

# Probing axionlike particles near the neutral pion mass with KOTO data

Reuven Balkin<sup>1</sup>,<sup>2</sup>,<sup>3</sup> Stefania Gori<sup>1</sup>,<sup>2</sup>,<sup>3</sup> Dean J. Robinson<sup>1</sup>,<sup>2</sup>,<sup>3</sup> and Christiane Scherb<sup>1</sup>,<sup>2</sup>,<sup>3</sup>

<sup>1</sup>*Department of Physics, University of California Santa Cruz and Santa Cruz Institute for Particle Physics, 1156 High Street, Santa Cruz, California 95064, USA*

<sup>2</sup>*Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA*

<sup>3</sup>*Berkeley Center for Theoretical Physics, Department of Physics, University of California, Berkeley, California 94720, USA*



(Received 22 August 2025; accepted 12 March 2026; published 22 April 2026)

We demonstrate that novel limits on prompt axionlike particles (ALPs) in the hard-to-probe mass range near the neutral pion—the so-called pion chimney—may be obtained from recasting  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  data taken by the J-PARC KOTO experiment, to search for  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$ . We also explore the power of KOTO  $6\gamma$  data to probe  $K_L \rightarrow 2\pi^0 a$  for a broader range of ALP masses, incorporating displaced decays.

DOI: [10.1103/hpzj-v6dl](https://doi.org/10.1103/hpzj-v6dl)

## I. INTRODUCTION

It is notoriously difficult to probe axionlike particles (ALPs) in a narrow interval of ALP masses nearby to the neutral pion, and with effective couplings  $c/f_a \gtrsim 10 \text{ TeV}^{-1}$ : the so-called pion chimney.<sup>1</sup> This difficulty arises for several interrelated reasons: (i) at the effective theory level, ALP-pion mixing is resonantly enhanced for  $m_a \simeq m_{\pi^0}$ . Thus, for couplings at which the ALP is usually a long-lived state, it instead becomes prompt, severely limiting the sensitivity of searches in which it is assumed to be invisible in the detector acceptance, as well as the sensitivity of direct searches for  $a \rightarrow \gamma\gamma$  (i.e., with a negligible electron branching ratio) that require some displacement of the ALP decay in order to separate the signal from  $\pi^0 \rightarrow \gamma\gamma$  backgrounds. This results in an upper limit on the ALP couplings that can be probed. Further, (ii) limits on prompt  $a \rightarrow \gamma\gamma$  from recasts of precision diphoton measurements are generically unable to probe ALP masses near  $m_{\pi^0}$ , because such measurements feature cuts on the diphoton invariant mass,  $m_{\gamma\gamma}$ , in order to avoid being overwhelmed by  $\pi^0 \rightarrow \gamma\gamma$  backgrounds. As an example, the KTeV data for  $K_L \rightarrow \pi^0 \gamma\gamma$  [1], which otherwise sets powerful limits on ALP couplings, applies an exclusionary cut on the invariant mass window  $100 < m_{\gamma\gamma} < 160 \text{ MeV}$ , and thus no limit can be derived in that mass regime.

<sup>1</sup>At least, so-called by S. Knapen.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

In this paper we show that one may leverage KOTO data for  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  to obtain novel limits on  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$  for prompt ALPs specifically in the pion chimney, which we take to be  $120 \leq m_a \leq 150 \text{ MeV}$ . We focus on a KOTO dataset corresponding to  $2 \times 10^{14}$  protons on target (POT) collected during the first run and used for the neutral flux calibration measurement, and for which  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  data are public. In contradistinction to most other search concepts, this approach specifically targets ALPs nearby the pion mass, because only their  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$  decays are able to pass the tight experimental selections for the  $K_L$  reconstruction. We show such an ALP signal would induce distortions in the precisely measured distribution of the  $K_L$  reconstructed mass,  $m_{K_L}^{\text{rec}}$ . In particular, we show the ALP signal generates single or multipeak structures that are shifted with respect to the peak at  $m_{K_L}$ , and thus enable novel limits on the underlying ALP couplings to be derived. (This approach might also be applied to the  $K_L \rightarrow 2\pi^0 \rightarrow 4\gamma$  flux calibration data, but it is likely much more challenging because of the lower statistics and  $K_L \rightarrow 3\pi^0$  backgrounds, plus likely degradation of the reconstruction efficiency under the details of the KOTO selection. See also Ref. [2] for a proposed  $K_L \rightarrow \pi^0 a$  search at KOTO outside the pion chimney region.) These limits are distinct from other KOTO searches, such as for  $K_L \rightarrow \pi^0 X$ , with  $X$  missing energy [3], or  $K_L \rightarrow XX \rightarrow 4\gamma$  [4]. The techniques developed in this work might also be used to derive limits from other  $K_L \rightarrow 3\pi^0$  datasets, such as from KTeV [5], provided sufficient detail regarding experimental selections and reconstruction is available to permit a recast. We leave the study of the adaptability and amenability of these techniques to other experiments for future study.

The  $K_L \rightarrow 2\pi^0 a$  decay probes a similar variety of new physics (NP) couplings as  $K_L \rightarrow \pi^0 a$ , but with differing sensitivities depending on the  $P$  and/or  $CP$  violating nature of the NP current. Very close to the neutral pion mass  $K_L \rightarrow 2\pi^0 a$  can provide a stronger probe of certain NP models than  $K_L \rightarrow \pi^0 a$  or even  $K^+ \rightarrow \pi^+ a$ , because of suppressed  $CP$  violation and weak interaction-induced resonant mixing, respectively ([6], see Fig. 5 therein). Thus, constraining ALP couplings via  $K_L \rightarrow 2\pi^0 a$  searches is phenomenologically well motivated. For further discussion on the connection of ALPs to flavor physics and flavor probes of ALPs, see, e.g., [7–31].

This work is structured as follows. In Sec. II we outline the KOTO detector simulation, event selection, and  $K_L$  reconstruction. To simulate  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$  contributions to the  $m_{K_L}^{\text{rec}}$  spectrum, we develop a simplified simulation of the KOTO selection and reconstruction, and verify that it faithfully reproduces the observed  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  signal peak. In Sec. III, we develop a phenomenological approach to parametrizing the line shape profile of the standard model (SM) backgrounds in the measured  $m_{K_L}^{\text{rec}}$  distribution, and obtain novel limits on  $\text{Br}[K_L \rightarrow 2\pi^0 a]$  within the pion chimney regime. In Sec. IV, we proceed to explore the power of KOTO to constrain ALPs lighter than the neutral pion, via direct searches in the reconstructed diphoton invariant mass, including the effects of displaced decays in flight of longer-lived ALPs. (See Ref. [32] for a discussion of KOTO bounds on long-lived ALPs produced instead in the KOTO target.) Section V concludes.

## II. SELECTION AND $K_L$ RECONSTRUCTION

### A. KOTO detector system

Details of the KOTO beamline and detector system configuration can be found in Refs. [33,34], which also detail the  $K_L \rightarrow 3\pi^0$  neutral flux calibration measurements that are central to the analysis in this paper (in particular, the  $m_{K_L}^{\text{rec}}$  distribution provided in Fig. 10 of Ref. [33]). For the purposes of this work, the salient features of the experiment are as follows:  $K_L$ 's are generated by a production target (T1) approximately 20.6 m upstream from the beam exit (BE) into the front barrel of the detector. The fiducial  $K_L$  decay volume (DV) is approximately 2.9–6.3 m downstream from the BE. The CsI electromagnetic calorimeter (CsIC) is located approximately 7.1 m downstream from the BE, see Fig. 1.<sup>2</sup> The  $K_L$  momentum distribution is measured at the BE, which is an  $8.5 \times 8.5 \text{ cm}^2$  square aperture. The geometric acceptance of the CsIC is a disc 1.9 m in diameter. It has a  $24 \times 24 \text{ cm}^2$  square beam hole at its center, and makes use of two

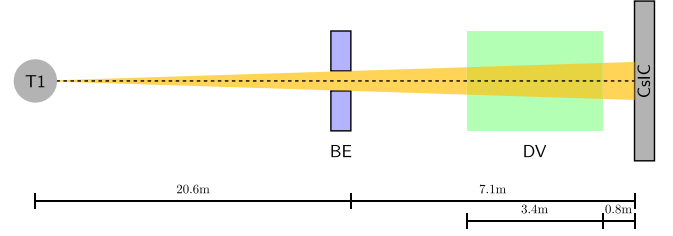


FIG. 1. Schematic profile of the KOTO production target (T1), beamline (yellow), beam exit (BE), and detector system, including the fiducial  $K_L$  decay volume (DV) and the CsI electromagnetic calorimeter (CsIC). Not to scale.

different crystal sizes: small  $25 \times 25 \text{ mm}^2$  crystals filling a square central region  $1.2 \times 1.2 \text{ m}^2$  in size, and large  $50 \times 50 \text{ mm}^2$  crystals in the remaining outer region.

### B. KOTO detector simulation

We make use of our own custom weighted Monte Carlo (MC) generator to emulate the core components of the KOTO detector simulation, including the event selection and  $K_L$  mass reconstruction [33,34]. We define the beamline as the  $z$ -axis with the CsIC at  $z = 0$ , such that  $z_{\text{T1}} = -27.7 \text{ m}$ ,  $z_{\text{BE}} = -7.1 \text{ m}$ , and the DV lies on the interval  $I_{\text{DV}} = [-4.2, -0.8] \text{ m}$ .

The  $K_L$  spatial profile over the aperture is well approximated as spherically symmetrically generated by a point source at T1. Just as in the KOTO simulation, we generate the kaon differential flux by drawing from a parametric Gaussian fit to the measured  $K_L$  momentum distribution at the BE [34]. As the energy dependence over the  $K_L \rightarrow 3\pi^0$  Dalitz space varies at the (sub)percent level [5,35], for the purposes of simulation we approximate the differential decay rate as pure phase space, noting also the  $\pi^0 \rightarrow 2\gamma$  decays proceed through an  $S$ -wave. (Note that the 6-body phase space measure itself is nontrivial, and thus the differential rate varies over phase space.) The MC weights are the normalized differential rates of the  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  process multiplied by the  $K_L$  decay and momentum probability distributions. These are reweighted from flat prior distributions in the natural helicity basis coordinates of the  $1 \rightarrow 3 \rightarrow 6$ -body phase space, and flat prior  $K_L$  spatial and momentum distributions.

