Probing axion like particles at the LHC with the CMS Experiment

PHENO 2026

Gonzalo Alonso Álvarez, Chiara Amendola, Matteo Cremonesi, Sena Durgut, Joerg Jaeckel, Matteo Marchegiani, Mariel Peczak

May 12, 2026

Benchmark ALP model

The benchmark model for many ALP studies is that of an ALP coupled electromagnetically to the SM via the Lagrangian:

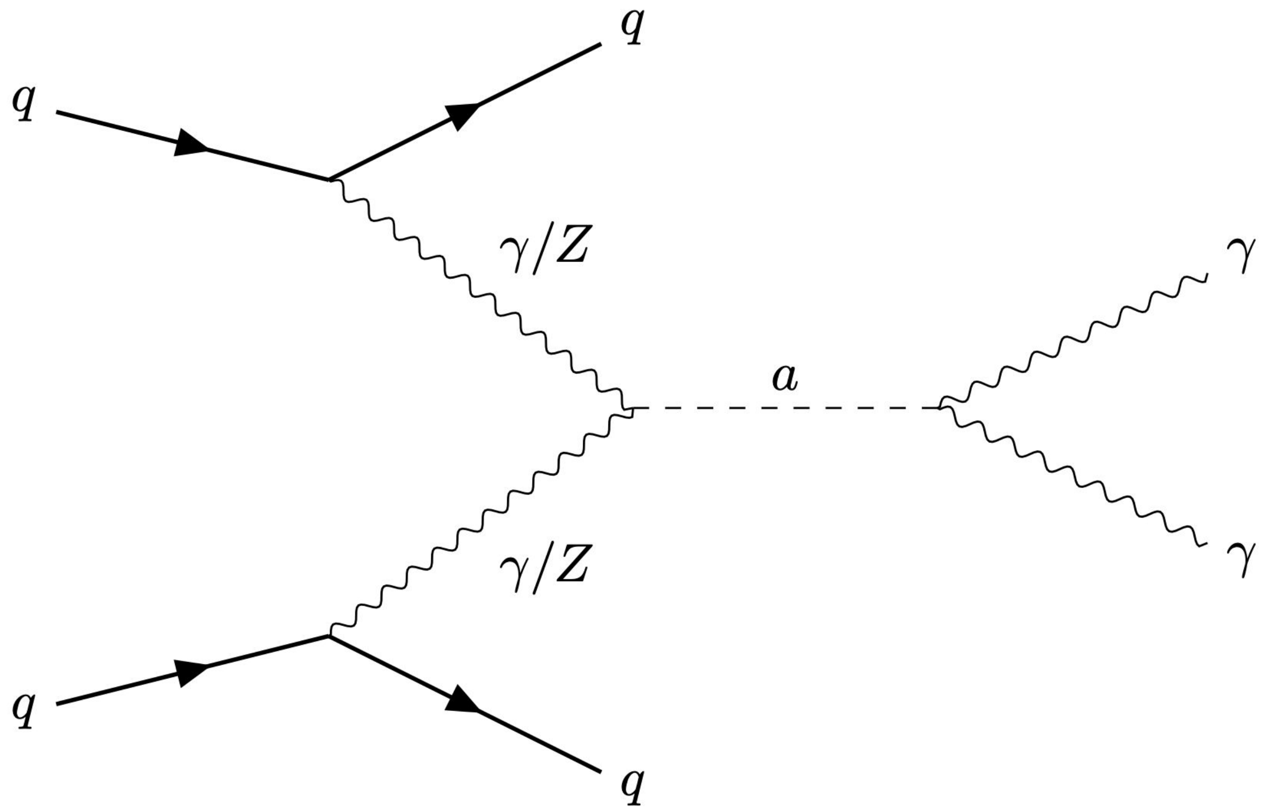
$$\mathcal{L} = \frac{1}{2}(\partial_\mu a)^2 - \frac{1}{2}m_a^2 - \frac{1}{4}g_{a\gamma\gamma}\, a\, F_{\mu\nu}\tilde{F}^{\mu\nu}$$

