

**Diphotons from tetraphotons in the decay of a 125 GeV Higgs boson at the LHC**Patrick Draper<sup>1,\*</sup> and David McKeen<sup>2,†</sup><sup>1</sup>*Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, California 95064, USA*<sup>2</sup>*Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia V8P 5C2, Canada*

(Received 4 May 2012; published 26 June 2012)

Recently the ATLAS and CMS experiments have presented data hinting at the presence of a Higgs boson at  $m_h \simeq 125$  GeV. The best-fit  $h \rightarrow \gamma\gamma$  rate averaged over the two experiments is approximately  $2.1 \pm 0.5$  times the standard model prediction. We study the possibility that the excess relative to the standard model is due to  $h \rightarrow aa$  decays, where  $a$  is a light pseudoscalar that decays predominantly into  $\gamma\gamma$ . Although this process yields  $4\gamma$  final states, if the pseudoscalar has a mass of the order tens of MeV, the two photons from each  $a$  decay can be so highly collimated that they may be identified as a single photon. Some fraction of the events then contribute to an effective  $h \rightarrow \gamma\gamma$  signal. We study the constraints on the parameter space where the net  $h \rightarrow \gamma\gamma$  rate is enhanced over the standard model by this mechanism and describe some simple models that give rise to the pseudoscalar-photon interaction. Further tests and prospects for searches in the near future are discussed.

DOI: [10.1103/PhysRevD.85.115023](https://doi.org/10.1103/PhysRevD.85.115023)

PACS numbers: 14.80.Ec, 14.70.Bh, 14.80.Va

**I. INTRODUCTION**

Determining the nature of electroweak symmetry breaking has been a primary goal of particle physics for several decades. In the standard model (SM), electroweak symmetry breaking occurs when the neutral component of a single scalar weak isospin doublet possesses a vacuum expectation value. The Higgs boson corresponds to the physical excitations of this field. Recent experimental advances make it clear that we are entering a new phase in searches for the Higgs boson, with tantalizing excesses appearing in several SM-like Higgs search channels at both the CERN Large Hadron Collider and the Fermilab Tevatron. These hints point to a relatively light state with a mass in the range  $122.5 \lesssim m_h \lesssim 127.5$  GeV [1,2]. In the SM, such a light Higgs boson has an extremely narrow width of the order  $10^{-5} \times m_h$ . The tiny width of the SM Higgs makes it a sensitive probe of new physics beyond the SM, especially in the only loosely constrained scalar sector [3]. New states coupled to a SM-like Higgs can have appreciable effect on its decays without ruining the excellent agreement of the SM with data observed elsewhere thus far (see, e.g., [4]).

In fact, the excesses seen at the LHC and Tevatron may already be pointing toward interesting deviations from the SM predictions for the Higgs branching ratios. The largest statistical power thus far comes from the searches for  $h \rightarrow \gamma\gamma$  at the LHC, where the ATLAS experiment finds a  $2.8\sigma$  excess at  $m_h \simeq 126$  GeV with a best-fit signal strength relative to the SM of approximately  $2.0 \pm 0.8$  [1,5]. The CMS experiment finds a  $3.1\sigma$  excess in  $h \rightarrow \gamma\gamma$  at  $m_h \simeq 124$  GeV with a best-fit signal strength relative to the SM of approximately  $2.1 \pm 0.6$  [2,6]. Under the assumption that these excesses are due to the same new particle, we can

naively combine the diphoton rates from the two experiments and estimate

$$\text{Rate}(h \rightarrow \gamma\gamma)/\text{SM} \simeq 2.1 \pm 0.5. \quad (1)$$

This estimate is not rigorous and the error bar is far from conclusive. However, the uncertainties will continue to decrease as data accumulate, and the central value may very well remain high. Therefore, it is of interest to classify models that can alter the  $h \rightarrow \gamma\gamma$  signal and study their other predictions. A number of recent papers have contributed to this program, including discussions of the effects of superpartners on the diphoton rate [7], the effects of more general new fermion and scalar states [8], the effects of singlet-doublet mixing [9], the predictions in the case of minimal universal extra dimensions [10], interference effects from charged Higgs contributions [11], and the possibility that the signal is due to the Randall-Sundrum radion [12], to name a few.

In this paper we consider a different scenario in which a 125 GeV Higgs boson can appear to have a larger branching ratio into  $\gamma\gamma$ . We introduce a very light pseudoscalar of mass in the range  $m_a \epsilon (\sim 10 \text{ MeV}, \sim m_\pi)$ , and we allow the Higgs boson to decay in the channel  $h \rightarrow aa$ . The pseudoscalars produced in these decays are extremely boosted in the lab frame, causing their decay products to be quite collimated. Furthermore, for such light pseudoscalars, there are no kinematically available hadronic decay modes. As first studied in [13], the decay  $a \rightarrow \gamma\gamma$  induced by the effective coupling  $aF_{\mu\nu}\tilde{F}^{\mu\nu}$  can easily dominate, leading to a  $h \rightarrow aa \rightarrow 4\gamma$  signature (in contrast to the “buried Higgs” scenario where the dominant decays are hadronic [14]). Because of the large boost for the pseudoscalars produced in  $h$  decays, the photon pairs from each  $a$  can be so highly collimated that a significant fraction may be identified as single photons, even in the ATLAS and CMS detectors which are very good at distinguishing two closely separated photons (e.g. originating from  $\pi^0$  or  $\eta$  decays).

\*pidraper@ucsc.edu  
†mckeen@uvic.ca

A related scenario, although not in the context of Higgs boson decays, was considered in [15], which studied decays of a new heavy vector through pseudoscalars to extremely collimated “photon jets.” Since their vector mass is much larger than 125 GeV, they consider much heavier pseudoscalars. Our analysis of the conditions necessary for photon jets to be identified as single photons is complementary to that of [15] and provides further motivation for the detailed study of photon jets. Previously, the  $h \rightarrow aa \rightarrow 4\gamma$  scenario was considered in [3,13,16] and, in particular, the relevance of this channel to  $h \rightarrow \gamma\gamma$  searches at the Tevatron was studied in [13]. While Ref. [13] focused mainly on pseudoscalars heavier than 200 MeV and heavier Higgs bosons, we will find that a significant  $4\gamma \rightarrow 2\gamma$  fake rate at the LHC with a 125 GeV Higgs boson requires a pseudoscalar lighter than about the mass of the  $\pi^0$ . In contrast to previous works, we also study the interplay between the net expected  $h \rightarrow \gamma\gamma$  signal and the signal in other Higgs search channels, provide an analysis of the compatibility of the light pseudoscalar scenario with the current Higgs data from the LHC and Tevatron, and study in detail the constraints from CERN LEP and low-energy experiments. For further discussion of Higgs decays to light pseudoscalars, see, e.g., [17].

This paper is organized as follows. In Sec. II we discuss the basic features of the model. In Sec. III we analyze the changes in the expected rates for SM Higgs search channels at the LHC. We give a detailed estimate of the probability for a collimated photon pair to be identified as a single photon, and perform a basic statistical analysis of the constraints on the light pseudoscalar parameter space coming from current Higgs searches at the Tevatron and LHC. In Sec. IV we study the constraints following from the muon anomalous magnetic moment, direct searches at LEP for  $e^+e^- \rightarrow \gamma + \text{inv}$ , meson decays, beam dump experiments, and nuclear physics. In Sec. V we survey model building for the light  $a$  and its coupling to photons, and consider the possibility that the latter is generated by  $\tau$  loops, or by  $a - \pi^0$  mixing. In Sec. IV we discuss the prospects for directly producing pseudoscalars in the mass and coupling range of interest at Primakoff-type experiments. In Sec. VII we conclude, and an Appendix generalizes our main results to the case where  $a$  has a substantial invisible branching fraction.

## II. MODEL

In addition to the SM, we consider a real pseudoscalar  $a$  which is the pseudo Nambu-Goldstone boson of some spontaneously broken approximate global symmetry.<sup>1</sup> It

<sup>1</sup>A frequently discussed model containing a light pseudo-Nambu-Goldstone boson is the next-to-minimal supersymmetric standard model in the approximate Peccei-Quinn- or R-symmetric limits. In that model the pseudoscalar has sufficiently large fermionic couplings that in the mass range we consider, it is completely ruled out by low-energy experiments such as those considered in Sec. IV [18].

has a small mass  $m_a$  coming from small explicit breaking of the symmetry, and it interacts with the SM via

$$\mathcal{L}_{\text{int}} = \frac{1}{\Lambda^2} (\partial^\mu a)^2 H^\dagger H - \frac{e^2}{4M} a F^{\mu\nu} \tilde{F}_{\mu\nu}, \quad (2)$$

where  $H$  is the SM Higgs doublet,  $F^{\mu\nu}$  is the photon field strength, and  $e$  is the positron charge.  $\Lambda$  and  $M$  are scales describing the strength of the higher-dimensional operators.

Expanding around the Higgs vacuum expectation value  $v \simeq 246$  GeV,  $H^0 = (v + h)/\sqrt{2}$ , leads to a rate for the Higgs to decay to pseudoscalars of

$$\Gamma(h \rightarrow aa) = \frac{v^2 m_h^3}{32\pi\Lambda^4} = 1.18 \text{ MeV} \left( \frac{m_h}{125 \text{ GeV}} \right)^3 \left( \frac{\Lambda}{\text{TeV}} \right)^{-4}, \quad (3)$$

where we have assumed that  $m_a \ll m_h$ . Using Eq. (2) we can also calculate the rate for  $a$  to decay to  $\gamma\gamma$ ,

$$\begin{aligned} \Gamma(a \rightarrow \gamma\gamma) &= \frac{\pi\alpha^2 m_a^3}{4M^2} \\ &= 2.68 \times 10^{-8} \text{ MeV} \left( \frac{M}{10 \text{ GeV}} \right)^{-2} \left( \frac{m_a}{40 \text{ MeV}} \right)^3. \end{aligned} \quad (4)$$

Pseudoscalars produced by the decay of 125 GeV Higgs at rest then have a decay length of

$$\gamma c\tau \simeq 1.15 \text{ cm} \left( \frac{M}{10 \text{ GeV}} \right)^2 \left( \frac{m_a}{40 \text{ MeV}} \right)^{-4} \times \left( \frac{m_h}{125 \text{ GeV}} \right). \quad (5)$$

If we require  $\geq 90\%$  of the decays to occur inside the electromagnetic calorimeters of ATLAS and CMS ( $\simeq 1$  m radius),  $\gamma c\tau$  should be less than about a half meter. We can express the scale  $M$  in terms of the decay length,  $m_a$ , and  $m_h$ ,

$$M = 9.3 \text{ GeV} \left( \frac{\gamma c\tau}{1 \text{ cm}} \right)^{1/2} \left( \frac{m_a}{40 \text{ MeV}} \right)^2 \times \left( \frac{m_h}{125 \text{ GeV}} \right)^{-1/2}. \quad (6)$$

In the Appendix, we expand this simple model to allow for  $a$  to additionally decay invisibly. We discuss the benefits and further constraints that this entails there.

