Chinese Physics C Vol. 50, No. 6 (2026) 063107

Discovery prospects for photophobic axion-like particles in the *WWjj* final state at the High-Luminosity LHC*

Jiaojiao Feng (冯娇娇)[†#] Ying-nan Mao (毛英男)[‡#] Kechen Wang (王科臣)[§#]

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

Abstract: We evaluate discovery prospects for photophobic axion-like particles (ALPs) in the $WWjj$ final state at the High-Luminosity LHC (HL-LHC; $\sqrt{s}$ = 14 TeV, $\mathcal{L}$ = 3 ab^{-1}). In the photophobic limit ($g_{a\gamma\gamma}=0$), ALPs couple to electroweak gauge bosons and are produced in association with two jets ($pp\to jj\,a$) via both s-channel electroweak exchange and vector boson fusion (VBF)–like topologies, followed by $a\to W^+W^-$ decay. We target the different-flavour dilepton mode $W^+W^-\to e^\pm\mu^\mp\nu\bar{\nu}$ with two jets and moderate missing transverse momentum. The analysis employs a two-step strategy: initial preselection that defines the signal-like final state, followed by a multivariate analysis (MVA) trained on dijet and dileptonic-WW kinematics from both the s-channel and VBF-like production mechanisms to separate signal from background; the MVA threshold is optimized independently at each m_a. We present 2σ and 5σ discovery sensitivities for the ALP–W coupling g_{aWW} across the 170–4000 GeV mass range. For 260 GeV $\le m_a\le$ 1500 GeV, the 2σ (5σ) sensitivity is approximately flat around 0.61(0.76) TeV^{-1}. At large masses/couplings, the mapping between an experimental rate and a single EFT parameter g_{aWW} can become UV sensitive, and we treat the extreme high-mass/high-coupling corner as an extrapolation; therefore, we also report model-independent discovery thresholds for the fiducial quantity $\sigma(pp\to jj\,a)\times\mathrm{Br}(\mathrm{a}\to\mathrm{W}^+\mathrm{W}^-)$ over the same mass range to enable reinterpretation for other models. These results indicate that the $WWjj$ topology offers competitive and complementary sensitivity to heavy photophobic ALPs at the HL-LHC.

Keywords: axion-like particles, collider phenomenology, the High-Luminosity LHC, the *WWjj* final state, physics beyond the Standard Model

DOI: 10.1088/1674-1137/ae5048 **CSTR:** 32044.14.ChinesePhysicsC.50063107

1. INTRODUCTION

Quantum chromodynamics (QCD) allows for a CP-violating θ term, yet no corresponding effect has been observed in measurements of the neutron electric dipole moment [1, 2]. This long–standing "strong CP problem" can be resolved by enlarging the Standard Model (SM) with a global $U(1)_{\rm PQ}$ symmetry, known as the Peccei-Quinn symmetry, whose spontaneous breaking generates a pseudo-Nambu–Goldstone boson, the axion [3−6]. In traditional QCD-axion models, the axion mass and its couplings are tightly linked to the symmetry-breaking scale [5−12].

However, in many extensions of the SM, pseudo-scalar states with axion-like couplings appear, whose masses and interaction strengths are not fixed by the strong CP solution. These states are collectively referred to as axion-like particles (ALPs) [13−16]. The decoupling of mass and couplings opens a large parameter space and has led to extensive search programs, ranging from astrophysical and cosmological probes to intensity-frontier experiments and high-energy colliders [17−68].

At colliders, ALPs can interact with SM gauge bosons and/or fermions and can be probed either as on-shell resonances or through off-shell exchange in high-energy processes. A large body of experimental work has already

---

Received 2 December 2025; Accepted 10 March 2026; Accepted manuscript online 11 March 2026

[*] Jiaojiao Feng and Kechen Wang are supported by the National Natural Science Foundation of China (11905162), the Excellent Young Talents Program of the Wuhan University of Technology (40122102, 40122103) and the research program of the Wuhan University of Technology (2020IB024, 104972025KFYjc0101). Ying-nan Mao is supported by the National Natural Science Foundation of China (12205227). 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: jiaojiao.feng@whut.edu.cn (Corresponding author)

[‡] E-mail: ynmao@whut.edu.cn (Corresponding author)

[§] E-mail: kechen.wang@whut.edu.cn (Corresponding author)

[#] These authors contributed equally as the first authors

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 SCOAP^3 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

constrained the ALP coupling to photons, $g_{a\gamma\gamma}$, using e^+e^- data and heavy-ion or proton–proton collisions, particularly via diphoton or light-by-light signatures [69−76]. The resulting bounds are strong over a wide mass range and motivate scenarios in which $g_{a\gamma\gamma}$ is tuned to be small while couplings to the electroweak gauge bosons remain sizeable. In such *photophobic* constructions, the $a\gamma\gamma$ vertex is suppressed, but interactions with $W^{\pm}$, Z, and γZ can persist [77−79], leading to collider signatures that differ qualitatively from generic diphoton ALP searches.

Heavy photophobic ALPs at the LHC have been explored in several ways. One line of work has reinterpreted SM triboson measurements at $\sqrt{s}=8$ TeV as signals of ALP-mediated processes, deriving limits up to a few hundred GeV in mass [77]. Other studies have mapped nonresonant ALP exchange onto vector-boson-scattering topologies [80] or recast Run-2 analyses targeting $W^{\pm}W^{\pm}W^{\mp}$ and $Z\gamma$ final states to constrain resonant ALP production and decay into electroweak boson pairs [81]. An overview of these and related collider bounds can be found in Ref. [82]. Because the original analyses were optimized for SM processes rather than for on-shell ALP production, differences in acceptance, kinematic selections, and background control often make the recast limits conservative and non-uniform across the parameter space.

More recently, dedicated projections have been carried out specifically for photophobic ALPs at the High-Luminosity LHC (HL-LHC) [83]. In Ref. [82], a detector-level study of $pp\to jja(\to Z\gamma)$ was performed, including both s-channel and vector-boson–fusion (VBF)–like production mechanisms, reconstructing the ALP resonance and using a multivariate analysis (MVA) to obtain discovery reaches on both g_{aWW} and $\sigma(pp\to jja)\times\mathrm{Br}(a\to Z\gamma)$ over $m_a\simeq(100\text{–}4000)$ GeV at $\sqrt{s}=14$ TeV. In parallel, Ref. [84] proposed a tri-W signature, $pp\to W^{\pm}X(\to W^+W^-)$, using a same-sign dimuon plus hadronic-W final state and an MVA tailored for background rejection; with heavy photophobic ALPs used as a benchmark, that study obtained sensitivities well beyond those of earlier triboson recasts while remaining applicable to more general neutral resonances coupled to W bosons. These works highlight two important lessons: clean electroweak final states with sharp kinematic handles and multivariate discrimination can dramatically sharpen sensitivity, and photophobic ALP searches should exploit multiple channels with different couplings and systematics.

The present study extends this program to the $WWjj$ topology. We consider photophobic ALPs that couple to the electroweak gauge bosons and are produced in association with two jets via both s-channel electroweak exchange and VBF-like topologies in pp collisions, $pp\to jja$, followed by $a\to W^+W^-$ decay. Both production mechanisms are simulated and analyzed on an equal footing within a single selection-and-MVA framework. This strategy utilizes the full photophobic electroweak structure (aWW, aZZ, $a\gamma Z$) and allows the analysis to adapt to whichever kinematic pattern dominates: s-channel–like configurations are more important at lower m_a, while VBF-like events with forward jets and large rapidity gaps become increasingly relevant at high m_a. As a clean and experimentally robust final state, we focus on the different-flavour dilepton mode $W^+W^-\to e^{\pm}\mu^{\mp}\nu\bar{\nu}$ accompanied by at least two jets, which efficiently suppresses Drell–Yan and multijet backgrounds. Our study provides projected discovery sensitivities on both the coupling g_{aWW} and fiducial quantity $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$ at the HL-LHC with $\sqrt{s}=14$ TeV and $\mathcal{L}=3\ \mathrm{ab}^{-1}$.

Throughout this study, unless stated otherwise, we impose the photophobic condition $g_{a\gamma\gamma}=0$, include both s-channel and VBF-like production of $pp\to jja$ with $a\to W^+W^-$ decay, and optimize the boosted decision tree (BDT) threshold *independently* for each assumed ALP mass. The quoted significances are purely statistical (systematic uncertainties are not included; cf. Sec. VI).

We note that our HL-LHC projection in the $pp\to jja(\to W^+W^-)$ topology should be read in the context of several recent and complementary studies, including those on the $jja(\to Z\gamma)$ channel [82], tri-W strategy $pp\to W^{\pm}X(\to W^+W^-)$ [84], and related Run-1/2 recasts [77, 80, 81]. Relative to the tri-W strategy [84], our $WWjj$ analysis typically achieves a stronger reach across the full mass range, with a particularly pronounced advantage for $m_a\gtrsim 700$ GeV, owing to the VBF-like production and kinematics. By contrast, $jja(\to Z\gamma)$ [82] attains a stronger reach overall—benefiting from a narrow $Z\gamma$ resonance and cleaner backgrounds. Meanwhile, $jja(\to W^+W^-)$ remains complementary: it *directly* probes the aWW vertex, thus enabling nontrivial cross-checks of the photophobic relations among aWW, $aZ\gamma$, and aZZ couplings when combined with previous results [82, 84]; it is subject to largely orthogonal systematics and backgrounds relative to $Z\gamma$; and our model-independent thresholds in $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$ facilitate straightforward reinterpretation for other neutral resonances that couple to W bosons [84].

The remainder of this paper is organized as follows. Sec. II introduces the effective description of photophobic ALPs and their electroweak couplings. In Secs. III and IV, we define the signal process and summarize the relevant SM backgrounds. The event selection, observable construction, and multivariate strategy are described in Sec. V. Our main results are presented in Sec. VI, and we conclude in Sec. VII. Additional distributions, BDT response shapes, and efficiency tables are presented in the appendices.

II. THEORY SETUP

We work with an ALP whose interactions with the SM gauge bosons arise solely through the $SU(2)_L$ and $U(1)_Y$ field strengths. At the level before electroweak symmetry breaking (EWSB), the relevant effective Lagrangian can be written as [85]

$$\mathcal{L}_{\rm ALP} \supset \frac{1}{2}\partial_\mu a\,\partial^\mu a - \frac{1}{2}m_a^2 a^2 - \frac{c_{\widetilde{W}}}{f_a}\, a\, W^b_{\mu\nu}\widetilde{W}^{b,\mu\nu} - \frac{c_{\widetilde{B}}}{f_a}\, a\, B_{\mu\nu}\widetilde{B}^{\mu\nu}, \tag{1}$$

where a and m_a denote the ALP field and its mass, respectively, and f_a is the associated decay constant. The tensors $W^b_{\mu\nu}$ (with $b = 1,2,3$) and $B_{\mu\nu}$ are the usual $SU(2)_L$ and $U(1)_Y$ field strengths, and the corresponding dual tensor is defined as $\widetilde{X}^{\mu\nu} \equiv 1/2\epsilon^{\mu\nu\alpha\beta}X_{\alpha\beta}$ with $X = W^b, B$. The dimensionless coefficients $c_{\widetilde{W}}$ and $c_{\widetilde{B}}$ encode the strengths of the ALP couplings to the electroweak gauge sector.

