# New Higgs decays to axion-like particles 

Anke Biekötter ${ }^{\text {a,* }}$, Mikael Chala ${ }^{\text {b }}$, Michael Spannowsky ${ }^{\text {a }}$<br>${ }^{a}$ Institute for Particle Physics Phenomenology, Department of Physics, Durham University, South Road, Durham DH1 3LE, United Kingdom<br>b Departamento de Física Teórica y del Cosmos, Universidad de Granada, E-18071 Granada, Spain

## ARTICLE INFO

## Article history:

Received 4 April 2022
Received in revised form 18 August 2022
Accepted 22 September 2022
Available online 26 September 2022
Editor: G.F. Giudice


#### Abstract

We investigate the interactions of a light scalar with the Higgs boson and second-generation fermions, which trigger new rare decays of the Higgs boson into $4 \mu, 2 \mu 2 \gamma, 6 \mu$ and $4 \mu 2 j$. We recast current LHC searches to constrain these decays and develop new collider analyses for those channels which are only poorly tested by existing studies. With the currently collected data we can probe branching ratios as small as $1.5 \times 10^{-5}, 8.7 \times 10^{-5}, 5.7 \times 10^{-8}$ and $1.6 \times 10^{-7}$, respectively. For the High-Luminosity LHC run, considered here to involve $3 \mathrm{ab}^{-1}$ of integrated luminosity, these numbers go down to $1.3 \times 10^{-5}$, $2.0 \times 10^{-6}, 3.0 \times 10^{-9}$ and $5.4 \times 10^{-9}$, respectively. We also comment on other channels that remain still unexplored.


© 2022 Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/). Funded by SCOAP ${ }^{3}$.

## 1. Introduction

Light neutral (pseudo-)scalar particles are ubiquitous in models of new physics, including theories with a spontaneously broken global symmetry such as the Peccei-Quinn solution to the strong CP problem [1-4], composite Higgs models [5-7] and others $[8,9]$. Moreover, they appear in different explanations for dark matter [10-12] as well as for the flavour problem [8,13-15] and the hierarchy problem [16]. With a little abuse of language, we will hereafter refer to these light scalars as axion-like particles (ALPs), irrespective of whether or not they fulfil a shift symmetry $a \rightarrow a+a_{0}$. In-depth analyses, including quantum effects, of the possible interactions between these ALPs and the SM fields, have been recently presented in Refs. [17-20].

Many of those couplings are constrained by data collected at low-energy and flavour facilities [21-33] colliders including LEP [34-37] and the LHC [36-48] or in astrophysical events [4955]. Some couplings have been shown to be only testable in Higgs physics, particularly through yet unexplored decays of the Higgs boson. Recent analyses in this line include flavour-violating ALPs $[56,57]$ and searches for Higgs decays to ALPs with subsequent decay to photons and jets explored in [58] and subsequent decays to electrons in [59]; see also Ref. [60] for an extensive review on the topic.

[^0]In this paper, despite continuing this research avenue, we focus on flavour-conserving Higgs decays in which not all final fermion pairs reconstruct narrow masses as our targeted signals ensue from Higgs-fermion-ALP contact interactions. The article is organised as follows. In section 2 we introduce the relevant ALP interactions with Standard Model (SM) particles; we discuss their possible origin and estimate the resulting Higgs branching ratios into the final states of interest. In section 3 we present a thorough analysis of the potential of current experimental data. We recast sensitive existing searches, develop new search strategies and set limits on the ALP interactions with the SM. In section 4 we briefly discuss the implications for the anomalous magnetic moment of the muon. We conclude in section 5 . We dedicate Appendix A to limitations of other potential analyses not considered in the main text.

## 2. Theory and experiment

From an agnostic point of view, and without exact knowledge of the nature of the electroweak symmetry breaking (EWSB), namely whether it is linearly or non-linearly realised, the interactions between the ALP and the SM particles must be described by an effective-field theory, the Lagrangian of which reads:

$$
\begin{align*}
L & =\frac{1}{2} \partial_{\mu} a \partial^{\mu} a-\frac{m_{a}^{2}}{2} a^{2} \\
& +c_{a} \frac{a}{f_{a}} A_{\mu \nu} \tilde{A}^{\mu \nu}+c_{a \psi} a \bar{\psi} \psi+c_{1}^{\psi} \frac{a}{f_{a}} h \bar{\psi} \psi+c_{2}^{\psi} \frac{a^{2}}{f_{a}^{2}} h \bar{\psi} \psi \tag{1}
\end{align*}
$$

plus other terms that are either suppressed by further powers $f_{a}$ or that involve other gauge bosons, which we are not interested in.


Fig. 1. Diagrams representing the Higgs decays of interest: $h \rightarrow 4 \mu$ (left), $h \rightarrow 2 \mu 2 \gamma$ (centre-left), $h \rightarrow 6 \mu$ (centre-right), $h \rightarrow 4 \mu 2 j$ (right).

Note that $h$ represents the physical Higgs boson, A stands for the photon, and $\psi$ can be any fermion; we ignore family indices as we assume flavour conservation. We only consider second-generation fermions in the following, since their couplings, particularly those involving leptons (muons), can almost exclusively be tested in Higgs decays at the LHC. There are currently no other facilities in which the intervening particles can collide at sufficiently high energies.

The Lagrangian above contains the first terms in the leadingorder ALP chiral Lagrangian. ${ }^{1}$ Different bounds on the parameters of this Lagrangian can be found, for example, in Ref. [40]. To the best of our knowledge no constraints have been reported on the equivalent of $c_{1}^{\psi}$ and $c_{2}^{\psi}$.

The $h a \bar{\psi} \psi$ interactions can be sizeable even if $a \bar{\psi} \psi$ couplings are small or absent. If the EWSB is linearly realised, then Eq. (1) arises from a more fundamental Lagrangian involving the full Higgs doublet $H=\left[G^{+},\left(h+i G^{0}\right) / \sqrt{2}\right]^{T}$ upon EWSB; $h \rightarrow h+v$ with $v \sim$ 246 GeV . Even in this case, $c_{1}^{\psi}$ is not necessarily related to $c_{a \psi}$, being therefore, in general, unconstrained. To see this, consider the following Lagrangian:
$L=\frac{a}{f_{a}} \overline{\psi_{L}} H \psi_{R}\left(c_{1}+c_{2} \frac{|H|^{2}}{f_{a}^{2}}\right)+$ h.c.
One can, without loss of generality, define the new coupling $c_{1}^{\prime}=$ $c_{1}+c_{2} v^{2} /\left(2 f_{a}^{2}\right)$, from where we obtain the following Lagrangian after EWSB:
$L=c_{1}^{\prime} \frac{v}{\sqrt{2} f_{a}} a \bar{\psi} \psi+\left(c_{1}^{\prime}+c_{2} \frac{v^{2}}{f_{a}^{2}}\right) \frac{a}{\sqrt{2} f_{a}} h \bar{\psi} \psi$.
Since $c_{1}$ was completely arbitrary, so is $c_{1}^{\prime}$. The fact that the coupling $c_{1}^{\prime}$, which induces decays of the ALP to fermions, is small as required by experimental data [36,61], implies some tuning between two couplings. More concretely, for $f_{a} \sim 1 \mathrm{TeV}, c_{1}^{\prime} \lesssim$ $10^{-2}-10^{-1}$ [61], which in turn requires a fine-tuning between $c_{1}$ and $c_{2} v^{2} / f_{a}^{2}$ of the order of $10-100$. It also implies that $c_{1} / c_{2}$ is smaller than naively anticipated based on $v / f_{a}$ power counting alone. Beyond accidental, this possibility could reflect that the two operators arise at different order in perturbation theory in the UV, similar to models in which the Weinberg operator is generated at loop level while its dimension-7 counterpart arises at tree level [62]. Be that as it may, it must be considered that it is possible to have sizeable $h a \bar{\psi} \psi$ couplings in the absence of large $a \bar{\psi} \psi$ interactions.