## III. HIGGS SIGNAL

The interactions described in Sec. II do not appreciably affect the production cross section of the SM Higgs boson at the LHC. In this section we study the effects of the decays  $h \rightarrow aa$ ,  $a \rightarrow \gamma\gamma$  on the expected Higgs-to-diphotons rate. We also estimate the constraints on such decays from existing SM Higgs searches at the Tevatron and LHC.

### A. Modified rates in SM Higgs search channels

We define the ratios of the Higgs branchings to photons and to SM fermions,  $f$ , or gauge bosons,  $V = W, Z$ , to their values in the SM as

$$\mathcal{B}(h \rightarrow \gamma\gamma)_{\text{eff}} = R_{\gamma\gamma} \times \mathcal{B}_{\text{SM}}(h \rightarrow \gamma\gamma), \quad (7)$$

$$\mathcal{B}(h \rightarrow f\bar{f}, VV) = R_{XX} \times \mathcal{B}_{\text{SM}}(h \rightarrow f\bar{f}, VV). \quad (8)$$

The presence of the  $h \rightarrow aa$  decay suppresses uniformly the rates into all SM final states, yielding

$$R_{XX} = 1 - \mathcal{B}(h \rightarrow aa). \quad (9)$$

Note that to avoid cluttering notation, and because the rescaling is the same in both cases, we use  $R_{XX}$  to represent both the fermionic and  $W, Z$  rescalings relative to the SM.

The suppression in Eq. (9) is also present for the pure  $\gamma\gamma$  final state; however, the subscript “eff” in Eq. (7) indicates that in pseudoscalar models there can be an additional, effective contribution to the measured  $h \rightarrow \gamma\gamma$  rate. Since the LHC detectors have finite resolution, a fraction of  $h \rightarrow aa \rightarrow 4\gamma$  events may have sufficiently boosted photon pairs that each pair is identified as a single photon. Assuming a SM production cross section for  $h$  and a 100% branching of  $a \rightarrow \gamma\gamma$ , the measured diphoton branching ratio will be

$$\mathcal{B}(h \rightarrow \gamma\gamma)_{\text{eff}} = \mathcal{B}(h \rightarrow \gamma\gamma) + \epsilon \times \mathcal{B}(h \rightarrow aa) \quad (10)$$

or

$$R_{\gamma\gamma} = 1 + \mathcal{B}(h \rightarrow aa) \left( \frac{\epsilon}{\mathcal{B}_{\text{SM}}(h \rightarrow \gamma\gamma)} - 1 \right). \quad (11)$$

In these formulas,  $\epsilon$  is the probability that both  $\gamma\gamma$  pairs are identified as single photons, so that four photons appear as two. We see that to achieve an effective diphoton rate greater than or equal to the SM rate requires

$$\epsilon \geq \mathcal{B}_{\text{SM}}(h \rightarrow \gamma\gamma) \simeq 0.0023 \quad (12)$$

for  $m_h = 125$  GeV.

### B. $4\gamma \rightarrow 2\gamma$ misidentification rate

Estimating  $\epsilon$  is complicated by several factors. We would like to be conservative in our estimate of the expected rate; on the other hand, underestimating the rate by too much might falsely indicate that some model points are allowed, when in fact the true rates are so large that the points are already ruled out by the LHC. We base our estimate of  $\epsilon$  on the ATLAS selection criteria used to identify isolated photons in their cut-based analysis, and we comment on differences with CMS.

ATLAS uses a number of calorimeter variables to parametrize the shape of an electromagnetic shower, which can then be used to discriminate true isolated photons from backgrounds. The background from isolated  $\pi^0 \rightarrow \gamma\gamma$  decays bears strong similarity to our  $a \rightarrow \gamma\gamma$  process,

and ATLAS efficiently vetoes isolated pions using information from the first layer of the calorimeter, which has finely segmented strips in  $\eta$  [19]. The primary discrimination variables in the first layer are  $\mathbf{E}_{\text{ratio}}$ , which is the difference in energies between two strips containing energy maxima normalized to their sum;  $\Delta\mathbf{E}$ , which measures the difference in energies between the strip with the second-largest energy maximum and the strip with the minimum energy between the first two maxima;  $\mathbf{F}_{\text{side}}$ , which is the fraction of the energy deposited in seven strips in  $\eta$  around the maximum that does not fall into the central three strips;  $\mathbf{w}_{s3}$ , which measures the energy deposition in the two strips adjacent in  $\eta$  to a strip with an energy peak, relative to the total energy in the three strips; and  $\mathbf{w}_{\text{stot}}$ , which generalizes  $w_{s3}$  to approximately 20 strips in  $\eta$  and two strips in  $\phi$  [20].

To simplify our analysis, we begin by restricting our attention to photons that do not convert to  $e^+e^-$  pairs in the tracker, and which are so highly collimated that a second energy maximum does not appear in the first-layer calorimeter strips. In this case  $\Delta E$  is set to 0,  $E_{\text{ratio}}$  is set to 1, and most of the energy will be deposited into just a few adjacent strips. Therefore we expect that the most sensitive variable will be  $w_{s3}$ , defined precisely as

$$w_{s3} \equiv \sqrt{\frac{\sum_i E_i (i - i_{\text{max}})^2}{\sum_i E_i}}, \quad (13)$$

where  $i$  labels the strips. On average the photon pair from a  $\pi^0$  decay generates a larger  $w_{s3}$  value than a true single-photon event and thus may be efficiently rejected. On the other hand, the probability that a photon pair passes the  $w_{s3}$  cut should increase with the collimation of the pair. To easily estimate  $\epsilon$ , we would like to approximate the cut on  $w_{s3}$  by a cut on the photon pair separation.

For unconverted events, ATLAS uses a weakly  $\eta$ -dependent cut on  $w_{s3}$  that is typically between 0.6–0.7, and is about 0.66 for the most central strips in the barrel. The average value of  $w_{s3}$  for true photons in these strips is about 0.58 [21]. We can reproduce this number with the following simple model. First, we assume that a single photon lands in the center of one of these most central strips, and deposits its energy according to a Gaussian distribution in  $\eta$  with standard deviation 0.52 times the smallest strip width. This value is chosen so that the photon deposits about 70% of the energy into the central strip and 15% into each adjacent strip, giving a  $w_{s3}$  value of 0.58. Subsequently, we compute  $w_{s3}$  for a pair of photons as a function of  $\eta$  separation, averaging over the impact point in the central strip. We find that for  $\Delta\eta_{\gamma\gamma} = 0.0015$ ,  $w_{s3} \simeq 0.66$ , equal to the ATLAS cut in these strips. On the other hand, the central strips have a width in  $\eta$  of  $\Delta\eta_{\text{strip}} = 0.0031$  (the smallest in the calorimeter.) Therefore, we estimate that the cut on  $w_{s3}$  can be simulated by a cut on the photon separation, given by

$$\Delta\eta_{\gamma\gamma} < 1/2 \times \Delta\eta_{\text{strip}}. \quad (14)$$

For larger  $\eta$ , where larger strips are present, we also require a separation less than 1/2 of the relevant strip size, which is probably mildly conservative since on average the energy leakage into adjacent strips from a single-photon hit will be less.

There are two simple ways in which  $w_{s3}$  may become insufficient. First, photons that are less highly collimated (such as those that appear in the decays of more massive pseudoscalars) may deposit energy in strips that are sufficiently separated (more than three strips apart) that  $w_{s3}$  becomes insensitive. We assume that such events are very efficiently rejected by the other first-layer discriminators described above, and therefore our cut on  $\Delta\eta_{\gamma\gamma}$  is still a good proxy for the cuts on those variables. Secondly, photons may be closely spaced in  $\eta$ , but more broadly spaced in  $\phi$ . Since the segmentation in  $\phi$  is much coarser than in  $\eta$ , such events must be very separated in  $\phi$  in order for cuts on the variable  $R_\phi$  (which measures the energy distribution across several second-layer  $\phi$  cells) to reject them. A cut on  $\Delta\phi_{\gamma\gamma}$  can simulate the cut on  $R_\phi$ , but since such events are geometrically rare, the net efficiency is quite insensitive to the precise value of this cut. For definiteness we set the cut on  $\Delta\phi_{\gamma\gamma}$  to be equal to the smallest second-layer cell size in  $\phi$  as a function of  $\eta$  ( $\Delta\phi_{\text{cell}} = 0.025$  for  $\eta < 1.4$ ):

$$\Delta\phi_{\gamma\gamma} < \Delta\phi_{\text{cell}}. \quad (15)$$

So far we have discussed only events which contain no converted photons. Conversions happen with an  $\eta$ - and  $E_T$ -dependent probability that ranges from about 10% at low  $\eta$  to more than 50% at larger  $\eta$  [20]. Since  $h \rightarrow aa$  events produce four photons in the final state instead of two, a larger fraction of our events will contain at least one converted photon than are contained in pure  $h \rightarrow \gamma\gamma$  events. In our conversion events, one cluster in the calorimeter may contain one  $e^+e^-$  pair and one  $\gamma$ , or two  $e^+e^-$  pairs, so we might imagine that these events would be easy to distinguish from pure single- $\gamma$  events and may even be vetoed. In the case with  $\gamma e^+e^-$ , there is a mismatch between the momentum measured by the tracker (sensitive only to the charged particles) and the energy deposit in the calorimeter. In the case with  $2e^+e^-$ , two conversion vertices may be reconstructed in the tracker. Both of these signatures have been considered for rejecting isolated  $\pi^0$  backgrounds [22]; however, at present, the ATLAS photon identification analysis does not use an  $E/p$  cut, and does not automatically veto events with multiple conversion vertices [20] (only the “best” reconstructed conversion vertex is used, and is determined by the conversion radius and the number of tracks associated with the vertex.) Furthermore, the cuts on the calorimeter variables are relaxed somewhat to accommodate the fact that the energy deposits for single-photon conversions tend to spread

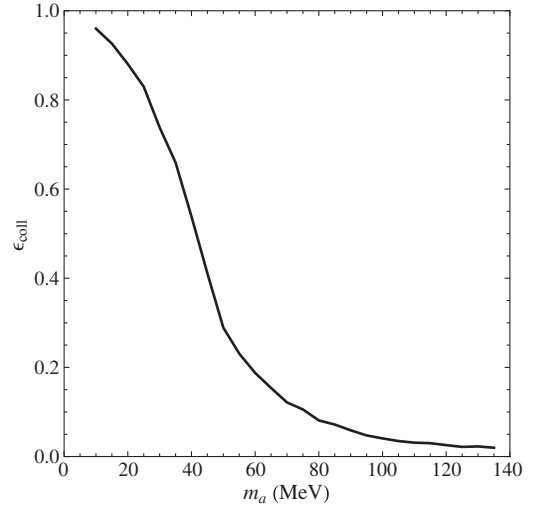


FIG. 1. Fraction of events where both photon pairs are sufficiently collimated to pass ATLAS isolation cuts, as a function of the pseudoscalar mass.

mildly in  $\eta$  and considerably in  $\phi$  (due to the magnetic field.) Therefore, we will make the approximation that the value of  $\epsilon$  relevant for  $4\gamma$  events containing conversions is the same as the value of  $\epsilon$  for the unconverted sample; namely, it is determined by applying the collimation cuts in Eqs. (14) and (15) to the sample of parent photons before they convert.

