

DRAFT

CMS Paper

The content of this note is intended for CMS internal use and distribution only

2025/06/18

Archive Hash: 7f1f256

Archive Date: 2025/06/18

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

The CMS Collaboration

Abstract

A search for a pair of light pseudoscalar bosons (a_1) produced in the decay of the 125 GeV Higgs (H) boson is presented. The analysis examines decay modes where one a_1 decays into a pair of tau (τ) leptons and the other decays into either another pair of τ leptons or a pair of muons (μ). The a_1 mass probed in this study ranges from 4 to 15 GeV. The data sample used was recorded by the CMS experiment in proton-proton collisions at a center-of-mass energy of 13 TeV and corresponds to an integrated luminosity of 138 fb^{-1} . No excess above standard model expectations is observed. The study combines the 4τ and $2\mu 2\tau$ channels to set upper limits at 95% confidence level on the product of the Higgs boson production cross section and the branching fraction to the 4τ final state, relative to the Higgs boson production cross section in the standard model. In this interpretation, the a_1 boson is assumed to have Yukawa-like couplings to fermions, with coupling strengths proportional to the respective fermion masses. The observed (expected) upper limits range between 0.007 (0.011) and 0.079 (0.066) across the mass range considered. The results are also interpreted in the context of models with two Higgs doublets and an additional complex singlet field.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	Lakshmi Priya Nair, Alexei Raspereza
PDFTitle:	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
PDFSubject:	CMS
PDFKeywords:	CMS, Higgs, pseudoscalars, 2HDM+S, muons, taus

Please also verify that the abstract does not use any user defined symbols

1 Introduction

Experimental verification of the standard model (SM), a quantum field theory describing the strong and electroweak interactions between elementary particles, culminated in the discovery of a scalar boson with a mass of 125 GeV by the ATLAS and CMS Collaborations at the CERN LHC [1–3]. Subsequent studies have demonstrated that the properties of the discovered boson, denoted henceforth as H, are compatible, at the current level of precision, with the predicted properties of the SM Higgs boson [4, 5].

Despite its striking success in describing the subatomic world, the SM is not considered a complete theory. It does not explain several key phenomena, such as the nature of dark matter, the origin of neutrino masses, and the matter-antimatter asymmetry in the universe. Various extensions to the SM have been proposed to address some of its shortcomings, featuring extended Higgs sectors in which the 125 GeV H boson appears as one of multiple Higgs boson states. Prominent examples of such extensions of the SM are the general Two Higgs Doublet models [6, 7], and Two Higgs Doublet models with an additional complex singlet field (2HD+S models) [8]. Extended Higgs sectors are embedded into more elaborate theoretical frameworks, such as the minimal supersymmetric SM (MSSM) [9, 10], which incorporates two Higgs doublets, or the next-to-MSSM (NMSSM) [11, 12], which solves the μ -problem of the MSSM by extending the two Higgs doublets by an additional complex singlet field [13, 14]. These supersymmetric extensions of the SM provide a candidate for dark matter and introduce additional sources of charge-parity violation that could explain the observed amount of matter-antimatter asymmetry.

The analysis presented in this paper probes the 2HD+S models, which predict seven physical Higgs boson states: three neutral CP-even, two neutral CP-odd, and two charged bosons. One of the scalars of the neutral sector can be identified as the 125 GeV Higgs boson. At the same time, one of the two pseudoscalars, denoted a_1 , can be light enough for $H \rightarrow a_1 a_1$ decays to be kinematically possible. Measurements of the Higgs boson couplings so far still allow for a substantial branching fraction of Higgs boson decays to beyond-the-SM (BSM) particles; the ATLAS and CMS experiments have established upper limits at 95% confidence level (CL) of 12% and 16%, respectively [4, 5]. The search for $H \rightarrow a_1 a_1$ decays, therefore, provides an effective probe for possible extensions of the SM and the discovery of new physics phenomena.

The 2HD+S models predict various decay channels for the a_1 boson, with significant branching fractions into fermion pairs. The decay patterns of a_1 are governed by parameters such as the ratio of the vacuum expectation values of the two Higgs doublets ($\tan \beta$) and the a_1 mass (m_{a_1}). The different variants of the 2HD+S models—Type I, II, III, and IV—are distinguished by the way the Higgs doublets couple to fermions, resulting in distinct decay patterns. Several of these models predict a significant branching fraction of the a_1 boson decay into a pair of tau (τ) leptons [8], making this channel a promising avenue for experimental searches.

Several searches for $H \rightarrow a_1 a_1$ decays have been carried out by the ATLAS and CMS Collaborations to date, exploring different decay modes of a_1 , covering the mass range $0.2 < m_{a_1} < 62.5$ GeV [15–33]. The studies did not reveal significant deviations from the SM background expectations, and upper limits were set on the signal rates, thereby constraining the parameters of the 2HD+S models.

In this paper, a search for a light a_1 boson via the decays $H \rightarrow a_1 a_1 \rightarrow 4\tau$ and $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$ is presented. The search is based on proton-proton (pp) collision data at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 138 fb^{-1} , recorded by the CMS detector in 2016–2018. The analysis covers the m_{a_1} range from 4 to 15 GeV and employs a spe-

cialized strategy to select and identify highly Lorentz-boosted muon or τ lepton pairs with overlapping decay products. Similar searches were performed by the CMS and ATLAS Collaborations [29, 33] using 36 fb^{-1} and 140 fb^{-1} of pp collision data, respectively. These searches established upper limits at 95% CL on the Higgs boson production cross section times branching fraction for $H \rightarrow a_1 a_1 \rightarrow 4\tau$, relative to the SM Higgs boson production cross section. The CMS analysis set limits in the range 0.022–0.23, depending on the mass of the a_1 boson. The ATLAS analysis, based on a larger analyzed data set, set limits in the range 0.03–0.10. The analysis presented here extends the previous CMS analysis by using approximately four times more data. Furthermore, the signal selection criteria have been optimized further, taking advantage of a more precise calibration of the CMS detector achieved towards the end of the LHC Run 2 data-taking period.

The signal topology targeted by the present analysis is illustrated in Fig. 1, where the Higgs boson decays into a pair of a_1 bosons. The $H \rightarrow a_1 a_1 \rightarrow 4\tau$ decay gives rise to τ leptons that have a softer spectrum of transverse momentum (p_T) compared to the $H \rightarrow \tau\tau$ decay. About 60% of τ leptons of the signal have a visible p_T lower than 20 GeV. Reconstruction and identification of τ leptons in this kinematic regime is challenging. In particular, the reconstruction of decay modes with neutral pions is hampered by a large flux of low- p_T photons originating from soft pp interactions in the same or neighboring bunch crossings (pileup). Conventional CMS algorithms [34] are not optimized for the identification of such soft τ leptons. As a consequence, the efficiency and purity of the τ reconstruction rapidly degrade with decreasing visible p_T . Overlap of τ leptons originating from the same a_1 boson can further complicate the identification of $a_1 \rightarrow \tau\tau$ candidates. These challenges are overcome by employing a dedicated search strategy.

Given the significant mass difference between the a_1 and Higgs bosons, the a_1 bosons are produced with a high Lorentz boost, leading to collimated decay products. Each a_1 boson is identified by the presence of a muon and a close-by additional charged particle, which could be an electron, muon, or charged hadron ($e^\pm/\mu^\pm/h^\pm$). These charged particles are characterized by the presence of one reconstructed track with a charge opposite to that of the muon. The analysis focuses on the $a_1 a_1 \rightarrow (\tau_\mu \tau_{1\text{-prong}})(\tau_\mu \tau_{1\text{-prong}})$ decays, where τ_μ denotes the muonic decay of a τ lepton, and $\tau_{1\text{-prong}}$ stands for its decay into an electron, muon, or 1-prong hadronic state. The 3-prong τ decays are excluded because of their very high quantum chromodynamics (QCD) multijet background and lower reconstruction efficiency. Although the analysis primarily targets the $(\tau_\mu \tau_{1\text{-prong}})(\tau_\mu \tau_{1\text{-prong}})$ decays, the decays $a_1 a_1 \rightarrow (\mu\mu)(\tau_\mu \tau_{1\text{-prong}})$ are also included as they can give rise to a similar topology. The two decay modes are combined to search for the $H \rightarrow a_1 a_1$ signal, assuming that a_1 has Yukawa-like couplings to fermions, with coupling strengths proportional to the respective fermion masses. Because of its significantly lower branching fraction, the $a_1 a_1 \rightarrow 4\mu$ decay mode has a negligible contribution to the search sensitivity and is not considered in the present analysis.

The analysis primarily targets the dominant gluon-gluon fusion (ggF) production mechanism, where the Higgs boson is produced with relatively small p_T and the a_1 bosons are emitted nearly back-to-back in the transverse plane, resulting in a large separation in azimuthal angle ($\Delta\phi$) between the decay products of the two a_1 bosons. When the Higgs boson is produced with high p_T , e.g., due to initial-state radiation in ggF, or through other production mechanisms, the azimuthal angle between the a_1 bosons is reduced, while the separation in pseudorapidity ($\Delta\eta$) can still be large. Therefore, the analysis focuses on the identification of same-sign dimuon events with significant angular separation, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, where each muon is accompanied by a nearby oppositely charged particle from the same a_1 decay. The requirement of same-sign muons significantly reduces backgrounds from top quark pairs, Drell-Yan processes, and diboson production, enhancing the sensitivity of the search.

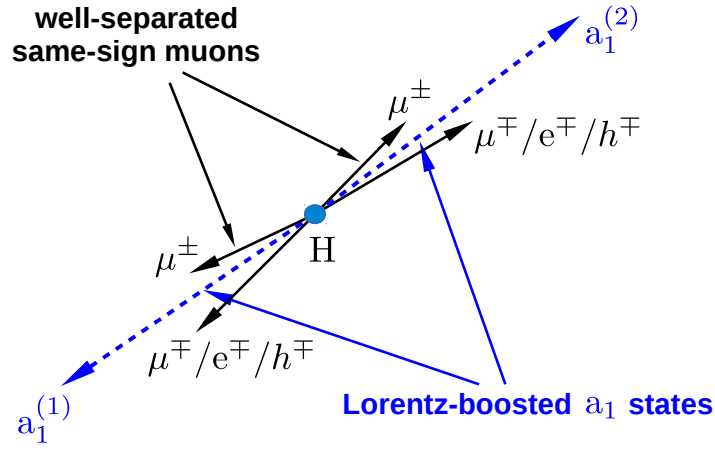


Figure 1: Illustration of the signal topology. The Higgs boson (H) decays into two a_1 bosons, where one a_1 boson decays into a pair of τ leptons, while the other decays into a pair of muons (μ) or τ leptons. In the case of $a_1 \rightarrow \tau\tau$ decays, one τ lepton is required to decay to a muon, and the other to a single charged particle—either an electron (e), a muon, or a charged hadron (h). The targeted final state thus consists of one muon and an oppositely charged track arising from each a_1 decay.

Outside the probed mass range of 4–15 GeV, the sensitivity of the present search decreases rapidly. At lower a_1 masses, the background from QCD multijet events overwhelms the signal. At higher masses, the angular separation between the decay products of the a_1 bosons increases, requiring a different analysis strategy. Tabulated results are provided in the HEP-Data record for this analysis [35].

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are reconstructed in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. More detailed descriptions of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Refs. [36, 37].

Events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of about 4 μ s [38]. The second level, known as the high-level trigger, consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, which reduces the event rate to around 1 kHz before data storage [39, 40].

3 Simulated samples

Signal events in the $H \rightarrow a_1 a_1 \rightarrow 4\tau$ channel are simulated at leading order (LO) using the PYTHIA (v.8.212) event generator [41], targeting three major Higgs boson production mechanisms: ggF, vector boson fusion (VBF), and Higgs-strahlung (VH). For the $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$ decay channel, events are generated at LO for the dominant ggF production mode with MADGRAPH5_aMC@NLO (v.2.6.5) [42]. The acceptance of the $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$ signal is corrected to account for subdominant contributions from the VBF and VH processes as detailed in Section 5. Simulation studies have demonstrated that the signal acceptance is highly sensitive to the p_T spectrum of the Higgs boson. Therefore, signal samples generated at LO are reweighted to match the Higgs boson p_T spectrum predicted by theoretical computations incorporating higher-order corrections. The p_T distribution of the Higgs boson produced via ggF is reweighted for both decay channels (4τ and $2\mu 2\tau$) using K -factors calculated with the program HQT (v.2.0) [43, 44], which computes the p_T spectrum of the Higgs boson at next-to-leading order (NLO) with the next-to-next-to-leading-logarithmic resummation at low p_T . For VBF and VH processes, the p_T distribution of the Higgs boson is reweighted using K -factors calculated with the NLO POWHEGBOX (v.2.0) event generator [45–47]. All K -factors are computed using the next-to-NLO (NNLO) NNPDF3.1 parton distribution functions (PDF) [48].

