CALT-TH-2023-001
DESY-2023-01670

Energetic long-lived particles in the CMS muon chambers

Andrea Mitridate,[1,2] Michele Papucci,[1] Christina W. Wang,[3] Cristián Peña,[4] and Si Xie[3,4]

[1] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*

[2] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*

[3] *California Institute of Technology, Pasadena, CA 91125, USA*

[4] *Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*

We present a recast in different benchmark models of the recent CMS search that uses the endcap muon detector system to identify displaced showers produced by decays of long-lived particles (LLPs). The exceptional shielding provided by the steel between the stations of the muon system drastically reduces the Standard Model background that limits other existing ATLAS and CMS searches. At the same time, by using the muon system as a sampling calorimeter, the search is sensitive to LLPs energies rather than masses. We show that, thanks to these characteristics, this new search approach is sensitive to LLPs masses even lighter than a GeV, and can be complementary to proposed and existing dedicated LLP experiments.

CONTENTS

I. INTRODUCTION

Despite its successes, the Standard Model (SM) of particle physics has several shortcomings. Most notably, it fails to explain the nature of Dark Matter, the baryon asymmetry in the Universe, and the origin of neutrino masses. Because of this, several extensions of the SM have been proposed, a common byproduct of which are long-lived particles (LLPs). This is especially true for *dark sectors* comprised of new particles with masses below the electroweak scale that interact with the SM through suppressed renormalizable couplings, heavy mediators, or both. In recent years, several searches for LLPs have been conducted by ATLAS, CMS, LHCb, and other dedicated LLP experiments and a vibrant search program is being developed [1–27]. However, most ATLAS and CMS searches are known to face challenges for LLPs with masses at or below a GeV. While some room for improvement may be present in existing analyses [14, 28–30], looking for tracks from displaced vertices will ultimately be limited by irreducible SM backgrounds. This is especially true in the long lifetime regime, where the few tracks produced by the decays of light LLPs are not enough to discriminate them against SM LLP decays such as those of K_L.

However, recently the CMS collaboration published a search [31] for neutral LLPs which uses the CMS endcap muon detectors (EMDs) as a sampling calorimeter. Thanks to the unique design of the CMS EMDs (made of stations of cathode strip chambers (CSCs) interleaved with steel return-yoke), LLPs decay products can induce hadronic and electromagnetic showers that give rise to high hit multiplicity in localized detector regions. Because of this, the LLPs signature tracks the LLP energy instead of its mass. This, together with the exceptional shielding provided by the CMS calorimeters and the steel in the front layers of the EMD, allows this search to be sensitive to LLPs with masses smaller than a GeV. Despite this search currently lacking a dedicated trigger to maximize its potential, its reach is competitive with many proposed dedicated LLP experiments [32–48] due to its large geometric acceptance.

The original CMS paper presented the results on a benchmark model motivated by the twin Higgs scenario, where the SM Higgs boson decays to a pair of neutral long-lived scalars, each of which decays in turn to a pair of bottom quarks, down quarks, or τ leptons. Masses of the scalar LLPs were probed down to 7 GeV. In this paper, we use the parameterized reconstruction and selection efficiency functions provided in the HEPData entry [49] of the CMS paper to recast the analysis in different benchmark models and explore its strengths and weaknesses. The goal of this recast is to inform future iterations of the CMS analysis and inform the choice of benchmarks for other proposed LLP experiments away from already covered regions of the parameter space.

The paper is structured as follows. In Section II, we briefly review the CMS analysis, we outline the generation and simulation framework, as well as the validation of the framework against the CMS result. In Section III, we discuss all the benchmark models that will be considered. Finally, in Section IV, we discuss the results of this paper.

II. ANALYSIS AND RECAST STRATEGY

In this section, we describe the details of the analysis. We start by summarizing the CMS analysis [31] and then we discuss and validate the recast procedure, including the event generation and detector simulation framework, signal selections, signal and background yield estimate, and statistical analysis to evaluate the upper limits.

A. CMS endcap muon system analysis

The missing transverse momentum, $p_{\mathrm{T}}^{\mathrm{miss}}$, is calculated by CMS by using only tracker and calorimeter information. Therefore, LLPs decaying beyond the calorimeters can produce large $p_{\mathrm{T}}^{\mathrm{miss}}$. Due to this feature and the lack of a dedicated trigger, events are triggered by requiring $p_{\mathrm{T}}^{\mathrm{miss}} > 120$ GeV. A further requirement of $p_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV is then applied offline. To further suppress the background from W and top quark production, the search also requires at least one jet from initial state radiation with transverse momentum $p_{\mathrm{T}} > 50$ GeV and pseudorapidity $|\eta| < 2.4$, and no leptons in the event.

By clustering CSC hits using the DBSCAN algorithm [50], the search then identifies displaced showers produced by LLPs decaying in the endcap muon system. To suppress background from punch-through jets and muon bremsstrahlung, clusters that are geometrically matched (*i.e.* within a cone of radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.4$) to jets (muons) with $p_T > 10\,(20)$ GeV are rejected. Furthermore, a number of active vetoes are applied to veto clusters with hits or track segments in the muon detector stations with the least amount of shielding. Finally, to suppress the muon bremsstrahlung background originating from muons beyond the detector acceptance, clusters with $|\eta| > 2.0$ are vetoed. Clusters are required to be consistent with an in-time interaction by restricting the cluster time ($-5.0 < t_{cluster} < 12.5$ ns). To reject clusters composed of hits from multiple bunch crossings, the root mean square spread of a cluster's hit times is required to be less than 20 ns.

A cut-based ID that distinguishes signal from background clusters is defined by using several features, including the cluster η position and the number of stations occupied by the cluster. The details required to simulate the cut-based ID efficiency have been provided in the HEPData entry of the CMS paper [49].

Finally, the number of hits in the cluster, N_{hits}, and the azimuthal angle between $p_{\mathrm{T}}^{\mathrm{miss}}$ and the cluster location, $\Delta\phi$, are used to make the final discrimination between signal and background. The signal is required to have large $N_{\mathrm{hits}} > 130$ and small $\Delta\phi < 0.75$. For the chosen signal the bulk of the $p_{\mathrm{T}}^{\mathrm{miss}}$ is produced by the LLP, while for the backgrounds, $\Delta\phi$ is independent of N_{hits}. The independence of the two variables for the background enables the use of the matrix (ABCD) method to predict the background yield. Using the matrix method and assuming no signal contribution, 2.0 ± 1.0 background events were predicted in the signal-enriched region, and

3 events were observed. No excess of events above the SM background was observed.

B. Event generation

We generated signal events using MadGraph5 v2.9.3 [51], and performed the parton shower and hadronization with Pythia v8.244 [52], while keeping the LLP stable. Samples with different jet multiplicities were merged according to the MLM algorithm [53, 54]. Generator-level cuts were applied to the events in order to increase the statistics in the phase space regions selected by $p_{\mathrm{T}}^{\mathrm{miss}}$ cut in the CMS analysis. Additional details on the samples, including the specific generator-level cuts, are given in Sec. III for each of the benchmark models considered in this work.

To efficiently decay the LLP, we used the fact that the reconstruction efficiency parameterization provided by the CMS search is spatially binned in a small number of regions with simple shapes (which are defined by the intersection of ranges in the radial direction, r, longitudinal direction, z, and pseudorapidity, η), for which the probability of decaying inside a region can be computed analytically. For a region determined by the conditions

$$\eta_0 \leq \eta \leq \eta_1, \qquad r_0 \leq r \leq r_1, \qquad z_0 \leq z \leq z_1, \tag{1}$$

the probability, P, to decay inside the region for a particle traveling with momentum p^μ and proper decay length $c\tau$ is

$$P = e^{-y_{min}} - e^{-y_{max}} \tag{2}$$

where we have defined

$$y_{min} = \frac{1}{\beta\gamma c\tau}\max\Big(\min\left(z_0\coth\eta, z_1\coth\eta\right), r_0\cosh\eta\Big), \tag{3}$$

$$y_{max} = \frac{1}{\beta\gamma c\tau}\min\Big(\max\left(z_0\coth\eta, z_1\coth\eta\right), r_1\cosh\eta, \beta\gamma c\, t_{cut}\left(\beta\gamma + \sqrt{1+(\beta\gamma)^2}\right)\Big), \tag{4}$$

with η being the pseudorapidity of the particle, and $\beta\gamma = |\vec{p}|/m$. We have further introduced a timing requirement such that $t_{decay} - d_{decay}/c < t_{cut}$ where t_{decay} and d_{decay} are the decay time and distance from the origin respectively.

The LLPs generated in the events were made to decay at fixed positions within each region using Pythia, as the presence of the decay vertex in a given region is the only geometrical information used by the detector simulation. For a given decay length $c\tau$, the probability for the LLP to decay in a given region was assigned to the event in the form of an event weight. Therefore, for each input event with an undecayed LLP, the decay program generated as many decayed events as the number of regions intersected by the LLP trajectory. We used the multi-weight capabilities of the HepMC event format to perform a scan in $c\tau$ without having to reprocess the events. In case there were multiple LLPs in the same event (as in the case of Higgs decays to pairs of dark vectors or scalars), we also included decays outside the signal regions. In the case of LLP decays in the inner

detector, where precise knowledge of the decay vertex position is used by the detector simulation, we generated the decay vertex position according to the decay probability distribution instead of keeping it fixed.

The events were subsequently passed to a simplified detector simulation based on DELPHES v3.4.2 [55] using the publicly available CMS configuration card for the reconstruction of prompt objects supplemented by a dedicated module, discussed in subsection II C, simulating the LLP decay reconstruction and selection using the information provided by the CMS search.

C. Detector simulation with dedicated Delphes modules

We based the fast simulation of the detector response to standard particle flow (PF) candidates on the CMS detector response provided by the CMS configuration card in the parametric DELPHES framework. The detector simulation of the hit clusters in the CSC endcap muon detector was based on a dedicated DELPHES module and class for the CSC cluster objects that we developed [56] based on the parameterized detector response functions provided in the HEPData entry [49]. Based on the recasting instructions provided in the HEPData entry, the simulation of cluster-level selection efficiencies was divided into three components.

The first component is cluster efficiency; which includes the cluster reconstruction efficiency, muon veto, active veto, time spread, and N_{hits} cut efficiency. This cluster efficiency is provided as a function of the LLP electromagnetic and hadronic energy in two separate LLP decay regions in the CSC detector. Building upon the existing `Efficiency` module that is already used by all other PF candidates, we implemented a dedicated `CscClusterEfficiency` module in DELPHES which encodes this parameterized function.

The second component is cluster identification efficiency. We implemented the `CscClusterID` module following the code function provided by the CMS HEPData entry.

The third and last component includes the model-dependent cluster time requirement, jet veto, and the $\Delta\phi$ requirement. The values of the three variables affected by these cuts were calculated using generator-level information and saved as part of the `CscCluster` class. Specifically, the cluster time was determined by calculating the LLP travel time from the production to the decay vertex in the lab frame. The jet veto was implemented by requiring no PF jets with $p_T > 10$ GeV within a cone of $\Delta = 0.4$ around the LLP. Finally, $\Delta\phi$ was computed as the azimuthal angle difference between the LLP momentum and MET using the result from the standard DELPHES simulation modules. All these requirements were then made at a later stage of the analysis workflow.

Finally, we modified the standard CMS configuration card from DELPHES to include the `CSCClusterEfficiency` and `CSCClusterID` module in the processing sequence. The modules require only generator-level LLP information and can be used for the recasting of the result for any other model. The implementation can be found in [56].

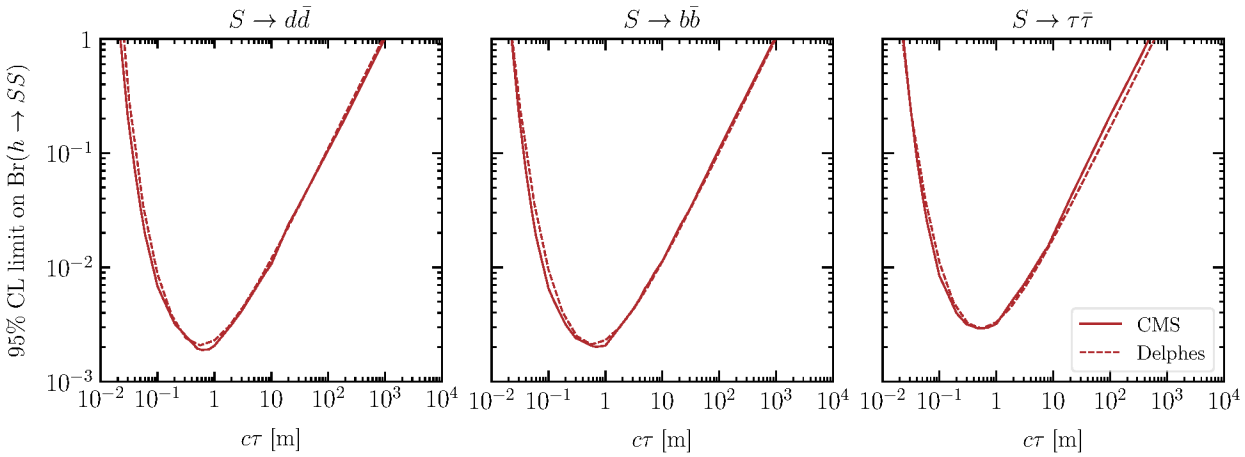

FIG. 1. Comparison of the 95% CL upper limits on the branching fraction Br($h \to SS$) as functions of $c\tau$ derived with the standalone workflow (dashed lines) and the CMS search (solid lines). In deriving these limits we have considered a 15 GeV LLP decaying into d-quark pairs (left), b-quark pairs (center), and τ pairs (right). The limits from this work are shown to agree with the CMS search to within 30%.

