

# DRAFT

## CMS Paper

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

2023/05/30

Archive Hash: 504b7fb

Archive Date: 2023/05/29

### Search for long-lived particles decaying in the CMS muon detectors in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

#### Abstract

A search for long-lived particles (LLPs) decaying in the CMS muon detectors is presented. The data sample consists of  $138 \text{ fb}^{-1}$  of proton-proton collisions at  $\sqrt{s} = 13$  TeV, recorded at the LHC in 2016–2018. A unique technique is employed to reconstruct decays of LLPs in the muon detectors. The search is sensitive to LLP masses from 0.01 to 55 GeV and a broad range of LLP decay modes, including decays to a pair of quarks, kaons, pions, electrons, photons, or  $\tau$  leptons. No excess of events above the standard model background is observed. The most stringent LHC limits to date on the branching fraction of the Higgs boson to a pair of LLPs with masses below 10 GeV are set. This search also provides the most stringent limit for proper decay lengths in the range 0.04–0.4 m and above 4 m for an LLP mass of 15 GeV, in the range 0.3–0.9 m and above 3 m for an LLP mass of 40 GeV, and above 0.8 m for an LLP mass of 55 GeV. Finally, this search sets the first LHC limits on a dark quantum chromodynamic sector that couples to the Higgs boson through gluon, Higgs, photon, vector, and dark photon portals, and is sensitive to branching fractions of the Higgs boson to dark quarks as low as  $10^{-3}$ .

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

PDFAuthor:	C. Pena, C. Wang
PDFTitle:	Search for long-lived particles decaying in the CMS muon detectors in proton-proton collisions at $\sqrt{s} = 13$ TeV
PDFSubject:	CMS
PDFKeywords:	CMS, long-lived particles, muon detector

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



# 1 Introduction

Many extensions of the standard model (SM) predict the existence of neutral, weakly-coupled particles that have long proper lifetimes. These long-lived particles (LLPs) naturally arise in models of split supersymmetry (SUSY) [1–6], SUSY with weak  $R$ -parity violation [7–10], SUSY with gauge-mediated supersymmetry breaking [11–13], stealth SUSY [14, 15], hidden valley scenarios [16–18], baryogenesis triggered by weakly interacting massive particles [19–21], inelastic dark matter [22], and twin Higgs models [23–25].

In this paper, we describe a search at the CERN LHC that uses muon detectors as a sampling calorimeter to identify particle showers produced by decays of LLPs. The search is based on proton-proton (pp) collision data at 13 TeV collected during 2016–2018 at the LHC, corresponding to an integrated luminosity of  $138 \text{ fb}^{-1}$ . The CMS muon detectors are composed of gaseous detector planes interleaved with the steel layers of the magnet flux-return yoke. Decays of LLPs in the muon detectors induce hadronic and electromagnetic showers, giving rise to a large multiplicity of hits in a localized detector region. The hadron calorimeter, solenoid magnet, and steel flux-return yoke together provide 12–27 nuclear interaction lengths of shielding [26, 27], which strongly suppresses particle showers from jets that are not fully contained within the calorimeters volume (punch-through).

This search has sensitivity to singly- or multiply-produced LLPs decaying to final states including hadrons, taus, electrons, or photons. The LLPs decaying to muons will very rarely produce a particle shower and will remain undetected by this search. While this search is sensitive to many models predicting LLPs, we interpret the result in two separate benchmark models. The first is a simplified model motivated by the twin Higgs scenario [16–18, 28–30] where the SM Higgs boson ( $H$ ) decays to a pair of neutral long-lived scalars ( $S$ ), each of which decays in turn to a pair of fermions or a pair of photons. We search for long-lived scalars with masses between 0.01 and 55 GeV and a wide range of decay modes, including decays resulting primarily in hadronic showers ( $b\bar{b}$ ,  $d\bar{d}$ ,  $K^+K^-$ ,  $K^0\bar{K}^0$ , and  $\pi^+\pi^-$ ), decays resulting primarily in electromagnetic showers ( $\pi^0\pi^0$ ,  $\gamma\gamma$ , and  $e^+e^-$ ), and decays to  $\tau^+\tau^-$ , which may result in hadronic or electromagnetic showers. The most stringent previous constraints for mean proper decay lengths  $c\tau < 0.3 \text{ m}$  are based on a search for displaced jets in the CMS tracker [31]. For  $c\tau > 0.3 \text{ m}$ , a search for displaced vertices in the ATLAS muon spectrometer [32, 33] and one using the CMS endcap muon detectors [34] set the most stringent previous limits.

We also interpret the search results in terms of a set of “dark shower” models with perturbative parton showers [35]. We consider production of an SM Higgs boson that decays to a pair of dark sector quarks, each of which hadronizes into a dark shower consisting of prompt and long-lived dark sector mesons that eventually decay back to SM particles. Depending on the symmetries and decay portal, the proper lifetime of the dark mesons and the final state SM particles can vary, resulting in a wide range of dark shower signatures. We interpret the search results in a framework with five different decay portals [35], namely the gluon portal producing hadron-rich showers, photon portal with photon showers, vector portal with semi-visible jets, Higgs portal with heavy flavor-rich showers, and dark photon portal with lepton-rich showers. This is the first LHC search with an interpretation in this framework. The Feynman diagrams of the two benchmark models are shown in Fig. 1.

There are two key advantages of the LLP search strategy presented in this paper over searches that employ displaced vertices.

- (i) The absorber material in front of the muon detectors acts as a shielding to maintain a sufficiently low level of background for the detection of a single LLP decay. This level

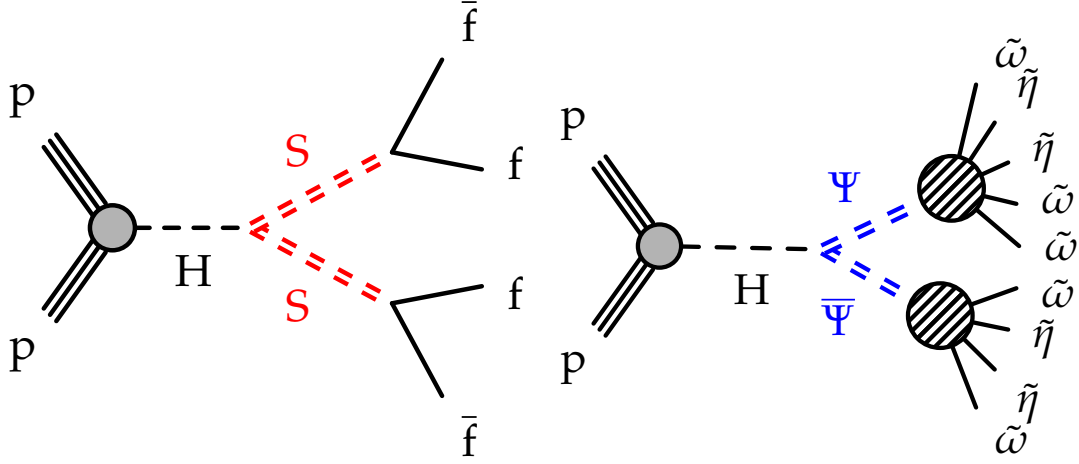


Figure 1: Feynman diagrams of twin Higgs model (left) and dark shower model (right). In the twin Higgs model, the SM Higgs boson decays to a pair of neutral long-lived scalars ( $S$ ), which then decay to two SM particles. Only fermions ( $f$ ) are shown in the Feynman diagram, but the LLP can also decay to a pair of photons. In the dark shower model, the SM Higgs boson ( $H$ ) decays to a pair of dark sector quarks ( $\Psi$ ), which then hadronize to form dark showers consisting of dark scalar ( $\tilde{\eta}$ ) and vector mesons ( $\tilde{\omega}$ ) that decay back to SM particles.

of background rejection could only be achieved in current displaced vertex searches by requiring the detection of two LLP decays.

- (ii) The calorimetric nature of the particle shower is sensitive to the LLP energy rather than its mass and renders this search equally sensitive to all LLP masses considered. Conversely, vertex reconstruction efficiency tends to decrease with the LLP mass because of the increasingly smaller opening angles.

Because of these advantages, the signal acceptance and sensitivity are improved relative to the previous best results [31–33] for all LLP masses and proper lifetimes.

This paper is organized as follows. We briefly describe the CMS detector in Section 2. Section 3 provides a summary of the simulated samples used in the analysis. The reconstructions of the particle-flow objects and muon detector shower clusters are discussed in Section 4 and Section 5, respectively. The event selections are described in Section 6. The background estimation methods are detailed in Section 7. The signal modeling and systematic uncertainties are discussed in Sections 8. We report and interpret the results in Section 9. Finally, a summary is given in Section 10.

## 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 (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are identified in gas detectors embedded in the steel flux-return yoke outside the solenoid using three technologies: drift tubes (DTs) in the barrel, cathode strip chambers (CSCs) in the endcaps, and resistive-plate chambers (RPCs) in the barrel and endcaps. A more

detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [27].

The DT and CSC detectors covering the barrel and endcaps, respectively, play critical roles in the search described in this paper. The barrel DT detectors cover a region of pseudorapidity  $|\eta| < 1.2$  and are organized into four stations of concentric cylinders around the beamline and five wheels along the beamline axis ( $z$ ). The four DT stations labeled MB1 to MB4 are located approximately 4, 5, 6, and 7 m away from the interaction point radially ( $r$ ) and interleaved with the layers of the steel flux-return yoke. Each DT station contains 8–12 layers of DT cells. As charged particles traverse the DT stations, they ionize the gas molecules and produce charges that drift to the anode wire at the center of the DT cells. A signal pulse measured at the anode wire is recorded as a hit. At the hit level, the drift direction is ambiguous, so each hit is assumed to be located at the center of the corresponding DT cell. The first three stations each contain twelve layers of DT cells, in three groups of four staggered layers called “superlayers” (SLs). The innermost and outermost SLs measure the hit coordinate in the  $r\phi$ -plane, and the central SL measures in the  $z$  direction, along the beamline. The fourth station only contains two SLs measuring the hit position in the  $r\phi$ -plane.

The CSC detectors cover a region of pseudorapidity between  $|\eta| = 0.9$  and 2.4 and comprise four stations in each endcap. The four CSC stations labeled ME1 to ME4 are located approximately 7, 8, 9.5, and 10.5 m away from the interaction point along the beamline axis on both ends of the detector, and are sandwiched between steel absorbers. Each chamber is composed of six thin layers containing cathode strips along the radial direction and anode wires perpendicular to the strips. Charged particles traversing the chambers ionize the gas molecules. The resulting electrons are accelerated towards the anode wires producing an avalanche, while the positive ions travel to the opposite end and induce signals in the cathode strips. By combining the information from signals on the anode wires and the cathode strips of each layer, we can determine the space-time coordinates of each hit with a resolution of 400–500  $\mu\text{m}$  and 5 ns [36].

Events of interest are selected using a two-tiered trigger system. The first level, 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  $\mu\text{s}$  [37]. 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 and reduces the event rate to around 1 kHz before data storage [38].

### 3 Simulated samples

The simulated  $H \rightarrow SS$  signal samples are generated at next-to-leading order (NLO) with the POWHEG 2.0 [39–42] generator for the five main production processes: gluon fusion, vector boson fusion, associated production with a vector boson, and associated production with a pair of top quarks. The Higgs boson mass is set to 125 GeV, while the  $S$  mass ( $m_S$ ) is set to 0.4, 1, 3, 7, 15, 40, or 55 GeV. The  $c\tau$  is varied between 1 mm and 100 m. We consider decays to  $b\bar{b}$ ,  $d\bar{d}$ ,  $K^+K^-$ ,  $K^0\bar{K}^0$ , and  $\pi^+\pi^-$ , hereafter referred to as the fully hadronic decay modes, decays to  $\pi^0\pi^0$ ,  $\gamma\gamma$ , and  $e^+e^-$ , referred to as the fully electromagnetic decay modes, and decays to  $\tau^+\tau^-$ .

The simulated signal samples that were generated were limited to LLP masses above 0.4 GeV. Therefore, to extrapolate the analysis sensitivity to lower masses, we validated that the signal efficiency is the same for LLPs that have the same product of LLP masses and proper lifetimes. The validation has been performed for 0.4 to 4 GeV LLP masses and both fully hadronic ( $S \rightarrow$

$\pi^+\pi^-$ ) and fully leptonic ( $S \rightarrow e^+e^-$ ) decays for all search categories. The opening angle of decay particles decreases as the LLP mass decreases, and since we show that there's no difference between 0.4 and 4 GeV, we reach the limit that the opening angle is sufficiently small that the detector response is identical. Therefore, we extrapolate the signal efficiency to LLP masses below 0.4 GeV, assuming the detector response is identical for all LLP masses below 4 GeV.

For the simulated signal samples for the dark shower models, the Higgs boson production is generated at NLO with the POWHEG 2.0 [39–42] generator and includes only gluon fusion production. The Higgs boson mass is set to 125 GeV. The Higgs boson decay and phenomenology of the dark showers are generated following the tools and theory priors presented in Ref. [35], using the PYTHIA 8 hidden valley module [43–45]. In the generation, the dark sector is reduced to a single dark quark ( $\Psi$ ), vector meson ( $\tilde{\omega}$ ), and scalar meson ( $\tilde{\eta}$ ), and there are three dark quantum chromodynamic (QCD) colors. As mentioned in Section 1, we generate signal samples of five different decay portals.