After EWSB, it is convenient to express the interactions in the mass eigenstate basis for the gauge bosons. The terms involving neutral and charged electroweak bosons can be recast as

$$\begin{aligned}\mathcal{L}_{\rm int} \supset & -\frac{1}{4}\, g_{a\gamma\gamma}\, a\, F_{\mu\nu}\widetilde{F}^{\mu\nu} - \frac{1}{2}\, g_{aZ\gamma}\, a\, Z_{\mu\nu}\widetilde{F}^{\mu\nu} \\ & -\frac{1}{4}\, g_{aZZ}\, a\, Z_{\mu\nu}\widetilde{Z}^{\mu\nu} - \frac{1}{2}\, g_{aWW}\, a\, W^+_{\mu\nu}\widetilde{W}^{-\mu\nu},\end{aligned} \tag{2}$$

with the tree-level relations

$$g_{a\gamma\gamma} = \frac{4}{f_a}\left(s_\theta^2 c_{\widetilde{W}} + c_\theta^2 c_{\widetilde{B}}\right), \tag{3}$$

$$g_{aZZ} = \frac{4}{f_a}\left(c_\theta^2 c_{\widetilde{W}} + s_\theta^2 c_{\widetilde{B}}\right), \tag{4}$$

$$g_{aZ\gamma} = \frac{2}{f_a}\, s_{2\theta}\left(c_{\widetilde{W}} - c_{\widetilde{B}}\right), \tag{5}$$

$$g_{aWW} = \frac{4}{f_a}\, c_{\widetilde{W}}. \tag{6}$$

Here, θ is the weak mixing angle, and we define $s_\theta \equiv \sin\theta$, $c_\theta \equiv \cos\theta$, and $s_{2\theta} \equiv \sin 2\theta$ for simplicity. Guided by the strong experimental constraints on the diphoton channel, we restrict ourselves to the photophobic case in which the $a\gamma\gamma$ coupling vanishes, $g_{a\gamma\gamma} = 0$ (see, *e.g.*, [77−82, 84]). Imposing

$$s_\theta^2 c_{\widetilde{W}} + c_\theta^2 c_{\widetilde{B}} = 0, \quad \text{or equivalently} \quad c_{\widetilde{B}} = -t_\theta^2\, c_{\widetilde{W}}, \tag{7}$$

with $t_\theta \equiv \tan\theta$, $g_{a\gamma\gamma}$ is removed at the tree level; thus, the remaining electroweak couplings correlate with each other. In particular, one finds

$$g_{aZZ} = \left(1 - t_\theta^2\right) g_{aWW}, \tag{8}$$

$$g_{aZ\gamma} = t_\theta\, g_{aWW}. \tag{9}$$

In the following, we use m_a and g_{aWW} as the primary parameters, with g_{aZZ} and $g_{aZ\gamma}$ fixed by these relations throughout our phenomenological analysis of the photophobic scenario.

Deviations from the strict photophobic limit can arise in realistic UV completions: $g_{a\gamma\gamma} = 0$ is a matching condition and can receive threshold corrections; therefore, a small tree-level $g_{a\gamma\gamma}$ at collider scales is possible [16, 77]. We parametrize a controlled departure from the strict relation via

$$\frac{c_{\widetilde{B}}}{c_{\widetilde{W}}} = -t_\theta^2 + \delta, \tag{10}$$

for which $g_{a\gamma\gamma} \propto \delta$ and thus ${\rm Br}(a \to \gamma\gamma) \propto \delta^2$ (at fixed g_{aWW}). For m_a above the WW threshold, $a \to W^+W^-$ remains dominant for modest departures, while larger δ enhances $a \to \gamma\gamma$ and dilutes ${\rm Br}(a \to W^+W^-)$ (see Appendix A.1). Independently, within the strict photophobic effective field theory (EFT), the photophobic condition is preserved by renormalization group (RG) evolution up to the two-loop level, and loop/RG-induced rare channels remain at the percent level (or below) in our mass range [82]. Therefore, our results are also reported as $\sigma(pp \to jja) \times {\rm Br}(a \to W^+W^-)$ to enable straightforward recasts.

We interpret the dimension-5 operators in Eq. (1) as part of an EFT valid below a model-dependent cutoff Λ; EFT-validity, unitarity, and width considerations become particularly relevant in the high-mass/high-coupling region. A weakly-coupled UV completion suggests $\Lambda \lesssim f_a/|c_{\widetilde{W}}|$ up to coupling/loop factors, while naive dimensional analysis (NDA) for strong dynamics suggests $\Lambda \lesssim 4\pi f_a/|c_{\widetilde{W}}|$ [86, 87]. Using $g_{aWW} = 4c_{\widetilde{W}}/f_a$, this corresponds to

$$\begin{aligned}\Lambda &\lesssim \frac{4}{|g_{aWW}|} \quad \text{(weakly coupled)}, \\ \Lambda &\lesssim \frac{16\pi}{|g_{aWW}|} \quad \text{(strongly coupled)}.\end{aligned} \tag{11}$$

We interpret our results in the region where m_a and representative hard scales (*e.g.*, momentum transfer) are below Λ. Partial-wave unitarity provides an independent (process-dependent) guide: in the gauge-invariant ALP

EFT the strongest bound arises from $VV \to Va$ (dominant $J=1$ eigenvalue), primarily constraining the $SU(2)_L$ coupling [88]; under our convention, this leads to Eq. (A2) in Appendix A.2, which we use as a conservative validity indicator.

Although m_a cannot be fully reconstructed in the $\ell\ell + E_T^{\rm miss}$ final state, our signal modelling assumes on-shell $pp \to jja$ and we present results in the factorised form $\sigma(pp \to jja) \times {\rm Br}(a \to WW)$. We therefore compute the total width in the strict photophobic benchmark with `MadGraph5_aMC@NLO` and report Γ_a/m_a along the $2\sigma/5\sigma$ reach curves in Appendix A.2. The width remains small in the sensitivity region for $m_a \lesssim 2$ TeV, while the extreme high-mass/high-coupling corner becomes broad and is treated as an extrapolation[1]; We therefore quote our primary results also as model-independent thresholds on $\sigma(pp \to jja) \times {\rm Br}(a \to WW)$ for reinterpretation; see Appendix A.2 for the EFT-validity, unitarity, and width checks.

III. SIGNAL PRODUCTION

In the photophobic limit $g_{a\gamma\gamma} = 0$, the ALP retains anomalous couplings to electroweak gauge bosons, notably aWW, aZZ, and $a\gamma Z$. We denote $V \in \{W, Z, \gamma\}$ and $V' \in \{W, Z\}$. Consequently, the ALP can be produced at proton colliders in association with one vector boson ($pp \to V'a$) through s-channel vector-boson exchange or with two jets ($pp \to jja$) via vector-boson-fusion (VBF)–like topologies, and it can decay as $a \to VV'$ with branching ratios varying with the ALP mass [82, 89]. For sufficiently large m_a, the $a \to W^+W^-$ mode dominates.

For decay channel choice, the strict photophobic condition guarantees the relations in Eqs. (8) and (9). Thus, for $m_a \gg m_W$, one has ${\rm Br}(a \to W^+W^-) : {\rm Br}(a \to ZZ) : {\rm Br}(a \to Z\gamma) \simeq 0.65 : 0.16 : 0.19$ [82]. Including leptonic decays, ${\rm Br}(WW \to e^\pm\mu^\mp + \not{E}_T) \simeq 2.3 \times 10^{-2}$ while ${\rm Br}(ZZ \to 4\ell) \simeq 4.5 \times 10^{-3}$ $(\ell = e, \mu)$ [90], implying $[{\rm Br}(a \to WW)\,{\rm Br}(WW \to e\mu + \not{E}_T)]/[{\rm Br}(a \to ZZ){\rm Br}(ZZ \to 4\ell)] \sim \mathcal{O}(20)$ in the heavy-mass regime. We thus focus on $W^+W^- \to e^\pm\mu^\mp + \not{E}_T$ as the rate-favored channel, noting that $ZZ \to 4\ell$ is highly complementary due to its excellent mass resolution and low backgrounds.

As illustrated in Fig. 1, we consider two representative production processes: (i) an s-channel electroweak-boson exchange yielding $pp \to jja$, with the two jets predominantly arising from the decay of the associated vector boson V', and (ii) an electroweak VBF-like topology in which the ALP is radiated off the t-channel gauge bosons. In both cases, the effective a–V–V' vertices include aWW, aZZ, and $a\gamma Z$. We focus on the different-flavour dilepton final state, $W^+W^- \to e^\pm\mu^\mp\nu\bar{\nu}$, which suppresses multijet backgrounds and Drell–Yan $Z/\gamma^* \to \ell^+\ell^-$ contamination while retaining high signal efficiency.

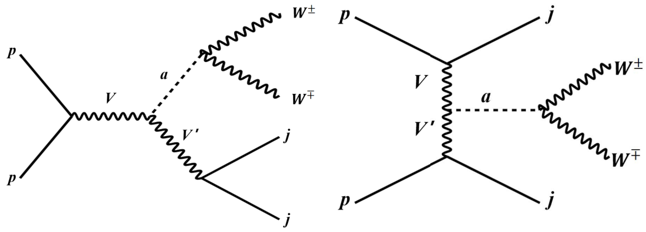

Fig. 1. Representative production mechanisms for heavy photophobic ALPs at pp colliders: s-channel exchange (left) and VBF-like topology (right), yielding $pp \to jja$ with a–V–V' vertices (aWW, aZZ, $a\gamma Z$). The ALP subsequently decays to two W bosons, defining the signal $pp \to jja(\to W^\pm W^\mp)$ with $W^\pm \to \ell^\pm\bar{\nu}/\nu$.

The VBF-like component is characterized by two forward tagging jets with large invariant mass m_{jj} and sizeable pseudo-rapidity difference $|\Delta\eta_{jj}|$, together with reduced central hadronic activity; meanwhile, the s-channel contribution typically yields a more central dijet event with smaller rapidity separation and an invariant mass near the parent-boson mass when the jets arise from $V' \to jj$. Accordingly, the MVA in Sec. V is trained on the inclusive (combined) signal and uses variables that are simultaneously sensitive to both topologies and kinematic observables that reconstruct the common decay $a \to W^+W^- \to e^\pm\mu^\mp\nu\bar{\nu}$ in the $e^\pm\mu^\mp + \not{E}_T$ final state.

Signal and background samples are generated with MadGraph5_aMC@NLO [91] using the NNPDF2.3 parton distribution functions (PDFs) [92]. The ALP interactions are implemented via a UFO model based on the linear Lagrangian [87, 93]. Parton showering and hadronization are performed with PYTHIA 8.3 [94], and the detector response is simulated with Delphes 3 [95] using HL-LHC CMS detector cards. The HL-LHC center-of-mass energy is set to $\sqrt{s} = 14$ TeV (7 TeV per beam). The hard-scattering process $pp \to jja$ is generated with MadGraph5_aMC@NLO, while the subsequent decays $a \to W^+W^-$ and $W^\pm \to \ell^\pm\nu/\bar{\nu}$ $(\ell = e, \mu)$, yielding $a \to W^+(\to \ell^+\nu)\,W^-(\to \ell^-\bar{\nu})\,jj$, are simulated with PYTHIA 8.3.