The different ways in which the ALP a can decay remain largely unconstrained. Thus, $a \rightarrow \mu^{+} \mu^{-}$can dominate if, for example, a $\mathbb{Z}_{2}$ symmetry $a \rightarrow-a$ is only broken in the muon sector while being exact in the quark sector. This forbids $h \rightarrow a q q$, but allows

[^1]$h \rightarrow a \mu \mu$. Likewise, it makes $c_{2}^{\psi}$ vanish for neither leptons nor quarks, so both $h \rightarrow a a q q$ and $h \rightarrow a a \mu \mu$ remain open in principle. However, the fact that both can coexist is of little relevance for our analysis, as the SM dominates the Higgs width; $\Gamma_{h} \sim 4 \mathrm{MeV}$ [63]. On the other hand, $a \rightarrow \gamma \gamma$ can be sizeable (or even the dominating decay) if the $\mathbb{Z}_{2}$ symmetry $a \rightarrow-a$ and $h \rightarrow-h$ is respected in good approximation by all fermions and provided $c_{a A}$ is nonzero (Note that even if $c_{a A} \sim \alpha /(4 \pi)$, as suggested by naive power counting, the decay $a \rightarrow \gamma \gamma$ is prompt.).

In light of this discussion, we focus on the Higgs decays to ALPs with subsequent decays of the ALP to muons or photons in this article. In particular, we investigate the following Higgs decay modes which are shown in Fig. 1: (i) $h \rightarrow a \mu \mu, a \rightarrow \mu \mu$; (ii) $h \rightarrow a \mu \mu, a \rightarrow \gamma \gamma$; (iii) $h \rightarrow a a \mu \mu, a \rightarrow \mu \mu$; (iv) $h \rightarrow a a q q, a \rightarrow$ $\mu \mu$. We restrict the ALP mass range to $m_{a} \in[10,50]$ in order to stay well below the $Z$ boson peak for di-muon resonances and to ensure a large enough angular separation between the photons in low-mass di-photon resonances. Apart from Higgs decays, the processes (i)-(iii) can only be tested at muon colliders in the considered ALP mass range. However, process (iv) can also be tested at the LHC through $p p \rightarrow h a a, a \rightarrow \mu \mu$ and we will indeed see in section 3.4 that this channel provides more substantial limits on the corresponding Wilson coefficient.

To test these Higgs decays, we recast current LHC analyses at $\sqrt{s}=13 \mathrm{TeV}$, where possible. However, we refrain from using 7 TeV or 8 TeV analyses due to the lower Higgs production cross-section and luminosity. We have performed a thorough and systematic study of experimental searches in final states involving at least two muons to find suitable analyses. We notice that most beyond SM searches are of little relevance. SUSY searches typically enforce substantial cuts on the missing energy $E_{T}^{\text {miss }}$ absent within our signals, whereas heavy resonance searches, on the other hand, often require large transverse momenta of the final-state particles, while the decay products of a four- or six-body decay of a 125 GeV resonance are rarely very energetic.

Before discussing search strategies for our exotic Higgs decays, let us estimate the corresponding partial decay widths. This allows us to project the number of events for current and future runs of the LHC and to establish the signal acceptance needed for an analysis to be sensitive to our ALP signals.

We concentrate first on the four-body decays of the Higgs, $h \rightarrow$ $4 \mu$ and $h \rightarrow 2 \mu 2 \gamma$. The partial width for the Higgs decay to an ALP and two fermions can be estimated to be:
$\Gamma(h \rightarrow a \bar{\psi} \psi) \sim \frac{\left(c_{1}^{\psi}\right)^{2}}{768 \pi^{3} f_{a}^{2}} m_{h}^{3}$.
For $\mathcal{O}(1)$ couplings at $f_{a}=1 \mathrm{TeV}$, the corresponding branching ratio is about $10^{-3}$. Taking into account the Higgs production cross section at $\sqrt{s}=13 \mathrm{TeV}$ which is approximately 48 pb [64] and the planned integrated luminosity of $\mathcal{L}=3 \mathrm{ab}^{-1}$ we expect $\sim 10^{6}$ $h \rightarrow a \bar{\psi} \psi$ events at the High-Luminosity LHC (HL-LHC). Optimistically assuming that at least 10 signal events are needed to distinguish the signal from background fluctuations at the $2 \sigma$ level (as these two channels are never background-free), it can be straightforwardly concluded that an analysis probing the two Higgs decays requires at least a signal acceptance of $10^{-5}$.

Let us now turn our attention to the six-body Higgs decays in the channels $h \rightarrow 6 \mu$ and $h \rightarrow 4 \mu 2 j$. These processes result from the decay of a Higgs into two ALPs and two fermions with an approximate width:
$\Gamma(h \rightarrow a a \bar{\psi} \psi) \sim \frac{\left(c_{2}^{\psi}\right)^{2}}{73728 \pi^{5} f_{a}^{4}} m_{h}^{5}$.
The corresponding branching ratio is of about $10^{-7}$ for $\mathcal{O}(1)$ couplings at $f_{a}=1 \mathrm{TeV}$ and we expect only $\sim 50 h \rightarrow a a \bar{\psi} \psi$ events to be produced at the HL-LHC. Even background-free experimental searches can thus only be sensitive to $\mathcal{O}(1)$ couplings triggering these decays provided their signal acceptance is at least $10 \%$.

## 3. Analysis

To set limits on the ALP interactions in our model, we compare the predicted fiducial cross-sections with current LHC data. We have implemented the interactions of the Lagrangian in Eq. (1) in FeynRules [65]. Events have been generated using MadGraphv2.7.3 [66], showered with Pythia8 [67] and analysed using Rivet-v2.7.0 [68].

Unless otherwise stated, all signal and background samples have been generated at leading order. For gluon fusion Higgs production, we include a $k$ factor to normalize the cross-section to its $\mathrm{N}^{3}$ LO QCD+NLO EW prediction of 48.58 pb [64].

### 3.1. Four muon final state

To constrain the Wilson coefficient $c_{1}^{\mu}$, we analyse contributions to Higgs decays in the four-lepton final state via $h \rightarrow a \mu \mu \rightarrow 4 \mu$ as represented by the Feynman diagram in the left panel of Fig. 1.

Several LHC searches analyse final states with four muons, rendering them potentially sensitive to our $h \rightarrow 4 \mu$ signal. However, most SM analyses focus on different invariant mass ranges for the four-lepton system, see e.g. Ref. [70] for a $Z \rightarrow 4 \mu$ analysis, and most exotic/SUSY searches focus on high-energy events with large total invariant or effective masses, see Refs. [71,72].

An analysis that targets explicitly our signal properties is the CMS search for low-mass di-lepton resonances in Higgs boson decays to four leptons [69]. The analysis requires two opposite-sign same-flavour (OSSF) lepton pairs with a four-lepton invariant mass close to the Higgs mass:

$$
\begin{gather*}
p_{T}^{\ell}>20,15,5,5 \mathrm{GeV}, \quad m_{4 \ell} \in[118,130] \mathrm{GeV} \\
m_{\ell^{+} \ell^{-}}>4 \mathrm{GeV}, \quad m_{\ell^{+} \ell^{-}} \notin[8.0,11.5] \mathrm{GeV}  \tag{6}\\
m_{Z_{1}}>40 \mathrm{GeV}, \quad \Delta R\left(\ell_{i}, \ell_{j}\right)>0.02
\end{gather*}
$$

For the $\mathrm{ZX}(\mu \mu)$ search region of the analysis, the leptons are paired such that $Z_{1}$ is the OSSF lepton pair closest to the $Z$ boson mass, i.e. for which $\left|m_{Z_{1}}-m_{Z}\right|$ is minimal. The second OSSF lepton pair $Z_{2}$ is required to consist of two muons. Events with an OSSF lepton pair in the invariant mass range $m_{\ell^{+} \ell^{-}} \in[8.0,11.5] \mathrm{GeV}$ are excluded to suppress the background from $\Upsilon(b b \overline{)}$ decays. We have implemented the cuts in Eq. (6) in Rivet.

