Future lepton collider prospects for a ubiquitous composite pseudoscalar

Alan S. Cornell[1,*], Aldo Deandrea[2,†], Benjamin Fuks,[3,4,‡] and Lara Mason[1,2,§]

[1]*Department of Physics, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa*
[2]*Université de Lyon, F-69622 Lyon, France: Université Lyon 1, Villeurbanne CNRS/IN2P3, UMR5822, Institut de Physique des 2 Infinis de Lyon*
[3]*Laboratoire de Physique Théorique et Hautes Energies (LPTHE), UMR 7589, Sorbonne Université et CNRS, 4 place Jussieu, 75252 Paris Cedex 05, France*
[4]*Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris, France*

Composite Higgs models feature new strong dynamics leading to the description of the Higgs boson as a bound state arising from the breaking of a global (flavor) symmetry. These models generally include light states generated by the same dynamics, the detection of which may present the first observable signs of compositeness. One such state is a pseudoscalar boson resulting from the breaking of a $U(1)$ symmetry common to most composite setups, and whose hints are expected to be visible through low-mass resonance searches at present and future hadron and lepton colliders. In this work we study the phenomenology of this pseudoscalar field. We show that, for a light state, bottom quark loop effects dominantly impact the production cross section and considerably modify the decay pattern. Moreover, we make a case for targeted low-mass analyses at future lepton colliders, with an emphasis on high-luminosity machines aiming to operate at low center-of-mass energies. We present a simplified outline of a search for this light pseudoscalar at one such machine, considering electron-positron collisions at the Z-pole. We focus on a signature arising from the pseudoscalar decay into a pair of hadronic taus and a production mode association with a pair of leptons of opposite electric charges, and compare cut and count methods with machine learning methods.

I. INTRODUCTION

One of the foremost goals of the current experimental high energy physics programme is the search for new resonances. The LHC is in a long shut-down following its 13 TeV run, preparing to operate at its 14 TeV design energy as well as for a high luminosity (HL-LHC) run. At these higher energies and luminosities, efforts will be focused predominantly on the search for resonances typically heavier than the Higgs boson. However, new physics may still be concealed at lower energies. Proposals for electron-positron colliders designed to be complementary to the LHC have been put forward, including the International Linear Collider (ILC) [1], with initial energies in the range of 250–500 GeV and ranging up to 1 TeV, the

Compact Linear Collider (CLIC) [2], which could reach up to 3 TeV, as well as the Future Circular Collider (FCC-ee) [3] and Circular Electron Positron Collider (CEPC) [4], which will operate around the Z pole, the WW and $t\bar{t}$ thresholds, and in a Higgs factory mode. These machines, with lower center-of-mass (c.m.) energies than the LHC, are designed to be "factories" for resonances such as the Z, W, and Higgs bosons and the top quark, offering the possibility of precision measurements of their couplings and related Lagrangian parameters. They can also be used for targeted low-mass resonance searches, providing a window to possible beyond the Standard Model (BSM) physics at lower energy scales through large integrated luminosities. In this article we investigate the potential of finding hints for a new light pseudoscalar a emerging from a composite Higgs model, which may be accessed at both hadron and lepton colliders.

Composite Higgs models [5–7] describe the Standard Model (SM) Higgs sector in terms of high-scale fundamental gauge dynamics by postulating the existence of a new strong sector. These models implement gauge and fermionic degrees of freedom, confining at low energies [8]. In the following we describe a class of composite Higgs models featuring fermionic matter [9,10], charged under a

*acornell@uj.ac.za
†deandrea@ipnl.in2p3.fr
‡fuks@lpthe.jussieu.fr
§mason@ipnl.in2p3.fr

global symmetry G and governed by a hypercolor gauge group G_{HC}. The breaking of G then leads to the appearance of resonances that are bound states of the underlying fermions [11]. Composite Higgs models [12–15] are an attractive class of BSM theories as they provide a solution to the hierarchy problem inherent to the SM whose Higgs sector is unstable with respect to quantum corrections. If the Higgs boson is not an elementary scalar but rather a bound state of strong dynamics, these quantum corrections may only contribute up to a finite scale, hence stabilizing the Higgs field dynamics [13]. Introducing this scale of compositeness is one of a few options for a natural generation of the Higgs boson mass, and offers an explanation for the scale of electroweak symmetry breaking [15].

In addition to the breaking of the electroweak symmetry [16], occurring at a scale $v \sim 246$ GeV, the global symmetries of the fundamental fermion sector are broken at some condensation scale on the order of 1 TeV [17]. The possibilities for global and gauge symmetries within a composite Higgs model, though subject to some constraints, are fairly broad. As such, we allow in our analysis the group structures and symmetries within the theory to vary, considering a spectrum of twelve possible models that have recently been proposed as the most minimal options for a composite high-scale dynamics featuring solely fundamental fermions [9]. These models, which are strongly coupled in the IR, employ a minimal set of fields and depend on fully computable parameters. Fermion mass generation is achieved via partial compositeness [18]; however, given constraints on asymptotic freedom, this mechanism is limited to the generation of the mass of the top quark [19]. While the global symmetries of the fermions and the hypercolor groups vary across the models, similarities across composite Higgs models of this nature remain, particularly at the effective low energy scale. In particular, in addition to the existence of a QCD symmetry, there always exists a nonanomalous $U(1)$ symmetry, acting on all species of underlying fermions in the theory [17,20,21]. The first evidence of new physics may therefore arise from direct searches for the additional light state a produced in association with the Higgs boson and associated with the breaking of this extra $U(1)$. Such a pseudoscalar state moreover features interactions with the SM gauge bosons via the Wess-Zumino-Witten anomaly [22,23], which is of particular phenomenological interest concerning its production at colliders.

In order to have evaded detection up to now, a light pseudoscalar would need to be weakly coupled to the SM particles, carrying no color or electric charge. Our candidate, the ubiquitous a, is considered here to have a mass between 10 and 100 GeV. This range warrants a particular investigation given the deficiency of LHC searches in the lower end of this mass range thus far [17]. In the considered set of models, the parameters of the theory and couplings to other states are fully specified and calculable, allowing for

the construction of a general analysis targeting a new light scalar. To this aim we have formulated a new unique implementation of the pseudoscalar a within state-of-the-art modeling tools, in order to describe simultaneously a range of composite Higgs models with a variety of group structures. We have generalised previous implementations of this class of theories [17,20], where only loops of SM top quarks were contributing to the interactions of the pseudoscalar with the SM gauge bosons, and the impact of all lighter quarks was taken to be negligible. On the contrary, our work includes both top and bottom quark contributions, which are shown to provide non-negligible effects, particularly for the phenomenology of very light a bosons. This is significant given the interest in low-mass resonance searches at the LHC, including dijet [24,25], dimuon [26,27], diphoton [28–30] and ditau [31] searches in recent years. These searches yield poor constraints in the low pseudoscalar mass regions, but the associated results rely on older cross section predictions, ignoring bottom-quark loop effects. The regions of interest are therefore better covered than initially thought. In particular, ditau searches consist of one of the golden channels for the considered pseudoscalar, as the corresponding branching ratio is usually quite large [20]. The latter are generally strictly restricted to the high-mass regime, due to the presence of the large QCD background that proves to be an obstacle to low mass searches. For this reason we have instead investigated the potential of future lepton colliders where, following an overview of possibilities across a range of proposed colliders, an analysis at a low c.m. energy of the signal induced by the production of the pseudoscalar is presented.

In this manuscript we begin, in Sec. II, with a theoretical motivation, describing composite Higgs models which are built on a theory of fundamental fermions and outlining the specifics of the models to be studied. In Sec. II A we describe the dynamics of the boson a, which is the subject of this work, and in Sec. II B, its anomalous couplings to the SM gauge bosons. Section III outlines a possible low mass search strategy at lepton colliders, investigating the production of a at a variety of future experiments, before constructing an analysis for a future electron-positron collider aiming at operating at the Z-pole. We in particular outline and compare the expectation of a cut and count analysis, described in Sec. III C, and a machine learning approach based on gradient boosted trees, described in Sec. III D. We summarize our work and conclude in Sec. IV.

II. THEORETICAL MOTIVATION

This work considers a class of models [9] with a variety of hypercolor groups and several of the most minimal cosets G/H characterizing the dynamics below the confinement scale. The symmetry breaking patterns in each model are determined by the properties of the underlying fermions. For a given model with N_f Dirac fermions of the

same species, we may only have one of two possible global flavor symmetries G, namely $SU(2N_f)$ for a (pseudo)real fermion representation, or $SU(N_f) \times SU(N_f)$ for a complex fermion representation [17]. The chiral symmetry breaking may then follow one of three patterns; $SU(2N_f) \to SO(2N_f)$ for a real representation, $SU(2N_f) \to Sp(2N_f)$ for a pseudoreal one, or $SU(N_f) \times SU(N_f) \to SU(N_f)$ in the case of a complex representation [32].

In a general composite Higgs model, the mass of the SM fermions is generated through either four-fermion interactions [33] or partial compositeness [18]. The latter presents a need for fermions in (at least) two different irreducible representations of the unbroken hypercolor group G_{HC}, leading to a rich spectrum within the theory [34]. All models considered in this work thus contain two species of underlying fermions which we denote ψ and χ following the notation of Ref. [9], and which belong to different irreducible representations of G_{HC} and G. The first fermion ψ gives rise to the Higgs boson through the breaking of the associated global symmetry G into the electroweak (EW) coset H, and carries electroweak charges. The misalignment of the Higgs field then drives the usual electroweak symmetry breaking process, the mass of the Higgs boson being generated through some explicit breaking in the global sector [35]. The second species of fermion χ is responsible for partial compositeness, that proceeds through a mixing of the top quark with a composite operator of the same quantum numbers [13]. The fermion χ hence carries QCD color and hypercharge quantum numbers, and the breaking of the global symmetry then generates the QCD coset. The traditionally searched-for spin-$\frac{1}{2}$ vectorlike top-partners are therefore composed of fermions in two representations of the hypercolor group, of the form $\psi\chi\chi$ or $\psi\psi\chi$, and are labelled chimera baryons.