We generate signal samples at representative ALP masses

$$m_a \in \{170,\ 185,\ 200,\ 230,\ 260,\ 350,\ 500,\ 750,\ 1000,\\ 1500,\ 2000,\ 2500,\ 3000,\ 4000\}\ {\rm GeV}.$$

At the HL-LHC, we generate 4.8, 0.8, and 4.0 million events for the benchmark points $m_a = 200, 260,$ and 350 GeV, respectively: 0.4 million events for each point with $m_a = 170, 185, 230$ GeV; 2.0 million events for each

---

[1] In many perturbative UV completions, the effective electroweak ALP couplings are loop suppressed, so reaching a given g_{aWW} may require sizable mediator multiplicity/charges or stronger dynamics, motivating complementary direct searches for electroweak-charged mediators [28, 77].

point with $m_a = 500, 750, 4000$ GeV; and 1.0 million events for each point with $m_a = 1000, 1500, 2000, 2500, 3000$ GeV. To remain inclusive at the generator level and to match experimental conditions, we employ the following loose parton-level cuts in the MadGraph5 configuration card:

(i) Minimal transverse momentum of jets, photons, and charged leptons:
$p_{\rm T}(j/\gamma/\ell^{\pm}) > 0.5$ GeV;

(ii) Maximal pseudorapidity:
$|\eta(j)| < 10$ for jets and $|\eta(\gamma/\ell^{\pm})| < 5$ for photons and charged leptons;

(iii) Minimal angular separation $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ between any two objects:
$\Delta R > 0.1$.

These generator-level requirements are deliberately much looser than the requirements for the analysis selection and are never applied as analysis cuts. They avoid generator-induced inefficiencies (*i.e.*, an artificial suppression of soft/forward configurations) and only regulate infrared-sensitive regions for technical stability [91]. Event realism is achieved downstream: parton showering and hadronization model QCD radiation. The detector simulation emulates experimental acceptance and resolution [94, 95]. Accordingly, the fiducial phase space and physics acceptance are defined by the detector simulation together with the selection in Sec. V, in line with LHC fiducial-measurement practice.

Figure 2 shows the production cross section for $pp \to jja(\to W^{\pm}W^{\mp})$ at $\sqrt{s} = 14$ TeV under the benchmark couplings $g_{a\gamma\gamma} = 0$ and $g_{aWW} = 1$ TeV^{-1}. The two-body WW decay channel opens rapidly once $m_a \gtrsim 2m_W \simeq 160$ GeV, causing a steep rise in $\text{Br}(a \to W^+W^-)$; for example, at $m_a = 185$ GeV, the branching fraction reaches ~49% [82, 89]. Meanwhile, the inclusive production cross section $\sigma(pp \to jja)$ decreases with increasing m_a due to diminishing PDFs, such that $\sigma(pp \to jja) \times \text{Br}(a \to W^+W^-)$ peaks around $m_a \sim 200$ GeV.

IV. BACKGROUND PROCESSES

The signal targets two *oppositely charged, different-flavour* leptons, with at least two jets and moderate missing transverse momentum, *i.e.*, $e^{\pm}\mu^{\mp} + jj + \not{E}_T$. In the $e\mu$ channel, Drell–Yan contamination is strongly suppressed; the dominant backgrounds are well known to include diboson and top-quark production, as well as processes with nonprompt or misidentified leptons. We group them by physics origin and reducibility at the detector level.

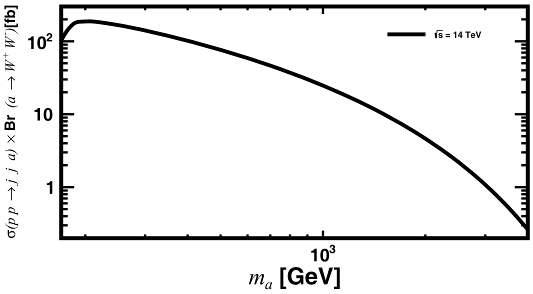

Fig. 2. Production cross section of $pp \to jja(\to W^{\pm}W^{\mp})$ at the HL-LHC ($\sqrt{s} = 14$ TeV) for $g_{a\gamma\gamma} = 0$ and $g_{aWW} = 1$ TeV^{-1} as a function of the ALP mass m_a.

Irreducible (same visible final state)

- W^+W^-jj **(QCD+EW).** Inclusive WW production accompanied by jets, dominated by $q\bar{q} \to WW$ with additional QCD radiation [96]. After leptonic W decays, the visible final state coincides with the signal ($e^{\pm}\mu^{\mp} + jj + \not{E}_T$).

- $t\bar{t}$ **(dileptonic).** When both W bosons decay leptonically, $t\bar{t} \to \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ can be indistinguishable from the signal up to global kinematics if b jets are not correctly tagged; thus, it is effectively irreducible in this channel.

Reducible (lost/misidentified objects)

- $WZjj$, $ZZjj$**.** Either when one prompt lepton is not reconstructed or fails identification/isolation, or when $Z \to \tau\tau$ produces a pair $e\mu$.

- $W^{\pm}jj$**.** Contributes via one prompt lepton from the W and one nonprompt or misidentified lepton (heavy-flavour decays, hadrons mis-identification, photon conversions).

- Zjj**.** Suppressed in $e\mu$ but nonzero through a misidentified lepton, a lost lepton plus a nonprompt lepton, or $Z \to \tau\tau \to e\mu + X$.

The $pp \to W^+W^-jj$ background is generated at matrix-element level with MadGraph5_aMC@NLO, while the subsequent decays $W \to \ell^{\pm}\nu_{\ell}$ $(\ell = e, \mu)$ and parton showering are modeled with PYTHIA 8.3. The $t\bar{t}$, $WZjj$, and $ZZjj$ samples are likewise produced with MadGraph5_aMC@NLO and interfaced to PYTHIA 8.3, which handles all kinematically allowed W- and Z-boson decays. Single-boson processes $W^{\pm}(\to \ell^{\pm}\nu_{\ell})jj$ and $Z(\to \ell^+\ell^-)jj$ with $\ell = e, \mu$ are generated fully differentially in MadGraph5_aMC@NLO, keeping the leptonic W and Z decays explicit in the matrix element.

The expected number of events scales with luminosity as

$$N = \sigma_{\rm pro} \times \mathcal{L} \times \varepsilon_{\rm sel}, \tag{12}$$

where $\sigma_{\rm pro}$ is the production cross section, $\mathcal{L}$ is the integrated luminosity, and $\varepsilon_{\rm sel}$ is the net acceptance efficiency multiplied by the efficiencies from object reconstruction and analysis selections. Table 1 lists the generator-level inclusive production cross sections of background processes at the HL-LHC ($\sqrt{s} = 14$ TeV). In practice, although the Wjj and Zjj processes dominate at the generator level (see Table 1), their contribution after the dilepton+ $\not{E}_T$ selections is strongly reduced, while $WWjj$ and $t\bar{t}$ processes remain the leading backgrounds in the $e\mu$ channel [96].

Table 1. Generator-level inclusive production cross-sections of background processes at the HL-LHC ($\sqrt{s} = 14$ TeV). Values are prior to analysis selections; the dominant Wjj and Zjj rates will be strongly reduced by the dilepton requirements discussed in Sec. V.

Background	σ/pb
$W^+(\to \ell^+\nu)W^-(\to \ell^-\bar{\nu})jj$	3.62×10^1
$WZjj$	3.44×10^2
$ZZjj$	1.03×10^2
$W^\pm(\to \ell^\pm\nu_\ell/\bar{\nu}_\ell)jj$	2.42×10^5
$Z(\to \ell^+\ell^-)jj$	2.21×10^4
$t\bar{t}$	5.97×10^2

V. DATA ANALYSIS

A. Preselection

We target final states with two oppositely charged, different-flavour leptons, with at least two jets and moderate missing transverse momentum. Throughout the analysis, objects are ordered by transverse momentum, and the leading (subleading) jet is denoted by j_1 (j_2). The following preselection is applied to select the $e^\pm\mu^\mp + jj + \not{E}_T$ final state and suppress backgrounds at the first stage:

(i) Exactly two leptons: $N(\ell) = 2$;

(ii) Lepton transverse momenta: $p_T(\ell) \geq 10$ GeV for both leptons;

(iii) Different flavour: exactly one electron and one muon, $N(e) = 1$ and $N(\mu) = 1$;

(iv) Opposite charge: $Q(e) + Q(\mu) = 0$;

(v) Jet multiplicity: at least two jets, $N(j) \geq 2$;

(vi) Leading-jet thresholds: $p_T(j_1) \geq 30$ GeV and $p_T(j_2) \geq 30$ GeV;

(vii) Top-quark suppression: zero b-tagged jets, $N(j_b) = 0$.

This opposite-sign $e\mu$ requirement strongly reduces Drell–Yan backgrounds, while the b-jet veto suppresses top-quark processes, in line with LHC $WW \to e^\pm\nu\mu^\mp\bar{\nu}$ measurements [96].

As summarized in Table 2, the Wjj and Zjj processes have the lowest preselection efficiencies because of the tight dilepton requirement. Diboson $WZjj$ and $ZZjj$ remain as subleading reducible backgrounds primarily through lost-lepton or τ-lepton decays, while $WWjj$ and $t\bar{t}$ are effectively irreducible up to global kinematics; the latter is also strongly reduced by the b-tag veto.

Table 3 lists the event yields after sequentially applying preselection criteria (i)–(vii) for the signal with benchmark $m_a = 750$ GeV and for background processes at the HL-LHC with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3$ ab^{-1}. With the opposite-sign $e\mu$ requirement, the large inclusive Wjj and Zjj rates are greatly reduced, while $WWjj$ and $t\bar{t}$ remain the leading backgrounds in the $e\mu$ channel [96].

B. Multivariate analysis

After the preselection, we employ a multivariate analysis (MVA) to further separate signal from background. The input observables are as follows:

Table 2. Preselection efficiencies for signal at representative m_a values and for background processes at the HL-LHC with $\sqrt{s} = 14$ TeV. Numbers are efficiencies after applying criteria (i)–(vii).

m_a/GeV	170	185	200	230	260	350	500
Efficiency	0.150	0.151	0.152	0.156	0.158	0.165	0.175
m_a/GeV	750	1000	1500	2000	2500	3000	4000
Efficiency	0.186	0.194	0.201	0.204	0.205	0.201	0.191
Background	$W^+(\to \ell^+\nu)W^-(\to \ell^-\bar{\nu})jj$	$WZjj$		$ZZjj$	$W^\pm(\to \ell^\pm\nu_\ell/\bar{\nu}_\ell)jj$	$Z(\to \ell^+\ell^-)jj$	$t\bar{t}$
Efficiency	2.65×10^{-2}	4.64×10^{-4}		2.82×10^{-4}	1.33×10^{-6}	3.16×10^{-6}	1.21×10^{-3}

Table 3. Event yields after sequentially applying preselection criteria (i)–(vii) for the signal with benchmark $m_a = 750$ GeV and for background processes at the HL-LHC with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3$ ab^{-1}.