D. Analysis strategy

In this section, we discuss the procedure used to recast the CMS Run 2 results and to make projections for high-luminosity LHC (HL-LHC).

For the recasting of the Run 2 results, we used the exact same selection and cuts used in the CMS paper and summarized in Sec. II A. For all the standard particle flow objects, we used the default CMS configuration card from DELPHES, which has been validated [55] to reproduce the object resolutions from Run 2. For the CSC cluster objects, we ran the `CSCClusterEfficiency` and `CscClusterID` modules to select clusters that would pass the corresponding selections. We then applied the $\Delta\phi < 0.75$ selection, jet veto, and time cut for CSC clusters. We used the number of signal events passing these signal selections as our estimate of the signal yield. Finally, by using this estimate for the signal yield together with the background yield (2 ± 1 background event) and observed data (3 observed events) obtained in the CMS analysis, we derived our constraints.

To further inform experimental studies and compare to other proposed LLP experiments, we projected the sensitivity of this analysis to Phase 2 conditions. To simulate the effect on signal yield from the increased number of pileup interactions during Phase 2, we modified the mean pileup number in the CMS configuration card from 32 to 200. We observed that, due to the larger number of jets from pileup interactions, the probability that a CSC cluster is accidentally matched to and vetoed by a pileup jet with $p_T > 10$ GeV is 20% higher. Furthermore, the resolution of p_T^{miss} is degraded by the large number of pileup jets, such that the MET cut efficiency increased by a factor of 2 due to fake MET. However, a more realistic assumption would be that with the help of the new MIP Timing Layer (MTD) and pileup removal algorithms, the MET resolution will be kept at the same level. Similarly, we assumed that with the help of the MTD and new reconstruction

algorithms, the efficiency and resolution would be kept at the same level for all PF candidates. Therefore, a simple projection for Phase 2 constraints can be derived by scaling the signal (and background) yield by the increased integrated luminosity, and applying an 80% correction to the signal yield per cluster while assuming the same efficiency and resolution for all PF candidates.

However, this simple recasting strategy significantly underestimates the potential sensitivity of a Phase 2 analysis. Realistically, given the larger dataset, we would apply tighter cuts to achieve near zero background. Therefore, we considered a second recasting strategy where we apply a tighter N_{hits} selection, the main discriminator of the analysis, until the expected background reaches zero. To estimate the signal and background yield with a tighter N_{hits} cut, we used the N_{hits} distributions in the auxiliary materials from the CMS analysis. We fitted the N_{hits} distribution for the background with an exponential function to extrapolate the background yield at higher N_{hits} cuts. We found that a N_{hits} of 200 would suppress the background yield to below 1 for the expected Phase 2 integrated luminosity. Similarly, we found that increasing the cut from 130 to 210 would give a signal efficiency of about 80%. Therefore, for this recast strategy, we scaled the signal yield by an additional 0.8 with respect to the simple recasting strategy previously described and assumed a background yield of 0.2.

Finally, we considered a search strategy that would be enabled by a new dedicated Level-1 and High Level Trigger targeting this signature starting from the beginning of Run 3. In this strategy, we remove the high $p_{\text{T}}^{\text{miss}}$ selection (which is necessary during Run 2 to trigger the events) and require at least two CSC clusters. In addition, we remove the requirement of at least one jet with $p_T > 50$ GeV and $\Delta\phi(\text{cluster, MET}) < 0.75$ that has high signal efficiency only in the high MET phase space. For this strategy, due to the double cluster requirement, we assumed that zero background can be achieved.

For all the analyses discussed here, we assigned 20% signal systematic uncertainty which is of the same order of signal systematic for the CMS result. There it is dominated by missing higher order QCD corrections, which have a size of 21% for the gluon fusion production mode. Same as the CMS result, we assigned no background systematic uncertainty.

The result of our recast, for both Run 2 and the projection for Phase 2, are shown for all the benchmark models in Section IV. Unless differently stated, the first recasting strategy for Phase 2 is indicated with solid lines, and the second (third) scenario with dot-dashed (dashed) lines.

E. Limit calculation and validation

Before moving to the discussion of the benchmark models, we present a validation of our recast analysis. Specifically, we derived the 95% confidence level (CL) limits on the branching fraction $\text{Br}(h \to SS)$ for different scenarios and compared them to the one derived in the CMS analysis. The observed 95% CL upper limits on the branching fraction $\text{Br}(h \to SS)$ for 15 GeV LLP as functions of $c\tau$ for the $S \to d\bar{d}$, $S \to b\bar{b}$, and $S \to \tau\bar{\tau}$ decay modes were compared against the CMS

results, as shown in Fig. 1. The limits evaluated using the fast simulation from DELPHES agree with the CMS result to within 30% for all lifetimes evaluated.

III. BENCHMARK MODELS

In this section, we briefly describe the benchmark models considered in this work. Each of these models has been chosen to showcase the strengths and limitations of the current analysis in concrete examples exhibiting different kinematics and signal topologies. Specifically, we want to investigate what happens to the analysis reach when lower values of LLP masses are chosen, when the LLP energy, E_{LLP}, is reduced, or when the LLP momentum is not correlated in magnitude or direction with the missing transverse energy. We also want to investigate what happens when there are multiple LLPs produced roughly in the same direction, potentially leading to failed isolation cuts in a non-trivial way.

Concretely, the models we consider are:

- **Exotic Higgs decays** into dark photons or light scalars. These are the closest models to the one considered in the original CMS analysis and are characterized by a production rate decoupled from the exclusive decay channels and LLP lifetime. Besides being commonly chosen benchmarks to compare the performance of different experiments in LLP searches, these benchmarks will allow us to probe the reach for LLP masses lighter than those presented in the CMS analysis, for a fixed production rate and using more realistic decay modes.

- **Axion-like particles (ALPs)** coupled to SM gauge bosons. In this model, the coupling to the SM is provided by a dimension five operator. A single parameter (the ALP decay constant) controls the production rate and lifetime. These models are characterized by a production cross section which is enhanced for energetic LLPs, irrespective of the light LLP mass.

- **Inelastic Dark Matter (DM)**. In these models, the LLP is provided by an almost degenerate partner of the DM, and the amount of energy carried out by the LLP is controlled by the DM-LLP mass splitting and decoupled from the missing transverse energy (MET). This allows us to probe the reach in the low E_{LLP} region while allowing the other selection requirements to be passed without much of a penalty.

- **Confining Hidden Valley models** where jets of LLPs are produced in perturbative hidden showers, analogously to the case of QCD. This benchmark allows us to study the impact of the jet veto in models where multiple LLPs are produced in the same detector region.

In addition to the models considered in this paper, the CMS analysis has also been shown to have non-trivial reach for Heavy Neutral Leptons (HNL) [57].

A. Light scalar singlet

The most minimal extension of the SM is provided by adding a real scalar singlet (S) that mixes with the SM Higgs through renormalizable operators. The Lagrangian for this model reads [58]:

$$\mathcal{L}_{SH} = \mathcal{L}_{\rm SM} + \mathcal{L}_{\rm DS} - \left(A_{HS}\,\hat{S} + \lambda_{HS}\,\hat{S}^2\right)\hat{H}^\dagger\hat{H} \tag{5}$$

where $\mathcal{L}_{\rm SM}$ is the SM Lagrangian, H is the complex Higgs doublet, and the dark sector Lagrangian is given by

$$\mathcal{L}_{\rm DS} = \frac{1}{2}\partial_\mu\hat{S}\,\partial^\mu\hat{S} - \frac{\mu_S^2}{2}\hat{S}^2 + ... \tag{6}$$

where we have omitted possible self-interactions of the scalar singlet, which we assume have been chosen in such a way that S does not have a vacuum expectation value. Here and in the following, we indicate with a hat the original fields with non-canonical kinetic terms, before any field redefinition is performed.

After electroweak symmetry breaking the Higgs scalar, $\hat{h}$, mixes with the singlet $\hat{S}$. The resulting physical states, h and S, obtained by diagonalizing the mass matrix, are given by the linear combination

$$\begin{pmatrix} h \\ S \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \hat{h} \\ \hat{S} \end{pmatrix}, \tag{7}$$

where the mixing angle is controlled by the parameters A_{HS}, and explicitly given by

$$\tan\theta = \frac{x}{1+\sqrt{1+x^2}} \qquad\qquad x = \frac{2vA_{HS}}{\mu_H^2 - \mu_S^2 - \lambda_{HS}\,v^2}\,, \tag{8}$$

with v being the Higgs vacuum expectation value (vev), and $\mu_H^2 = \lambda_H v^2$ with λ_H the Higgs quartic coupling. The mass eigenvalues can also be expressed in terms of the small parameter x as

$$m_{h,S}^2 = \left(\frac{\mu_H^2 + \mu_S^2 + \lambda_{HS}\,v^2}{2}\right) \pm \left(\frac{\mu_H^2 - \mu_S^2 - \lambda_{HS}\,v^2}{2}\right)\sqrt{1+x^2}\,, \tag{9}$$

which for $x \ll 1$ reduce to $m_h^2 \simeq \mu_H^2$ and $m_s^2 \simeq \mu_S^2 + \lambda_{HS}\,v^2$.

Due to the mixing in Eq. (7), S inherits all the couplings of the SM Higgs, modulo a suppression factor, $\sin\theta$, which is controlled by the parameter A_{HS}. Therefore, the decay width of the singlet can be obtained by rescaling the one of a SM Higgs of the same mass. Specifically, we follow references [59, 60] to derive the singlet branching ratios used in this work.

In general, the production cross section is fixed by a combination of the parameters A_{HS} and λ_{HS}. The former controls the production via the $b \to s$ penguin diagram (allowed for $m_S < m_B - m_K$) [61–63] and $s \to d$ penguins for (allowed for $m_S < m_K - m_\pi$); while the latter fixes the double S production via the $b \to s$ penguin diagram with an off-shell SM Higgs [64], or through

direct Higgs decay. Indeed, in presence of a non-vanishing λ_{HS} (and for $2m_S < m_h$) the Higgs can decay into a couple of S with a width given by [58]

$$\Gamma_{h\to SS} = \frac{\lambda_{HS}^2 v^2}{8\pi\, m_h}\sqrt{1-4m_S^2/m_h^2}\,. \tag{10}$$

When $b \to s$ transitions dominate the production channel, decay and production are controlled by the same parameter, θ, and the model parameter space is given by $\{\sin\theta,\, m_S\}$. However, the analysis discussed in this work has no reach for the products of $b \to s$ transitions, as the LLPs would be mostly produced inside (or near) b-jets and fail isolation cuts. Therefore, we will concentrate on the limit where the production is dominated by Higgs decays to two S, which is controlled by the parameter λ_{HS}. Therefore, production and decay channels will be decoupled and the model parameter space given by $\{\lambda_{HS},\, \sin\theta,\, m_S\}$.

Concretely, we generated events for Higgs production from gluon fusion in association with up to two jets and decayed the Higgs to two scalars. No generator level cuts are imposed and the Higgs p_T distribution is reweighed to the NNLO prediction.

We conclude by noticing that, given $m_S^2 \simeq \mu_S^2 + \lambda_{HS}\, v^2$, some level of fine-tuning is required for $m_S^2 < \lambda_{HS}\, v^2$. Measuring the degree of fine-tuning in terms of the parameter $\Delta \equiv m_S^2/(\lambda_{HS}\, v^2)$, we can write the branching ratio for the exotic decay $h \to SS$ as

$$\mathrm{Br}\,(h\to SS) \simeq \frac{\Gamma_{h\to SS}}{\Gamma_h^{\mathrm{SM}}} \simeq 6\cdot 10^{-3}\left(\frac{m_S}{2.5\,\mathrm{GeV}}\right)^4\left(\frac{0.1}{\Delta}\right)^2, \tag{11}$$

where Γ_h^{SM} is the total SM Higgs width.