We present the mass distribution of the second OSSF lepton pair $m_{Z_{2}}$ in the top panel of Fig. 2. A Higgs decay via the $h \rightarrow$ $a \mu \mu \rightarrow 4 \mu$ channel leads to a resonance at $m_{Z_{2}} \sim m_{a}$ (in the considered $m_{a}$ mass range). Therefore, we can use this distribution to set limits on the Wilson coefficient $c_{1}^{\mu}$. The dominant backgrounds for this analysis are $h \rightarrow Z Z^{*}$ and $Z Z$ production, and we take their prediction directly from Fig. 2a of Ref. [69] along with a $10 \%$ systematic uncertainty on the total per-bin background prediction. The $h \rightarrow Z Z^{*} \rightarrow 4 \ell$ background prediction was used to validate our implementation of the CMS analysis in Rivet. We agree with


Fig. 2. Top: $m_{Z_{2}}$ distribution in the $h \rightarrow 4 \mu$ analysis for data, background and signals at different masses $\left(\mathcal{L}=137 \mathrm{fb}^{-1}\right)$. The grey dashed line at $m_{Z_{2}}=36 \mathrm{GeV}$ represents the upper limit of the search region considered in the CMS analysis [69]; see text for details. Bottom: Upper 95\% CLs limits on $c_{1}^{\mu} / f_{a}$ at different luminosities. Limits marked with an " $x$ " are observed limits, other limits are expected limits.
the number of events in each of the 64 bins at the $10 \%$ level after including an experimental efficiency of $74 \%$.

The corresponding observed (expected) $95 \%$ CLs limits [73] at $\mathcal{L}=137 \mathrm{fb}^{-1}$ are:
$\frac{c_{1}^{\mu}}{f_{a}}<\begin{array}{ll}0.027(0.021) / \mathrm{TeV} & \text { for } m_{s}=15 \mathrm{GeV} \\ 0.027(0.025) / \mathrm{TeV} & \text { for } m_{s}=30 \mathrm{GeV}\end{array}$,
assuming that all ALPs decay to muons only.
While the CMS search only considers the region $m_{Z_{2}} \leq 36 \mathrm{GeV}$, we can in principle extend the search region to higher $Z_{2}$ masses. The background at these masses is dominated by SM Higgs decays and we rely on our own background prediction in this region. The expected $95 \%$ CLs limit for a $m_{a}=50 \mathrm{GeV}$ ALP and $137 \mathrm{fb}^{-1}$ of data are:
$\frac{c_{1}^{\mu}}{f_{a}}<0.037 / \mathrm{TeV}$ for $m_{s}=50 \mathrm{GeV}$.
In the bottom panel of Fig. 2, we show the limits on $c_{1}^{\mu} / f_{a}$ for different ALP masses. We present the limits for the current luminosity $137 \mathrm{fb}^{-1}$ as well as for the planned luminosities after LHC Run III, $300 \mathrm{fb}^{-1}$, and the HL-LHC, $3 \mathrm{ab}^{-1}$. For higher ALP masses, the limits slightly decrease due to the larger phase-space suppression; also, compare the top panel of Fig. 2. However, the reduced production cross-section is only partially compensated by a smaller background contribution and a higher signal acceptance as a result of the muons carrying on average higher transverse momenta.

The sensitivity of the current CMS analysis to our ALP signal could be further improved by requiring that all four leptons, and not only the ones forming $Z_{2}$, are muons. This would decrease the background contribution by an approximate factor of $1 / 2$ and tighten the expected limits by around $10 \%$.


Fig. 3. Top: $m_{\gamma \gamma}$ distribution in the $h \rightarrow a \mu \mu \rightarrow 2 \mu 2 \gamma$ analysis for signal and backgrounds at $\mathcal{L}=137 \mathrm{fb}^{-1}$. Bottom: Upper $95 \%$ CLs limits on $c_{1}^{\mu} / f_{a}$ for different luminosities.

### 3.2. Two muon two photon final state

We now consider the possibility that the ALP is produced in a Higgs decay in association with two muons and that it decays into two photons. The corresponding Feynman diagram is shown in the second panel of Fig. 1.

The most promising LHC analysis to constrain our $2 \mu 2 \gamma$ ALP signal is the CMS search targeting events with a leptonically decaying $Z$ boson and two photons [74]. ${ }^{2}$ However, its cut on the transverse momentum of the two photons, $p_{T}^{\gamma}>20 \mathrm{GeV}$, is only rarely passed for the photons resulting from the decay of a lowmass ALP $a \rightarrow \gamma \gamma$. To increase the sensitivity on Higgs decays to ALPs via the $h \rightarrow a 2 \mu$ channel, we design our own analysis based on Ref. [74]. We relax the cuts $m_{\ell \ell}>55 \mathrm{GeV}$ and $p_{T}^{\gamma}>$ 20 GeV in the experimental reference to make the analysis more inclusive:

$$
\begin{gather*}
p_{T}^{\mu}>30,15 \mathrm{GeV}, \quad\left|\eta^{\mu}\right|<2.4 \mathrm{GeV}, \quad m_{\mu \mu}>12 \mathrm{GeV} \\
\left|\eta^{\gamma}\right|<2.5 \mathrm{GeV}\left(\text { excl. } 1.44<\left|\eta^{\gamma}\right|<1.57\right)  \tag{9}\\
p_{T}^{\gamma}>15, \mathrm{GeV}, \quad \Delta R(\gamma, \mu / \gamma)>0.4
\end{gather*}
$$

To select the Higgs decay region we further require $m_{\ell \ell \gamma \gamma} \in$ [110, 140] GeV.

To set limits, we consider the invariant mass distribution of the two photons $m_{\gamma \gamma}$, where the ALP signal will produce a resonance at its mass; compare the top panel of Fig. 3.

[^2]The dominant backgrounds are $Z \gamma \gamma$ and $Z+$ jets production, where the jets fake photons. To estimate the size of these backgrounds, we have generated $\mu^{+} \mu^{-} \gamma \gamma$ production at NLO in QCD using MadGraph. To account for the $Z+$ jets background, we have multiplied the generated background distribution by a factor 1.7. This factor is based on the $Z+$ jets production background contribution in $Z \gamma \gamma$ production [74], which is $\sim 62 \%$ the size of the signal. We conservatively assume a $25 \%$ systematic uncertainty on the background prediction based on the uncertainty of $Z+$ jets production background contribution; see Ref. [74].

The $95 \%$ CLs limits resulting from the $m_{\gamma \gamma}$ distribution for $137 \mathrm{fb}^{-1}$ are:
$\frac{c_{1}^{\mu}}{f_{a}}<\begin{array}{ll}0.066 / \mathrm{TeV}^{2} & \text { for } m_{s}=15 \mathrm{GeV} \\ 0.083 / \mathrm{TeV}^{2} & \text { for } m_{s}=50 \mathrm{GeV}\end{array}$.
In the bottom panel of Fig. 3, we show the limits on $c_{1}^{\mu} / f_{a}$ for different ALP masses and luminosities. The limits depend only very weakly on the ALP mass $m_{a}$ as a result of three competing effects: First, the Higgs branching ratio into ALPs decreases with $m_{a}$ due to the phase-space suppression. Second, the signal acceptance increases with $m_{a}$. And third, the background distribution is not flat but peaks around 40 GeV ; see the upper panel of Fig. 3 .