HL-LHC	initial	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Signal	1.00×10^6	5.53×10^5	5.36×10^5	2.65×10^5	2.65×10^5	2.64×10^5	2.25×10^5	1.87×10^5
$W^+(\to\ell^+\nu)W^-(\to\ell^-\bar{\nu})jj$	1.09×10^8	3.49×10^7	3.36×10^7	1.65×10^7	1.65×10^7	1.51×10^7	3.41×10^6	2.88×10^6
$WZjj$	1.03×10^9	2.71×10^7	2.53×10^7	2.51×10^6	1.36×10^6	1.31×10^6	5.77×10^5	4.78×10^5
$ZZjj$	3.10×10^8	1.40×10^7	1.37×10^7	3.14×10^5	2.23×10^5	2.18×10^5	1.15×10^5	8.75×10^4
$W^\pm(\to\ell^\pm\nu_\ell/\bar{\nu}_\ell)jj$	7.26×10^{11}	7.78×10^8	3.10×10^7	1.37×10^7	7.50×10^6	6.87×10^6	1.28×10^6	9.64×10^5
$Z(\to\ell^+\ell^-)jj$	6.62×10^{10}	1.81×10^{10}	1.78×10^{10}	2.10×10^6	1.06×10^6	9.88×10^5	2.71×10^5	2.09×10^5
$t\bar{t}$	1.79×10^9	4.35×10^7	3.78×10^7	1.85×10^7	1.84×10^7	1.84×10^7	1.53×10^7	2.16×10^6

(i) Object kinematics: energy, transverse momentum, pseudorapidity, and azimuth of the final-state objects, $E(O)$, $p_T(O)$, $\eta(O)$, $\phi(O)$ with $O = \mu, e, j_1, j_2$.

(ii) Missing transverse momentum (magnitude and azimuth): $\not{E}_T$ and $\phi(\not{E}_T)$.

(iii) Among all jet pairs, the pair with the minimal angular separation: $\Delta R(j, j')_{\min}$ and the corresponding dijet mass $m(j+j')_{\min\Delta R}$, where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}$.

(iv) Among all jet pairs, the pair with the maximal pseudorapidity separation: $\Delta\eta(j, j')_{\max}$ and the corresponding dijet mass $m(j+j')_{\max\Delta\eta}$.

(v) The invariant mass of the first two leading jets and the jet multiplicity: $m(j_1+j_2)$ and $N(j)$.

(vi) Angular separations between key objects: $\Delta R(j_1, j_2)$, $\Delta\eta(j_1, j_2)$, $\Delta R(e,\mu)$, $\Delta R(\mu, j_1)$, $\Delta R(\mu, j_2)$, $\Delta R(e, j_1)$, and $\Delta R(e, j_2)$.

(vii) Transverse masses of leptons and $\not{E}_T$: $m_{\rm T}(\mu+\not{E}_T)$, $m_{\rm T}(e+\not{E}_T)$, and $m_{\rm T}(\mu+e+\not{E}_T)$. Here, the transverse mass $m_{\rm T}$ of a visible system and the missing transverse momentum is defined as $m_{\rm T}^2 = \left(E_{\rm T}^{\rm vis} + \not{E}_T\right)^2 - \left|\vec{p}_{\rm T}^{\rm \,vis} + \vec{p}_{\rm T}^{\rm \,miss}\right|^2$, where $E_{\rm T}^{\rm vis} = \sqrt{|\vec{p}_{\rm T}^{\rm \,vis}|^2 + (m^{\rm vis})^2}$ is the transverse energy of the visible system and $\not{E}_T = |\vec{p}_{\rm T}^{\rm \,miss}|$ assumes a massless invisible system.

Observables (iv)–(vi) exploit the VBF-like topology: two forward tagging jets, large $\Delta\eta_{jj}$, and high m_{jj} with reduced central hadronic activity, while (iii) captures s-channel topologies where jets tend to be closer in phase space. This division follows standard LHC practices for VBF/di-jet tagging. Observables (vii) are related to the kinematics of the decay $a \to W^+W^- \to e^\pm\mu^\mp\nu\bar{\nu}$ in the $e^\pm\mu^\mp + \not{E}_T$ final state.

We implement the MVA using a BDT via the TMVA package integrated in ROOT, adopting the default BDT configuration for classification [97]. This setup is used to maximize background rejection while retaining high signal efficiency. Figure 3 shows the BDT responses for two benchmark masses, $m_a = 750$ GeV and $m_a = 2000$ GeV. As m_a increases, the underlying kinematics becomes more boosted and the separation between signal and background events improves, as shown by the shift of the signal (background) distribution towards higher (lower) BDT scores.

The BDT score threshold is optimized independently at each m_a by maximizing the statistical significance

$$\sigma_{\rm stat} = \sqrt{2\left[(N_s + N_b)\ln\left(1 + \frac{N_s}{N_b}\right) - N_s\right]}, \tag{13}$$

where N_s and N_b denote the expected signal and total

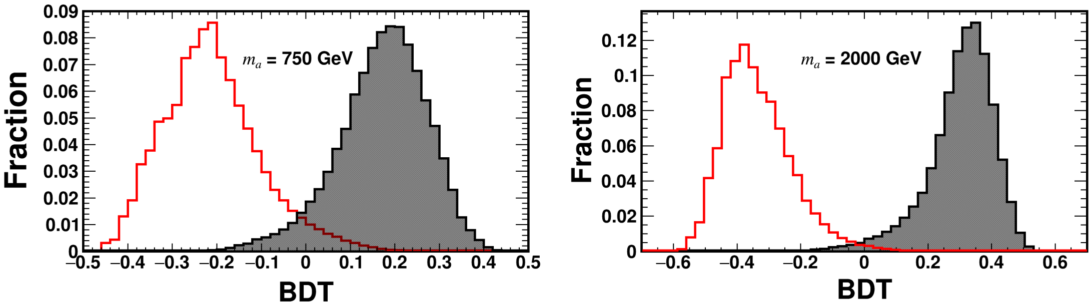

Fig. 3. (color online) BDT response distributions for signal (shaded) and total background (outline) at the HL-LHC with $\sqrt{s} = 14$ TeV, for $m_a = 750$ GeV (left) and $m_a = 2000$ GeV (right).

background yields after preselection and the BDT cut, respectively [98].

For completeness, Appendix B presents representative input-observable distributions after preselection for benchmark $m_a = 750$ GeV, while Appendix C shows BDT responses across the full m_a range for various m_a cases. Appendix D reports the preselection and BDT efficiencies for signals and backgrounds at $\sqrt{s} = 14$ TeV. In that table, a dash ("−") indicates that the surviving yield of a given background is negligible at $\mathcal{L} = 3$ ab^{-1}. Despite their large inclusive cross sections, the $W^{\pm}(\to \ell^{\pm}\nu)jj$ and $Z(\to \ell^{+}\ell^{-})jj$ background events are almost completely removed by the opposite-sign $e\mu$ requirement and multivariate selection, so they are omitted from the appendix table for brevity.

VI. RESULTS

Building on the strategy above, we present the projected discovery sensitivities at the HL-LHC with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3$ ab^{-1}. We show both the discovery reach on the coupling g_{aWW} and corresponding model-independent reach on $\sigma(pp \to jja) \times \mathrm{Br}(a \to W^+W^-)$ across $m_a \in [170, 4000]$ GeV. Statistical significances are evaluated with Eq. (13). All results assume the *photophobic* limit $g_{a\gamma\gamma} = 0$, as defined in Ref. [77]; the BDT score threshold is optimized independently at each mass point (see Sec. V.B). Unless stated otherwise, the quoted sensitivities are statistical only; the impact of systematics can be estimated in reinterpretations by introducing background-normalization nuisances in a profile-likelihood fit following Ref. [98].

Figure 4 reports the corresponding 2σ and 5σ sensitivities on the model-independent fiducial production cross section $\sigma(pp \to jja) \times \mathrm{Br}(a \to W^+W^-)$, facilitating reinterpretations in alternative theory models. Near the threshold ($m_a \sim 170$–200 GeV), the sensitivity varies quickly as the W^+W^- mode turns on and the event kinematics evolve; meanwhile, in the larger m_a region, the signal–background separation improves (cf. Sec. V.B), which offsets the falling production rate and allows competitiveness up to multi-TeV masses.

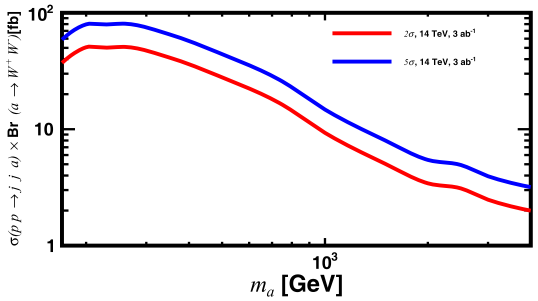

Fig. 4. (color online) Projected 2σ and 5σ discovery sensitivities on the model-independent fiducial production cross-section $\sigma(pp \to jja) \times \mathrm{Br}(a \to W^+W^-)$ versus m_a at the HL-LHC ($\sqrt{s} = 14$ TeV, $\mathcal{L} = 3$ ab^{-1}). The projection assumes $g_{a\gamma\gamma} = 0$ and uses per-mass BDT optimization. Red and blue curves correspond to 2σ and 5σ, respectively.

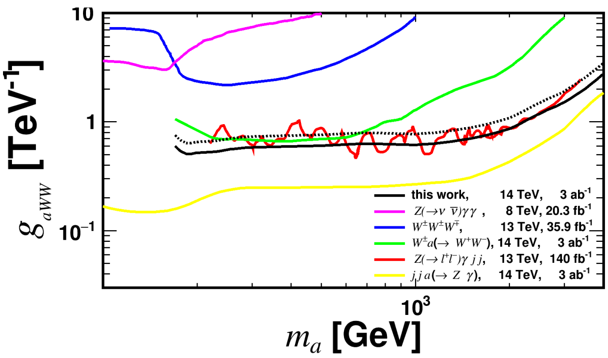

Fig. 5. (color online) Projected discovery reach on g_{aWW} for photophobic ALPs as a function of m_a (170–4000 GeV) at the HL-LHC ($\sqrt{s} = 14$ TeV, $\mathcal{L} = 3$ ab^{-1}). Solid (dashed) black curves denote 2σ (5σ) significances. The projection assumes $g_{a\gamma\gamma} = 0$ and uses per–mass-point BDT optimization; significances are statistical only. For comparison, we overlay 95% CL limits recast from triboson $pp \to Z(\to \nu\bar{\nu})\gamma\gamma$ at 8 TeV with 20.3 fb^{-1} [77] and from Run-2 analyses at 13 TeV: $pp \to W^{\pm}W^{\pm}W^{\mp}$ with 35.9 fb^{-1} and $pp \to Z(\to \ell^+\ell^-)\gamma jj$ with 140 fb^{-1} [81]. Also shown are HL-LHC (14 TeV, 3 ab^{-1}) 2σ projections for $pp \to W^{\pm}a(\to W^+W^-)$ [84] and $pp \to jja(\to Z\gamma)$ [82].

Figure 5 summarizes the 2σ and 5σ discovery reaches on g_{aWW} for heavy photophobic ALPs. At $m_a = 170$ GeV, the $2\sigma(5\sigma)$ discovery threshold is $g_{aWW} = 0.60\,(0.76)$ TeV^{-1}. As the $a \to W^+W^-$ channel opens and production cross-section of $pp \to jja(\to W^{\pm}W^{\mp})$ grows rapidly between 170 and 200 GeV (cf. Fig. 2), the required coupling decreases; over $260 \leq m_a \leq 1500$ GeV, it remains near $0.61\,(0.76)$ TeV^{-1}. At very high masses, the reach weakens due to falling PDFs and reduced acceptance, with the $2\sigma(5\sigma)$ discovery threshold $g_{aWW} = 2.75\,(3.48)$ TeV^{-1} at $m_a = 4$ TeV.