In the vector portal, the dark vector meson is long-lived and couples to SM particles, while the dark scalar meson is undetectable. For the dark photon portal, the dark scalar meson decays into a pair of long-lived dark photons, which then decay into SM particles. For all other portals, the dark scalar meson is long-lived and couples to SM particles, while the dark vector meson is undetectable. The LLP mass is varied between 2–20 GeV. The minimum LLP mass considered depends on the theoretically motivated minima discussed in Ref. [35]. The  $c\tau$  is varied between 1 mm and 10 m. Characteristics of the models are  $x_{i\omega}$ , the  $\tilde{\omega}$  to  $\tilde{\eta}$  mass ratio, and  $x_{i\Lambda}$ , the ratio of the dark sector QCD scale to the dark scalar meson mass. We consider three sets of numerical values:  $(x_{i\omega}, x_{i\Lambda}) = (2.5, 2.5)$ ,  $(2.5, 1)$ , and  $(1, 1)$ . These three scenarios present a wide range of signatures with different LLP multiplicities, visible decay product multiplicities, and missing transverse energies.

For both signal models, parton showering, hadronization, and the underlying event are modeled by PYTHIA 8.205 and 8.230 [45] with parameters set by the CUETP8M1 [46] and CP5 tunes [47] used for samples simulating the 2016 and 2017–2018 datasets, respectively. The NNPDF3.0 [48] and 3.1 [49] parton distribution functions are used in the generation of all simulated samples. The GEANT4 [50] package is used to model the response of the CMS detector. Simulated minimum-bias events are mixed with the hard interactions in simulated events to reproduce the effect of additional proton-proton interactions within the same or neighboring bunch crossings as the recorded event (pileup). Events are weighted such that the distribution of the number of interactions per bunch crossing agrees with that observed during each data-taking period.

## 4 Object reconstruction

The particle-flow algorithm [51] aims to reconstruct and identify each individual particle in an event, with an optimized combination of information from the various elements of the CMS detector. The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is



obtained from the corresponding corrected ECAL and HCAL energies.

Muons are measured in the pseudorapidity range  $|\eta| < 2.4$ , with detection planes made using three technologies: DTs, CSCs, and resistive plate chambers (RPCs). The efficiency to reconstruct and identify muons is greater than 96%. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution, for muons with  $p_T$  up to 100 GeV, of 1% in the barrel and 3% in the endcaps. The  $p_T$  resolution in the barrel is better than 7% for muons with  $p_T$  up to 1 TeV [52]. In this search, muons are used to veto muon detector shower clusters in order to suppress background from muon bremsstrahlung. Depending on the search categories, different muon  $p_T$  and identification requirements are applied.

For each event, hadronic jets are clustered from these reconstructed particles using the infrared and collinear safe anti- $k_T$  algorithm [53, 54] with a distance parameter of 0.4. The jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole  $p_T$  spectrum and detector acceptance. Pileup interactions can produce additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, charged particles identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions. Jet energy corrections are derived from simulation to bring the measured response of jets to that of particle-level jets on average. In situ measurements of the momentum balance in dijet, photon + jet, Z + jet, and multijet events are used to account for any residual differences in the jet energy scale between data and simulation [55]. Additional selection criteria are applied to each jet to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures.

The missing transverse momentum vector  $\vec{p}_T^{\text{miss}}$  is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as  $p_T^{\text{miss}}$ . The  $\vec{p}_T^{\text{miss}}$  is modified to account for corrections to the energy scale of the reconstructed jets in the event [56].

## 5 Muon detector showers

When charged particles traverse the CSCs, signal pulses are collected on the anode wires and the cathode strips. The signals from wire groups are combined with signals from the cathode strips to form a two-dimensional point on each chamber layer called a CSC hit. When charged particles traverse the DT chambers, signal pulses are collected on the anode wires at the center of the DT cells. Since the DT hits only provide measurement in one dimension, the DT hit position is assumed to be at the center of each DT chamber in the orthogonal direction. For LLPs that decay within or just in front of the muon system, the material in the iron return yoke structure will induce a hadronic or electromagnetic shower, creating a geometrically localized and isolated cluster of signal hits in the muon detectors.

Because the CSC and DT hits contain different information, they are handled separately. The CSC and DT hits are clustered in  $\eta$  and the azimuthal angle  $\phi$  (in radians) using the DBSCAN algorithm [57], which groups hits by high-density regions. A minimum of 50 hits and a distance parameter  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$  of 0.2 is used. The minimum of 50 hits is chosen because it is larger than the number of hits that a muon is expected to create in the CSC or DT detectors. A spatial position is associated with each cluster by taking the geometric center of the hits in the cluster. From this, we can calculate the  $\eta$  and  $\phi$  coordinates of each cluster. Nearby clusters are merged if they satisfy  $\Delta R < 0.6$ . This is repeated until all clusters within an event are isolated. This merging procedure ensures that clusters coming from the same source are

reconstructed as one object. In the overlap region of the muon detectors, with pseudorapidity  $0.9 < |\eta| < 1.2$ , if both CSC and DT clusters are reconstructed with  $\Delta R < 0.4$ , the CSC clusters are given precedence and the corresponding DT clusters are removed.

The cluster reconstruction efficiencies, including both DT and CSC clusters, as a function of the simulated  $r$  and  $|z|$  decay positions of the particle S are shown in Fig. 2. The DT (CSC) cluster reconstruction efficiencies are shown separately as a function of the simulated  $r$  ( $|z|$ ) in Fig. 3. The cluster efficiencies are shown for events satisfying  $p_T^{\text{miss}} > 200 \text{ GeV}$  as required in the search described in Section 6. The cluster reconstruction efficiency strongly depends on the LLP decay position. The efficiency is highest when the LLP decays near the edges of the shielding absorber material, where there is enough material to induce the shower, but not so much that it stops the shower secondary particles.

The cluster reconstruction efficiency also depends on whether the LLP decays hadronically or leptonically. In general, hadronic showers have higher efficiency because they are more likely to penetrate through the steel in between stations. Showers induced by electromagnetic decays generally occupy just one station and are stopped by the steel between stations. When the LLP decays close to or in the CSCs, defined as the union of the two regions: (1)  $500 < |z| < 661 \text{ cm}$ ,  $r < 270 \text{ cm}$ , and (2)  $|\eta| < 2.4$  and (2)  $660 < |z| < 1100 \text{ cm}$ ,  $r < 695.5 \text{ cm}$ , and  $|\eta| < 2.4$ , the inclusive CSC cluster reconstruction efficiency is approximately 80% for fully hadronic decays, 55% for  $\tau^+\tau^-$  decays, and 35% for fully leptonic decays. When the LLP decays close to or in the DTs, defined as the region  $380 < r < 738 \text{ cm}$  and  $|z| < 661 \text{ cm}$ , the inclusive DT cluster reconstruction efficiency is approximately 80% for fully hadronic decays, 60% for  $\tau^+\tau^-$  decays, and 45% for fully leptonic decays.

The accuracy of the simulation modeling of the cluster reconstruction efficiency has been studied using a  $Z \rightarrow \mu^+\mu^-$  data sample, where clusters are produced when one of the muons undergoes bremsstrahlung and the associated photon produces an electromagnetic shower. The discrepancy between data and simulation is taken as a systematic uncertainty on the cluster reconstruction efficiencies, as detailed in Section 8.

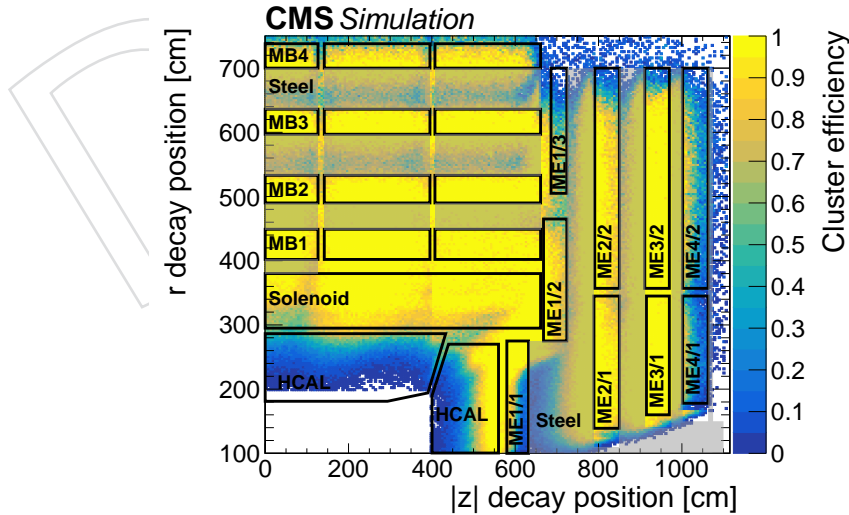


Figure 2: The cluster reconstruction efficiencies, including both DT and CSC clusters, as a function of the simulated  $r$  (left) and  $|z|$  (right) decay positions of the particle S decaying to  $d\bar{d}$ , for a mass of  $40 \text{ GeV}$  and a range of  $c\tau$  values between 1 and 10 m. The barrel and endcap muon stations are drawn as black boxes and labeled by their station names. Regions occupied by steel shielding are shaded in gray.



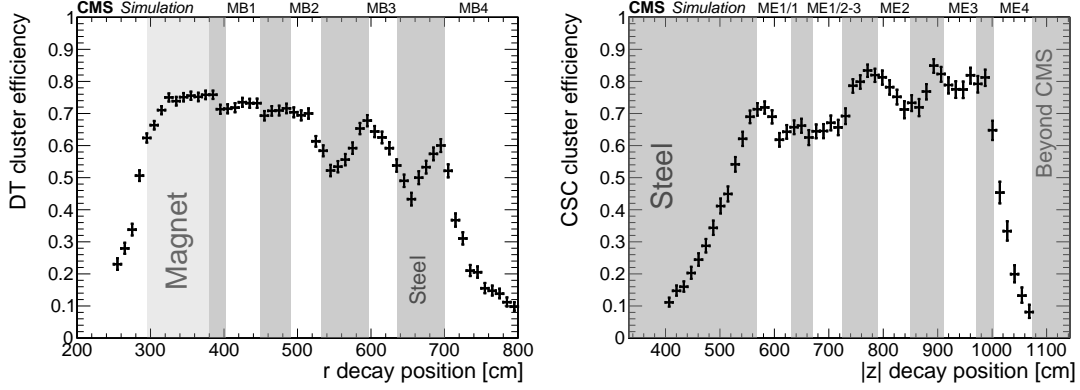


Figure 3: The DT (left) and CSC (right) cluster reconstruction efficiency as a function of the simulated  $r$  or  $z$  decay positions of  $S$  decaying to  $d\bar{d}$ , for a mass of 40 GeV and a range of  $c\tau$  values between 1 and 10 m. Regions occupied by steel shielding are shaded in gray.

## 6 Search strategy and event selection

An LLP that decays after it has traversed the calorimeter systems can produce large  $p_T^{\text{miss}}$  because its momentum will not be properly measured by the tracker, calorimeter, or muon systems. We exploit this feature by triggering on events with online  $p_T^{\text{miss}} > 120$  GeV [38], and subsequently requiring offline  $p_T^{\text{miss}} > 200$  GeV, to ensure that the selected events are in the plateau region of the high-level trigger efficiency. We require at least one jet with  $p_T > 30$  GeV and pseudorapidity  $|\eta| < 2.4$ , because signal events passing the  $p_T^{\text{miss}}$  requirement are always produced together with a jet from initial-state radiation. To suppress noncollision backgrounds, we apply filters that remove events containing beam-halo muons or calorimeter noise [58]. The event-level selections are kept minimal to be as model independent as possible.

After the event-level selection, the events are separated into 3 mutually exclusive categories based on the number and location of the clusters: (1) events with two clusters in the muon detectors, (2) events with exactly one CSC cluster, and (3) events with exactly one DT cluster. Events with two clusters are further categorized into categories with 2 CSC clusters, 2 DT clusters, and 1 CSC and 1 DT cluster. The second category is based on a previous search using the endcap muon detectors [34] with a few changes, including explicitly excluding overlapping double cluster events and loosening the event-level selections to be consistent with other categories. The geometric acceptance and the efficiency of the  $p_T^{\text{miss}}$  selection for each category are shown in Fig. 4.

The main SM backgrounds are similar among the 3 categories and include punch-through jets, muons that undergo bremsstrahlung, and isolated hadrons from pileup, recoils, or underlying events. To suppress background from punch-through jets or muon bremsstrahlung, we reject CSC and DT clusters that have a jet or muon within  $\Delta R < 0.4$  in all categories. However, depending on the category, the  $p_T$  thresholds and identification requirements of jets and muons are different. Furthermore, additional tighter vetoes are applied to the single cluster categories to reject background, as detailed in the following subsections.