### 3.3. Six muon final state

The ALP coupling $c_{2}^{\mu}$ induces a Higgs decay to six muons via the $h \rightarrow a a \mu \mu$ channel, where the ALP subsequently decays to muons.

Several ATLAS and CMS searches are sensitive to $6 \mu$ final states. However, as discussed in Sec. 2, even in background-free regions testing $\mathcal{O}(1) c_{2}^{\mu}$ couplings for $f_{a}=1 \mathrm{TeV}$ requires a signal acceptance of at least $10 \%$. Analyses that require at least four muons are either insensitive, as a result of the larger background contributions [77], or they require cuts on the lepton transverse momenta and sufficient energy, which render their signal acceptance below $10 \%$ [71,78].

Analyses targeting a six-muon final state are currently focusing on specific resonances producing di-muon pairs. Ref. [79] explores the production of three $Z$ bosons in final states with six leptons. Since the leptons from $Z$ boson decays typically have relatively high transverse momenta, the analysis selects events in which the scalar sum of the $p_{T}$ of all leptons is above 250 GeV , where only about $2 \%$ of our signal events survive. Ref. [80] concentrates on sixmuon events ensuing from the production of three $J / \Psi$ mesons. The cut on the invariant mass of each opposite-sign pair of muons, $m_{\mu \mu} \in[2.9,3.3] \mathrm{GeV}$, reduces our signal to insignificant levels (for the range of $m_{a}$ of interest). On the positive side, the search provides evidence that the high-level trigger, together with mild offline data requirements, can efficiently select multi-muon events with muon transverse momenta as low as 6 GeV . Our following analysis relies on this fact.

We define our $h \rightarrow 6 \mu$ signal region through the minimal cuts:
$p_{T}^{\mu}>5 \mathrm{GeV}, \quad m_{6 \mu} \in[110,140] \mathrm{GeV}$.
Potential backgrounds from low-mass $J / \psi$ or $\Upsilon$ resonances decaying to two muons are suppressed by the additional cut on the six muon invariant mass. At $\sqrt{s}=13 \mathrm{TeV}$, the SM cross section for $6 \mu$ production (before requiring $m_{6 \mu} \in[110,140] \mathrm{GeV}$ ) is only 0.24 ab . Therefore, we can safely neglect the SM background for LHC luminosities below $3 \mathrm{ab}^{-1}$.

We estimate the experimental efficiency of the $6 \mu$ final state based on the $4 \mu$ analysis in Ref. [69]. We approximate the total efficiency per lepton as $93 \%$ and exponentiate this with six. The total deduced efficiency is $64 \%$. To determine the ALP mass from


Fig. 4. Top: $m_{12}$ distribution in the $h \rightarrow a a \mu \mu \rightarrow 6 \mu$ analysis for different ALP masses at $\mathcal{L}=137 \mathrm{fb}^{-1}$. Bottom: Upper $95 \%$ CLs limits on $c_{2}^{\mu} / f_{a}^{2}$ for different luminosities.


Fig. 5. Feynman diagrams for $p p \rightarrow 4 \mu 2 j$ involving a five-point interaction of two quarks, two ALPs and the Higgs.
the $6 \mu$ analysis, we assign opposite-sign lepton pairs such that the difference $\Delta m=m_{12}-m_{34}$ between the invariant mass of two of these pairs is minimal, with $m_{12}>m_{34}$. The resulting $m_{12}$ distribution for different ALP masses, which peaks at $m_{12} \sim m_{a}$, is shown in the top panel of Fig. 4.

We use the $m_{12}$ distribution to constrain $c_{2}^{\mu} / f_{a}^{2}$ and find the following $95 \%$ CLs limits at $\mathcal{L}=137 \mathrm{fb}^{-1}$ :
$\frac{c_{2}^{\mu}}{f_{a}^{2}}<\begin{aligned} & 6.5 / \mathrm{TeV}^{2} \\ & 320 / \mathrm{TeV}^{2}\end{aligned}$ for $m_{s}=10 \mathrm{GeV}$ for $m_{s}=50 \mathrm{GeV}$.
In the bottom panel of Fig. 4, we show the limits on $c_{2}^{\mu} / f_{a}^{2}$ for different ALP masses and luminosities. For high ALP masses, the limits on $c_{2}^{\mu}$ become very weak due to the significant phase-space suppression. Since this channel is background free, the limits scale with $\mathcal{L}^{-1 / 2}$.

### 3.4. Four muon two jet final state

Finally, we consider a five-point interaction of two ALPs with the Higgs boson and two quarks. The corresponding Wilson coefficient $c_{2}^{q}$ denotes a universal coupling to both up-type and down-type second-generation quarks. Assuming that the ALPs decay to muons, a non-zero $c_{2}^{q}$ coupling produces a $4 \mu 2 j$ signature at the LHC via two different channels shown in Fig. 5: First, the


Fig. 6. Top: $M_{\mathrm{OSSF}}^{20}$ distribution in the $p p \rightarrow h a a \rightarrow 4 \mu 2 j$ analysis for backgrounds, CMS data [77] and ALP signal for different masses ( $\mathcal{L}=137 \mathrm{fb}^{-1}$ ). Bottom: Upper $95 \%$ CLs limits on $c_{2}^{q} / f_{a}^{2}$ at different luminosities. Observed limits are marked with an "x".
signature can result from a Higgs decaying into two ALPs and two jets $p p \rightarrow h \rightarrow 2 a 2 j$. And second, the coupling to secondgeneration quarks on the production side induces the $p p \rightarrow$ haa (with $h \rightarrow j j$ ) channel. The Higgs decay channel $h \rightarrow a a j j$ suffers from a large phase-space suppression and is only open for ALP masses $m_{a} \lesssim 36 \mathrm{GeV}$ when we require the two jets to have transverse momenta greater than 25 GeV . Since the $h$ and haa production channels do not interfere, we can distinguish their respective cross sections and find that the haa cross section is 250 $\left(24 \times 10^{6}\right)$ times larger than the $h$ channel cross section for an ALP with mass $m_{a}=15 \mathrm{GeV}\left(m_{a}=30 \mathrm{GeV}\right)$. Focusing on the Higgs decay region only, for instance by cutting on the invariant mass of the 4 -muon-2-jets system would thus significantly weaken the limits on $c_{2}^{q} / f_{a}^{2}$. Here, we investigate a generic $4 \mu 2 j$ ALP signature and present an analysis focusing on Higgs decays only in Appendix A.