The presence of fermions in two irreducible representations in the theory means that there will always exist two $U(1)$ axial symmetries resulting from the full symmetry breaking pattern, one combination of these being non-anomalous with respect to the confining hypercolor group [36]. As a result, one of the numerous pseudo-Nambu-Goldstone bosons always turns out to be light, i.e., lighter than the confinement scale. This contrasts with the anomalous axial current in QCD. In a composite Higgs model we thus expect a low energy spectrum in which the Higgs boson is accompanied by exotic composite scalars and fermions, some of which are ubiquitous to all composite Higgs models and are of the $\langle\psi\psi\rangle$, $\langle\chi\chi\rangle$, $\langle\psi\chi\chi\rangle$ or $\langle\psi\psi\chi\rangle$ forms. Notably, the condensation of the underlying fermions also breaks the pervading non-anomalous $U(1)$ symmetry, leading to two massive singlet physical eigenstates that we denote a and η', where a is the lightest state. There exists some non-trivial mixing between the two corresponding gauge eigenstates, which depends on the characteristics of the underlying fermionic sector. The mixing angle α_{dec} reads, in the decoupling limit [17],

$$\csc\alpha_{dec} = -\sqrt{1 + \frac{q_\psi^2 N_\psi f_\psi^2}{q_\chi^2 N_\chi f_\chi^2}}, \qquad (2.1)$$

where q_ψ (q_χ) is the charge of the fermion ψ (χ) under the non-anomalous $U(1)$ symmetry, N_ψ (N_χ) its multiplicity, and f_ψ (f_χ) its decay constant. In the decoupling limit that we consider in this work, all other states decouple so that one solely focuses on the phenomenology of the light pseudoscalar a. The range of models according to the most minimal choices for the gauge structure, numbered M1–M12 [9], are presented in Table I in which we summarize their properties. Each model is there defined by a hypercolor symmetry group G_{HC}, given together with the corresponding irreducible representations of the two species of fermions. The table also specifies the EW and QCD cosets for each model. A great advantage to these models is the computability of all low-energy parameters, which are completely determined by the underlying fermion construction. The models are further described in detail in Refs. [17,20,36].

A. A light $U(1)$ pseudoscalar

In order to study the phenomenology of the light pseudoscalar a at colliders, we have constructed a new FeynRules [37] implementation of a simplified model describing the dynamics of the a state in a general way. This allows for the generation of UFO model files [38] that can be used further within the MadGraph5_aMC@NLO (MG5_aMC) framework [39] for the calculation of predictions at colliders. We extend the SM by a composite pseudoscalar a that exhibits small couplings to the SM fermions, gauge bosons, and the Higgs boson, and is a singlet under the SM symmetries. The pseudoscalar is modelled as having a mass of less than 100 GeV, and we consider a parametrization in which the couplings and mass are independent. In practice, we augment the SM Lagrangian with the effective Lagrangian $\mathcal{L}_a$ [20],

$$\mathcal{L}_a = \frac{1}{2}(\partial_\mu a)(\partial^\mu a) - \frac{1}{2}M_a^2 a^2 - \sum_f \frac{iC_f m_f}{f_a} a\bar{\Psi}_f\gamma^5\Psi_f$$

$$+ \frac{g_s^2\kappa_g a}{16\pi^2 f_a}G_{\mu\nu}^a\tilde{G}^{a\mu\nu} + \frac{g^2\kappa_W a}{16\pi^2 f_a}W_{\mu\nu}^i\tilde{W}^{i\mu\nu}$$

$$+ \frac{g'^2\kappa_B a}{16\pi^2 f_a}B_{\mu\nu}\tilde{B}^{\mu\nu}, \qquad (2.2)$$

where M_a is the mass of the pseudoscalar and the sum indicates a sum over all SM fermion fields Ψ_f with masses m_f. The C_f and κ_V (with $V = g$, W, B) parameters are model-specific and control the couplings of a to fermions and gauge bosons respectively. The C_f coefficient is universal for the lighter fermions whose masses arise from four-fermion interactions. Although the coupling for the

TABLE I.   Key features of the models studied in this work and described further in Refs. [17,20,36]. The first column contains the model naming convention, and the second indicates the confining hypercolor gauge group, followed by the EW and QCD cosets (third column). We then provide the irreducible representations of the fermions ψ (fourth column) and χ (fifth column) under the coset choice. The final column includes the ratio of the fermion charges under the nonanomalous $U(1)$ symmetry.

	G_{HC}	EW and QCD coset	ψ	χ	q_χ/q_ψ
M1	$SO(7)$	$\frac{SU(5)}{SO(5)} \times \frac{SU(6)}{SO(6)}$	$5 \times \mathbf{F}$	$6 \times \mathbf{Sp}$	$-5/6$
M2	$SO(9)$				$-5/12$
M3	$SO(7)$	$\frac{SU(5)}{SO(5)} \times \frac{SU(6)}{SO(6)}$	$5 \times \mathbf{Sp}$	$6 \times \mathbf{F}$	$-5/6$
M4	$SO(9)$				$-5/3$
M5	$Sp(4)$	$\frac{SU(5)}{SO(5)} \times \frac{SU(6)}{SO(6)}$	$5 \times \mathbf{A}_2$	$6 \times \mathbf{F}$	$-5/3$
M6	$SU(4)$	$\frac{SU(5)}{SO(5)} \times \frac{SU(3)^2}{SU(3)}$	$5 \times \mathbf{A}_2$	$3 \times (\mathbf{F}, \overline{\mathbf{F}})$	$-5/3$
M7	$SO(10)$		$5 \times \mathbf{F}$	$3 \times (\mathbf{Sp}, \overline{\mathbf{Sp}})$	$-5/12$
M8	$Sp(4)$	$\frac{SU(4)}{Sp(4)} \times \frac{SU(6)}{SO(6)}$	$4 \times \mathbf{F}_2$	$6 \times \mathbf{A}_2$	$-1/3$
M9	$SO(11)$		$4 \times \mathbf{Sp}$	$6 \times \mathbf{F},$	$-8/3$
M10	$SO(10)$	$\frac{SU(4)^2}{SU(4)} \times \frac{SU(6)}{SO(6)}$	$4 \times (\mathbf{Sp}, \overline{\mathbf{Sp}})$	$6 \times \mathbf{F}$	$-8/3$
M11	$SU(4)$		$4 \times (\mathbf{F}, \overline{\mathbf{F}})$	$6 \times \mathbf{A}_2$	$-2/3$
M12	$SU(5)$	$\frac{SU(4)^2}{SU(4)} \times \frac{SU(3)^2}{SU(3)}$	$4 \times (\mathbf{F}, \overline{\mathbf{F}})$	$4 \times (\mathbf{A}_2, \overline{A}_2)$	$-4/9$

TABLE II.   Relevant couplings [20] in the twelve models [36] of interest used as benchmarks in this study. We distinguish the coefficients controlling the couplings of the pseudoscalar to the SM gauge bosons (κ_V) and fermions (C_f).

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
κ_g	-7.2	-8.7	-6.3	$-11.$	-4.9	-4.9	-8.7	-1.6	-10	-9.4	-3.3	-4.1
κ_W	7.6	$12.$	8.7	$12.$	3.6	4.4	$13.$	1.9	5.6	5.6	3.3	4.6
κ_B	2.8	5.9	-8.2	$-17.$	$.40$	1.1	7.3	-2.3	$-22.$	$-19.$	-5.5	-6.3
C_f	2.2	2.6	2.2	1.5	1.5	1.5	2.6	1.9	$.70$	$.70$	1.7	1.8

top, C_t, may take several values depending on the representation of the top partner in the partial compositeness mechanism [20], we have followed the convention of Ref. [20] and taken $C_f = C_t$. For the 12 benchmark scenarios under consideration, we present the coefficients dictating the coupling of the pseudoscalar in Table II.

In our notation, g_s, g and g' refer to the strong, weak and hypercharge coupling constants, and $G_{\mu\nu}$, $W_{\mu\nu}$ and $B_{\mu\nu}$ ($\tilde{G}_{\mu\nu}$, $\tilde{W}_{\mu\nu}$ and $\tilde{B}_{\mu\nu}$) stand for the associated (dual) field strength tensors. The decay constant f_a of the pseudoscalar a, that drives the strength of the pseudoscalar couplings to the SM particles, is defined as

$$f_a = \sqrt{\frac{q_\psi^2 N_\psi f_\psi^2 + q_\chi^2 N_\chi f_\chi^2}{q_\psi^2 + q_\chi^2}}, \tag{2.3}$$

which we set to 1 TeV in this analysis, as previous studies [20] show that the lower bound on f_a is always below 1 TeV for the models under consideration. This description

being effective, we recall that we can only rely upon it for energies or momenta below a cut-off scale $\Lambda \sim 4\pi f_a$.