The major source of background for the analysis is QCD multijet production, followed by subdominant contributions from Drell–Yan processes with Z or W boson production accompanied by jets (Z+jets and W+jets), top quark-antiquark pair production with additional jets ($t\bar{t}$ +jets), single top quark production, and vector boson pair production (diboson). The QCD multijet and diboson backgrounds are simulated using PYTHIA (v.8.212), whereas the Z+jets and W+jets backgrounds are generated at LO with MADGRAPH5_aMC@NLO (v.2.6.5). The single top quark and $t\bar{t}$ +jets processes are generated at NLO using POWHEGBOX (v.2.0).

Parton showering and fragmentation for all Monte Carlo (MC) samples are executed using PYTHIA (v.8.212). The CP5 tune is applied to describe the underlying event [49]. All simulated samples utilize the NNLO NNPDF3.1 PDFs. For samples produced at LO with MADGRAPH5_aMC@NLO, the MLM jet matching scheme is employed [50]. The detector response is simulated with the GEANT4 package [51, 52]. The contribution of pileup is replicated by simulation of minimum bias interactions, which are then superimposed on the primary hard-scattering event. The simulated events are then reweighted to reflect the observed pileup distribution in the experimental data. The average number of pileup collisions was 27 (38) in the 2016 (2017–2018) data.

4 Event reconstruction

Events containing two muon candidates and two additional tracks are selected in the analysis. The track reconstruction is performed using the combinatorial track finder algorithm [53], which employs Kalman filtering to refine the track estimates; the analysis selects the “high-purity” tracks defined in Ref. [53]. The primary vertex (PV) is identified as the vertex corresponding to the hardest scattering in the event, evaluated using tracking information alone, as described in Section 9.4.1 of Ref. [54].

The particle-flow (PF) algorithm [55] reconstructs and identifies individual particles (PF candidates) in an event using an optimized combination of information from the various elements of the CMS detector. Muons are identified as tracks in the central tracker consistent with either a track or several hits in the muon system and associated with calorimeter deposits compatible with the expectation for minimum-ionizing particles. These PF muons are further required

to pass dedicated identification requirements, which depend on several parameters, including track quality, impact parameter significance, and number of muon chamber hits, to suppress contributions from nonprompt decays of hadrons into muons and their punchthrough to the muon detectors. The analysis employs the medium identification criterion [56], which yields an overall efficiency between 98 and 99% for muons with $p_T > 20$ GeV. Scale factors, derived by comparing the muon identification efficiencies between data and simulation, are applied to the MC samples to match the performance observed in data.

Jets are reconstructed by clustering PF candidates using the anti- k_T algorithm [57, 58] with a distance parameter of 0.4. Charged particles not associated with the PV are excluded from the clustering using the charged hadron subtraction method [55]. The reconstructed jet energies are corrected in both data and simulation to account for effects from nonlinear detector response and contamination from pileup [59]. The jet energy resolution is corrected in simulation to improve the agreement with data.

Jets resulting from the fragmentation and hadronization of bottom quarks (b jets) are identified using the DEEPJET algorithm [60, 61]. The tight working point of the b tagging discriminator is used, which corresponds to a rate of 0.1% for misidentifying a light jet, i.e., a gluon or light-quark jet, as a b jet [62]. The corresponding b tagging efficiency, measured from $t\bar{t}$ +jets events, is around 65%.

5 Event selection

The search, targeting nonisolated same-sign muon pairs, requires a specialized triggering strategy. Events of interest are recorded using a set of dimuon triggers. In 2016 and 2017, the trigger required the highest- p_T (leading) and second-highest- p_T (subleading) muons to have transverse momenta of at least 17 and 8 GeV, respectively. In 2018, these p_T thresholds were raised to 18 and 9 GeV, respectively. Additionally, in 2016 and 2018, the trigger required the two muons to have the same sign. In 2016, it was also required that the two muon tracks have their points of closest approach to the beam axis within 0.2 cm of each other along the z direction. The triggers employed in 2016 and 2018 did not impose any isolation requirements on the muons. However, in 2017, a nonisolated same-sign dimuon trigger was not available. Thus, a dimuon trigger imposing a loose isolation criterion on both muons was used instead. This criterion required the ratio of the p_T sum of charged hadrons, within an isolation cone of size $\Delta R = 0.3$ around the muon, to the p_T of the muon to be less than 0.4. Even with this requirement, the trigger was efficient in selecting the signal with a relative loss of efficiency, estimated from simulation studies, ranging from 5% at $m_{a_1} = 15$ GeV to 20% at $m_{a_1} = 4$ GeV.

As the first step in the offline selection, a b jet veto is imposed, rejecting events that contain one or more b-tagged jets with $p_T > 20$ GeV and $|\eta| < 2.4$. This veto is essential to suppress background from QCD multijet events with heavy-flavor hadrons that decay into muons. The remaining events that pass the trigger selection criteria are required to contain at least two same-sign muons, each having $|\eta| < 2.4$. The p_T of the leading and subleading muon is required to be greater than 19 and 10 GeV, respectively. The transverse (longitudinal) impact parameter of the muons with respect to the PV is required to be $|d_0| < 0.05$ ($|d_z| < 0.1$) cm. The angular separation between the muons must be $\Delta R > 1.5$. If more than one pair of same-sign muons meets these criteria, the pair with the highest scalar sum of p_T is selected.

The next selection step employs information from the tracks associated with the reconstructed charged PF objects, excluding the pair of same-sign muons. The information is used to identify and isolate the candidates for the $a_1 \rightarrow \tau_\mu \tau_{1\text{-prong}}$ or $a_1 \rightarrow \mu\mu$ decays, hereafter referred to

as a_1 candidates. Two types of tracks are considered, namely “isolation” and “signal” tracks. The signal tracks are a subset of the isolation tracks, subject to stricter selection criteria. The purpose of each track type and its selection criteria are given in Table 1.

Table 1: The purpose and selection criteria for the two types of tracks used in the selection procedure. The variables ($|d_0|$) and ($|d_z|$) are defined in the text.

Type of track	p_T	$ \eta $	$ d_0 $	$ d_z $	Purpose
Isolation	$> 1.0 \text{ GeV}$	< 2.4	$< 0.2 \text{ cm}$	$< 0.3 \text{ cm}$	Isolation criterion for a_1 candidates
Signal	$> 2.5 \text{ GeV}$	< 2.4	$< 0.02 \text{ cm}$	$< 0.04 \text{ cm}$	Build a_1 candidates

Each selected muon of the same-sign muon pair is required to have exactly one isolation track within a ΔR cone of 0.5 around the muon. The background components, especially QCD multijet events, tend to have higher track multiplicity with respect to the signal and, hence, are rejected by imposing such an isolation requirement. A muon-track system is accepted as an a_1 candidate if the isolation track around each muon meets the signal track criteria. The event is selected in the final sample if it contains two a_1 candidates. The set of selection requirements outlined above defines the signal region (SR).

The number of observed events selected in the SR is 7803. Simulation studies indicate that QCD multijet events are the predominant background in the SR. Contributions from other backgrounds only comprise about 1% of the selected events in the SR. In Table 2, the expected signal acceptance times selection efficiency ($\mathcal{A}\epsilon$) and the expected yield of signal events in the SR from simulation are reported for a few representative values of m_{a_1} . For each of the two considered decay channels, namely 4τ and $2\mu 2\tau$, the value of $\mathcal{A}\epsilon$ is evaluated as the expected yield of signal events from simulation relative to the total number of signal events expected in the respective decay channel. The expected signal yields are calculated assuming a benchmark value of the branching fraction, $\mathcal{B}(H \rightarrow a_1 a_1) \mathcal{B}^2(a_1 \rightarrow \tau\tau) = 0.05$, and the SM predictions for the Higgs boson production cross sections: 48.58 pb for ggF, 3.78 pb for VBF, and 2.26 pb for VH [63]. The contributions from the ggF, VBF, and VH processes are summed to determine the 4τ yield. The expected $2\mu 2\tau$ signal yield is estimated under the assumption that the a_1 has Yukawa-like couplings to fermions, with coupling strengths proportional to the fermion masses. Under this assumption, the ratio of the $a_1 \rightarrow \mu\mu$ and $a_1 \rightarrow \tau\tau$ partial widths satisfies the relation [64]

$$\frac{\Gamma(a_1 \rightarrow \mu\mu)}{\Gamma(a_1 \rightarrow \tau\tau)} = \frac{m_\mu^2}{m_\tau^2} \frac{\sqrt{1 - (2m_\mu/m_{a_1})^2}}{\sqrt{1 - (2m_\tau/m_{a_1})^2}}. \quad (1)$$

Using Eq. 1, the ratio of the branching fractions for $a_1 a_1 \rightarrow 2\mu 2\tau$ and $a_1 a_1 \rightarrow 4\tau$ decays is computed as

$$\frac{\mathcal{B}(a_1 a_1 \rightarrow 2\mu 2\tau)}{\mathcal{B}(a_1 a_1 \rightarrow 4\tau)} = 2 \frac{\mathcal{B}(a_1 \rightarrow \mu\mu)}{\mathcal{B}(a_1 \rightarrow \tau\tau)} = 2 \frac{\Gamma(a_1 \rightarrow \mu\mu)}{\Gamma(a_1 \rightarrow \tau\tau)}. \quad (2)$$

The factor of 2 in Eq. 2 accounts for the possibility of either a_1 decaying into $\mu\mu$ or $\tau\tau$. The value of the ratio in Eq. 2 decreases from 0.0155 at $m_{a_1} = 4 \text{ GeV}$ to 0.0073 at $m_{a_1} = 15 \text{ GeV}$. To

account for subdominant contributions to the $2\mu 2\tau$ final state from VBF and VH production, the analysis assumes that the ratios of $\mathcal{A}\epsilon$ between the $2\mu 2\tau$ and 4τ channels are identical across all production mechanisms. The value of $\mathcal{A}\epsilon$ of the ggF process is then scaled accordingly to estimate the total $2\mu 2\tau$ yield. This assumption has been verified with simulated signal samples for two representative mass hypotheses of $m_{a_1} = 5$ and 10 GeV. For both masses, the values of $\mathcal{A}\epsilon$ for the two decay channels are found to be consistent within 3–5%.

Table 2: Signal acceptance times selection efficiency ($\mathcal{A}\epsilon$), defined in the text, and the number of expected signal events after selection in the SR, computed using simulated signal samples for representative mass hypotheses. The Higgs boson cross sections from ggF, VBF, and VH production mechanisms are set to the SM predictions [63]. The number of expected signal events is computed for a benchmark value of the branching fraction $\mathcal{B}(\text{H} \rightarrow a_1 a_1) \mathcal{B}^2(a_1 \rightarrow \tau\tau) = 0.05$. The quoted uncertainties in the predictions from simulation include only the statistical component. The last row reports the total event yield in data.

Signal m_{a_1} [GeV]	$\mathcal{A}\epsilon$ [10^{-4}]		Number of expected signal events	
	4τ	$2\mu 2\tau$	4τ	$2\mu 2\tau$
5	3.52 ± 0.10	103 ± 2	134 ± 4	39.7 ± 0.4
8	2.55 ± 0.09	76.0 ± 1.0	97.2 ± 3.3	23.0 ± 0.3
12	1.37 ± 0.06	35.6 ± 0.7	52.1 ± 2.4	10.1 ± 0.2
15	0.32 ± 0.03	7.5 ± 0.3	12.3 ± 1.1	2.1 ± 0.1
Data	7803			

6 Final discriminant

A two-dimensional (2D) distribution of the invariant masses of the muon-track systems constituting the a_1 candidates is chosen as the final discriminant and employed to distinguish between signal and background. This 2D distribution is populated with pairs of muon-track invariant masses (m_1, m_2) , ordered by their value, $m_2 > m_1$. Figure 2 illustrates the binning of the 2D distribution used. Each bin is labeled (i, j) , where i and j denote the bin indices along the m_1 and m_2 axes, respectively. Since m_2 is required to be greater than m_1 , only the bins (i, j) where $j \geq i$ are filled, resulting in a total of $6(6+1)/2 = 21$ independent bins. The bin boundaries along each axis are at 0, 1, 2, 3, 4, and 5.2 GeV. Bins $(i, 6)$ with $i = 1, \dots, 5$ include all events with $m_2 > 5.2$ GeV, while bin $(6, 6)$ contains all events with m_1 and $m_2 > 5.2$ GeV. The binning is optimized to achieve the highest sensitivity to the signal for all tested mass hypotheses of a_1 while ensuring reliable and robust estimation of the background, as detailed in Section 7. With the chosen binning, the statistical uncertainty in the background estimate is kept between 1 and 40% across all bins of the (m_1, m_2) distribution.