Figure 5 also displays previous constraints/projections for context. The magenta curve corresponds to the 95% C.L. limit obtained by reinterpreting the Run-1 triboson channel $pp \to Z(\to \nu\bar{\nu})\gamma\gamma$ at $\sqrt{s} = 8$ TeV with 20.3 fb^{-1} [77]. The blue and red curves show Run-2 recasts at $\sqrt{s} = 13$ TeV: $pp \to W^{\pm}W^{\pm}W^{\mp}$ with 35.9 fb^{-1} and $pp \to Z(\to \ell^+\ell^-)\gamma jj$ with 140 fb^{-1}, respectively, as reported in Ref. [81]. The green curve is an HL-LHC 2σ projection for the tri-W topology $pp \to W^{\pm}a(\to W^+W^-)$ using a same-sign dimuon plus hadronic-W selection and an MVA [84]. The yellow curve shows the HL-LHC 2σ projection for $pp \to jja(\to Z\gamma)$, which exploits a clean $Z(\to \ell^+\ell^-)+\gamma$ resonance with both the VBF-like and s-channel productions and detector-level MVA [82].

Relative to the tri-W strategy (green), our dedicated $jja(\to W^+W^-)$ analysis achieves a stronger reach (lower

g_{aWW}) over all mass ranges, with a particularly visible improvement in the high mass region ($m_a \gtrsim 700$ GeV), where the VBF-like kinematics sharpen signal–background separation. Compared with that of the $jja(\to Z\gamma)$ projection (yellow), our reach is weaker across the mass range. The reason is clear: the $jja(\to Z\gamma)$ final state benefits from a narrow mass peak, excellent photon/lepton resolution, and lower irreducible backgrounds [82]; by contrast, although the branching ratio of $a\to W^+W^-$ is larger than that of $a\to Z\gamma$, the dileptonic-WW channel involves missing energy, larger SM backgrounds (WW, $t\bar{t}$), and no fully reconstructible resonance, which intrinsically reduces sensitivity. Despite this, the $jja(\to W^+W^-)$ topology remains important and complementary. First, it *directly* probes the aWW vertex via $a\to W^+W^-$, while $a\to Z\gamma$ tests the $aZ\gamma$ interaction; a joint observation/non-observation across channels would enable a nontrivial check of the photophobic relations $g_{aZ\gamma}=t_\theta\, g_{aWW}$ and $g_{aZZ}=(1-t_\theta^2)\,g_{aWW}$ and thus of the underlying electroweak structure [77, 81]. Second, systematics and backgrounds are largely orthogonal between $jja(\to W^+W^-)$ and $jja(\to Z\gamma)$, providing robustness against analysis-specific effects and a richer basis for global fits. Third, we provide model-independent discovery thresholds for $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$, which enable straightforward reinterpretation for other neutral resonances decaying to WW (including scenarios where $aZ\gamma$ is suppressed but aWW remains sizable) [84]. Taken together, these considerations warrant a dedicated $jja(\to W^+W^-)$ search parallel to $Z\gamma$ analyses.

To gauge the impact of background systematics, we take the statistical-only reach as a baseline and estimate a background-normalization uncertainty using an Asimov profile-likelihood treatment (Appendix A.3) [98, 99]. For the benchmark $m_a=750$ GeV, Table A2 indicates that the required $\sigma\times\mathrm{Br}$ corresponding to 2σ (5σ) significance increases by $\sim 1.82\,(1.86)$ for $\delta_b=5\%$ and by $\sim 3.45\;(3.58)$ for δ_b = 10% compared with those in the statistical-only case. In realistic analyses, dominant backgrounds are typically constrained with dedicated control regions, which can reduce the effective δ_b; our $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$ thresholds therefore facilitate straightforward rescaling under alternative systematic assumptions.

Regarding EFT validity and interpretation, our g_{aWW} reach is derived in the minimal photophobic ALP EFT truncated at dimension-5. At large masses/couplings, the relevant partonic energies in the selected $jjWW$ events can approach the regime where perturbative unitarity of this truncation is lost; therefore, we regard the $J=1$ partial-wave bound from Ref. [88] (Appendix A.2) as an indication of where additional UV dynamics or unitarisation effects must enter. For clarity, we do not translate Eq. (A2) into a mass-dependent line in Fig. 5, as the relevant $\sqrt{\hat{s}}$ is process- and selection-dependent and cannot be identified with m_a; we use it only as a conservative validity guide (Appendix A.2). In this region, the mapping between an experimental rate and a single EFT parameter g_{aWW} becomes UV-sensitive (*e.g.*, via higher-dimensional operators, resonances, or form factors); therefore, we emphasize our model-agnostic sensitivity in $\sigma(pp\to jja)\times\mathrm{Br}(a\to WW)$ as the primary benchmark for reinterpretation.

VII. CONCLUSIONS

We assessed the discovery potential for heavy *photophobic* ALPs at the HL-LHC ($\sqrt{s}=14$ TeV), where the ALP–diphoton coupling vanishes, i.e., $g_{a\gamma\gamma}=0$, and interactions with electroweak gauge bosons [77]. In this limit, the ALP is produced in association with two jets, $pp\to jja$, via either s-channel electroweak-boson exchange or VBF-like topologies; for $m_a\gtrsim 2m_W$, the decay channel $a\to W^+W^-$ dominates. We target the opposite-sign, different-flavour dilepton final state $e^\pm\mu^\mp+jj+\not{E}_T$, which strongly suppresses Drell–Yan contamination and multijet backgrounds, and we adopt the HL-LHC baseline of $\sqrt{s}=14$ TeV and $\mathcal{L}=3$ ab^{-1} [83].

Events are generated with MadGraph5_aMC@NLO, showered and hadronised with PYTHIA 8.3, and passed through a Delphes 3 fast detector simulation [91, 94, 95]. The analysis applies a physics-motivated preselection for the $e\mu+jj+\not{E}_T$ topology and then a multivariate classification based on a BDT implemented in TMVA [97]. The BDT input features capture both VBF-like jet–gap kinematics (large m_{jj} and $\Delta\eta_{jj}$, reduced central activity) and s-channel configurations (small ΔR_{jj} and reconstructed $m_{jj}\sim m_V$), and the BDT score threshold is optimised independently at each m_a.

Our results provide the projected discovery reaches at 2σ and 5σ statistical significances on the ALP–W coupling g_{aWW} and, in a model-independent form, on $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$, over the mass range $m_a\in[170,4000]$ GeV. Specifically, for $260\leqslant m_a\leqslant 1500$ GeV, the sensitivity remains approximately constant at approximately $0.61\,(0.76)$ TeV^{-1} at the $2\sigma\,(5\sigma)$ level. We also report discovery sensitivities on the fiducial production cross sections $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$ across $m_a\in[170,4000]$ GeV, enabling straightforward reinterpretations in alternative theory models.

For orientation, we compare our $pp\to jja(\to W^+W^-)$ projection with recent HL-LHC studies and Run-1/2 recasts: the tri-W strategy $pp\to W^\pm a(\to W^+W^-)$ [84], $jja(\to Z\gamma)$ channel [82], and related recasts [77, 81]. In short, our reach improves on the tri-W strategy (notably at high m_a) but is weaker than the clean $Z\gamma$ resonance channel (see Fig. 5). The $WWjj$ topology remains complementary: it directly probes aWW, carries largely orthogonal systematics to $Z\gamma$, and provides model-independent thresholds in $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$; full details and a quantitative comparison are deferred to Sec. VI.

In summary, the $e\mu + jj + \not{E}_T$ topology at the HL-LHC offers competitive sensitivity to heavy photophobic ALPs, with robust performance in the multi-TeV regime owing to improved signal–background separation at large m_a. The model-independent $\sigma\times\mathrm{Br}$ projections are intended to facilitate recasts beyond the specific benchmark ALP model investigated in this study.

APPENDIX A: ROBUSTNESS AND CONSISTENCY CHECKS

A.1. Photophobic deviations

In the strict photophobic limit, one imposes $g_{a\gamma\gamma} = 0$ at the matching scale, which corresponds to the coupling relation $c_{\widetilde{B}}/c_{\widetilde{W}} = -t_\theta^2$ with $t_\theta \equiv \tan\theta$. To quantify a controlled departure, we introduce a dimensionless parameter δ in Eq. (10), which directly measures the mismatch of the cancellation in the photon direction.

Using the standard relations of Eqs. (3) and (6) between the gauge-eigenstate operators and the mass-eigenstate couplings, one finds that Eq. (10) implies

$$g_{a\gamma\gamma} = g_{aWW}\, c_\theta^2\, \delta. \tag{A1}$$

Therefore, for fixed g_{aWW}, the diphoton partial width scales as $\Gamma(a\to\gamma\gamma) \propto |g_{a\gamma\gamma}|^2 \propto \delta^2$, and the effect of $\delta \neq 0$ is to open/enhance $a\to\gamma\gamma$ and correspondingly dilute the dominant electroweak modes.

Figure A1 shows the branching ratios of the ALP into W^+W^-, ZZ, $Z\gamma$, and $\gamma\gamma$ as functions of m_a for representative values $\delta = 0.1, 0.3, 0.5$ (with other assumptions as in the main text). Above the WW threshold, $a\to W^+W^-$ remains dominant for modest departures from the strict photophobic relation, while increasing δ enhances the diphoton mode and induces a corresponding reduction in $\mathrm{Br}(a\to W^+W^-)$. In the heavy-mass region of interest for our HL-LHC study, $\delta \lesssim 0.3$ keeps $\mathrm{Br}(a\to\gamma\gamma)$ at the percent level or below, implying only mild changes to our $jja(\to W^+W^-)$ sensitivity; for larger deviations (e.g. $\delta\sim 0.5$), the $\mathrm{Br}(a\to\gamma\gamma)$ reaches $\sim 5\%$ level and diphoton searches can become increasingly competitive.

Finally, we emphasize that the above discussion concerns *tree-level* departures from the strict photophobic relation. Even if $\delta = 0$ is imposed at a high scale, radiative effects induce rare decay modes, such as $a\to\gamma\gamma$ (from W loops) and $a\to f\bar{f}, gg$ (from renormalization group running and/or loop effects). As shown explicitly in Ref. [82], in the mass region relevant to this study, these radiatively induced channels remain below the percent level and thus only negligibly affect the dominant $WW/ZZ/Z\gamma$ pattern.

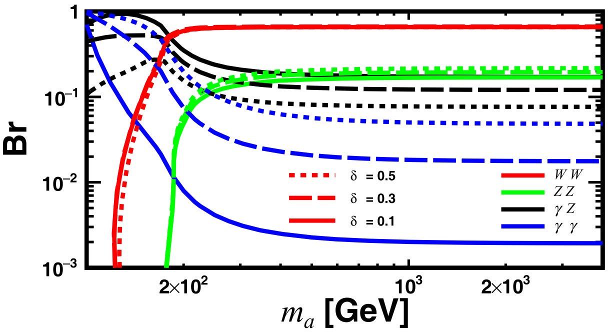

Fig. A1. (color online) Branching ratios of the ALP into W^+W^-, ZZ, γZ, and $\gamma\gamma$ as functions of the ALP mass m_a for representative deviations from the strict photophobic relation, $c_{\widetilde{B}}/c_{\widetilde{W}} = -t_\theta^2 + \delta$ (Eq. (10), with $t_\theta \equiv \tan\theta$). Solid, dashed, and dotted curves denote $\delta = 0.1$, 0.3, and 0.5, respectively. Increasing δ enhances the diphoton mode and correspondingly reduces the $WW/ZZ/Z\gamma$ branching fractions, while $a\to W^+W^-$ remains dominant above threshold in the mass range shown.