The full KOTO detector simulation incorporates showering and clustering in the CsIC, with isolation and timing requirements. These lead to finite energy and spatial resolution. Based, however, on the intuition that the finite granularity of the CsIC crystals is the leading source of smearing in reconstructed observables, and that the showering and clustering does not bias the spatial reconstruction, we instead approximate photon hits on the CsIC as single crystal hits, whose location is reconstructed to be at the center of the hit crystal. We further treat the truth-level photon energy as the reconstructed energy of each hit, as the raw energy resolution of the CsIC is very high. We will

<sup>2</sup>References [33,34] use multiple overlapping coordinate systems. The specified dimensions are based on our best understanding of the specifications therein.

see below that these approximations lead to an excellent description of the  $m_{K_L}^{\text{rec}}$  distribution, and supply a faithful representation of the KOTO reconstruction of the  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  signal. Per the KOTO selection, we require six reconstructed photon hits in the CsIC acceptance, each with energy  $E \geq 50$  MeV, and isolated by at least 150 mm.

The KOTO reconstruction of the  $K_L$  makes use of a perturbative procedure, whose first step assumes the  $K_L$  decay vertex is constrained to the  $z$ -axis, at  $z = z_v < 0$ . The subsequent  $\pi^0$  decays are treated as prompt. For a pairing of hits of energy  $E_i$  and  $E_j$  into a putative  $\pi^0$  decay, the inferred opening angle of the photon momenta is determined by the pion mass relation

$$m_{\pi^0}^2 = 2E_i E_j (1 - \cos \theta_{ij}). \quad (1)$$

Combining  $\cos \theta_{ij}$  with the pair's reconstructed hits on the CsIC at  $\vec{r}_i$  and  $\vec{r}_j$ , respectively, uniquely defines the locus of a self-intersecting torus, whose central axis lies along  $\vec{d}_{ij} = \vec{r}_i - \vec{r}_j$ . The intersection of this torus with the negative  $z$ -axis determines the reconstructed vertex  $z_v^{ij}$  of the pair, for which there may be 0, 1, or (very rarely) 2 solutions: for zero solutions the pairing is rejected as not reconstructing a  $\pi^0$  vertex, while for two the solution with the largest  $|z_v^{ij}|$  is chosen (as in this rare case it is more probable for the  $K_L$  to have decayed earlier).

There are  $6!/(2^3 \times 3!) = 15$  possible sets of pairwise combinations of the six photon hits, labeled by  $\alpha = 1, 2, \dots, 15$ . For each “allowed” set, i.e., one whose three pairs each reconstruct a vertex, KOTO computes a weighted average vertex

$$\bar{z}_v^\alpha = \frac{\sum_{k=1,2,3} z_v^{\alpha k} / (\sigma_v^{\alpha k})^2}{\sum_{k=1,2,3} 1 / (\sigma_v^{\alpha k})^2}, \quad (2)$$

in which  $\sigma_v^{\alpha k}$  is the uncertainty of  $z_v^{\alpha k}$  within the reconstruction framework. The allowed set whose weighted average has the smallest  $\chi_\alpha^2 = \sum_k (z_v^{\alpha k} - \bar{z}_v^\alpha)^2 / (\sigma_v^{\alpha k})^2$  is selected as the reconstructed photon pair combination for the event, and its weighted average vertex is assigned to be the reconstructed  $K_L$  vertex,  $z_v$ . [No upper bound on  $\min(\chi_\alpha^2)$  appears to be imposed in the KOTO selection.] Only events with  $z_v \in I_{\text{DV}}$  pass selection.

The second step of the KOTO reconstruction procedure perturbs the constraint that the  $K_L$  vertex lies on the  $z$ -axis. (Per the geometry of the square aperture at the BE, the truth-level transverse displacement of the  $K_L$  vertex is small, at most  $\sim 8$  cm.) The “center of energy” of the reconstructed photon hits on the CsIC,  $\vec{r}_E = \sum_{i=1}^6 \vec{r}_i E_i / \sum_i E_i$ , is used by KOTO to approximate the direction of the parent  $K_L$  momentum and thus determines the transverse displacement  $\vec{r}_v$  of the parent  $K_L$  vertex at  $z = z_v$ , under the assumption that the kaon was produced from a T1 point source at  $z_{T1}$

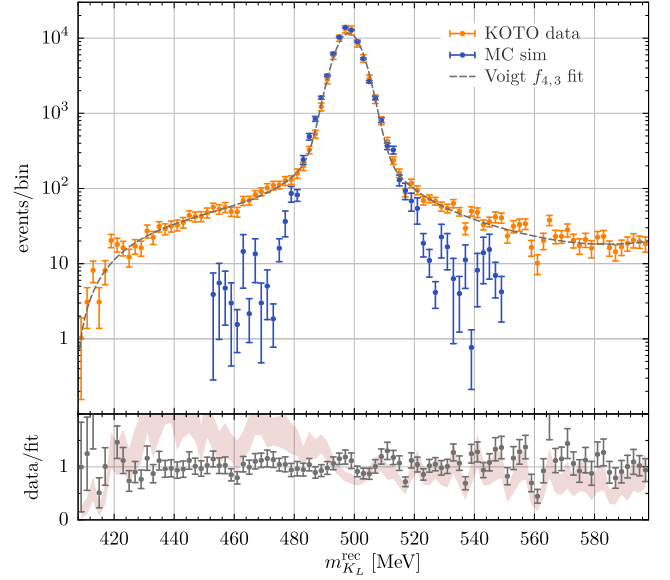


FIG. 2. Top: KOTO data (orange) for the reconstructed  $K_L$  mass,  $m_{K_L}^{\text{rec}}$ , from the  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  dataset for  $2 \times 10^{14}$  POT, versus our MC simulation (blue). Uncertainties in the latter are due to MC statistics only. Also shown is the line shape SM background fit using the  $f_{4,3}$  Voigt-like model, see Sec. III A. Bottom: ratio of data versus the  $f_{4,3}$  fit (dark gray points), and the 68% CL band for the ratio of data versus an approximate simulation of the  $K_L \rightarrow 3\pi^0$  peak plus continuum contributions (light red band), see Sec. IV B.

using  $z_v$ , as determined in the first step of this procedure.<sup>3</sup> Using  $\{\vec{r}_v, z_v\}$  as the  $K_L$  vertex location, one may then reconstruct all photon momenta, and thus the  $K_L$  invariant mass (as well as the pairwise pion invariant masses).

In our MC we apply the same selection and reconstruction procedure, but with the following simplifications: based on the intuition that the  $\sigma_v^{ij}$  uncertainty is mainly driven by detector CsIC crystal granularity, which in turn scales with the areas of CsIC crystals, we assume that all pairs with both hits on the central (small crystal) region have the same vertex uncertainty  $\sigma_{v0}$ . For pairs with one or two hits in the outer (large crystal) region, we then scale the uncertainty by a factor of 4 and 16, respectively, per the relative areas of the small and large crystals. As  $\bar{z}_v^\alpha$  in Eq. (2) and the relative ratios of  $\chi_\alpha^2$  are unaffected by the overall scaling of the uncertainties, under this approximation we may compute each weighted average vertex and identify  $z_v$  from  $\min(\chi_\alpha^2)$  using the relative uncertainties alone, without needing to estimate  $\sigma_{v0}$  itself.

In Fig. 2 we show the resulting  $m_{K_L}^{\text{rec}}$  distribution (blue) from a MC sample with  $10^7$  initial  $K_L$ 's at the BE, compared to KOTO data (orange) from the  $2 \times 10^{14}$  POT dataset (Fig. 10 of Ref. [33]). The simulated

<sup>3</sup>The correct quantity for this is the center of longitudinal momentum, but this is not directly reconstructible at KOTO.

distribution is normalized to the total number of events in the KOTO  $m_{K_L}^{\text{rec}}$  data—approximately  $70 \times 10^3$  postselection—by applying an overall scaling, noting that the MC statistical errors over the simulated peak are negligible compared to those in the measured  $m_{K_L}^{\text{rec}}$  distribution. The excellent agreement over the peak verifies that our MC simulation provides a faithful representation of the KOTO reconstruction of the  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  signal. The long tails in the KOTO data predominantly arise from other background activity in the detector, such as halo neutron scattering, coinciding with  $K_L \rightarrow 3\pi^0$  decays in which one or more photon pairs miss the calorimeter acceptance. This “accidental overlay” results in six-photon-clusters events that generate a continuum background in the  $m_{K_L}^{\text{rec}}$  distribution. We need not simulate these tails to characterize the  $3\pi^0$  signal peak, nor that of  $2\pi^0 a$  considered below. Simulation of such backgrounds will be relevant in Sec. IV, which considers a modified KOTO selection to search for displaced ALP decays with masses outside the pion chimney.

The acceptance in our simulation is approximately  $19 \times 10^{-4}$ , which is somewhat larger than the quoted  $4.3 \times 10^{-4}$  acceptance in the full KOTO simulation [33,34]. The latter incorporates additional selections as well as the abovementioned showering, clustering, and timing effects, which may account for the difference. The results in Fig. 2 demonstrate that these effects do not appear to significantly bias the shape of the  $m_{K_L}^{\text{rec}}$  peak distribution from  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$ .