The CMS search for beyond the SM phenomena in multi-lepton final states [77] provides a signal region with good sensitivity to our signal process. ${ }^{3}$ Since the Higgs boson predominantly decays to bottom quarks, the final-state jets in our haa analysis are likely to be $b$ jets. Therefore, we can recast the $4 L(\mu \mu) 1 B$ signal region of the CMS search, which requires four leptons and at least one $b$-tagged jet. To suppress the $Z Z+$ jets background, the invariant mass of OSSF lepton pairs must not be within 15 GeV of the $Z$ boson mass:

[^3]

Fig. 7. Two-loop contributions to the anomalous magnetic moment of the muon and their effective one-loop contribution after integrating out the Higgs boson.

$$
\begin{gather*}
N_{\ell} \geq 4, \quad N_{\mathrm{OSSF}} \geq 1, \quad p_{T}^{\ell}>10 \mathrm{GeV}, \\
\left|\eta^{\mu(e)}\right|<2.4(2.5), \quad p_{T}^{\mu_{1}\left(e_{1}\right)}>26(35) \mathrm{GeV}, \\
\left|m_{\ell^{+} \ell^{-}}-m_{Z}\right|<15 \mathrm{GeV}, \quad M_{\mathrm{OSSF}}^{20}(\mu \mu) \in[12,77] \mathrm{GeV},  \tag{13}\\
p_{T}^{j}>30 \mathrm{GeV}, \quad\left|\eta^{j}\right|<2.1, \quad N_{b} \geq 1, \\
\Delta R\left(\ell_{i}, \ell_{j} / j\right)>0.4, \quad m_{\ell \ell, S F}>12 \mathrm{GeV} .
\end{gather*}
$$

The variable $M_{\mathrm{OSSF}}^{20}$ denotes the invariant mass of the two OSSF leptons with an invariant mass closest to the target mass of 20 GeV . Since the same target mass of 20 GeV is applied for the pairing of the OSSF leptons, high-mass ALPs can contribute to more than one bin of the $M_{\text {OSSF }}^{20}$ distribution when leptons from different ALP decays are paired. We show the distribution of $M_{\text {OSSF }}^{20}$ for the SM backgrounds and different mass ALP signals in the top panel of Fig. 6. The background, which is dominated by $t t \bar{Z}$ and $Z Z$ production, is directly taken from the HepData [82] entry for Fig. 10e of Ref. [77]. We use the same reference for the systematic uncertainty on the background distribution, which is approximately $25 \%$ in each bin. We have estimated the detector efficiency to be $64 \%$, analogous to the $6 \mu$ final state. To ensure that the EFT description is valid, we require the invariant mass of all selected muons and jets to be smaller than 1 TeV .

In Fig. 6, we present the limits on $c_{2}^{q} / f_{a}^{2}$ as a function of the ALP mass for different luminosities. The limits for low $m_{a}$ are weaker as a result of the decreased signal acceptance, in particular for the muon transverse momentum cuts. At $\mathcal{L}=137 \mathrm{fb}^{-1}$ luminosity the observed (expected) 95\% CLs limits are:

$$
\frac{c_{2}^{q}}{f_{a}^{2}}<\begin{array}{ll}
1.8(1.8) / \mathrm{TeV}^{2} & \text { for } m_{s}=15 \mathrm{GeV}  \tag{14}\\
0.5(0.5) / \mathrm{TeV}^{2} & \text { for } m_{s}=50 \mathrm{GeV}
\end{array}
$$

## 4. Muon anomalous magnetic moment

Measurements of the anomalous magnetic moment of the muon $a_{\mu}$ indicate a difference between its experimental value and its SM prediction. The most recent experimental average of the results by the BNL E821 and FNAL collaborations finds a discrepancy of $\Delta a_{\mu}^{\exp }=a_{\mu}^{\exp }-a_{\mu}^{\text {SM }}=(251 \pm 59) \times 10^{-11}$ [83].

Following Ref. [84], this discrepancy could be explained if the Wilson coefficient of the dipole operator
$\mathcal{O}_{\mu \nu}=\overline{\mu_{L}} \sigma^{\mu \nu} \mu_{R} A_{\mu \nu}$
fulfils $c_{\mu \gamma} \sim 2 \times 10^{-9} \mathrm{GeV}^{-1}$ at low energies, $E \sim \mathrm{GeV}$.
Several works focus on the explanation of the muon $g-$ 2 anomaly by loop effects of (scalar or pseudo-scalar) axions with couplings to photons and muons ${ }^{4}$ [17,18,36,85-89]. While these couplings exist in the model considered here, we do not obtain new bounds on $c_{a A}$ and $c_{a \mu}$. Therefore, we not obtain new constraints on previous explanations of the muon anomalous magnetic moment. However, within our framework, there exist additional two-loop contributions to the muon $g-2$ from the two diagrams shown in Fig. 7. Both diagrams can be matched to a one-loop diagram including an effective four-muon interaction after integrating

[^4]

Fig. 8. One-loop $h-a$ contribution resulting in an effective four-muon coupling.
out the Higgs boson. Therefore, we can estimate the dominant (large-logarithmic) contribution to $C_{\mu \gamma}$ from its running induced by the four-fermion term
$\mathcal{O}_{4 \mu}=(\bar{\mu} \mu)(\bar{\mu} \mu)$.
An explicit computation of the loop involving $h$ and $a$, see Fig. 8, in the $\overline{\mathrm{MS}}$ scheme to zero external momentum and taking $m_{h} \gg m_{a}$ gives:

$$
\begin{align*}
i \mathcal{M}_{\text {loop }} & =i^{2}\left(c_{1}^{\mu} / f_{a}\right)^{2} \int \frac{d^{D}}{(2 \pi)^{D}} \frac{i}{k^{2}-m_{h}^{2}} \frac{i}{k^{2}} \times \text { four-fermion } \\
& =\frac{i\left(c_{1}^{\mu} / f_{a}\right)^{2}}{16 \pi^{2}} \times \text { four-fermion } \tag{15}
\end{align*}
$$

with $D=4-2 \epsilon$. We use "four-fermion" to label the product of external spinors. We can match this to tree-level contribution induced by the operator $\mathcal{O}_{4 \mu}$ :
$i \mathcal{M}_{\text {tree }}=2 c_{4 \mu} \times$ four-fermion.
Thus, at energies $E \ll m_{h}$, the effects of the loop in Fig. 8 can be captured by the effective operator $\mathcal{O}_{4 \mu}$ whose Wilson coefficient is $c_{4 \mu}=\left(c_{1}^{\mu} / f_{a}\right)^{2} /\left(32 \pi^{2}\right)$. We have cross checked the matching using Matchmakereft [90]. The operator $\mathcal{O}_{4 \mu}$ mixes into $\mathcal{O}_{\mu \gamma}$ at one loop; see Ref. [91] for the corresponding anomalous dimensions. At energies $E \sim m_{\mu}$, we have:
$c_{\mu \gamma} \sim \frac{e m_{\mu}\left(c_{1}^{\mu} / f_{a}\right)^{2}}{128 \pi^{4}} \log \frac{m_{h}}{m_{\mu}} \sim 10^{-5}\left(c_{1}^{\mu} / f_{a}\right)^{2} \mathrm{GeV}$,
with $e=\sqrt{4 \pi \alpha}$. Thus, the two-loop contributions in Fig. 7 have the correct sign to potentially explain the anomalous $\Delta a_{\mu}$ with a value of $c_{1}^{\mu} / f_{a} \sim 11 /(1 \mathrm{TeV})$. This is well within the reach of our suggested analyses and with the available $137 \mathrm{fb}^{-1}$ data we can already rule out that $c_{1}^{\mu}$ alone could explain the observed discrepancy in $\Delta a_{\mu}$. Obviously, other operators could also indirectly contribute to $\Delta a_{\mu}$. Precisely because of this, direct searches like the ones we propose, are necessary and complementary.

## 5. Conclusions

This paper investigates search strategies for a light pseudoscalar particle in the $m_{a} \in[10,50] \mathrm{GeV}$ range coupling to the Higgs boson and fermions. These interactions induce rare new decays of the Higgs boson. We have focused on second-generation leptons and have considered the possibility that the pseudo-scalar decays into two muons or two photons.