A.2. Consistency checks: EFT, unitarity, width

Partial-wave unitarity (conservative guide). Following Ref. [88], which derives partial-wave unitarity bounds in the gauge-invariant ALP EFT up to dimension-6, we use its dimension-5 truncation as a baseline guide. The strongest perturbative constraint arises from $VV\to Va$ with one external ALP (dominant $J=1$ eigenvalue), primarily bounding the $SU(2)_L$ coefficient $c_{\widetilde{W}}$. Imposing $|T_J|\leq 1$ in the high-energy limit yields

$$|c_{\widetilde{W}}| \lesssim 0.14\left(\frac{f_a}{\mathrm{TeV}}\right)\left(\frac{\mathrm{TeV}}{\sqrt{\hat{s}}}\right)^3 \;\Rightarrow\; \sqrt{\hat{s}} \lesssim 0.82\ \mathrm{TeV}\left|\frac{\mathrm{TeV}^{-1}}{g_{aWW}}\right|^{1/3}, \tag{A2}$$

using $g_{aWW} = 4c_{\widetilde{W}}/f_a$. Because $\sqrt{\hat{s}}$ is not directly observable event-by-event, we treat Eq. (A2) as a conservative guide for where the minimal dimension-5 truncation is expected to require UV unitarisation, rather than as a strict cut on our event selection.

Unitarity vs. EFT cutoff and UV sensitivity. The cutoff Λ controls the breakdown of the derivative expansion and is UV dependent. Parametrically one expects $\Lambda \lesssim 4/|g_{aWW}|$ for weakly-coupled UV completions, while NDA for strong dynamics suggests $\Lambda \lesssim 16\pi/|g_{aWW}|$ (cf. Eq. (11)). By contrast, Eq. (A2) constrains the perturbative description of specific high-energy $2\to2$ channels in the truncated EFT and can be saturated already below the NDA cutoff estimate for the minimal dimension-5 truncation [88]. In that regime, additional UV ingredients (new states, higher-dimensional operators, or form factors) are expected to unitarise the amplitude, and the numerical bound can be relaxed once such effects are included (see, e.g., the modified limits when allowing dimension-6 contributions and cancellations in Ref. [88]). For collider applications, it is thus useful to treat the unitarity curve as a conservative validity guide and, if desired, to assess robustness by restricting to events below a chosen hard scale (*e.g.*, an upper cut on a diboson invariant mass as in ALP-VBS studies [80]). Finally, because the high-energy rate in a UV-complete (unitarised) realization need not map one-to-one onto a single EFT coupling, we also provide the sensitivity in terms of $\sigma(pp\to jja)\times$ $\mathrm{Br}(a\to WW)$ for reinterpretation.

Total width estimate. Although m_a is not directly reconstructible due to the two neutrinos, the on-shell (factorised) signal modeling motivates verifying whether Γ_a/m_a remains small in the quoted sensitivity region. For dimension-5 gauge couplings, one expects $\Gamma_a\propto g_{aWW}^2 m_a^3$ at large m_a (and hence $\Gamma_a/m_a\propto g_{aWW}^2 m_a^2$). We compute the total width Γ_a in the strict photophobic benchmark with `MadGraph5_aMC@NLO` and report Γ_a/m_a along the 2σ and 5σ reach curves in Table A1. We find $\Gamma_a/m_a\lesssim 5\%$ (8%) on the 2σ (5σ) curve for $m_a\lesssim 2$ TeV, supporting the narrow-width factorization of $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$. For $m_a\gtrsim 2.5$ TeV, the required couplings lead to a rapidly growing width ($\Gamma_a/m_a\gtrsim 0.1$ and up to $O(1)$ in the extreme corner), where finite-width and/or UV-completion effects are expected to be important. In channels where the resonance mass cannot be reconstructed (as in $VV\to 2\ell 2\nu$), off-shell contributions and signal–background interference can become relevant once the width is not small [100]; hence, the broad-width corner should be interpreted as an extrapolation. Accordingly, our main results are also presented as model-independent thresholds on $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$, which can be mapped onto specific UV completions once the relevant scales and Wilson coefficients are specified.

A.3. Impact of systematics

We estimate the impact of a residual background-normalization uncertainty by treating the post-selection background yield N_b as a Gaussian-constrained nuisance parameter with width $\sigma_b\equiv\delta_b N_b$. For a single-bin counting experiment, the median (Asimov) discovery significance from the profile-likelihood-ratio test statistic is [98, 99]

$$\sigma_{\rm stat+syst}=\left[2\left((N_s+N_b)\ln\frac{(N_s+N_b)(N_b+\sigma_b^2)}{N_b^2+(N_s+N_b)\sigma_b^2}-\frac{N_b^2}{\sigma_b^2}\ln\left(1+\frac{\sigma_b^2 N_s}{N_b(N_b+\sigma_b^2)}\right)\right)\right]^{1/2}, \tag{A3}$$

which reduces for $N_s\ll N_b$ and $\sigma_b^2\ll N_b$ to the familiar approximation $\sigma_{\rm stat+syst}\simeq N_s/\sqrt{N_b+\sigma_b^2}$ [99].

As an illustration, we repeat the analysis at the benchmark point $m_a=750$ GeV and, for each assumed δ_b, choose the BDT score threshold by maximizing $\sigma_{\rm stat+syst}$ in Eq. (A3) (rather than the statistical-only significance). As a result, the optimal post-selection background yield N_b depends on δ_b. Table A2 reports the corresponding 2σ

Table A1. Total ALP width in the strict photophobic benchmark from `MadGraph5_aMC@NLO`. For each m_a, we list the g_{aWW} values on the 2σ (5σ) reach curve in Fig. 5 and the corresponding Γ_a/m_a; the last column lists Γ_a/m_a for $g_{aWW}=1$ TeV^{-1}. Along the reach curves, the width remains small up to $m_a\lesssim 2$ TeV and increases rapidly at higher masses/couplings.

m_a/GeV	2σ		5σ		$g_{aWW}=1$ TeV^{-1}
	g_{aWW}/(TeV^{-1})	Γ_a/m_a	g_{aWW}/(TeV^{-1})	Γ_a/m_a	Γ_a/m_a
170	0.601	1.49×10^{-5}	0.757	2.36×10^{-5}	4.13×10^{-5}
185	0.509	2.23×10^{-5}	0.640	3.54×10^{-5}	8.63×10^{-5}
200	0.519	4.04×10^{-5}	0.652	6.39×10^{-5}	1.50×10^{-4}
230	0.532	8.96×10^{-5}	0.669	1.42×10^{-4}	3.17×10^{-4}
260	0.562	1.65×10^{-4}	0.707	2.61×10^{-4}	5.23×10^{-4}
350	0.589	4.64×10^{-4}	0.741	7.35×10^{-4}	1.34×10^{-3}
500	0.605	1.20×10^{-3}	0.761	1.90×10^{-3}	3.28×10^{-3}
750	0.626	3.17×10^{-3}	0.788	5.01×10^{-3}	8.08×10^{-3}
1000	0.614	5.57×10^{-3}	0.772	8.84×10^{-3}	1.48×10^{-2}
1500	0.704	1.68×10^{-2}	0.886	2.67×10^{-2}	3.40×10^{-2}
2000	0.860	4.51×10^{-2}	1.08	7.17×10^{-2}	6.10×10^{-2}
2500	1.18	0.134	1.49	0.212	9.56×10^{-2}
3000	1.51	0.316	1.91	0.502	0.138
3500	2.02	0.766	2.55	1.22	0.188
4000	2.75	1.86	3.48	2.97	0.246

Table A2. Projected 2σ and 5σ sensitivities to g_{aWW} and the fiducial rate $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$ for the benchmark $m_a = 750$ GeV. For assumed relative background-yield, systematic uncertainties δ_b = 2%, 5%, 8%, 10%. $\delta_b = 0$ corresponds to the statistical-only case.

δ_b	2σ		5σ	
	g_{aWW}/(TeV^{-1})	$\sigma\times$ Br/fb	g_{aWW}/(TeV^{-1})	$\sigma\times$ Br/fb
0%	0.626	16.3	0.788	25.8
2%	0.642	17.1	0.811	27.3
5%	0.847	29.7	1.07	47.9
8%	1.04	45.2	1.33	73.7
10%	1.16	56.3	1.49	92.3

and 5σ sensitivities to g_{aWW} and to the fiducial rate $\sigma(pp\to jja)\times\mathrm{Br}(a\to W^+W^-)$ for δ_b = 0, 2, 5, 8, 10%. In realistic experimental analyses, dominant backgrounds are typically constrained with dedicated control regions, which can reduce the effective δ_b; therefore, we use statistical-only reaches as a transparent baseline in this work and quote this benchmark study only to indicate the possible degradation from percent-level systematics.