B. Abelian hidden sector

The next benchmark model that we consider consists in extending the SM by adding a *dark* U(1) gauge symmetry which is spontaneously broken by a dark Higgs field, S. The dark U(1) is mediated by a *dark* photon, X, which kinetically mixes with the SM hypercharge as:

$$\mathcal{L}_{SH} = \mathcal{L}_{\mathrm{SM}} + \mathcal{L}_{\mathrm{DS}} - \lambda_{HS}\,\hat{S}^2\hat{H}^\dagger\hat{H} - \frac{\epsilon}{2\cos\theta_W}\hat{X}_{\mu\nu}\hat{B}^{\mu\nu}\,, \tag{12}$$

where $\hat{B}_{\mu\nu}$ and $\hat{X}_{\mu\nu}$ are the field strengths of the hypercharge and the new U(1) gauge group respectively. The dark sector Lagrangian is

$$\mathcal{L}_{\mathrm{DS}} = -\frac{1}{4}\hat{X}_{\mu\nu}\hat{X}^{\mu\nu} + \mu_S^2\hat{S}^2 - \lambda_S\hat{S}^4 + |(\partial_\mu + ig_D\hat{X}_\mu)\hat{S}|^2\,. \tag{13}$$

As before, we indicate with a hat the original fields with non-canonical kinetic terms, before any field redefinition is performed. The dark U(1) is spontaneously broken by the vev of the dark Higgs, $\langle S\rangle = v_S/\sqrt{2}$, which generate a mass for the dark photon $m_{X,0} = g_D v_S$.

After electroweak symmetry breaking the kinetic mixing between the dark photon and the hypercharge induces a coupling of the dark photon to the SM fermions which, in the $m_X^2 << m_Z^2$ limits, reads

$$\mathcal{L}_{Xf\bar{f}} = \epsilon e\, Q_f X_\mu \bar{f}\gamma^\mu f\,, \tag{14}$$

where Q_f is the fermion electric charge. This coupling, controlled by the small parameter ϵ, provides the decay channel in visible states for the dark photon. Specifically, we compute the dark photon branching ratios by using the package provided in [65].

The diagonalization of the scalar sector proceeds similarly to what was discussed in the previous section, with the only difference that now we are interested in the regime where $m_S \gg m_h$, so that the dark Higgs decouples from the phenomenology of the model. Given the non vanishing coupling between S and X, the mixing between the SM and dark Higgs generates a non-zero hXX coupling which gives rise to the exotic Higgs decay $h \to XX$, with a width given by

$$\Gamma(h \to XX) = \frac{\lambda_{HS}^2}{32\pi}\frac{m_h v^2}{m_S^2}\sqrt{1-\frac{4m_X^2}{m_h^2}}\,\frac{(m_h^2+2m_X^2)^2 - 8(m_h^2-m_X^2)m_X^2}{m_h^4}\,. \tag{15}$$

In the limit of small ϵ (which will be the relevant limit for our analysis), this dominates over Drell-Yan and $h \to ZX$ production and becomes the dominant dark photon production channel. In this limit the decay channel, controlled by ϵ, and the production channel, controlled by λ_{HS}, are decoupled; and the model parameter space is given by $\{\epsilon, \lambda_{HS}, m_X\}$. The event generation for this benchmark was performed similarly to the light scalar singlet case.

C. Inelastic Dark Matter

Inelastic Dark Matter (iDM) models are characterized by a DM candidate that couples with the SM only through interactions with a nearly degenerate state. A simple realization of this scenario can be obtained by adding to the model discussed in the previous section a Dirac pair of Weyl fermions, η and ξ, that couple to the dark photon, X, with opposite charges. As before, the Higgs provides a source of U(1) breaking, generating a mass for the dark photon and a Majorana mass, δ, for the two Weyl fermions. A Dirac mass, m_D, involving the two Weyl fermions is also allowed, so that at energies below the dark U(1) breaking scale, the mass terms for the dark fermions are

$$\mathcal{L} \supset -m_D\,\eta\xi - \frac{\delta}{2}(\eta^2+\xi^2) + \text{h.c.}. \tag{16}$$

For a technically natural small Majorana mass, these mass terms can be perturbatively diagonalized to give the physical states

$$\chi_1 \simeq \frac{i}{\sqrt{2}}(\eta-\xi) \qquad\qquad \chi_2 \simeq \frac{1}{\sqrt{2}}(\eta+\xi)\,, \tag{17}$$

which have nearly degenerate masses $m_{1,2} \simeq m_D \pm \delta$. These mass eigenstates couple off-diagonally to the dark photon, *i.e.*

$$\mathcal{L} \supset ie_D \hat{X}_\mu \, \bar{\chi}_1 \gamma^\mu \chi_2 + \mathcal{O}\left(\frac{\delta}{m_D}\right), \tag{18}$$

where we have written $\chi_{1,2}$ as Majorana spinors using four-component notation. Therefore, if $m_X > m_1 + m_2$ and $\alpha_D \gg \epsilon \alpha_{\rm em}$, once produced dark photons decay into $\chi_1\chi_2$ pairs with a rate given by $\Gamma_{X\to\chi_1\chi_2} \simeq \alpha_D m_X$, and provide the dominant production channel for $\chi_1\chi_2$ pairs at LHC. For the values of ϵ we are interested in this analysis, the dominant production channel for dark photons is provided by Drell-Yan processes and scales as ϵ^2.

The lightest state, χ_1, is stable and once produced leaves the detector as missing energy; χ_2 can decay into χ_1 plus a pair of SM particles through an off-shell dark photon, possibly leaving a detectable signature. The rate for decays with leptonic final states is given by [66]:

$$\Gamma_{\chi_2\to\chi_1 l\bar{l}} = \epsilon^2 \alpha_{\rm em}\alpha_D \int_{4m_l^2}^{(m_1\Delta)^2} ds \frac{|\vec{p}_1|(m_1^2\Delta^2 - s)(2s + m_1^2(2+\Delta)^2)(s+2m_l^2)(s-4m_l^2)^{1/2}}{6\pi m_2^2 s^{3/2}(s-m_X^2)^2} \tag{19}$$

where s is the invariant mass of the lepton pair, $\vec{p}_1$ is the momentum of χ_1 in the rest frame of χ_2, and we have introduced the dimensionless parameter $\Delta \equiv (m_2 - m_1)/m_1$. The rate for decays involving hadronic final states can be derived by setting $m_l = m_\mu$ and multiplying the integrand of Eq. (19) by the experimentally measured quantity $R(s) \equiv \sigma(e^+e^- \to \text{hadrons})/\sigma(e^+e^- \to \mu^+\mu^-)$.

For this benchmark, events were generated using a MadGraph5 Z' model for X via production in association with up to three jets. A generator level cut $p_T > 100$ GeV was applied on the X, as the truth-level $p_{\rm T}^{\rm miss}$ is given by p_T of the Z' (its decay products are one DM particle and an LLP decaying in the muon chambers).

D. Axion-like particles

Axion-like particles (ALPs) extend the axion scenario to include any pseudoscalar particle that couples to the SM through dimension five operators. The naturally suppressed couplings make them a natural candidate for LLP searches. The general Lagrangian for these kinds of models is given by

$$\begin{aligned}\mathcal{L} =& \mathcal{L}_{SM} + \frac{1}{2}(\partial_\mu a)^2 - \frac{1}{2}m_a^2 a^2 + \frac{c_q^{ij}}{2f}(\partial_\mu a)\bar{q}_i\gamma^\mu\gamma^5 q_j + \frac{c_\ell^{ij}}{2f}(\partial_\mu a)\bar{\ell}_i\gamma^\mu\gamma^5\ell_j \\ &+ \frac{a}{4\pi f}\left(\alpha_s c_{GG}\, G^a_{\mu\nu}\widetilde{G}^{a,\mu\nu} + \alpha_2 c_{WW}\, W^a_{\mu\nu}\widetilde{W}^{a,\mu\nu} + \alpha_1 c_{BB}\, B_{\mu\nu}\widetilde{B}^{\mu\nu}\right) + \cdots\end{aligned} \tag{20}$$

where $\widetilde{G}_{\mu\nu} = 1/2\,\epsilon_{\mu\nu\rho\sigma}G^{\rho\sigma}$ where $G^{\rho\sigma}$ is the gluon field strength, and similarly for $\widetilde{W}$ and $\widetilde{B}$. In the broken, phase the couplings to W and B bosons induce couplings to photons and Z-bosons which are given by:

$$c_{ZZ} = c_{WW} + c_{BB} \qquad c_{\gamma Z} = c_w^2 c_{WW} - s_w^2 c_{BB} \qquad c_{\gamma\gamma} = c_w^4 c_{WW} + s_w^4 c_{BB}\,. \tag{21}$$

In this work, we will focus on benchmark models in which the ALP couples only to gauge bosons ($c_q^{ij} = c_\ell^{ij} = 0$). Since the focus is on the production of energetic, isolated LLPs, this choice is sufficient to capture most of the dominant production modes at the LHC. Specifically, we will consider the three following scenarios: ALP coupled to W ($c_{WW} \neq 0$, $c_{GG} = c_{BB} = 0$), *photophilic* ALP ($c_{\gamma\gamma} \neq 0$, $c_{\gamma Z} = c_{GG} = 0$), and ALP coupled to gluons ($c_{GG} \neq 0$, $c_{BB} = c_{WW} = 0$). The latter is a well-studied benchmark model in the context of light LLP searches, yielding the highest production rate at the LHC. The photophilic model chosen here is one of the (infinite) possible choices of UV-completion at LHC energies of the well-studied low-energy benchmark of "ALP coupled to photons". The conservative choice $c_{\gamma Z} = 0$ is to focus on the parameter region where the existing LEP bounds are the weakest. Finally, the $c_W \neq 0$ benchmark provides a better UV-motivated benchmark than the photophilic choice, where associated ALP production with all the gauge bosons is allowed.

For the ALP coupled to gluons, we generated events where the LLP is produced in association with up to 3 jets, and imposed a $p_T > 100$ GeV and a $|\eta| < 3$ generator-level cuts on the transverse momentum and pseudorapidity of the ALP. The MadGraph5 model used here has been described in [67], and we have only adapted the normalization of the couplings to the one used above. We did not include ALP production in the shower (i.e. where the ALP is produced at intermediate scales between the hard process collision and the QCD confinement scale) which was first estimated in [42], as there are not yet reliable event generators that can be used to keep track of the angular separation between the ALP and QCD jets (necessary for the jet veto requirements of the analysis) [68]. Therefore, for this benchmark, our limits should be considered conservative estimates for the reach of this CMS analysis, as they miss an important production channel. Production from meson mixing and meson decay was also neglected because it yields softer and non-isolated ALPs, for which this analysis has no sensitivity. For the lifetime and exclusive decay branching ratios of this benchmark, we used the estimates of [69].

For the case of the other two ALP benchmarks, we considered ALP production in association with either a W, a Z, or a photon and up to 2 extra jets. We kept the same generator-level pseudorapidity cut but lowered the p_T cut to 50 GeV as some of the missing transverse energy can be produced by the decay products of the W and Z bosons. In these two benchmarks, the ALP decays predominantly into two photons.

E. Hidden Valley

Confining Hidden Valleys (HV) [70], with a perturbative evolution below the scale mediating the interactions producing hidden sector particles, are a generic hidden sector extension of the SM on which we can have some theoretical control based on our knowledge of QCD-like theories. In general, one expects jets of hidden sector partons to hadronize in HV particles, some of which may decay back into SM final states, potentially as LLPs. Still, large freedom exists in defining a specific

model. From the field content of the hidden sector and its symmetries, to the portal interactions mediating both the production of HV states and their decay back to the SM [71]. Many studies of search strategies at the LHC have been performed for different incarnations of this paradigm [21].