The expected upper limits on the effective four-point interactions at $3 \mathrm{ab}^{-1}$ luminosity and assuming $f_{a}=1 \mathrm{TeV}$ are of the order of $c_{1}^{\mu} \lesssim 0.01$ for the $h \rightarrow 4 \mu$ analysis and $c_{1}^{\mu} \lesssim 0.03$ for the $h \rightarrow 2 \mu 2 \gamma$ analysis. Couplings inducing a five-particle interaction are less constrained and have a stronger mass dependence. We find $c_{2}^{\mu} \lesssim 1-80$ for the $h \rightarrow 6 \mu$ channel and $c_{2}^{q} \lesssim 2.5-6.0$ for the $h \rightarrow 4 \mu 2 j$ channel (see the Appendix A for this later one), depending on the ALP mass. In terms of Higgs branching ratios these numbers correspond to $1.3 \times 10^{-5}, 2.0 \times 10^{-6}, 3.0 \times 10^{-9}$ and $5.4 \times 10^{-9}$ respectively.

For $h \bar{q} q a a$ couplings, it turns out that despite the PDF suppression, the channel $p p \rightarrow a a h$ is by far more constraining than the Higgs decay $p p \rightarrow h \rightarrow q q a a$. Using the $p p \rightarrow a a h \rightarrow 4 \mu 2 j$ channel, the upper limits on $c_{2}^{q}$ can be reduced to $c_{2}^{q} \lesssim 0.2-1$ at $3 \mathrm{ab}^{-1}$ luminosity and assuming $f_{a}=1 \mathrm{TeV}$.

There are four channels involving contact interactions of the Higgs with two fermions and ALPs with ALP decays to muons and photons which we did not consider here:

- $h \rightarrow a q q, a \rightarrow \mu \mu(2 \mu 2 j)$ or $p p \rightarrow h a \rightarrow b \bar{b} \mu \mu$,
- $h \rightarrow a q q, a \rightarrow \gamma \gamma(2 \gamma 2 j)$ or $p p \rightarrow h a \rightarrow b \bar{b} \gamma \gamma$,
- $h \rightarrow a a \mu \mu, a \rightarrow \gamma \gamma(4 \gamma 2 \mu)$,
- $h \rightarrow a a q q, a \rightarrow \gamma \gamma(4 \gamma 2 q)$ or $p p \rightarrow h a a \rightarrow b \bar{b}+4 \gamma$.

We have neglected these channels due to the poorer reconstruction efficiency of low-energy photons compared to muons and the extensive backgrounds for events with two muons/photons and jets.

Dedicated searches for events with two leptons and more than two photons currently do not exist. The most promising search for a $2 \mu 4 \gamma$ signature would thus still be the $Z \gamma \gamma$ search in Ref. [74], which already proved to have too stringent requirements on the photons for the $2 \mu 2 \gamma$ signature studied in Sec. 3.2.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgements

This work has been supported by the Newton International Fellowship Alumni follow-on funding 2020/21 under grant number AL211013/4. A.B. gratefully acknowledges support from the Alexander-von-Humboldt foundation as a Feodor Lynen Fellow. MC is supported by the Spanish MINECO under the Ramón y Cajal programme as well as by the SRA under grant number PID2019-106087GB-C21/C22 (10.13039/501100011033), and by the Junta de Andalucía grants FQM 101, A-FQM-211-UGR18 and P18-FR-4314 (FEDER).

## Appendix A. Four muon and two jet final state in Higgs decays

For completeness, we present an analysis for Higgs decays to two ALPs and two jets induced by a non-zero $c_{2}^{q}$. We select the Higgs decay region by requiring that the invariant mass of the four muons and two jets is within 25 GeV of the Higgs mass. While this selection ensures a better validity of the EFT approach, it also significantly decreases the signal cross section as discussed in Section 3.4.

As our signal is characterized by two resonant $a \rightarrow \mu \mu$ decays, we assign the two opposite-sign muon pairs such that the difference between their invariant mass, $\left(m_{12}-m_{34}\right) /\left(m_{12}+m_{34}\right)$, is minimal and require $\Delta m=\left|m_{12}-m_{34}\right|<4 \mathrm{GeV}$. The total signal selection is

$$
\begin{gathered}
N_{\mu} \geq 4, \quad N_{\mathrm{OSSF}} \geq 2, \quad p_{T}^{\mu}>5 \mathrm{GeV}, \quad m_{\mu \mu}>12 \mathrm{GeV}, \\
p_{T}^{j}>25 \mathrm{GeV}, \quad\left|\eta^{j}\right|<4.5, \quad N_{j} \geq 2 \\
\left|m_{\mu^{+} \mu^{-}}-m_{Z}\right|<15 \mathrm{GeV}, \quad \Delta m=\left|m_{12}-m_{34}\right|<4 \mathrm{GeV} \\
m_{4 \mu 2 j} \in[100,150] \mathrm{GeV} .
\end{gathered}
$$




Fig. 9. Top: Invariant mass distribution of the heavier opposite-sign muon pair $m_{12}$ in the $h \rightarrow a a j j$ analysis at $\mathcal{L}=137 \mathrm{fb}^{-1}$. Bottom: Upper $95 \%$ CLs limits on $c_{2}^{q} / f_{a}^{2}$.

The dominant backgrounds from $Z Z+$ jets, $Z \rightarrow 2 \mu$ production, $Z+$ jets, $Z \rightarrow 4 \mu$ production and $h+$ jets, $h \rightarrow 4 \mu$ production are significantly suppressed by the cuts on $m_{4 \mu 2 j}$ and $\Delta m$. After the selection cuts we expect less than one background event at the HL-LHC and we therefore consider our signal region as background free.

To set limits we use the invariant mass pair of the heavier opposite-sign muon pair $m_{12}$; see the upper panel of Fig. 9. We find the following upper CLs limits on $c_{2}^{2} / f_{a}^{2}$ at $\mathcal{L}=137 \mathrm{fb}^{-1}$ :
$\frac{c_{2}^{q}}{f_{a}^{2}}<\begin{aligned} & 12 / \mathrm{TeV}^{2} \\ & 27 / \mathrm{TeV}^{2}\end{aligned}$ for $m_{s}=15 \mathrm{GeV}$ for $m_{s}=50 \mathrm{GeV}$.
In the bottom panel of Fig. 9, we present the $95 \%$ CLs limits on $c_{2}^{q} / f_{a}^{2}$ at different ALP masses and luminosities. Comparing these limits to those shown in Fig. 6 for a generic $4 \mu 2 j$ signature analysis, we find that the limits obtained by focusing on Higgs decays only are worse by a factor 6 - 40 depending on the ALP mass.