B. Representative observable distributions

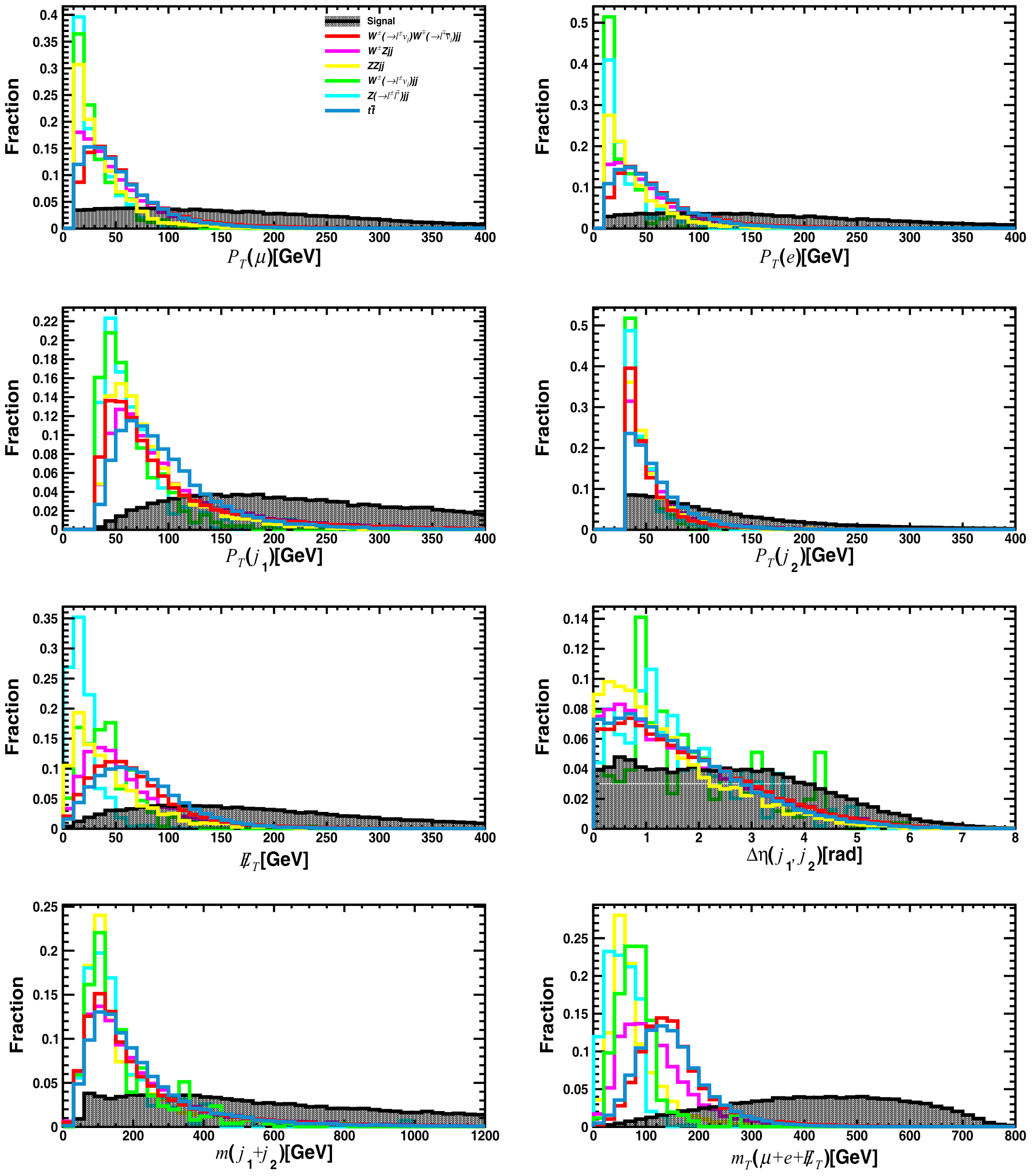

Fig. B1. (color online) Representative observable distributions for the signal (black, dashed) and the six background processes at the HL-LHC ($\sqrt{s}$ = 14 TeV), assuming m_a = 750 GeV and after the preselection of Section V.A.

C. BDT response distributions

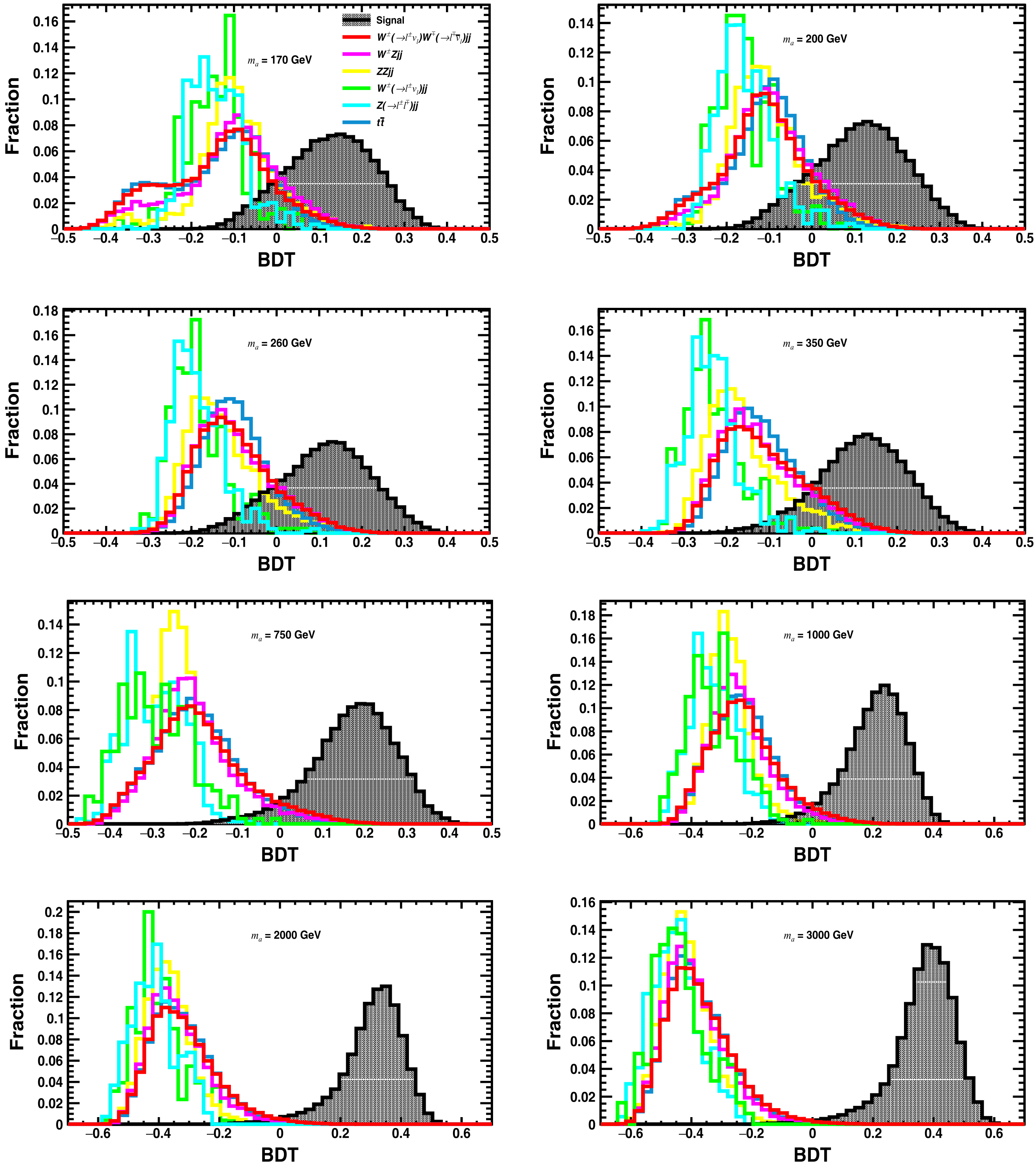

Fig. C1. (color online) BDT response distributions for the signal (shaded) and the six background processes at the HL-LHC ($\sqrt{s}$ = 14 TeV), for representative ALP masses after the preselection of Section V.A.

D. Selection efficiency table

Table D1. BDT working points and post-BDT selection efficiencies at the HL-LHC ($\sqrt{s} = 14\,\text{TeV}$) for various m_a. The second column gives the lower BDT-score threshold (cut value) used at each mass point; entries report the fraction of events surviving the BDT *after the preselection* of Section V.A. A dash ("−") indicates a negligible yield at $\mathcal{L} = 3\ \text{ab}^{-1}$.

m_a/GeV	BDT cut	Signal	$W^+(\to \ell^+\nu)W^-(\to \ell^-\bar{\nu})jj$	$WZjj$	$ZZjj$	$t\bar{t}$
170	0.165	3.62×10^{-1}	4.91×10^{-3}	4.74×10^{-3}	3.85×10^{-3}	2.77×10^{-3}
185	0.144	4.06×10^{-1}	9.07×10^{-3}	8.22×10^{-3}	6.93×10^{-3}	6.58×10^{-3}
200	0.141	4.14×10^{-1}	1.24×10^{-2}	1.04×10^{-2}	5.39×10^{-3}	8.57×10^{-3}
230	0.146	3.91×10^{-1}	1.13×10^{-2}	9.96×10^{-3}	4.62×10^{-3}	9.35×10^{-3}
260	0.144	4.02×10^{-1}	1.28×10^{-2}	8.99×10^{-3}	4.24×10^{-3}	9.52×10^{-3}
350	0.152	3.85×10^{-1}	9.02×10^{-3}	4.54×10^{-3}	7.70×10^{-4}	7.45×10^{-3}
500	0.166	4.03×10^{-1}	5.33×10^{-3}	2.22×10^{-3}	−	3.12×10^{-3}
750	0.176	5.19×10^{-1}	3.48×10^{-3}	1.16×10^{-3}	−	1.99×10^{-3}
1000	0.204	5.56×10^{-1}	1.43×10^{-3}	2.90×10^{-4}	−	3.46×10^{-4}
1500	0.231	6.92×10^{-1}	7.41×10^{-4}	1.93×10^{-4}	−	1.73×10^{-4}
2000	0.243	7.75×10^{-1}	4.48×10^{-4}	9.67×10^{-5}	−	8.66×10^{-5}
2500	0.230	8.74×10^{-1}	4.66×10^{-4}	1.93×10^{-4}	−	8.66×10^{-5}
3000	0.228	9.07×10^{-1}	2.93×10^{-4}	9.67×10^{-5}	−	8.66×10^{-5}
4000	0.235	9.43×10^{-1}	2.07×10^{-4}	9.67×10^{-5}	−	−

Note. The inclusive *Wjj* and *Zjj* backgrounds are omitted from this table: after the opposite-sign *eμ* preselection and the per-mass BDT optimisation (Section 5.2), their expected yields are negligible at $\mathcal{L} = 3\ \text{ab}^{-1}$; see also Tables 2 and 3.

ACKNOWLEDGMENTS

We thank Bin Diao, Zilong Ding, Yiheng Xiong and Ye Lu for helpful discussions.

Note added. Following standard practice in high energy physics, authors are listed in strict alphabetical order by surname. This ordering should not be interpreted as indicating any ranking of contribution, seniority, or leadership.

References

[1] C. A. Baker *et al.*, Phys. Rev. Lett. **97**, 131801 (2006), arXiv: hep-ex/0602020

[2] C. Abel *et al.*, Phys. Rev. Lett. **124**, 081803 (2020), arXiv: 2001.11966

[3] R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977)

[4] R. D. Peccei and H. R. Quinn, Phys. Rev. D **16**, 1791 (1977)

[5] S. Weinberg, Phys. Rev. Lett. **40**, 223 (1978)

[6] F. Wilczek, Phys. Rev. Lett. **40**, 279 (1978)

[7] J. E. Kim, Phys. Rev. Lett. **43**, 103 (1979)

[8] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. B **166**, 493 (1980)

[9] A. R. Zhitnitsky, Sov. J. Nucl. Phys. **31**, 260 (1980)

[10] M. Dine, W. Fischler, and M. Srednicki, Phys. Lett. B **104**, 199 (1981)

[11] J. Quevillon and C. Smith, Eur. Phys. J. C **79**, 822 (2019), arXiv: 1903.12559

[12] A. Hook, S. Kumar, Z. Liu *et al.*, Phys. Rev. Lett. **124**, 221801 (2020), arXiv: 1911.12364

[13] P. Svrcek and E. Witten, JHEP **06**, 051 (2006), arXiv: hep-th/0605206

[14] I. G. Irastorza and J. Redondo, Prog. Part. Nucl. Phys. **102**, 89 (2018), arXiv: 1801.08127

[15] L. Di Luzio, M. Giannotti, E. Nardi *et al.*, Phys. Rept. **870**, 1 (2020), arXiv: 2003.01100

[16] K. Choi, S.H. Im, and C. Sub Shin, Ann. Rev. Nucl. Part. Sci. **71**, 225 (2021), arXiv: 2012.05029

[17] M. I. Vysotsky, Y. B. Zeldovich, M. Y. Khlopov *et al.*, Pisma Zh. Eksp. Teor. Fiz. **27**, 533 (1978)

[18] Z. G. Berezhiani and M. Y. Khlopov, Z. Phys. C **49**, 73 (1991)

[19] Z. G. Berezhiani, A. S. Sakharov, and M. Y. Khlopov, Sov. J. Nucl. Phys. **55**, 1063 (1992)

[20] A. S. Sakharov and M. Y. Khlopov, Phys. Atom. Nucl. **57**, 485 (1994)