In the context of this reinterpretation study, we choose one particular realization as an example model generating the LLP-jet signature, aiming at maximizing the multiplicity of LLP produced in a jet, while keeping a high level of simplicity of reinterpretation. Therefore the example chosen is by no means generic *per se*, although the lessons learned about the CMS analysis are. Specifically, we used the Hidden Valley module [72] implemented in Pythia to generate events and choose a perturbative hidden sector with an $SU(N_c)$ asymptotically free gauge group with N_f hidden quark flavors, fixing $N_c = 3$ and $N_f = 1$. The choice of $N_f = 1$ is to guarantee the absence of stable hidden mesons, therefore reducing the amount of collider stable particles produced and maximizing the number of LLPs in a hidden jet. This has a drawback, namely the lack of knowledge of the mass spectrum of such a theory as it lacks chiral symmetry breaking which is an important handle used in lattice simulation. In particular, the mass ratio between the first (pseudo-)scalar, η_V, and vector, ω_V, resonances are poorly known, but expected to be $\mathcal{O}(1)$ [73–75]. Again, motivated by maximizing LLP multiplicity, we choose $m_{\omega_V} = 2.5 m_{\eta_V} = \Lambda_{HV}$ and assumed that the lowest scalar state (which Pythia will not use in the hadronization of the HV partons) is also able to decay to pairs of η_V. In this way, vector resonances can promptly decay to pseudoscalar mesons η_V, which will be the LLPs. For portals, we decide to decouple production and HV meson decays so that we can study the effects of varying the LLP lifetime on the limits for a fixed production rate. Specifically, we will produce hidden quarks in Higgs decays and will decay back the hidden spin-0 mesons 100% into pair of photons. The latter choice is purely driven by the fact that the CMS analysis is not too sensitive to the relative amount of hadronic vs electromagnetic energy in LLP decays. At the same time, existing limits on light LLPs decaying to pair of photons are quite weak, so we can focus on reinterpreting this analysis without worrying about recasting other existing searches[1]. From a model building point of view, these portals can be easily generated by introducing a heavy scalar and pseudoscalar states S and A, having Yukawa interactions with the HV vector-like quark q_V. The scalar S can then interact with the SM Higgs via a $|H|^2 S$ cubic interaction, generating a Yukawa coupling between q_V and the SM Higgs and a q_V mass after electroweak symmetry breaking. At the same time, the pseudoscalar A can have a coupling to the SM photons $AF\tilde{F}$ which in turn will induce a small decay width for η_V via $\eta_V - A$ mixing.

[1] The case of a recent CMS search for trackless jets [76] provides likely the strongest constraint for low values of $c\tau$ where a significant fraction of LLPs decays in the inner detector. However, that analysis explicitly vetoes signatures compatible with loose photons and photon conversion. The efficiency for one or more light LLPs decaying into pairs of photons being identified as loose photons is hard to recast, therefore we do not consider such analysis when presenting our limits.

IV. RESULTS

In this section, we present the results for the benchmark models discussed in the previous section. We present both the current constraints, derived from data collected from 2016 to 2018, corresponding to an integrated luminosity of 137 fb^{-1} and the projected constraints for Phase 2. The different projections for Phase 2 are derived by using the three different search strategies discussed in Section II D. Specifically, solid lines correspond to the search with the same selections as the CMS paper and a background rescaled according to the higher luminosity, dot-dashed lines correspond to the search with a higher N_{hits} cut and zero background, and the dashed lines to the search with a dedicated trigger (that no longer require the MET and isolation cuts, but the presence of two separate LLP decays in the muon chambers) and zero background (see Section II D for a more detailed discussion of the search strategies). Other existing and projected limits shown in the following plots are all taken from the literature, as referenced in the figure captions. The only exception is a limit originating from an ATLAS mono-jet search for the case of the gluon-coupled ALP, Fig. 8, whose mass dependence was derived in this work as described in Appendix A.

In Fig. 2 we show the reach for the light scalar model (discussed in Section III A) with $\lambda = 1.6 \times 10^{-3}$. This choice of λ corresponds to an exotic Higgs branching fraction of $\text{Br}(h \to SS) = 0.01$, which is roughly the future reach for the Higgs branching into invisible final states. The present constraints are shown in the left panel, where we see that for low masses the analysis probes a previously unconstrained region of the parameter space, while at higher masses the constraints are similar to the ones of the ATLAS search for displaced vertices in the muon chambers (indicated as ATLAS mu-ROI in Fig. 2), whose reach was presented for $m_S > 5$ GeV. In the right panel, we show the projections for Phase 2 and compare them with the projected constraints from other future experiments and upgrades. We can see that, thanks to the different distance from the interaction point (IP), the projected results are complementary to dedicated LLPs experiments such as CODEX-b, FASER2, and MATHUSLA; all of which are positioned further away from the IP. To give an idea of how the constraints depend on the value of $\text{Br}(h \to SS)$, in Fig. 3 we show the same constraints of Fig. 2 but for different values of $\text{Br}(h \to SS)$. We see that the current search start to lose sensitivity for $\text{Br}(h \to SS) \lesssim 3 \times 10^{-3}$, while for the future Phase 2 search we start to lose sensitivity for $\text{Br}(h \to SS) \lesssim 3 \times 10^{-4}$. In all the plots we present also the values of the LLP mass (function of $\text{Br}(h \to SS)$) below which tuning of more than 10% is present. Alternatively, in Fig. 4 we show the limits for a different slicing of the parameter space of this model, where the tree-level mass for S is absent and the LLP mass is fully controlled by λ_{HS} and therefore by $\text{Br}(h \to SS)$. In this case, there is no tuning even for lower masses, but the production rate varies with m_S and searches for $H \to$ inv. set an upper bound on m_S. Finally, to compare with the results of the CMS analysis, in Fig. 5 we report the present and future limits on $\text{Br}(h \to SS)$ as a function of the scalar lifetime.

In Fig. 6 we report the constraints for the Abelian hidden sector discussed in Section III B. As

before, the value of the exotic Higgs branching ratio is fixed to $\mathrm{Br}(h \to A'A') = 0.01$. We see that our current constraints (left panel) cover a mostly unconstrained region of the parameter space, except for the overlap with the ATLAS mu-ROI search at high masses. As for the scalar model, our projected constraints (right panel) well complement dedicated LLPs searches thanks to the different baselines. To investigate which is the lowest value of $\mathrm{Br}(h \to A'A') = 0.01$ that we can probe, in Fig. 7 we show present and future constraints for different values of the exotic Higgs branching. For the current search we see that we start to lose sensitivity for $\mathrm{Br}(h \to A'A') = 3 \times 10^{-3}$, while for the Phase 2 the constraints start to disappear for $\mathrm{Br}(h \to A'A') = 3 \times 10^{-4}$. This is consistent with what was found for the singlet scalar model and shows the relative insensitivity of the analysis to the specific exclusive decay modes. The only significant differences are around resonance mixing with hadronic resonances, which differ between the scalar and vector LLPs, and affect the LLP lifetime; and in the region between $200\,\mathrm{MeV} \lesssim m \lesssim 300\,\mathrm{MeV}$ where the 2μ final state, to which this analysis is not sensitive to, contributes to $\mathcal{O}(50\%)$ of the dark photon branching ratios.

The constraints for the three ALP models that we consider are shown in Fig. 8 - 9. For both the gluon (Fig. 8) and electroweak (Fig. 9 and Fig. 10) coupled scenarios, we find that the reinterpretation of the CMS analysis covers new territory beyond previous monojet [77] and fixed target [78] searches while being complementary to dedicated LLP experiments. Moreover, one can expect the projections shown here to be underestimated, as dedicated searches using the fact the ALP is produced in association with a photon or a vector boson may allow us to relax some of the cuts and access softer LLPs that are produced with higher rates, pushing the estimated limits towards higher ALP masses and decay constants.

We now turn to the inelastic DM model results. The reinterpretation of this model is fairly sensitive to the LLP energy, E_{LLP}, via the mass splitting, Δ. Unfortunately, the efficiency tables provided by the CMS Collaboration in HepData are not granular enough at low deposited energies (E_{em}, E_{had}) to resolve the turn-on shoulder of the 2D efficiency surface (the first bin is between 0 and 25 GeV). Therefore, our ability to reliably recast this model is hampered by the lack of knowledge about the minimal energy threshold for which the LLP visible decay products can produce $\mathcal{O}(20-30)$ charged particles emerging from a steel layer into the muon stations. To estimate this energy threshold, we impose an additional cut $E_{LLP} > 5\,\mathrm{GeV}$ (which is approximately the energy needed for an electron to produce $\mathcal{O}(20)$ charged particles at the shower development maximum). The constraints for this choice of cut and using the model parameters $\Delta = 0.005$, $\alpha_D = 0.1$, and $m_{A'} = 3m_1$, are reported in Fig. 11. We see that the analysis covers previously unconstrained regions of the parameter space near the Z-resonance at $m_{A'} = 3m_1 = m_Z$. We have also derived the constraints for a model with $\Delta = 0.01$ but decided not to show them since they are weaker than already existing limits, as larger mass splittings produce lifetimes too short to reach the CMS muon chambers. To further estimate the sensitivity of these results to the lower cut on the LLP energy, we show in Fig. 12 the effect of varying it between 0 and 10 GeV.

Finally, in Fig. 13 we report the limits on the exotic Higgs branching ratio $\mathrm{Br}(h \to Q\bar{Q})$ for the

hidden valley model discussed in Section III E. We specifically choose a value for the HV confining scale $\Lambda_{HV} = 20\,\text{GeV}$, which correspond to a pseudoscalar mass $m_{\eta_V} = 8\,\text{GeV}$. Since in this model LLPs are produced within dark-showers in LLP jets, we expect the jet veto to reduce the sensitivity of the analysis. To quantify this effect, in the lower panels of Fig. 13 we show the ratio of the signal efficiency of the CMS analysis divided by the signal efficiency of the same analysis without the jet veto. As expected, this ratio rapidly approaches zero for small LLP lifetimes, when it is more likely for multiple LLPs to decay within the inner detector regions and the calorimeters in front of the cluster in the muon chambers selected as a signal by the analysis. Conversely, in the long lifetime area, the higher LLP multiplicity renders the limit more stringent than the case of Higgs decay to pairs of LLPs. Lowering the hidden confinement scale will increase the meson multiplicity inside hidden jets and therefore amplify this behavior.

V. DISCUSSION

We have explored some of the strengths and limitations of a recent search for LLPs using the muon chambers at CMS. The reinterpretation was made possible by the additional information provided by the collaboration in HEPData, which was embedded into DELPHES modules.

We have shown that this analysis proves very effective at constraining light LLPs, $m_{LLP} < \mathcal{O}(\text{GeV})$, as long as they can be produced energetically in LHC hadronic collisions and have $c\tau \lesssim \mathcal{O}(\text{m})$. In fact, we found that the current version of such a search strategy not only provides a counterexample to the lore that LLP searches at ATLAS and CMS are limited at low masses by irreducible SM backgrounds, but it is already able to cover previously unconstrained parameter space in many models, see Figs. 2, 6, 8, 10, 11, 13, competing with and complementing the reach of dedicated LLP detectors.

Still, various avenues for improvement exist. As mentioned in Sec. IV, producing signal categories with lower MET requirements but in association with another object such as a photon, lepton(s) or b-jet, may improve limits on specific models such as ALPs and HNLs. This will greatly increase the coverage of the search for many other models. This is especially true given the particular simplicity and reliability of the recasting provided by the publicly released information in HepData [49]. In this respect, we encourage the CMS Collaboration to provide more finely spaced efficiency maps at low (E_{had}, E_{em}) to fully capture the turn-on shoulders, which is required in models where LLPs are producing less visible energy such as in the inelastic Dark Matter benchmark shown here.

Perhaps the most important avenue of improvement may be the investigation of how much the cluster isolation requirement can be relaxed. Many models, and production modes within a model, produce LLPs inside (b-)jets. Examples include the case of a light scalar model, where S can be produced efficiently in b decays and would yield muon chamber clusters not isolated from a b-jet; the case of ALPs produced in b-flavored hadron decays or in hadronic showers via π^0- $\eta^{(\prime)}$

mixing; or the case of emerging jets [79] where showering within QCD and a Hidden Valley happens concurrently. Extending this kind of searches into the non-isolated regime will inevitably require some characterization and understanding of the origin of SM backgrounds mimicking clusters in the CSC. This effort has also implications and synergies beyond CMS itself. In fact, the amount of (instrumented) shielding provided by the calorimeters and the steel layers in the muon chambers is about $20-27$ nuclear interaction lengths, not far from the required shielding of other proposals for dedicated LLP experiments, such as e.g. the 30λ of active shield estimated to be necessary for CODEX-b [37]. Therefore, any characterization of SM backgrounds for CMS LLP searches would also benefit and inform the ongoing shielding design and simulation of other experimental proposals such as CODEX-b.

Many benchmarks chosen here correspond to some of those selected to compare present and future efforts in the LLP search program such as within the CERN Physics Beyond Collider (PBC) [6]. Given the relevance of this novel type of CMS search on the LHC reach for LLPs, we encourage the Collaboration to produce official limits that can be included in the PBC comparison plots. The capabilities of this kind of CMS search in probing light LLPs greatly extend what was considered possible for general-purpose existing LHC experiments to achieve and will undoubtedly complement and inform the broader future search program for LLPs beyond the Standard Model.