Additionally, the difference in azimuthal angle ( $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$ ) between the cluster location and the  $\vec{p}_T^{\text{miss}}$  is used as a discriminating variable in all 3 categories. For signal,  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  peaks near zero because the large  $p_T^{\text{miss}}$  requirement tends to select highly energetic Higgs bosons for which the  $S$  and  $H$  momentum vectors nearly coincide. For the backgrounds,  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  is uniformly distributed because the cluster and  $\vec{p}_T^{\text{miss}}$  are independent. The exact threshold on the variable is different for the barrel and endcap, and looser for the double

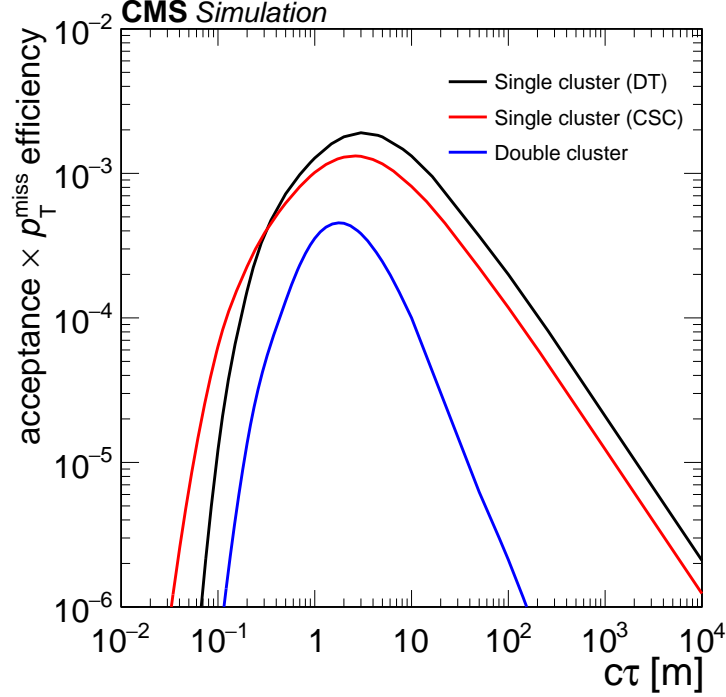


Figure 4: The geometric acceptance and the efficiency of the  $p_T^{\text{miss}} \geq 200 \text{ GeV}$  selection as a function of  $S$  proper decay length for a mass of  $40 \text{ GeV}$ .

cluster category, as detailed in the following subsections.

Finally, the number of hits in the clusters ( $N_{\text{hits}}$ ) is used to discriminate signal and background. Signal clusters tend to have large  $N_{\text{hits}}$ , while background clusters are expected to have small  $N_{\text{hits}}$ . The exact threshold on  $N_{\text{hits}}$  is different for the barrel and endcap, and looser for the double cluster category, as detailed in the following subsections that describe the detailed event selections for each category.

## 6.1 Double clusters

In the double cluster category, we search for events containing two clusters satisfying the selection criteria described below. Events are separated into 3 categories depending on whether there are two DT clusters, two CSC clusters, or one CSC and one DT cluster present. Requiring two muon system clusters significantly reduces the expected background, so the selection requirements in the double cluster category are much looser compared to the single cluster categories.

CSC clusters are rejected if any jet with  $p_T > 30 \text{ GeV}$  or global muon [52], built by matching muon tracks in the muon detectors and tracker, with  $p_T > 30 \text{ GeV}$  is found within  $\Delta R < 0.4$ . Similarly, DT clusters are rejected if any jet with  $p_T > 50 \text{ GeV}$  or muon passing loose identification criteria [52, 59] with  $p_T > 10 \text{ GeV}$  is found within  $\Delta R < 0.4$ .

We veto CSC clusters that are entirely contained in the innermost rings of the ME1 station (ME1/1) and veto DT clusters that have more than 90% of the hits contained in the innermost station (MB1), both of which have the least absorber material between them and the interaction point, and larger background contamination.

We further reject CSC clusters produced by adjacent bunch crossings, known as out-of-time (OOT) pileup, by requiring that the cluster time  $t_{\text{cluster}}$ , defined as the average time of the hits in

the cluster relative to the LHC clock, is consistent with an in-time interaction ( $-5.0 < t_{\text{cluster}} < 12.5 \text{ ns}$ ). An asymmetric time window is used to capture signal clusters with longer delays from slower-moving LLPs. To reject clusters composed of hits from multiple bunch crossings, the root-mean-square of the hit times of each cluster is required to be less than 20 ns.

Tracks from muons in the barrel are likely to deposit a similar number of hits in all 4 DT stations, while showers from LLP decays are likely to have hits concentrated in 1 or 2 stations. Therefore, to reject DT clusters from muon bremsstrahlung we veto clusters that contain hits in all 4 stations and that have a ratio of the minimum and maximum number of hits per station less than 0.4.

Cosmic ray muons produce hits in both the upper and lower hemispheres of the muon barrel system. To suppress this background, we reject DT clusters if there are at least 6 segments, which are straight-line tracks built within each DT chamber, and at least 1 segment in every station found in the opposite hemisphere ( $|\Delta\phi| > 2$ ) from the cluster. In addition, cosmic ray muon showers produce hits in multiple regions of the CMS detector. Thus, we reject any event in which more than a quarter of the DT and CSC  $\phi$ -rings, consisting of chambers along  $\phi$  with the same  $r$  and  $z$  coordinates, contain 50 or more hits. Finally, we require  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster}) < 1$  (1.2) for CSC (DT) clusters.

For signal events with two clusters, the  $\Delta R$  between two LLPs, thus between the clusters, is typically small. For background, the  $\Delta R$  between clusters can be large because the two clusters generally come from separate processes especially for the CSC-CSC and DT-CSC categories. Therefore, we require the  $\Delta R$  between the two clusters to be less than 2 (2.5) for the CSC-CSC (DT-CSC) subcategory. There is already an implicit  $\Delta\phi$  selection between the two clusters by requiring both clusters to pass the  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  selection. The  $\Delta R$  selection additionally requires the two clusters to be close in  $\eta$ .

Finally,  $N_{\text{hits}}$  is used to discriminate signal and background. We require  $N_{\text{hits}} \geq 100$  (80) for CSC (DT) clusters.

## 6.2 Single CSC cluster

In the single CSC cluster category, we search for events in which only one LLP decay produces a displaced cluster in the endcap muon system. In this category, the expected background yield is significantly higher than in the double cluster category, so we apply much tighter cluster veto requirements to achieve the same near-zero background level.

The cluster veto requirements are the same as in Ref. [34], except the event-level selections are changed to align with the other two categories, and events with two clusters are only included in the double cluster category, which has higher sensitivity. Clusters that have a jet with  $p_T > 10 \text{ GeV}$  or muon with  $p_T > 20 \text{ GeV}$  within  $\Delta R < 0.4$  are rejected. We veto clusters that have any hits in the two innermost rings of the ME1 station (ME1/1 and ME1/2), which have the least absorber material between them and the interaction point, or match any hit (with  $\Delta R(\text{cluster}, \text{hit}) < 0.4$ ) in the RPCs located immediately next to ME1/2. In the region where the barrel and endcap muon detectors overlap ( $0.9 < |\eta| < 1.2$ ), we veto any cluster matched to any track segment reconstructed in the innermost station of the DT detectors (MB1), or any hit in the RPCs situated in front of and behind MB1 matched to within  $\Delta R(\text{cluster}, \text{segment or hit}) < 0.4$ . Finally, we reject clusters with  $|\eta| > 2.0$  to suppress the muon bremsstrahlung background that evaded the muon veto because of the decreasing muon reconstruction and identification efficiencies at larger  $|\eta|$ .

As in the double cluster category, we reject any event in which more than a quarter of the DT

and CSC  $\phi$ -rings contain 50 or more hits.

After the veto requirements are applied, the dominant background source consists of decays of SM LLPs, which are predominantly produced by pileup interactions and are independent of the primary interaction that yielded the large  $p_T^{\text{miss}}$ . These pileup interactions may be in-time or OOT with the primary interaction, as shown in Fig. 5. Clusters produced by OOT pileup are rejected by requiring  $-5.0 < t_{\text{cluster}} < 12.5$  ns, as in the double cluster category. The time window requirement suppresses the background by a factor of 5. Similarly, the root-mean-square of the hit times of each cluster is required to be less than 20 ns.

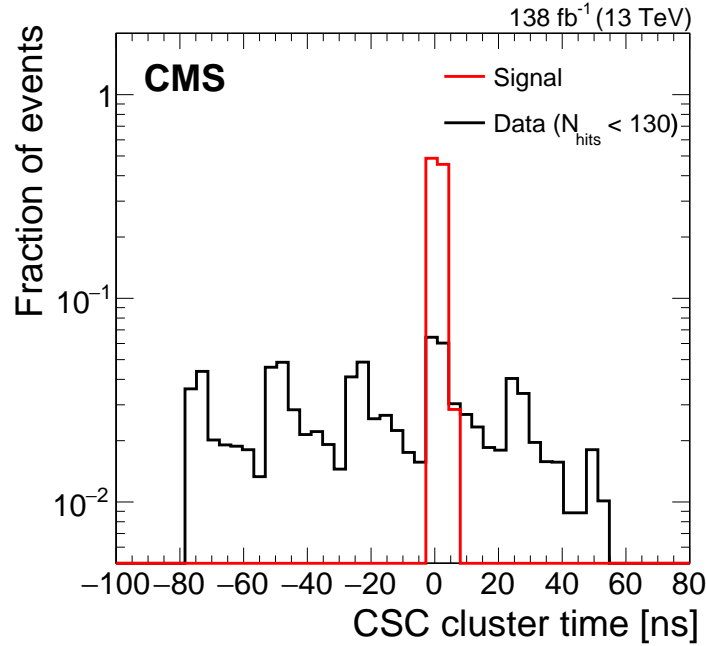


Figure 5: The shapes of the cluster time for signal, where  $S$  decaying to  $d\bar{d}$  for a lifetime of 1 m and mass of 40 GeV and for a background-enriched sample in data, selected by inverting the  $N_{\text{hits}}$  requirement.

Several features distinguish signal and background clusters. Clusters from all background processes occur more often at larger values of  $|\eta|$ , as the effectiveness of the jet and muon vetoes decreases because of decreasing reconstruction efficiencies. Signal clusters often occupy more than one CSC station ( $N_{\text{stations}} > 1$ ) and occur more frequently in stations farther away from the primary interaction point. A cluster identification algorithm was devised that makes more restrictive  $|\eta|$  requirements for clusters that occupy only one CSC station ( $N_{\text{stations}} = 1$ ) and are closer to the primary interaction point. The  $|\eta|$  requirements are:

- $|\eta| < 1.9$  if  $N_{\text{stations}} > 1$ ,
- $|\eta| < 1.8$  if  $N_{\text{stations}} = 1$  and the cluster is in station 4, then require  $|\eta| < 1.8$ ,
- $|\eta| < 1.6$  if  $N_{\text{stations}} = 1$  and the cluster is in station 3 or station 2, and
- $|\eta| < 1.1$  if  $N_{\text{stations}} = 1$  and the cluster is in station 1 because of an implicit selection from the ME1/1 and ME1/2 vetoes.

The cluster identification algorithm has  $\sim 80\%$  efficiency for simulated clusters originating from  $S$  decays, and suppresses the background by a factor of 3. The events that pass the cluster identification criteria are used to define the search region, and those that fail are used as an additional in-time validation region.

Both  $N_{\text{hits}}$  and  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  are used to discriminate signal and background. The signal and background shapes of the two discriminants are shown in Fig. 6. For the backgrounds,  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  is independent of  $N_{\text{hits}}$ , enabling the use of the matrix method to predict the background yield in the signal-enriched bin, as detailed in Section 7.

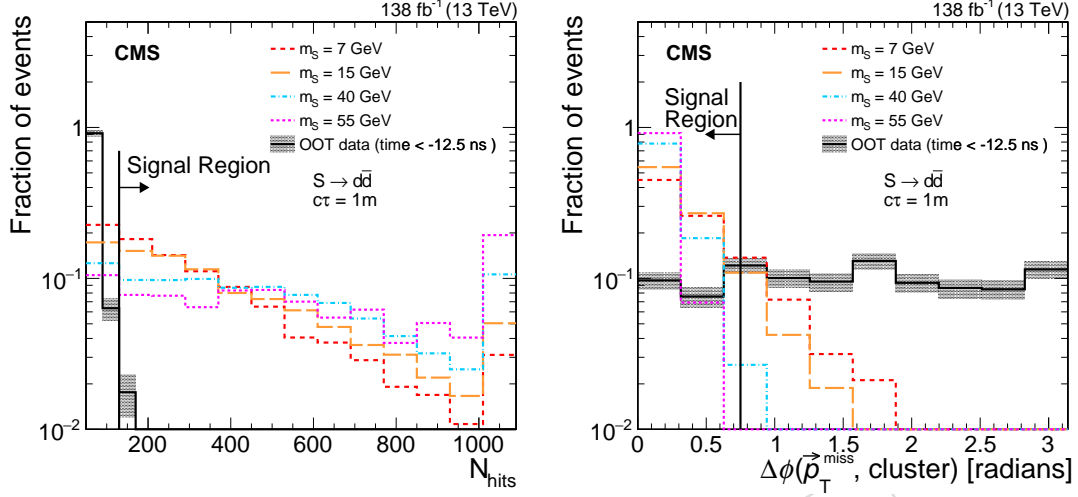


Figure 6: The shape of  $N_{\text{hits}}$  (left) and  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  (right) are shown for  $S$  decaying to  $d\bar{d}$  for a lifetime of 1 m and various masses compared to the OOT background ( $t_{\text{cluster}} < -12.5$  ns). The OOT background is representative of the overall background shape, because the background passing all the selections described above is dominated by pileup and underlying events.

### 6.3 Single DT cluster

The single DT cluster category targets events in which only one LLP decay produces a displaced cluster in the barrel muon system. Events passing the selection criteria for the double cluster or single CSC cluster categories are not considered in this category, to give precedence to the category with higher sensitivity (double cluster) and to minimize differences with the previously published search (single CSC cluster).

First, to remove high  $p_T^{\text{miss}}$  events due to mismeasured jets, we require the minimum of  $|\Delta\phi(\text{jet}, \vec{p}_T^{\text{miss}})|$  over all the jets with  $p_T > 30$  GeV to be greater than 0.6. This requirement reduces the background from SM events composed uniquely of jets produced through the strong interaction, referred to as QCD multijet events, and is only applied to the single DT cluster category because this category is dominated by punch-through jet background.