B. Pseudoscalar couplings to gauge bosons

Leading-order couplings of the form aVV', where V, V' stand for gauge bosons which may or may not be different, proceed via the Wess-Zumino-Witten anomaly and are depicted through an effective vertex in the Lagrangian of Eq. (2.2). However, an SM component where the vertex is constructed from loops of SM fermions, an example of which is shown in Fig. 1 (left), is generally significant and should be included. In order to access each gauge-pseudoscalar vertex, we rewrite the gauge-boson interaction part $\mathcal{L}_V$ of the Lagrangian of Eq. (2.2) in terms of the physical gauge bosons,

$$\mathcal{L}_V = \frac{a}{16\pi^2 f_a} \left(g_s^2 \kappa_g G_{\mu\nu} \tilde{G}^{\mu\nu} + g^2 \kappa_W W_{\mu\nu}^+ \tilde{W}^{-\mu\nu} + e^2 \kappa_{\gamma\gamma} F_{\mu\nu} \tilde{F}^{\mu\nu} \right.$$
$$\left. + \frac{e^2 \kappa_{ZZ}}{s_W^2 c_W^2} Z_{\mu\nu} \tilde{Z}^{\mu\nu} + \frac{2e^2 \kappa_{Z\gamma}}{s_W c_W} F_{\mu\nu} \tilde{Z}^{\mu\nu} \right), \tag{2.4}$$
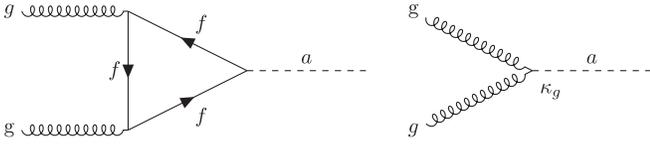
FIG. 1. The SM (left) and BSM (right) components of the pseudoscalar coupling to gluons, relevant for its production at hadron colliders. The BSM vertex consists of an effective Wess-Zumino-Witten structure, as shown in Eq. (2.2).
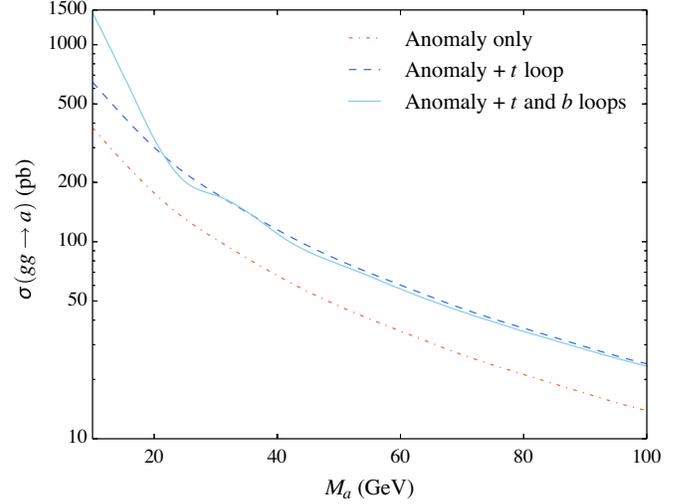

FIG. 2. Gluon fusion production cross section for the pseudoscalar a with an effective gluon-pseudoscalar coupling of 1 for $M_a \in [10, 100]$ GeV. We distinguish the contributions arising solely from the anomalous interactions (dash-dotted red), the sum of the latter with the top quark loop contributions (dashed blue) and the full predictions including the bottom loop contributions as well (solid teal).

where c_W and s_W denote the cosine and sine of the EW mixing angle, and e is the electromagnetic coupling constant. While the anomalous couplings read

$$\kappa_{\gamma\gamma} = \kappa_W + \kappa_B, \qquad \kappa_{Z\gamma} = c_W^2 \kappa_W - s_W^2 \kappa_B \quad \text{and}$$

$$\kappa_{ZZ} = c_W^4 \kappa_W + s_W^4 \kappa_B, \tag{2.5}$$

where contributions originating from the SM fermion loops should additionally be included for all existing interactions ($gga, \gamma\gamma a, ZZa, W^+W^- a$ and $Z\gamma a$). It is, however, expected that the role of the leptons and of the five light flavors of quarks is negligible, their couplings to the pseudoscalar being suppressed by the smallness of their masses, as are the contributions from the electroweak bosons that are suppressed by the heavy propagators running into the loops. In the following, we however stress the importance of the bottom quark, whose contributions are in fact not so negligible.

As an example, we focus on the gga vertex and calculate the partonic gluon-fusion production cross section of a pseudoscalar,

$$\sigma_0 = \frac{1}{256\pi f_a^2} \frac{g_s^4}{16\pi^2} \left| \kappa_g + \sum_f C_f^2 A(\tau_f) \right|^2 \quad \text{with} \quad \tau_f = \frac{4m_f^2}{M_a^2}. \tag{2.6}$$

Such an expression includes the anomaly contribution, as well as the sum over the contributions from each fermion species. The function $A(\tau)$ is defined, for a given fermion, by

$$A(\tau) = \tau \begin{cases} -\frac{1}{4}\left[\log\frac{1+\sqrt{1-\tau}}{1-\sqrt{1-\tau}} - i\pi\right]^2 & \text{if } \tau < 1, \\ \arcsin^2(\frac{1}{\sqrt{\tau}}) & \text{if } \tau \geq 1, \end{cases} \tag{2.7}$$

which results from the three-point scalar function of the quark loop propagator. In the case of top quarks, $\tau_t \geq 1$ and $A(\tau_t)$ is approximately constant (≈ 1) throughout the pseudoscalar mass range. This thus leads to an approximately constant increase in the gluon fusion production cross section relative to the pure anomalous component, as illustrated in Fig. 2. In the latter, we convolute the expression of Eq. (2.6) with the leading order set of NNPDF 2.3

parton densities NNPDF23_lo_as_0130_qed [40] with $M_a \in [10, 100]$ GeV. In this analysis we ignore lower masses, where a light pseudoscalar is subject to strong experimental bounds [20]. The behavior is quite different in the case of b quarks, where logarithmic effects produce an oscillation in the contribution to the cross section (see Fig. 2). At low masses of a, the bottom quark contribution substantially increases the cross section. For higher masses, there is a small destructive interference between the top and bottom contributions, leading to a slight decrease in cross section. This shape of the bottom quark contribution arises due to the form of the three point scalar function $A(\tau)$ for small τ, where the interplay between the real and imaginary part of the loop-integral leads to the observed undulation in the cross section behavior. For a detailed investigation into the corresponding modifications to the pseudoscalar decay pattern due to the inclusion of the b-quark loop, we refer the reader to Ref. [41].

We do not include the contributions of any quarks lighter than the b to run in the fermion loop, as the considered mass range of the pseudoscalar leads to lighter quarks having a negligible impact on the cross section. This can be supported by the observation that the c quark, the next heaviest quark with $M_c = 1.275$ GeV, would impact the cross section at most by the amount that the b quark contributes at $M_a \approx 33$ GeV, due to the form of $A(\tau_c)$ for the small τ_c value. Given that the bottom quark contribution to the cross section is already negligible at this point, all other quark contributions can therefore be ignored. As a consequence, we will consider throughout this work all bottom and top quark loop-contributions to the aVV

couplings for all SM gauge bosons V, additionally to the tree-level BSM contributions.

III. A LOW MASS PSEUDOSCALAR AT LEPTON COLLIDERS

In this section we estimate the prospects of future e^+e^- collider searches for this additional light scalar as an alternative to searches at hadronic colliders. We are interested, in particular, in designing an analysis that addresses the parameter space region in which $M_a \in [10, 60]$ GeV, where constraints on possible light scalars are particularly weak [20]. This mass window has indeed been left relatively open due to few direct searches performed thus far, the dominant constraint originating from searches for novel SM Higgs decay modes ($h \to aa$).

We begin this section with a study of pseudoscalar production at lepton colliders, which differs from that at hadron colliders as the pseudoscalar is produced predominantly in association with other states. We consider a variety of future e^+e^- colliders operating at different c.m. energies and, in the case of linear colliders, different polarization options. We then focus on a circular electron-positron collider aiming to operate at the Z-pole and investigate a possible search channel targeting pseudoscalar decays into a pair of tau leptons.

A. Pseudoscalar production at lepton colliders

We consider the production of the pseudoscalar in association with a virtual photon or with a (virtual or real) Z-boson, that we generically denote as V, and that 'decays' into any pair of fermions. This mainly proceeds via the Feynman diagrams shown in the first row of Fig. 3, where one distinguishes a tree-level contribution in which an e^+e^- pair annihilates into a Va system through the Wess-Zumino-Witten anomalous $\kappa_{\gamma Z}$ coupling (central diagram), a loop-induced contribution (left diagram) where SM top

and bottom quarks are running in the loop (see Sec. II B), and a t-channel contribution (right diagram). In the case where the V-boson leads to an e^+e^- final state, extra nonresonant diagrams additionally contribute (second row of Fig. 3). We identify two potentially appealing signatures that differ by the considered final state: the production of the pseudoscalar in association either with a pair of opposite-sign first or second generation leptons $\ell = e$, μ, or with a pair of light jets j (i.e., not originating from the fragmentation of a b-quark),

$$e^+e^- \to \ell^+\ell^- a, \qquad e^+e^- \to jja. \qquad (3.1)$$

Across all models, the branching ratios of the pseudoscalar a into $\tau^+\tau^-$ and $b\bar{b}$ pairs are the highest, as the coupling of the pseudoscalar to fermions is proportional to the corresponding fermion mass. For further insights into the branching modes of the pseudoscalar under the influence of the bottom-quark loop effects, the reader is referred to Ref. [41], where the rankings of the branching ratios are demonstrated. Across the models considered, the $b\bar{b}$ and $\tau^+\tau^-$ branching ratios are consistently the most promising avenues. We therefore choose to design an analysis dedicated to probing those decay modes, as they are not only the most abundant, but also feature final-state objects not too difficult to reconstruct, and offer several handles to extract a signal from the background, as will be shown below.