ACKNOWLEDGMENTS

We would like to thank the CMS Collaboration and especially the CMS Exotica group, where part of this work was presented in 2021, for their positive feedback. CW and SX are partially supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0011925. This work has been discussed in the relevant Snowmass community study subgroups in 2022 and is part of CW's PhD thesis along with additional reinterpretation studies based solely on CMS published results and Delphes public codes [31, 49, 56]. CW, CP, and SX are grateful to the organizers and participants of the "New ideas in detecting long-lived particles at the LHC" workshop at LBNL in the Summer of 2018 where experimentalists and theorists gathered to generate new ideas on triggers and analysis strategies for long-lived particles searches at the LHC [80] as well as the Fermilab LPC LLP group.

AM was supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award No. DE-SC0021431, the Quantum Information Science Enabled Discovery (QuantISED) for High Energy Physics (KA2401032), and the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy - EXC 2121 Quantum Universe - 390833306.

MP was supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0011632 and by the Walter Burke Institute for Theoretical Physics. MP would like to thank the Aspen Center for Physics where part of this work was performed.

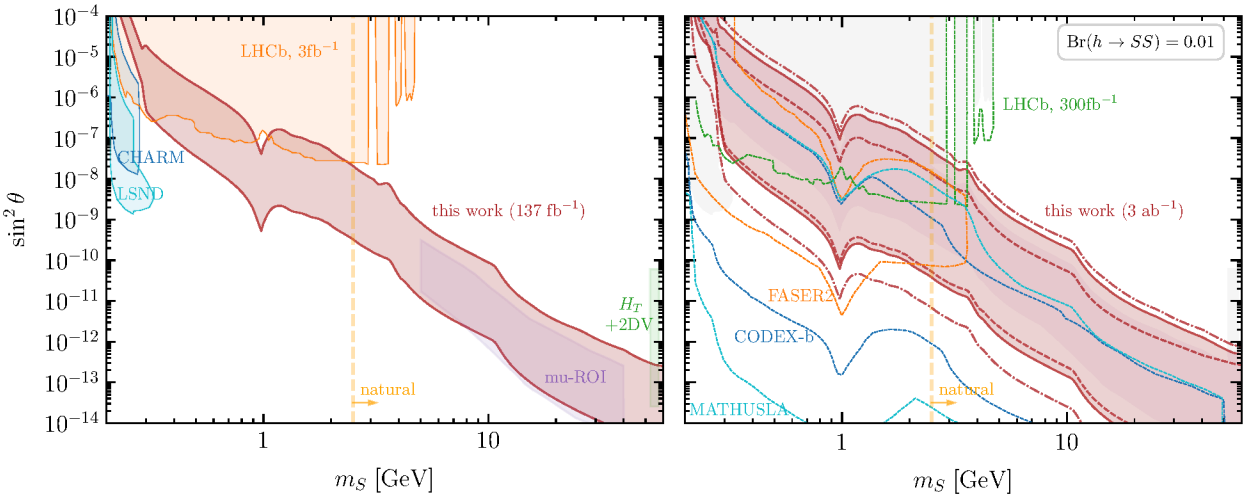

FIG. 2. Constraints on light scalars produced in Higgs decays for $\mathrm{Br}(h \to SS) = 0.01$. **Left:** Comparison of our current reach (red region) with existing limits from LHCb (orange) [81], LSND (azure) [82], reinterpretation [59] of the CHARM experiment (blue) [83], CMS "HT +2DV" search (green) [10, 60], and reinterpretation of ATLAS mu-ROI (purple) [60, 84]. **Right:** Projections of our constraints for a luminosity of $3\,\mathrm{ab}^{-1}$ (red region). The three red contours (solid, dashed, and dot-dashed) correspond to the three search strategies discussed in the main text (rescaled CMS analysis, dedicated trigger, and higher N_{hits}). We compare our results with current constraints (gray shaded region) and projections for MATHUSLA [85], CODEX-b [42], FASER2 [86], and LHCb 300fb^{-1} [87]. The vertical orange line indicates the scalar mass below which the model needs to be fine-tuned (see discussion around Eq. (11)).

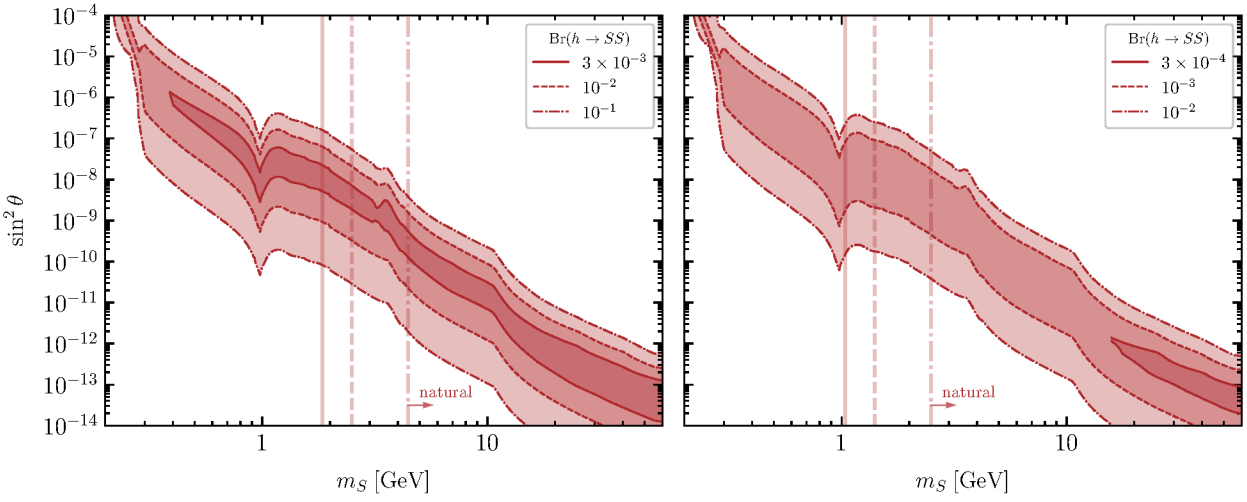

FIG. 3. Our limits for the light scalar model for different values of $\mathrm{Br}(h \to SS)$. In the left panel we show the current reach, while in the right panel we present the $3\,\mathrm{ab}^{-1}$ projections assuming that the same selections of the original CMS analysis are used. As in the previous plot, the vertical lines indicate the scalar mass below which the model needs to be fine-tuned (see discussion around Eq. (11)).

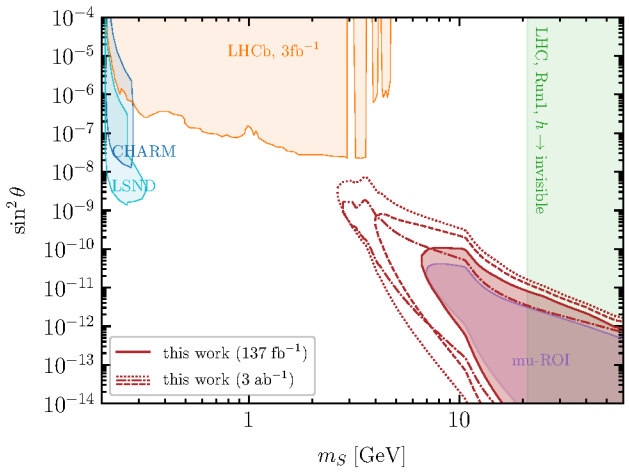

FIG. 4. Constraints on the singlet scalar model in absence of a tree-level mass for S ($\mu_S = 0$). The solid red line shows the current constraints, while the other contours (dashed, dot-dashed, and dotted) show 3 ab^{-1} projections derived by using the three recast strategies discussed in Section II D (rescaled CMS analysis, dedicated trigger, and higher $N_{\rm hits}$). The other existing constraints appearing on the plot are the same of Figure 2.

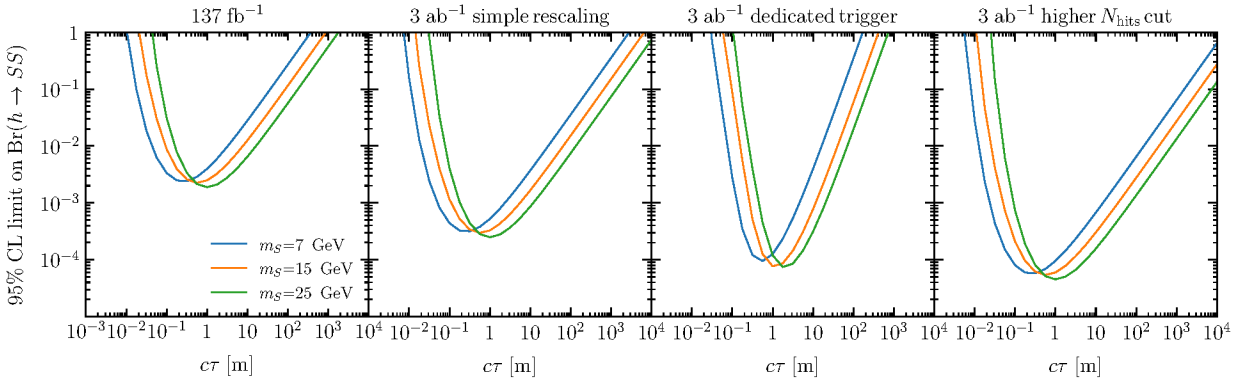

FIG. 5. Upper limits on the branching fraction Br($h \to SS$) as functions of $c\tau$. In the left panel, we report the current constraints set by our analysis. In the other panels, we show the projected constraints for a luminosity of 3 ab^{-1} derived from the three different search strategies discussed in the main text.

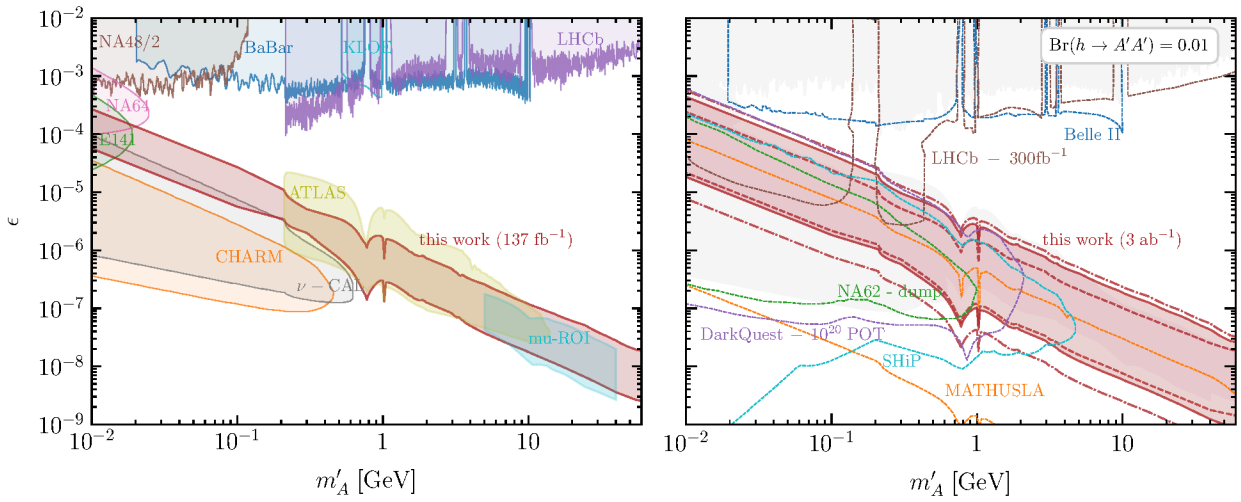

FIG. 6. Constraints on dark-photons produced in Higgs decays for $\mathrm{Br}(h \to SS) = 0.01$. **Left:** Comparison of our current reach (red region) with existing limits from BaBar (blue) [88], KLOE (azure) [89], LHCb (purple) [90], NA48 (brown) [91], reinterpretation of ATLAS μ-ROI (yellow) [84], ATLAS search for displace dark-photon jets (yellow) [92], and beam dump experiments (orange, gray, green, pink) [83, 93–95]. Most of the experimental constraints appearing in this plot have been digitized with the help of `darkcast` [65]. **Right:** Projections of our constraints for a luminosity of $3\,\mathrm{ab}^{-1}$ (red region). The three red contours (solid, dashed, and dot-dashed) correspond to the three search strategies discussed in the main text (rescaled CMS analysis, dedicated trigger, and higher N_{hits}). We compare our results with current constraints (gray shaded region) and projections for MATHUSLA (orange) [85], SHiP (azure) [38], DarkQuest (purple) [13, 96], NA62 in dump mode (green) [13, 97], LHCb upgrade (brown) [13, 98], and Belle II (blue) [99].