### C. $K_L \rightarrow 2\pi^0 a$ signal

We simulate the  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$  signal by varying the mass of one pseudoscalar in our treatment of  $K_L \rightarrow 3\pi^0$ . We treat the chimney range as a  $\pm 15$  MeV window around  $m_{\pi^0}$ ,

$$120 \text{ MeV} \leq m_a \leq 150 \text{ MeV}. \quad (3)$$

We use the basis-independent chiral perturbation theory (ChPT) description of the ALP production and decay amplitudes from Ref. [6] throughout. In the chimney regime the  $K_L \rightarrow 2\pi^0 a$  amplitude is dominated by ALP-pion mixing, so that corrections to the energy dependence of the  $2\pi^0 a$  versus  $3\pi^0$  Dalitz spaces are subleading, being estimated as  $\lesssim 10\%$  [6], and may be safely neglected. Similarly, we treat the ALP as prompt over the entire chimney and assume that the  $a \rightarrow \gamma\gamma$  branching ratio dominates those of other decay modes, as generically expected. Some of our projected limits below extend to regions in which the ALP lifetime  $c\tau_a \sim \text{mm} - \text{cm}$ . This remains smaller than our estimate of the KOTO vertex resolution, which we find to be  $\mathcal{O}(10 \text{ cm})$  by propagating the KOTO photon energy and hit resolutions (Eqs. (6.1) and (6.2) of Ref. [34]) through our simulation of the KOTO

reconstruction. Thus, the ALPs may still be treated as prompt.

The  $K_L \rightarrow 2\pi^0 a$  signal is generated with the same overall scaling as derived for the  $K_L \rightarrow 3\pi^0$  sample, so that the ALP signal distributions are normalized to the  $K_L \rightarrow 3\pi^0$  branching ratio. We generate simulated samples with  $3 \times 10^6$  initial  $K_L$ 's at the BE—sufficient for sub-leading statistical uncertainties versus the KOTO data for  $2 \times 10^{14}$  POT—for ALP masses in 2 MeV intervals over the chimney range, as well as for  $m_a = m_{\pi^0}$ . For each  $m_a$ , the signal event sample is passed through our simulation of the KOTO selection and  $K_L$  mass reconstruction to generate a  $m_{K_L}^{\text{rec}}$  distribution. Critically, this procedure applies the pion mass constraint (1) to all pairs of photons, even those originating from the ALP.

Some intuition is helpful for anticipating the morphology of the resulting ALP signal line shapes in the  $m_{K_L}^{\text{rec}}$  distribution. In general, for a lighter (heavier) ALP, one expects photon pairs to be produced in a typically narrower (broader) cone versus the pion. Under a reconstruction that assumes the ALP is a pion, one should expect the variation in the cone to appear to be the result of a typically more (less) energetic pion, and thus a typically heavier (lighter) reconstructed kaon. This results in a peak that is shifted upwards (downwards) with respect to  $m_{K_L}$ .

We may approximately quantify how this expectation manifests through the application of Eq. (1) by taking the small opening angle limit and constraining the decay vertices to the  $z$ -axis for simplicity. It follows from Eq. (1) that for the  $a \rightarrow \gamma\gamma$  vertex, the variation induced in the reconstructed decay vertex,  $z_v^a$ , and in the reconstructed pairwise opening angle,  $\theta$ , by the shift  $\delta m = m_a - m_{\pi^0}$ , obeys  $\delta z_v^a / z_v^a \simeq -\delta\theta / \theta \simeq \delta m / m_{\pi^0}$ . Further assuming the simplified case that the three reconstructed vertices—i.e.,  $z_v^a$  and  $z_v^{1,2}$  for the ALP and two  $\pi^0$ 's, respectively—have equal uncertainties and that the combinatoric set corresponding to  $\min(\chi^2)$  remains the same, then misreconstruction of the ALP vertex distorts the weighted average such that  $\delta \bar{z}_v = \delta z_v^a / 3$ . As  $z_v^a \simeq z_v^{1,2}$  up to corrections typically parametrically smaller than  $\delta z_v^a$ , then also  $\bar{z}_v \simeq z_v^a$  and so

$$\delta \bar{z}_v / \bar{z}_v \simeq \delta z_v^a / 3 z_v^a \simeq \delta m / 3 m_{\pi^0}. \quad (4)$$

Because the  $K_L$  momentum is aligned with the beamline under the assumption that the decay is constrained to the  $z$ -axis, then the sum of photons' transverse momentum is zero, while their total longitudinal momentum  $\sum_i k_{L,i} \simeq \sum_i E_i (1 - r_i^2 / (2\bar{z}_v^2))$ , in which  $r_i$  is the hit location of the  $i$ th photon on the CsIC and we take the limit  $r_i \ll \bar{z}_v$ . Hence, the reconstructed mass  $m_{K_L}^{\text{rec}2} \simeq \sum_{i,j} E_i E_j r_j^2 / \bar{z}_v^2$ , whence  $\delta m_{K_L}^{\text{rec}} / m_{K_L}^{\text{rec}} \simeq -\delta \bar{z}_v / \bar{z}_v$ . Combining this with Eq. (4), one expects that the application of Eq. (1) should result in shifts in the peak in the  $m_{K_L}^{\text{rec}}$  distribution, such that

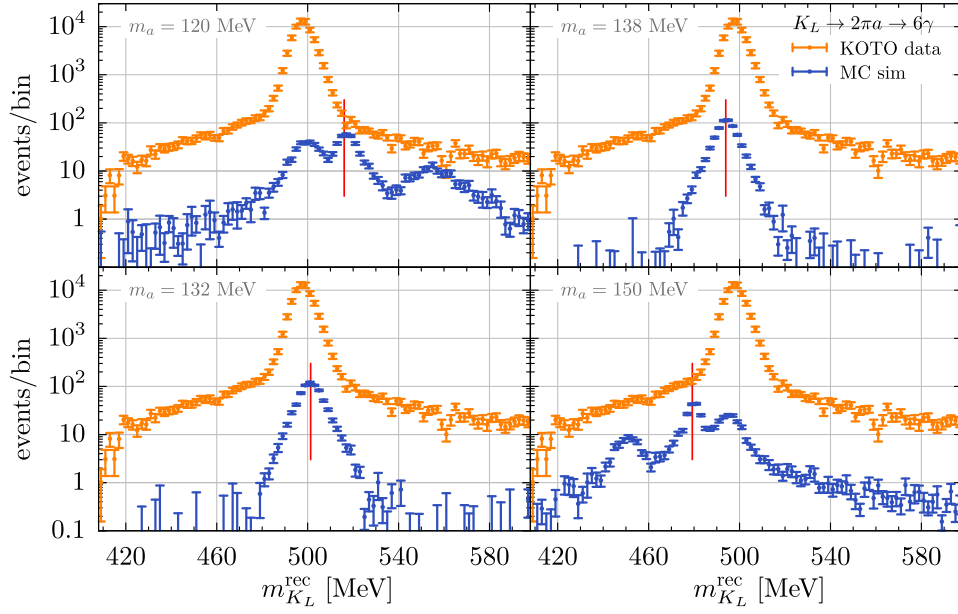


FIG. 3. Simulated distributions for the reconstructed  $K_L$  mass,  $m_{K_L}^{\text{rec}}$ , from  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$  under our emulation of the KOTO selection and reconstruction (in blue), compared to the KOTO  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  data (in orange) for  $2 \times 10^{14}$  POT. For visual clarity, the ALP signal distributions are normalized such that  $\text{Br}[K_L \rightarrow 2\pi^0 a]/\text{Br}[K_L \rightarrow 3\pi^0] = 0.01$ . The red vertical line indicates the expected location of the shifted  $m_{K_L}^{\text{rec}}$  peak from Eq. (5).

$$\delta m_{K_L}^{\text{rec}} \simeq -[m_{K_L}/3m_{\pi^0}]\delta m \simeq -1.2\delta m. \quad (5)$$

Beyond the small perturbation regime, one may expect combinatoric misreconstructions to play a role, leading to another peak near  $m_{K_L}$ , or elsewhere.

In Fig. 3 we show the  $m_{K_L}^{\text{rec}}$  distributions arising from a  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$  signal (blue), versus the KOTO data (orange), for the four cases  $m_a = 120, 132, 138$ , and  $150$  MeV. In each case, there is a shifted peak matching the predictions from Eq. (5) (red vertical lines). For ALP masses near the edges of the considered mass window, peaks at  $m_{K_L}$  and in the far tails also arise, leading to a multipeak morphology for the ALP signal line shape.

### III. FITS AND LIMITS

#### A. Line shape model

The measured  $m_{K_L}^{\text{rec}}$  spectrum includes the  $K_L \rightarrow 3\pi^0$  peak as well as a continuum contribution that results in the extended high and low tails. To set limits on the  $K_L \rightarrow 2\pi^0 a$  contribution from data requires comparison to a prior for these SM backgrounds from either theory and/or Monte Carlo, in order to obtain a background-subtracted dataset. No first principles calculation of the  $m_{K_L}^{\text{rec}}$  spectrum is likely possible. Further, one would require a precision MC simulation that is self-consistently implemented to avoid tunings to other measurements that may in principle be sensitive to the same ALP phenomenology being probed.

To avoid this, we instead adopt a phenomenological approach to obtain a line shape description of the collective SM background. Based on the assumption that the true backgrounds should be smooth, we employ a smooth functional form in terms of  $m_{K_L}^{\text{rec}} \equiv \mu$  (for ease of reading) that parametrizes the expected morphology: we attenuate resulting model dependence via information theoretic techniques, described below. In particular, we choose a form that incorporates the expected  $3\pi^0$  kinematic threshold singularity plus a Voigt profile description of the peak, modified by rational polynomial prefactors that vanish at  $\mu = 3m_{\pi^0}$ , and have simple large  $\mu$  behavior.<sup>4</sup> Explicitly, we consider a line shape

$$f_{2n_a, n_b}(\mu) = \sqrt{\mu^2 - 9m_{\pi^0}^2} \left\{ \sum_{k=1}^{n_a} a_{2k} \frac{[\mu^2 - 9m_{\pi^0}^2]^{2k-2}}{\mu^{2k}} + \left[ \sum_{k=0}^{n_b} b_k \frac{[\mu^2 - 9m_{\pi^0}^2]^k}{\mu^k} \right] V_{\delta, \sigma}(\mu - m_{K_L} + \delta m) \right\}, \quad (6)$$

in which  $V_{\delta, \sigma}(\mu)$  is the Voigt profile distribution, proportional to the real part of  $e^{-\zeta^2} \text{erfc}[i\zeta]$ ,  $\zeta = (\mu - i\delta)/(\sqrt{2}\sigma)$ .