The future colliders currently under study include both linear (ILC [1] and CLIC [2]) and circular (FCC-ee [3] and CEPC [4]) possibilities. Electron-positron colliders may provide a promising avenue through which to search for a light pseudoscalar. Many future lepton colliders indeed offer low c.m. energies associated with high luminosities, which may allow for the detection of weakly interacting light particles. While linear colliders offer the benefit of beam polarization, where some processes such as W-boson fusion Higgs production (which depend on the chirality of
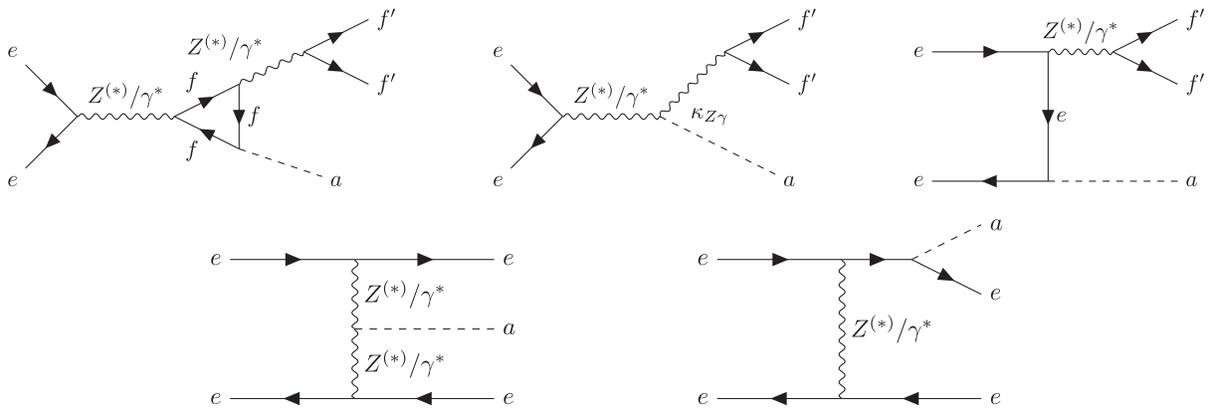

FIG. 3.   Representative Feynman diagrams depicting the production of a pseudoscalar in association with a virtual photon, or with a (virtual or real) Z-boson. We include the latter decay into a fermion-antifermion pair (first line) and additionally consider extra nonresonant diagrams relevant for a V decays into an e^+e^- pair (second line). The first two diagrams denote the SM and BSM components of the same process.
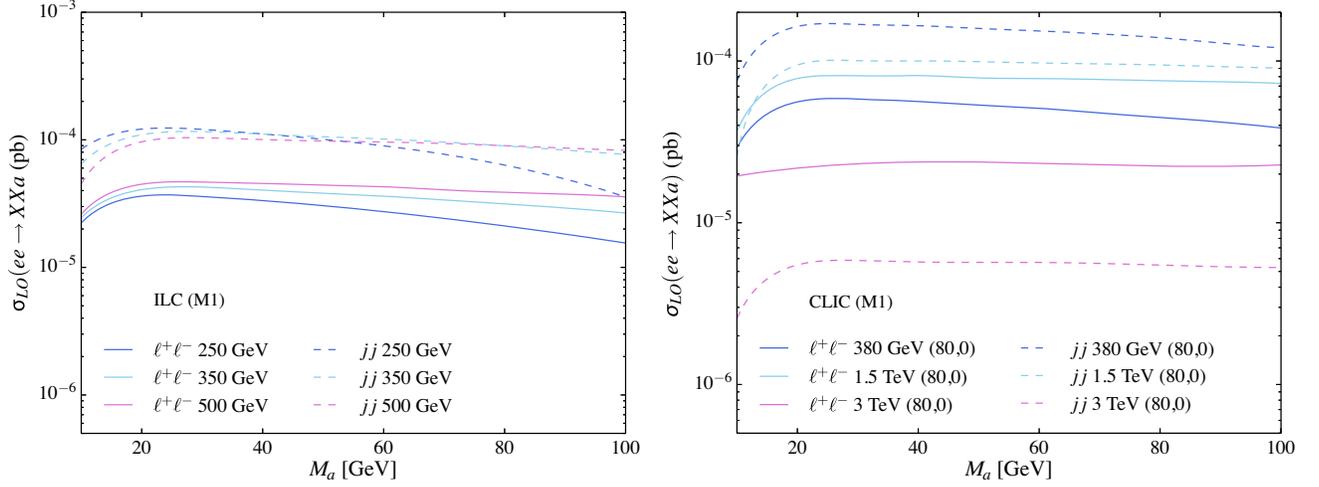
FIG. 4.   Total cross section associated with the production of the pseudoscalar a in association with leptons (solid lines) or jets (dashed lines) at the ILC and CLIC, for various c.m. energies and including beam polarizations where relevant.

the colliding leptons) may be amplified, circular colliders allow for the accumulation of higher luminosities, useful when searching for particles that couple timidly or are not copiously produced. They moreover offer the possibility of hadron collider upgrades.

We first proceed with the calculation of the production cross section associated with all processes of interest, for each of the aforementioned future lepton collider options and the varied c.m. energy choices that have been proposed in the literature. Using MG5_aMC, we hence present in Fig. 4 total cross sections for the two processes of Eq. (3.1) for the ILC and CLIC linear colliders. We rely, as a benchmark, on the model M1 and show leading-order (LO) predictions for the production of the pseudoscalar a in association with leptons (solid lines) and jets (dashed lines). We also considered a variety of beam polarizations relevant for the ILC collider, however, there was no appreciable change

in behavior for different polarizations. In Fig. 5 we focus instead on the CEPC and FCC-ee circular colliders, our predictions for pseudoscalar production in association with leptons being multiplied by a factor of 15 for legibility. Our results include basic selections on the final-state leptons and jets, their transverse momentum p_T and pseudorapidity η being enforced to satisfy

$$p_T(j) > 20 \text{ GeV}, \qquad |\eta(j)| < 5;$$
$$p_T(\ell) > 5 \text{ GeV}, \qquad |\eta(\ell)| < 2.5. \qquad (3.2)$$

Moreover, jets and leptons are required to be well separated in the transverse plane, by an angular distance ΔR of at least 0.4,

$$\Delta R(\ell, \ell') > 0.4, \qquad \Delta R(j, j') > 0.4. \qquad (3.3)$$
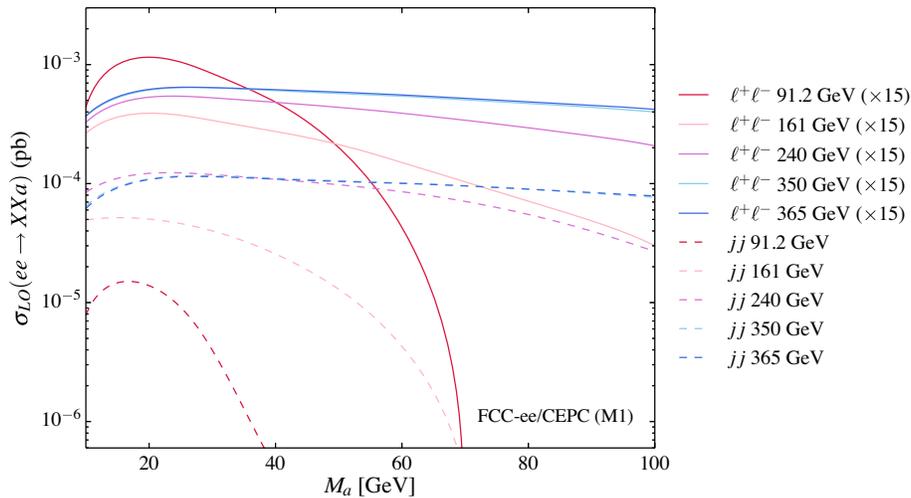

FIG. 5.   Same as in Fig. 4, but for the FCC-ee and CEPC circular colliders.

From these figures we are able to gain an understanding of the potential pseudoscalar abundance at future lepton colliders.

Though linear colliders enjoy a relatively constant production cross section of about 0.01–0.1 fb throughout the entire probed pseudoscalar mass range ($M_a < 100$ GeV), circular colliders, operating at lower c.m. energies, are subject to a fall-off of production cross section at higher masses as the available phase space decreases. However, for the parameter space region in which we are interested (featuring $M_a < 60$ GeV), the production rates lie in the same ballpark regardless of the collider under consideration. As linear colliders are subject to much lower expected integrated luminosities, we focus, in the following analysis outline, on a circular collider. Their larger expected luminosities may prove crucial in a search for weakly coupled particles, while operating at relatively low c.m. energies is useful in reducing otherwise significant backgrounds, such as originating from $t\bar{t}$ and di-boson processes.

B. A case study at a circular collider operating at the Z-pole

We propose an analysis at a future high luminosity lepton collider aiming at operating at the Z-pole and an integrated luminosity of 150 ab^{-1}, which corresponds to the FCC-ee expectation. While the signal production cross section is subject to a steep decline for higher masses of the pseudoscalar a, the mass range of interest ($M_a \in [10, 60]$ GeV) is still relatively well covered, as shown in the previous subsection (see Fig. 5).

As a choice for the detector parametrization, we consider the IDEA detector concept of the FCC-ee project. This detector, as any detector project at any future circular electron-positron machine, is designed using recent technological advances to take advantage of the exceptionally large data samples due to be delivered by the forecast integrated luminosities. IDEA is planned to be constructed with a short drift wire chamber and calorimeter, and includes a low mass superconducting solenoid coil. The drift chamber will allow for high precision momentum measurements and good tracking capabilities, as well as excellent particle identification performance through cluster counting when combined with the dual readout calorimeter [3]. In particular, this aims at achieving a much improved impact parameter resolution over that of LEP, as well as a better momentum resolution and electromagnetic calorimeter resolution, and a finer electromagnetic calorimeter transverse granularity [42]. Precise measurements of charged objects properties at lower energies are therefore clearly achievable.