7 Background modeling

As mentioned in Section 5, QCD multijets constitute the dominant source of background, with small contributions coming from processes such as $t\bar{t}$, Z+jets, W+jets, and diboson production. To model the shape of the (m_1, m_2) distribution of the background in the SR, a binned template is constructed as:

$$f_{2D}(i, j) = C(i, j) [f_{1D}(i) f_{1D}(j)]^{\text{sym}}, \quad (3)$$

where

- $f_{2D}(i, j)$ represents the content of bin (i, j) in the (m_1, m_2) distribution, normalized to unity.
- $f_{1D}(i)$ is the content of bin i in the one-dimensional (1D) muon-track invariant mass distribution, normalized to unity. Each event contributes two entries to this distribution.
- $C(i, j)$ is a symmetric matrix, accounting for possible correlation between m_1 and m_2 . The condition $C(i, j) = 1$ for all bins (i, j) would indicate an absence of correlation between m_1 and m_2 . The elements of the matrix $C(i, j)$ are referred to as “correlation factors” henceforth.

Equation 3 includes a symmetrization operation, denoted by ‘sym,’ which is applied to the product of the one-dimensional distributions $f_{1D}(i)$ and $f_{1D}(j)$ and defined as follows:

$$[f_{1D}(i) f_{1D}(j)]^{\text{sym}} = \begin{cases} 2f_{1D}(i)f_{1D}(j), & \text{if } i \neq j, \\ f_{1D}(i)f_{1D}(i), & \text{if } i = j. \end{cases} \quad (4)$$

This symmetrization ensures that the contributions from the nondiagonal bins (i, j) and (j, i) in the Cartesian product $f_{1D}(i)f_{1D}(j)$ are correctly accounted for in the 2D (m_1, m_2) distribution, given the ordering imposed on the muon-track invariant masses. The normalization of the background is left unconstrained prior to the extraction of the signal.

In order to derive and validate the modeling of $f_{1D}(i)$ and $C(i, j)$, multiple control regions (CRs), disjoint from the SR, are defined based on variations in the isolation criteria applied to one or both muon-track pairs. The isolation criteria are defined by the number of tracks within

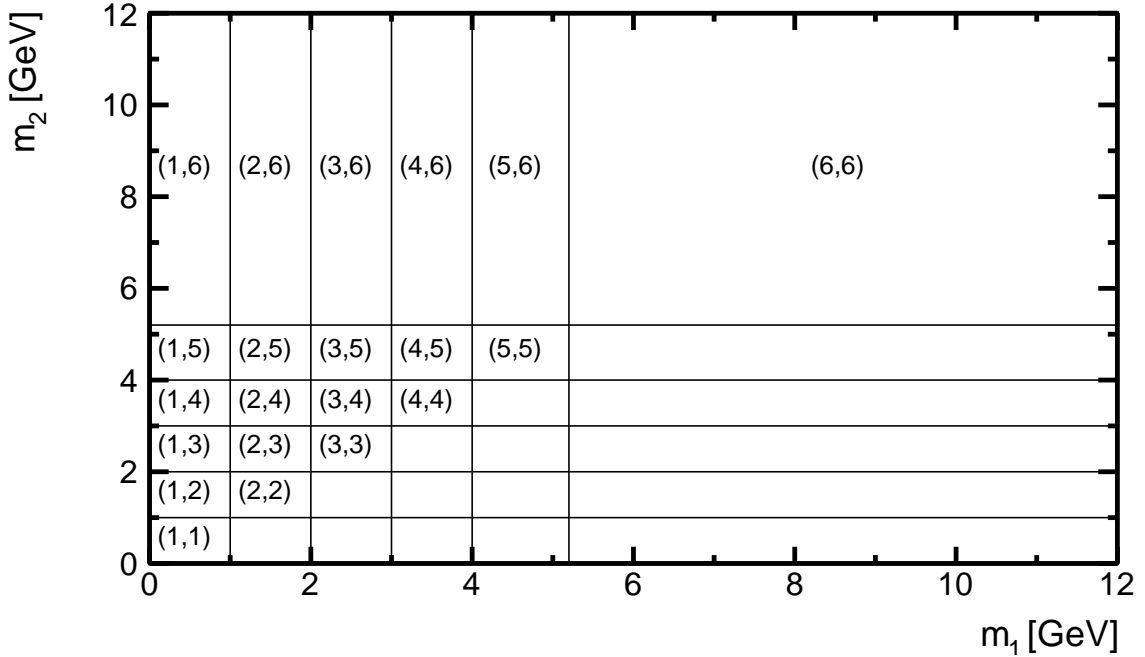


Figure 2: Binning of the 2D (m_1, m_2) distribution. Each bin is labeled (i, j) , where i is the bin index along m_1 (x axis) and j is the bin index along m_2 (y axis).

a cone of $\Delta R = 0.5$ around the muon momentum direction. The two muons are labeled as the “first muon” and “second muon”. The first muon serves as a “probe” used to assess $f_{1D}(i)$ as a function of the isolation criteria of the second muon (the “tag”), as described in Section 7.1 below. In the so-called loose-iso CR, employed to derive $C(i, j)$, both muons play the role of probes. A summary of the CRs and SR, along with the specifications for the first and second muon, is presented in Table 3. In all CRs, the expected signal contamination constitutes 0.1–1% (0.5–6%) of the background yield across all bins of the $f_{1D}(i)$ ($f_{2D}(i, j)$) distributions for all tested mass hypotheses of a_1 , and thus has negligible impact on the background modeling and final result.

Table 3: Definition of the CRs used to construct and validate the background model. The last row defines the SR. The symbols N_{iso} and N_{sig} denote the number of isolation and signal tracks, respectively, within a cone of $\Delta R = 0.5$ around the muon momentum direction. In cases where N_{sig} is not mentioned, there is no explicit requirement on the number of signal tracks.

Region	First μ	Second μ	Purpose
N_{23}	$N_{\text{sig}} = 1, N_{\text{iso}} = 1$	$N_{\text{iso}} = 2, 3$	Determination of $f_{1D}(i)$
$N_{\text{iso},2} = 1$ $N_{\text{iso},2} = 2, 3$	$N_{\text{sig}} \geq 1, N_{\text{iso}} > 1$ $N_{\text{sig}} \geq 1, N_{\text{iso}} > 1$	$N_{\text{sig}} = 1, N_{\text{iso}} = 1$ $N_{\text{iso}} = 2, 3$	Validation and systematic uncertainty estimate of $f_{1D}(i)$
Loose-Iso	$N_{\text{sig}} = 1, N_{\text{iso}} = 3, 4$	$N_{\text{sig}} = 1, N_{\text{iso}} = 3, 4$	Determination of $C(i, j)$
Signal region	$N_{\text{sig}} = 1, N_{\text{iso}} = 1$	$N_{\text{sig}} = 1, N_{\text{iso}} = 1$	Signal extraction

7.1 Modeling of $f_{1D}(i)$

The $f_{1D}(i)$ distribution is modeled using the N_{23} CR. This CR comprises events that pass the same-sign dimuon selection criteria and include only one a_1 candidate, formed by an isolated signal track and a muon (referred to as the first muon). The invariant mass of this first muon and its associated track is used to construct the $f_{1D}(i)$ distribution. The second muon in the event must be accompanied by either two or three nearby isolation tracks. Simulations indicate that the N_{23} CR is enriched in QCD events, with less than 5% of the events coming from non-QCD backgrounds. The modeling of $f_{1D}(i)$ is based on the assumption that the kinematic distributions of the muon-track system forming the a_1 candidate are weakly affected by the isolation criteria applied to the second muon, therefore implying that the $f_{1D}(i)$ distribution in the SR is similar to that in the N_{23} CR.

A direct test of this assumption is not conclusive because of the limited size of the simulated samples. Therefore, the hypothesis is verified using additional CRs labeled $N_{\text{iso},2} = 1$ and $N_{\text{iso},2} = 2, 3$. Events are selected in these CRs if the first muon has more than one isolation track ($N_{\text{iso}} > 1$), with at least one of these isolation tracks also fulfilling the criteria for signal tracks. As more than one of these tracks can qualify the requirements to be a signal track, two scenarios are evaluated, using either the signal track with the lowest p_T (“softest”) or the highest p_T (“hardest”) to compute the muon-track invariant mass. If only one signal track is found around the first muon, it serves as both the “softest” and “hardest” track. For the second muon, two isolation scenarios are considered: one where the second muon has only one signal track, making the muon-track system isolated as in the SR (CR $N_{\text{iso},2} = 1$) and another where

it has two or three isolation tracks as in the N_{23} region ($\text{CR } N_{\text{iso},2} = 2,3$). The invariant mass distributions of the first muon and its softest or hardest accompanying track are then compared between the two different isolation scenarios of the second muon. The results of this study are illustrated in Fig. 3. In both cases, the invariant mass distributions differ in each bin by less than 7%, suggesting that the invariant mass of the muon-track system forming an a_1 candidate is not highly sensitive to the isolation requirement on the second muon. To address any systematic effects on the modeling of the $f_{1D}(i)$ distribution in the N_{23} region, the observed differences are treated as a shape uncertainty in the normalized $f_{1D}(i)$ template.

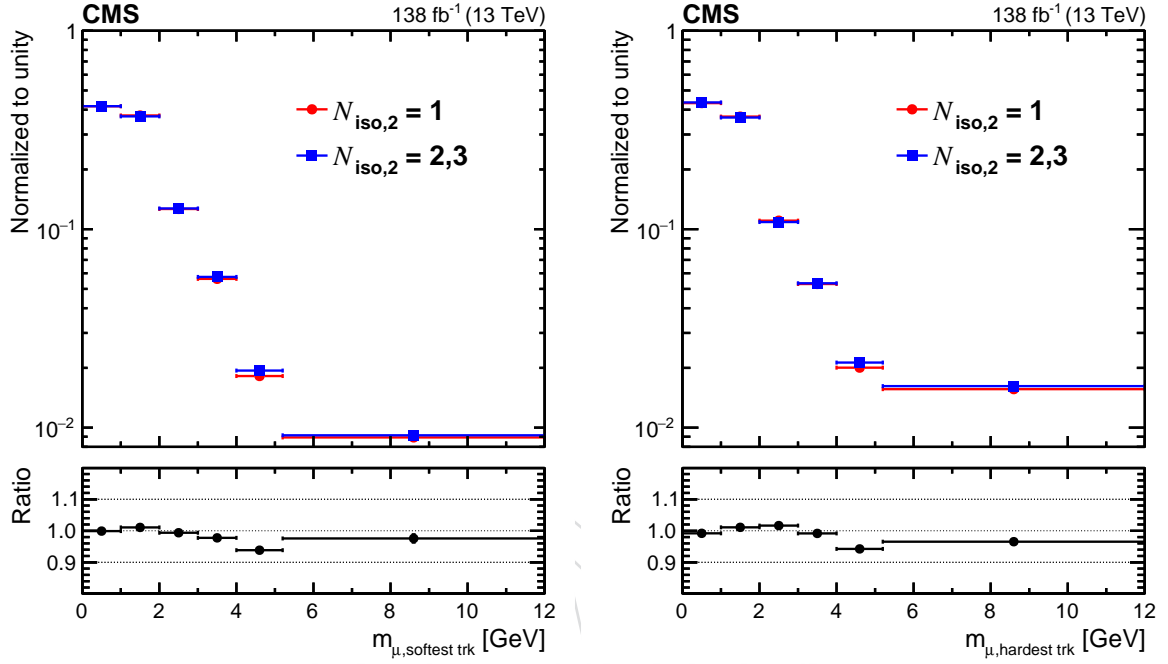


Figure 3: The observed invariant mass distribution, normalized to unity, of the first muon and the softest (left) or hardest (right) accompanying signal track for different isolation requirements imposed on the second muon: one isolation track ($N_{\text{iso},2} = 1$; circles) or two to three isolation tracks ($N_{\text{iso},2} = 2,3$). The lower panels show the ratio of the distribution in the $N_{\text{iso},2} = 2,3$ region to that in the $N_{\text{iso},2} = 1$ region.

Figure 4 presents the unit-normalized invariant mass distribution of the muon-track systems for data selected in the SR and for the background model derived from the N_{23} CR. The data and background distributions are compared to the signal distributions obtained from simulation for four representative mass hypotheses, $m_{a_1} = 5, 8, 12$, and 15 GeV . Distributions in data and simulated samples contain two entries per event, corresponding to two selected muon-track systems. The invariant mass of the muon-track system demonstrates higher discrimination power between the background and the signal at higher m_{a_1} . For lower masses, the signal shape becomes more similar to the background, thereby reducing the discrimination power.