<sup>4</sup>The Voigt profile [36], a convolution of a Lorentzian and Gaussian, is very widely used as a descriptor of an experimentally smeared resonance peak. It is often combined in experimental particle physics with *ad hoc* descriptions of continuum background distributions. Here, we use the nonrelativistic version, though relativistic versions can also be contemplated, see, e.g., Ref. [37].

For the purpose of numerical stability in the fits, we take starting values of  $\delta = 1$  MeV and  $\sigma = 4$  MeV in fits to Eq. (6). The  $a_{2k}$  and  $b_k$  terms have large- $\mu$  dependence scaling as  $\mu^{2k-3}$  and  $\mu^{k+1}$ , respectively. The free parameters of this model are the three Voigt parameters  $\delta$ ,  $\sigma$ , and  $\delta m$ , which control the height, spread, and maximum of the peak, respectively, as well as the rational polynomial prefactor parameters  $a_{2k}$ ,  $k = 1, \dots, n_a$ , and  $b_k$ ,  $k = 0, \dots, n_b$ .

In all fits we consider, we assume that bin-bin correlations in the  $m_{K_L}^{\text{rec}}$  spectrum are negligible: assuming some degree of (anti)correlation will typically weaken (strengthen) the power of the limits we derive. Thus, one should interpret the limits as prospective, not only because of our phenomenological description of the assumed smooth line shape, but also pending a fully fledged analysis within an experimental framework that incorporates effects of correlations. In order to avoid over- or under-fitting we use information theoretic techniques to select the appropriate number of the polynomial prefactors. Considering all possible models with  $n_a \leq 3$  and  $n_b \leq 3$ , in order to keep the complexity somewhat low, the Akaike information criterion (AIC) (see, e.g., Ref. [38]) selects  $f_{4,3}$ , with nine parameters:  $a_{2,4}$ ,  $b_{0,1,2,3}$ ,  $\delta$ ,  $\sigma$ , and  $\delta m$ . The resulting “SM background fit” is shown in Fig. 2 along with the ratio of the data versus the fit in the bottom panel. The fit  $\chi^2/\text{dof} = 124.3/(95 - 9) = 1.45$ .

## B. Approach

Caution and care must be taken in treating the residuals of the SM background line shape fit as background-subtracted data to which a  $K_L \rightarrow 2\pi^0 a$  signal distribution can be compared, because of several complications. First, just as for the SM backgrounds, there is no closed-form theory description for the multipeak morphology of the ALP signal itself, and it cannot be simply modeled as a narrow peak on top of the fit residuals. Second, the chosen model of the line shape may not be fully orthogonal to the morphology of the ALP signal, such that fits to the model may in principle absorb some of the signal line shape features, leading to weaker limits. That is, the limits may be overestimated if this effect is not included.

These complications motivate a combined signal plus background fit approach, as follows: we introduce an additional parameter

$$\rho = \text{Br}[K_L \rightarrow 2\pi^0 a] / \text{Br}[K_L \rightarrow 3\pi^0] \geq 0. \quad (7)$$

For a given ALP mass and  $\rho$  hypothesis, we subtract the simulated ALP signal from the measured data and then reperform the fit to the  $f_{4,3}$  line shape template, yielding a “subtracted” fit statistic  $\chi_{\text{sub}}^2(m_a, \rho)$ . Constructing a likelihood ratio with respect to the SM background fit, which has  $\chi_0^2 = \chi_{\text{sub}}^2(m_a, 0)$ , yields

$$\delta\chi^2(m_a, \rho) = \chi_{\text{sub}}^2(m_a, \rho) - \chi_0^2. \quad (8)$$

which is strictly positive provided that the data currently set only limits on  $\rho$  (verified below), equivalent to the hypothesis that the true value is at  $\rho = 0$ . Because  $\rho \geq 0$  only, this  $\delta\chi^2$  is then distributed as a 1:1 admixture of a 0 and 1 degree of freedom  $\chi^2$  [39], i.e., with probability density function  $p(x) = 1/2[\delta(x) + p_{\chi_1^2}(x)]$ . Thus, a 90% CL on  $\rho$  corresponds to  $\delta\chi^2(m_a, \rho) < 1.64$ .

Although we have argued the residuals of the SM background fit in Sec. III A should not be used directly to obtain a limit, we use them to obtain a “naive” limit for  $\rho$ . Over a range of  $\rho$  values covering this naive limit, we then refit the  $f_{4,3}$  line shape to determine  $\delta\chi^2(m_a, \rho)$ , and interpolate to obtain the 90% CL. (In principle, one should reperform the AIC model selection for each refit. However, the shift in the criterion value from the naive to full treatment is smaller than the shift in the AIC to other models, such that  $f_{4,3}$  remains the optimal model in the considered model space.)

## C. Results

As above, we consider ALPs with masses in the pion chimney mass range, in steps of 2 MeV. Noting the 2 MeV binning of the KOTO  $m_{K_L}^{\text{rec}}$  data, and with reference to Eq. (5), we do not set limits within a  $\pm 1$  MeV window around  $m_a = m_{\pi^0}$ . Within this window  $\delta m_{K_L}^{\text{rec}}$  is below the binning resolution, so that one may then expect that the line shape fit becomes mainly sensitive to higher-order details of the shape of the simulated  $m_{K_L}^{\text{rec}}$  peak, which we expect is less well controlled in our approximation of the detector simulation versus the location of the peak itself. When interpreting our results in the context of limits on Wilson coefficients of particular models (see Sec. III D), we similarly exclude a  $\sim 2$  MeV window around  $m_{\pi^0}$  because of the breakdown of the ChPT description.

In Fig. 4 we show the resulting limits on  $\text{Br}[K_L \rightarrow 2\pi^0 a] (= \rho \times \text{Br}[K_L \rightarrow 3\pi^0])$  at 90% CL, using the known value  $\text{Br}[K_L \rightarrow 3\pi^0] = 0.1952(12)$  [35]. For  $m_a$  close to  $m_{\pi^0}$ , such that the ALP signal distribution is a single shifted peak, the naive limit (in gray) for  $\rho$  coincides reasonably closely with the interpolated value from the refit treatment (in orange). By contrast, near the boundaries of the chimney, the interpolated limit can sometimes be  $\mathcal{O}(1)$  weaker. These results suggest that any model dependence from our choice of line shape model for the backgrounds is at most an  $\mathcal{O}(1)$  effect, and that the limits are reliable at that level. From the figure, limits on branching ratios as small as  $\text{few} \times 10^{-4}$  may be extracted already, using a luminosity as small as  $2 \times 10^{14}$  POT.

To gain a sense of the size of the ALP branching ratio limit versus the fit residuals, we show in Fig. 5 the residuals of the  $f_{4,3}$  line shape fit to the KOTO  $m_{K_L}^{\text{rec}}$  data, after subtracting the ALP signal component evaluated at the 90% CL branching ratio limit or four different ALP mass values. As expected, the subtracted ALP component is

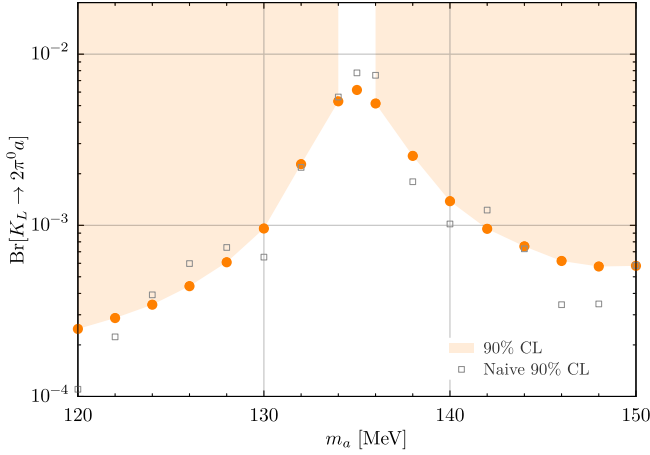


FIG. 4. Interpolated limits (orange) on  $\text{Br}[K_L \rightarrow 2\pi^0 a]$  at 90% CL from the KOTO measurement of  $m_{K_L}^{\text{rec}}$  in  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$ , with  $2 \times 10^{14}$  POT. Limits evaluated at specific  $m_a$  values are shown by orange data points. The limit at  $m_a = m_{\pi^0}$  is shown for reference, but no limit is inferred within  $\pm 1$  MeV of  $m_{\pi^0}$ , see the text. Also shown are naive limits at 90% CL (gray squares) derived from the residuals of the SM background fit.

comparable in size to and closely bounded by the fit residuals.

### D. Projected limits

The branching ratio limits may be reinterpreted in the context of a particular set of ALP-SM Wilson coefficients, from which the  $K_L \rightarrow 2\pi^0 a$  branching ratio may be computed. Per the analysis and conventions of Ref. [6], we consider three scenarios, in which the nonzero Wilson

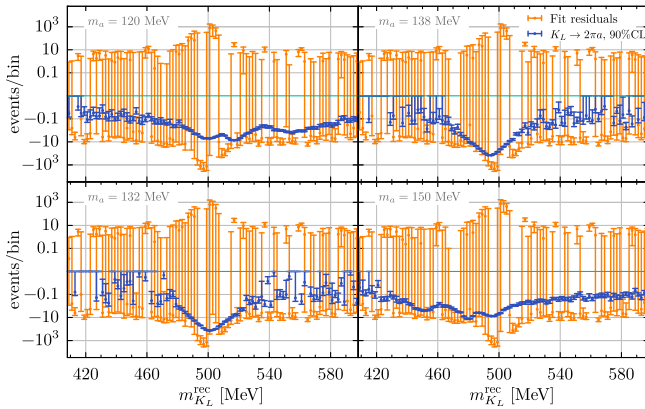


FIG. 5. The fit residuals (orange) for a fit of the  $f_{4,3}$  line shape to the KOTO  $m_{K_L}^{\text{rec}}$  data with subtracted ALP signal component (blue), evaluated at the derived limit for the branching ratio at 90% CL from Fig. 4, for the same four ALP masses as in Fig. 3. Fit residuals may be of either sign, or may have uncertainty ranges that fluctuate through zero (in either direction). We, therefore, show signed logarithmic plots, with the logarithmic ranges soldered at  $\pm 10^{-3}$  events/bin, indicated by the green line.

coefficients are (i)  $c = c_{gg} = -(3/2)c_{d_r}$ , which exhibits a large enhancement of the  $K_L \rightarrow 2\pi a$  rates versus  $K_L \rightarrow \pi a$ , (ii)  $c_{gg}$ , and (iii)  $c_{d_r}$ . The latter two scenarios are commonly examined in ALP literature. The resulting limits are shown by gray shaded regions in Fig. 6, compared to various other known limits from NA62 [40,41], NA48 [42], E391A [43], BarBar [44], E949 [45], KOTO [46], and CMS [47]. For a discussion, see Ref. [6]. While we emphasize again that one should interpret these limits as prospective, we see from our analysis that the KOTO calibration data ( $2 \times 10^{14}$  POT) can prospectively set novel limits in the pion chimney for all three scenarios. If this technique is not systematically limited for the larger datasets comparable to the sizes of the 2016–2021 runs (roughly  $6 \times 10^{19}$  POT) or the projected full run ( $10^{21}$  POT), and assuming the reach scales as  $1/\sqrt{\text{POT}}$ , then one sees from the solid black lines in Fig. 6 that one may probe most of the pion chimney.