In summary, for all classes of photon events, we make the approximation

$$\epsilon \simeq \epsilon_{\text{coll}}, \quad (16)$$

where  $\epsilon_{\text{coll}}$  is the rate at which the  $4\gamma$  events are expected to pass the  $\eta$ -dependent collimation cuts  $\Delta\eta_{\gamma\gamma} < 1/2 \times \Delta\eta_{\text{strip}}$ ,  $\Delta\phi_{\gamma\gamma} < \Delta\phi_{\text{cell}}$ .

The collimation of a photon pair produced in an  $a$  decay is controlled by the ratio  $2m_a/E_a$ , where  $E_a$  is determined by the momentum of the Higgs boson. We simulate 7 TeV  $gg \rightarrow h \rightarrow aa \rightarrow 4\gamma$  events in Madgraph 5 [23], and on the subsample of events with all photons satisfying  $|\eta| \in (0, 1.37) \cup (1.52, 2.37)$  (the region in which ATLAS reconstructs photon candidates for the  $h \rightarrow \gamma\gamma$  analysis), and which pass the ATLAS photon  $p_T$  cuts [5], we compute the efficiency  $\epsilon_{\text{coll}}$ . In Fig. 1 we plot  $\epsilon_{\text{coll}}$  as a function of  $m_a$ .<sup>2</sup>

To cross-check that our cuts effectively reproduce the rejection power of  $w_{s3}$  and the other discrimination variables, we use the known isolated  $\pi^0$  rejection power of ATLAS as a function of  $(E_T, \eta)$  (given in Fig. 11 of

<sup>2</sup>Production mechanisms other than gluon fusion result in somewhat different kinematic distributions for  $h$ . However, gluon fusion dominates the production with O(10%) contribution from all other channels [5], so the effects on  $\epsilon_{\text{coll}}$  from these processes will be second-order. We neglect them in this study. In addition, we have checked that the effect of raising  $\sqrt{s}$  to 8 TeV is negligible.

Sec. V of [19]) to estimate  $\epsilon$  from our simulation at the mass point  $m_a = m_{\pi^0}$ . We find good agreement between the  $\epsilon$  derived from this method and  $\epsilon_{\text{coll}}$  obtained from the collimation cuts. Since the ATLAS  $\pi^0$  rejection power was defined on events containing both converted and unconverted photons, we are reassured that Eq. (16) is plausible.

Unlike the ATLAS analysis, the CMS analysis does place cuts on the ratio of the calorimeter energy measurement to the momentum measured in the tracker in order to isolate single photons from  $\pi^0$  decays [24]. The cut may reduce  $\epsilon$  relative to  $\epsilon_{\text{coll}}$ . However, the CMS electromagnetic calorimeter has a barrel granularity  $\simeq 6\times$  larger in  $\eta$  than that of ATLAS [25], increasing  $\epsilon_{\text{coll}}$ . For this analysis we assume that the net efficiencies  $\epsilon$  will be comparable between the two detectors, but in principle any difference between them translates into an experiment-dependent and possibly conversion-dependent prediction for the  $\gamma\gamma$  rate and could be used to probe the value of  $m_a$ .

To summarize this section:

- (1) High collimation is sufficient for a significant number of  $2\gamma$  clusters to be misidentified as single isolated photons by the experimental analyses.
- (2) In order to produce an interesting misidentification rate,  $m_a \lesssim m_{\pi^0}$ . For larger  $m_a$  there may still be misidentified events in the  $h \rightarrow \gamma\gamma$  signal, particularly if  $\mathcal{B}(h \rightarrow aa)$  is not small, but the scenario should be more properly studied in the context of an  $h \rightarrow 4\gamma$  signal.
- (3) Contamination of  $h \rightarrow \gamma\gamma$  by misidentified  $4\gamma$  events will be visible as an excess of the total number of conversion events relative to the  $\gamma\gamma$  expectation, as an excess of conversion events with multiple reconstructed vertices, and as an excess of conversion events with a mismatch between the track  $p_T$  and the energy deposit in the electromagnetic calorimeter.

### C. Current constraints and fits

In this study we are interested in the possibility that the current excesses in Higgs searches at ATLAS, CMS, and the Tevatron are due to a Higgs boson with mass near 125 GeV. Under this hypothesis, some parts of the  $(m_a, \mathcal{B}(h \rightarrow aa))$  parameter space are disfavored: they produce too few vector boson or fermion events, or too many or too few diphoton-like events. We estimate the constraints with a matched filter, taking each point in parameter space as a template and estimating a best-fit amplitude  $\hat{R}$  for the signal strength at the point.<sup>3</sup> In the Gaussian limit  $\hat{R}$  can be estimated as

$$\hat{R} = \sigma^2 t_i C_{ij}^{-1} d_j, \quad (17)$$

where  $d_j$  is the ‘‘data,’’ which we take to be the best-fit amplitudes relative to the SM for the channels  $h \rightarrow \gamma\gamma$  [5],

<sup>3</sup>For other recent studies of the Higgs best-fit cross section data using similar  $\chi^2$  analyses, see [26].

$h \rightarrow ZZ \rightarrow 4l$  [27],  $h \rightarrow WW \rightarrow l\nu l\nu$  [28],  $(W/Z)h \rightarrow (ll, l\nu, \nu\nu)b\bar{b}$  [29], and  $h \rightarrow \tau\tau$  [30], presented by ATLAS at  $m_h = 126$  GeV [31], the channels  $gg \rightarrow h \rightarrow \gamma\gamma$  [6,32],  $qqh \rightarrow qq\gamma\gamma$  [6,32],  $h \rightarrow ZZ \rightarrow 4l$  [33],  $h \rightarrow WW \rightarrow l\nu l\nu$  [34],  $(W/Z)h \rightarrow (ll, l\nu, \nu\nu)b\bar{b}$  [35], and  $h \rightarrow \tau\tau$  [36], presented by CMS at  $m_h = 124$  GeV [37], and the channels  $h \rightarrow b\bar{b}$  and  $h \rightarrow WW$  presented by CDF and DZero at  $m_h = 125$  GeV [38]. For ATLAS and CMS we chose slightly different mass points based on where their respective  $\gamma\gamma$  excesses are most significant; we assume that the current experimental resolution is large enough that both excesses can come from the same Higgs-like particle. The vector  $t_i$  gives the theoretical prediction for each data point (either  $R_{\gamma\gamma}$  or  $R_{XX}$ ), and  $C_{ij}^{-1}$  is the inverse covariance matrix set by the squared symmetrized error bars (we include also correlations which we estimate from the luminosity uncertainty and the theoretical uncertainty on the gluon fusion cross section). The error on the estimate  $\hat{R}$  is given by

$$\sigma \equiv (t_i C_{ij}^{-1} t_j)^{-1/2}. \quad (18)$$

Model points where  $R = 1$  is outside the 90% C.L. band around  $\hat{R}$  are then disfavored. Of course, this statistical procedure is quite approximate at this stage, and serves only to get an estimated exclusion region if the data sets are large and  $m_h$  actually is at 125 GeV. A precise calculation would construct the full Poisson likelihood with all correlations. With these caveats in mind, in Fig. 2 we present the regions of  $(m_a, \mathcal{B}(h \rightarrow aa))$  parameter space disfavored by the current LHC and Tevatron results, including contours of  $R_{\gamma\gamma}$  and  $R_{XX}$ . The contours make clear the fact that the disfavored regions are mainly controlled by the excesses in the diphoton channels: increasing it above a few times the SM rate is in tension with the current excesses. We also maximize the likelihood over the plane, fixing the signal strength parameter to  $R = 1$  and constraining  $m_a < m_{\pi^0}$ . The best-fit point lies near the  $R_{\gamma\gamma} = 2$  contour at the  $m_{\pi^0}$  boundary; relaxing the constraint, it would move beyond to near the intersection of the  $R_{\gamma\gamma} = 2$  and  $R_{XX} = 0.5$  contours. However, the data are over-fit,  $\chi^2/\text{d.o.f.} = 7.3/10$ , and the statistic is very shallow, particularly in the direction of constant  $R_{\gamma\gamma}$ . Therefore, the precise location of the best-fit point has little meaning, and we do not show it in Fig. 2. For comparison, we find that the SM  $\chi^2/\text{d.o.f.} = 12.2/12$ .

### IV. DIRECT CONSTRAINTS ON $a$

We now discuss existing experimental constraints on the model in Eq. (2). Similar considerations have been undertaken with light pseudoscalars in [18,39].