We veto clusters that have a jet with  $p_T > 10$  GeV or a muon with  $p_T > 10$  GeV passing loose identification criteria [52, 59] within  $\Delta R < 0.4$ . In addition, we reject clusters that are within  $\Delta R < 1.2$  from the leading- $p_T$  jet. Furthermore, DT clusters that are within  $\Delta R < 0.4$  of two or more hits in the innermost station MB1 are rejected. Additionally, clusters with maximum hit counts in MB3 or MB4 are rejected if they are within  $\Delta R < 0.4$  of two or more MB2 hits. Each cluster is associated with a wheel based on the average  $z$  position of its hits. To reject clusters from noise in the DTs, we require clusters to be matched to at least 1 RPC hit from the same wheel and within  $\Delta\phi < 0.5$ .

To suppress background from cosmic ray muons, we veto clusters that have more than 8 hits in MB1 within  $\Delta\phi < \pi/4$  in either adjacent wheel. In addition, we veto clusters with maximum hit counts in MB3 and MB4 that have more than 8 hits in MB2 within  $\Delta\phi < \pi/4$  in either adjacent wheel. Furthermore, we look for DT segments that are far from the clusters with



$\Delta R > 0.4$  in the upper or lower hemisphere of the DT wheels. We veto the cluster if more than 14 segments are found in either hemisphere or more than 10 segments are found in both hemispheres.

As in the single CSC cluster category,  $N_{\text{hits}}$  and  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  are used to discriminate signal and background. We require  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster}) < 1$  and  $N_{\text{hits}} > 100$ . The signal and background distributions of the two discriminants are shown in Fig. 7.

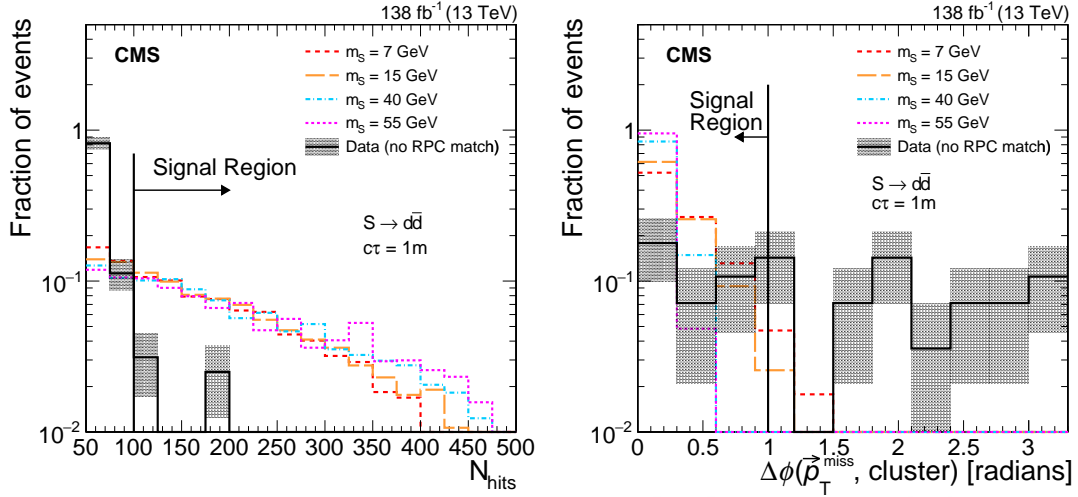


Figure 7: The shapes of  $N_{\text{hits}}$  (left) and  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  (right) for DT clusters are shown for  $S$  decaying to  $d\bar{d}$  for a proper decay length of 1 m and various masses compared to the shape of background in a selection where the cluster is not matched to any RPC hit.

Finally, the DT clusters are categorized into 3 exclusive categories according to the station that contains the most hits: MB2, MB3, or MB4. These categories have different background compositions, where the punch-through jet background is more prominent in the stations that are closer to the interaction point.

## 7 Background estimation

The data-driven matrix (ABCD) method is used for background estimation for all 3 categories. The matrix method requires two variables that discriminate signal and background and are independent of one another for the background. Two separate requirements, one on each variable, partition the two-dimensional space into four bins, A, B, C, and D, where bin A contains events that pass both signal-like requirements, events in bins B and D only pass one of the requirements, and events in bin C pass neither requirement. Because of the independence of the two variables, the expected background event rate in the signal-enriched bin A can be related to the other three bins by  $\lambda_A = (\lambda_B \lambda_D) / \lambda_C$ , where  $\lambda_X$  is the expected background event rate (i.e. Poisson mean) in each bin X. In the double-cluster categories, the two variables that are used are the  $N_{\text{hits}}$  of each of the clusters, while in the single DT and CSC cluster categories, the two variables are  $N_{\text{hits}}$  and  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$ .

In addition to the background predicted by the matrix method, the single DT cluster category estimates the punch-through jet background separately, while the punch-through jet background in the other categories is negligible. Other noncollision backgrounds, including cosmic ray muons, have been suppressed by dedicated filters described in Section 6 and have been demonstrated to be negligible in the signal region.

The following subsections detail for each category the main background component, the background estimation method, and its validation.

### 7.1 Double clusters

For the DT-CSC category, the two independent discriminating variables are the  $N_{\text{hits}}$  of the DT and CSC cluster respectively. The four bins are defined as follows and shown in the left panel of Fig. 8:

- (A) Bin A includes events with the CSC cluster with  $N_{\text{hits}} > 100$  and the DT cluster with  $N_{\text{hits}} > 80$ ;
- (B) Bin B includes events with the CSC cluster with  $N_{\text{hits}} > 100$  and the DT cluster with  $N_{\text{hits}} \leq 80$ ;
- (C) Bin C includes events with the CSC cluster with  $N_{\text{hits}} \leq 100$  and the DT cluster with  $N_{\text{hits}} \leq 80$ ;
- (D) Bin D includes events with the CSC cluster with  $N_{\text{hits}} \leq 100$  and the DT cluster with  $N_{\text{hits}} > 80$ .

For the CSC-CSC and DT-DT categories, the two variables are symmetric, so we combine bins B and D and define the combined expected background rate as  $\lambda_{BD}$ . Bins A, BD, and C contain events with 2, 1, and 0 clusters passing the  $N_{\text{hits}}$  selection, respectively, as shown in the right panel of Fig. 8. The expected background yield in the signal-enriched bin A is related to the other two bins as  $\lambda_A = (\lambda_{BD}/2)^2/\lambda_C$ .

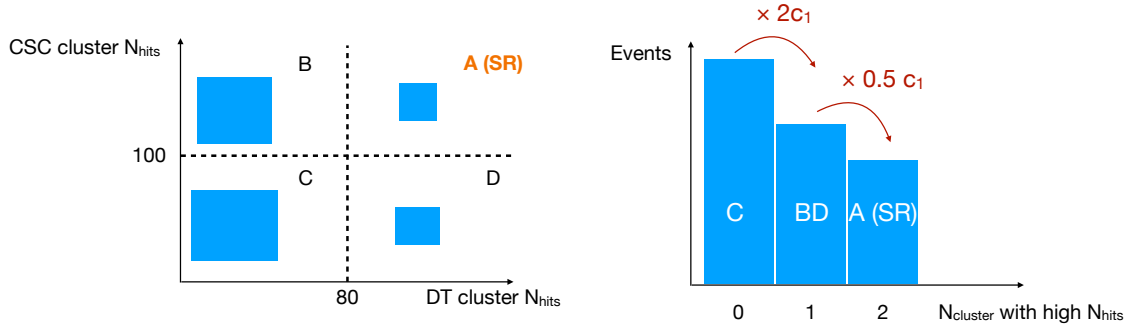


Figure 8: Diagrams illustrating the ABCD plane for the DT-CSC category (left) and for the DT-DT and CSC-CSC categories (right). The variable  $c_1$  is the pass-fail ratio of the  $N_{\text{hits}}$  selection for the background cluster. Bin A is the signal region for all categories.

To validate the background estimation method, we define two validation regions: the inverted  $N_{\text{hits}}$  region and the inverted  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  region. The inverted  $N_{\text{hits}}$  validation region is defined by inverting the  $N_{\text{hits}}$  requirements for both clusters while maintaining all the other cluster-level selections. The  $N_{\text{hits}}$  threshold used in the inverted  $N_{\text{hits}}$  validation region is 70 (80) for DT (CSC) clusters. Similarly, the inverted  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  validation region is defined by inverting the  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  requirement of both clusters while maintaining all the other cluster-level selections. To probe signal-like events in the inverted  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  validation region, we additionally require  $\Delta\phi(\text{cluster1}, \text{cluster2}) < 2$  for the DT-DT category. For the CSC-CSC and DT-CSC categories, the two clusters are close to each other because of the  $\Delta R(\text{cluster1}, \text{cluster2})$  requirement. The  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  validation region allows us to test

for any nonnegligible backgrounds at high  $N_{\text{hits}}$  that cannot be accessed in the inverted  $N_{\text{hits}}$  validation region. The background estimate agrees with the number of observed background events in both validation regions and all 3 categories, as shown in Table 1.

Table 1: Validation of the ABCD method for the double cluster category in both validation regions. The uncertainty of the prediction is the statistical uncertainty propagated from bins B, C, and D or bins BD and C. The symbol  $\lambda_X$  is the expected background event rate in bin X, while  $N_X$  is the observed number of events in bin X.

Category	Validation region	$\lambda_A$	$N_A$	$N_B$	$N_C$	$N_D$	$N_{BD}$
DT-DT	Inverted $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$	$0.02 \pm 0.05$	0	—	11	—	1
	Inverted $N_{\text{hits}}$	$0.12 \pm 0.27$	0	—	2	—	1
CSC-CSC	Inverted $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$	$0.12 \pm 0.18$	0	—	8	—	2
	Inverted $N_{\text{hits}}$	$0.25 \pm 0.38$	0	—	4	—	2
DT-CSC	Inverted $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$	$0 \pm 0.3$	0	0	19	3	—
	Inverted $N_{\text{hits}}$	$0.18 \pm 0.23$	0	2	11	1	—

## 7.2 Single CSC cluster

For the single CSC category, the two discriminating variables are  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  and  $N_{\text{hits}}$ . For the backgrounds,  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  is independent of  $N_{\text{hits}}$  enabling the use of the matrix method. As shown in Fig. 9, the four bins are defined as follows.

- (A) Bin A includes events with  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster}) < 0.75$  and  $N_{\text{hits}} > 130$ ;
- (B) bin B includes events with  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster}) \geq 0.75$  and  $N_{\text{hits}} > 130$ ;
- (C) bin C includes events with  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster}) \geq 0.75$  and  $N_{\text{hits}} \leq 130$ ; and
- (D) bin D includes events with  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster}) < 0.75$  and  $N_{\text{hits}} \leq 130$ .

The background estimation procedure is validated using events in two separate validation regions: one in the OOT region, where  $t_{\text{cluster}} < -12.5 \text{ ns}$ , and one in the in-time region, where the clusters that fail the cluster identification criteria are selected, as shown in Table 2.

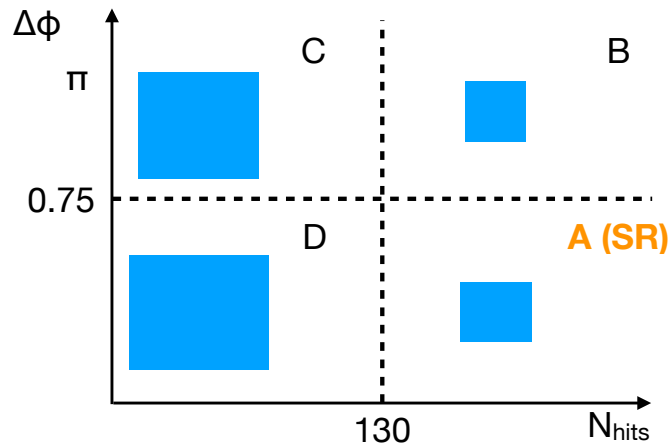


Figure 9: Diagram illustrating the ABCD plane, where bin A is the signal region.

Table 2: Validation of the ABCD method for the single CSC cluster category in both validation regions. The uncertainty of the prediction is the statistical uncertainty propagated from bins B, C, and D. The symbol  $\lambda_X$  is the expected background event rate in bin X, while  $N_X$  is the observed number of events in bin X.

Validation region	$\lambda_A$	$N_A$	$N_B$	$N_C$	$N_D$
Out-of-time region	$2.2 \pm 0.8$	3	8	442	121
In-time region	$2.2 \pm 0.8$	2	8	317	87

### 7.3 Single DT cluster

The dominant backgrounds in the single DT cluster category are punch-through jets and low- $p_T$  particles from pileup. To estimate the background from the low- $p_T$  particles, the matrix method is used. Like the single CSC category, the same discriminating variables  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  and  $N_{\text{hits}}$  are used, except the thresholds for the signal-like selection are  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster}) < 1$  and  $N_{\text{hits}} > 100$ .

The background prediction is validated in a pileup-enriched validation region, as shown in Table 3. The pileup-enriched region is defined by inverting the loose identification criterion on the leading jet, and removing the RPC matching criteria,  $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$  requirement, and filters that reject noncollision background. To reduce the statistical uncertainties when estimating the background rate in bins C and D, we merge the MB3 and MB4 categories. The final ABCD fit in the signal region is also performed with those categories merged.