7.2 Modeling of $C(i,j)$

The correlation factors $C(i,j)$ are determined using a different CR, labeled Loose-Iso, which does not overlap with the SR. This CR consists of events containing two same-sign muons that meet the identification and kinematic selection criteria detailed in Section 5. In this CR, each muon must have three or four nearby tracks, one of which must be a signal track and the rest isolation tracks. Simulation predicts that QCD multijet events dominate in this CR, comprising

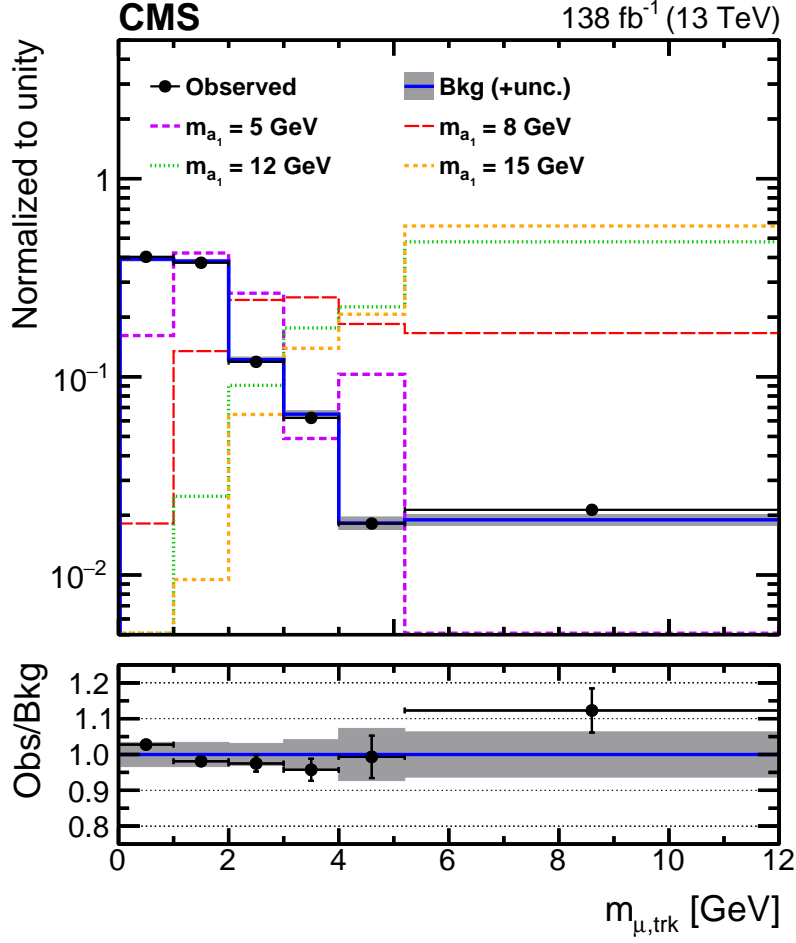


Figure 4: Normalized invariant mass distribution of the muon-track system for events passing the signal selection. Observed events are shown as black points with error bars. The background model in blue is derived from the N_{23} control region. Also shown are normalized distributions from signal simulations for four mass hypotheses, $m_{a_1} = 5, 8, 12,$ and 15 GeV (dashed histograms). Signal distributions include both the $2\mu 2\tau$ and 4τ contributions. Each event in the observed and the simulated signal distributions contributes two entries, corresponding to the two muon-track systems that pass the selection. The lower panel shows the ratio between observed and expected background events in each bin. The grey shaded area represents the uncertainty in the background model, including systematic and statistical components.

approximately 99% of the selected events. The events selected in this CR are used to construct the normalized distributions, $f_{2D}(i, j)$. The correlation factors $C(i, j)$ are then derived using Eq. 3 as:

$$C(i, j) = \frac{f_{2D}(i, j)}{(f_{1D}(i)f_{1D}(j))^{\text{sym}}}, \quad (5)$$

where $f_{1D}(i)$ is the 1D normalized distribution with two entries per event. Figure 5 shows the correlation factors $C(i, j)_{\text{data}}^{\text{CR}}$ obtained from the data in the Loose-Iso CR.

To estimate $C(i, j)$ in data in the SR, the correlation factors derived with data in the Loose-Iso CR are corrected for the difference in $C(i, j)$ between the SR and the Loose-Iso CR by comparing

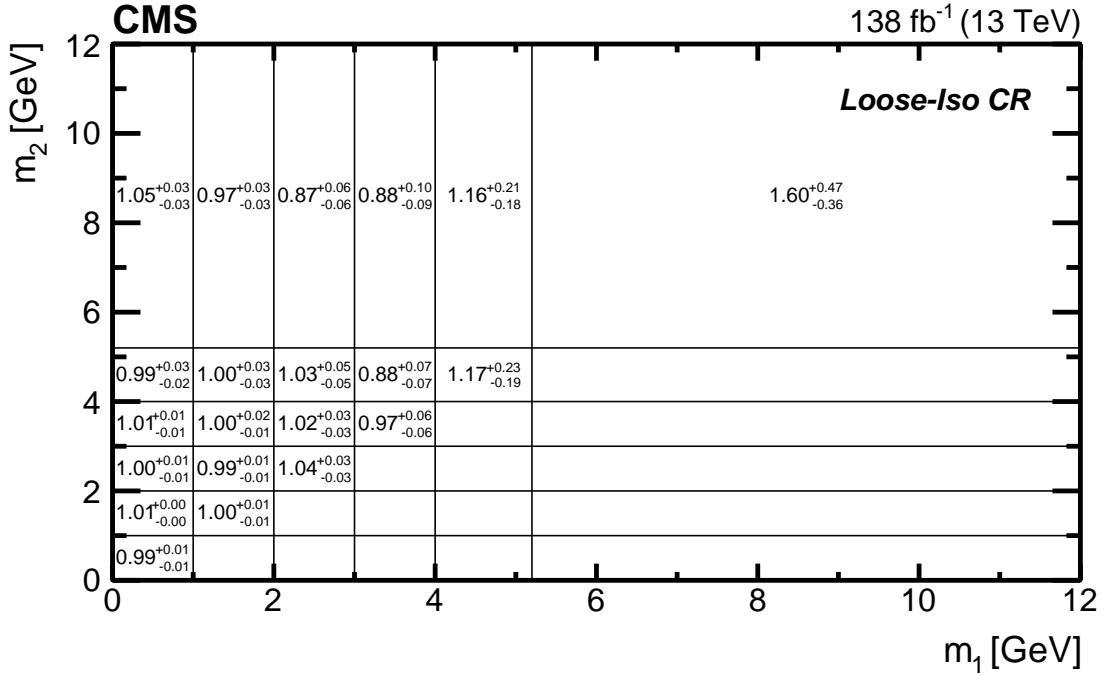


Figure 5: The correlation factors derived in the Loose-Iso CR, $C(i, j)_{\text{data}}^{\text{CR}}$, with their statistical uncertainties.

337 samples of simulated background events. The correlation factors estimated from simulation in
 338 the SR, $C(i, j)_{\text{MC}}^{\text{SR}}$, and the Loose-Iso CR, $C(i, j)_{\text{MC}}^{\text{CR}}$, are shown in Fig. 6.

339 The correlation factors in data in the SR, $C(i, j)_{\text{data}}^{\text{SR}}$, are then computed as

$$C(i, j)_{\text{data}}^{\text{SR}} = C(i, j)_{\text{data}}^{\text{CR}} \frac{C(i, j)_{\text{MC}}^{\text{SR}}}{C(i, j)_{\text{MC}}^{\text{CR}}}, \quad (6)$$

340 Finally, $C(i, j)_{\text{data}}^{\text{SR}}$ are multiplied by a common scale factor of 1.02 to ensure that $f_{2\text{D}}(i, j)$ inte-
 341 grates to unity in the SR according to Eq. 3. The introduction of this scale factor does not affect
 342 the final results, as the overall normalization of the background for each data taking period is
 343 left unconstrained prior to the signal extraction.

344 8 Signal modeling

345 The signal model is constructed using simulated samples of the $H \rightarrow a_1 a_1 \rightarrow 4\tau$ and $H \rightarrow$
 346 $a_1 a_1 \rightarrow 2\mu 2\tau$ decays. The analysis probes the signal strength modifier, defined as the product
 347 of the measured signal cross section times branching fraction, $\mathcal{B}(H \rightarrow a_1 a_1) \mathcal{B}^2(a_1 \rightarrow \tau\tau)$,
 348 relative to the value predicted by the SM for the inclusive H boson production cross section.
 349 The relative contributions from different Higgs boson production modes are determined by
 350 the corresponding cross sections predicted by the SM. The contribution from the $H \rightarrow a_1 a_1 \rightarrow$
 351 $2\mu 2\tau$ decay is computed under the assumption that the partial widths of the $a_1 \rightarrow \tau\tau$ and
 352 $a_1 \rightarrow \mu\mu$ decays satisfy Eq. 1.

353 The invariant mass distribution of the muon-track system in the $a_1 \rightarrow \mu\mu$ decay channel peaks
 354 at the nominal a_1 boson mass. In contrast, the reconstructed mass of the muon-track system
 355 in the $a_1 \rightarrow \tau\tau$ decay is generally lower and has a larger dispersion due to the presence of

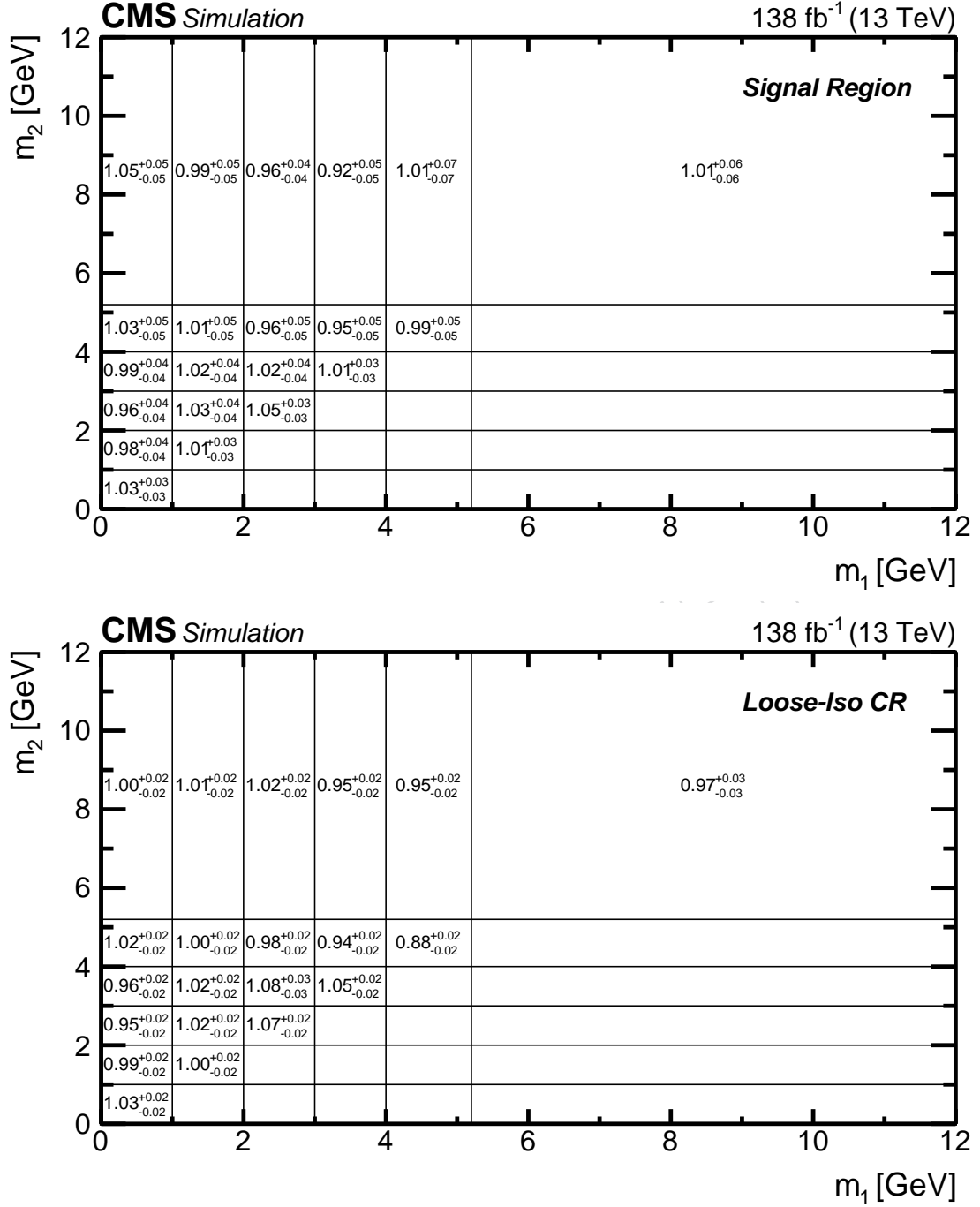


Figure 6: The correlation factors $C(i,j)_{MC}^{SR}$ (upper) and $C(i,j)_{MC}^{CR}$ (lower) with their statistical uncertainties.

undetected neutrinos. As a result, the $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$ signal samples have a considerably different shape in the (m_1, m_2) distribution compared to the $H \rightarrow a_1 a_1 \rightarrow 4\tau$ signal samples. Figure 7 illustrates the (m_1, m_2) distributions, unrolled into a one-dimensional array, for the $H \rightarrow a_1 a_1 \rightarrow 4\tau$ and $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$ signal samples for the mass hypotheses $m_{a_1} = 5, 8, 12$, and 15 GeV. These signal distributions are normalized assuming the SM H boson production rate and a branching fraction $\mathcal{B}(H \rightarrow a_1 a_1) \mathcal{B}^2(a_1 \rightarrow \tau\tau)$ of 0.05.