Constraints from  $e^+e^- \rightarrow \gamma a$ , quarkonia decays, and beam dump experiments are robust in the sense that they are controlled largely or entirely by the coupling of  $a$  to photons. Constraints from the muon anomalous magnetic moment and meson decays involving flavor-changing neutral currents are sensitive not only to the  $aF\tilde{F}$  coupling, but

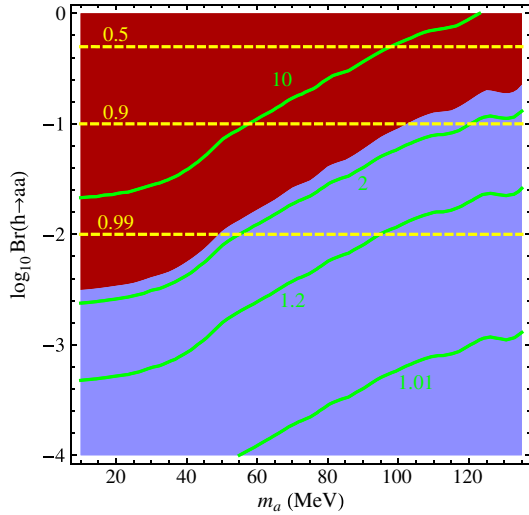


FIG. 2 (color online). Regions of light pseudoscalar parameter space that are favored (blue/light) and disfavored (red/dark) by the current best-fit signal strengths in the  $h \rightarrow \gamma\gamma, ZZ, WW, bb, \tau\tau$  channels. Contours are overlaid for the net diphoton (solid green lines) and  $ZZ, WW, bb, \tau\tau$  rates (dashed yellow lines) expected at the LHC relative to the SM rates.

also any small coupling of  $a$  to SM fermions that might be present in the Lagrangian. These contributions may interfere constructively or destructively with the  $aF\tilde{F}$  terms—we can even envision situations where some constraints are mitigated by finely tuned tree-level couplings of  $a$  to SM fermions. We assume for the purpose of this section that the fermionic couplings at the scale  $M$  are small enough that the leading contributions to the constraints come from photon loops.

Since  $aF\tilde{F}$  is dimension-5, the amplitudes are typically logarithmically divergent. To remove the divergence, we cut the integrals off at momenta  $\sim M$ , which is motivated by the view that the  $aF\tilde{F}$  interaction is an effective coupling resulting from some new physics at a scale  $M$ —whether  $a$  is a composite with constituents with masses of that order or is fundamental and coupled to photons via particles with such masses. The phenomenological need for  $M$  to be tens of GeV poses some model-building puzzles in this respect and we will sketch a few potential scenarios leading to such a scale in Sec. V.

Taken together, these low-energy constraints should be viewed mainly illustratively. We attempt to be as conservative as possible, erring on the side of presenting stronger limits on the parameter space of the model and comment that these limits can likely be circumvented with relatively straightforward extensions of the model under consideration.

### A. Muon anomalous magnetic moment

A light pseudoscalar that couples to two photons will contribute to the anomalous magnetic moment of the muon. The leading correction, shown in Fig. 3, occurs at

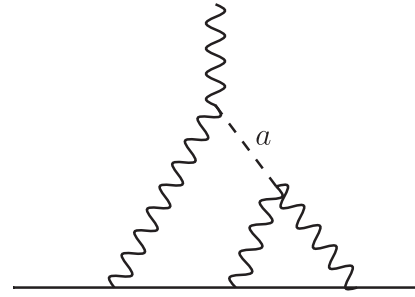


FIG. 3. A representative diagram of the leading contribution of the pseudoscalar,  $a$ , to  $(g-2)_\mu$ .

order  $\alpha^3$  and is analogous to the  $\pi^0$ -pole portion of the hadronic light-by-light contribution to  $(g-2)_\mu$ .

To estimate this effect as a function of  $m_a$  and  $M$ , we make use of the expression for the  $\pi^0$ -pole contribution to  $(g-2)_\mu$  in [40], rescale it by the strength of the photon coupling relative to that of  $\pi^0$ ,  $(4\pi^2 F_\pi/M)^2$ , replace the pion mass with  $m_a$ , and cut off the logarithmic divergence at the scale  $M$  instead of at  $\sim 4\pi^2 F_\pi$ . To set limits we require that this contribution is less than the current deviation between experiment and theory on  $\Delta a_\mu = (g-2)_\mu/2$  of about  $25 \times 10^{-10}$  [41]. This sets a lower limit on  $M$  around  $4\pi^2 F_\pi = 3.64$  GeV that is not very sensitive to  $m_a$  for the range of pseudoscalar masses we consider.

Because of the lepton mass dependence of corrections to  $g-2$ , the electron anomalous magnetic moment is less constraining and does not appear in our limits.

### B. $e^+e^- \rightarrow \gamma a$ limits

The reaction  $e^+e^- \rightarrow \gamma a$  can proceed through the  $s$ -channel exchange of a virtual photon. The cross section for this process is independent of the center-of-mass energy,

$$\frac{d\sigma}{d\cos\theta} = \frac{2\pi^2\alpha^3}{3M^2} \frac{3(1+\cos^2\theta)}{8}, \quad (19)$$

where  $\theta$  is the angle between the photon and the beam axis in the center of mass. If  $a$  decays without being detected, then this process is subject to limits from  $e^+e^- \rightarrow \gamma + \text{inv}$ . We apply limits from the DELPHI Collaboration [42] using  $650 \text{ pb}^{-1}$  of data at center-of-mass energies  $\sqrt{s} = 180 \text{ GeV} - 209 \text{ GeV}$ .<sup>4</sup> In the photon energy range of interest for  $m_a \ll \sqrt{s}$ ,  $E_\gamma \approx \sqrt{s}/2$ , we require that there are fewer than 20 events and assume that the product of

<sup>4</sup>We note here that the limit on  $e^+e^- \rightarrow \gamma + \text{inv}$  from the ASP experiment [43] that has been used in previous studies to constrain light pseudoscalars coupled to photons does not actually apply. This is because their analysis imposed a cut on the photon's energy of  $E_\gamma < \sqrt{s}/2 - 4.5 \text{ GeV} = 10 \text{ GeV}$  to eliminate backgrounds from  $e^+e^- \rightarrow \gamma\gamma$ , rendering it insensitive to invisibly decaying particles produced in association with a photon with a mass less than 16 GeV.

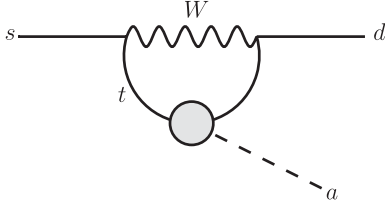


FIG. 4. Diagram that gives the leading contribution to  $s \rightarrow d + a$  from an effective interaction between  $a$  and the top quark.

angular acceptance and efficiency is  $\sim 0.5$ . Assuming that the  $a$  decay length must be larger than 2 m to be unseen at DELPHI and a boost of  $\gamma \sim 90 \text{ GeV}/m_a$  excludes the following region of parameter space:

$$110 \text{ GeV} \left( \frac{m_a}{40 \text{ MeV}} \right)^2 \lesssim M \lesssim 125 \text{ GeV}. \quad (20)$$

### C. Meson decays

The interactions in Eq. (2) lead to effective couplings of  $a$  to SM fermions at the one-loop level. This can have effects on rare meson decays involving photons or missing energy if  $a$  decays too late to be detected.

Effective couplings to up-type quarks can then lead to flavor-changing transitions in the down sector such as  $s \rightarrow d + a$  as in Fig. 4. This leads to the flavor-changing decay  $K^\pm \rightarrow \pi^\pm a$  which must be confronted with experimental data on  $K^\pm \rightarrow \pi^\pm \gamma \gamma$ . The amplitude for  $s \rightarrow d + a$  can be very roughly estimated as

$$\mathcal{M}(s \rightarrow d + a) \sim \frac{\alpha^2}{M} \log^2 \left( \frac{M^2}{m_t^2} \right) \left( \frac{G_F m_t^2}{4\sqrt{2}\pi^2} \right) \times V_{td}^* V_{ts} (m_s \bar{d}_L s_R - m_d \bar{d}_R s_L), \quad (21)$$

where we have again cut off the logarithmic divergence at the scale  $M$ . Using this we arrive at a rate for  $K^\pm \rightarrow \pi^\pm a$  of

$$\Gamma(K^\pm \rightarrow \pi^\pm a) \sim \frac{\alpha^4 m_s^2}{M^2} \log^4 \left( \frac{M^2}{m_t^2} \right) \left( \frac{G_F^2 m_t^4}{2^{10} \pi^5} \right) \frac{p_\pi}{m_K^2} |V_{td}^* V_{ts}|^2 \times \left( \frac{m_K^2 - m_\pi^2}{m_s - m_d} \right)^2 |f_0^{K\pi}(0)|^2, \quad (22)$$

where the form factor  $f_0^{K\pi}(0) \simeq 1$  [44] and  $p_\pi$  is the pion momentum in the kaon rest frame. We apply the limits on hypothetical particles with masses less than 100 MeV decaying to two photons from [45] (generally at the  $10^{-7}$  level) to  $\mathcal{B}(K^+ \rightarrow \pi^+ a)$ , resulting in  $M \gtrsim 1\text{--}2 \text{ GeV}$ . In addition, if  $a$  is sufficiently long-lived so that its decay is not detected, it will give a contribution to the process  $K^+ \rightarrow \pi^+ + \text{inv}$ . The current experimental upper limit on the branching ratio for  $K^+$  to decay to  $\pi^+$  and a light invisible particle is  $7.3 \times 10^{-11}$  [46]. For this limit to apply,  $a$  must have a decay length larger than about 1 m when produced with a  $\pi^+$  in the decay of a  $K^+$  at rest. The excluded region is then

$$1500 \text{ GeV} \left( \frac{m_a}{40 \text{ MeV}} \right)^2 \lesssim M \lesssim 17 \text{ GeV}. \quad (23)$$

Analogous limits from  $B^\pm \rightarrow K^\pm$  transitions are less stringent in the pseudoscalar mass range of interest.

