Rare decays of the SM Higgs to light pseudoscalars in the CMS experiment

Anagha Aravind for the CMS collaboration

University of Wisconsin-Madsion, USA

E-mail: aaravind@wisc.edu, aaravind@cern.ch

Abstract: An overview of selected searches where the SM Higgs boson decays to a pair of light pseudoscalars performed with p-p collisions data recorded during Run 2 by the CMS detector at $\sqrt{s}$ = 13 TeV is presented. Analyses with b quarks, leptons or photons in the final state are highlighted.

Presented at the 32nd International Symposium on Lepton Photon Interactions at High Energies, Madison, Wisconsin, USA, August 25-29, 2025

1. Introduction

A recent CMS measurement sets a limit of 12% at 95% CL on the decay of the Standard Model (SM) Higgs boson to beyond SM particles [1, 2]. The SM Higgs boson has a narrow decay width $\simeq 4.07$ MeV, which implies that even exotic processes with small branching fractions become relevant for new physics searches [3]. Interpreting extended Higgs-sector models provides a simple and well-motivated approach to exploring new physics. Such models may address several shortcomings of the SM, including the hierarchy problem, matter–antimatter asymmetry, and the nature of dark matter.

This article will focus on exotic searches in the CMS detector where the SM Higgs boson decays to two light pseudoscalars, referred to as a, with each a boson further decaying to SM particles. One popular theory for the interpretation of these searches is the 2 Higgs Doublet model (2HDM+S) [4].

2. Recent $h \to aa$ results at CMS Experiment

CMS has a rich program of $h \to aa$ searches. In this article, selected recent results that analyze the full Run-2 data for 10 MeV $< m_a <$ 60 GeV are discussed.

2.1 $h \to aa \to 4e$

This analysis probes $h \to aa$ with $4e$ in the final state, i.e, each $a \to ee$ where $10 < m_a < 100$ MeV [5]. The results are interpreted in Axion-Like Particle (ALP) models, where the a pseudoscalars are potential dark matter candidates.

Since the a's have very low masses, the final state e's are highly collimated, and the search is limited by the resolution of the ECAL detector of CMS. Selected events are required to have two Merged Electron Pairs (MEPs), which are reconstructed from pairs of tracks in the silicon tracker matched to ECAL super clusters within a distance of $\Delta R < 0.015$. In addition, an Adaptive Vertex Fitter is used to fit the vertex of the selected MEPs. Isolation cuts are applied to the selected MEPs to further suppress the background from QCD activity.

Major backgrounds in the analysis are SM photon conversions, SM processes with electrons and QCD multijet. To optimize signal extraction and background suppression, a multivariate identification technique is used, taking as input various characteristics of the MEPs.

The signal is extracted from the reconstructed Higgs boson mass of the two MEP system by fitting the data to a set of parametric shapes depending on the physics process. Stringent limits are set on the B($h \to aa \to 4e$) for different lifetimes of the ALP, with lower lifetimes being the most sensitive as shown in Fig. 1.

2.2 $h \to aa \to 4\gamma$

A search is performed for $h \to aa \to 4\gamma$, where one pair of $\gamma\gamma$ is reconstructed as fully resolved photons and the other pair is reconstructed as a merged photon [6]. The analysis is interpreted in NMSSM, ALP, and 2HDM theories.

Events are required to pass diphoton triggers. Selected events require exactly 3 reconstructed γ, where the invariant mass of the γ, $m_{\gamma\gamma\gamma} > 90$ GeV. The analysis strategy is to extract the $h \to aa \to 4\gamma$ signal from the 2D m_a spectra in the signal region defined by $\Delta m_a \leq 2$GeV and

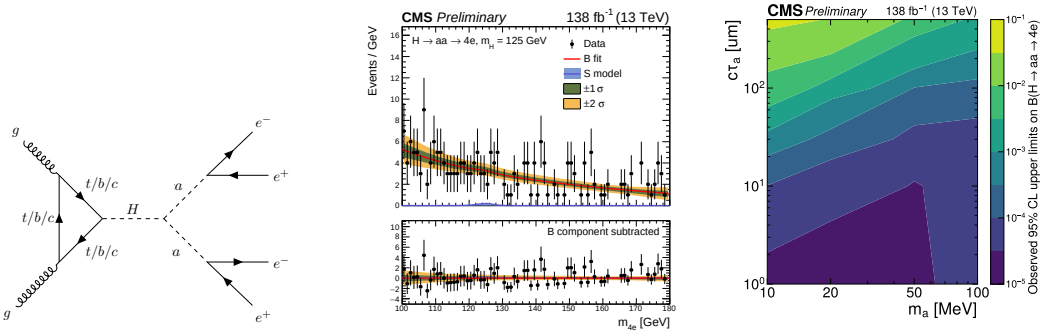

Figure 1: Left: Feynmann diagram of $h \to aa \to 4e$ where the Higgs is produced by gluon-gluon fusion. Middle: Invariant mass distribution of the 2 MEPs ($4e$) system compared to a fit of the background (red) and signal (blue) for $m_a = 20$ GeV and $c\tau = 10\mu m$. Right: The observed limits at 95% CL on the B($h \to aa \to 4e$) in the 2D decay length vs m_a plane [5].