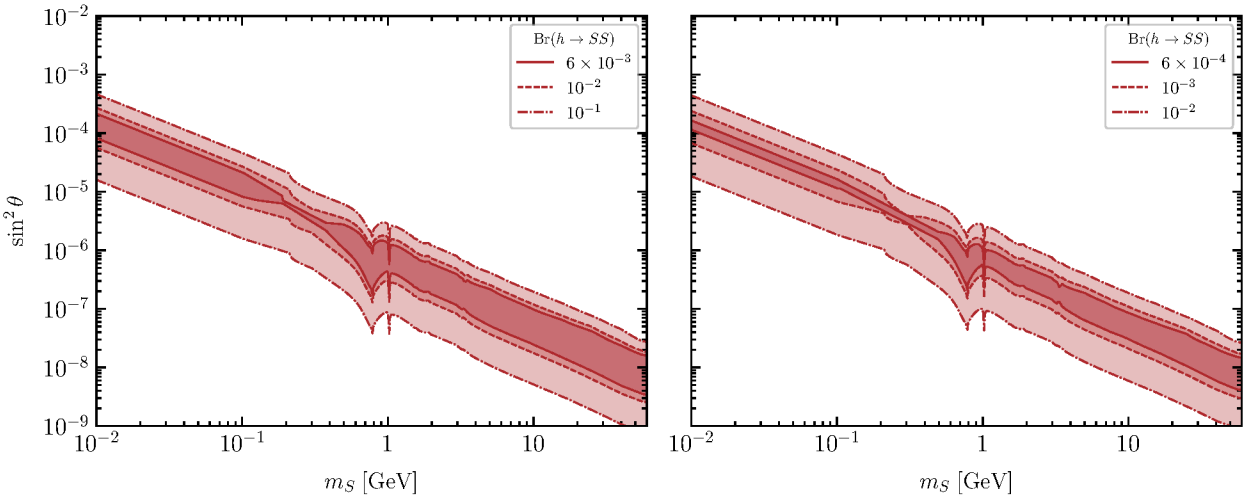

FIG. 7. Our limits for the dark photon model for different values of $\mathrm{Br}(h \to SS)$. In the left panel we show the current reach, while in the right panel we present the projections for 3 ab^{-1} assuming the same selections of the original CMS analysis are used.

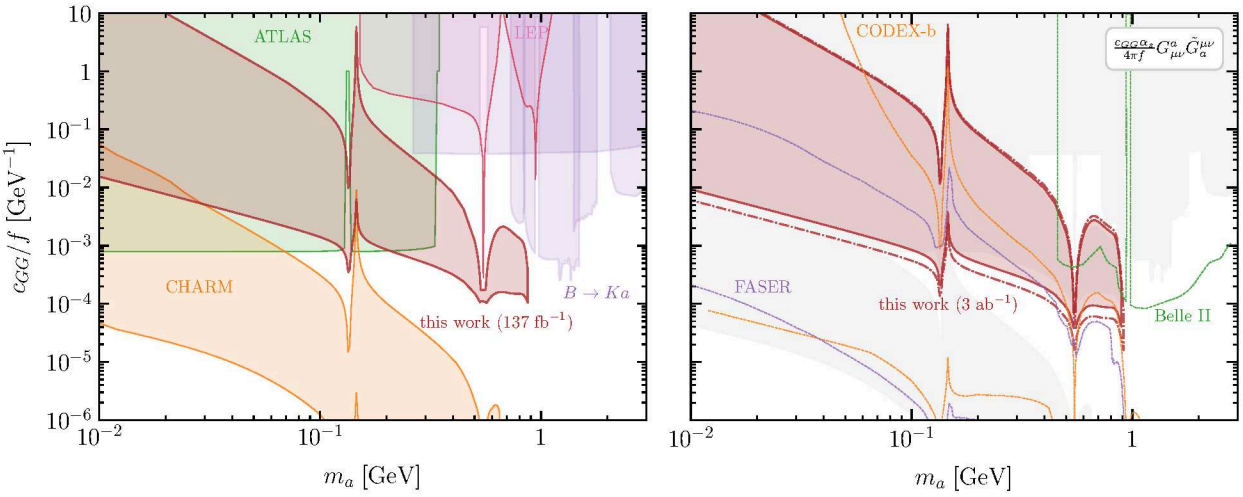

FIG. 8. Constraints on ALPs coupled to gluons. **Left:** Comparison of our current reach (red region) with existing limits from CHARM (orange) [42, 78], our reinterpretation of ATLAS (green) [77], LEP [69], and flavor probes (purple) [100–104]. **Right:** Projections of our constraints for a luminosity of 3 ab^{-1} (red region). The solid and dot-dashed red contours correspond to the projections derived by using the same selections of the original CMS analysis, and the one derived by using a higher $N_{\rm hits}$ cut and assuming zero background. We compare our results with current constraints (gray shaded region) and projections for FASER (purple) [36], CODEX-b (orange) [42], and Belle II (green) [100, 102]

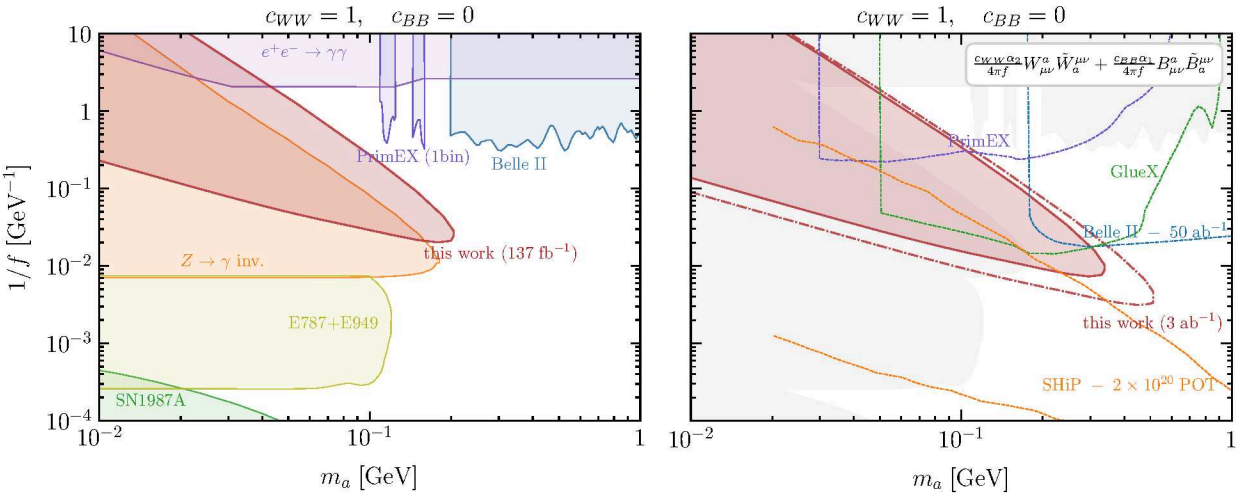

FIG. 9. Constraints on ALPs coupled to W bosons. **Left:** Comparison of our current reach (red region) with existing constraints from star cooling constraints (green) [105], beam dump experiments (yellow) [105], Z invisible branching ratio (orange) [106, 107], limits on $e^+e^- \to \gamma\gamma$ from LEP data (violet) [106, 108–110], PrimeEX (purple) [111, 112], and Belle II (blue) [113]. **Right:** Projections of our constraints for a luminosity of 3ab^{-1} (red region). The solid and dot-dashed red contours correspond to the projections derived by using the same selections of the original CMS analysis, and the one derived by using a higher $N_{\rm hits}$ cut and assuming zero background. We compare our results with current constraints (gray shaded region) and projections for SHiP (orange) [6, 38], PrimEX (purple) [111, 112], GlueX (violet) [111, 114], , and Belle II (blue) [6, 115].

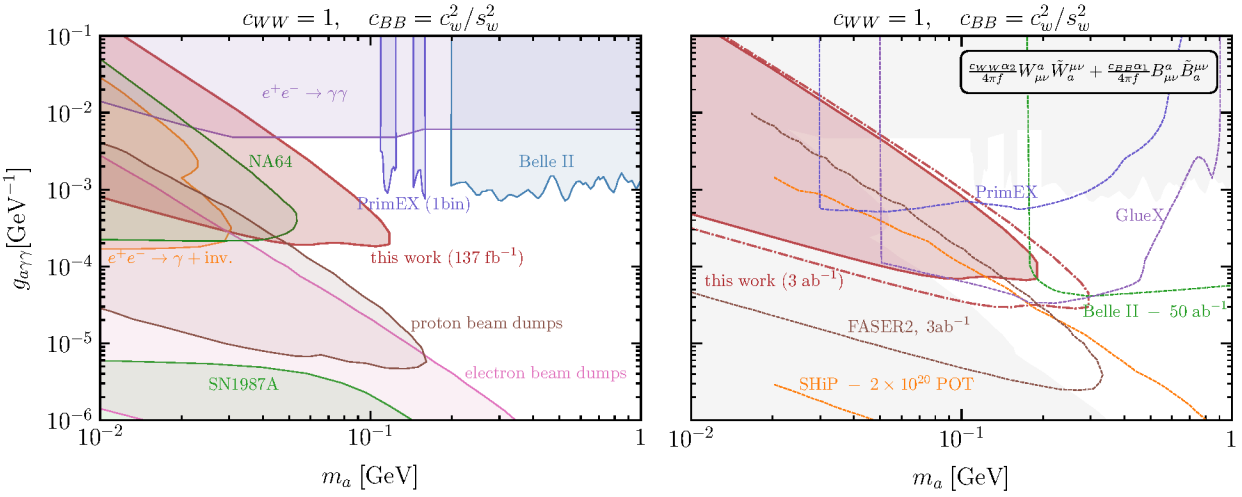

FIG. 10. Constraints on ALPs coupled to electroweak gauge bosons, and with $c_{\gamma Z} = 0$. **Left:** Comparison of our current reach (red region) with existing constraints from star cooling constraints (green) [105], electron [94, 106, 116] and proton [6, 78, 117] beam dump experiments (pink and brown), limits from monophoton searches at LEP (orange) [106, 118], NA64 (green) [95], PrimEX (purple) [111, 112], and Belle II (blue) [113]. **Right:** Projections of our constraints for a luminosity of 3ab^{-1} (red region). The solid and dot-dashed red contours correspond to the projections derived by using the same selections of the original CMS analysis, and the one derived by using a higher N_{hits} cut and assuming zero background. We compare our results with current constraints (gray shaded region) and projections for FASER (brown) [36], SHiP (orange) [38], PrimEX (purple) [111, 112], GlueX (green) [111, 114], and Belle II (green) [115].

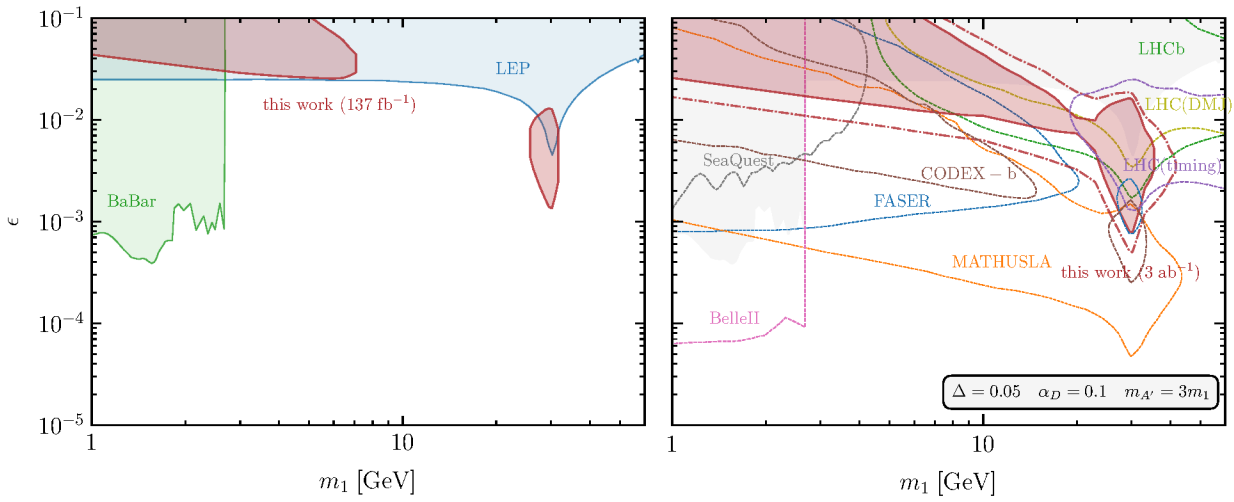

FIG. 11. Constraints on inelastic DM models, assuming a normalized mass splitting of $\Delta = 0.05$, a dark coupling $\alpha_D = 0.1$, and mediator mass given by $m_{A'} = 3m_1$. **Left:** Comparison of our current reach(red region) with existing constraints from BaBar (green) [119, 120] and LEP (blue) [98, 119, 121, 122]. **Right:** Projections of our constraints for a luminosity of 3ab^{-1} (red region). The solid and dot-dashed red contours correspond to the projections derived by assuming the same selections of the original CMS analysis, and the one derived by using a higher $N_{\rm hits}$ cut and assuming zero background. We compare our results with current constraints (gray shaded region) and projections for BelleII (pink) [115], SeaQuest (gray) [96], FASER (blue) [35], MATHUSLA (orange) [32], CODEX-b (brown) [37], LHC (yellow and purple) [119, 123, 124], and LHCb (green) [119, 125].