$^3S_1$  quarkonium states can decay to  $\gamma a$  through an  $s$ -channel virtual photon, just as in  $e^+e^-$  annihilation. In the  $\bar{c}c$  and  $\bar{b}b$  systems, assuming  $m_a \ll m_{c,b}$ , the branching ratios are

$$\begin{aligned} \mathcal{B}(J/\psi \rightarrow \gamma a) &\simeq \frac{2\pi\alpha m_c^2}{M^2} \mathcal{B}(J/\psi \rightarrow e^+e^-) \\ &\simeq 4.6 \times 10^{-6} \left( \frac{10 \text{ GeV}}{M} \right)^2, \end{aligned} \quad (24)$$

$$\begin{aligned} \mathcal{B}(Y(1S) \rightarrow \gamma a) &\simeq \frac{2\pi\alpha m_b^2}{M^2} \mathcal{B}(Y(1S) \rightarrow e^+e^-) \\ &\simeq 2.3 \times 10^{-5} \left( \frac{10 \text{ GeV}}{M} \right)^2. \end{aligned} \quad (25)$$

We require that the contribution to  $J/\psi \rightarrow 3\gamma$  from  $J/\psi \rightarrow \gamma a$ ,  $a \rightarrow \gamma\gamma$  is less than the experimental result  $\mathcal{B}(J/\psi \rightarrow 3\gamma) = (1.2 \pm 0.4) \times 10^{-5}$  [47]. This translates into a limit  $M \gtrsim 6.2 \text{ GeV}$ . Current limits on  $\mathcal{B}(Y(1S) \rightarrow \gamma a)$  at the  $10^{-6}$  to  $10^{-5}$  level assume  $a$  decays to leptons or hadrons [48] and therefore are not constraining on a pseudoscalar that decays promptly to  $\gamma\gamma$ . In addition, if  $a$  is long-lived enough to decay after passing through the detector these decays are also subject to limits on  $J/\psi, Y(1S) \rightarrow \gamma + \text{inv}$  as in the  $e^+e^- \rightarrow \gamma + \text{inv}$  case in Sec. IV B. The current experimental limits are  $\mathcal{B}(J/\psi \rightarrow \gamma a \rightarrow \gamma + \text{inv}) < 4.6 \times 10^{-6}$  for  $m_a < 150 \text{ MeV}$  [49] and  $\mathcal{B}(Y(1S) \rightarrow \gamma a \rightarrow \gamma + \text{inv}) < 1.4 \times 10^{-5}$  for  $m_a < 5 \text{ GeV}$  [50]. To estimate the region of parameter space that these limits are sensitive to, we assume that the  $a$  decay length needs to be greater than 1 m to go undetected in these searches. This excludes the regions

$$590 \text{ GeV} \left( \frac{m_a}{40 \text{ MeV}} \right)^2 \lesssim M \lesssim 10 \text{ GeV}, \quad (26)$$

from  $J/\psi \rightarrow \gamma + \text{inv}$  and

$$330 \text{ GeV} \left( \frac{m_a}{40 \text{ MeV}} \right)^2 \lesssim M \lesssim 13 \text{ GeV}, \quad (27)$$

from  $Y(1S) \rightarrow \gamma + \text{inv}$ .

We note that the regions of  $(m_a, M)$  ruled out by  $K^+ \rightarrow \pi^+ + \text{inv}$  and  $J/\psi, Y(1S) \rightarrow \gamma + \text{inv}$  are all contained within the region excluded by  $e^+e^- \rightarrow \gamma + \text{inv}$  in Sec. IV B.

### D. Beam dump experiments

Beam dump experiments, where weakly coupled particles are searched for in the collision of proton or electron beams with fixed targets, also provide limits on light

pseudoscalars (often called axionlike particles in this context). To be successfully probed, the particles under consideration need to decay visibly and live long enough that they escape the target, but not so long that they do not decay before passing the detectors downstream, typically tens of meters away from the target. This results in mass-dependent exclusion bands in the pseudoscalar's lifetime.

Light pseudoscalars are produced in beam dump experiments either through the Primakoff process via their coupling to two photons or by being radiated via direct couplings to SM fermions as the beam constituents are stopped, in a manner analogous to bremsstrahlung. Each mode typically provides an  $\mathcal{O}(1)$  fraction of the total production cross section. The final states that are considered in searches for light pseudoscalars include  $\gamma\gamma$ ,  $e^+e^-$ , and  $\mu^+\mu^-$ .

In the model we consider, the pseudoscalars are predominantly produced by the Primakoff process since they do not directly couple to SM fermions and the dominant decay mode is  $\gamma\gamma$ . In placing limits on our parameter space, however, we simply apply the experiments' reported exclusions [51] to generic axionlike particles, ignoring subtleties in the slightly different production cross sections and different branching fractions. This approximation results in conservative exclusion regions since reducing the direct couplings to SM fermions should not make for more stringent limits. In the region of parameter space of interest we find upper limits on the pseudoscalar lifetime, corresponding to prompt decays inside the targets, or, equivalently, upper limits on the inverse coupling  $M$  in Eq. (2). The values of these upper limits tend to roughly coincide with those coming from the requirement that the pseudoscalar also decays promptly at the LHC.

### E. Further limits

The coupling of  $a$  to SM fermions becomes negligible in the nonrelativistic limit. Therefore, the strict constraints light scalars and vectors are subject to from diffractive low-energy neutron scattering on nuclei do not apply in this case. For the same reason, these particles are also not constrained by measurements of D–P transitions in muonic Si and Mg.

Additionally, light pseudoscalars can also be produced in the decays of excited nuclear states in nuclear reactors. However, this production mode is kinematically limited to  $m_a \lesssim 10$  MeV, and does not extend the exclusions we have considered.

### F. Allowed regions

In Fig. 5, we show regions of  $(M, m_a)$  excluded by the constraints outlined in Secs. IVA, IVB, IV C, IV D, and IV E. We also display contours of 1 cm and 50 cm decay lengths, assuming  $\gamma = m_h/2m_a$  with  $m_h = 125$  GeV.

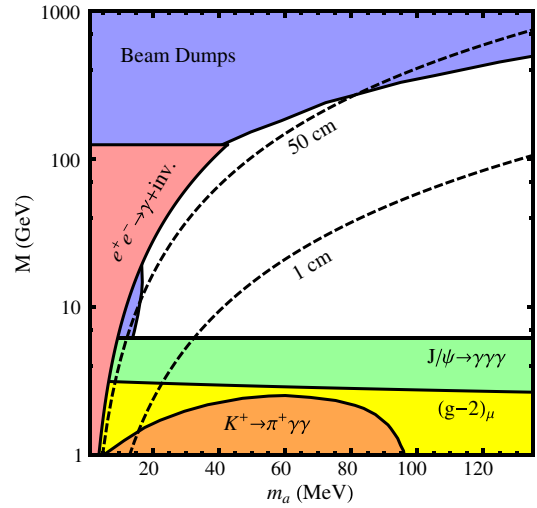


FIG. 5 (color online). Regions of pseudoscalar mass and inverse  $a\gamma\gamma$  coupling  $M$  excluded by the constraints in Secs. IVA, IVB, IV C, IV D, and IV E. The dashed curves are contours of 1 cm and 50 cm decay lengths for the pseudoscalar, assuming a boost  $\gamma = m_h/2m_a$  with  $m_h = 125$  GeV.

## V. MODEL BUILDING

Having investigated the phenomenology associated with a light pseudoscalar coupled to the Higgs, we speculate on a few scenarios where the  $aF\tilde{F}$  interaction in Eq. (2) could result. As we have seen, given a 125 GeV Higgs,  $m_a$  needs to be on the order of tens of MeV for the photons from its decay to be collimated enough to plausibly fake a single photon at the LHC. Using this fact in Eqs. (5) and (6), we observe that  $M$  cannot be at the TeV scale for  $a$  to decay promptly as well to evade constraints from beam dump experiments. Obtaining a relatively low scale for a dimension-5 operator such as this involving interactions beyond the SM poses a model-building challenge.

A simple scenario where an interaction  $aF\tilde{F}$  is generated is if  $a$  interacts with particles of mass  $m$  that have electric charge  $q$  in units of the electron's charge. One can imagine that  $a$  is either a composite made up of these charged particles or is fundamental and has a renormalizable coupling to them. After integrating out these charged particles, the dimension-5 interaction between  $a$  and two photons is generated with a scale  $M \sim 4\pi^2 m/q^2$ . Given the requirement that  $a$  decay promptly seems to require  $a$  to couple to charged particles with a mass not much larger than tens of GeV.<sup>5</sup> Candidates for such charged particles

<sup>5</sup>We note that a scenario where  $a \rightarrow \gamma\gamma$  is mediated by a coupling of  $a$  to a new, light vector boson,  $V$ , that kinetically mixes with the photon (thereby lifting the requirement that  $a$  couple to electrically charged particles) is not viable because the amplitude for such a process is proportional to  $q_1^2 q_2^2 (q_1^2 - m_V^2)^{-1} (q_2^2 - m_V^2)^{-1}$  where  $q_{1,2}$  are the photons' momenta, which vanishes for on-shell photons. If  $m_V$  vanishes, then  $a \rightarrow VV$  decays will strongly dominate the branching ratios.



are likely limited to SM fermions other than the  $t$  quark. Significant direct couplings to electrons or muons can be ruled out by anomalous magnetic moment measurements. Couplings to heavy quarks  $c$  and  $b$  likely cause issues with well-studied quarkonium transitions. We are left with  $\tau$  and light quarks as SM mediators of the  $aF\tilde{F}$  interaction. For a recent analysis of new light bosons decaying to  $\gamma\gamma$  see Ref. [52].

### A. Coupling to $\tau$

We can also imagine that the effective  $aF\tilde{F}$  interaction is due to a coupling of  $a$  to  $\tau$  leptons, writing this coupling as<sup>6</sup>

$$\mathcal{L}_{a\tau} = -ig_\tau a \bar{\tau} \gamma^5 \tau. \quad (28)$$

This leads to an effective operator mediating  $a \rightarrow \gamma\gamma$  in Eq. (2) with (assuming  $m_a \ll m_\tau$ )

$$M = \frac{4\pi^2 m_\tau}{g_\tau}. \quad (29)$$

The experimental information about the anomalous magnetic moment of  $\tau$  [53]

$$-0.052 < \frac{(g-2)_\tau}{2} < 0.013, \quad (30)$$

limits  $g_\tau \lesssim 2.5$  for  $m_a > 10$  MeV.

A similar limit of  $g_\tau \lesssim 2-3$  also comes from the OPAL Collaboration's result [54]

$$\mathcal{B}(Z \rightarrow \tau^+ \tau^- \gamma) < 7.3 \times 10^{-4}, \quad (31)$$

where  $Z \rightarrow \tau^+ \tau^- a \rightarrow \tau^+ \tau^- \gamma\gamma$  can mimic  $Z \rightarrow \tau^+ \tau^- \gamma\gamma$  events [52].

Without a dedicated search for these pseudoscalars in  $\tau$  decays, it is difficult to be certain about constraints on this scenario coming from  $\tau$  decay. This is especially true if  $a$  has a significant invisible width, a scenario which we discuss in the Appendix. For example, the branching ratio of  $\tau$  to  $a$  and  $\mu$  can be estimated to be [52]

$$\mathcal{B}(\tau^\pm \rightarrow a \mu^\pm \bar{\nu}) \simeq 3.5 \times 10^{-5} g_\tau^2. \quad (32)$$

Some of these decays could appear as  $\tau^\pm \rightarrow \mu^\pm \bar{\nu} \nu \gamma$  decays, given the decay  $a \rightarrow \gamma\gamma$ . Even with the large Yukawa allowed by  $(g-2)_\tau$  and if all of the  $\tau^\pm \rightarrow a \mu^\pm \bar{\nu}$  were registered as  $\tau^\pm \rightarrow \mu^\pm \bar{\nu} \nu \gamma$  decays, the contribution is still below the uncertainty on the measurement of this branching of  $(3.61 \pm 0.38) \times 10^{-3}$  [55] which agrees well with expectations. Limits from  $\tau^\pm \rightarrow e^\pm \bar{\nu} \nu \gamma$  are less stringent and hadronic  $\tau$  decays pose a bit more difficulty in prediction.