$100 < m_{\gamma\gamma\gamma} < 140$ GeV. The a decaying to resolved $\gamma\gamma$ is reconstructed from the four-momentum sum of the resolved photons. To reconstruct the merged $\gamma\gamma$ pair, a Graphical Neural Network (GNN) is used, with energy deposit maps from the ECAL as input.

A side-band region is used for estimating the dominant backgrounds, such as QCD multijet, photon+jet and prompt diphoton, from data. This analysis sets an upper limit at 95% CL on the B($h \to aa \to 4\gamma$) of 0.264 to 0.008 pb for $1 < m_a < 15$ GeV as shown in Fig. 2.

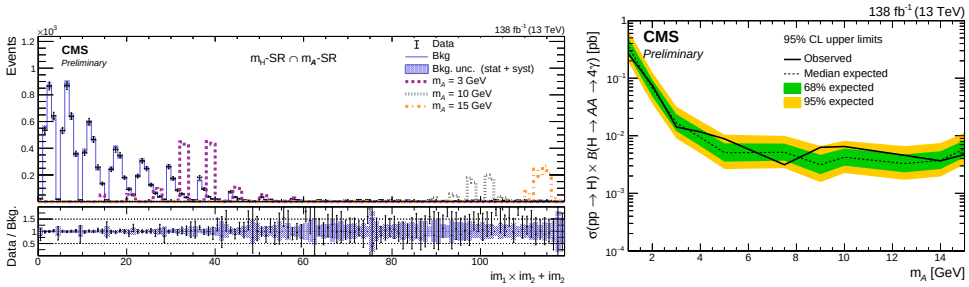

Figure 2: Left: Unrolled 2D-m_a spectra in the final signal region obtained by fixing $m_{A,1}$, while scanning along increasing bins of $m_{A,2}$. Right: Observed and median expected upper limits at 95% CL for $h \to aa \to 4\gamma$ [6].

2.3 $h \to aa \to 4\tau/2\mu 2\tau$

This analysis probes $h \to aa$ decay, where one a decays to $\tau\tau$ and the other a decays to either $\tau\tau$ or $\mu\mu$ [7]. The final state is $(\tau_\mu \tau_{\text{1-prong}})(\tau_\mu \tau_{\text{1-prong}})$, i.e., for each a, one τ decays into a μ and the other is a leptonic or 1-prong hadronic τ decay. Since the light a bosons are Lorentz boosted, the final products are highly collimated.

Events must pass dimuon trigger. Selected events require two same-signed muons that are well-separated, i.e. $\Delta R > 1.5$ and a charged track in the vicinity of each of these muons. Depending on the isolation and kinematics, these tracks are classified into "isolation" and "signal" (with tighter cuts) tracks. The "isolation" tracks are used for background estimation and "signal" tracks are used

in signal region. The signal is extracted from a 2D map filled with ordered pairs of (m_1,m_2) where $m_2 > m_1$.

The major background is QCD multijet, with minor contributions from $t\bar{t}$, W+Jets, Z+Jets, and diboson processes. These are modeled in the 2D distribution from 1-D distributions of the μ-track system transformed by a matrix dependent on the correlation between the 2 reconstructed a bosons.

The model independent observed limits at 95% CL on the $\sigma(pp \to H+X)B(h \to aa)B(a \to \tau\tau)$ relative to the SM Higgs boson production cross section (σ_{SM}) ranges from 0.007 to 0.079 for $4 \leq m_a \leq 15$ GeV. Most stringent limits are set in the Type III of the 2HDM+S interpretation as shown in Fig. 3.