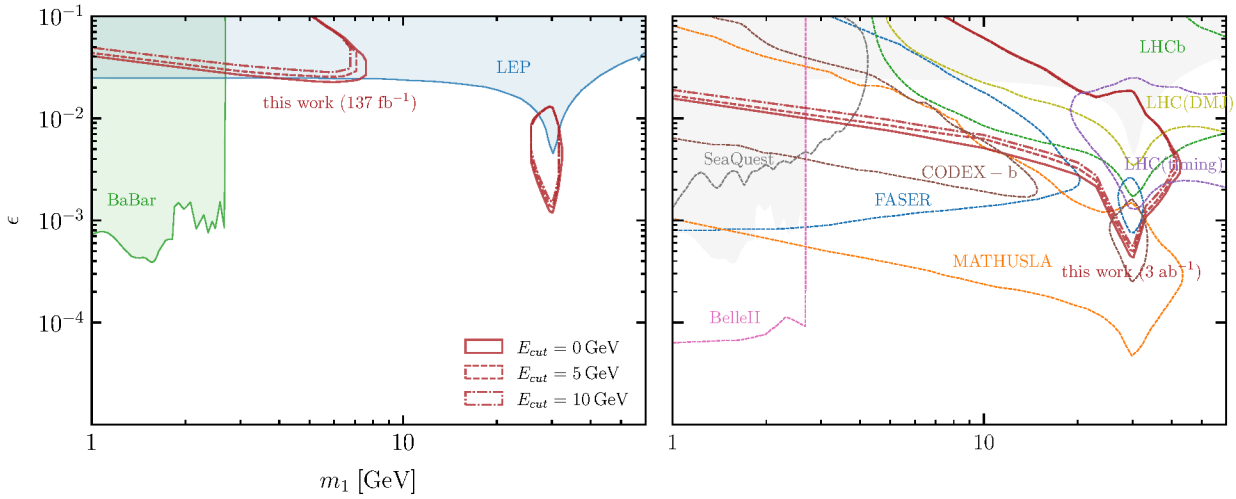

FIG. 12. Constraints on the inelastic DM model for different choices of the lower cut on the LLP energy. The other constraints appearing in the plot are the same reported in Figure 11. The projections in the right panel are for a luminosity of 3ab^{-1} luminosity, and assuming a tighter $N_{\rm hits}$ cut and zero background.

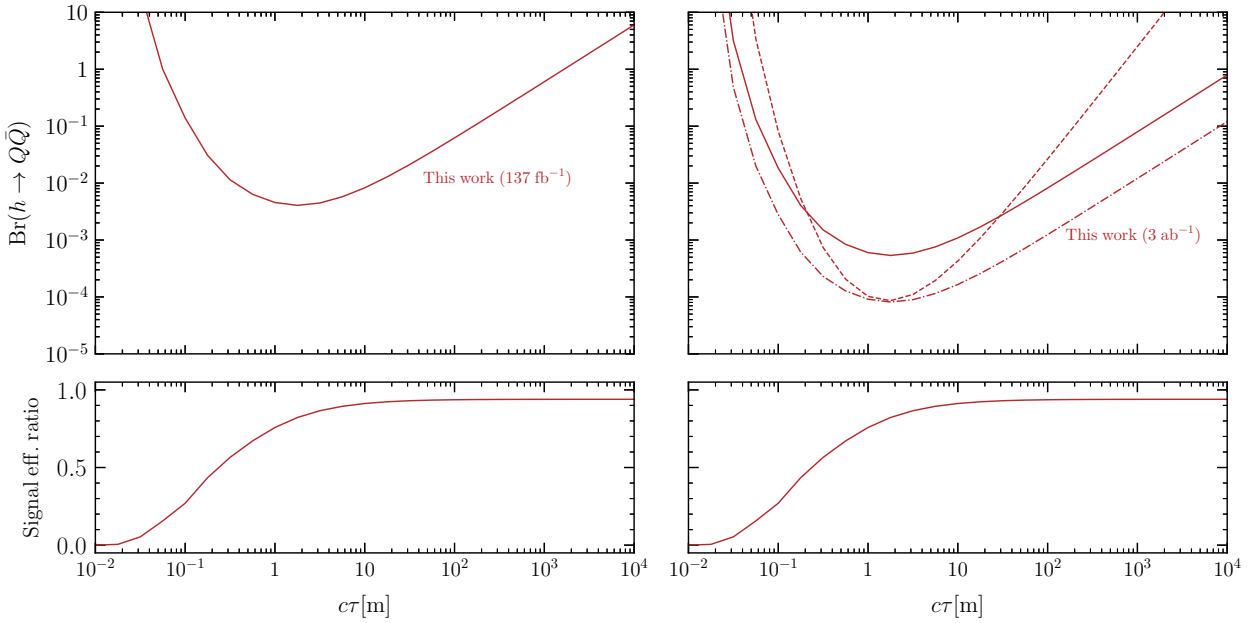

FIG. 13. Current (upper left panel) and projected (upper right panel) constraints on the Higgs exotic decay into dark quarks of a confining hidden valley model. For the projections we report (solid, dashed, and dot-dashed lines) the results obtained by using the three search strategies discussed in the main text (rescaled CMS analysis, dedicated trigger, and higher $N_{\rm hits}$).

Appendix A: ATLAS mono-jet limit for Axion-like Particles coupled to gluons

Here we summarize the procedure used to reinterpret the ATLAS monojet limit on ALPs coupled to gluons [77]. The ATLAS collaboration already provides a lower limit on the ALP decay constant at a fixed ALP mass $m_a = 1\,\text{MeV}$ in this particular model and claims that such limit should hold for ALP masses up to approximately 1 GeV. This claim is motivated from ALP literature prior to the improved estimates on ALP lifetimes and branching ratios provided in Ref. [69] and it is modified in the region $0.1-1\,\text{GeV}$ due to the non-trivial behavior from ALP mixing with the neutral pseudoscalar mesons. To estimate the limit curve in this region we use our ALP+jet simulation to extract the 2D LLP energy and pseudorapidity distributions, convolve that with the lifetime model of [69], and require that the ALP does not decay in the ATLAS detector volume, for a fixed value of m_a and f. We then rescale the ATLAS limit for the ratio of the two efficiencies described above computed at $m_a = 1\,\text{MeV}$ and at a different mass point. This produces a function of $(m_a/1\,\text{MeV}, f/f_{\text{limit,1MeV}})$. We then invert this function to solve for the limit on f as function of m_a as shown in Fig. 8. As expected the limit is fairly flat at low ALP masses but gets cut off earlier than 1 GeV due to the ALP lifetime significantly changing after the m_η threshold. The steepness of the turn-off renders this limit curve a little sensitive to the specific geometric dimensions considered for the ATLAS detector.

[1] M. Aaboud *et al.* (ATLAS), Phys. Rev. D **97**, 052012 (2018), arXiv:1710.04901 [hep-ex].
[2] R. Aaij *et al.* (LHCb), Eur. Phys. J. C **77**, 812 (2017), arXiv:1705.07332 [hep-ex].
[3] L. Lee, C. Ohm, A. Soffer, and T.-T. Yu, Prog. Part. Nucl. Phys. **106**, 210 (2019), [Erratum: Prog.Part.Nucl.Phys. 122, 103912 (2022)], arXiv:1810.12602 [hep-ph].
[4] M. Aaboud *et al.* (ATLAS), Phys. Rev. D **99**, 012001 (2019), arXiv:1808.03057 [hep-ex].
[5] J. Alimena *et al.*, J. Phys. G **47**, 090501 (2020), arXiv:1903.04497 [hep-ex].
[6] J. Beacham *et al.*, J. Phys. G **47**, 010501 (2020), arXiv:1901.09966 [hep-ex].
[7] A. M. Sirunyan *et al.* (CMS), Phys. Rev. D **100**, 112003 (2019), arXiv:1909.06166 [hep-ex].
[8] A. M. Sirunyan *et al.* (CMS), Phys. Lett. B **797**, 134876 (2019), arXiv:1906.06441 [hep-ex].
[9] A. M. Sirunyan *et al.* (CMS), Phys. Lett. B **806**, 135502 (2020), arXiv:2004.05153 [hep-ex].
[10] A. M. Sirunyan *et al.* (CMS), Phys. Rev. D **104**, 012015 (2021), arXiv:2012.01581 [hep-ex].
[11] G. Aad *et al.* (ATLAS), Phys. Rev. D **102**, 032006 (2020), arXiv:2003.11956 [hep-ex].
[12] G. Aad *et al.* (ATLAS), Phys. Rev. Lett. **127**, 051802 (2021), arXiv:2011.07812 [hep-ex].
[13] P. Agrawal *et al.*, Eur. Phys. J. C **81**, 1015 (2021), arXiv:2102.12143 [hep-ph].
[14] D. Acosta *et al.*, (2021), arXiv:2110.14675 [hep-ex].
[15] M. Borsato *et al.*, Rept. Prog. Phys. **85**, 024201 (2022), arXiv:2105.12668 [hep-ph].
[16] A. Tumasyan *et al.* (CMS), Eur. Phys. J. C **82**, 153 (2022), arXiv:2110.04809 [hep-ex].
[17] A. Tumasyan *et al.* (CMS), JHEP **04**, 062 (2022), arXiv:2112.13769 [hep-ex].
[18] A. M. Sirunyan *et al.* (CMS), Phys. Rev. D **104**, 052011 (2021), arXiv:2104.13474 [hep-ex].
[19] R. Aaij *et al.* (LHCb), Eur. Phys. J. C **82**, 373 (2022), arXiv:2110.07293 [hep-ex].

[20] R. Aaij *et al.* (LHCb), Eur. Phys. J. C **81**, 261 (2021), arXiv:2012.02696 [hep-ex].

[21] S. Knapen and S. Lowette, (2022), arXiv:2212.03883 [hep-ph].

[22] ATLAS Collaboration, (2022), arXiv:2209.01029 [hep-ex].

[23] G. Aad *et al.* (ATLAS), JHEP **06**, 005 (2022), arXiv:2203.01009 [hep-ex].

[24] G. Aad *et al.* (ATLAS), Phys. Rev. D **106**, 032005 (2022), arXiv:2203.00587 [hep-ex].

[25] A. Tumasyan *et al.* (CMS), JHEP **07**, 081 (2022), arXiv:2201.05578 [hep-ex].

[26] CMS Collaboration, (2022), arXiv:2205.08582 [hep-ex].

[27] ATLAS Collaboration, (2023), arXiv:2301.13866 [hep-ex].

[28] Y. Gershtein, Phys. Rev. D **96**, 035027 (2017), arXiv:1705.04321 [hep-ph].

[29] Y. Gershtein and S. Knapen, Phys. Rev. D **101**, 032003 (2020), arXiv:1907.00007 [hep-ex].

[30] E. Bernreuther, J. C. Mejia, F. Kahlhoefer, M. Krämer, and P. Tunney, JHEP **04**, 210 (2021), arXiv:2011.06604 [hep-ph].

[31] A. Tumasyan *et al.* (CMS), Phys. Rev. Lett. **127**, 261804 (2021), arXiv:2107.04838 [hep-ex].

[32] J. P. Chou, D. Curtin, and H. J. Lubatti, Phys. Lett. B **767**, 29 (2017), arXiv:1606.06298 [hep-ph].

[33] D. Curtin *et al.*, Rept. Prog. Phys. **82**, 116201 (2019), arXiv:1806.07396 [hep-ph].

[34] C. Alpigiani *et al.* (MATHUSLA), (2018), arXiv:1811.00927 [physics.ins-det].

[35] J. L. Feng, I. Galon, F. Kling, and S. Trojanowski, Phys. Rev. D **97**, 035001 (2018), arXiv:1708.09389 [hep-ph].