To estimate the punch-through jet background in the signal region, which is not accounted for in the ABCD method, we measure the number of observed events in excess of the ABCD prediction in the region with the inner DT station hit veto inverted, and multiply it by  $\epsilon / (1 - \epsilon)$  where  $\epsilon$  is the corresponding veto efficiency. For clusters in MB2, only the MB1 veto is applied, so only the MB1 veto is inverted. For clusters in MB3 and MB4, both MB1 and MB2 vetoes are inverted. The number of excess events in the inverted region is measured to be  $22 \pm 7$ ,  $7 \pm 3$ , and  $2.0 \pm 1.7$  for clusters in MB2, MB3, and MB4, respectively. The veto efficiency  $\epsilon$  is measured to be  $0.23 \pm 0.02$ ,  $0.38 \pm 0.07$ ,  $0.29 \pm 0.14$  for clusters in MB2, MB3, and MB4, respectively. The statistical uncertainties in the measured veto efficiencies and the number of excess events in the inverted region are propagated as an additional systematic uncertainty on the background prediction.

The inner DT station (MB1 or MB1 plus MB2) hit veto efficiency is measured in a separate punch-through jet enriched region, by selecting clusters that have a jet with  $p_T > 10$  GeV within  $\Delta R < 0.4$ . The measured veto efficiency is fitted as a function of the matched jet  $p_T$  with the sum of an exponential and constant function for clusters from each station separately. The veto efficiencies are then extrapolated by evaluating the fit function at 0 GeV to predict the veto efficiencies for clusters passing the jet veto.

The punch-through jet background prediction method is validated, as shown in Table 4, by predicting background clusters that are matched to 2–5 MB1 or 2–5 MB2 hits, instead of  $< 2$  MB1 or  $< 2$  MB2 hits in the signal region.

## 8 Systematic uncertainties

The dominant source of uncertainty in the signal prediction is missing higher order QCD corrections, which amounts to 21% for the gluon fusion production. The other main sources of uncertainties include the signal modeling of the cluster reconstruction and selections, as detailed

Table 3: Validation of the ABCD method for the single DT cluster category in a pileup-enriched region. The uncertainty of the prediction is the statistical uncertainty propagated from bins B, C, and D. Bins C and D for MB3 and MB4 categories are combined to reduce statistical uncertainty in the two regions. The final ABCD fit in the signal region will also be performed with those bins combined.

Cluster station	$\lambda_A$	$N_A$	$N_B$	$N_C$	$N_D$
MB2	$1.3 \pm 0.9$	3	2	130	82
MB3	$0.6 \pm 1.0$	1	1	20	11
MB4	$0.0 \pm 1.1$	1	0		

Table 4: Validation of the punch-through jet background prediction method for the single DT cluster category. The uncertainty of the prediction is the statistical uncertainty propagated from the extrapolated MB1/MB2 hit veto efficiency.

Cluster station	$\lambda_A$	$N_A$
MB2	$4.7 \pm 1.5$	6
MB3	$1.5 \pm 0.9$	0
MB4	$1.0 \pm 0.9$	0

in the following paragraphs, jet energy scale (3–6%) [55], parton distribution functions (PDFs) (3%), pileup modeling (2%), integrated luminosity (1.6%) [60–62],  $p_T^{\text{miss}}$  trigger efficiency (5% downward correction and 1% uncertainty) and simulation sample size (3–5%).

The only source of systematic uncertainty in the background prediction is the punch-through jet background in the single DT cluster category. The size of the uncertainty is 32, 50, and 100% for clusters in MB2, MB3, and MB4, respectively. No additional background systematic uncertainties are assigned for the background predicted by the ABCD method, because the data-driven background estimation method is validated, as detailed in Section 7.

We studied the simulation modeling of the cluster reconstruction efficiency, cluster-level selections, and veto efficiencies. The accuracy of the simulation prediction for the clusters depends on its modeling of the response of the muon detectors in an environment with multiple particles, each producing many secondary shower particles. This aspect is validated by measuring clusters produced in  $Z \rightarrow \mu^+\mu^-$  data events, where one of the muons undergoes bremsstrahlung in the muon detectors and the associated photon produces an electromagnetic shower. The discrepancy between data and simulation is taken as an additional systematic uncertainty in the cluster reconstruction and selection efficiencies.

The simulation modeling of the veto efficiencies, including the jet, muon, ME1, MB1, and RPC hit vetoes, is predominantly determined by the simulation of jets and muons, the presence of pileup particles, and random noise. The veto efficiencies are measured by randomly sampling the  $(\eta, \phi)$  locations of clusters from the signal distribution and evaluating whether a jet, muon, or ME1, MB1, or RPC hit from  $Z \rightarrow \mu^+\mu^-$  events has been observed within  $\Delta R < 0.4$  of the cluster’s location.

For CSC clusters in the single CSC cluster category, the systematic uncertainty is dominated by the cluster reconstruction efficiency and the cluster identification efficiency. This uncertainty is determined by measuring the difference in efficiencies in simulation and data, which amounts to an 8% relative uncertainty. Additionally, in signal simulations, the CSC hits are always assumed to be read out, while in real data acquisition, only those hits that have at least two cathode hits at different CSC layers and match predefined hit patterns are read out. This could



lead to an underestimation of the efficiency of ME1/1 or ME1/2 vetoes in simulation, which we estimated to have an uncertainty of 1%.

The signal loss from the vetoes is dominated by the muon veto, which is affected by muon segments produced by particles resulting from the LLP decay itself. The simulation modeling of this effect is measured using a control sample of clusters matched to trackless jets made to resemble the signal LLP decay by requiring the neutral energy fraction to be larger than 95%. A 10% downward correction is applied to the signal efficiency to account for the simulation's mismodeling of the vetoes.

For CSC clusters in the double cluster category, looser selections are applied, resulting in the systematic uncertainty being dominated by the cluster time spread requirement, which amounts to 10% in the CSC-CSC category and 5% in the DT-CSC category. Furthermore, in the double cluster category, the jet and muon vetoes are implemented with tighter identification criteria. As a result, the presence of jets and muons is well modeled and no corrections and uncertainty are assigned.

For the DT clusters in the single DT cluster category, the systematic uncertainty is dominated by the cluster reconstruction efficiency, which is measured to be 15%. The MB1 and MB2 veto efficiencies are also measured separately by randomly sampling the locations of the clusters from the signal distribution and evaluating whether an MB1 or MB2 hit has been matched. No correction or uncertainty is assigned for the MB2 veto. A 10% downward correction with a 7% uncertainty is applied to the signal efficiency to account for additional noise MB1 hits in data.

For the DT clusters in the double cluster category, the signal systematic uncertainty is dominated by the cluster reconstruction efficiency, which is measured to be 3% and 1% in DT-DT and DT-CSC categories, respectively.

## 9 Results and interpretation

In this section, we report the search results in each category and the interpretation of the combined results of all three categories in the twin Higgs and dark shower models.

### 9.1 Double cluster category

The result of the search in the double cluster category is shown in Table 5. We observed no statistically significant deviation with respect to the SM prediction. The signal and data distributions of  $N_{\text{clusters}}$  passing the  $N_{\text{hits}}$  selection are shown in Fig. 10.

Table 5: Number of predicted background and observed events in the double cluster category.

	Category	A	B	C	D
Background-only fit	DT-DT	$0.06 \pm 0.06$	$0.9 \pm 0.7$	$3.1 \pm 1.6$	—
	CSC-CSC	$0.7 \pm 0.4$	$3.6 \pm 1.5$	$4.7 \pm 2.0$	—
	DT-CSC	$0.12 \pm 0.12$	$1.9 \pm 1.2$	$14.1 \pm 3.8$	$0.9 \pm 0.7$
Observation	DT-DT	0	1	3	—
	CSC-CSC	2	1	6	—
	DT-CSC	0	2	14	1

### 9.2 Single CSC cluster category

The corresponding search result in the single CSC cluster category is shown in Table 6. No statistically significant deviation with respect to the SM prediction is observed. The signal

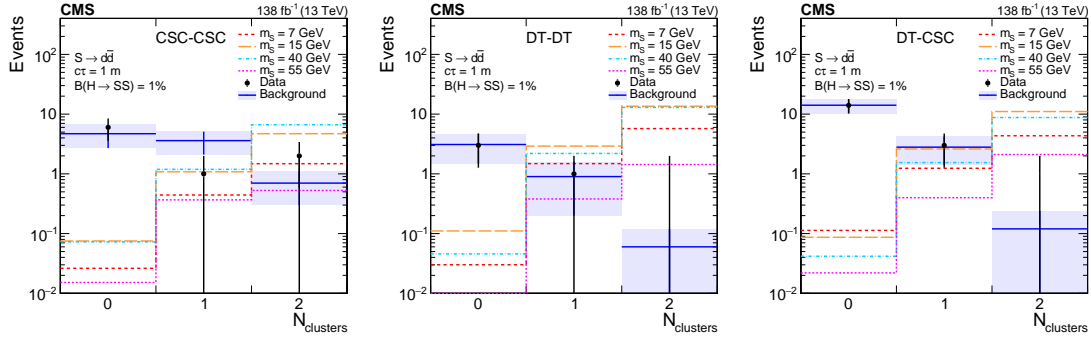


Figure 10: The signal (assuming  $\mathcal{B}(H \rightarrow SS) = 1\%$ ,  $S \rightarrow d\bar{d}$ , and  $c\tau = 1$  m) and data distributions of  $N_{\text{clusters}}$  passing the  $N_{\text{hits}}$  selection in the search region for CSC-CSC (left), DT-DT (center), and DT-CSC (right) categories.

552 and data distributions of  $N_{\text{hits}}$  ( $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$ ) in the bins A and D (A and B) are shown in  
 553 Fig. 11.

Table 6: Number of predicted background and observed events in the single CSC cluster category.

	A	B	C	D
Background-only fit	$1.8 \pm 0.8$	$4.2 \pm 1.7$	$120 \pm 11$	$51 \pm 7$
Observed	3	3	121	50

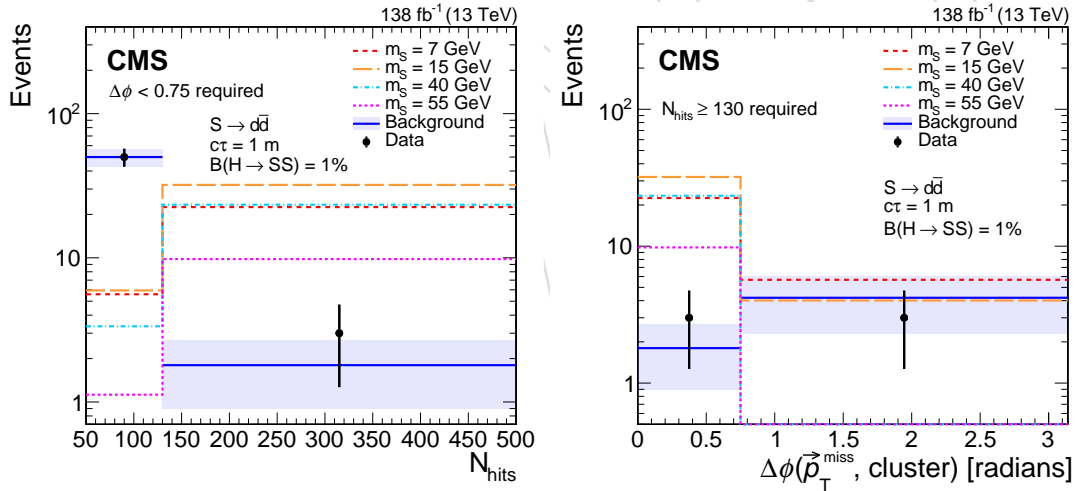


Figure 11: Distributions of  $N_{\text{hits}}$  (left) and  $\Delta\phi_c$  (right) in the search region of single CSC cluster category. The background predicted by the fit is shown in blue with the shaded region showing the fitted uncertainty. The expected signal with  $\mathcal{B}(H \rightarrow SS) = 1\%$ ,  $S \rightarrow d\bar{d}$ , and  $c\tau = 1$  m is shown for  $m_S$  of 7, 15, 40, and 55 GeV in various colors and dotted lines. The  $N_{\text{hits}}$  distribution includes only events in bins A and D, while the  $\Delta\phi_c$  one includes only events in bins A and B. The last bin in the  $N_{\text{hits}}$  distribution includes overflow events.

### 9.3 Single DT cluster category

555 Using the methods detailed in Section 7, the background from punch-through jets and low-  
 556  $p_T$  pileup particles are estimated in the single DT cluster category. The result of the search is  
 557 shown in Table 7. We observed no statistically significant deviation with respect to the SM  
 558 prediction. The signal and data distributions of  $N_{\text{hits}}$  ( $\Delta\phi(\vec{p}_T^{\text{miss}}, \text{cluster})$ ) in the bins A and D  
 559 (A and B) are shown in Fig. 12.

Table 7: Number of predicted background and observed events in the single DT cluster category.