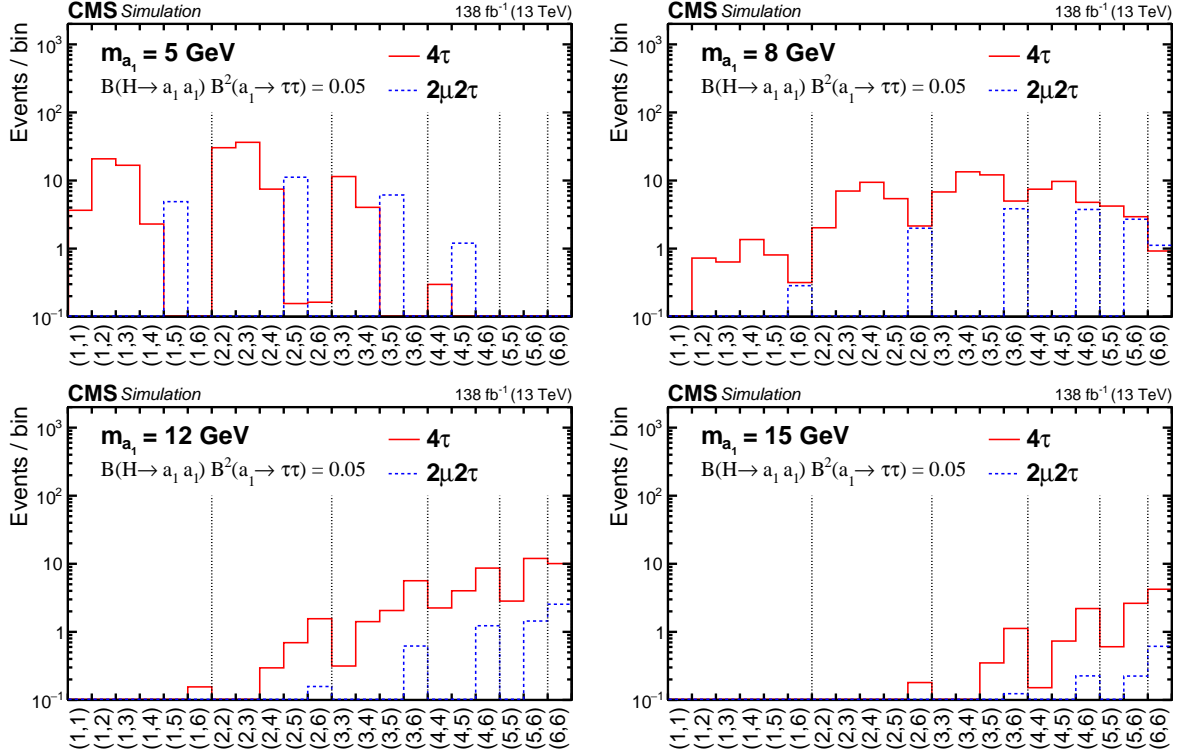


Figure 7: The signal $f_{2D}(i, j)$ templates for the mass hypotheses $m_{a_1} = 5$ GeV (upper left), 8 GeV (upper right), 12 GeV (lower left) and 15 GeV (lower right). The $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$ (blue histogram) and $H \rightarrow a_1 a_1 \rightarrow 4\tau$ (red histogram) contributions are shown. The distributions are normalized assuming SM H production cross section and $\mathcal{B}(H \rightarrow a_1 a_1) \mathcal{B}^2(a_1 \rightarrow \tau\tau) = 0.05$. The bin notation follows that of Fig. 2.

9 Systematic uncertainties

Several sources of uncertainty, of both statistical and systematic origin, are considered in the statistical analysis. A summary of the uncertainties is provided in Table 4. Statistical uncertainties arise from the limited size of the data samples in the CRs used for background modeling and simulated signal samples. They are incorporated into the analysis on a bin-by-bin basis using the Barlow–Beeston “lite” method [65]. Additionally, multiple systematic uncertainties are included, which can be classified into two categories: uncertainties related to the background, and uncertainties related to the signal.

9.1 Uncertainties related to background

The shape of the background in the 2D (m_1, m_2) distribution is modeled according to Eq. 3. This distribution is affected by the shape uncertainty in $f_{1D}(i)$, described in Section 7.1, which accounts for any potential bias introduced by estimating the 1D distribution in the N_{23} CR. The uncertainty, derived by comparing $f_{1D}(i)$ between the $N_{\text{iso},2} = 1$ and $N_{\text{iso},2} = 2, 3$ CRs, is found to vary between 2 and 7% across individual bins of $f_{1D}(i)$. These variations are propagated to $f_{2D}(i, j)$ in Eq. 3 via six nuisance parameters, one per bin of the $f_{1D}(i)$ template.

The background shape is further impacted by uncertainties related to the extrapolation of $C(i, j)$ from the Loose-Iso CR to the SR using simulated background samples. Variations in the modeling of initial- and final-state radiation (ISR and FSR) during parton showering can impact the reconstructed muon-track masses and their correlations in events where an a_1 candidate

Table 4: Summary of systematic uncertainties affecting the estimation of signal and background. The impact of shape-altering and bin-by-bin uncertainties is quoted in terms of relative variations of yields across all bins of the modeled (m_1, m_2) distributions. For the normalization (norm.) uncertainties, the impact on the overall estimated yield is reported. The last column indicates how a given uncertainty is correlated across the data-taking years. Bin-by-bin statistical uncertainties for simulated signal samples are quoted for the most populated bins containing 80% of the total yield of selected signal events.

Source	Type	Impact	Correlation
Uncertainties affecting background estimate			
Modeling of $f_{1D}(i)$ in CRs	shape	2–7%	no
ISR/FSR scales	shape	1–2%	yes
Non-QCD contribution	shape	6%	yes
Limited size of data samples in CRs	bin-by-bin	1–40%	no
Uncertainties affecting signal estimate			
Luminosity	norm.	<3%	partial
Muon ID and trigger efficiency	norm.	3%	yes
Track ID/isolation efficiency	shape	10–20%	no
Prefiring weights	norm.	<3%	no
b tagging	norm.	<0.5%	partial
Limited size of simulated samples	bin-by-bin	1–2%	no
μ_R and μ_F scales ($\mathcal{A} \times \epsilon$)	norm.	1–3%	yes
μ_R and μ_F scales (cross sections)	norm.	1–5%	yes
PDF ($\mathcal{A} \times \epsilon$)	norm.	1–2%	yes
PDF (cross sections)	norm.	1–3%	yes

is mimicked by a hadronic jet, thus leading to potential deviations in $C(i, j)$. To account for this, the ISR and FSR parton shower scales are varied independently up and down by a factor of 2 and 0.5, respectively. These variations impact the estimate of $C(i, j)$ in the Loose-Iso CR and the SR by 0.5–3% depending on (i, j) . Uncertainties in the estimate of $C(i, j)$ also arise from limited understanding of contributions from non-QCD events. The non-QCD background fraction is varied by $\pm 50\%$, and this leads to variations between 0.2 and 5%, depending on (i, j) . Shape-altering uncertainties in the estimation of $C(i, j)_{\text{data}}^{\text{SR}}$ are incorporated by varying the correlation factors $C(i, j)_{\text{MC}}^{\text{SR}}$ and $C(i, j)_{\text{MC}}^{\text{CR}}$ based on the aforementioned systematic shifts. The associated shape uncertainties are determined by comparing the correlation factors derived from simulated events in the SR and the Loose-Iso CR, after applying these systematic variations. Overall, the systematic uncertainties associated to ISR, FSR, and non-QCD contributions induce variations of up to 1, 2, and 6%, respectively, in the background yields for individual bins. Other uncertainties related to the simulation of background samples have a negligible impact on $C(i, j)$ and affect the final results only marginally.

9.2 Uncertainties related to signal

The integrated luminosities for the 2016, 2017, and 2018 data-taking years have individual uncertainties in the range 1.2–2.5% [66–68], while the overall uncertainty for the 2016–2018 period is 1.6%.

Uncertainties in muon identification and trigger efficiencies, as determined using a “tag-and-probe” method [69], are estimated to be 1.5% per muon. The efficiencies of track selection and

muon-track isolation are evaluated in a study of $Z \rightarrow \tau\tau$ events where one τ lepton decays to a muon while the other decays into an isolated track in a 1-prong decay. This track meets the same selection criteria as in the nominal analysis. In this study, scale factors, accounting for differences in efficiency between data and simulation, are derived and applied to the simulated samples. The uncertainties in scale factors affect the shape of the signal estimate and alter the overall signal yield by 10–20%.

During the 2016–2017 data-taking periods, a timing shift in the ECAL L1 trigger inputs in the forward endcap region ($|\eta| > 2.4$) caused inefficiencies by incorrectly associating events with the previous bunch crossing [38]. A correction for this effect, determined using an unbiased data sample, is applied to the simulation, accompanied by normalization-altering uncertainties ranging between 0.1 and 2.8%, depending on the mass point and sample.

The uncertainties in measuring the b tagging efficiency are applied separately to heavy-flavor and light-flavor jets in the simulated samples as described in Ref. [61]. These uncertainties are divided into components specific to the data-taking period and components correlated across periods. The b tagging uncertainties lead to variations in the yield between 0.2 and 0.5%.

Theoretical uncertainties impact the kinematic distributions of the Higgs boson, particularly its p_T spectrum, thereby affecting the signal acceptance. The uncertainty due to missing higher-order corrections in the ggF process is estimated using the HQT program by varying independently the renormalization (μ_R) and factorization (μ_F) scales by factors of 0.5 and 2. The p_T -dependent K -factors are recomputed according to these variations and applied to the simulated signal samples. The resulting effect on the signal acceptance varies between 2.5 and 3%, depending on m_{a_1} . Similarly, uncertainties in the signal acceptance for the VBF and VH processes are computed, with impacts ranging from 1 to 3%, depending on the process and m_{a_1} .

The HQT program is also used to evaluate the uncertainties arising from the choice of PDFs for the ggF process. Nominal K -factors for the p_T spectrum of the Higgs boson are computed using the NNPDF3.1 PDF set [70]. Variations within the uncertainties of the NNPDF3.1 PDFs alter the signal acceptance by approximately 1%. The impact of the PDF uncertainties on the acceptance of the VBF and VH processes is estimated in a similar way, resulting in a 2% uncertainty.

Systematic uncertainties in the theoretical predictions for the H boson production cross section are driven by variations of the μ_F and μ_R scales and PDF uncertainties. Uncertainties related to scale variations range from 1 to 5%, depending on the H boson production mode. Uncertainties related to PDF vary between 1 and 3%.

Bin-by-bin statistical uncertainties in $C(i, j)$, related to the limited size of the data sample in the Loose-Iso CR, constitute the dominant uncertainty across all probed mass hypotheses. Additionally, shape uncertainties related to the modeling of $f_{1D}(i)$ have a substantial effect. For higher a_1 boson masses, the uncertainty associated with the track selection and muon-track isolation efficiency also becomes significant.

10 Results

The signal is extracted with a binned maximum-likelihood fit applied to the 2D (m_1, m_2) distribution, using the CMS statistical analysis tool COMBINE [71], based on the ROOFIT [72] and ROOSTATS [73] frameworks. The normalizations of both the signal and background are allowed to float freely in the fit. Systematic uncertainties affecting the normalization of the signal templates are incorporated via nuisance parameters with log-normal distributions. Shape-

altering systematic uncertainties are modeled by nuisance parameters whose variations cause continuous morphing of the signal or background template shapes, as discussed in Section 4.2.1 of Ref. [71], and are assigned Gaussian prior probability density functions. For each probed m_{a_1} , the (m_1, m_2) distribution is fitted with the sum of five templates: one for the background model and four for the signal. The signal templates correspond to the ggF, VBF, and VH production modes with $H \rightarrow a_1 a_1 \rightarrow 4\tau$, and inclusive H boson production followed by $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$. The normalization of the signal templates is scaled by a common signal strength modifier. The yields of the $2\mu 2\tau$ and 4τ signal are related according to Eqs. 1 and 2. No significant excess of events over the SM background prediction is observed. The compatibility of the observed (m_1, m_2) distribution with the background-only model is quantified with a goodness-of-fit test based on the saturated model for the test statistics [74, 75], yielding a p -value of 0.45.

Figure 8 displays the distribution of (m_1, m_2) , where the notation for the bins follows that of Fig. 2. For illustrative purposes, the background distribution is normalized by fitting the observed data under the background-only hypothesis. Expectations for the signal for $m_{a_1} = 5, 8, 12$, and 15 GeV are also shown. The signal normalization is calculated assuming the SM prediction for the cross sections of the ggF, VBF, and VH processes and a branching fraction of 5% for the $H \rightarrow a_1 a_1 \rightarrow 4\tau$ decays.