In building our analysis, we will moreover only consider the channel in which the pseudoscalar a is produced in association with a pair of opposite-charge leptons. This allows us to avoid the difficulty in dealing with the multijet
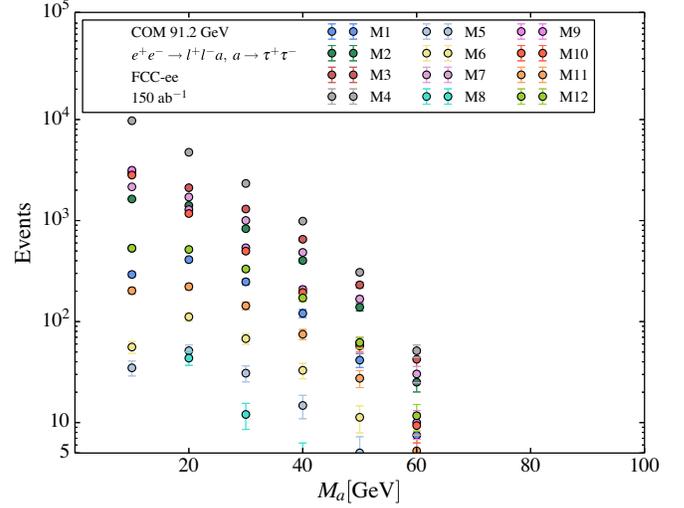

FIG. 6. Expected number of $\ell\ell\tau\tau$ signal events in 150 ab^{-1} of e^+e^- collisions at the Z-pole, for the 12 benchmark models under consideration in this work.

background still reasonably present at lepton colliders. Moreover, we will focus on the case where the pseudoscalar decays into a pair of hadronic tau leptons for two reasons. Hadronic tau decays account for approximately 2/3 of all tau decays, and future e^+e^- collider detectors are expected to have excellent handles on the associated decay vertices. This hence allows for a very efficient tau reconstruction. In addition, events with electrons or muons in the final state are expected to be reconstructed with a very good resolution [43]. Basing our estimations on previous detectors, we hence expect a typical systematic uncertainty on lepton identification of around 1%, together with a systematic uncertainty on hadronic tau identification of around 2–5% for taus with a transverse momentum $p_T > 20$ GeV, and up to 15% otherwise [44].

In Fig. 6 we display the number of signal events expected from the production of the pseudoscalar a in association with a pair of oppositely charged leptons, with a subsequent pseudoscalar decay into a pair of hadronic tau leptons,

$$e^+e^- \to \ell^+\ell^- a \to \ell^+\ell^- \tau^+\tau^-. \qquad (3.4)$$

We consider the entire mass range of interest and include statistical error bars. This illustrates the motivation to expect a significant number of new physics events in electron-positron collisions at the Z-pole. It moreover attests that the sensitivity of the machine will depend on the exact details of the model, as there is a significant range in the expected number of new physics events across the considered models.

We now design our analysis from Monte Carlo simulations of both the signal and background processes, using MG5_aMC employed in conjunction with PYTHIA8 [45] to describe parton showering (which includes both QED and QCD initial state radiation effects) and hadronization.

The IDEA detector response has been simulated by relying on the DELPHES3 [46] software package, that makes use of the anti-k_T algorithm [47] as implemented in FastJet3 [48] for event reconstruction. Both these last codes are driven through their interface with the MadAnalysis 5 platform [49,50], that we also use to carry on our phenomenological analysis. We begin with a cut-and-count analysis, aiming at unravelling the signal from the overwhelming backgrounds. However, as a consequence of the low statistical significance, we then employ a novel machine learning algorithm based on boosted decision trees in an attempt to improve the significance, using the XGBOOST toolkit [51].

In identifying backgrounds to the signal process, we consider both processes which lead to a true $\ell\ell\tau\tau$ final state, and those containing fakes. Background events of the first category feature prompt taus that are most likely to arise from the production of an $\ell^+\ell^-$ pair in association with a pair of opposite-sign tau leptons through one or two (virtual) Z bosons and photons. Given the relatively low c.m. energy of 91.2 GeV, potential background events originating from two virtual weak bosons or top quark decays are not expected to contribute much. Background events of the second category arise from jets faking hadronic taus. They appear in processes such as vector boson production in association with a pair of jets, where the boson then decays to a pair of leptons, as well as in processes where the intermediate boson is virtual or the mediation occurs through virtual photons,

$$e^+e^- \to jj\ell^+\ell^-. \tag{3.5}$$

In the following, we refer to those background as $Z+$jets, although contributions from virtual photons and their interferences are accounted for as well. The corresponding cross section is 6.65×10^{-4} pb, that should then be multiplied by the appropriate fake rate factor. The IDEA detector parametrization shipped with DELPHES includes a 0.1% misidentification rate for jets faking hadronic taus and a tau identification efficiency of 60% [52]. This tau identification efficiency is based on efficiencies at existing colliders, and is expected to be a conservative estimate. We therefore expect a fake contribution of the order of 10^{-10} pb, which can thus safely be neglected.

C. A cut-and-count analysis

Our selection closely follows the pattern of the final state under consideration. We require events to contain at least two leptons ($N_\ell \geq 2$), each with a minimum p_T of 10 GeV to ensure good reconstruction, and at least two hadronic taus ($N_\tau \geq 2$), each with $p_T > 5$ GeV. We moreover enforce a minimum invariant mass $m_{\ell\ell} > 12$ GeV on the lepton pair produced in association with the ditau system, which is necessary to eliminate nonprompt leptons and avoid low mass hadronic resonances. Similarly, the invariant mass of the tau pair $M_{\tau\tau}$ is constrained to be at

least 10 GeV. Given that the hadronic tau decay mode of the pseudoscalar leads to neutrinos which carry away momentum, the ditau invariant mass spectrum is expected to be soft and peak below the pseudoscalar mass M_a. It is therefore useful to maintain low momentum thresholds for the tau pair where possible. To summarize, our preselection imposes that

$$\begin{aligned} N_\ell &\geq 2 \quad \text{with} \quad p_T(\ell) > 10 \text{ GeV;} \\ N_\tau &\geq 2 \quad \text{with} \quad p_T(\tau) > 5 \text{ GeV;} \\ M_{\ell\ell} &> 12 \text{ GeV;} \qquad M_{\tau\tau} > 10 \text{ GeV.} \end{aligned} \tag{3.6}$$

After this preselection, we expect about 50,000 background events for signal event counts ranging from below 1 ($M_a \lesssim 10$ GeV) or a few ($M_a \gtrsim 50$ GeV) to 10–40 (10 GeV $< M_a <$ 50 GeV).

In order to reject the background while keeping a large signal efficiency, we investigate first the properties of the two final-state leptons. As in the case of the signal they are produced together with a low mass resonance, we expect the presence of potentially discriminating features in various kinematic distributions, the exact details behind those features being related to the resonance mass. We present, in Fig. 7, the angular separation between the two leptons $\Delta R(\ell^+, \ell^-)$ (left) as well as the dilepton invariant mass distribution $M_{\ell\ell}$ (right). We consider both the $Z+$ jets background (red dashed) as well as five signal hypotheses from different models and pseudoscalar masses below 60 GeV (i.e., our pseudoscalar mass range of interest). More precisely, we have chosen a selection of five models (M2, M4, M7, M10 and M12) exhibiting a variety of hypercolor group and coset structures (see Table I). This allows us to largely explore the possibilities of the considered class of composite scenarios. The results depicted in the figures demonstrate that the $\Delta R(\ell^+, \ell^-)$ and $M_{\ell\ell}$ spectra tend to peak at higher values for the background than for the illustrative signal hypotheses. This suggests two interesting cuts to isolate the pseudoscalar signal,

$$\Delta R(\ell^+, \ell^-) < 3; \qquad M_{\ell\ell} < 40 \text{ GeV.} \tag{3.7}$$

In addition, we also make use of the properties of the ditau system to extract our composite signal from the background. In the signal case, the pair of hadronic tau leptons is issued from the decay of a resonance, so that its properties are expected to be largely different from the background case. In particular, the invariant mass of the ditau system is expected to peak at a value just below the mass of the pseudoscalar, as a result of the presence of the neutrinos originating from the tau decays and carrying away some momentum [53]. The resulting distribution is shown in the left panel of Fig. 8. As expected, the invariant mass of the ditau system is shifted relative to
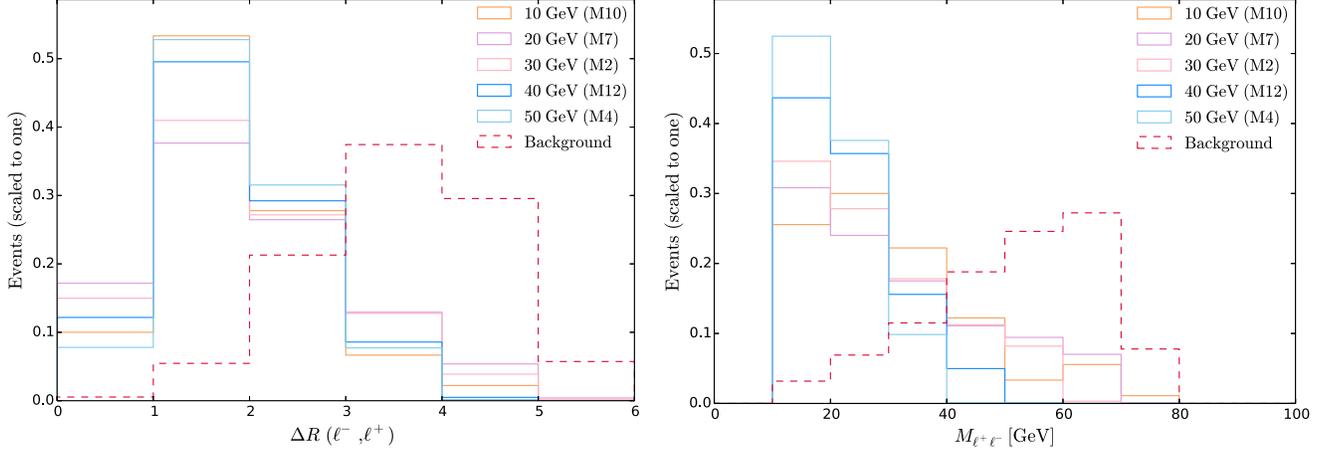
FIG. 7. Properties of the dilepton system for signal and background events originating from electron-positron collisions at a c.m. energy of 91.2 GeV. We consider various representative signal hypotheses, and focus on the angular distance in the transverse plane between the two leptons (left), and their invariant-mass spectrum (right).

the mass of the pseudoscalar, and peaks just below the true value of the latter. The background distribution is, however, not so well differentiated from the signal one, as the ditau system features predominantly a low invariant mass, as driven by the selection cuts of Eq. (3.7). We nevertheless define five signal regions, each of them being dedicated to one specific pseudoscalar mass hypothesis, and respectively impose

$$M_{\tau\tau} < 10, 20, 30, 40, 50 \text{ GeV}. \tag{3.8}$$

This allows us to eliminate some background in the heavier pseudoscalar cases. The minimum requirement on $M_{\tau\tau}$ of Eq. (3.6) that protects us from the contamination of QCD resonances makes us, however, unable to get further handles on the signal for pseudoscalar masses of 10 GeV or smaller. Similarly, we investigate the potential of the angular separation between the two taus (right panel of Fig. 8). Although shape differences are visible, they do not

allow for a clear separation of the signal and the background. Any related cut will therefore be omitted from our analysis.