	Category	A (total)	A (punch-through)	A (ABCD pred.)	B	C	D
Background-only fit	MB2	$9.5 \pm 1.9$	$6.3 \pm 1.7$	$3.1 \pm 1.3$	$4.8 \pm 1.9$	$119.2 \pm 11.5$	$76.8 \pm 8.1$
	MB3	$3.7 \pm 1.5$	$3.1 \pm 1.1$	$0.6 \pm 1.1$	$0.5 \pm 0.5$	$6.5 \pm 2.5$	$7.5 \pm 2.6$
	MB4	$1.2 \pm 0.9$	$1.2 \pm 0.9$	$0.1 \pm 0.5$	$0.06 \pm 0.22$		
Observation	MB2	9	—	—	5	119	77
	MB3	1	—	—	1	6	8
	MB4	2	—	—	0		

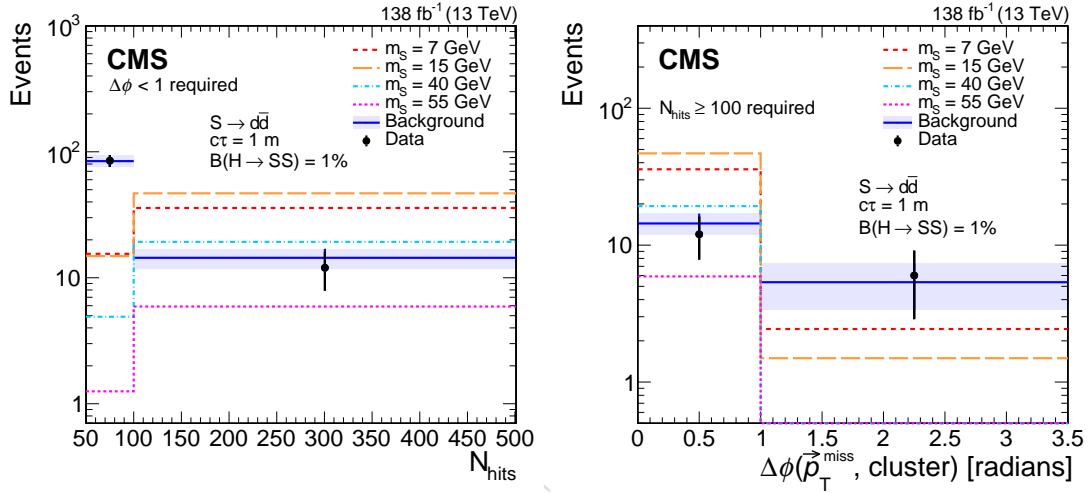


Figure 12: Distributions of  $N_{\text{hits}}$  (left) and  $\Delta\phi_c$  (right) in the search region of the single DT cluster category. The background predicted by the fit is shown in blue with the shaded region showing the fitted uncertainty. The expected signal with  $\mathcal{B}(H \rightarrow SS) = 1\%$ ,  $S \rightarrow d\bar{d}$ , and  $c\tau = 1$  m is shown for  $m_S$  of 7, 15, 40, and 55 GeV in various colors and dotted lines. The  $N_{\text{hits}}$  distribution includes only events in bins A and D, while the  $\Delta\phi_c$  one includes only events in bins A and B. The last bin in the  $N_{\text{hits}}$  distribution includes overflow events.

#### 9.4 Interpretations

In this section, we present the combination of the double cluster, single CSC cluster, and single DT cluster categories and the interpretations in the context of the twin Higgs and dark shower models. The ABCD planes of the 3 categories are defined to be mutually exclusive and in the combination, we fit the ABCD plane of each category simultaneously. Because of the better sensitivity of the double cluster category, and to minimize changes to the previously published single CSC cluster category, overlapping events are kept in only one category with the following order of precedence: double cluster, single CSC cluster, and single DT cluster category.

All theoretical uncertainties assigned to signal simulations are fully correlated. Experimental uncertainties that are not related the cluster, such as luminosity, jet energy scale, PDFs, pileup modeling, are fully correlated. Experimental uncertainties associated with cluster selections are assumed to be fully uncorrelated. All uncertainties are incorporated into the analysis via nuisance parameters and treated according to the frequentist paradigm.

We evaluate 95% confidence level (CL) upper limits on the branching fractions  $\mathcal{B}(H \rightarrow SS)$  and  $\mathcal{B}(h^0 \rightarrow \Psi\Psi)$  for both the twin Higgs and dark shower models using the modified frequentist criterion  $\text{CL}_s$  [63, 64] with the profile likelihood ratio test statistic [65]. The limits are obtained using asymptotic formulae [66].

We show the predicted number of signal events for the twin Higgs model in a few benchmark decay, mass, and lifetime scenarios in Table 8.

Table 8: Number of signal events in bin A for each category for a few benchmark signal models assuming  $\mathcal{B}(H \rightarrow SS) = 1\%$ .

LLP decay mode, mass, lifetime	CSC-CSC	DT-DT	DT-CSC	single CSC	single DT
$d\bar{d}, 7\text{ GeV}, c\tau = 1\text{ m}$	1.5	5.7	4.3	22.5	35.8
$d\bar{d}, 15\text{ GeV}, c\tau = 1\text{ m}$	4.7	13.6	11.1	32.0	46.8
$d\bar{d}, 40\text{ GeV}, c\tau = 1\text{ m}$	6.6	12.9	8.8	23.4	19.3
$d\bar{d}, 55\text{ GeV}, c\tau = 1\text{ m}$	0.5	1.4	2.1	9.8	5.9
$\tau^+\tau^-, 7\text{ GeV}, c\tau = 1\text{ m}$	0.6	1.8	1.6	14.2	22.5
$\tau^+\tau^-, 15\text{ GeV}, c\tau = 1\text{ m}$	1.7	5.2	3.9	20.1	28.9
$\tau^+\tau^-, 40\text{ GeV}, c\tau = 1\text{ m}$	3.3	4.5	3.3	21.3	17.0
$\tau^+\tau^-, 55\text{ GeV}, c\tau = 1\text{ m}$	0.3	0.9	1.0	10.6	6.0
$\pi^0\pi^0, 0.4\text{ GeV}, c\tau = 0.1\text{ m}$	0.1	0.4	0.4	6.8	19.2
$\pi^0\pi^0, 1\text{ GeV}, c\tau = 0.1\text{ m}$	0.4	1.3	1.1	11.6	30.7

The upper limits for the twin Higgs model are shown in Fig. 13 for the  $S \rightarrow d\bar{d}$ ,  $S \rightarrow \pi^0\pi^0$ , and  $S \rightarrow \tau^+\tau^-$  decay modes, as a function of  $c\tau$  for a selection of  $m_S$  values. Additional upper limits for other decay modes, including  $S \rightarrow b\bar{b}$ ,  $S \rightarrow \pi^+\pi^-$ ,  $S \rightarrow KK$ ,  $S \rightarrow b\bar{b}$ ,  $S \rightarrow e^+e^-$  and  $S \rightarrow \gamma\gamma$ , are shown in Appendix A. Due to the differences in geometric acceptance, the sensitivity at shorter proper decay lengths is dominated by the single CSC cluster category, at longer proper decay lengths by the single DT cluster category, and at intermediate proper decay lengths, where the analysis is most sensitive, by the double cluster category. The addition of the single DT and double cluster categories improves upon the previous result based on only CSC clusters [34] by a factor of 2 for  $c\tau$  above 0.2, 0.5, 2, and 18 m for LLP masses 7, 15, 40, and 55 GeV, respectively.

The upper limits as a function of both mass and  $c\tau$  for the  $S \rightarrow d\bar{d}$ ,  $S \rightarrow \pi^0\pi^0$ , and  $S \rightarrow \tau^+\tau^-$  decay modes are shown in Fig. 14 and extend as low as 10 MeV.

Finally, we consider the scenario where mixing between  $S$  and the Higgs boson causes  $S$  to inherit all the couplings of the Higgs boson evaluated at mass  $m_S$  [67]. In Fig. 15, we show the upper limits as a function of both mass and  $c\tau$  using the branching fractions for  $S$  calculated in Ref. [67]. This search provides the most stringent constraint on  $\mathcal{B}(H \rightarrow SS)$  for proper decay lengths in the range 0.04–0.4 m and above 4 m for an LLP mass of 15 GeV, in the range 0.3–0.9 m and above 3 m for an LLP mass of 40 GeV, and above 0.8 m for an LLP mass of 55 GeV. For LLP masses below 10 GeV, this search provides the most stringent constraints on LLPs decaying to particles other than muons.

The upper limits for the dark shower models are shown in Fig. 16 for the gluon, photon, and vector portals, as a function of  $c\tau$  for a selection of  $m_\Psi$  values and for scenarios with  $(x_{i\omega}, x_{i\Lambda}) = (2.5, 2.5)$ ,  $(2.5, 1)$ , and  $(1, 1)$ . The upper limits for all additional portals,  $x_{i\omega}$  and  $x_{i\Lambda}$ , are shown in Appendix A. This search sets the first LHC limits on models of dark showers produced via Higgs boson decay, and is sensitive to the branching fraction of the Higgs boson to dark quarks as low as  $10^{-3}$ .

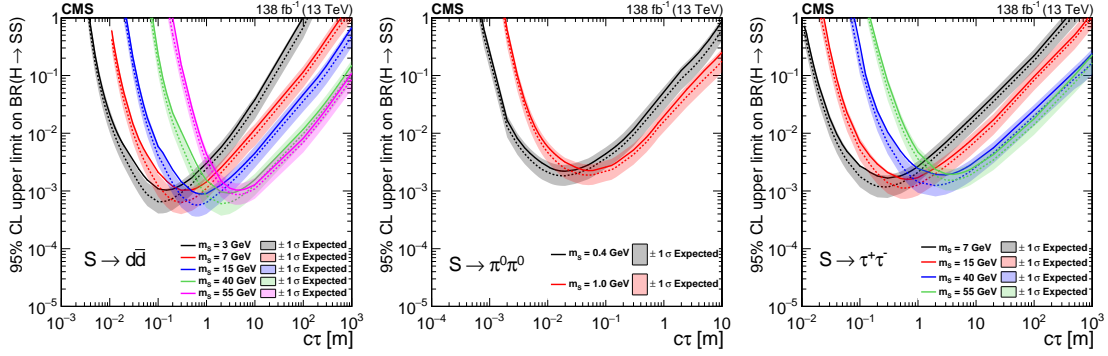


Figure 13: The 95% CL expected (dotted curves) and observed (solid curves) upper limits on the branching fraction  $\mathcal{B}(H \rightarrow SS)$  as a function of  $c\tau$  for the  $S \rightarrow d\bar{d}$  (left),  $S \rightarrow \pi^0\pi^0$  (center), and  $S \rightarrow \tau^+\tau^-$  (right) decay modes. The exclusion limits are shown for different mass hypotheses: 0.4, 1, 3, 7, 15, 40, and 55 GeV.

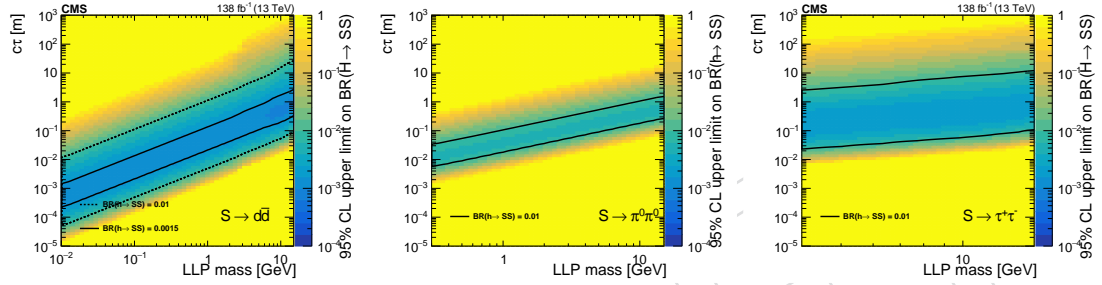


Figure 14: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(H \rightarrow SS)$  as a function of mass and  $c\tau$  for the  $S \rightarrow d\bar{d}$  (left),  $S \rightarrow \pi^0\pi^0$  (center), and  $S \rightarrow \tau^+\tau^-$  (right) decay modes.

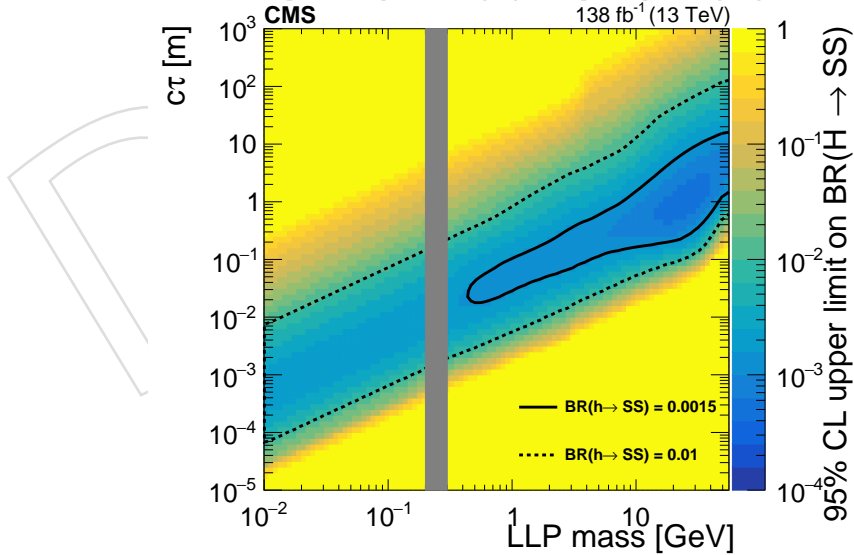


Figure 15: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(H \rightarrow SS)$  as a function of mass and  $c\tau$ , assuming the branching fractions for  $S$  calculated in Ref. [67]. We do not set limits for  $S$  mass between 0.2 and 0.3 GeV (gray region), where  $S$  decays to dimuons.