The validity of the ChPT description for the ALP interactions around  $m_a \sim m_{\pi^0}$  in Fig. 6 holds in the small-angle approximation

$$\theta_{\pi a} \sim f_{\pi}/f \times c_X/|\delta_{m_a}| \ll 1, \quad (9)$$

where  $c_X$  is the Wilson coefficient,  $\theta_{\pi a}$  is the mixing angle between  $\pi^0$  and the ALP, and the tuning parameter  $\delta_{m_a} \equiv m_a^2/m_{\pi^0}^2 - 1$ . For example, even for the moderately low UV scale  $c_X/f = 0.1 \text{ GeV}^{-1}$ —corresponding to the typical best reach of the limits from the  $2 \times 10^{14}$  POT KOTO calibration dataset—a few percent-level mass tuning  $\delta_{m_a} \sim 0.03$  falls within the small-angle approximation, with  $\theta_{\pi a} \sim 0.3$ . This corresponds to excluding  $|m_a - m_{\pi}| \lesssim 2 \text{ MeV}$  from the regime of validity, which is similar to the already excluded interval shown in our Figs. 4 and 6, arising from finite binning resolution. There is, in addition, a constraint on  $\theta_{a\pi}$  from the  $\pi^0$  lifetime (see also, e.g., [51]), which is modified by a  $\cos^2 \theta_{a\pi}$  factor in the small-angle regime. Comparing the next-to-leading-order ChPT prediction [52] to the measured value [35] implies similarly  $\theta_{a\pi} < 0.27$  at  $2\sigma$  CL (the measured value itself exceeds the prediction at just over  $2\sigma$  CL, favoring some degree of mixing.)

More generally, the prospective limits on  $c_X/f$  from our analysis are substantially model dependent. We show in Fig. 7 example upper bounds for which the ChPT description is expected to remain valid for  $\theta_{\pi a} < 0.04$ ,  $\theta_{\pi a} < 0.1$ , and  $\theta_{\pi a} < 0.27$ . In some models, the prospective limits for the  $2 \times 10^{14}$  POT KOTO calibration dataset may be close to or exceed the upper bounds outside the  $134 \text{ MeV} \leq m_a \leq 136 \text{ MeV}$  window, depending on  $\theta_{\pi a}$ . However, the projected stronger limits on  $c_X/f$  (black lines in Fig. 6) that may be obtained from more recent or future, larger KOTO datasets, will fall well within the regime of validity of ChPT up to very narrow mass windows around  $m_{\pi^0}$ .

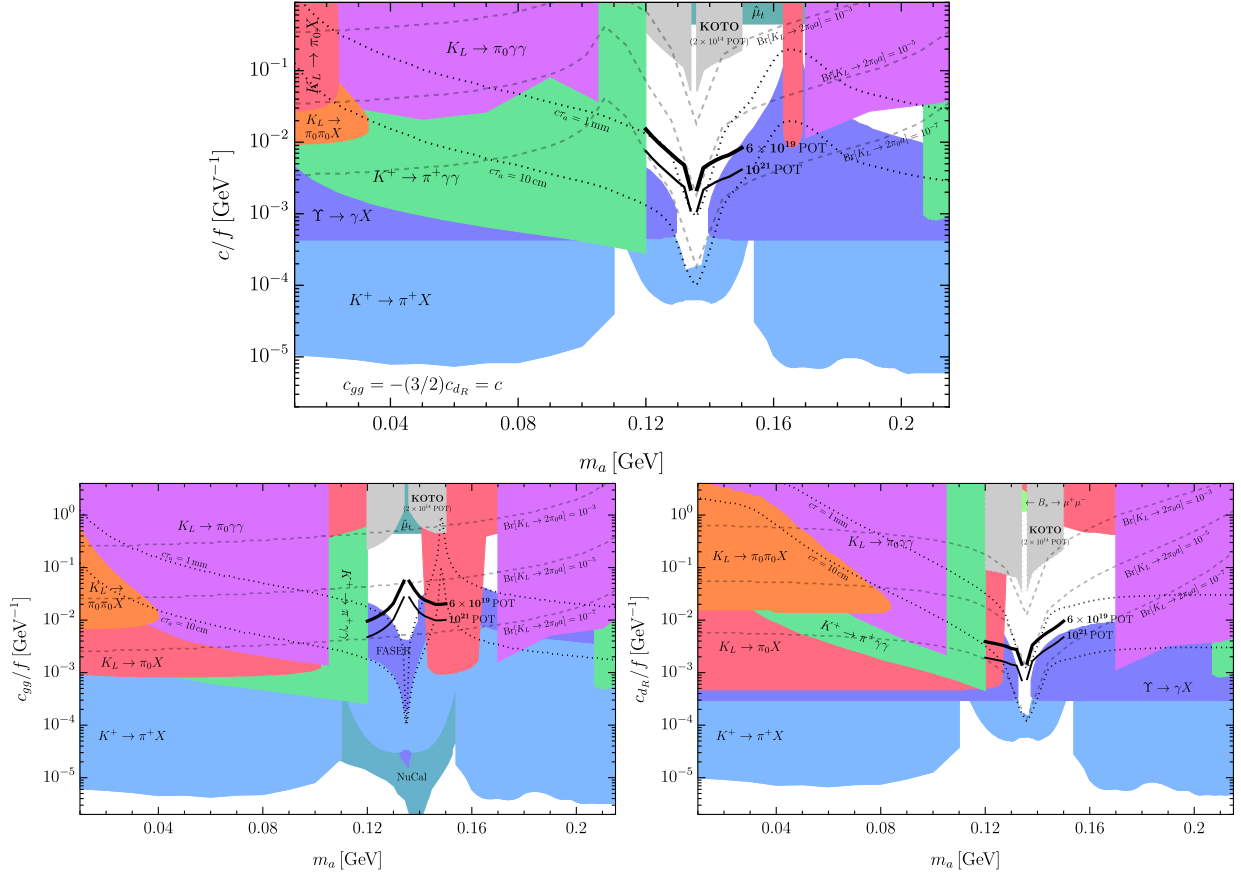


FIG. 6. Prospective 90% CL bounds on ALP couplings in the pion chimney from this analysis (gray shaded) for  $c_{gg} = -3/2c_{dR} = c$  (top),  $c_{gg}$  (bottom left), and  $c_{dR}$  (bottom right), per Ref. [6]. Previous bounds include limits from  $K^+ \rightarrow \pi^+ X$  (blue) [40],  $K^+ \rightarrow \pi^+ \gamma \gamma$  (green) [41,45],  $K_L \rightarrow \pi^0 X$  (red) [46],  $K_L \rightarrow \pi^0 \gamma \gamma$  (purple) [42],  $K_L \rightarrow \pi^0 \pi^0 X$  (orange) [43],  $B_s \rightarrow \mu^+ \mu^-$  (light green) [48], FASER (dark purple) [49], NuCal (teal) [50], and the top chromomagnetic dipole moment  $\hat{\mu}_t$  (dark green) [47] using the expression in Ref. [11]. Details on the calculation of the bounds can be found in [6]. The heavy and light solid black lines represent our naive projections for  $6 \times 10^{19}$  and  $10^{21}$  POT, respectively. A fully fledged analysis within an experimental framework is required to validate the precise CLs.

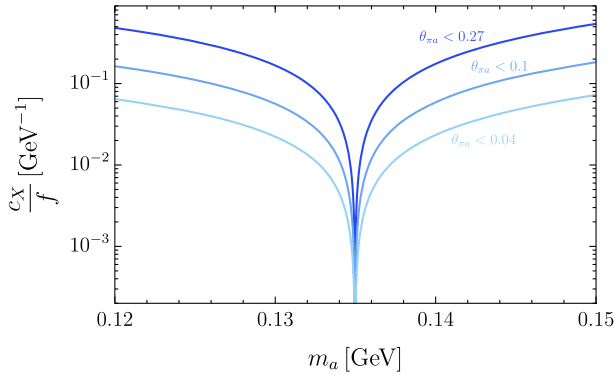


FIG. 7. Upper bounds for the validity of the ChPT description of the ALP interactions for  $\theta_{\pi a} < 0.1$  and  $\theta_{\pi a} < 0.3$ .

#### IV. BEYOND THE CHIMNEY

So far we have considered limits for  $K_L \rightarrow 2\pi^0 a$  signals that are able to misreconstruct as  $K_L \rightarrow 3\pi^0$  under the KOTO selection described in Sec. II B, which imposes

Eq. (1) on all photon pairs. One may consider, however, an adaptation of the KOTO reconstruction for ALPs in which Eq. (1) is relaxed for one photon pair, and one instead sets limits on the  $K_L \rightarrow 2\pi^0 a$  branching ratio via “bump hunting” in the spectrum of the measured invariant diphoton mass,  $m_a^{\text{rec}}$ . This technique may set more powerful ALP limits for  $m_a$  far outside the chimney compared to line shape fits to the  $m_{K_L}^{\text{rec}}$  spectrum (though any such limits will currently not be competitive with constraints from other  $K \rightarrow \pi a$  searches).