<sup>6</sup>We write the interaction in terms of a pseudoscalar instead of an axial vector coupling here for simplicity.

In this scenario, the scale  $M$  is bounded by

$$M = \frac{4\pi^2 m_\tau}{g_\tau} \gtrsim 35 \text{ GeV}. \quad (33)$$

### B. Mixing with $\pi^0$

Coupling  $a$  to light quarks so that  $a$  mixes with  $\pi^0$  could provide a solution. Given the couplings (ignoring  $s$  quarks for clarity)

$$\mathcal{L}_{aq} = -ig_u a \bar{u} \gamma^5 u - ig_d a \bar{d} \gamma^5 d, \quad (34)$$

using the partially conserved axial current, one can estimate the mixing angle,

$$\theta \sim \frac{(g_u - g_d) \langle \bar{q}q \rangle}{F_\pi (m_{\pi^0}^2 - m_a^2)} \simeq (g_u - g_d) \frac{(400 \text{ MeV})^2}{(m_{\pi^0}^2 - m_a^2)}, \quad (35)$$

where  $\langle \bar{q}q \rangle$  is the value of the light quark condensate and  $F_\pi = 92.2$  MeV is the pion decay constant. Here, the scale  $M$  is given by its analogue in the  $\pi^0$  case scaled by the mixing angle,

$$M = \frac{4\pi^2 F_\pi}{\sin\theta}. \quad (36)$$

If there are no symmetry conditions on  $g_u$  and  $g_d$ , then we will in general expect additional isospin violation due to interactions involving  $a$ . The relative size of these effects is of the order  $g_{u,d}^2 / (m_{u,d}^2 / F_\pi^2)$ . If we demand that these not exceed  $\sim 1\%$  then we require  $g_{u,d} \lesssim 10^{-3}$ . Furthermore, one must then contend with corrections to pion beta decay  $\pi^\pm \rightarrow \pi^0 e^\pm \nu$  which is shifted by

$$\Gamma(\pi^\pm \rightarrow \pi^0 e^\pm \nu) = \cos^2 \theta \Gamma_{\text{SM}}(\pi^\pm \rightarrow \pi^0 e^\pm \nu). \quad (37)$$

The 0.6% agreement of theory and experiment for this rate [56] implies that  $\theta \lesssim 0.08$ . Such a bound is compatible with the limits above on the Yukawa couplings from isospin violation.

The current measurement of the  $\pi^0$  lifetime is at the 3% level [57], and therefore does not pose any stricter constraints on the strength of this mixing. Furthermore, given a mixing angle of this size, the contribution to  $\pi^0 \rightarrow e^+ e^-$  from  $a-\pi^0$  mixing is negligible.

There are potentially strict limits in this scenario coming from the  $K^+ \rightarrow \pi^+ \gamma\gamma$  limits described in Sec. IV C since the rate for  $K^+ \rightarrow \pi^+ a$  is enhanced. The rate can be estimated as

$$\begin{aligned} \mathcal{B}(K^+ \rightarrow \pi^+ a) &\simeq \sin^2 \theta \mathcal{B}(K^+ \rightarrow \pi^+ \pi^0) \\ &= \sin^2 \theta \times 0.21. \end{aligned} \quad (38)$$

The limits on this branching at the  $10^{-7}$  level [45] imply that  $\theta \lesssim 10^{-3}$ , or equivalently,  $M \gtrsim 1$  TeV. This does not allow for prompt decays at the LHC if  $a$  is produced in the decay of a 125 GeV Higgs and is problematic with respect to beam dump constraints. However, the analysis in Ref. [45] only selected events corresponding to

$m_a < 100$  MeV because of the very large background from  $K^+ \rightarrow \pi^+ \pi^0$ . Thus, to avoid the strict limit on  $M$  that results from this analysis,  $a$  must have a mass greater than 100 MeV if its coupling to photons is mediated by a mixing with  $\pi^0$ .

The limit from pion beta decay implies that  $M \geq 47$  GeV in this case.

### C. Other scenarios

In addition to the two scenarios described above we briefly mention a few other potential models that could lead to the interactions in Eq. (2) with a scale  $M$  on the order of tens to a few hundred GeV. As in [13], the pseudoscalar could be coupled to heavy vectorlike matter that mediates its interaction with photons. However, ensuring that the scale  $M$  is small enough in this situation so that  $a$  decays promptly is difficult. This could be solved by giving the heavy matter a large electric charge or gauging it under a non-Abelian group with a large number of colors. We also speculate that sterile neutrinos with very large transition magnetic moments coupled to  $a$  could also offer a solution since their masses need not be at the weak scale.

## VI. OUTLOOK

As discussed in Sec. III B, the scenario analyzed in this paper can be tested in  $h \rightarrow \gamma\gamma$  events, since the expected rate depends sensitively on the photon identification criteria used in the analysis. Furthermore, the presence of more photons implies that the fraction of events with a photon conversion is higher than in a pure  $\gamma\gamma$  sample. Although CMS has already presented the best-fit rates in conversion and unconverted events separately, the statistics remain low, and we leave such analysis for the future. We also note that in the case where the  $a$  is long-lived,  $\gamma c\tau \sim 50$  cm, it is imaginable that the decay length may be determined from the distribution of conversion radii, even though since the states are highly boosted they do not display displaced vertices in the traditional sense.

The light pseudoscalar hypothesis can also be tested by looking for a  $3\gamma$  signal with invariant mass matching that of the Higgs boson. If the Higgs is produced with a sizable boost, some decays will yield one slow  $a$  and one fast  $a$  in the lab frame, and the former will produce widely separated photons. Alternatively, if  $m_a$  is much larger than the values studied here, the pseudoscalar can still contribute some events to  $h \rightarrow 2\gamma$  if  $\mathcal{B}(h \rightarrow aa)$  is large enough, but then the particle may be more easily seen in searches for  $h \rightarrow 4\gamma$  [16].

A different way to search for a light pseudoscalar coupling to photons is to produce the  $a$  directly in Primakoff-type experiments,  $\gamma + \text{Nuc} \rightarrow a + \text{Nuc}$ , which would allow the mass of  $a$  to be directly reconstructed. A proposal for an upgrade to the PrimEx Experiment at Jefferson Lab to measure  $\Gamma(\eta \rightarrow \gamma\gamma)$  envisions collisions of photons with an energy  $E_\gamma \simeq 11$  GeV on a liquid hydrogen target

with a luminosity of  $\mathcal{L} \sim 10^{-2} \text{ nb}^{-1} \text{ s}^{-1}$  with a run of 45 days [58]. Using the interaction in Eq. (2), the cross section for the Primakoff production of  $a$  in the collision of a photon with a proton can be found to be

$$\frac{d\sigma}{dQ^2} \simeq \frac{2\pi^2\alpha^3}{M^2} \frac{|F_{\text{em}}(Q^2)|^2}{Q^2}, \quad (39)$$

where  $Q$  is the momentum transferred to the proton and  $F_{\text{em}}$  is its electromagnetic form factor. Using a simple dipole form factor and a proton charge radius of 0.87 fm (the results do not depend sensitively on the form factor and charge radius—using 0.84 fm as measured in muonic hydrogen does not change the estimate) and a detection efficiency of 60%, this cross section would lead to

$$N(a) \simeq 10^4 \left( \frac{10 \text{ GeV}}{M} \right)^2 \quad (40)$$

pseudoscalars collected with  $m_a = 40$  MeV. For comparison, there would be about  $10^4$   $\eta$ 's produced via the Primakoff process and subsequently decaying to two photons under the same conditions.

It appears likely, at least statistically, that pseudoscalars with parameters in the range that is interesting in the context of the  $h \rightarrow \gamma\gamma$  signal can be probed at future Primakoff experiments.

## VII. CONCLUSIONS

In this paper we have investigated a simple model that can give rise to an apparent excess relative to the SM in the  $h \rightarrow \gamma\gamma$  channel. In our study the two-photon excess comes from Higgs decays to two light, boosted pseudoscalars,  $h \rightarrow aa$ , and each pseudoscalar decays into two photons. For very light pseudoscalars, each pair of photons is highly collimated and a non-negligible fraction of the pairs can appear as single photons, even at high-resolution detectors like those at the LHC. This scenario serves as an example where photon jets [15] are produced and helps to motivate the future experimental study of these objects in more detail.

We have estimated the fraction of  $4\gamma$  events that appear as  $2\gamma$ , taking into account the fine granularity of the ATLAS detector and the tight photon selection criteria. Additionally, we have investigated subtleties that may arise when one photon or more converts. We have assumed that the different analyses used at CMS and ATLAS do not give rise to large deviations between the  $h \rightarrow \gamma\gamma$  measurements, but we note that the amount of  $4\gamma$  contamination can be analysis-dependent and thus the differing analyses may be used as a test of the hypothesis.

In addition to a diphoton excess, the model also predicts a deficit relative to the SM for Higgs decays into other final states. Although the uncertainties are large, the data that have been collected so far at the LHC and Tevatron are consistent with both of these predictions, and with more data the model will be easily tested. The relative size of a

potential excess in the  $\gamma\gamma$  channel and a decrement in the remaining channels will be a direct probe of the mass of the pseudoscalar, with a lighter pseudoscalar better able to accommodate a larger signal in the non- $\gamma\gamma$  channels for a fixed  $\gamma\gamma$  rate. Lower bounds around  $m_a \simeq 10$  MeV from  $\gamma + \text{inv}$  searches and beam dump experiments offer a complimentary sensitivity to the model.

Finally, the mass of the light pseudoscalar can be well measured at Primakoff experiments. As we discussed in Sec. IV, future Primakoff experiments appear well poised to test models involving new light bosons coupled to photons like the one we have considered, underscoring the complementarity of experiments at the intensity frontier with high-energy studies of the Higgs boson.

## ACKNOWLEDGMENTS

The authors would like to thank J. Albert, V. Bansal, K. Jensen, M. Lefebvre, J. Mitrevski, J. Nielsen, M. Pospelov, J. Redondo, and A. Ritz for helpful discussions. The work of D.M. was supported by NSERC, Canada. P.D. is supported by the DOE under Grant No. DE-FG02-04ER41286.