[21] A. S. Sakharov, D. D. Sokoloff, and M. Y. Khlopov, Phys. Atom. Nucl. **59**, 1005 (1996)

[22] M. Y. Khlopov, A. S. Sakharov, and D. D. Sokoloff, Nucl. Phys. B Proc. Suppl. **72**, 105 (1999)

[23] M. Y. Khlopov, S. G. Rubin, and A. S. Sakharov, Astropart. Phys. **23**, 265 (2005), arXiv: astro-ph/0401532

[24] K. Mimasu and V. Sanz, JHEP **06**, 173 (2015), arXiv: 1409.4792

[25] J. Jaeckel and M. Spannowsky, Phys. Lett. B **753**, 482 (2016), arXiv: 1509.00476

[26] G. Galanti and M. Roncadelli, Universe **8**, 253 (2022), arXiv: 2205.00940

[27] J. Liu, L. T. Wang, X. P. Wang *et al.*, Phys. Rev. D **97**, 095044 (2018), arXiv: 1712.07237

[28] M. Bauer, M. Neubert, and A. Thamm, JHEP **12**, 044 (2017), arXiv: 1708.00443

[29] M. J. Dolan, T. Ferber, C. Hearty *et al.*, JHEP **12**, 094 (2017), arXiv: 1709.00009

[30] M. Bauer, M. Heiles, M. Neubert *et al.*, Eur. Phys. J. C **79**, 74 (2019), arXiv: 1808.10323

[31] M. B. Gavela, J. M. No, V. Sanz *et al.*, Phys. Rev. Lett. **124**, 051802 (2020), arXiv: 1905.12953

[32] D. d'Enterria, *Collider constraints on axion-like particles*, in *Workshop on Feebly Interacting Particles*, 2, (2021), arXiv: 2102.08971

[33] P. Agrawal *et al.*, Eur. Phys. J. C **81**, 1015 (2021), arXiv: 2102.12143

[34] H. Y. Zhang, C. X. Yue, Y. C. Guo *et al.*, Phys. Rev. D **104**, 096008 (2021)

[35] M. Tian, Z. S. Wang, and K. Wang, *Search for long-lived axions with far detectors at future lepton colliders*, arXiv: 2201.08960

[36] F. A. Ghebretinsaea, Z. S. Wang, and K. Wang, JHEP **07**, 070 (2022), arXiv: 2203.01734

[37] A. de Giorgi, L. Merlo, and J. L. Tastet, Fortsch. Phys. **71**, 2300027 (2023), arXiv: 2212.11290

[38] H. Wang, C. X. Yue, Y. C. Guo *et al.*, J. Phys. G **49**, 115002 (2022)

[39] R. Schäfer, F. Tillinger, and S. Westhoff, Phys. Rev. D **107**, 076022 (2023), arXiv: 2202.11714

[40] K. Cheung and C. J. Ouseph, Phys. Rev. D **108**, 035003 (2023), arXiv: 2303.16514

[41] C. Antel *et al.*, Eur. Phys. J. C **83**, 1122 (2023), arXiv: 2305.01715

[42] P. S. B. Dev, J. F. Fortin, S. P. Harris *et al.*, Phys. Rev. Lett. **132**, 101003 (2024), arXiv: 2305.01002

[43] T. Biswas, JHEP **05**, 081 (2024), arXiv: 2312.05992

[44] C. X. Yue, H. Wang, X. J. Cheng *et al.*, Phys. Rev. D **107**, 115025 (2023), arXiv: 2305.19561

[45] C. X. Yue, H. Wang, and Y. Q. Wang, Phys. Lett. B **848**, 138368 (2024), arXiv: 2311.16768

[46] R. Balkin, O. Hen, W. Li *et al.*, JHEP **02**, 123 (2024), arXiv: 2310.08827

[47] F. Esser, M. Madigan, V. Sanz *et al.*, JHEP **09**, 063 (2023), arXiv: 2303.17634

[48] Y. Lu, Y. n. Mao, K. Wang *et al.*, *LAYCAST: LAYered CAvern Surface Tracker at future electron-positron colliders*, arXiv: 2406.05770

[49] Q. Qiu, Y. Gao, H. j. Tian *et al.*, JCAP **02**, 001 (2025), arXiv: 2404.18099

[50] K. Cheung, F. T. Chung, G. Cottin *et al.*, JHEP **07**, 209 (2024), arXiv: 2404.06126

[51] F. Esser, M. Madigan, A. Salas-Bernardez *et al.*, JHEP **10**, 164 (2024), arXiv: 2404.08062

[52] Z. Wang and Y. Gao, Phys. Rev. D **111**, 043042 (2025), arXiv: 2409.02468

[53] M. B. Marcos, A. de Giorgi, L. Merlo *et al.*, SciPost Phys. **18**, 084 (2025), arXiv: 2407.14970

[54] G. Yang, T. Wang, and G. L. Wang, Chin. Phys. C **49**, 013109 (2025), arXiv: 2407.05298

[55] H. Li, Q. Wu, and G. Zhang, Mod. Phys. Lett. A **39**, 2450113 (2024)

[56] S.C. İnan and A.V. Kisselev, *Analysis of axion-like particles in a top-quark pair production at the CLIC*, arXiv: 2501.15500

[57] C. X. Yue, X. Y. Li, M. S. Y. Wang *et al.*, Phys. Rev. D **111**, 075015 (2025), arXiv: 2503.03179

[58] X. Ai *et al.*, *New Physics Search at the CEPC: a General Perspective*, arXiv: 2505.24810

[59] S. Y. Wang, Y. P. Jiao, H. H. Zhang *et al.*, *Probing Long-Lived Photophobic Axion-Like Particles via Prompt Leptons and Mono-gamma at FCC-ee and CEPC*, arXiv: 2509.17718

[60] S. s. Bao, Y. Ma, Y. Wu *et al.*, JHEP **10**, 122 (2025), arXiv: 2505.10023

[61] J. Alda, M. Fuentes Zamoro, L. Merlo *et al.*, *Comprehensive ALP Searches in Meson Decays*, arXiv: 2507.19578

[62] Z. Wang and Y. Gao, *Axion Star Bosenova in Axion Miniclusters*, arXiv: 2508.14535

[63] R. Bedi, T. Gherghetta, S. Kumar *et al.*, *Heavy QCD Axions at High-Energy Muon Colliders*, arXiv: 2509.10605

[64] M. Figliolia, F. Grippa, G. Lambiase *et al.*, *Gravitational Signatures of Axion Dark Matter via Parity-Violating Interactions*, arXiv: 2509.12038

[65] C. X. Yue, X. Y. Li, S. Yang *et al.*, *Searching for long-lived axion-like particles via displaced vertices at the HL-LHC*, arXiv: 2510.24097

[66] F. Esser, A. Salas-Bernardez, M. Ubiali *et al.*, *A global analysis of ALP-mediated multiboson production at the LHC*, arXiv: 2510.24873

[67] J. Butterworth, M. Cullingworth, J. Egan *et al.*, *Probing the coupling of axions to tops and gluons with LHC measurements*, arXiv: 2508.21660

[68] S. Bhattacharya, S. Jahedi, S.K. Manna *et al.*, *Probing ALP-portal Fermionic Dark Matter at the* $e.+e.-$ *Colliders*, arXiv: 2505.00478

[69] CMS Collaboration, Phys. Lett. B **797**, 134826 (2019), arXiv: 1810.04602

[70] Belle-Ⅱ Collaboration, Phys. Rev. Lett. **125**, 161806 (2020), arXiv: 2007.13071

[71] ATLAS Collaboration, JHEP **03**, 243 (2021), arXiv: 2008.05355

[72] ATLAS Collaboration, JHEP **07**, 155 (2023), arXiv: 2211.04172

[73] BESIII Collaboration, Phys. Lett. B **838**, 137698 (2023), arXiv: 2211.12699

[74] BESIII Collaboration, Phys. Rev. D **110**, L031101 (2024)

[75] CMS Collaboration, JHEP **08**, 006 (2025), arXiv: 2412.15413

[76] C. O'HARE, *cajohare/AxionLimits: AxionLimits*, https://doi.org/10.5281/zenodo.3932429, (2024)

[77] N. Craig, A. Hook, and S. Kasko, JHEP **09**, 028 (2018), arXiv: 1805.06538

[78] N. Fonseca, E. Morgante, and G. Servant, JHEP **10**, 020 (2018), arXiv: 1805.04543

[79] A. Hook and G. Marques-Tavares, JHEP **12**, 101 (2016), arXiv: 1607.01786

[80] J. Bonilla, I. Brivio, J. Machado-Rodríguez *et al.*, JHEP **06**, 113 (2022), arXiv: 2202.03450

[81] M. Aiko and M. Endo, JHEP **06**, 194 (2024), arXiv:

2401.13323

[82] Z. Ding, Y. n. Mao, and K. Wang, JHEP **06**, 087 (2025), arXiv: 2411.08660

[83] A. Dainese, M. Mangano, A.B. Meyer *et al*., *Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC*, CERN Yellow Reports: Monographs, CERN, Geneva (2019), DOI: 10.23731/CYRM-2019-007

[84] Y. n. Mao, K. Wang, and Y. Xiong, Chin. Phys. C **49**, 083106 (2025), arXiv: 2411.14041

[85] H. Georgi, D. B. Kaplan, and L. Randall, Phys. Lett. B **169**, 73 (1986)

[86] H. Georgi, Phys. Lett. B **298**, 187 (1993), arXiv: hep-ph/9207278

[87] I. Brivio, M. B. Gavela, L. Merlo *et al*., Eur. Phys. J. C **77**, 572 (2017), arXiv: 1701.05379

[88] I. Brivio, O. J. P. Éboli, and M. C. Gonzalez-Garcia, Phys. Rev. D **104**, 035027 (2021), arXiv: 2106.05977

[89] M. Aiko and M. Endo, JHEP **05**, 147 (2023), arXiv: 2302.11377

[90] Particle Data Group Collaboration, Phys. Rev. D **110**, 030001 (2024)

[91] J. Alwall, R. Frederix, S. Frixione *et al*., JHEP **07**, 079 (2014), arXiv: 1405.0301

[92] R.D. Ball *et al*., Nucl. Phys. B **867**, 244 (2013), arXiv: 1207.1303

[93] C. Degrande, C. Duhr, B. Fuks *et al*., Comput. Phys. Commun. **183**, 1201 (2012), arXiv: 1108.2040

[94] C. Bierlich *et al*., SciPost Phys. Codeb. **2022**, 8 (2022), arXiv: 2203.11601

[95] DELPHES 3 Collaboration, JHEP **02**, 057 (2014), arXiv: 1307.6346

[96] CMS Collaboration, Phys. Lett. B **861**, 139231 (2025), arXiv: 2406.05101

[97] TMVA Collaboration, *TMVA - Toolkit for Multivariate Data Analysis*, arXiv: physics/0703039

[98] G. Cowan, K. Cranmer, E. Gross *et al*., Eur. Phys. J. C **71**, 1554 (2011), arXiv: 1007.1727

[99] G. Cowan, *Discovery sensitivity for a counting experiment with background uncertainty*, https://www.pp.rhul.ac.uk/cowan/stat/medsig/medsigNote.pdf, (2012)

[100] N. Kauer and G. Passarino, JHEP **08**, 116 (2012), arXiv: 1206.4803