Consider an ALP coupled to the hypercharge gauge boson field strength Bμν:

$$\begin{aligned}\mathcal{L} &\supset -\frac{1}{4}\frac{1}{c_W^2} g_{a\gamma\gamma}\, a\, B_{\mu\nu}\tilde{B}^{\mu\nu} \\ &= -\frac{1}{4} g_{a\gamma\gamma}\, a\, F_{\mu\nu}\tilde{F}^{\mu\nu} - \frac{1}{4}\frac{s_W^2}{c_W^2} g_{a\gamma\gamma}\, a\, Z_{\mu\nu}\tilde{Z}^{\mu\nu} + \frac{1}{4}\frac{2s_W}{c_W} g_{a\gamma\gamma}\, a\, F_{\mu\nu}\tilde{Z}^{\mu\nu}.\end{aligned}$$

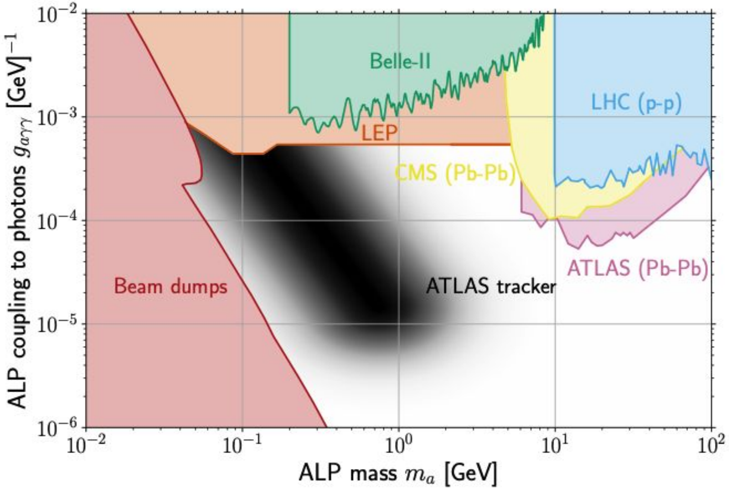
This generates γγ→a, ZZ→a, and γZ→a, all contribute comparably at high energies to VBF production at the LHC.