## APPENDIX A: INVISIBLE DECAYS

In this Appendix, we describe a simple extension of the model in Sec. II to allow for  $a$  to decay invisibly. The Higgs phenomenology and direct constraints change somewhat from the case where  $a$  decays purely to  $\gamma\gamma$ . As before, we present fits to the Higgs data and show allowed regions of parameter space.

### 1. Framework

In addition to the two-photon decay channel, the pseudoscalar  $a$  could have an appreciable invisible decay width. As a concrete example, we can add an interaction with a stable (on collider scales) SM singlet Dirac fermion  $\chi$ ,

$$\mathcal{L}_{a\chi} = \frac{\partial_\mu a}{M'} \bar{\chi} \gamma^\mu \gamma^5 \chi, \quad (\text{A1})$$

where  $M'$  is another scale describing the strength of this interaction. The rate for  $a \rightarrow \chi\bar{\chi}$  is then

$$\Gamma(a \rightarrow \bar{\chi}\chi) = \frac{m_\chi^2 m_a}{8\pi M'^2} \beta_\chi. \quad (\text{A2})$$

When this channel is open, the branching ratio for  $a \rightarrow \gamma\gamma$  becomes

$$\mathcal{B}(a \rightarrow \gamma\gamma) = \left[ 1 + \frac{\beta_\chi}{2} \left( \frac{1}{\pi\alpha} \frac{m_\chi M'}{m_a} \right)^2 \right]^{-1}, \quad (\text{A3})$$

where  $\beta_\chi^2 = 1 - 4m_\chi^2/m_a^2$ . The decay length (for a boost  $\gamma = m_h/2m_a$ ) is shortened by a factor of this branching ratio,

$$\begin{aligned} \gamma c\tau &\simeq 1.15 \text{ mm} \left( \frac{M}{10 \text{ GeV}} \right)^2 \left( \frac{m_a}{40 \text{ MeV}} \right)^{-4} \\ &\times \left( \frac{m_h}{125 \text{ GeV}} \right) \left( \frac{\mathcal{B}(a \rightarrow \gamma\gamma)}{0.1} \right), \end{aligned} \quad (\text{A4})$$

and the scale  $M$  can then be expressed as

$$\begin{aligned} M &= 29 \text{ GeV} \left( \frac{\gamma c\tau}{1 \text{ cm}} \right)^{1/2} \left( \frac{m_a}{40 \text{ MeV}} \right)^2 \left( \frac{m_h}{125 \text{ GeV}} \right)^{-1/2} \\ &\times \left( \frac{\mathcal{B}(a \rightarrow \gamma\gamma)}{0.1} \right)^{-1/2}. \end{aligned} \quad (\text{A5})$$

For a fixed decay length, a sizable invisible branching fraction for  $a$  allows the scale  $M$  to be somewhat larger than it would be in the case where  $a$  only decays to photons.

### 2. Higgs signal

The primary changes to Higgs phenomenology are in  $\mathcal{B}(h \rightarrow \gamma\gamma)_{\text{eff}}$  and  $R_{\gamma\gamma}$ , now given by

$$\mathcal{B}(h \rightarrow \gamma\gamma)_{\text{eff}} = \mathcal{B}(h \rightarrow \gamma\gamma) + \epsilon \mathcal{B}(h \rightarrow aa) \mathcal{B}(a \rightarrow \gamma\gamma)^2 \quad (\text{A6})$$

and

$$R_{\gamma\gamma} = 1 + \mathcal{B}(h \rightarrow aa) \left( \frac{\epsilon \mathcal{B}(a \rightarrow \gamma\gamma)^2}{\mathcal{B}_{\text{SM}}(h \rightarrow \gamma\gamma)} - 1 \right). \quad (\text{A7})$$

Fixing  $\mathcal{B}(a \rightarrow \gamma\gamma) = 0.1$ , we regenerate the exclusion contours and contours of fixed  $R_{\gamma\gamma, XX}$  in Fig. 6. The best-fit  $\chi^2/\text{d.o.f.} = 6.5/10$ , and the best-fit point is close to the intersection of the  $R_{\gamma\gamma} = 2$  and  $R_{XX} = 0.5$  contours. Like the  $\mathcal{B}(a \rightarrow \gamma\gamma) = 1$  case, the  $\chi^2$  is shallow and the best-fit

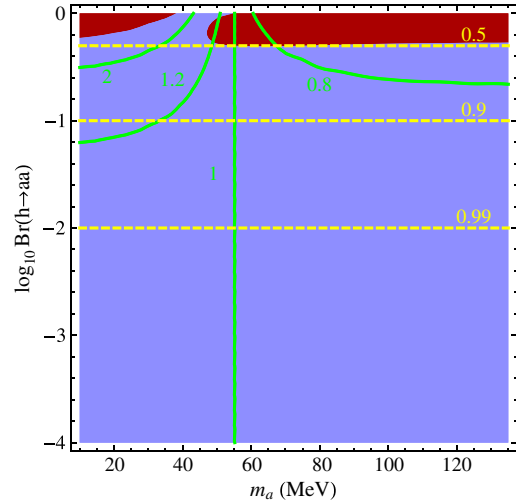


FIG. 6 (color online). Regions of light pseudoscalar parameter space that are favored (blue/light) and disfavored (red/dark) by the current best-fit signal strengths in the  $h \rightarrow \gamma\gamma, ZZ, WW, bb, \tau\tau$  channels. Contours are overlaid for the net diphoton (solid green lines) and  $ZZ, WW, bb, \tau\tau$  rates (dashed yellow lines) expected at the LHC relative to the SM rates.  $\mathcal{B}(a \rightarrow \gamma\gamma)$  is fixed to 0.1, indicating a 90% branching of  $a \rightarrow \text{invisible}$ .

point is not yet a meaningful quantity, so we do not show it in Fig. 6.

In addition to an increased apparent branching to  $\gamma\gamma$ , there are contributions to the invisible width of the Higgs, which can be related to the increased diphoton branching ratio,

$$\begin{aligned} \mathcal{B}(h \rightarrow \text{invisible}) &= \mathcal{B}(h \rightarrow aa)\mathcal{B}(a \rightarrow \bar{\chi}\chi)^2 \\ &= \mathcal{B}(h \rightarrow aa)[1 - \mathcal{B}(a \rightarrow \gamma\gamma)]^2. \end{aligned} \quad (\text{A8})$$

There is also a ‘‘monophoton’’ branching,

$$\begin{aligned} \mathcal{B}(h \rightarrow \gamma + \text{invisible}) &= \mathcal{B}(h \rightarrow aa)\mathcal{B}(a \rightarrow \gamma\gamma) \\ &\quad \times \mathcal{B}(a \rightarrow \bar{\chi}\chi) \times 2\epsilon' \\ &= 2\epsilon'\mathcal{B}(h \rightarrow aa)\mathcal{B}(a \rightarrow \gamma\gamma) \\ &\quad \times [1 - \mathcal{B}(a \rightarrow \gamma\gamma)], \end{aligned} \quad (\text{A9})$$

where  $\epsilon'$  is the efficiency for the one photon jet in the event to be reconstructed as a single photon. Future LHC monophoton searches with low MET cuts may be able to probe this channel.

### 3. Direct constraints on $a$

We now analyze changes to the direct limits given in Secs. IVA, IVB, IVC, IVD, and IVE that arise if  $a$  has an appreciable invisible branching fraction.

Signals involving missing energy now apply regardless of  $a$ 's decay length. The  $e^+e^- \rightarrow \gamma + \text{inv}$  limit then becomes the strongest lower limit on  $M$ .

The limits on  $M$  from beam dumps must be rescaled by a factor of  $\mathcal{B}(a \rightarrow \gamma\gamma)^{-1/2}$  since they are limits on the  $a$  lifetime. We conservatively ignore any shrinking of the exclusion region due to the smaller visible branching ratio.

An invisibly decaying  $a$  could contribute to star cooling by providing an additional channel for energy to leave a star. However, such limits are not important in the mass range that we consider here.

If  $a$  mixes with the  $\pi^0$  and can decay invisibly, there are stringent limits on the invisible decay rate of the  $\pi^0$ ,

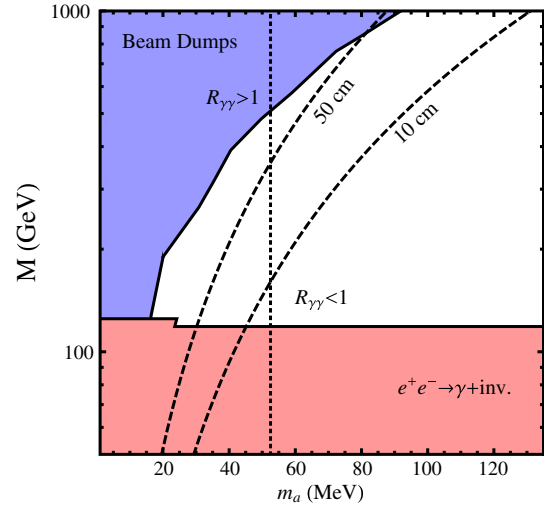


FIG. 7 (color online). Same as Fig. 5 in the case that  $\mathcal{B}(a \rightarrow \gamma\gamma) = 0.1$ . The dashed curves are contours of 10 and 50 cm decay lengths for  $\gamma = m_h/2m_a$  and  $m_h = 125$  GeV. To the left of the vertical dotted line,  $R_{\gamma\gamma} > 1$  [Eqs. (7) and (11)] while to its right,  $R_{\gamma\gamma} < 1$ . In this case, only the beam dump and  $e^+e^- \rightarrow \gamma + \text{inv}$  limits, which are the most constraining, are shown.

with the collider upper limit measured to be  $2.7 \times 10^{-7}$  [59] and a limit from big bang nucleosynthesis that is several orders of magnitude stronger [60]. These are quite constraining and make the scenario where the  $aF\tilde{F}$  coupling is generated by a mixing with  $\pi^0$  described in Sec. VB unlikely if  $a$  has an appreciable invisible branching fraction.

A large invisible branching fraction for  $a$  renders the model where the coupling to photons is mediated by a coupling to  $\tau$  more plausible. This is because  $\tau$  decays necessarily involve at least one neutrino in the final state so observing missing energy in a  $\tau$  decay does not signal new physics or a rare SM decay.

We reproduce the low-energy exclusions of Fig. 5 in Fig. 7 for  $\mathcal{B}(a \rightarrow \gamma\gamma) = 0.1$ , showing only the most stringent limits from beam dumps and  $e^+e^- \rightarrow \gamma + \text{inv}$ .