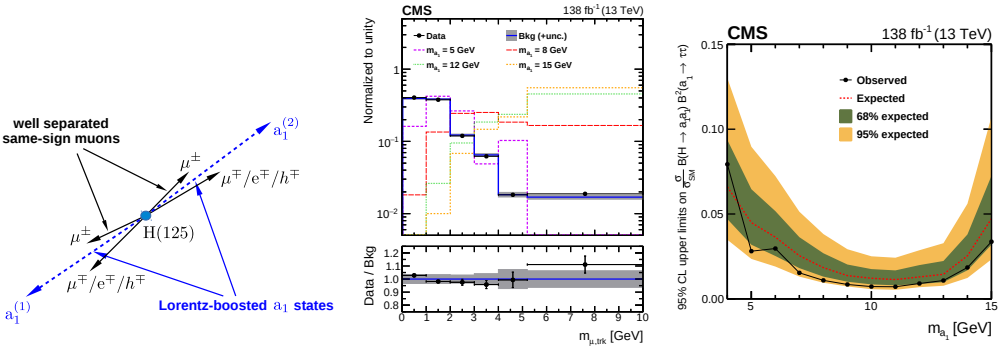

Figure 3: Left: Signal topology for $h \to aa \to 4\tau/2\mu2\tau$. Middle: Normalized invariant mass distribution of the μ-track system for events passing signal selection. Right: Observed and median expected upper limits at 95% CL for $\sigma(pp \to H+X)B(h \to aa)B(a \to \tau\tau)$ relative to the SM Higgs boson production cross section (σ_{SM}) [7].

2.4 $h \to aa \to bb\tau\tau$

A search is performed for $h \to aa \to bb\tau\tau$ where $12 < m_a < 60$ GeV [8]. The τs decay further and hence, three channels are considered: $e\mu$, $e\tau_h$, and $\mu\tau_h$. The $bb\tau\tau$ final state is chosen for it's favourable predicted branching fraction when interpreted in the 2HDM+S theory.

Selected events must pass either single lepton or di-lepton triggers depending on the channel. Each event must have exactly one $\tau\tau$ pair and at least one b-tagged jet. The two legs of the $\tau\tau$ pair must have opposite charges and a $\Delta R > 0.4$. The major backgrounds are QCD multijet background, jets misidentified as τ_h fakes, $Z \to \tau\tau$, $t\bar{t}$ and are estimated from a combination of data-driven methods and simulation.

Deep Neural Networks (DNN) are trained for each channel for the purpose of signal optimization, taking advantage of several distinguishing kinematic features between signal and background. The signal regions are defined by cutting on the transformed DNN score. The signal is then extracted from the invariant mass distribution of the reconstructed $\tau\tau$ system in the signal region using a maximum likelihood fit method.

This analysis sets an upper limit at 95% CL on the B($h \to aa \to bb\tau\tau$) at 1.8–7.7% for $12 < m_a < 60$ GeV for observed data. The results are interpreted in the 2HDM+S theory and also shown in combination with the $bb\mu\mu$ final state as shown in Fig. 4.

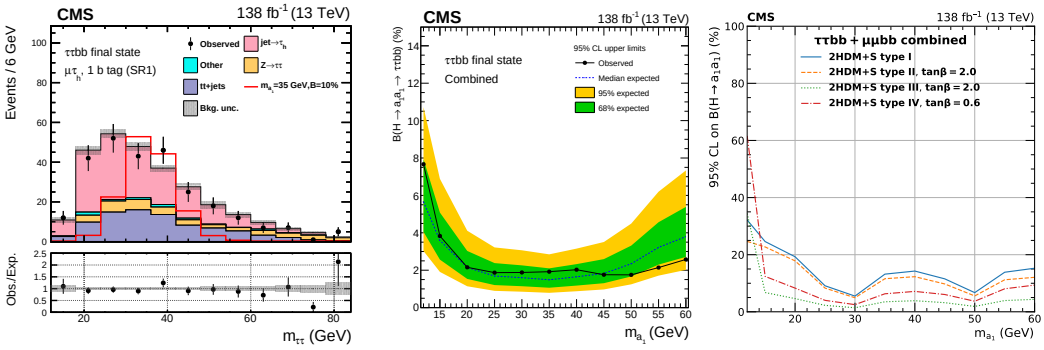

Figure 4: Left: $m_{\tau\tau}$ distribution for $\mu\tau$ channel in one of the signal regions. Middle: Observed and median expected upper limits at 95% CL for $h \to aa \to 2b2\tau$ for the combination of all 3 channels: $e\mu$, $e\tau_h$ and $\mu\tau_h$. Right: Observed upper limits at 95% CL for $h \to aa$ obtained from the combination of $2b2\tau$ and $\mu\mu bb$ final states for 2HDM+S Type I (independent of $\tan b$), Type II ($\tan b = 2.0$%), Type III ($\tan b = 2.0$), and Type IV ($\tan b = 0.6$), respectively [8].

3. Conclusion

The $h \to aa \to 4e$ result is the first $h \to aa$ search in CMS to set limits for $10 < m_a < 100$ MeV. The analysis with $h \to aa \to 4\gamma$ bridges the gap between two previous analyses in CMS (one with fully merged and the other with fully resolved final state [9, 10]). The $h \to aa \to 4\tau/2\mu2\tau$ result shows improvements over the previous CMS result [11] by a factor of 2. Similarly, $h \to aa \to bb\tau\tau$ in combination with $bb\mu\mu$ sets stringent limits for B($h \to aa \to bbll$) in the 2HDM+S model interpretation, for 12–60 GeV.