Benchmark ALP model



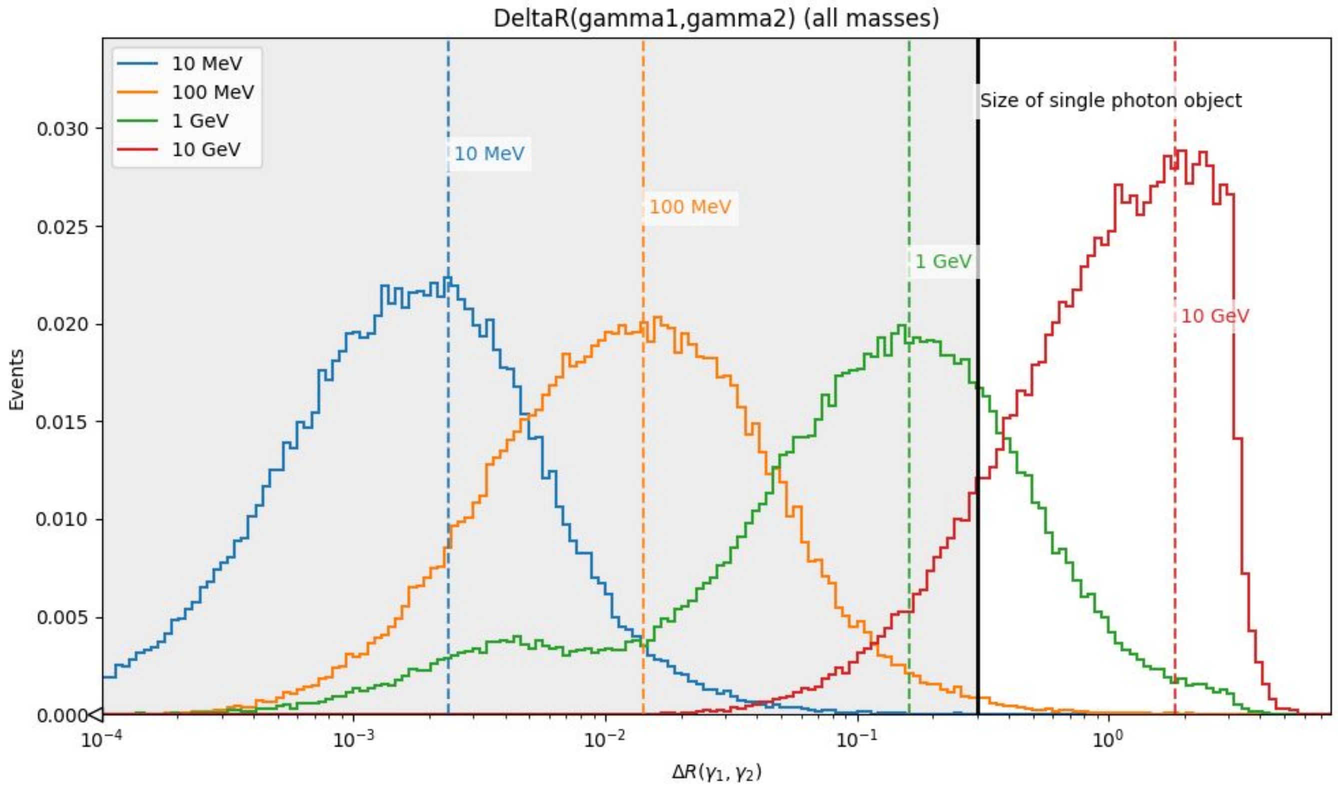
The ALP Landscape: An Open Window

- Mass range 10 MeV - 10 GeV with g ~ 10^{-6} - 10^{-3} GeV^{-1}: wide-open parameter space between beam dump and collider sensitivities
- LHC produces several × 10^5 ALPs in this range via VBF at Run 3 luminosity
- Since the ALPs have low mass, the two decay photons end up being collimated and merged.



https://arxiv.org/abs/2302.12262

The Collimation Challenge



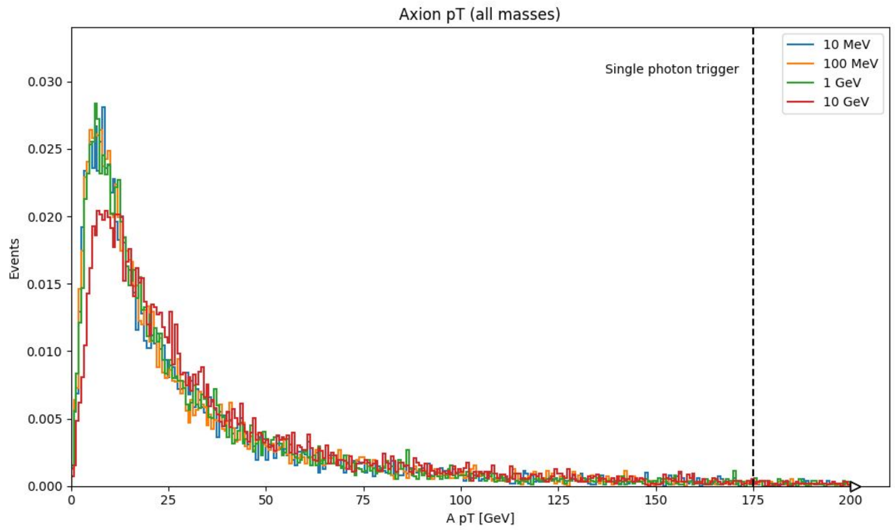
VBF Trigger

Problem:

- CMS two-tier (Level-1(L1) and High Level Trigger (HLT)) trigger reduces 30 MHz collision rate to ~O(1) kHz for storage
- High trigger thresholds needed to manage SM backgrounds ⟶ signal events with low pT are discarded
- For ALPs: Most of the signal is below 50 GeV, using single photon triggers is not ideal

⟶ Trigger on the jets using dedicated VBF triggers to enable the ALP searches at CMS.

But how to make the threshold on the jets as soon as possible?



Alternative Trigger Strategies

Data parking: Defer reconstruction →record at **higher rate**

- Store full raw data but skip prompt reconstruction
- Process later when computing resources are available (months/years)
- Full offline-quality event content preserved