In Table III we present the expected sensitivity of our cut-and-count analysis in terms of standard deviations defined by an $S/\sqrt{S+B}$ figure of merit, S and B respectively representing the number of selected signal and background events. We find that, given the relative rareness of the signal events among an abundance of background, it is difficult to obtain any hope to observe even a 1σ deviation from the background-only hypothesis across the entire mass range considered. It is, however, possible that this could be ameliorated by designing more appropriate and dedicated variables like the missing mass. As reflected in the M_a-dependence of the cross section shown in Fig. 5, the significance is maximized at $M_a = 20$ GeV. For larger pseudoscalar masses, the steep fall-off of the cross section indeed reduces S to too large a level.
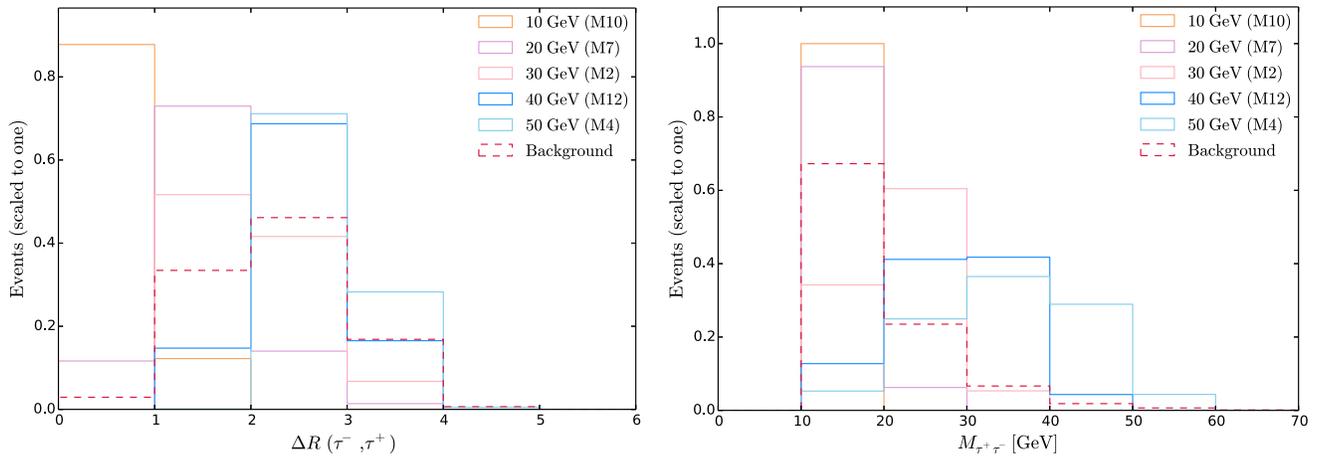

FIG. 8. Same as in Fig. 7 but for the ditau system and after the cuts of Eq. (3.7).

TABLE III.   Sensitivity of our cut-and-count analysis expressed as the significance $S/\sqrt{S+B}$ for a selection of considered composite scenarios and 150 ab^{-1} of electron-positron collisions at the Z-pole.

Model	$M_a = 10$ GeV	$M_a = 20$ GeV	$M_a = 30$ GeV	$M_a = 40$ GeV	$M_a = 50$ GeV
M2	0.0015	0.13	0.090	0.049	0.020
M4	0.0013	0.42	0.26	0.12	0.040
M7	0.0024	0.14	0.11	0.061	0.023
M10	0.0042	0.11	0.055	0.023	0.0078
M12	0.00061	0.047	0.035	0.021	0.017

D. A machine-learning-based analysis

In order to improve the figure of merit of our analysis, we move on with considering a machine learning algorithm. We rely on the XGBOOST toolkit [51] that allows for utilizing gradient boosted tree methods [54] while offering fast training speed coupled with a good accuracy [55].

In general, a machine learning algorithm employing a tree ensemble uses a series of additive optimizations computed from a given set of variables to predict an output, i.e., in our case the classification of an event as a signal or a background event. At each stage of the training process, gradient boosting modifies the existing constraints in order to correct the classification errors made by the current best set of optimizations, continuing until no further improvement can be made in considering the residuals and errors of the prior stages. The XGBOOST toolkit includes a novel algorithm geared toward the handling of sparse data, which is useful in our case as both signal and background events may not fully populate the event space.

The performance of the algorithm for a given set of optimizations can be evaluated by a quantity denoted as the area under the curve (auc). This corresponds to the integral of the receiver operating characteristic (roc) depicting the dependence of the signal purity of the events selected by the algorithm, $S/(S+B)$, on the signal selection efficiency S/S_0, where S_0 stands for the total number of signal events provided to the algorithm. The auc metric hence represents the degree of separability between background and signal. In addition, we use the approximate median discovery significance (ams) to estimate the sensitivity of the analysis to our signal. It is defined by [56]

$$ams = \sqrt{2\left((S+B)\ln\left(1+\frac{S}{B}\right) - S\right)}, \quad (3.9)$$

where S and B can also be seen as the true and false positives respectively. While the ams provides the discovery potential of the analysis, its usage as an evaluation metric and learning objective is unstable and may lead to overfitting. The performance of the algorithm was therefore optimised using the auc quantity, following which the corresponding ams was calculated.

After applying the preselection of Eq. (3.6), we derive a set of uncorrelated kinematic variables to be used as input to our machine learning algorithm. They consist of a combination of primary variables (the tau and lepton transverse momenta, pseudorapidities and azimuthal angles) and derived variables (the ditau invariant mass $M_{\tau\tau}$ and angular separation $\Delta R(\tau, \tau)$, as well as the invariant mass of the $\ell\ell\tau\tau$ system) that have been chosen such that their importance to the machine learning algorithm was maximised while removing any variables that were too strongly correlated with the others. The variables and their correlations are depicted in Fig. 9. The objective of the XGBOOST learning task was set to a logistic regression for binary classification. At each step (also known as splitting), the tree booster constructs new classifiers by combining and weighting the classifiers obtained at the previous step, the initial classifiers being the input variables. The hyperparameters that were found to affect the performance of the method were the learning rate, the maximum tree depth and the minimum child weight. The learning rate controls data over-fitting by varying the learning step size, the maximum depth of a tree indicates how many times a tree can split
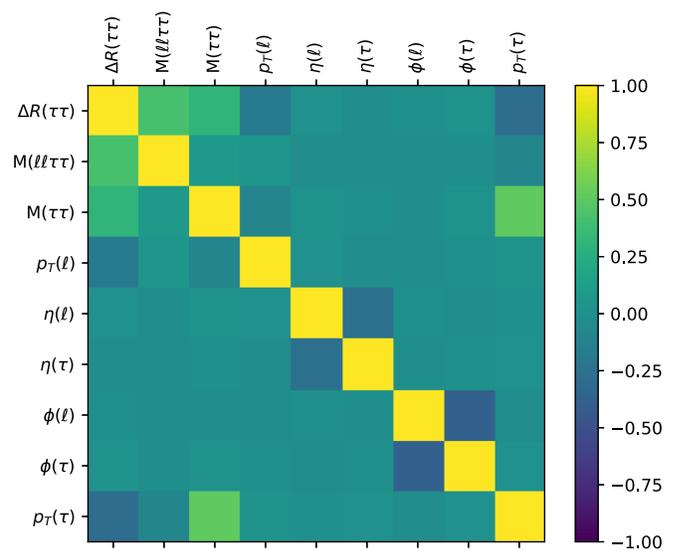

FIG. 9.   Correlations among the nine kinematic variables employed in our machine learning exercise. The set of variables includes the angular separation of the hadronic tau pair $\Delta R(\tau\tau)$, the invariant mass of the combined lepton-tau system $\ell\ell\tau\tau$ as well as of the di-tau system, and the transverse momenta p_T, pseudorapidities η and azimuthal angle ϕ of the taus and leptons.

TABLE IV. XGBOOST evaluation metric and significance obtained for a representative set of composite scenarios and pseudoscalar masses, using 150 ab^{-1} of electron-positron collisions at the Z-pole.

Model	Metric	$M_a = 10$ GeV	$M_a = 20$ GeV	$M_a = 30$ GeV	$M_a = 40$ GeV	$M_a = 50$ GeV
M2	*auc*	0.98 ± 0.003	0.87 ± 0.006	0.84 ± 0.0013	0.94 ± 0.0058	0.95 ± 0.0066
	ams	0.22	2.96	2.41	0.29	0.11
M4	*auc*	0.98 ± 0.0045	0.95 ± 0.0029	0.87 ± 0.020	0.88 ± 0.042	0.89 ± 0.061
	ams	1.16	2.83	1.69	0.54	0.15
M7	*auc*	0.98 ± 0.0018	0.86 ± 0.0082	0.88 ± 0.0011	0.90 ± 0.0012	0.94 ± 0.019
	ams	0.22	3.20	2.58	0.27	0.14
M10	*auc*	0.98 ± 0.003	0.92 ± 0.0057	0.90 ± 0.019	0.96 ± 0.0078	0.96 ± 0.0050
	ams	0.37	4.08	2.35	0.14	0.042
M12	*auc*	0.98 ± 0.0075	0.92 ± 0.003	0.92 ± 0.013	0.95 ± 0.0044	0.96 ± 0.0082
	ams	0.066	1.26	0.98	0.11	0.046

(hence controlling the algorithm complexity), and the minimum child weight controls the minimum weight that can be assigned when designing a new classifier.