The results of the analysis are used to set upper limits at 95% CL on the product of the cross section and branching fraction, $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)\mathcal{B}^2(a_1 \rightarrow \tau\tau)$, relative to the inclusive SM H boson production cross section, σ_{SM} . The CL_s criterion [76, 77] is used to set the upper limits, using the asymptotic approximation [78]. The test statistic employed in the statistical inference is the profile likelihood ratio modified for upper limits [79]. Figure 9 shows the obtained observed and expected upper limits. The observed limits range from 0.007 at $m_{a_1} = 11$ GeV to 0.079 at $m_{a_1} = 4$ GeV. The expected upper limits in the absence of signal range from 0.011 at $m_{a_1} = 11$ GeV to 0.066 at $m_{a_1} = 4$ GeV. The observed limits are compatible with the expected limits within two standard deviations in the entire range of m_{a_1} considered.

The inclusion of the $H \rightarrow a_1 a_1 \rightarrow 2\mu 2\tau$ channel in the signal model consistently improves both the expected and observed limits across all mass points, with improvements ranging from 25 to 35% in expected limits and up to 30% in observed limits. The sizable improvement of sensitivity with the inclusion of the $2\mu 2\tau$ channel is driven by two factors. Firstly, $\mathcal{A} \times \epsilon$ for the $2\mu 2\tau$ channel is higher compared to the 4τ channel because of the substantially harder p_T spectrum of muons in $a_1 \rightarrow \mu\mu$ decays compared to $a_1 \rightarrow \tau_\mu \tau_{1\text{-prong}}$ decays. Secondly, the $a_1 \rightarrow \mu\mu$ decay produces resonant structures in the distribution of m_1 , making the discrimination between signal and background in the $2\mu 2\tau$ channel more efficient than in the 4τ channel.

The degradation of the sensitivity towards lower values of m_{a_1} is caused by the increase of the background towards low invariant masses of the muon-track systems, as illustrated in Figs. 4 and 8. As m_{a_1} increases, the average angular separation between the decay products of the a_1 boson increases. Consequently, the efficiency of the signal selection decreases due to the requirement that the muon and the track from the $a_1 \rightarrow \tau_\mu \tau_{1\text{-prong}}$ or $a_1 \rightarrow \mu\mu$ decay must be within a cone of $\Delta R = 0.5$. This explains the reduced sensitivity at higher values of m_{a_1} .

The (m_1, m_2) distribution, presented in Fig. 8, shows a deficit in observed yields compared to the background-only expectations in two regions:

- bins (2,2) and (2,3), which are sensitive to the signal with $m_{a_1} = 5$ GeV;
- and bins (4,4), (4,5), and (4,6), where a significant contribution is expected for the signal with m_{a_1} in the range 7–12 GeV.

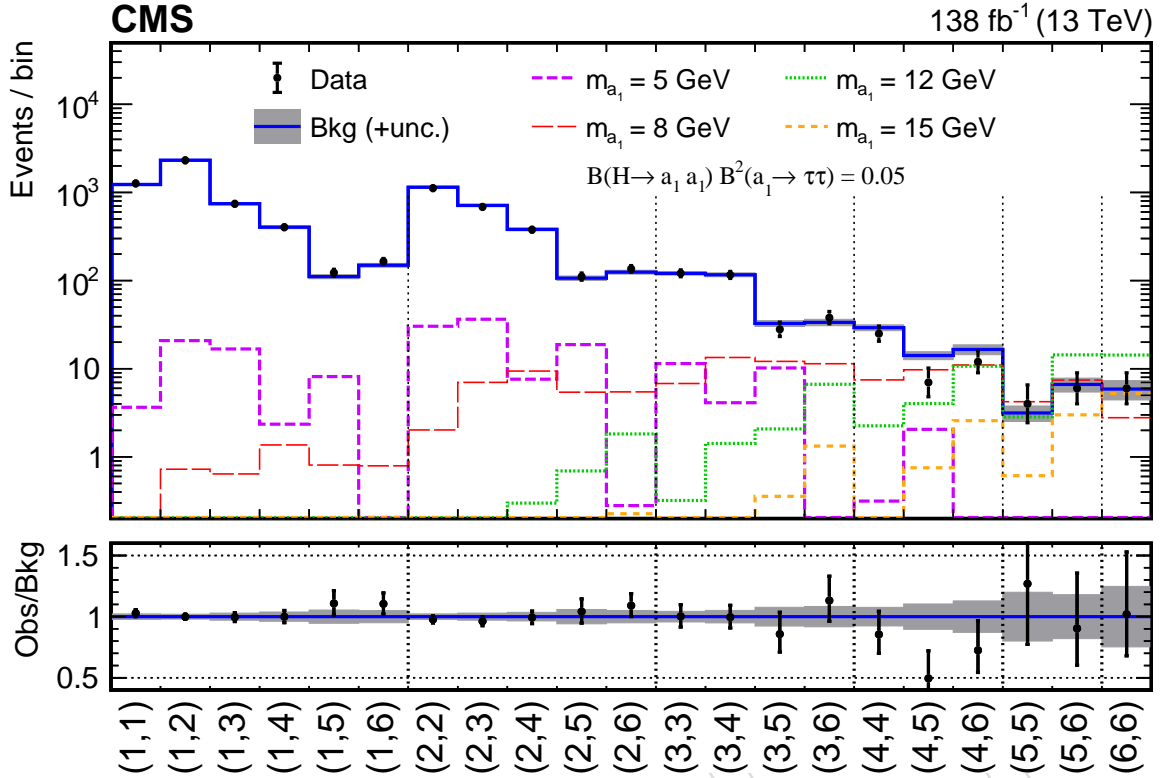


Figure 8: The unrolled (m_1, m_2) distribution used to extract the signal. The observed number of events is represented by data points with error bars. The background with its uncertainty is shown as the blue histogram with the shaded error band. The normalization for the background is obtained by fitting the observed data under the background-only hypothesis. Signal expectations for the 4τ and $2\mu 2\tau$ final states are shown as dashed histograms for the mass hypotheses $m_{a_1} = 5, 8, 12$, and 15 GeV. The relative normalization of the 4τ and $2\mu 2\tau$ final states are given by Eq. (1) as explained in Section 8. The signal normalization is computed assuming that the H boson is produced in pp collisions with a rate predicted by the SM and decays into $a_1 a_1 \rightarrow 4\tau$ final state with the branching fraction of 5%. The lower plot shows the ratio of the observed data events to the expected background yield in each bin of the (m_1, m_2) distribution.

Owing to the limited mass resolution of the final discriminant, which ranges from 1.5 GeV at low a_1 masses up to 3 GeV at high masses, these deficits result in the observation of stronger limits compared to the expectation in the absence of signal for all probed a_1 masses with the exception of $m_{a_1} = 4$ GeV. The largest discrepancy between the observed and expected limits amounts to 1.7 standard deviations, at $m_{a_1} = 8$ GeV.

The results of the search are also interpreted in the context of the 2HD+S models. The upper limits on the signal strength are translated into constraints on $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)$ by scaling them with the theoretically predicted values of $\mathcal{B}^2(a_1 \rightarrow \tau\tau)$. The branching fraction, which depends on the model type, m_{a_1} , and $\tan\beta$, is calculated using the decay width expressions from Ref. [80]. Figure 10 shows the observed and expected 95% CL upper limits on $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)$ obtained for the four types of 2HD+S models for benchmark values of $\tan\beta$, corresponding to scenarios where the $a_1 \rightarrow \tau\tau$ decay has a sizable branching fraction. Among the scenarios considered, the Type III 2HD+S model with $\tan\beta = 2$ provides the most stringent limits across all m_{a_1} values between 4 and 15 GeV. The observed upper limits range from 0.01 at $m_{a_1} = 9$ GeV to 0.36 at $m_{a_1} = 4$ GeV. For the Type II model, for $\tan\beta = 5$, tight

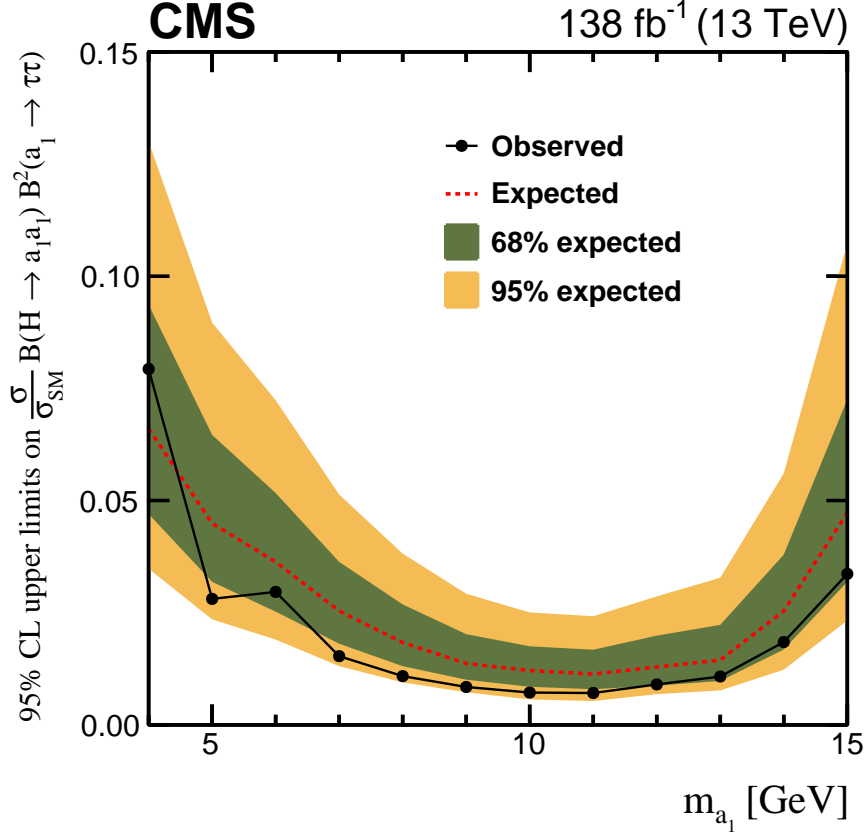


Figure 9: The observed (points) and expected (red line) 95% CL upper limits on the product of the signal cross section and the branching fraction $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)\mathcal{B}^2(a_1 \rightarrow \tau\tau)$, relative to the inclusive Higgs boson production cross section σ_{SM} predicted in the SM. The green and yellow bands indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis.

constraints between 0.013 and 0.09 are obtained for m_{a_1} up to 9 GeV. Above 9 GeV, the decay of a_1 to bottom quarks overwhelms the decay to τ leptons, making the analysis less sensitive. In the Type I 2HD+S model, a_1 couplings to fermions are independent of $\tan\beta$, so the upper limit depends only on m_{a_1} . The branching fraction to $\tau\tau$ is less enhanced since it has to compete with other channels, such as decays to charm and bottom quark pairs. Thus, the analysis provides less stringent constraints on $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)$ for Type I 2HD+S models. For the Type IV model, the analysis is only sensitive for $\tan\beta < 1$, as for higher values of $\tan\beta$, the decays to quarks dominate in the considered mass range.

The peak-like structures seen in the expected limits of Fig. 10 occur in m_{a_1} regions where quarkonium states, such as η_c and η_b , are found. In these regions, the mixing of a_1 -quarkonium states plays a crucial role, leading to a sudden increase in the hadronic decay width due to nonperturbative QCD effects. This results in a significant decrease of the branching fractions to unbound systems, such as $\tau\tau$. The mixing of a_1 with η_c (η_b) is significantly amplified in scenarios where the coupling of a_1 to charm (bottom) quarks is enhanced. Further details of the a_1 -quarkonium mixing can be found in Refs. [80, 81]. As a result of the mixing, the analysis fails to provide tight constraints in these mass regions.