#### A. Modified reconstruction

For  $m_a$  well below  $m_{\pi^0}$ , it is no longer natural to treat the ALP as prompt, and thus one must also generalize the KOTO reconstruction technique to accommodate displaced-in-flight ALP decays. One may also consider the case of  $m_a > m_{\pi^0}$  with small enough coupling that the ALP is long-lived. In all that follows we will consider only  $m_a \leq m_{\pi^0}$ , keeping in mind that the proposed approach may also be applied to larger ALP masses.

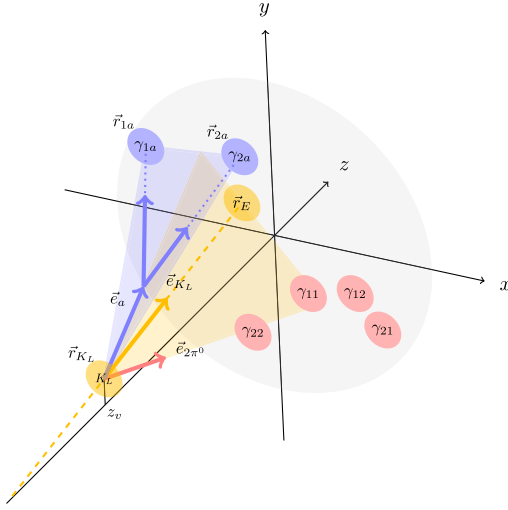


FIG. 8. Geometry of reconstructed kinematics for  $K_L \rightarrow 2\pi^0 a$ , applying Eq. (1) to only two diphoton pairs (red), followed by the selection of Sec. II B to determine the  $K_L$  decay vertex (orange). Under the modified selection to determine  $m_a^{\text{rec}}$ , we require the reconstructed ALP momentum  $p_a$  to be coplanar with that of  $K_L$  and  $2\pi^0$  (orange plane) and also lie in the plane defined by the  $a \rightarrow \gamma\gamma$  photon hits (blue) and the  $K_L$  decay vertex (blue plane).

We adapt the KOTO reconstruction to examine the  $6!/(2^3 \times 2!) = 45$  possible sets of pairwise combinations of the six photon hits into two pairs that are constrained by Eq. (1), plus one unconstrained pair. These are labeled by  $\alpha = 1, 2, \dots, 45$ . For each “allowed” set, i.e., one whose two constrained pairs both reconstruct a vertex, one computes a weighted average vertex, as in Eq. (2). One again determines the reconstructed photon pair combination by finding the minimal  $\chi_a^2$ , using just the two vertices from the two constrained pairs in each combination. The weighted average vertex of the combination with  $\min(\chi_a^2)$  is assigned to be the location of the reconstructed  $K_L$  vertex,  $z_v$ , on the  $z$ -axis, and we determine the perturbative transverse displacement  $\vec{r}_v$  from the center-of-energy  $\vec{r}_E$  on the CsIC.

With reference to Fig. 8, the direction of the  $K_L$  momentum  $\vec{e}_{K_L}$  must lie on  $\vec{r}_E - \vec{r}_{T1}$ , the location of the T1 target, which is treated as pointlike (see Sec. II B for the definition of  $\vec{r}_E$ ), while the summed spatial momentum of the two pions  $\vec{p}_{2\pi^0}$  is also known, and thus the corresponding unit direction  $\vec{e}_{2\pi^0}$ . The ALP spatial momentum should be coplanar with  $\vec{e}_{K_L}$  and  $\vec{e}_{2\pi^0}$ , and it should also be coplanar with the plane formed by the reconstructed  $K_L$  vertex at  $\vec{r}_{K_L} = \{\vec{r}_v, z_v\}$  and the two calorimeter hits for the unconstrained photon pair ( $\gamma_{1a,2a}$  in Fig. 8) at  $\vec{r}_{ia}$ ,  $i = 1, 2$ , respectively. Thus the direction of the ALP momentum  $\vec{e}_a$  is known via

$$\vec{e}_a \sim [(\vec{r}_{1a} - \vec{r}_{K_L}) \times (\vec{r}_{2a} - \vec{r}_{K_L})] \times [\vec{e}_{K_L} \times \vec{e}_{2\pi^0}], \quad (10)$$

in which we use the notation “ $\vec{e} \sim \vec{k}$ ” to mean the unit vector  $\vec{e} = \vec{k}/|\vec{k}|$ .

Defining an ALP decay displacement parameter,  $\lambda$ , then the unit directions of the ALP daughter photon momenta  $\vec{e}_{ia}(\lambda) \sim \vec{r}_{ia} - (\lambda \vec{e}_a + \vec{r}_{K_L})$ , and the momenta  $\vec{k}_{ia}(\lambda) = E_{ia} \vec{e}_{ia}(\lambda)$ . At truth level,  $\vec{e}_a \sim \vec{k}_{1a} + \vec{k}_{2a}$ , so the spatial momenta obey the relation  $(\vec{k}_{1a} + \vec{k}_{2a})^2 - [(\vec{k}_{1a} + \vec{k}_{2a}) \cdot \vec{e}_a]^2 = 0$ . However,  $\vec{e}_a$  is necessarily always imperfectly reconstructed via Eq. (10) because of finite detector resolution effects. Instead, we seek a value for  $\lambda$ ,  $\lambda_{K_L}$ , such that the total invariant mass

$$\sqrt{[p_{2\pi^0} + k_{1a}(\lambda_{K_L}) + k_{2a}(\lambda_{K_L})]^2} = m_{K_L}, \quad (11)$$

whence the reconstructed ALP mass  $m_a^{\text{rec}} = \sqrt{[k_{1a}(\lambda_{K_L}) + k_{2a}(\lambda_{K_L})]^2}$ , the diphoton invariant mass. We do not reconstruct events for which Eq. (11) has no solution. (Alternatively, in the case of no solution one may select  $\lambda_{K_L} = 0$ , which also appears to work well for reconstructing  $m_a$ , perhaps as zero is an unbiased estimator for the true  $\lambda_{K_L}$ .)

## B. Signal and background estimation

Unlike in Sec. III, in which we used the measured  $m_{K_L}^{\text{rec}}$  distribution to set limits, to illustrate the potential KOTO limits from  $m_a^{\text{rec}}$  we must estimate both the  $K_L \rightarrow 3\pi^0$  contribution as well as the continuum from accidental overlay. To simulate the  $K_L \rightarrow 3\pi^0$  contribution, we generate a simulated sample with  $3 \times 10^7$  initial  $K_L$ 's at the BE, and pass it through the above-described modified reconstruction, which we emphasize imposes different requirements versus those for the  $m_{K_L}^{\text{rec}}$  KOTO data in Sec. II. In particular, for the third diphoton pair, we set the small but finite lifetime  $c\tau = c\tau_\pi$  and require Eq. (11) to have a solution. As before, we assume that the differential decay rate is well approximated as pure phase space.

The statistics of a sample of this size, when passed through the selection of Sec. II B, roughly corresponds to the statistics in the KOTO  $m_{K_L}^{\text{rec}}$  data for  $2 \times 10^{14}$  POT, i.e.,  $70 \times 10^3$  events. The acceptance of the modified reconstruction is approximately five times lower, such that one obtains a sample with statistics equivalent to  $12 \times 10^3$  events. The resulting  $m_a^{\text{rec}}$  distribution is shown in Fig. 9 (orange). We also show in Fig. 9 an example distribution for  $m_a = 100$  MeV with  $c\tau = 1$  cm (blue) and for  $m_a = 50$  MeV with  $c\tau = 1$  cm (light red). In the latter examples, typical ALP displacement  $c\tau\beta\gamma$  can be  $\gtrsim 10$  cm. For the  $\pi^0$  and 100 MeV cases, the pseudoscalar mass is faithfully reconstructed, while for 50 MeV the peak starts to degrade.

To estimate the accidental overlay continuum is more challenging. As an approximate treatment we imagine the continuum can be well approximated by  $K_L \rightarrow 3\pi^0$ ,

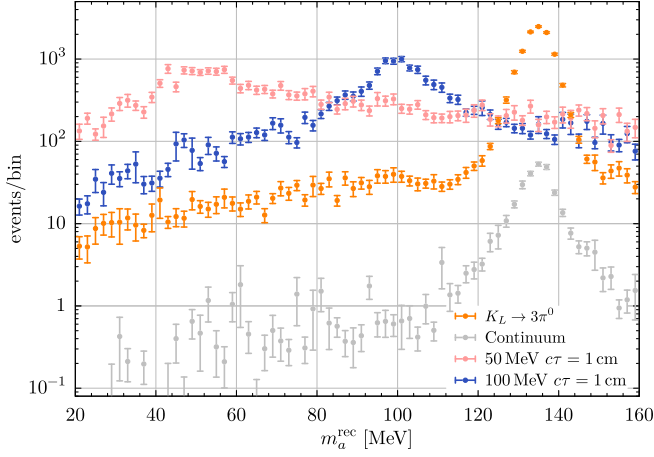


FIG. 9. Distributions of the reconstructed ALP mass,  $m_a^{\text{rec}}$ , per the selection of Sec. IV and with statistics roughly corresponding to  $2 \times 10^{14}$  POT for three cases:  $m_a = m_\pi$  with  $c\tau = c\tau_\pi$  (orange), corresponding to the background from  $K_L \rightarrow 3\pi^0$ ,  $m_a = 100$  MeV with  $c\tau = 1$  cm (blue), and  $m_a = 50$  MeV with  $c\tau = 1$  cm (light red). Also shown is the contribution to the continuum background (light gray).

but allowing the parent  $K_L$  mass to vary uniformly between  $3m_\pi^0$  and 600 MeV, per the range of the  $m_{K_L}^{\text{rec}}$  KOTO data. To approximate the relative contributions of  $K_L \rightarrow 3\pi^0$  events ( $B_{3\pi^0}$ ) versus the continuum ( $B_{\text{cont}}$ ) to the total background  $B_{\text{tot}}$ , i.e., to determine the fraction  $f$  in

$$B_{\text{tot}} = (1 - f)B_{3\pi^0} + fB_{\text{cont}}, \quad (12)$$

we simulate normalized  $m_{K_L}^{\text{rec}}$  distributions for  $B_{\text{cont}}$  and  $B_{\text{tot}}$ , to be fit to the measured  $m_{K_L}^{\text{rec}}$  spectrum. We generate a  $B_{\text{cont}}$  sample with  $\simeq 5 \times 10^6$  initial  $K_L$ 's. Fitting the linear combination (12) to the measured  $m_{K_L}^{\text{rec}}$  spectrum, we obtain  $f \simeq 0.03$ . As  $\simeq 99\%$  of the  $K_L \rightarrow 3\pi^0$  contribution lies within an approximately  $\pm 16$  MeV window of the peak, then the continuum contribution dominates the background outside this window, as expected from Fig. 2. The 68% CL band for the ratio of this MC simulation versus the  $m_{K_L}^{\text{rec}}$  data is shown in Fig. 2 (bottom panel, light red band) and exhibits generally good agreement across the entire  $m_{K_L}^{\text{rec}}$  range. The corresponding continuum contribution to the  $m_a^{\text{rec}}$  distribution is shown in Fig. 9 (light gray). It features the same morphology as the  $K_L \rightarrow 3\pi^0$  background (in orange) and can be safely neglected for the purpose of estimating limits.