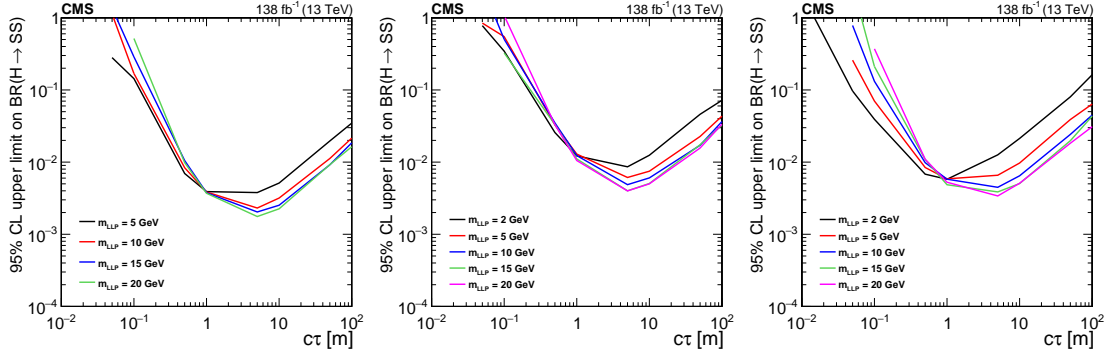


Figure 16: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(h^0 \rightarrow \Psi\Psi)$  as functions of  $c\tau$  for the gluon portal (left) assuming  $(x_{i\omega}, x_{i\Lambda}) = (2.5, 1)$ , photon portal (center) assuming  $(x_{i\omega}, x_{i\Lambda}) = (2.5, 1)$ , and vector portal (right) assuming  $(x_{i\omega}, x_{i\Lambda}) = (1, 1)$ . The exclusion limits are shown for different LLP mass hypotheses: 2, 5, 10, 15, and 20 GeV.

## 10 Summary

In summary, proton-proton collision data at  $\sqrt{s} = 13$  TeV recorded by the CMS experiment in 2016–2018, corresponding to an integrated luminosity of  $138 \text{ fb}^{-1}$ , have been used to conduct the first search that uses both the barrel and endcap CMS muon detectors to detect hadronic and electromagnetic showers from long-lived particle (LLP) decays. Based on the unique detector signature, the search is largely model-independent, with sensitivity to a broad range of LLP decay modes and masses below the GeV scale. With the excellent shielding provided by the inner CMS detector, the CMS magnet, and the steel return yoke, the background is suppressed to a low level and a search for both single and double LLP decays is possible.

No significant deviation from the standard model background is observed. The most stringent LHC constraints to date are set on the branching fraction of the Higgs boson to LLPs with masses below 10 GeV and decaying to particles other than muons. The search provides the most stringent branching fraction limit for proper decay lengths in the range 0.04–0.4 m and above 4 m for an LLP mass of 15 GeV, in the range 0.3–0.9 m and above 3 m for an LLP mass of 40 GeV, and above 0.8 m for an LLP mass of 55 GeV. Finally, the first LHC limits on models of dark showers produced via Higgs boson decay are set, and constrain branching fractions of the Higgs boson decay to dark quarks as low as  $10^{-3}$ .

## Acknowledgments

## References

- [1] G. F. Giudice and A. Romanino, “Split supersymmetry”, *Nucl. Phys. B* **699** (2004) 65, doi:10.1016/j.nuclphysb.2004.08.001, arXiv:hep-ph/0406088. [Erratum: doi:10.1016/j.nuclphysb.2004.11.048].
- [2] J. L. Hewett, B. Lillie, M. Masip, and T. G. Rizzo, “Signatures of long-lived gluinos in split supersymmetry”, *JHEP* **09** (2004) 070, doi:10.1088/1126-6708/2004/09/070, arXiv:hep-ph/0408248.
- [3] N. Arkani-Hamed, S. Dimopoulos, G. F. Giudice, and A. Romanino, “Aspects of split supersymmetry”, *Nucl. Phys. B* **709** (2005) 3, doi:10.1016/j.nuclphysb.2004.12.026, arXiv:hep-ph/0409232.

- [4] P. Gambino, G. F. Giudice, and P. Slavich, “Gluino decays in split supersymmetry”, *Nucl. Phys. B* **726** (2005) 35, doi:10.1016/j.nuclphysb.2005.08.011, arXiv:hep-ph/0506214.
- [5] A. Arvanitaki, N. Craig, S. Dimopoulos, and G. Villadoro, “Mini-split”, *JHEP* **02** (2013) 126, doi:10.1007/JHEP02(2013)126, arXiv:1210.0555.
- [6] N. Arkani-Hamed et al., “Simply unnatural supersymmetry”, 2012. arXiv:1212.6971.
- [7] 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.
- [8] G. R. Farrar and P. Fayet, “Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry”, *Phys. Lett. B* **76** (1978) 575, doi:10.1016/0370-2693(78)90858-4.
- [9] S. Weinberg, “Supersymmetry at ordinary energies. Masses and conservation laws”, *Phys. Rev. D* **26** (1982) 287, doi:10.1103/PhysRevD.26.287.
- [10] R. Barbier et al., “R-parity violating supersymmetry”, *Phys. Rept.* **420** (2005) 1, doi:10.1016/j.physrep.2005.08.006, arXiv:hep-ph/0406039.
- [11] G. F. Giudice and R. Rattazzi, “Theories with gauge mediated supersymmetry breaking”, *Phys. Rept.* **322** (1999) 419, doi:10.1016/S0370-1573(99)00042-3, arXiv:hep-ph/9801271.
- [12] P. Meade, N. Seiberg, and D. Shih, “General gauge mediation”, *Prog. Theor. Phys. Suppl.* **177** (2009) 143, doi:10.1143/PTPS.177.143, arXiv:0801.3278.
- [13] M. Buican, P. Meade, N. Seiberg, and D. Shih, “Exploring general gauge mediation”, *JHEP* **03** (2009) 016, doi:10.1088/1126-6708/2009/03/016, arXiv:0812.3668.
- [14] J. Fan, M. Reece, and J. T. Ruderman, “Stealth supersymmetry”, *JHEP* **11** (2011) 012, doi:10.1007/JHEP11(2011)012, arXiv:1105.5135.
- [15] J. Fan, M. Reece, and J. T. Ruderman, “A stealth supersymmetry sampler”, *JHEP* **07** (2012) 196, doi:10.1007/JHEP07(2012)196, arXiv:1201.4875.
- [16] M. J. Strassler and K. M. Zurek, “Echoes of a hidden valley at hadron colliders”, *Phys. Lett. B* **651** (2007) 374, doi:10.1016/j.physletb.2007.06.055, arXiv:hep-ph/0604261.
- [17] M. J. Strassler and K. M. Zurek, “Discovering the Higgs through highly-displaced vertices”, *Phys. Lett. B* **661** (2008) 263, doi:10.1016/j.physletb.2008.02.008, arXiv:hep-ph/0605193.
- [18] T. Han, Z. Si, K. M. Zurek, and M. J. Strassler, “Phenomenology of hidden valleys at hadron colliders”, *JHEP* **07** (2008) 008, doi:10.1088/1126-6708/2008/07/008, arXiv:0712.2041.
- [19] Y. Cui, L. Randall, and B. Shuve, “A WIMPy baryogenesis miracle”, *JHEP* **04** (2012) 075, doi:10.1007/JHEP04(2012)075, arXiv:1112.2704.
- [20] Y. Cui and R. Sundrum, “Baryogenesis for weakly interacting massive particles”, *Phys. Rev. D* **87** (2013) 116013, doi:10.1103/PhysRevD.87.116013, arXiv:1212.2973.

- [21] Y. Cui and B. Shuve, “Probing baryogenesis with displaced vertices at the LHC”, *JHEP* **02** (2015) 049, doi:10.1007/JHEP02(2015)049, arXiv:1409.6729.
- [22] D. Smith and N. Weiner, “Inelastic dark matter”, *Phys. Rev. D* **64** (2001) 043502, doi:10.1103/PhysRevD.64.043502, arXiv:hep-ph/0101138.
- [23] Z. Chacko, H.-S. Goh, and R. Harnik, “Natural electroweak breaking from a mirror symmetry”, *Phys. Rev. Lett.* **96** (2006) 231802, doi:10.1103/PhysRevLett.96.231802, arXiv:hep-ph/0506256.
- [24] D. Curtin and C. B. Verhaaren, “Discovering uncolored naturalness in exotic Higgs decays”, *JHEP* **12** (2015) 072, doi:10.1007/JHEP12(2015)072, arXiv:1506.06141.
- [25] H.-C. Cheng, S. Jung, E. Salvioni, and Y. Tsai, “Exotic quarks in twin Higgs models”, *JHEP* **03** (2016) 074, doi:10.1007/JHEP03(2016)074, arXiv:1512.02647.
- [26] CMS Collaboration, J. G. Layter, “The CMS muon project: Technical Design Report”. Technical design report. CMS. CERN, Geneva, 1997.
- [27] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [28] N. Craig, A. Katz, M. Strassler, and R. Sundrum, “Naturalness in the dark at the LHC”, *JHEP* **07** (2015) 105, doi:10.1007/JHEP07(2015)105, arXiv:1501.05310.
- [29] M. J. Strassler, “On the phenomenology of hidden valleys with heavy flavor”, 2008. arXiv:0806.2385.
- [30] J. E. Juknevich, D. Melnikov, and M. J. Strassler, “A pure-gluon hidden valley I. states and decays”, *JHEP* **07** (2009) 055, doi:10.1088/1126-6708/2009/07/055, arXiv:0903.0883.
- [31] CMS Collaboration, “Search for long-lived particles using displaced jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, *Phys. Rev. D* **104** (2021), no. 1, 012015, doi:10.1103/PhysRevD.104.012015, arXiv:2012.01581.
- [32] ATLAS Collaboration, “Search for long-lived particles produced in  $pp$  collisions at  $\sqrt{s} = 13$  TeV that decay into displaced hadronic jets in the ATLAS muon spectrometer”, *Phys. Rev. D* **99** (2019) 052005, doi:10.1103/PhysRevD.99.052005, arXiv:1811.07370.
- [33] ATLAS Collaboration, “Search for events with a pair of displaced vertices from long-lived neutral particles decaying into hadronic jets in the ATLAS muon spectrometer in  $pp$  collisions at  $\sqrt{s}=13$  TeV”, *Phys. Rev. D* **106** (2022), no. 3, 032005, doi:10.1103/PhysRevD.106.032005, arXiv:2203.00587.
- [34] CMS Collaboration, “Search for Long-Lived Particles Decaying in the CMS End Cap Muon Detectors in Proton-Proton Collisions at  $\sqrt{s}=13$  TeV”, *Phys. Rev. Lett.* **127** (2021), no. 26, 261804, doi:10.1103/PhysRevLett.127.261804, arXiv:2107.04838.
- [35] S. Knapen, J. Shelton, and D. Xu, “Perturbative benchmark models for a dark shower search program”, *Phys. Rev. D* **103** (2021), no. 11, 115013, doi:10.1103/PhysRevD.103.115013, arXiv:2103.01238.

- [36] CMS Collaboration, “Performance of the CMS cathode strip chambers with cosmic rays”, *JINST* **5** (2010) T03018, doi:10.1088/1748-0221/5/03/T03018, arXiv:0911.4992.
- [37] 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.
- [38] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.
- [39] 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.
- [40] 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.
- [41] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [42] E. Re, “Single-top  $Wt$ -channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [43] L. Carloni, J. Rathsmann, and T. Sjostrand, “Discerning Secluded Sector gauge structures”, *JHEP* **04** (2011) 091, doi:10.1007/JHEP04(2011)091, arXiv:1102.3795.
- [44] L. Carloni and T. Sjostrand, “Visible Effects of Invisible Hidden Valley Radiation”, *JHEP* **09** (2010) 105, doi:10.1007/JHEP09(2010)105, arXiv:1006.2911.
- [45] 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.
- [46] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [47] 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.
- [48] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [49] 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.
- [50] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

- [51] 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.
- [52] 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.
- [53] 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.
- [54] 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.
- [55] 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.
- [56] CMS Collaboration, “Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 13$  TeV using the CMS detector”, *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.
- [57] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise”, in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, p. 226. Association for the Advancement of Artificial Intelligence, 1996.
- [58] CMS Collaboration, “Missing transverse energy performance of the CMS detector”, *JINST* **6** (2011) P09001, doi:10.1088/1748-0221/6/09/P09001, arXiv:1106.5048.
- [59] CMS Collaboration, “Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, *JINST* **15** (2020) P02027, doi:10.1088/1748-0221/15/02/P02027, arXiv:1912.03516.
- [60] 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.
- [61] CMS Collaboration, “CMS luminosity measurements for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2017.
- [62] CMS Collaboration, “CMS luminosity measurements for the 2018 data-taking period at  $\sqrt{s} = 13$  TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2018.
- [63] 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.
- [64] 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.
- [65] The ATLAS Collaboration, The CMS Collaboration, The 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.



- 793 [66] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for  
794 likelihood-based tests of new physics”, *Eur. Phys. J. C* **71** (2011) 1554,  
795 doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum:  
796 doi:10.1140/epjc/s10052-013-2501-z].
- 797 [67] Y. Gershtein, S. Knapen, and D. Redigolo, “Probing naturally light singlets with a  
798 displaced vertex trigger”, *Phys. Lett. B* **823** (2021) 136758,  
799 doi:10.1016/j.physletb.2021.136758, arXiv:2012.07864.

DRAFT

## A Additional Interpretations

We present additional decay modes, including  $S \rightarrow b\bar{b}$ ,  $S \rightarrow \pi^+\pi^-$ ,  $S \rightarrow K^+K^-$ ,  $S \rightarrow b\bar{b}$ ,  $S \rightarrow e^+e^-$  and  $S \rightarrow \gamma\gamma$ , for twin Higgs model. The upper limits as a function of  $c\tau$  for a selection of values of  $m_S$  are shown in Fig. 17. The upper limits as a function of both  $m_S$  and  $c\tau$  are shown in Fig. 18.