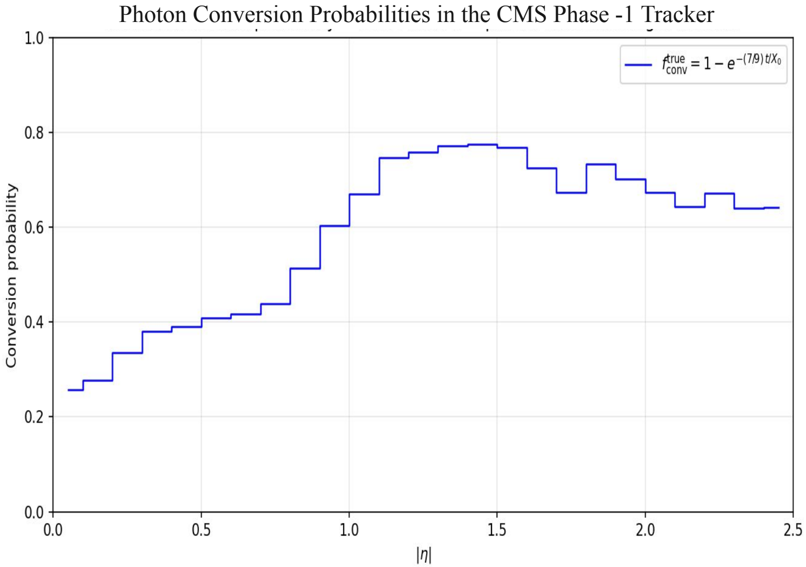
Data Scouting: reduce event size ⟶ record at **higher rate**

- Store only trigger-level objects, not full raw data
- ~10 KB/event vs ~1 MB/event ⟶ 10–100× higher rate for same bandwidth
- In HL-LHC, we will have data scouting available at L1 Trigger.

CMS-EXO-23-007

Run 3 Analysis Strategy (Data Parking)

- Data parking allows us to keep the VBF thresholds as low as possible
- Since data parking preserves the full raw detector readout, a transformer based reconstruction method (like "ClusTEX" (arXiv:2603.18172) can exploit rich information, and:
  - reconstruct ALP decay displacement dd and diphoton invariant mass $m_{\gamma\gamma}$
  - exploit tracking information from converted photons to differentiate them from single photons, especially for low mass (<100 MeV) ALPs .



Phase-2 Analysis Strategy (L1 Data Scouting)

HL-LHC Trigger problem:

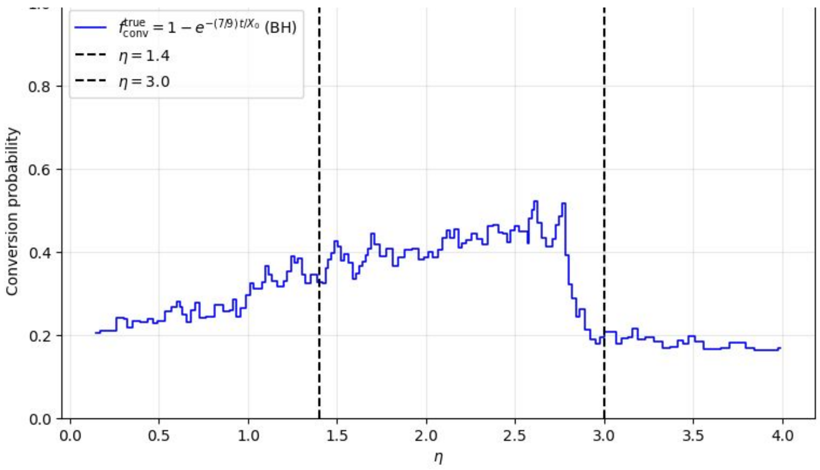
- Instantaneous luminosity 7.5 x 10^{34} cm^{-2} s^{-1}, 200 pileup (vs ~50 in Run 3) → standard triggers are even more restrictive to control rates

Proposed approach: Use L1 Data Scouting stream → processes every bunch crossing at 40 MHz, **no trigger thresholds**