While several $h \to aa$ searches with different final states have been performed in CMS, no significant deviation from the SM have been observed. However, there is potential for other final states to be explored in the future, including searches with asymmetric decays, i.e. $h \to a_1 a_2$ where a_1, a_2 have different masses.

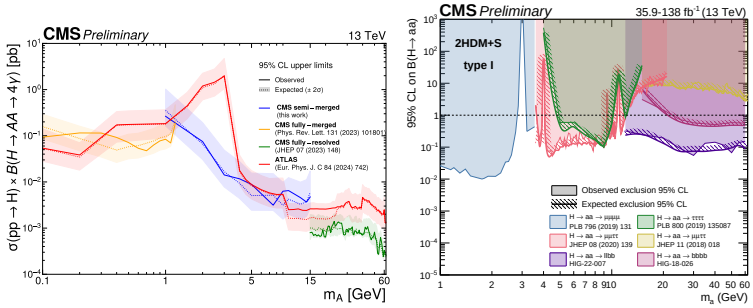

Figure 5: Left: CMS results for $h \to aa \to 4\gamma\gamma$ in different mass ranges shown in comparison with results from the ATLAS experiment [6, 9, 10]. Right: CMS summary plot of recent $h \to aa$ results for 2HDM+S Type I interpretation [12].

References

[1] CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3 (08):S08004, Aug 2008. doi: 10.1088/1748-0221/3/08/S08004. URL https://doi.org/

10.1088/1748-0221/3/08/S08004.

[2] CMS Collaboration. A portrait of the Higgs boson by the CMS experiment ten years after the discovery. *Nature*, 607(7917):60–68, 2022. doi: 10.1038/s41586-022-04892-x. URL http://dx.doi.org/10.1038/s41586-022-04892-x.

[3] David Curtin et al. Exotic decays of the 125 GeV Higgs boson. *Phys. Rev. D*, 90:075004, Oct 2014. doi: 10.1103/PhysRevD.90.075004. URL https://link.aps.org/doi/10.1103/PhysRevD.90.075004.

[4] G.C. Branco et al. Theory and phenomenology of two-Higgs-doublet models. *Physics Reports*, 516(1):1–102, 2012. ISSN 0370-1573. doi: https://doi.org/10.1016/j.physrep.2012.02.002. URL https://www.sciencedirect.com/science/article/pii/S0370157312000695.

[5] CMS Collaboration. Search for light pseudoscalar a bosons in H $\rightarrow$ aa $\rightarrow$ 4e decays in pp collisions at $\sqrt{s}$ = 13 TeV. *CMS Physics Analysis Summary CMS-PAS-EXO-24-031*, 2025. URL https://cds.cern.ch/record/2932415.

[6] CMS Collaboration. Search for exotic Higgs boson decays $H \rightarrow \mathcal{A}\mathcal{A}$ with $\mathcal{A} \rightarrow \gamma\gamma$ in events with a partially merged topology in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *CMS Physics Analysis Summary CMS-PAS-EXO-24-025*, 2025. URL https://cds.cern.ch/record/2937956.

[7] CMS Collaboration. Search for light pseudoscalar boson pairs produced from Higgs boson decays using the 4τ and $2\mu 2\tau$ final states in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Submitted to JHEP*, 2025. URL https://arxiv.org/abs/2508.06947.

[8] CMS Collaboration. Search for exotic decays of the Higgs boson to a pair of pseudoscalars in the $\mu\mu$bb and $\tau\tau$bb final states. *Eur. Phys. J. C*, 84:493, 2024. doi: 10.1140/epjc/s10052-024-12727-4. URL https://cds.cern.ch/record/2889907.

[9] CMS Collaboration. Search for exotic Higgs boson decays H $\rightarrow \mathcal{A}\mathcal{A} \rightarrow 4\gamma$ with events containing two merged diphotons in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Phys. Rev. Lett.*, 131:101801, 2023. doi: 10.1103/PhysRevLett.131.101801. URL https://cds.cern.ch/record/2826997.

[10] CMS Collaboration. Search for the exotic decay of the Higgs boson into two light pseudoscalars with four photons in the final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *JHEP*, 2307:148, 2023. doi: 10.1007/JHEP07(2023)148. URL https://cds.cern.ch/record/2823906.

[11] CMS Collaboration. Search for light pseudoscalar boson pairs produced from decays of the 125 GeV Higgs boson in final states with two muons and two nearby tracks in pp collisions at $\sqrt{s}$ = 13 TeV. *Phys. Lett. B*, 800:135087, 2020. doi: 10.1016/j.physletb.2019.135087. URL https://cds.cern.ch/record/2682977.

[12] CMS Collaboration. Summary of 2HDM+S searches at 13 TeV (Run 2). *CMS Public TWiki*, 2025. https://twiki.cern.ch/twiki/bin/view/CMSPublic/Summary2HDMSRun2.