Upper limits at 95% CL are also set on $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)$, relative to σ_{SM} , as

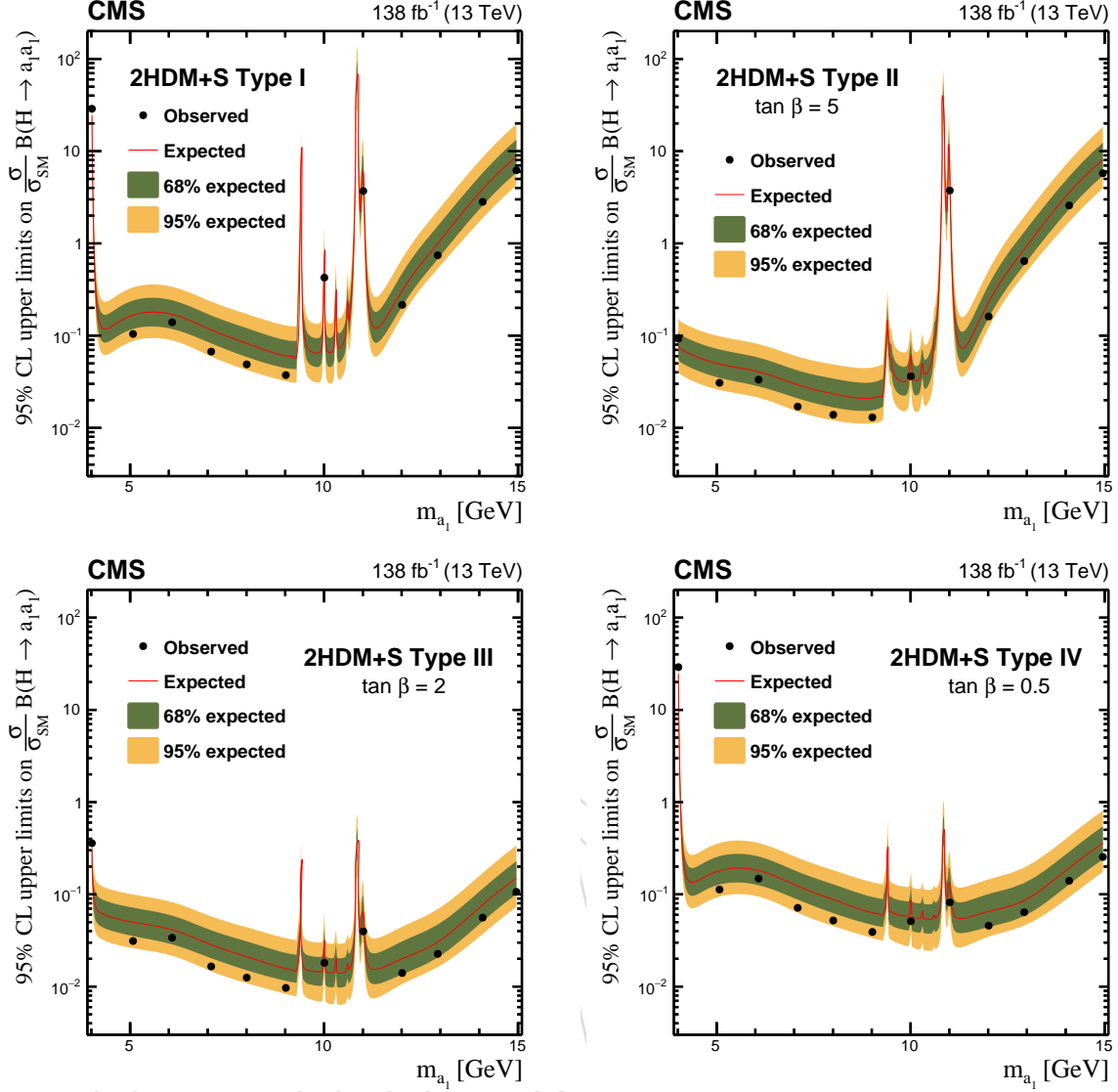


Figure 10: The observed (points) and expected (red line) 95% CL upper limits on $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)$, relative to σ_{SM} , as a function of m_{a_1} for different 2HDM+S models for benchmark $\tan\beta$ values: Type I ($\tan\beta$ independent; upper left), Type II ($\tan\beta = 5$; upper right), Type III ($\tan\beta = 2$; lower left), and Type IV ($\tan\beta = 0.5$; lower right).

a function of $\tan\beta$ for benchmark a_1 masses. Figures 11 and 12 present the limits obtained for Type II and Type III 2HDM+S models, respectively, for $m_{a_1} = 5, 8, 12$, and 15 GeV. For both models, the analysis sets stringent constraints for $\tan\beta > 1$, where the coupling to τ leptons is enhanced. For $\tan\beta < 1$, the a_1 decays to quarks dominate and suppress the $\tau\tau$ decay, resulting in weaker limits. The Type III 2HDM+S model provides the best constraints for $\tan\beta > 1$ across all considered m_{a_1} values. In the Type II model, tight constraints are obtained for $5 < m_{a_1} < 8$ GeV, but the limits deteriorate for higher masses due to the enhanced Yukawa coupling to b quarks. The observed deterioration of the limits at lower and higher a_1 masses for both models can be directly related to the trends observed in the model-independent results. The enhanced b quark coupling plays an additional role in weakening the limits for higher masses in the Type II model.

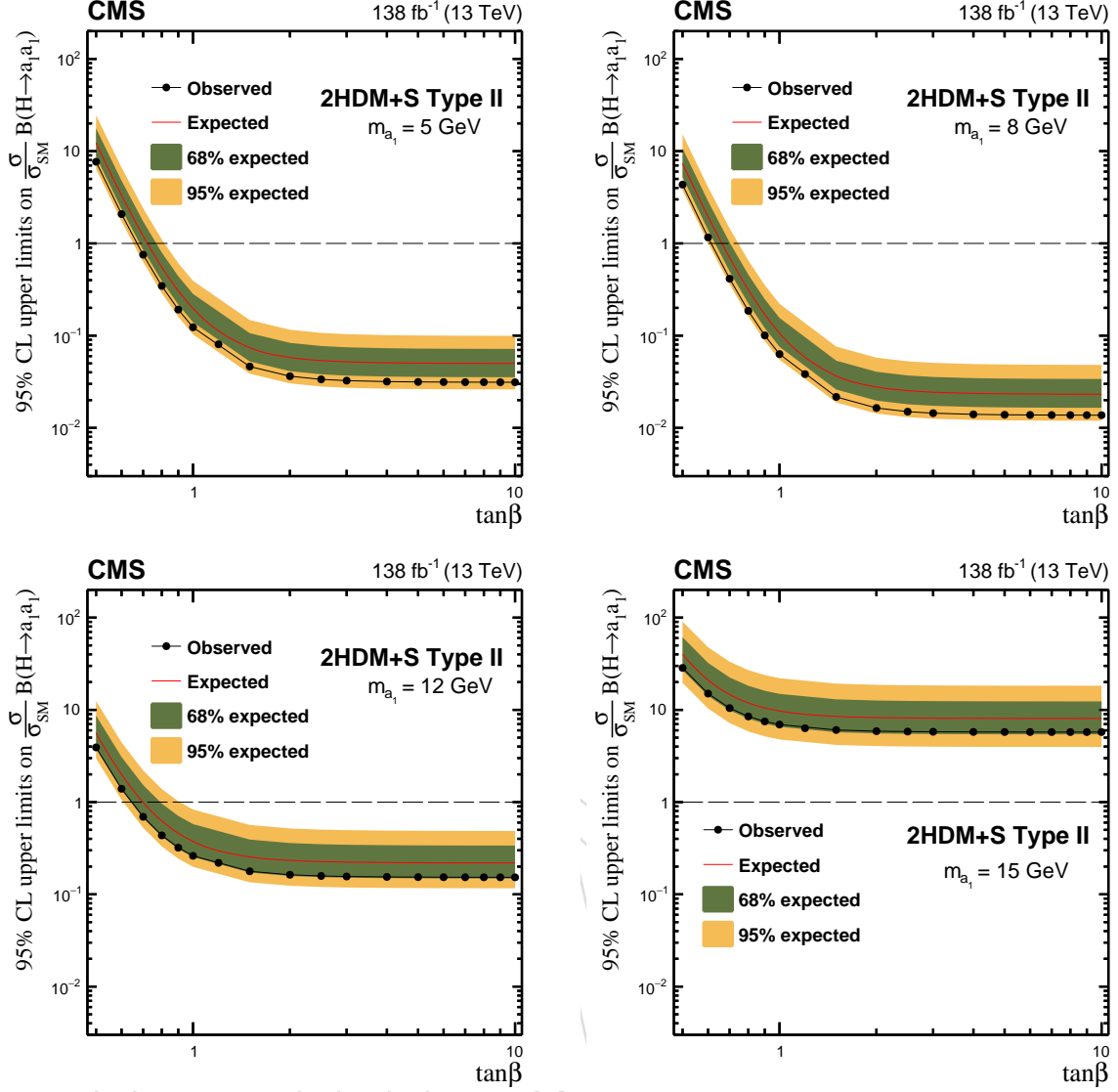


Figure 11: The observed and expected 95% CL upper limits on $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)$, relative to σ_{SM} , as a function of $\tan \beta$ for the Type II 2HDM+S model for: $m_{a_1} = 5$ GeV (upper left), $m_{a_1} = 8$ GeV (upper right), $m_{a_1} = 12$ GeV (lower left), and $m_{a_1} = 15$ GeV (lower right).

11 Summary

A search for a pair of light pseudoscalar bosons (a_1) produced in decays of the 125 GeV Higgs (H) boson, $H \rightarrow a_1 a_1$, in final states with two muons and two charged particle tracks is presented. The search is performed using data from proton-proton collisions at a center-of-mass energy of 13 TeV, collected by the CMS experiment at the LHC between 2016 and 2018, and corresponding to an integrated luminosity of 138 fb^{-1} . The analysis exploits the gluon-gluon fusion (ggF), vector boson fusion (VBF), and Higgs-strahlung (VH) production modes and targets the $H \rightarrow a_1 a_1 \rightarrow 4\tau$ and $2\mu 2\tau$ decay channels. Masses of the a_1 boson (m_{a_1}) in the range 4 to 15 GeV are examined. No excess of data above the standard model (SM) background prediction is found. Upper limits on the product of the inclusive signal cross section and the branching fraction, $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)\mathcal{B}^2(a_1 \rightarrow \tau\tau)$, relative to the SM Higgs boson production cross section σ_{SM} , are set at 95% confidence level by combining the 4τ and $2\mu 2\tau$

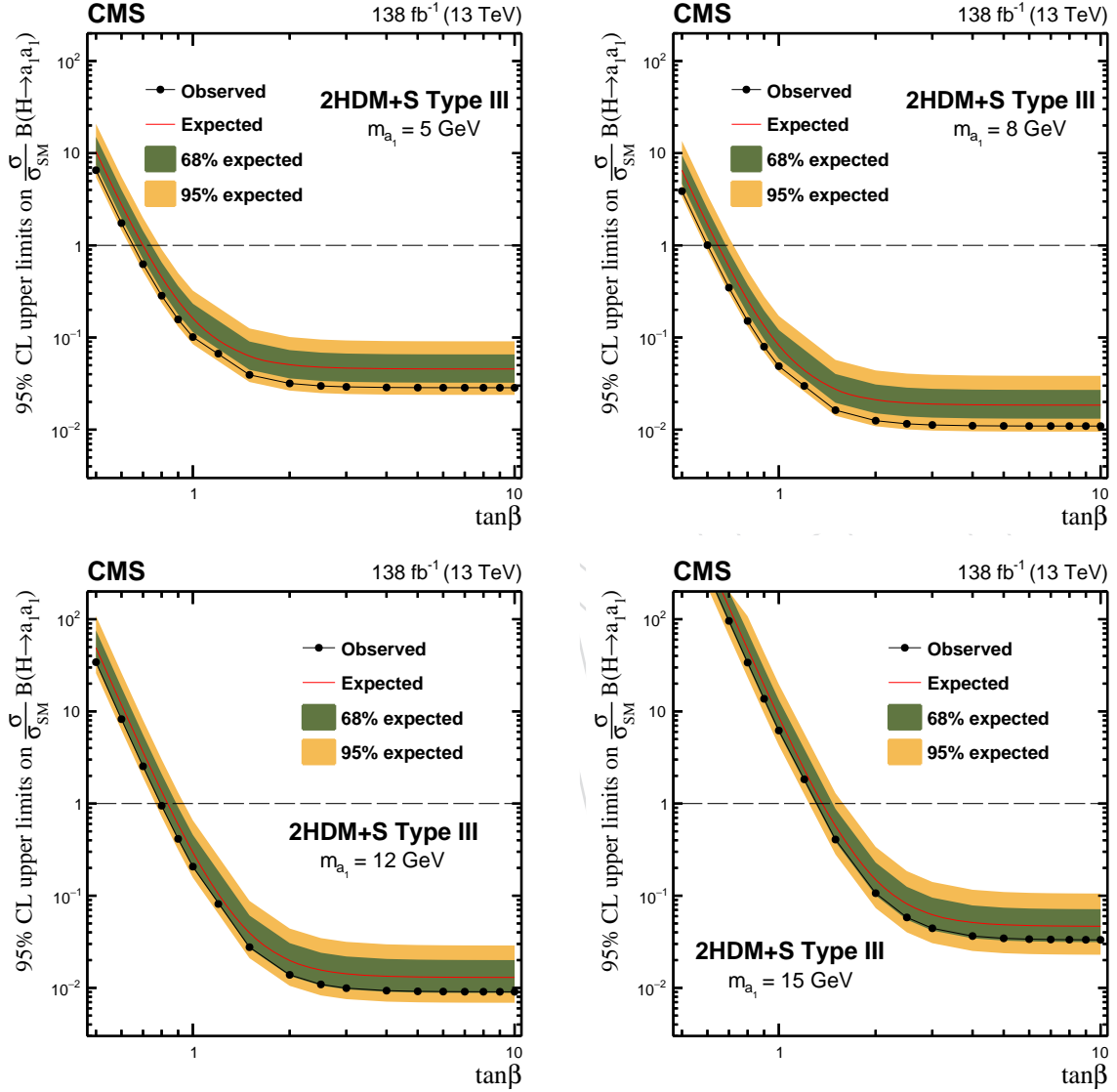


Figure 12: The observed and expected 95% CL upper limits on $\sigma(pp \rightarrow H + X)\mathcal{B}(H \rightarrow a_1 a_1)$, relative to σ_{SM} , as a function of $\tan \beta$ for the Type III 2HDM+S model for: $m_{a_1} = 5$ GeV (upper left), $m_{a_1} = 8$ GeV (upper right), $m_{a_1} = 12$ GeV (lower left), and $m_{a_1} = 15$ GeV (lower right).

decay channels and assuming Yukawa-like couplings of a_1 to fermions. The observed limits range from 0.007 at $m_{a_1} = 11$ GeV to 0.079 at $m_{a_1} = 4$ GeV. The expected limits in the absence of signal range from 0.011 at $m_{a_1} = 11$ GeV to 0.066 at $m_{a_1} = 4$ GeV. The results indicate significant improvement compared to the earlier CMS analysis at 13 TeV, exceeding the anticipated improvement from the larger data sample alone. The sensitivity is enhanced by a factor of 2 to 4, depending on the probed mass hypothesis, which can be attributed to the introduction of a veto for b-tagged jets and further optimization of the selection criteria targeting the $a_1 \rightarrow \tau\tau$ and $\mu\mu$ decays.