[36] A. Ariga *et al.* (FASER), Phys. Rev. D **99**, 095011 (2019), arXiv:1811.12522 [hep-ph].

[37] V. V. Gligorov, S. Knapen, M. Papucci, and D. J. Robinson, Phys. Rev. D **97**, 015023 (2018), arXiv:1708.09395 [hep-ph].

[38] M. Anelli *et al.* (SHiP), (2015), arXiv:1504.04956 [physics.ins-det].

[39] M. Bauer, O. Brandt, L. Lee, and C. Ohm, (2019), arXiv:1909.13022 [physics.ins-det].

[40] J. L. Feng *et al.*, (2022), arXiv:2203.05090 [hep-ex].

[41] V. V. Gligorov, S. Knapen, B. Nachman, M. Papucci, and D. J. Robinson, Phys. Rev. D **99**, 015023 (2019), arXiv:1810.03636 [hep-ph].

[42] G. Aielli *et al.*, Eur. Phys. J. C **80**, 1177 (2020), arXiv:1911.00481 [hep-ex].

[43] S. Dreyer *et al.*, (2021), arXiv:2105.12962 [hep-ph].

[44] A. Haas, C. S. Hill, E. Izaguirre, and I. Yavin, Phys. Lett. B **746**, 117 (2015), arXiv:1410.6816 [hep-ph].

[45] S. Cerci *et al.*, JHEP **2022**, 110 (2022), arXiv:2201.00019 [hep-ex].

[46] C. Ahdida *et al.* (SHiP), (2020), arXiv:2002.08722 [physics.ins-det].

[47] A. Boyarsky, O. Mikulenko, M. Ovchynnikov, and L. Shchutska, JHEP **03**, 006 (2022), arXiv:2104.09688 [hep-ph].

[48] B. Acharya *et al.*, in *2022 Snowmass Summer Study* (2022) arXiv:2209.03988 [hep-ph].

[49] CMS Collaboration, (2021), https://doi.org/10.17182/hepdata.104408.v2.

[50] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, Data Mining and Knowledge Discovery , 226 (1996).

[51] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, JHEP **07**, 079 (2014), arXiv:1405.0301 [hep-ph].

[52] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015), arXiv:1410.3012 [hep-ph].

[53] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07**, 001 (2003), arXiv:hep-ph/0206293.

[54] J. Alwall, S. de Visscher, and F. Maltoni, JHEP **02**, 017 (2009), arXiv:0810.5350 [hep-ph].

[55] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3), JHEP **02**, 057 (2014), arXiv:1307.6346 [hep-ex].

[56] C. Wang, "Dedicated Delphes Module: `https://github.com/delphes/delphes/pull/103`," (2022).

[57] G. Cottin, J. C. Helo, M. Hirsch, C. Peña, C. Wang, and S. Xie, JHEP **02**, 011 (2023), arXiv:2210.17446 [hep-ph].

[58] D. O'Connell, M. J. Ramsey-Musolf, and M. B. Wise, Phys. Rev. D **75**, 037701 (2007), arXiv:hep-ph/0611014.

[59] M. W. Winkler, Phys. Rev. D **99**, 015018 (2019), arXiv:1809.01876 [hep-ph].

[60] Y. Gershtein, S. Knapen, and D. Redigolo, Phys. Lett. B **823**, 136758 (2021), arXiv:2012.07864 [hep-ph].

[61] R. S. Willey and H. L. Yu, Phys. Rev. D **26**, 3287 (1982).

[62] R. S. Chivukula and A. V. Manohar, Phys. Lett. B **207**, 86 (1988), [Erratum: Phys.Lett.B 217, 568 (1989)].

[63] B. Grinstein, L. J. Hall, and L. Randall, Phys. Lett. B **211**, 363 (1988).

[64] B. Batell, M. Pospelov, and A. Ritz, Phys. Rev. D **83**, 054005 (2011), arXiv:0911.4938 [hep-ph].

[65] P. Ilten, Y. Soreq, M. Williams, and W. Xue, JHEP **06**, 004 (2018), arXiv:1801.04847 [hep-ph].

[66] M. Duerr, T. Ferber, C. Garcia-Cely, C. Hearty, and K. Schmidt-Hoberg, JHEP **04**, 146 (2021), arXiv:2012.08595 [hep-ph].

[67] I. Brivio, M. B. Gavela, L. Merlo, K. Mimasu, J. M. No, R. del Rey, and V. Sanz, Eur. Phys. J. C **77**, 572 (2017), arXiv:1701.05379 [hep-ph].

[68] A. Mitridate, M. Papucci, and Y. Wang, "Work in progress," (2022).

[69] D. Aloni, Y. Soreq, and M. Williams, Phys. Rev. Lett. **123**, 031803 (2019), arXiv:1811.03474 [hep-ph].

[70] M. J. Strassler and K. M. Zurek, Phys. Lett. B **651**, 374 (2007), arXiv:hep-ph/0604261.

[71] S. Knapen, J. Shelton, and D. Xu, Phys. Rev. D **103**, 115013 (2021), arXiv:2103.01238 [hep-ph].

[72] L. Carloni and T. Sjostrand, JHEP **09**, 105 (2010), arXiv:1006.2911 [hep-ph].

[73] F. Farchioni, I. Montvay, G. Munster, E. E. Scholz, T. Sudmann, and J. Wuilloud, Eur. Phys. J. C **52**, 305 (2007), arXiv:0706.1131 [hep-lat].

[74] M. Creutz, Annals Phys. **322**, 1518 (2007), arXiv:hep-th/0609187.

[75] A. Armoni, M. Shifman, and G. Veneziano, Phys. Rev. Lett. **91**, 191601 (2003), arXiv:hep-th/0307097.

[76] A. Tumasyan *et al.* (CMS), Eur. Phys. J. C **82**, 213 (2022), arXiv:2105.09178 [hep-ex].

[77] G. Aad *et al.* (ATLAS), Phys. Rev. D **103**, 112006 (2021), arXiv:2102.10874 [hep-ex].

[78] F. Bergsma *et al.* (CHARM), Phys. Lett. B **157**, 458 (1985).

[79] P. Schwaller, D. Stolarski, and A. Weiler, JHEP **05**, 059 (2015), arXiv:1502.05409 [hep-ph].

[80] "New ideas in detecting long-lived particles at the LHC Workshop," `https://indico.physics.lbl.gov/event/633/`.

[81] R. Aaij *et al.* (LHCb), Phys. Rev. D **95**, 071101 (2017), arXiv:1612.07818 [hep-ex].

[82] S. Foroughi-Abari and A. Ritz, Phys. Rev. D **102**, 035015 (2020), arXiv:2004.14515 [hep-ph].

[83] F. Bergsma *et al.* (CHARM), Phys. Lett. B **157**, 458 (1985).

[84] M. Aaboud *et al.* (ATLAS), Phys. Rev. D **99**, 052005 (2019), arXiv:1811.07370 [hep-ex].

[85] C. Alpigiani *et al.* (MATHUSLA), (2020), arXiv:2009.01693 [physics.ins-det].

[86] J. L. Feng, I. Galon, F. Kling, and S. Trojanowski, Phys. Rev. D **97**, 055034 (2018), arXiv:1710.09387 [hep-ph].

[87] R. Aaij *et al.* (LHCb), Phys. Rev. D **95**, 071101 (2017), arXiv:1612.07818 [hep-ex].

[88] J. P. Lees *et al.* (BaBar), Phys. Rev. Lett. **113**, 201801 (2014), arXiv:1406.2980 [hep-ex].

[89] A. Anastasi *et al.* (KLOE-2), Submitted to: Phys. Lett. B (2018), arXiv:1807.02691 [hep-ex].

[90] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **124**, 041801 (2020), arXiv:1910.06926 [hep-ex].

[91] J. R. Batley *et al.* (NA48/2), Phys. Lett. **B746**, 178 (2015), arXiv:1504.00607 [hep-ex].

[92] ATLAS Collaboration (ATLAS), (2022), arXiv:2206.12181 [hep-ex].

[93] J. Blumlein *et al.*, Int. J. Mod. Phys. **A7**, 3835 (1992).

[94] E. M. Riordan *et al.*, Phys. Rev. Lett. **59**, 755 (1987).

[95] D. Banerjee *et al.* (NA64), Phys. Rev. Lett. **125**, 081801 (2020), arXiv:2005.02710 [hep-ex].

[96] A. Berlin, S. Gori, P. Schuster, and N. Toro, Phys. Rev. D **98**, 035011 (2018), arXiv:1804.00661 [hep-ph].

[97] NA62 Collaboration, (2019).

[98] P. Ilten, Y. Soreq, J. Thaler, M. Williams, and W. Xue, Phys. Rev. Lett. **116**, 251803 (2016), arXiv:1603.08926 [hep-ph].

[99] W. Altmannshofer *et al.* (Belle-II), PTEP **2019**, 123C01 (2019), [Erratum: PTEP 2020, 029201 (2020)], arXiv:1808.10567 [hep-ex].

[100] S. Chakraborty, M. Kraus, V. Loladze, T. Okui, and K. Tobioka, Phys. Rev. D **104**, 055036 (2021), arXiv:2102.04474 [hep-ph].

[101] P. A. Zyla *et al.* (Particle Data Group), PTEP **2020**, 083C01 (2020).

[102] V. Chobanova *et al.* (Belle), Phys. Rev. D **90**, 012002 (2014), arXiv:1311.6666 [hep-ex].

[103] B. Aubert *et al.* (BaBar), Phys. Rev. Lett. **101**, 091801 (2008), arXiv:0804.0411 [hep-ex].

[104] J. P. Lees *et al.* (BaBar), Phys. Rev. D **84**, 012001 (2011), arXiv:1105.5159 [hep-ex].

[105] F. Ertas and F. Kahlhoefer, JHEP **07**, 050 (2020), arXiv:2004.01193 [hep-ph].

[106] M. J. Dolan, T. Ferber, C. Hearty, F. Kahlhoefer, and K. Schmidt-Hoberg, JHEP **12**, 094 (2017), [Erratum: JHEP 03, 190 (2021)], arXiv:1709.00009 [hep-ph].

[107] M. Acciarri *et al.* (L3), Phys. Lett. B **412**, 201 (1997).

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

[109] E. Anashkin *et al.* (DELPHI), (1999).

[110] M. Acciarri *et al.* (L3), Phys. Lett. B **353**, 136 (1995).

[111] D. Aloni, C. Fanelli, Y. Soreq, and M. Williams, Phys. Rev. Lett. **123**, 071801 (2019), arXiv:1903.03586 [hep-ph].

[112] I. Larin *et al.* (PrimEx), Phys. Rev. Lett. **106**, 162303 (2011), arXiv:1009.1681 [nucl-ex].

[113] F. Abudinén *et al.* (Belle-II), Phys. Rev. Lett. **125**, 161806 (2020), arXiv:2007.13071 [hep-ex].

[114] M. Dugger *et al.* (GlueX), (2012), arXiv:1210.4508 [hep-ex].

[115] W. Altmannshofer *et al.* (Belle-II), PTEP **2019**, 123C01 (2019), [Erratum: PTEP 2020, 029201 (2020)], arXiv:1808.10567 [hep-ex].

[116] J. D. Bjorken, S. Ecklund, W. R. Nelson, A. Abashian, C. Church, B. Lu, L. W. Mo, T. A. Nunamaker, and P. Rassmann, Phys. Rev. D **38**, 3375 (1988).

[117] J. Blumlein *et al.*, Z. Phys. C **51**, 341 (1991).

[118] J. Abdallah *et al.* (DELPHI), Eur. Phys. J. C **60**, 17 (2009), arXiv:0901.4486 [hep-ex].

[119] A. Berlin and F. Kling, Phys. Rev. D **99**, 015021 (2019), arXiv:1810.01879 [hep-ph].

[120] J. P. Lees *et al.* (BaBar), Phys. Rev. Lett. **119**, 131804 (2017), arXiv:1702.03327 [hep-ex].

[121] A. Hook, E. Izaguirre, and J. G. Wacker, Adv. High Energy Phys. **2011**, 859762 (2011), arXiv:1006.0973 [hep-ph].

[122] A. Pierce, B. Shakya, Y. Tsai, and Y. Zhao, Phys. Rev. D **97**, 095033 (2018), arXiv:1708.05389 [hep-ph].

[123] E. Izaguirre, G. Krnjaic, and B. Shuve, Phys. Rev. D **93**, 063523 (2016), arXiv:1508.03050 [hep-ph].

[124] J. Liu, Z. Liu, and L.-T. Wang, Phys. Rev. Lett. **122**, 131801 (2019), arXiv:1805.05957 [hep-ph].

[125] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **120**, 061801 (2018), arXiv:1710.02867 [hep-ex].