### C. Naive projections

For the purpose of demonstrating the approximate reach of KOTO, we use the purely statistical uncertainties of the combined background  $B_{\text{tot}}$  to naively approximate the full uncertainties in a KOTO measurement with statistics corresponding to  $2 \times 10^{14}$  POT, and treat them as residuals

to which a signal hypothesis may be compared. To obtain naive limits, we generate simulated samples for the  $K_L \rightarrow 2\pi^0 a$  signal including finite ALP displacements, with  $3 \times 10^6$  initial  $K_L$ 's, which is again sufficient for subleading statistical uncertainties versus the background. We consider  $m_a = 40, 50, \dots, 130$  MeV and  $c\tau = c\tau_\pi, 1$  cm, and 10 cm, using a 2 MeV binning for the  $m_a^{\text{rec}}$  distribution between 20 MeV and 160 MeV.

Assuming that the signal MC uncertainties remain negligible and that there are no bin-bin correlations, we set a 90% CL limit for a given mass and lifetime hypothesis via  $\chi^2(m_a, c\tau) \simeq \sum_i [\rho S_i(m_a, c\tau) / \sigma_{B_{\text{tot},i}}]^2$ , with  $\rho$  defined as in Eq. (7),  $S_i$  the  $i$ th signal bin for  $K_L \rightarrow 2\pi^0 a$ , and  $\sigma_{B_{\text{tot},i}}$  the background uncertainty for that bin. Because  $\rho \geq 0$  the  $\chi^2$  statistic is a 1:1 admixture of a 0 and 1 degree of freedom  $\chi^2$ , as above, the 90% CL corresponds to  $\chi^2(m_a, c\tau) \leq 1.64$ . The projected branching ratio limits are shown in Fig. 10 for  $2 \times 10^{14}$  POT (solid) and are similar to the limits obtained at the boundary of the pion chimney. The dependence on  $m_a$  is somewhat weak over the considered mass range, except within the pion chimney range in which they begin to weaken: whether future limits set inside the pion chimney via the  $m_a^{\text{rec}}$  spectrum are more powerful than those from  $m_{K_L}^{\text{rec}}$  requires a dedicated study within the experimental framework itself. We also show limits for  $6 \times 10^{19}$  POT, roughly corresponding to the size of the 2016–2021 runs, and the expected future  $10^{21}$  POT (dotted), in both cases simply scaling by  $1/\sqrt{\text{POT}}$ .

While one might expect that signals with larger ALP lifetimes should exhibit cleaner separation versus the prompt  $3\pi^0$  background, we see in Fig. 10 that the  $c\tau = 1$  and 10 cm benchmarks have slightly weaker limits versus the prompt case. This may in part be because of isolation requirements that favor ALPs decaying farther

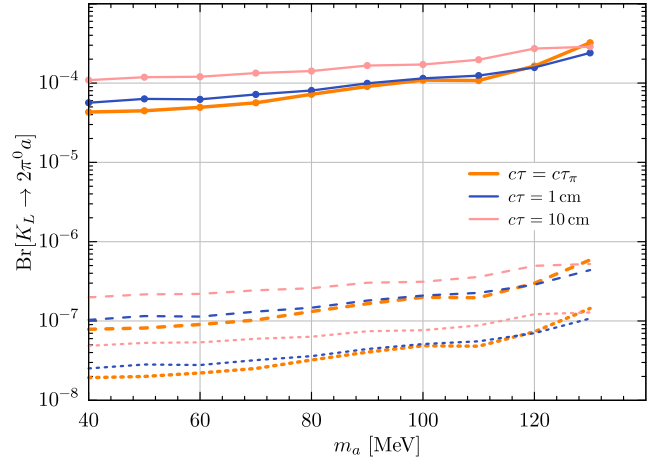


FIG. 10. Naive estimates for the 90% CL limits for  $\text{Br}[K_L \rightarrow 2\pi^0 a]$  with  $2 \times 10^{14}$  POT (solid),  $6 \times 10^{19}$  POT (dashed), and  $10^{21}$  POT (dotted) for ALPs with lifetime  $c\tau = c\tau_\pi$  (orange), 1 cm (blue), and 10 cm (magenta).

from the CsIC. Similarly, for larger lifetimes one expects more and more ALPs to decay beyond the DV and after the CsIC, reducing the acceptance. In the long-lifetime regime, e.g.,  $c\tau = 1$  m, for which the typical displacement is well beyond the DV, the fraction of decays in the DV scales as  $1/c\tau$ , so that the branching fraction limits should correspondingly weaken. We anticipate that further study of the selection and reconstruction for  $m_a^{\text{rec}}$  within the full experimental framework, including a study of tight timing requirements, may be able to achieve improved efficiencies for displaced ALPs.

Comparing our estimate of the branching ratio reach in Fig. 10 to the predicted branching ratios and corresponding lifetimes in the different benchmark scenarios in Fig. 6, we observe that a search for  $K_L \rightarrow 2\pi^0 a \rightarrow 6\gamma$  at KOTO could be competitive with searches for the two-body decay  $K_L \rightarrow \pi^0 a$  in the mass range  $m_a < m_{\pi^0}$ , particularly in the benchmark model shown in the top of Fig. 6. In general, however, these neutral kaon searches have less constraining power than those involving the corresponding charged kaon decay,  $K^+ \rightarrow \pi^+ a$ .

## V. SUMMARY AND CONCLUSIONS

In this work we showed that  $K_L \rightarrow 2\pi^0 a$  decays, with ALP masses within  $\mathcal{O}(10)$  MeV of the neutral pions, are able to pass the KOTO selection for the  $K_L \rightarrow 3\pi^0 \rightarrow 6\gamma$  calibration data with  $2 \times 10^{14}$  POT. These decays induce additional single or multippeak structures in the reconstructed  $K_L$  mass spectrum,  $m_{K_L}^{\text{rec}}$ , that are shifted with respect to the peak at  $m_{K_L}$ . This allows for the KOTO calibration data to be used in principle to set limits on  $\text{Br}[K_L \rightarrow 2\pi^0 a]$ , for ALP mass ranges that are typically hard to probe by other means—a mass range we have called the pion chimney.

In the absence of priors for SM background contributions to the  $m_{K_L}^{\text{rec}}$  line shape, we showed that one may nonetheless derive limits on  $\text{Br}[K_L \rightarrow 2\pi^0 a]$  via phenomenological Voigt-style parametrizations of the line shape profile, combined with standard information theoretic techniques and hypothesis testing. In the ALP mass range  $120 \leq m_a \leq 150$  MeV, we find branching ratio limits of order  $\text{few} \times 10^{-4}$  to  $\text{few} \times 10^{-3}$ , which correspond to novel limits in the  $m_a$  vs ALP coupling plane for several

commonly considered simplified ALP models. These results should be considered prospective, in the sense that model dependence effects from our phenomenological line shape description, that appear to be at most  $\mathcal{O}(1)$ , require quantification from precision simulations of the accidental overlay continuum contributions and other backgrounds. Further, a fully fledged analysis within an experimental framework is required to quantify and incorporate any correlations in the measured  $m_{K_L}^{\text{rec}}$  spectrum.

In addition, we explored the potential for KOTO to set branching ratio limits on  $\text{Br}[K_L \rightarrow 2\pi^0 a]$  via a modified selection that incorporates the possibility of ALP displaced decays-in-flight. This technique can be applied to ALP masses outside the pion chimney. Naive limits are obtained via the reconstructed diphoton mass spectrum from  $a \rightarrow \gamma\gamma$ , i.e.,  $m_a^{\text{rec}}$ , that are comparable to the limits obtained via  $m_{K_L}^{\text{rec}}$  within the chimney. These beyond-chimney limits may be competitive with other neutral kaon modes for a larger POT dataset depending on the NP model, though they generally cannot be competitive with limits from  $K^+ \rightarrow \pi^+ X$  searches. However, extension of either the  $m_{K_L}^{\text{rec}}$ - or  $m_a^{\text{rec}}$ -based searches to larger (future) KOTO datasets may permit almost the entirety of the pion chimney itself to be probed.

## ACKNOWLEDGMENTS

We thank Florian Bernlochner, Vladimir Gligorov, Yuval Grossman, and Simon Knapen for discussions. The research of R.B. and S.G. is supported in part by the U.S. Department of Energy, Grant No. DE-SC0010107. This work was performed in part at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-2210452. This research was supported in part by grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP). D.J.R. and C.S. are supported by the Office of High Energy Physics of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

## DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

- 
- [1] E. Abouzaid *et al.* (KTeV Collaboration), Final results from the KTeV experiment on the decay  $K_L \rightarrow \pi^0 \gamma\gamma$ , *Phys. Rev. D* **77**, 112004 (2008).  
 [2] S. Gori, G. Perez, and K. Tobioka, KOTO vs. NA62 dark scalar searches, *J. High Energy Phys.* **08** (2020) 110.