The results of the search are also interpreted in the context of various models with two Higgs doublets and an additional complex singlet field (2HD+S models). The tightest constraints on $\sigma(\text{pp} \rightarrow \text{H} + \text{X})\mathcal{B}(\text{H} \rightarrow a_1 a_1)$, relative to σ_{SM} are provided for the Type III 2HD+S model. For this scenario, upper limits at 95% CL between 0.01 and 0.05 are obtained for probed mass hypotheses in the ranges $5 < m_{a_1} < 8$ GeV and $12 < m_{a_1} < 14$ GeV for $\tan\beta > 2$. For the Type II 2HD+S scenario, realized in the next-to-minimal supersymmetric standard model, limits between 0.01 and 0.1 are set for m_{a_1} between 4 and 9 GeV and $\tan\beta > 5$.

DRAFT

References

- [1] ATLAS Collaboration, “Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [2] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [3] CMS Collaboration, “Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV”, *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.
- [4] ATLAS Collaboration, “A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery”, *Nature* **607** (2022) 52, doi:10.1038/s41586-022-04893-w, arXiv:2207.00092. [Erratum: doi:10.1016/0370-2693(84)91890-2].
- [5] CMS Collaboration, “A portrait of the Higgs boson by the CMS experiment ten years after the discovery”, *Nature* **607** (2022) 60, doi:10.1038/s41586-022-04892-x, arXiv:2207.00043. [Author correction: doi:10.1038/s41586-023-06164-8].
- [6] T. D. Lee, “A theory of spontaneous T violation”, *Phys. Rev. D* **8** (1973) 1226, doi:10.1103/PhysRevD.8.1226.
- [7] G. C. Branco et al., “Theory and phenomenology of two-Higgs-doublet models”, *Phys. Rept.* **516** (2012) 1, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.
- [8] D. Curtin et al., “Exotic decays of the 125 GeV Higgs boson”, *Phys. Rev. D* **90** (2014) 075004, doi:10.1103/PhysRevD.90.075004, arXiv:1312.4992.
- [9] P. Fayet, “Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino”, *Nucl. Phys. B* **90** (1975) 104, doi:10.1016/0550-3213(75)90636-7.
- [10] P. Fayet, “Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions”, *Phys. Lett. B* **69** (1977) 489, doi:10.1016/0370-2693(77)90852-8.
- [11] U. Ellwanger, C. Hugonie, and A. M. Teixeira, “The next-to-minimal supersymmetric standard model”, *Phys. Rept.* **496** (2010) 1, doi:10.1016/j.physrep.2010.07.001, arXiv:0910.1785.
- [12] M. Maniatis, “The next-to-minimal supersymmetric extension of the standard model reviewed”, *Int. J. Mod. Phys. A* **25** (2010) 3505, doi:10.1142/S0217751X10049827, arXiv:0906.0777.
- [13] J. E. Kim and H. P. Nilles, “The μ -problem and the strong CP-problem”, *Phys. Lett. B* **138** (1984) 150, doi:10.1016/0370-2693(84)91890-2.
- [14] S. Ramos-Sanchez, “The μ -problem, the NMSSM and string theory”, *Fortsch. Phys.* **58** (2010) 748, doi:10.1002/prop.201000058, arXiv:1003.1307.

- [15] ATLAS Collaboration, “Search for new light gauge bosons in Higgs boson decays to four-lepton final states in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC”, *Phys. Rev. D* **92** (2015) 092001, doi:10.1103/PhysRevD.92.092001, arXiv:1505.07645.
- [16] CMS Collaboration, “Search for a non-standard-model Higgs boson decaying to a pair of new light bosons in four-muon Final States”, *Phys. Lett. B* **726** (2013) 564, doi:10.1016/j.physletb.2013.09.009, arXiv:1210.7619.
- [17] CMS Collaboration, “A search for pair production of new light bosons decaying into muons”, *Phys. Lett. B* **752** (2016) 146, doi:10.1016/j.physletb.2015.10.067, arXiv:1506.00424.
- [18] CMS Collaboration, “Search for light bosons in decays of the 125 GeV Higgs boson in proton-proton collisions at $\sqrt{s} = 8$ TeV”, *JHEP* **10** (2017) 076, doi:10.1007/JHEP10(2017)076, arXiv:1701.02032.
- [19] CMS Collaboration, “Search for a very light NMSSM Higgs boson produced in decays of the 125 GeV scalar boson and decaying into τ leptons in pp collisions at $\sqrt{s} = 8$ TeV”, *JHEP* **01** (2016) 079, doi:10.1007/JHEP01(2016)079, arXiv:1510.06534.
- [20] ATLAS Collaboration, “Search for new phenomena in events with at least three photons collected in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”, *Eur. Phys. J. C* **76** (2016) 210, doi:10.1140/epjc/s10052-016-4034-8, arXiv:1509.05051.
- [21] ATLAS Collaboration, “Search for Higgs boson decays to beyond-the-standard-model light bosons in four-lepton events with the ATLAS detector at $\sqrt{s} = 13$ TeV”, *JHEP* **06** (2018) 166, doi:10.1007/JHEP06(2018)166, arXiv:1802.03388.
- [22] CMS Collaboration, “A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 TeV”, *Phys. Lett. B* **796** (2019) 131, doi:10.1016/j.physletb.2019.07.013, arXiv:1812.00380.
- [23] ATLAS Collaboration, “Search for Higgs bosons decaying to aa in the $\mu\mu\tau\tau$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS experiment”, *Phys. Rev. D* **92** (2015) 052002, doi:10.1103/PhysRevD.92.052002, arXiv:1505.01609.
- [24] CMS Collaboration, “Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state of two muons and two τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **11** (2018) 018, doi:10.1007/JHEP11(2018)018, arXiv:1805.04865.
- [25] CMS Collaboration, “Search for a light pseudoscalar Higgs boson in the boosted $\mu\mu\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **08** (2020) 139, doi:10.1007/JHEP08(2020)139, arXiv:2005.08694.
- [26] ATLAS Collaboration, “Search for Higgs boson decays into a pair of light bosons in the $bb\mu\mu$ final state in pp collision at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Lett. B* **790** (2019) 1, doi:10.1016/j.physletb.2018.10.073, arXiv:1807.00539.
- [27] 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** (2024) 493, doi:10.1140/epjc/s10052-024-12727-4, arXiv:2402.13358.

- [28] ATLAS Collaboration, “Search for Higgs boson decays into pairs of light (pseudo)scalar particles in the $\gamma\gamma j j$ final state in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector”, *Phys. Lett. B* **782** (2018) 750, doi:10.1016/j.physletb.2018.06.011, arXiv:1803.11145.
- [29] 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** (2020) 135087, doi:10.1016/j.physletb.2019.135087, arXiv:1907.07235.
- [30] CMS Collaboration, “Search for the decay of the Higgs boson to a pair of light pseudoscalar bosons in the final state with four bottom quarks in proton-proton collisions at $\sqrt{s}=13$ TeV”, *JHEP* **06** (2024) 097, doi:10.1007/JHEP06(2024)097, arXiv:2403.10341.
- [31] 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* **07** (2023) 148, doi:10.1007/JHEP07(2023)148, arXiv:2208.01469.
- [32] 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** (2023) 101801, doi:10.1103/PhysRevLett.131.101801, arXiv:2209.06197.
- [33] ATLAS Collaboration, “Search for Higgs boson exotic decays into Lorentz-boosted light bosons in the four- τ final state at $\sqrt{s}=13$ TeV with the ATLAS detector”, arXiv:2503.05463.
- [34] CMS Collaboration, “Identification of hadronic tau lepton decays using a deep neural network”, *JINST* **17** (2022) P07023, doi:10.1088/1748-0221/17/07/P07023, arXiv:2201.08458.
- [35] HEPData record for this analysis, 2025. doi:10.17182/hepdata.158360.
- [36] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [37] CMS Collaboration, “Development of the CMS detector for the CERN LHC Run 3”, *JINST* **19** (2024) P05064, doi:10.1088/1748-0221/19/05/P05064.
- [38] CMS Collaboration, “Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s}=13$ TeV”, *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.
- [39] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.
- [40] CMS Collaboration, “Performance of the CMS high-level trigger during LHC Run 2”, *JINST* **19** (2024) P11021, doi:10.1088/1748-0221/19/11/P11021, arXiv:2410.17038.
- [41] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

- [42] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [43] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, “Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC”, *Nucl. Phys. B* **737** (2006) 73, doi:10.1016/j.nuclphysb.2005.12.022, arXiv:hep-ph/0508068.
- [44] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, “Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC”, *JHEP* **11** (2011) 064, doi:10.1007/JHEP11(2011)064, arXiv:1109.2109.
- [45] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG”, *JHEP* **04** (2009) 002, doi:10.1088/1126-6708/2009/04/002, arXiv:0812.0578.
- [46] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [47] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with parton shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [48] NNPDF Collaboration, “Parton distributions from high-precision collider data”, *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.
- [49] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements”, *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.
- [50] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [51] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [52] J. Allison et al., “GEANT4 developments and applications”, *IEEE Trans. Nucl. Sci.* **53** (2006) 270, doi:10.1109/TNS.2006.869826.
- [53] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [54] CMS Collaboration, “Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid”, CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015.
- [55] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

- [56] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.
- [57] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_T jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [58] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- [59] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- [60] E. Bols et al., “Jet Flavour Classification Using DeepJet”, *JINST* **15** (2020) P12012, doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.
- [61] CMS Collaboration, “Performance summary of AK4 jet b tagging with data from proton-proton collisions at 13 TeV with the CMS detector”, CMS Detector Performance Summary CMS-DP-2023-005, 2023.
- [62] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [63] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector”, CERN Report CERN-2017-002-M, 2016. doi:10.23731/CYRM-2017-002, arXiv:1610.07922.
- [64] M. Lisanti and J. G. Wacker, “Discovering the Higgs boson with low mass muon pairs”, *Phys. Rev. D* **79** (2010) 115006, doi:10.1103/PhysRevD.79.115006, arXiv:0903.1377.
- [65] R. J. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples”, *Comput. Phys. Commun.* **77** (1993) 219, doi:10.1016/0010-4655(93)90005-W.
- [66] CMS Collaboration, “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS”, *Eur. Phys. J. C* **81** (2021) 800, doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.
- [67] CMS Collaboration, “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.
- [68] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.
- [69] CMS Collaboration, “Measurement of the inclusive W and Z production cross sections in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **10** (2011) 132, doi:10.1007/JHEP10(2011)132, arXiv:1107.4789.
- [70] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [71] CMS Collaboration, “The CMS statistical analysis and combination tool: Combine”, *Comput. Softw. Big Sci.* **8** (2024), no. 1, 19, doi:10.1007/s41781-024-00121-4, arXiv:2404.06614.

- [72] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling”, *eConf* **C0303241** (2003) MOLT007, arXiv:physics/0306116.
- [73] L. Moneta et al., “The RooStats Project”, *PoS* **ACAT2010** (2010) 057, doi:10.22323/1.093.0057, arXiv:1009.1003.
- [74] J. K. Lindsey, “Parametric Statistical Inference”. Oxford University Press, 05, 1996. doi:10.1093/oso/9780198523598.001.0001, ISBN 9780198523598.
- [75] R. D. Cousins, “Lectures on statistics in theory: prelude to statistics in practice”, arXiv:1807.05996.
- [76] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.
- [77] A. L. Read, “Presentation of search results: the CL_s technique”, *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.
- [78] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].
- [79] ATLAS and CMS Collaborations, and LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011”, Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.
- [80] U. Haisch, J. F. Kamenik, A. Malinauskas, and M. Spira, “Collider constraints on light pseudoscalars”, *JHEP* **03** (2018) 178, doi:10.1007/JHEP03(2018)178, arXiv:1802.02156.
- [81] M. Baumgart and A. Katz, “Implications of a new light scalar near the bottomonium regime”, *JHEP* **08** (2012) 133, doi:10.1007/JHEP08(2012)133, arXiv:1204.6032.