In our analysis, 80% of the available Monte Carlo data was used for training purposes, and the remaining 20% for testing. For each model and M_a value, we tuned the hyperparameters using a k-fold cross-validation, so that the choice maximizing the *auc* was adopted. In particular, the maximum depth parameter was kept low and early stopping was employed in order to control over-fitting. It was found that a maximum depth of 3, a minimum child weight of 1 and a learning rate of 0.3 gave the most desirable result across the entire range of considered models. The *auc* metric and the corresponding significances obtained for a representative set of models are indicated in Table IV.

The results indicated in Table IV display a general improvement over the traditional cut-and-count method, but also an important variation across models and pseudoscalar masses. In particular, the significance peaks at $M_a = 20$ GeV for all models, as this corresponds to the maximum of the signal cross section (see Fig. 6). However, there are large differences in the trends across the models. For example, the performance for the model M10 quickly falls to one of the lowest for $M_a = 50$ GeV. These behaviors reflect not only the varying production cross sections across the models but also the variations in the kinematics resulting from differing Lagrangian parameters. On the other hand, we find a low significance for $M_a = 10$ GeV, where despite a relatively large cross section, the preselection cuts (and in particular the $M_{\tau\tau}$ requirement) rejects a large potion of the signal.

The best performance of our analysis is found for scenarios featuring $M_a = 20$, 30 GeV. For all models, the performance then drops off quickly for $M_a = 40$, 50 GeV, and it falls more sharply than it does in the cut-and-count case. Such a drop in significance at these pseudoscalar masses is expected, as the cross section decreases with M_a. Moreover, some signal kinematic distributions exhibit important variations with the pseudoscalar mass. An example can be taken from the $\Delta R(\tau^-, \tau^+)$ spectrum (see the right panel of Fig. 8), where scenarios with higher a masses ($M_a = 40$, 50 GeV) lead to very similar signal and background distribution shapes. Finally, we have trained the gradient boosting algorithm using one hyperparameter choice across all considered masses within a given model, and the choice of kinematic variables on which to train the models was guided by a focus on the lower mass setups. This training has been done in isolation for each model in order to find the hyperparameters best suited in each case. This path has clearly optimized the 20/30 GeV scenarios, as they were expected to yield the highest significance by virtue of the larger associated cross sections. The potential price to pay could be a less efficient training for higher masses of a.

In comparing the significance trends of the gradient boosting results with those of the cut-and-count method, we observe the same ranking of performance among the different models, with the exception of the model M4. This model corresponds to the highest cross sections, and may therefore have been expected to be better performing. However, our framework leads to overfitting for the $M_a = 20$, 30 GeV cases, which had to be carefully controlled by using an early stopping of the algorithm. This resulted in a lower significance.

In Table V, we translate our results in terms of the luminosity that is needed in order to achieve a significance of 2σ (to preclude the existence of the new resonance) or 3σ (to claim evidence for the resonance) at a future electron-positron collider aiming at operating at the Z-pole. The table also shows the gain obtained by using the gradient boosting algorithm over a more traditional cut-and-count method. Very importantly, our findings show that for certain models, an achievable integrated luminosity would yield a 2σ or even 3σ significance. In all cases, larger pseudoscalar masses remain likely out of reach at a c.m. energy of 91.2 GeV, as does the $M_a = 10$ GeV case.

TABLE V.   Required luminosities, in ab^{-1}, to obtain a 2σ and 3σ significance to the pseudoscalar signal at a future electron-positron collider operating at the Z-pole. We present results for our cut-and-count (third and fourth columns) and gradient boosting (fifth and sixth columns) methods, for an illustrative selection of models.

Model	M_a [GeV]	Cut and count		Machine learning	
		2σ	3σ	2σ	3σ
M2	10	2.67×10^8	6.00×10^8	1.24×10^4	2.79×10^4
	20	3.55×10^4	7.99×10^4	68.5	154
	30	7.41×10^4	1.67×10^5	103	232
	40	2.50×10^5	5.62×10^5	7.13×10^3	1.61×10^4
	50	1.50×10^6	3.38×10^6	4.96×10^4	1.12×10^5
M4	10	3.55×10^8	7.99×10^8	446	1.00×10^3
	20	3.40×10^3	7.65×10^3	74.9	169
	30	8.88×10^3	2.00×10^4	210	473
	40	4.17×10^4	9.38×10^4	2.06×10^3	4.63×10^3
	50	3.75×10^5	8.44×10^5	2.67×10^4	6.00×10^4
M7	10	1.04×10^8	2.34×10^8	1.24×10^4	2.79×10^4
	20	3.06×10^4	6.89×10^4	58.5	132
	30	4.96×10^4	1.12×10^5	90.1	203
	40	1.61×10^5	3.63×10^5	8.23×10^3	1.85×10^4
	50	1.13×10^6	2.55×10^6	3.06×10^4	6.89×10^4
M10	10	3.40×10^7	7.65×10^7	4.38×10^3	9.86×10^3
	20	4.96×10^4	1.12×10^5	36.0	81.1
	30	1.98×10^5	4.46×10^5	109	244
	40	1.13×10^6	2.55×10^6	3.06×10^4	6.89×10^4
	50	9.86×10^6	2.22×10^7	3.40×10^5	7.65×10^5
M12	10	1.61×10^9	3.63×10^9	1.38×10^5	3.10×10^5
	20	2.72×10^5	6.11×10^5	378	850
	30	4.90×10^5	1.10×10^6	624	1.41×10^3
	40	1.36×10^6	3.06×10^6	4.96×10^4	1.12×10^5
	50	2.08×10^6	4.67×10^6	2.84×10^5	6.38×10^5

IV. CONCLUSION

In this work we have designed an analysis targeting a light pseudoscalar, ubiquitous to composite Higgs models, at a future electron-positron collider aiming at collecting a large luminosity at the Z-pole. In our predictions we have considered the pseudoscalar couplings to gauge bosons to full leading order, i.e., by including relevant effects stemming from loops of b quarks. The latter have a significant impact for low mass pseudoscalars, unlike what is traditionally assumed, and should be considered both at present and future hadron and lepton colliders.

We have demonstrated the possibility of actually getting hints for a low mass pseudoscalar at a future lepton collider operating at a center-of-mass energy of 91.2 GeV, focusing on the production mode in which the pseudoscalar is produced in association with a pair of electrons or muons and decays into a pair of hadronic taus. The corresponding Standard Model background has been found difficult to

reduce via a standard cut-and-count analysis, which resulted in a poor sensitivity and a rare signal entirely hidden within the large background. In an attempt to improve these findings we have made use of a machine learning algorithm based on boosted decision trees. It yielded an improvement in sensitivity in almost all cases. In particular, we have observed a marked improvement for scenarios in which the pseudoscalar mass $M_a = 20, 30$ GeV, where the related significance approaches 3σ at an integrated luminosity of 150 ab^{-1}. Lighter configurations ($M_a \lesssim 10$ GeV) are not promising, given that the signal is expected to be dominated by the background and mostly annihilated by any decent event preselection. The significance also drops off for higher pseudoscalar masses, by virtue of a decreasing signal cross section and key kinematic properties becoming very similar to the background ones. It is, however, possible that this could be ameliorated by designing more appropriate and dedicated variables. Additionally, while this analysis has focused on the four-lepton final state, a further work may investigate the $e^+e^- \to a\gamma$ channel which may be competitive, where similar studies [57] have shown the $e^+e^- \to a\gamma(a \to \ell\ell)$ signature to be more constraining for axionlike particles. In focusing on the four lepton final state we have shown that, using machine learning capabilities, even weaker signals are accessible.

We have used, as an example, the framework of a future electron-positron collider aiming at operating at the Z-pole and collecting a very large integrated luminosity. In this context, we have demonstrated that search avenues for weakly coupled particles are promising. The analysis that we have proposed complements earlier works, such as in Ref. [20], focusing on potential options for the LHC, and fills a gap for what concerns light states. It has indeed been shown that there is a scarcity of searches for such a light pseudoscalar in the mass region that has been considered in this paper [17]. In particular, the Higgs to BSM branching ratio often provides the sole (indirect) constraint in the mass region $M_a \in [10, 60]$ GeV [17,20]. Additionally, investigations into future collider prospects for light pseudoscalars have been previously presented in Refs. [57,58]. In those broad comprehensive reviews of existing bounds on axionlike particles, our channel is complementary to those studied, and falls within the bands of unconstrained parameter space. The proposed analysis would be complementary to existing diphoton searches [28], where a drop in sensitivity in the diphoton channel corresponds to increased sensitivity within the ditau channel [20] (and vice versa) for the models M1-12 across the mass range here considered. Additionally, this analysis could reach below the existing 90 GeV lower bound for ditau searches [59]. In this we find our pseudoscalar to be a candidate which evades all existing bounds, and present this channel as an additional one which may be accessed through machine learning capabilities.

From our findings, we demonstrated that a direct search for a light composite pseudoscalar at high integrated luminosity lepton colliders should be seriously considered.

While our generic analysis covers the parameter space region in which the mass of the pseudoscalar is less than 60 GeV, it is certainly less sensitive to M_a values of 40 GeV or more. Future works should determine whether it could be optimised for these heavier configurations, perhaps by considering future lepton colliders operating at higher center-of-mass energies. Among the avenues to be explored, one could benefit from a gain in sensitivity by relying on the spin-0 nature of the pseudoscalar and assessing the potential of various angular distributions between pairs of final-state objects. For the same reason, it may also be useful to make use of tau polarization in order to separate signal from the background [60]. Finally, other options may rely on the presence of a second heavier pseudoscalar η', that is common to many composite models.