## References

[1] R.D. Peccei, H.R. Quinn, Phys. Rev. Lett. 38 (1977) 1440.
[2] R.D. Peccei, H.R. Quinn, Phys. Rev. D 16 (1977) 1791.
[3] S. Weinberg, Phys. Rev. Lett. 40 (1978) 223.
[4] F. Wilczek, Phys. Rev. Lett. 40 (1978) 279.
[5] B. Gripaios, A. Pomarol, F. Riva, J. Serra, J. High Energy Phys. 04 (2009) 070, arXiv:0902.1483 [hep-ph].
[6] B. Gripaios, M. Nardecchia, T. You, Eur. Phys. J. C 77 (2017) 28, arXiv:1605. 09647 [hep-ph].
[7] M. Chala, G. Durieux, C. Grojean, L. de Lima, O. Matsedonskyi, J. High Energy Phys. 06 (2017) 088, arXiv:1703.10624 [hep-ph].
[8] F. Wilczek, Phys. Rev. Lett. 49 (1982) 1549.
[9] Y. Chikashige, R.N. Mohapatra, R.D. Peccei, Phys. Lett. B 98 (1981) 265.
[10] J. Preskill, M.B. Wise, F. Wilczek, Phys. Lett. B 120 (1983) 127.
[11] L.F. Abbott, P. Sikivie, Phys. Lett. B 120 (1983) 133.
[12] M. Dine, W. Fischler, Phys. Lett. B 120 (1983) 137.
[13] A. Davidson, K.C. Wali, Phys. Rev. Lett. 48 (1982) 11.
[14] Y. Ema, K. Hamaguchi, T. Moroi, K. Nakayama, J. High Energy Phys. 01 (2017) 096, arXiv:1612.05492 [hep-ph].
[15] L. Calibbi, F. Goertz, D. Redigolo, R. Ziegler, J. Zupan, Phys. Rev. D 95 (2017) 095009, arXiv:1612.08040 [hep-ph].
[16] P.W. Graham, D.E. Kaplan, S. Rajendran, Phys. Rev. Lett. 115 (2015) 221801, arXiv:1504.07551 [hep-ph].
[17] M. Chala, G. Guedes, M. Ramos, J. Santiago, Eur. Phys. J. C 81 (2021) 181, arXiv: 2012.09017 [hep-ph].
[18] M. Bauer, M. Neubert, S. Renner, M. Schnubel, A. Thamm, J. High Energy Phys. 04 (2021) 063, arXiv:2012.12272 [hep-ph].
[19] A.M. Galda, M. Neubert, S. Renner, J. High Energy Phys. 06 (2021) 135, arXiv: 2105.01078 [hep-ph].
[20] J. Bonilla, I. Brivio, M.B. Gavela, V. Sanz, J. High Energy Phys. 11 (2021) 168, arXiv:2107.11392 [hep-ph].
[21] R. Povey, J. Hartnett, M. Tobar, Phys. Rev. D 82 (2010) 052003, arXiv:1003.0964 [hep-ex].
[22] A. Wagner, et al., ADMX, Phys. Rev. Lett. 105 (2010) 171801, arXiv:1007.3766 [hep-ex].
[23] M. Betz, F. Caspers, M. Gasior, M. Thumm, S.W. Rieger, Phys. Rev. D 88 (2013) 075014, arXiv:1310.8098 [physics.ins-det].
[24] J.D. Bjorken, R. Essig, P. Schuster, N. Toro, Phys. Rev. D 80 (2009) 075018, arXiv: 0906.0580 [hep-ph].
[25] S. Andreas, C. Niebuhr, A. Ringwald, Phys. Rev. D 86 (2012) 095019, arXiv:1209. 6083 [hep-ph].
[26] R. Essig, et al., in: Community Summer Study 2013: Snowmass on the Mississippi, 2013, arXiv:1311.0029 [hep-ph].
[27] L. Harland-Lang, J. Jaeckel, M. Spannowsky, Phys. Lett. B 793 (2019) 281, arXiv: 1902.04878 [hep-ph].
[28] M.B. Gavela, R. Houtz, P. Quilez, R. Del Rey, O. Sumensari, Eur. Phys. J. C 79 (2019) 369, arXiv:1901.02031 [hep-ph].
[29] L. Calibbi, D. Redigolo, R. Ziegler, J. Zupan, J. High Energy Phys. 09 (2021) 173, arXiv:2006.04795 [hep-ph].
[30] J. Martin Camalich, M. Pospelov, P.N.H. Vuong, R. Ziegler, J. Zupan, Phys. Rev. D 102 (2020) 015023, arXiv:2002.04623 [hep-ph].
[31] M. Bauer, M. Neubert, S. Renner, M. Schnubel, A. Thamm, arXiv:2110.10698 [hep-ph], 2021.
[32] A. Carmona, C. Scherb, P. Schwaller, J. High Energy Phys. 08 (2021) 121, arXiv: 2101.07803 [hep-ph].
[33] J. Jerhot, B. Döbrich, F. Ertas, F. Kahlhoefer, T. Spadaro, arXiv:2201.05170 [hepph], 2022.
[34] J. Jaeckel, M. Spannowsky, Phys. Lett. B 753 (2016) 482, arXiv:1509.00476 [hepph].
[35] K. Mimasu, V. Sanz, J. High Energy Phys. 06 (2015) 173, arXiv:1409.4792 [hepph ].
[36] M. Bauer, M. Neubert, A. Thamm, J. High Energy Phys. 12 (2017) 044, arXiv: 1708.00443 [hep-ph].
[37] N. Craig, A. Hook, S. Kasko, J. High Energy Phys. 09 (2018) 028, arXiv:1805. 06538 [hep-ph].
[38] J. Jaeckel, M. Jankowiak, M. Spannowsky, Phys. Dark Universe 2 (2013) 111, arXiv:1212.3620 [hep-ph].
[39] S. Knapen, T. Lin, H.K. Lou, T. Melia, Phys. Rev. Lett. 118 (2017) 171801, arXiv: 1607.06083 [hep-ph].
[40] I. Brivio, M.B. Gavela, L. Merlo, K. Mimasu, J.M. No, R. del Rey, V. Sanz, Eur. Phys. J. C 77 (2017) 572, arXiv:1701.05379 [hep-ph].
[41] M. Bauer, M. Neubert, A. Thamm, Phys. Rev. Lett. 119 (2017) 031802, arXiv: 1704.08207 [hep-ph].
[42] G. Alonso-Álvarez, M.B. Gavela, P. Quilez, Eur. Phys. J. C 79 (2019) 223, arXiv: 1811.05466 [hep-ph].
[43] J. Ebadi, S. Khatibi, M. Mohammadi Najafabadi, Phys. Rev. D 100 (2019) 015016, arXiv:1901.03061 [hep-ph].
[44] M.B. Gavela, J.M. No, V. Sanz, J.F. de Trocóniz, Phys. Rev. Lett. 124 (2020) 051802, arXiv:1905.12953 [hep-ph].
[45] R.O. Coelho, V.P. Goncalves, D.E. Martins, M.S. Rangel, Phys. Lett. B 806 (2020) 135512, arXiv:2002.06027 [hep-ph].
[46] G. Haghighat, D. Haji Raissi, M. Mohammadi Najafabadi, Phys. Rev. D 102 (2020) 115010, arXiv:2006.05302 [hep-ph].
[47] V.P. Goncalves, W.K. Sauter, Phys. Lett. B 811 (2020) 135981, arXiv:2006.16716 [hep-ph].
[48] A. Carmona, F. Elahi, C. Scherb, P. Schwaller, arXiv:2202.09371 [hep-ph], 2022.
[49] G.G. Raffelt, Lect. Notes Phys. 741 (2008) 51, arXiv:hep-ph/0611350.
[50] S. Aune, et al., CAST, Phys. Rev. Lett. 107 (2011) 261302, arXiv:1106.3919 [hepex].
[51] J.S. Lee, arXiv:1808.10136 [hep-ph], 2018.
[52] J.H. Chang, R. Essig, S.D. McDermott, J. High Energy Phys. 09 (2018) 051, arXiv: 1803.00993 [hep-ph].
[53] J. Jaeckel, L.J. Thormaehlen, Phys. Rev. D 100 (2019) 123020, arXiv:1908.10878 [astro-ph.SR].
[54] G. Alonso-Álvarez, R.S. Gupta, J. Jaeckel, M. Spannowsky, J. Cosmol. Astropart. Phys. 03 (2020) 052, arXiv:1911.07885 [hep-ph].
[55] F. Ertas, F. Kahlhoefer, J. High Energy Phys. 07 (2020) 050, arXiv:2004.01193 [hep-ph].
[56] J.A. Evans, P. Tanedo, M. Zakeri, J. High Energy Phys. 01 (2020) 028, arXiv:1910. 07533 [hep-ph].
[57] H. Davoudiasl, R. Marcarelli, N. Miesch, E.T. Neil, Phys. Rev. D 104 (2021) 055022, arXiv:2105.05866 [hep-ph].
[58] A. Alves, A.G. Dias, D.D. Lopes, J. High Energy Phys. 10 (2021) 012, arXiv:2105. 01095 [hep-ph].
[59] J. Liu, N. McGinnis, C.E.M. Wagner, X.-P. Wang, J. High Energy Phys. 05 (2021) 138, arXiv:2102.10118 [hep-ph].
[60] M. Cepeda, S. Gori, V.M. Outschoorn, J. Shelton, arXiv:2111.12751 [hep-ph], 2021.
[61] J. Liu, C.E.M. Wagner, X.-P. Wang, J. High Energy Phys. 03 (2019) 008, arXiv: 1810.11028 [hep-ph].
[62] G. Anamiati, O. Castillo-Felisola, R.M. Fonseca, J.C. Helo, M. Hirsch, J. High Energy Phys. 12 (2018) 066, arXiv:1806.07264 [hep-ph].
[63] D. de Florian, et al., LHC Higgs Cross Section Working Group, 2/2017 (2016), https://doi.org/10.23731/CYRM-2017-002, arXiv:1610.07922 [hep-ph].
[64] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos, B. Mistlberger, J. High Energy Phys. 05 (2016) 058, arXiv:1602.00695 [hep-ph].
[65] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, T. Reiter, Comput. Phys. Commun. 183 (2012) 1201, arXiv:1108.2040 [hep-ph].
[66] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.S. Shao, T. Stelzer, P. Torrielli, M. Zaro, J. High Energy Phys. 07 (2014) 079, arXiv:1405. 0301 [hep-ph].
[67] T. Sjostrand, S. Mrenna, P.Z. Skands, Comput. Phys. Commun. 178 (2008) 852, arXiv:0710.3820 [hep-ph].
[68] A. Buckley, J. Butterworth, D. Grellscheid, H. Hoeth, L. Lonnblad, J. Monk, H. Schulz, F. Siegert, Comput. Phys. Commun. 184 (2013) 2803, arXiv:1003.0694 [hep-ph].
[69] arXiv:2111.01299 [hep-ex], 2021.
[70] A.M. Sirunyan, et al., CMS, Phys. Lett. B 792 (2019) 345, arXiv:1808.03684 [hepex].
[71] M. Aaboud, et al., ATLAS, Phys. Rev. D 98 (2018) 032009, arXiv:1804.03602 [hep-ex].
[72] A.M. Sirunyan, et al., CMS, Eur. Phys. J. C 81 (2021) 629, arXiv:2010.02984 [hepex].
[73] A.L. Read, J. Phys. G 28 (2002) 2693.
[74] A. Tumasyan, et al., CMS, J. High Energy Phys. 10 (2021) 174, arXiv:2105.12780 [hep-ex].
[75] G. Aad, et al., ATLAS, Phys. Lett. B 809 (2020) 135754, arXiv:2005.05382 [hepex].
[76] A.M. Sirunyan, et al., CMS, J. High Energy Phys. 11 (2018) 152, arXiv:1806. 05996 [hep-ex].
[77] A.M. Sirunyan, et al., CMS, J. High Energy Phys. 03 (2020) 051, arXiv:1911.04968 [hep-ex].
[78] A.M. Sirunyan, et al., CMS, Phys. Rev. D 100 (2019) 052003, arXiv:1905.10853 [hep-ex].
[79] A.M. Sirunyan, et al., CMS, Phys. Rev. Lett. 125 (2020) 151802, arXiv:2006. 11191 [hep-ex].
[80] A. Tumasyan, et al., CMS, arXiv:2111.05370 [hep-ex], 2021.
[81] A.M. Sirunyan, et al., CMS, J. High Energy Phys. 03 (2020) 065, arXiv:1910. 11634 [hep-ex].
[82] E. Maguire, L. Heinrich, G. Watt, J. Phys. Conf. Ser. 898 (2017) 102006, arXiv: 1704.05473 [hep-ex].
[83] B. Abi, et al., Muon g-2, Phys. Rev. Lett. 126 (2021) 141801, arXiv:2104.03281 [hep-ex].
[84] J. Aebischer, W. Dekens, E.E. Jenkins, A.V. Manohar, D. Sengupta, P. Stoffer, J. High Energy Phys. 07 (2021) 107, arXiv:2102.08954 [hep-ph].
[85] D. Chang, W.-F. Chang, C.-H. Chou, W.-Y. Keung, Phys. Rev. D 63 (2001) 091301, arXiv:hep-ph/0009292.
[86] W.J. Marciano, A. Masiero, P. Paradisi, M. Passera, Phys. Rev. D 94 (2016) 115033, arXiv:1607.01022 [hep-ph].
[87] M. Bauer, M. Neubert, S. Renner, M. Schnubel, A. Thamm, Phys. Rev. Lett. 124 (2020) 211803, arXiv:1908.00008 [hep-ph].
[88] C. Cornella, P. Paradisi, O. Sumensari, J. High Energy Phys. 01 (2020) 158, arXiv: 1911.06279 [hep-ph].
[89] M.A. Buen-Abad, J. Fan, M. Reece, C. Sun, J. High Energy Phys. 09 (2021) 101, arXiv:2104.03267 [hep-ph].
[90] A. Carmona, A. Lazopoulos, P. Olgoso, J. Santiago, SciPost Phys. 12 (2022) 198, arXiv:2112.10787 [hep-ph].
[91] E.E. Jenkins, A.V. Manohar, P. Stoffer, J. High Energy Phys. 01 (2018) 084, arXiv: 1711.05270 [hep-ph].


