 (2020), arXiv: [2003.01100](#)
- [22] G. Galanti and M. Roncadelli, *Universe* **8**, 253 (2022), arXiv: [2205.00940](#)
- [23] K. Choi, S. H. Im and C. Sub Shin, *Ann. Rev. Nucl. Part. Sci.* **71**, 225 (2021), arXiv: [2012.05029](#)
- [24] Q. Qiu, Y. Gao, H. j. Tian *et al.*, *Wide Binary Evaporation by Dark Solitons: Implications from the GAIA Catalog*, arXiv: [2404.18099](#)
- [25] M. Bauer, M. Neubert, and A. Thamm, *JHEP* **12**, 044 (2017), arXiv: [1708.00443](#)
- [26] M. J. Dolan, T. Ferber, C. Hearty *et al.*, *JHEP* **12**, 094 (2017), arXiv: [1709.00009](#)
- [27] M. Bauer, M. Heiles, M. Neubert *et al.*, *Eur. Phys. J. C* **79**, 74 (2019), arXiv: [1808.10323](#)
- [28] H. Y. Zhang, C. X. Yue, Y. C. Guo *et al.*, *Phys. Rev. D* **104**, 096008 (2021), arXiv: [2103.05218](#)
- [29] D. d'Enterria, *Collider constraints on axion-like particles*, in *Workshop on Feebly Interacting Particles*, 2, (2021), arXiv: [2102.08971](#)
- [30] P. Agrawal *et al.*, *Eur. Phys. J. C* **81**, 1015 (2021), arXiv: [2102.12143](#)
- [31] M. Tian, Z. S. Wang, and K. Wang, *Search for long-lived axions with far detectors at future lepton colliders*, arXiv: [2201.08960](#)
- [32] F. A. Ghebretinsaea, Z. S. Wang, and K. Wang, *JHEP* **07**, 070 (2022), arXiv: [2203.01734](#)
- [33] C. Antel *et al.*, *Eur. Phys. J. C* **83**, 1122 (2023), arXiv: [2305.01715](#)
- [34] T. Biswas, *JHEP* **05**, 081 (2024), arXiv: [2312.05992](#)
- [35] Y. Lu, Y.-n. Mao, K. Wang *et al.*, *LAYCAST: LAYered CAvern Surface Tracker at future electron-positron colliders*, arXiv: [2406.05770](#)
- [36] S. Navas *et al.* (Particle Data Group), *Phys. Rev. D* **110**, 030001 (2024)
- [37] C. O'hare, *cajohare/AxionLimits: AxionLimits*, (2024), <https://doi.org/10.5281/zenodo.3932429>
- [38] BELLE-II Collaboration, *Phys. Rev. Lett.* **125**, 161806 (2020), arXiv: [2007.13071](#)
- [39] BESIII collaboration, *Phys. Lett. B* **838**, 137698 (2023), arXiv: [2211.12699](#)
- [40] BESIII Collaboration, *ALPs searches at BESIII*, in *57th Rencontres de Moriond on Electroweak Interactions and Unified Theories*, 5, (2023), arXiv: [2305.08043](#)
- [41] CMS Collaboration, *Phys. Lett. B* **797**, 134826 (2019), arXiv: [1810.04602](#)
- [42] ATLAS Collaboration, *JHEP* **03**, 243 (2021), arXiv: [2008.05355](#)
- [43] J. Bonilla, I. Brivio, J. Machado-Rodríguez *et al.*, *JHEP* **06**, 113 (2022), arXiv: [2202.03450](#)
- [44] M. Aiko, M. Endo, and K. Fridell, *JHEP* **06**, 194 (2024), arXiv: [2401.13323](#)
- [45] ATLAS Collaboration, *Eur. Phys. J. C* **77**, 141 (2017), arXiv: [1610.05088](#)
- [46] CMS Collaboration, *Phys. Lett. B* **809**, 135710 (2020), arXiv: [2005.01173](#)
- [47] CMS Collaboration, *Phys. Rev. D* **100**, 012004 (2019), arXiv: [1905.04246](#)
- [48] Z. Ding, Y.-n. Mao, and K. Wang, *Search for the γZ decay mode of heavy photophobic axion-like particles at the LHC*, arXiv: [2411.08660](#)
- [49] M. Aiko and M. Endo, *JHEP* **05**, 147 (2023), arXiv: [2302.11377](#)
- [50] I. Brivio, M. B. Gavela, L. Merlo *et al.*, *Eur. Phys. J. C* **77**, 572 (2017), arXiv: [1701.05379](#)
- [51] J. Alwall, R. Frederix, S. Frixione *et al.*, *JHEP* **07**, 079 (2014), arXiv: [1405.0301](#)
- [52] R. D. Ball *et al.*, *Nucl. Phys. B* **867**, 244 (2013), arXiv: [1207.1303](#)
- [53] T. Sjöstrand, S. Ask, J. R. Christiansen *et al.*, *Comput. Phys. Commun.* **191**, 159 (2015), arXiv: [1410.3012](#)
- [54] DELPHES 3 Collaboration, *JHEP* **02**, 057 (2014), arXiv: [1307.6346](#)
- [55] C. Degrande, C. Duhr, B. Fuks *et al.*, *Comput. Phys. Commun.* **183**, 1201 (2012), arXiv: [1108.2040](#)
- [56] TMVA Collaboration, *TMVA - Toolkit for Multivariate Data Analysis*, arXiv: [physics/0703039](#)
- [57] J. Thaler, <http://madgraph.phys.ucl.ac.be/Manual/lhco.html>
- [58] T. Han, *Collider phenomenology: Basic knowledge and techniques*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in $D \geq 4$* , pp. 407–454, 8, (2005), arXiv: [hep-ph/0508097](#)
- [59] L. Breiman, J. Friedman, R. A. Olshen *et al.*, *Classification and Regression Trees*, Chapman and Hall/CRC (10, 2017), [10.1201/9781315139470](#).