[1] T. Behnke, J. E. Brau, B. Foster, J. Fuster, M. Harrison, J. M. Paterson, M. Peskin, M. Stanitzki, N. Walker, and H. Yamamoto, The international linear collider technical design report—Volume 1: Executive summary, arXiv:1306.6327.

[2] T. K. Charles *et al.* (CLICdp, CLIC Collaborations), The compact linear collider (CLIC)—2018 Summary report, CERN Yellow Rep. Monogr. **2/2018,** 1 (2018).

[3] A. Abada *et al.* (FCC Collaboration), FCC-ee: The lepton collider, Eur. Phys. J. Special Topics **228,** 261 (2019).

[4] M. Dong *et al.* (CEPC Study Group Collaboration), CEPC conceptual design report: Volume 2—Physics & detector, arXiv:1811.10545.

[5] D. B. Kaplan and H. Georgi, SU(2) × U(1) Breaking by vacuum misalignment, Phys. Lett. **136B,** 183 (1984).

[6] D. B. Kaplan, H. Georgi, and S. Dimopoulos, Composite Higgs scalars, Phys. Lett. **136B,** 187 (1984).

[7] M. J. Dugan, H. Georgi, and D. B. Kaplan, Anatomy of a composite Higgs model, Nucl. Phys. **B254,** 299 (1985).

[8] G. Cacciapaglia and F. Sannino, Fundamental composite (goldstone) Higgs dynamics, J. High Energy Phys. 04 (2014) 111.

[9] G. Ferretti and D. Karateev, Fermionic UV completions of Composite Higgs models, J. High Energy Phys. 03 (2014) 077.

[10] J. Barnard, T. Gherghetta, and T. S. Ray, UV descriptions of composite Higgs models without elementary scalars, J. High Energy Phys. 02 (2014) 002.

[11] A. Arbey, G. Cacciapaglia, H. Cai, A. Deandrea, S. Le Corre, and F. Sannino, Fundamental composite electroweak dynamics: Status at the LHC, Phys. Rev. D **95,** 015028 (2017).

[12] K. Agashe, R. Contino, and A. Pomarol, The minimal composite Higgs model, Nucl. Phys. **B719,** 165 (2005).

[13] R. Contino, The higgs as a composite Nambu-Goldstone boson, in *Physics of the large and the small, TASI 09, Proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009* (2011), pp. 235–306, https://doi.org/10.1142/9789814327183_0005.

[14] B. Bellazzini, C. Csáki, and J. Serra, Composite Higgses, Eur. Phys. J. C **74,** 2766 (2014).

[15] G. Panico and A. Wulzer, The composite Nambu-Goldstone Higgs, Lect. Notes Phys. **913,** 1 (2016).

[16] S. Weinberg, Implications of dynamical symmetry breaking, Phys. Rev. D **13,** 974 (1976).

[17] G. Cacciapaglia, G. Ferretti, T. Flacke, and H. Serôdio, Light scalars in composite Higgs models, Front. Phys. **7,** 22 (2019).

[18] D. B. Kaplan, Flavor at SSC energies: A New mechanism for dynamically generated fermion masses, Nucl. Phys. **B365,** 259 (1991).

[19] G. Ferretti, UV completions of partial compositeness: The case for a SU(4) gauge group, J. High Energy Phys. 06 (2014) 142.

[20] G. Cacciapaglia, G. Ferretti, T. Flacke, and H. Serodio, Revealing timid pseudo-scalars with taus at the LHC, Eur. Phys. J. C **78,** 724 (2018).

[21] D. Buarque Franzosi, G. Cacciapaglia, and A. Deandrea, Sigma-assisted natural composite Higgs, Eur. Phys. J. C **80,** 28 (2020).

[22] E. Witten, Global aspects of current algebra, Nucl. Phys. **B223,** 422 (1983).

[23] J. Wess and B. Zumino, Consequences of anomalous Ward identities, Phys. Lett. **37B,** 95 (1971).

[24] A. M. Sirunyan *et al.* (CMS Collaboration), Search for low-mass resonances decaying into bottom quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. D **99,** 012005 (2019).

[25] M. Aaboud *et al.* (ATLAS Collaboration), Search for Low-Mass Dijet Resonances using Trigger-Level Jets with the ATLAS Detector in pp Collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **121,** 081801 (2018).

[26] A. M. Sirunyan *et al.* (CMS Collaboration), Search for a Narrow Resonance Lighter than 200 GeV Decaying to a Pair

of Muons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **124**, 131802 (2020).

[27] G. Aad *et al.* (ATLAS Collaboration), Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Lett. B **796**, 68 (2019).

[28] A. Mariotti, D. Redigolo, F. Sala, and K. Tobioka, New LHC bound on low-mass diphoton resonances, Phys. Lett. B **783**, 13 (2018).

[29] A. M. Sirunyan *et al.* (CMS Collaboration), Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV, Phys. Lett. B **793**, 320 (2019).

[30] M. Aaboud *et al.* (ATLAS Collaboration), Search for resonances in diphoton events at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 09 (2016) 001.

[31] M. Aaboud *et al.* (ATLAS Collaboration), Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 01 (2018) 055.

[32] G. Ferretti, Gauge theories of Partial Compositeness: Scenarios for Run-II of the LHC, J. High Energy Phys. 06 (2016) 107.

[33] G. von Gersdorff, E. Pontón, and R. Rosenfeld, The dynamical composite Higgs, J. High Energy Phys. 06 (2015) 119.

[34] O. Witzel, Review on composite Higgs models, Proc. Sci., LATTICE2018 (**2019**) 006 [arXiv:1901.08216].

[35] G. Cacciapaglia, The dynamics of composite Higgses, J. Phys. Conf. Ser. **623**, 012006 (2015).

[36] A. Belyaev, G. Cacciapaglia, H. Cai, G. Ferretti, T. Flacke, A. Parolini, and H. Serodio, Di-boson signatures as standard candles for partial compositeness, J. High Energy Phys. 01 (2017) 094; Erratum, J. High Energy Phys. 12 (2017) 088.

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

[38] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, UFO—The universal FeynRules output, Comput. Phys. Commun. **183**, 1201 (2012).

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

[40] R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, and J. Rojo (NNPDF Collaboration), Parton distributions with QED corrections, Nucl. Phys. **B877**, 290 (2013).

[41] L. Mason, A. S. Cornell, A. Deandrea, and B. Fuks, Bottom-quark contributions to composite pseudo–scalar couplings at LHC, in *Strong Dynamics for Physics within and Beyond the Standard Model at LHC and Future Colliders* (2019), pp. 110–115.

[42] M. Dam, Tau-lepton physics at the FCC-ee circular e^+e^- collider, SciPost Phys. Proc. **1**, 041 (2019).

[43] J. Yan, S. Watanuki, K. Fujii, A. Ishikawa, D. Jeans, J. Strube, J. Tian, and H. Yamamoto, Measurement of the Higgs boson mass and $e^+e^- \to ZH$ cross section using $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ at the ILC, Phys. Rev. D **94**, 113002 (2016).

[44] A. M. Pingel, Tau lepton identification and studies of associated Higgs boson production with the ATLAS detector, 2015, https://cds.cern.ch/record/2224284. Presented 18 Aug 2015.

[45] T. Sjostrand, S. Mrenna, and P. Z. Skands, A brief introduction to PYTHIA8.1, Comput. Phys. Commun. **178**, 852 (2008).

[46] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), DELPHES3, A modular framework for fast simulation of a generic collider experiment, J. High Energy Phys. 02 (2014) 057.

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

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

[49] E. Conte, B. Fuks, and G. Serret, MadAnalysis 5, A user-friendly framework for collider phenomenology, Comput. Phys. Commun. **184**, 222 (2013).

[50] E. Conte and B. Fuks, Confronting new physics theories to LHC data with MadAnalysis 5, Int. J. Mod. Phys. A **33**, 1830027 (2018).

[51] T. Chen and C. Guestrin, Xgboost: A scalable tree boosting system, arXiv:1603.02754.

[52] E. Fontanesi, L. Pezzotti, and M. Selvaggi, Fast (and full) simulation of the idea detector, 2019, https://agenda.infn.it/event/18976/contributions/88573/attachments/62454/74955/CEPC-Oxford2019_v1.pdf.

[53] P. Bärtschi, C. Galloni, C. Lange, and B. Kilminster, Reconstruction of τ lepton pair invariant mass using an artificial neural network, Nucl. Instrum. Methods Phys. Res., Sect. A **929**, 29 (2019).

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

[55] T. Chen and T. He, Higgs boson discovery with boosted trees, in *Proceedings of the 2014 International Conference on High-Energy Physics and Machine Learning—Volume 42*, HEPML'14 (Journal of Machine Learning Research, 2014), p. 69–80.

[56] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71**, 1554 (2011); Erratum, Eur. Phys. J. C **73**, 2501 (2013).

[57] M. Bauer, M. Heiles, M. Neubert, and A. Thamm, Axion-like particles at future colliders, Eur. Phys. J. C **79**, 74 (2019).

[58] M. Bauer, M. Neubert, and A. Thamm, Collider probes of axion-like particles, J. High Energy Phys. 12 (2017) 044.

[59] CMS Collaboration, Search for a neutral MSSM Higgs boson decaying into $\tau\tau$ at 13 TeV, CERN Report No. CMS-PAS-HIG-16-006, 2016, https://cds.cern.ch/record/2160252.

[60] K. A. Assamagan and A. Deandrea, The Hadronic tau decay of a heavy charged Higgs in models with singlet neutrino in large extra dimensions, Phys. Rev. D **65**, 076006 (2002).