[^0]:    * Corresponding author.

    E-mail addresses: abiekoet@uni-mainz.de, biekoetter.anke@gmail.com
    (A. Biekötter), mikael.chala@ugr.es (M. Chala), michael.spannowsky@durham.ac.uk (M. Spannowsky).

[^1]:    ${ }^{1}$ The Higgs-ALP-fermion interactions in the leading-order chiral Lagrangian reads [40]:
    $L_{\mathrm{LO}}=-i \sqrt{2 v} \frac{a}{f_{a}}\left(1+a_{1} \frac{h}{v}+a_{2} \frac{h^{2}}{v^{2}}+\cdots\right) \bar{\psi} \mathbf{U} \sigma^{3} Y \psi+$ h.c.,
    with $Y$ representing the Yukawa coupling of $\psi ; \mathbf{U}$ is defined as $\mathbf{U}=\exp \left(i \sigma_{a} \pi^{a} / v\right)$ with $\pi^{a}$ denoting the gauge boson longitudinal degrees of freedom and $\sigma_{a}$ the Pauli matrices. The crucial point is that the parameters $a_{1}, a_{2}$, etc. are all independent.

[^2]:    2 Searches for Higgs decays in the $h \rightarrow Z \gamma \rightarrow \ell \ell \gamma$ channel also offer some sensitivity to our $h \rightarrow 2 \mu 2 \gamma$ signal. However, due to the fact these analysis target events with a single photon, we cannot use $h \rightarrow Z \gamma$ searches to constrain the ALP mass which shows up as a resonance in the $m_{\gamma \gamma}$ spectrum. Moreover, the ATLAS search [75] uses a boosted decision tree (BDT) which makes it impossible to recast without detailed information on the BDT. Moreover, the CMS search [76] is currently only published for a low luminosity ( $\mathcal{L}=35.9 \mathrm{fb}^{-1}$ ).

[^3]:    ${ }^{3}$ Most experimental searches testing exotic signals in $4 \mu 2 j$ final states are insensitive to our signal since they apply relatively strong cuts on the lepton transverse momenta or on their effective mass [71,78]. The CMS search for pseudo-scalars decaying to a $Z$ boson and a Higgs, in events with four leptons and two jets [81] is insensitive to our ALP signal since it requires at least one electron.

[^4]:    4 The contribution from diagrams with only axion-muon couplings has the wrong sign to explain the experimental deviation from the SM prediction.