We also show additional upper limits for dark shower models in this section. For each portal, we show the upper limits assuming  $(x_{i\omega}, x_{i\Lambda}) = (1, 1)$  (left),  $(x_{i\omega}, x_{i\Lambda}) = (2.5, 1)$  (center), and  $(x_{i\omega}, x_{i\Lambda}) = (2.5, 2.5)$  (right) and for different mass and proper decay lengths hypotheses. The 95% CL upper limits are shown for the gluon portal in Fig. 19, photon portal in Fig. 20, Higgs portal in Fig. 21, and dark photon portal in Fig. 22.

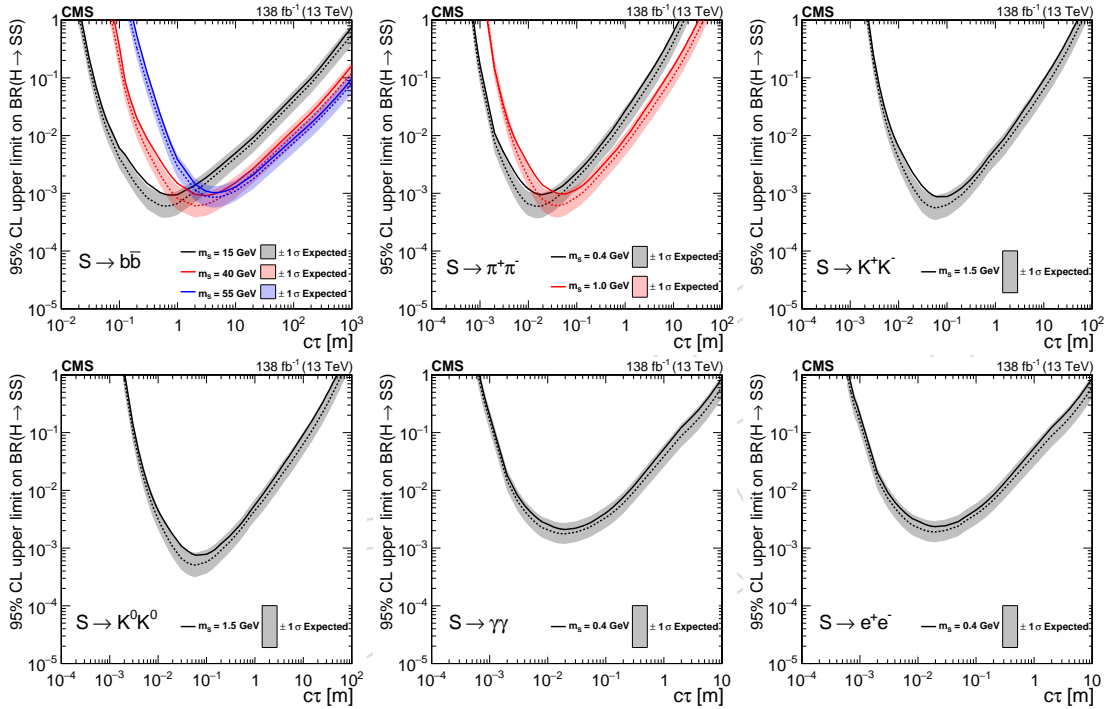


Figure 17: The 95% CL expected (dotted curves) and observed (solid curves) upper limits on the branching fraction  $\mathcal{B}(H \rightarrow SS)$  as functions of  $c\tau$  for the  $S \rightarrow b\bar{b}$  (top left),  $S \rightarrow \pi^+\pi^-$  (top center), and  $S \rightarrow K^+K^-$  (top right),  $S \rightarrow K^0\bar{K}^0$  (bottom left),  $S \rightarrow \gamma\gamma$  (bottom center),  $S \rightarrow e^+e^-$  (bottom right) decay modes. The exclusion limits are shown for different mass hypotheses: 0.4, 1, 3, 7, 15, 40, and 55 GeV.

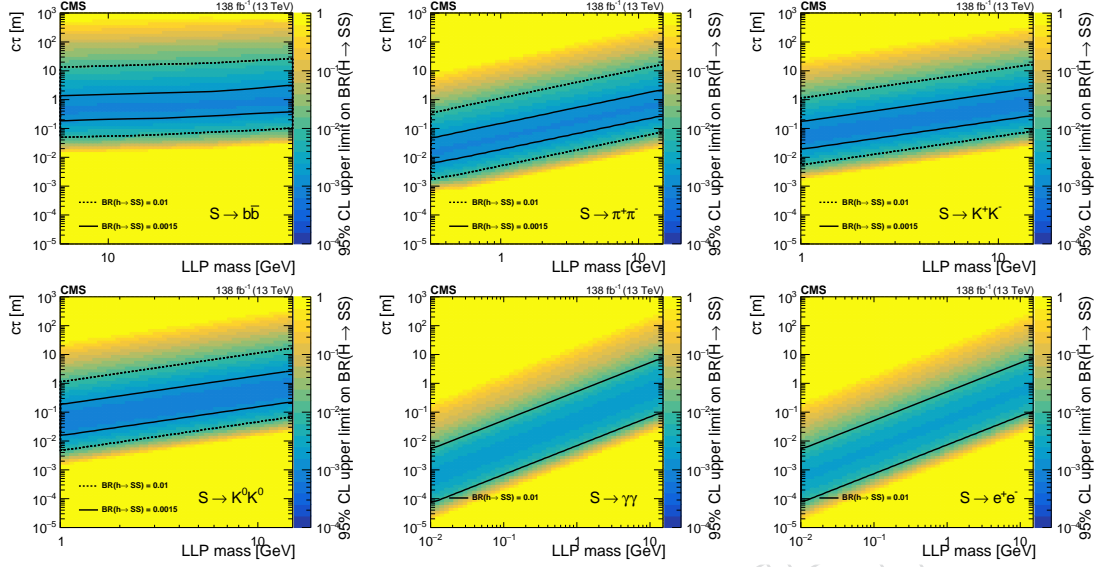


Figure 18: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(H \rightarrow SS)$  as a function of mass and  $c\tau$  for the  $S \rightarrow b\bar{b}$  (top left),  $S \rightarrow \pi^+\pi^-$  (top center), and  $S \rightarrow K^+K^-$  (top right),  $S \rightarrow K^0\bar{K}^0$  (bottom left),  $S \rightarrow \gamma\gamma$  (bottom center),  $S \rightarrow e^+e^-$  (bottom right) decay modes. and  $S \rightarrow \tau^+\tau^-$  (right) decay modes.

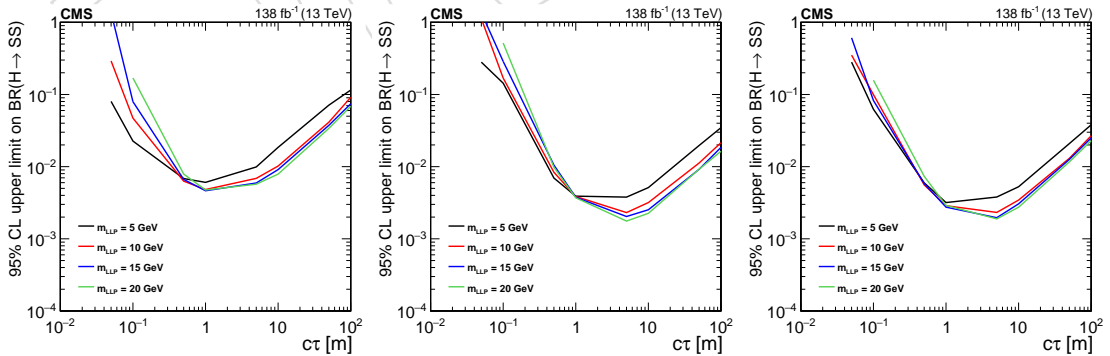


Figure 19: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(h^0 \rightarrow \Psi\Psi)$  as functions of  $c\tau$  for the gluon portal, assuming  $x_{i\omega} = x_{i\Lambda} = 1.0$  (left),  $x_{i\omega} = 2.5, x_{i\Lambda} = 1.0$  (center), and  $x_{i\omega} = x_{i\Lambda} = 2.5$  (right). The exclusion limits are shown for different mass hypotheses: 5, 10, 15, and 20 GeV. The center panel is shown again in the appendix, for completeness.

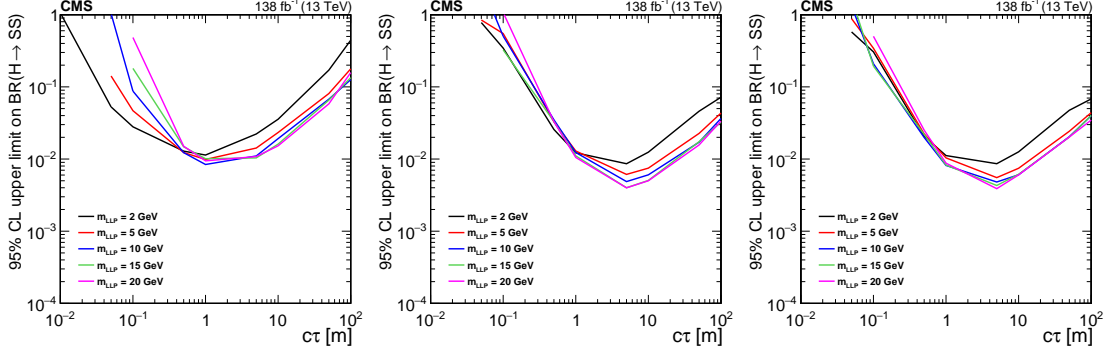


Figure 20: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(h^0 \rightarrow \Psi\Psi)$  as functions of  $c\tau$  for the photon portal, assuming  $x_{i\omega} = x_{i\Lambda} = 1.0$  (left),  $x_{i\omega} = 2.5, x_{i\Lambda} = 1.0$  (center), and  $x_{i\omega} = x_{i\Lambda} = 2.5$  (right). The exclusion limits are shown for different mass hypotheses: 2, 5, 10, 15, and 20 GeV. The center limit plot is shown again in the appendix, for completeness.

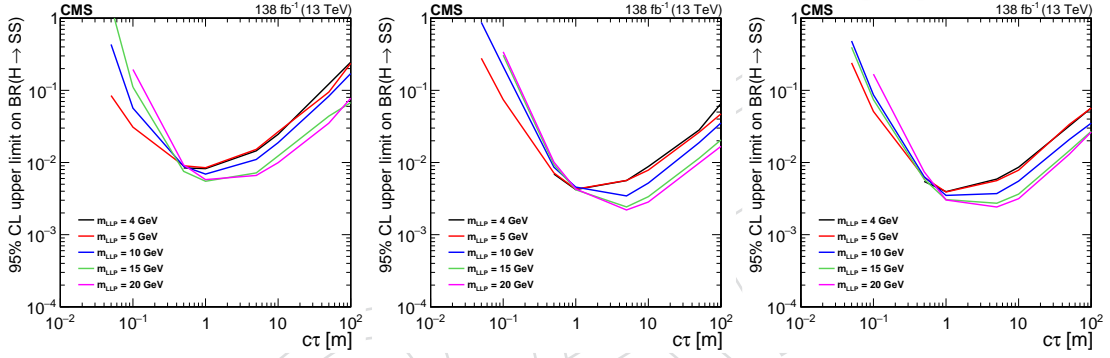


Figure 21: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(h^0 \rightarrow \Psi\Psi)$  as functions of  $c\tau$  for the Higgs portal, assuming  $x_{i\omega} = x_{i\Lambda} = 1.0$  (left),  $x_{i\omega} = 2.5, x_{i\Lambda} = 1.0$  (center), and  $x_{i\omega} = x_{i\Lambda} = 2.5$  (right). The exclusion limits are shown for different mass hypotheses: 4, 5, 10, 15, and 20 GeV.

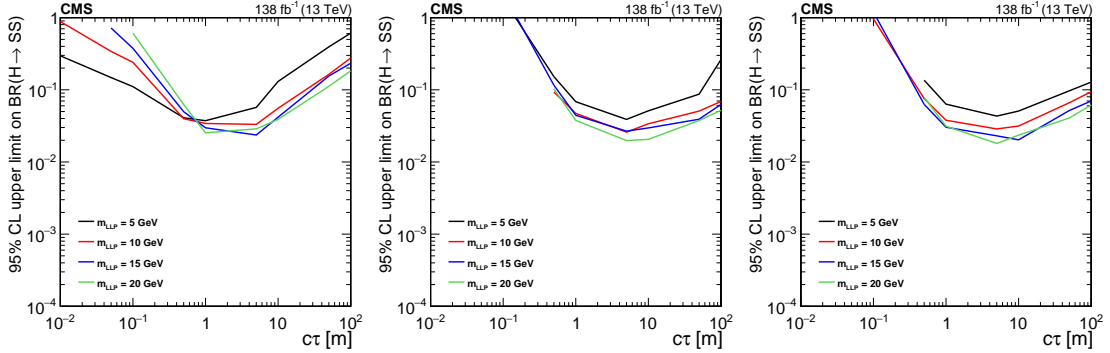


Figure 22: The 95% CL observed upper limits on the branching fraction  $\mathcal{B}(h^0 \rightarrow \Psi\Psi)$  as functions of  $c\tau$  for the dark photon portal, assuming  $x_{i\omega} = x_{i\Lambda} = 1.0$  (left),  $x_{i\omega} = 2.5, x_{i\Lambda} = 1.0$  (center), and  $x_{i\omega} = x_{i\Lambda} = 2.5$  (right). The exclusion limits are shown for different mass hypotheses: 5, 10, 15, and 20 GeV.