- 
- [1] S. Kortner, in Proceedings of Rencontres de Moriond, 2012 (unpublished).
- [2] M. Pieri, in Proceedings of Rencontres de Moriond, 2012 (unpublished).
- [3] S. Chang, R. Dermisek, J.F. Gunion, and N. Weiner, *Annu. Rev. Nucl. Part. Sci.* **58**, 75 (2008).
- [4] R. Dermisek and J.F. Gunion, *Phys. Rev. D* **76**, 095006 (2007).
- [5] G. Aad *et al.* (ATLAS Collaboration), *Phys. Rev. Lett.* **108**, 111803 (2012).
- [6] S. Chatrchyan *et al.* (CMS Collaboration), *Phys. Lett. B* **710**, 403 (2012).
- [7] M. Carena, S. Gori, N. R. Shah, and C. E. Wagner, *J. High Energy Phys.* **03** (2012) 014; S. Heinemeyer, O. Stal, and G. Weiglein, *Phys. Lett. B* **710**, 201 (2012); A. Arbey, M. Battaglia, and F. Mahmoudi, *Eur. Phys. J. C* **72**, 1847 (2012); J. Cao, Z. Heng, J. M. Yang, Y. Zhang, and J. Zhu, *J. High Energy Phys.* **03** (2012) 086; D. A. Vasquez, G. Belanger, C. Boehm, J. Da Silva, P. Richardson *et al.*, arXiv:1203.3446.
- [8] V. Barger, M. Ishida, and W.-Y. Keung, arXiv:1203.3456.
- [9] U. Ellwanger, *J. High Energy Phys.* **03** (2012) 044; Z. Kang, J. Li, and T. Li, arXiv:1201.5305.

- [10] G. Belanger, A. Belyaev, M. Brown, M. Kakizaki, and A. Pukhov, [arXiv:1201.5582](https://arxiv.org/abs/1201.5582).
- [11] A. Arhrib, R. Benbrik, M. Chabab, G. Moultaqa, and L. Rahili, *J. High Energy Phys.* **04** (2012) 136; A. Arhrib, R. Benbrik, and N. Gaur, *Phys. Rev. D* **85**, 095021 (2012).
- [12] K. Cheung and T.-C. Yuan, *Phys. Rev. Lett.* **108**, 141602 (2012).
- [13] B. A. Dobrescu, G. L. Landsberg, and K. T. Matchev, *Phys. Rev. D* **63**, 075003 (2001).
- [14] E. L. Berger, C.-W. Chiang, J. Jiang, T. M. P. Tait, and C. E. M. Wagner, *Phys. Rev. D* **66**, 095001 (2002); B. Bellazzini, C. Csaki, A. Falkowski, and A. Weiler, *ibid.* **80**, 075008 (2009); A. Falkowski, D. Krohn, L.-T. Wang, J. Shelton, and A. Thalapillil, *ibid.* **84**, 074022 (2011).
- [15] N. Toro and I. Yavin, [arXiv:1202.6377](https://arxiv.org/abs/1202.6377).
- [16] S. Chang, P. J. Fox, and N. Weiner, *Phys. Rev. Lett.* **98**, 111802 (2007).
- [17] B. A. Dobrescu and K. T. Matchev, *J. High Energy Phys.* **09** (2000) 031.
- [18] S. Andreas, O. Lebedev, S. Ramos-Sanchez, and A. Ringwald, *J. High Energy Phys.* **08** (2010) 003.
- [19] G. Aad *et al.* (The ATLAS Collaboration), [arXiv:0901.0512](https://arxiv.org/abs/0901.0512).
- [20] G. Aad *et al.* (Atlas Collaboration), *Phys. Rev. D* **83**, 052005 (2011).
- [21] ATLAS Collaboration Public Note, Report No. ATLAS-PHYS-PUB-2011-007, 2011 [<http://cdsweb.cern.ch/record/1345329/files/ATLAS-PHYS-PUB-2011-007.pdf>].
- [22] ATLAS Collaboration, “*Atlas Calorimeter Performance: Technical Design Report*,” CERN, Geneva, 1996.
- [23] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *J. High Energy Phys.* **06** (2011) 128.
- [24] CMS Collaboration Public Note, Report No. CMS-PAS-EGM-10006, 2012 [<http://cdsweb.cern.ch/record/1324545/files/EGM-10-006-pas.pdf>].
- [25] G. Bayatian *et al.* (CMS Collaboration), *J. Phys. G* **34**, 995 (2007).
- [26] D. Carmi, A. Falkowski, E. Kuik, and T. Volansky, [arXiv:1202.3144](https://arxiv.org/abs/1202.3144); J. Espinosa, C. Grojean, M. Muhlleitner, and M. Trott, *J. High Energy Phys.* **05** (2012) 097; P. P. Giardino, K. Kannike, M. Raidal, and A. Strumia, [arXiv:1203.4254](https://arxiv.org/abs/1203.4254); A. Azatov, R. Contino, and J. Galloway, *J. High Energy Phys.* **04** (2012) 127.
- [27] G. Aad *et al.* (ATLAS Collaboration), *Phys. Lett. B* **710**, 383 (2012).
- [28] ATLAS Collaboration Public Note, Report No. ATLAS-CONF-2012012, 2012 [<http://cdsweb.cern.ch/record/1429660/files/ATLAS-CONF-2012-012.pdf>].
- [29] ATLAS Collaboration Public Note, Report No. ATLAS-CONF-2012015, 2012 [<http://cdsweb.cern.ch/record/1429664/files/ATLAS-CONF-2012-015.pdf>].
- [30] ATLAS Collaboration Public Note, Report No. ATLAS-CONF-2012-014, 2012 [<http://cdsweb.cern.ch/record/1429662/files/ATLAS-CONF-2012-014.pdf>].
- [31] ATLAS Collaboration Public Note, Report No. ATLAS-CONF-2012-019, 2012 [<http://cdsweb.cern.ch/record/1430033/files/ATLAS-CONF-2012-019.pdf>].
- [32] CMS Collaboration Public Note, Report No. CMS-PAS-HIG-12-001, 2012 [<http://cdsweb.cern.ch/record/1429931/files/HIG-12-001-pas.pdf>].
- [33] S. Chatrchyan *et al.* (CMS Collaboration), *Phys. Rev. Lett.* **108**, 111804 (2012).
- [34] S. Chatrchyan *et al.* (CMS Collaboration), *Phys. Lett. B* **710**, 91 (2012).
- [35] S. Chatrchyan *et al.* (CMS Collaboration), *Phys. Lett. B* **710**, 284 (2012).
- [36] S. Chatrchyan *et al.* (CMS Collaboration), *Phys. Lett. B* **713**, 68 (2012).
- [37] S. Chatrchyan *et al.* (CMS Collaboration), *Phys. Lett. B* **710**, 26 (2012).
- [38] TEVNPHWorking Group, CDF, and D0 Collaborations, [arXiv:1203.3774](https://arxiv.org/abs/1203.3774).
- [39] E. Masso and R. Toldra, *Phys. Rev. D* **52**, 1755 (1995); J. Redondo, [arXiv:0807.4329](https://arxiv.org/abs/0807.4329).
- [40] M. Knecht and A. Nyeler, *Phys. Rev. D* **65**, 073034 (2002).
- [41] T. Teubner, K. Hagiwara, R. Liao, A. D. Martin, and D. Nomura, *Chinese Phys. C* **34**, 728 (2010); M. Davier, A. Hoecker, B. Malaescu, and Z. Zhang, *Eur. Phys. J. C* **71**, 1 (2011).
- [42] J. Abdallah *et al.* (DELPHI Collaboration), *Eur. Phys. J. C* **38**, 395 (2005); **60**, 17 (2009).
- [43] C. Hearty, J. Rothberg, K. Young, A. Johnson, J. S. Whitaker *et al.*, *Phys. Rev. D* **39**, 3207 (1989).
- [44] W. J. Marciano and Z. Parsa, *Phys. Rev. D* **53**, R1 (1996).
- [45] P. Kitching *et al.* (E787 Collaboration), *Phys. Rev. Lett.* **79**, 4079 (1997).
- [46] V. Anisimovsky *et al.* (E949 Collaboration), *Phys. Rev. Lett.* **93**, 031801 (2004).
- [47] G. Adams *et al.* (CLEO Collaboration), *Phys. Rev. Lett.* **101**, 101801 (2008).
- [48] W. Love *et al.* (CLEO Collaboration), *Phys. Rev. Lett.* **101**, 151802 (2008); D. McKeen, *Phys. Rev. D* **79**, 015007 (2009).
- [49] J. Insler *et al.* (CLEO Collaboration), *Phys. Rev. D* **81**, 091101 (2010).
- [50] R. Balest *et al.* (CLEO Collaboration), *Phys. Rev. D* **51**, 2053 (1995).
- [51] F. Bergsma *et al.* (CHARM), *Phys. Lett.* **157B**, 458 (1985); E. M. Riordan *et al.*, *Phys. Rev. Lett.* **59**, 755 (1987); M. Davier and H. Nguyen Ngoc, *Phys. Lett. B* **229**, 150 (1989); A. Bross, M. Crisler, S. Pordes, J. Volk, S. Errede, and J. Wrbanek, *Phys. Rev. Lett.* **67**, 2942 (1991).
- [52] D. McKeen, M. Pospelov, and J. M. Roney, *Phys. Rev. D* **85**, 053002 (2012).
- [53] J. Abdallah *et al.* (DELPHI Collaboration), *Eur. Phys. J. C* **35**, 159 (2004).
- [54] P. Acton *et al.* (OPAL Collaboration), *Phys. Lett. B* **273**, 338 (1991).
- [55] T. Bergfeld *et al.* (CLEO Collaboration), *Phys. Rev. Lett.* **84**, 830 (2000).
- [56] D. Pocanic, E. Frlez, V. A. Baranov, W. H. Bertl, C. Bronnimann *et al.*, *Phys. Rev. Lett.* **93**, 181803 (2004).
- [57] I. Larin *et al.* (PrimEx Collaboration), *Phys. Rev. Lett.* **106**, 162303 (2011).
- [58] A. Gasparian *et al.* (PrimEx Collaboration), [http://www.jlab.org/primex/primex\\_notes/PR-09-015.pdf](http://www.jlab.org/primex/primex_notes/PR-09-015.pdf).
- [59] A. V. Artamonov *et al.* (E949 Collaboration), *Phys. Rev. D* **72**, 091102 (2005).
- [60] W. P. Lam and K. W. Ng, *Phys. Rev. D* **44**, 3345 (1991).