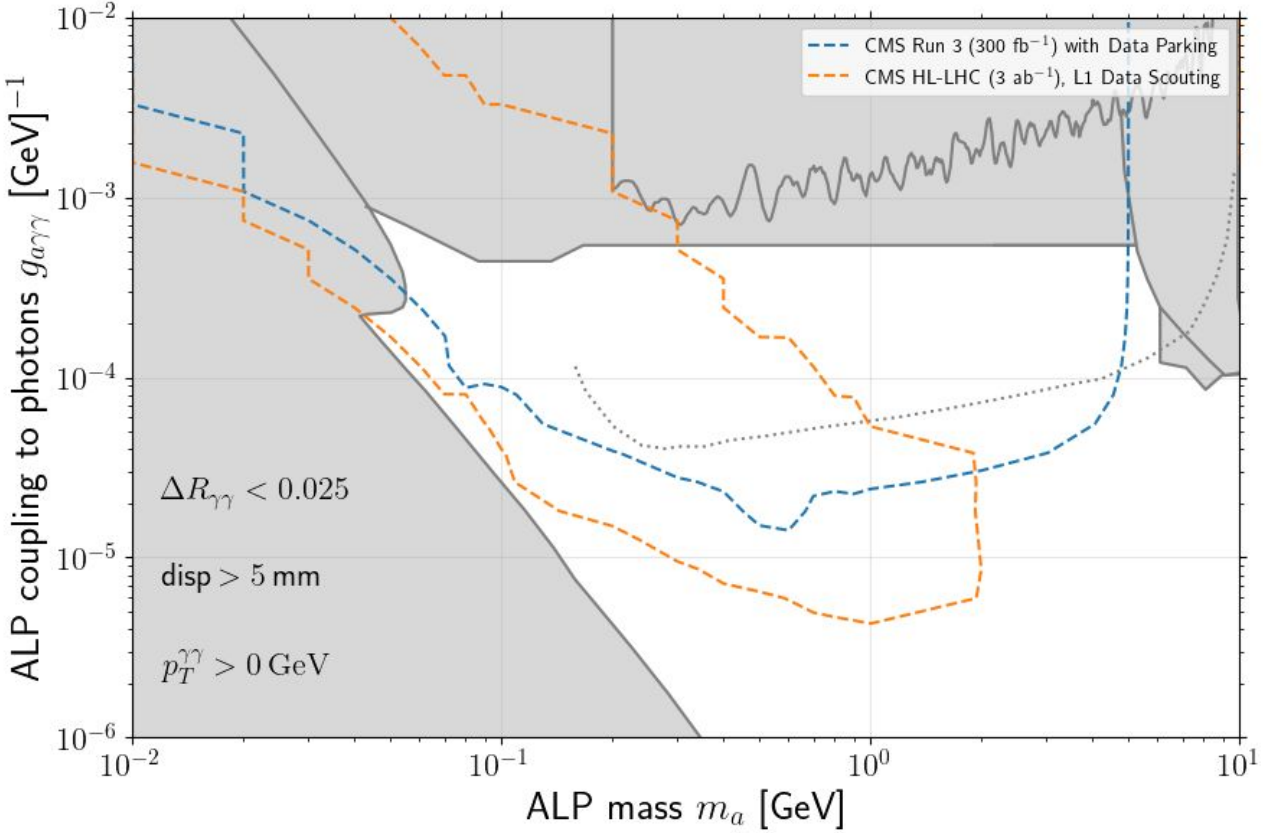
- Only L1-quality reconstruction → we don't have access to information on tracks and energy deposits at the same time → need a way to make the reject background
- Leverage the relationship between the lifetime and mass of the ALP, by using tracks to reconstruct the displaced vertex:

$$\Gamma(a \to \gamma\gamma) = \frac{g_{a\gamma\gamma}^2 m_a^3}{64\pi}$$

Photon Conversion Probabilities in the CMS Phase -1 Tracker



Projected Sensitivity: CMS Run 3 and Phase-2



Outlook and Path Forward

- First CMS study of VBF production of ALPs addressed by using alternative trigger strategies: Data Parking and Data Scouting!
- Very well aligned with progress with alternative trigger strategies at CMS. L1 Scouting is a unique capability we will have in Phase-2!
- Next step is to finalize the study and finalize the model that can be used for a real CMS search.

Backup

Run 3 Analysis Strategy (Data Parking)

Precedent: Transformer-based overlapping shower reconstruction: ClusTEX (arXiv:2603.18172)

- A graph transformer ("ClusTEX") trained on GEANT4 simulations of CMS ECAL barrel geometry reconstructs energy and position of overlapping EM showers directly from crystal-level readout
- Uses only calorimeter information, 7×7 crystal energy windows, no tracking
- Solves merged diphoton problem with 98.7% efficiency

Directly relevant to our ALP search, with room to improve:

- **Our ALPs provide two additional handles:**
    - Displaced decay vertex: conversion tracks from a displaced ALP point back to a non-IP origin
    - Full tracking from data parking: conversion track geometry (impact parameters, hit patterns, vertex fit) provides direct measurement of the two-photon topology, complementary to shower-shape discrimination, extends eta coverage
    - By explicitly feeding these features into the NN, we can exploit the full raw detector readout that parking provides