- [3] J. K. Ahn *et al.* (KOTO Collaboration), Search for the  $K_L \rightarrow \pi^0 \nu \bar{\nu}$  and  $K_L \rightarrow \pi^0 X^0$  decays at the J-PARC KOTO experiment, *Phys. Rev. Lett.* **122**, 021802 (2019).  
 [4] C. Lin *et al.* (KOTO Collaboration), Search for the pair production of dark particles  $X$  with  $K_L^0 \rightarrow XX$ ,  $X \rightarrow \gamma\gamma$ , *Phys. Rev. Lett.* **130**, 111801 (2023).

- [5] E. Abouzaid *et al.* (KTeV Collaboration), Detailed study of the  $K_L \rightarrow \pi^0 \pi^0 \pi^0$  dalitz plot, *Phys. Rev. D* **78**, 032009 (2008).
- [6] R. Balkin, S. Gori, and C. Scherb, Probing  $CP$  and flavor violation in neutral kaon decays with ALPs, *J. High Energy Phys.* **02** (2026) 011.
- [7] E. Goudzovski *et al.*, New physics searches at kaon and hyperon factories, *Rep. Prog. Phys.* **86**, 016201 (2023).
- [8] T. Kitahara, T. Okui, G. Perez, Y. Soreq, and K. Tobioka, New physics implications of recent search for  $K_L \rightarrow \pi^0 \nu \bar{\nu}$  at KOTO, *Phys. Rev. Lett.* **124**, 071801 (2020).
- [9] E. Izaguirre, T. Lin, and B. Shuve, Searching for axionlike particles in flavor-changing neutral current processes, *Phys. Rev. Lett.* **118**, 111802 (2017).
- [10] M. Freytsis, Z. Ligeti, and J. Thaler, Constraining the axion portal with  $B \rightarrow K l^+ l^-$ , *Phys. Rev. D* **81**, 034001 (2010).
- [11] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, Flavor probes of axion-like particles, *J. High Energy Phys.* **09** (2022) 056.
- [12] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, Consistent treatment of axions in the weak chiral lagrangian, *Phys. Rev. Lett.* **127**, 081803 (2021).
- [13] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, The low-energy effective theory of axions and ALPs, *J. High Energy Phys.* **04** (2021) 063.
- [14] M. Chala, G. Guedes, M. Ramos, and J. Santiago, Running in the ALPs, *Eur. Phys. J. C* **81**, 181 (2021).
- [15] J. Martin Camalich, M. Pospelov, P.N.H. Vuong, R. Ziegler, and J. Zupan, Quark flavor phenomenology of the QCD axion, *Phys. Rev. D* **102**, 015023 (2020).
- [16] K. Choi, S.H. Im, C.B. Park, and S. Yun, Minimal flavor violation with axion-like particles, *J. High Energy Phys.* **11** (2017) 070.
- [17] A. Celis, J. Fuentes-Martín, and H. Serôdio, An invisible axion model with controlled fcncs at tree level, *Phys. Lett. B* **741**, 117 (2015).
- [18] Y.H. Ahn, Flavored Peccei-Quinn symmetry, *Phys. Rev. D* **91**, 056005 (2015).
- [19] J.L. Feng, T. Moroi, H. Murayama, and E. Schnapka, Third generation familons,  $b$  factories, and neutrino cosmology, *Phys. Rev. D* **57**, 5875 (1998).
- [20] K.S. Babu and S.M. Barr, Family symmetry, gravity, and the strong  $CP$  problem, *Phys. Lett. B* **300**, 367 (1993).
- [21] A. Davidson and A. Vozmediano, The horizontal axion alternative: The interplay of vacuum structure and flavor interactions, *Nucl. Phys. B* **248**, 647 (1984).
- [22] A. Davidson, V.P. Nair, and K.C. Wali, Peccei-Quinn symmetry as flavor symmetry and grand unification, *Phys. Rev. D* **29**, 1504 (1984).
- [23] D.B. Reiss, Can the family group be a global symmetry?, *Phys. Lett.* **115B**, 217 (1982).
- [24] A. Davidson and K.C. Wali, Minimal flavor unification via multigenerational Peccei-Quinn symmetry, *Phys. Rev. Lett.* **48**, 11 (1982).
- [25] Y. Ema, K. Hamaguchi, T. Moroi, and K. Nakayama, Flaxion: A minimal extension to solve puzzles in the standard model, *J. High Energy Phys.* **01** (2017) 096.
- [26] L. Calibbi, F. Goertz, D. Redigolo, R. Ziegler, and J. Zupan, Minimal axion model from flavor, *Phys. Rev. D* **95**, 095009 (2017).
- [27] C. Cornella, A. M. Galda, M. Neubert, and D. Wyler,  $K^\pm \rightarrow \pi^\pm a$  at next-to-leading order in chiral perturbation theory and updated bounds on ALP couplings, *J. High Energy Phys.* **06** (2024) 029.
- [28] C. Cornella, P. Paradisi, and O. Sumensari, Hunting for ALPs with lepton flavor violation, *J. High Energy Phys.* **01** (2020) 158.
- [29] F. Björkeröth, E. J. Chun, and S. F. King, Flavourful axion phenomenology, *J. High Energy Phys.* **08** (2018) 117.
- [30] F. Ertas and F. Kahlhoefer, On the interplay between astrophysical and laboratory probes of MeV-scale axion-like particles, *J. High Energy Phys.* **07** (2020) 050.
- [31] M. B. Gavela, R. Houtz, P. Quilez, R. Del Rey, and O. Sumensari, Flavor constraints on electroweak ALP couplings, *Eur. Phys. J. C* **79**, 369 (2019).
- [32] Y. Afik, B. Döbrich, J. Jerhot, Y. Soreq, and K. Tobioka, Probing long-lived axions at the KOTO experiment, *Phys. Rev. D* **108**, 055007 (2023).
- [33] T. Masuda *et al.*, Long-lived neutral-kaon flux measurement for the KOTO experiment, *Prog. Theor. Exp. Phys.* **2016**, 013C03 (2016).
- [34] T. Masuda, Development and experimental study of the KOTO detector system using three  $K_L$  neutral decay modes, Ph.D. thesis, Kyoto University, 2014.
- [35] S. Navas *et al.* (Particle Data Group), Review of particle physics, *Phys. Rev. D* **110**, 030001 (2024).
- [36] W. Voigt, *Über das Gesetz der Intensitätsverteilung innerhalb der Linien eines Gasspektrums*, Sitzungsberichte, (München, 1912), Vol. 1912, p. 25, <https://publikationen.baw.de/de/003395768>.
- [37] R. A. Kycia and S. Jadach, Relativistic Voigt profile for unstable particles in high energy physics, *J. Math. Anal. Appl.* **463**, 1040 (2018).
- [38] K. Burnham and D. Anderson, *Model Selection and Multi-model Inference: A Practical Information-Theoretic Approach* (Springer Verlag, Berlin, 2002).
- [39] H. Chernoff, On the distribution of the likelihood ratio, *Ann. Math. Stat.* **25**, 573 (1954).
- [40] E. Cortina Gil *et al.* (NA62 Collaboration), Measurement of the very rare  $K^+ \rightarrow \pi^+ \nu \bar{\nu}$  decay, *J. High Energy Phys.* **06** (2021) 093.
- [41] E. Cortina Gil *et al.* (NA62 Collaboration), Measurement of the  $K \rightarrow \pi^+ \gamma \gamma$  decay, *Phys. Lett. B* **850**, 138513 (2024).
- [42] A. Lai *et al.* (NA48 Collaboration), Precise measurement of the decay  $K_L \rightarrow \pi^0 \gamma \gamma$ , *Phys. Lett. B* **536**, 229 (2002).
- [43] R. Ogata *et al.* (E391a Collaboration), Study of the  $K_L^0 \rightarrow \pi^0 \pi^0 \nu \bar{\nu}$  decay, *Phys. Rev. D* **84**, 052009 (2011).
- [44] P. del Amo Sanchez *et al.* (BABAR Collaboration), Search for production of invisible final states in single-photon decays of  $\Upsilon(1S)$ , *Phys. Rev. Lett.* **107**, 021804 (2011).
- [45] A. V. Artamonov *et al.* (E949 Collaboration), Search for the decay  $K^+ \rightarrow \pi^+ \gamma \gamma$  in the  $\pi^+$  momentum region  $P > 213$  MeV/c, *Phys. Lett. B* **623**, 192 (2005).
- [46] J. K. Ahn *et al.* (KOTO Collaboration), Search for the  $K_L \rightarrow \pi^0 \nu \bar{\nu}$  decay at the J-PARC KOTO experiment, *Phys. Rev. Lett.* **134**, 081802 (2025).
- [47] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of the top quark polarization and  $t\bar{t}$  spin correlations using dilepton final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV, *Phys. Rev. D* **100**, 072002 (2019).

- [48] Combination of the ATLAS, CMS and LHCb results on the  $B_{(s)}^0 \rightarrow \mu^+\mu^-$  decays (2020).
- [49] R. Mammen Abraham *et al.* (FASER Collaboration), Shining light on the dark sector: Search for axion-like particles and other new physics in photonic final states with FASER, *J. High Energy Phys.* **01** (2025) 199.
- [50] J. Blumlein *et al.*, Limits on the mass of light (pseudo)scalar particles from Bethe-Heitler  $e^+e^-$  and  $\mu^+\mu^-$  pair production in a proton—iron beam dump experiment, *Int. J. Mod. Phys. A* **07**, 3835 (1992).
- [51] M. Ovchinnikov and A. Zaporozhchenko, Advancing the phenomenology of GeV-scale axionlike particles, *Phys. Rev. D* **112**, 015001 (2025).
- [52] J. L. Goity, A. M. Bernstein, and B. R. Holstein, The decay  $\pi^0 \rightarrow \gamma\gamma$  to next to leading order in chiral perturbation theory, *Phys. Rev. D* **66**, 076014 (2002).