CMS PAPER EXO-23-005

DRAFT
CMS Paper

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

2024/05/22
Archive Hash: 49f809f
Archive Date: 2024/05/22

Dark sector searches with the CMS experiment

The CMS Collaboration

Abstract

Astrophysical observations provide compelling evidence for gravitationally interacting dark matter in the universe that cannot be explained by the standard model of particle physics. The extraordinary amount of data from the CERN LHC presents a unique opportunity to shed light on the nature of dark matter at unprecedented collision energies. This Report comprehensively reviews the most recent searches with the CMS experiment for particles and interactions belonging to a dark sector and for dark-sector mediators. Models with invisible massive particles are probed by searches for signatures of missing transverse momentum recoiling against visible standard model particles. Searches for mediators are also conducted via fully visible final states. The results of these searches are compared with those obtained from direct-detection experiments. Searches for alternative scenarios predicting more complex dark sectors with multiple new particles and new forces are also presented. Many of these models include long-lived particles, which could manifest themselves with striking unconventional signatures with relatively small amounts of background. Searches for such particles are discussed and their impact on dark-sector scenarios is evaluated. Many results and interpretations have been newly obtained for this Report.

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

PDFAuthor:	The EXO Community
PDFTitle:	Physics of Dark Sectors in CMS
PDFSubject:	CMS
PDFKeywords:	CMS, Dark Matter, Dark Sectors, BSM

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

Contents

1 Introduction

There is strong astrophysical and cosmological evidence that dark matter (DM) exists and makes up approximately 26% of the total mass-energy budget of the universe [1, 2]. This evidence is based on numerous observations of its gravitational interaction on galactic scales. The rotation curves of most galaxies do not match the expected behavior from visible matter [3, 4]. Recently, several galaxies have been observed whose rotation curves do match the expectation [5, 6], suggesting DM is unevenly distributed. Strong lensing observations of galaxy cluster collisions [7] and weak gravitational lensing from large-scale structures [8] both indicate the presence of DM at super-galactic scales. Accurate modeling of the cosmic microwave background power spectrum [1] and the matter power spectrum of the universe [9, 10] requires the presence of DM. Various scenarios beyond the standard model (BSM) that contain DM particle candidates may also resolve discrepancies in the standard model (SM), such as the predictions for light-element abundances from Big Bang nucleosynthesis [11].

A range of complementary approaches [12] can study potential interactions of DM particles with the SM. Direct-detection (DD) experiments directly probe DM scattering from ordinary matter, usually nuclei. The search for such scattering is the basis of experiments such as XENON [13], LUX-ZEPLIN [14], PandaX [15], PADME [16], and others (a review can be found in Ref. [17]). This approach is very sensitive to low values of the scattering cross section, down to the zeptobarn scale, but may face difficulties detecting DM-lepton interactions or light DM particles ($\lesssim$1 GeV in mass). These difficulties arise from the fact that the liquid xenon and liquid argon energy resolutions are poor for low-energy recoils. To probe low recoil energies, different technologies are needed. Conversely, the indirect-detection (ID) approach looks for signals of DM-DM annihilation into SM particles, which are being searched for by experiments such as AMS-02 [18], EGRET [19], Fermi-LAT [20], and IceCube [21]. This approach is sensitive to the coupling of DM to SM particles, while also probing the nature of the DM-DM annihilation process that plays a fundamental role in the observed thermal relic density. The main difficulty is the need for accurate modeling of the astrophysical background sources and of the DM density profile in the region of interest. There also exist beam-dump experiments that could potentially produce DM [22], which are beyond the scope of this Report.

Many BSM scenarios predict the existence of a *dark sector* (DS) that can be probed with proton-proton (pp) and heavy ion (HI) collisions in the CMS experiment at the CERN LHC. At particle colliders, searches for DM often involve the production of a pair of DM candidates, leading to a signature of missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) recoiling against an SM particle. Simplified benchmark models have been put forward by the community to guide these searches [23], together with recommendations on the presentation of experimental results [24] and guidelines for the comparison between the collider and DD/ID experiments [25]. These benchmark models usually have a DM candidate and a mediator particle, which may also be a BSM state. Collider searches generally present their results in terms of the masses and spins of both of these particles. As will be shown in this Report, the collider approach can provide sensitivity that is complementary to those of the DD and ID experiments. In the particular case of simplified models, certain assumptions on the mediator couplings to both SM and DM particles allow us to compare collider and DD searches. Given those assumptions, the collider experiment limits are usually stronger than the limits from other approaches for lighter DM particles (masses down to a few GeV) and for models where the nuclear interaction is spin dependent.

Going beyond the simplified-model picture entails the construction of an extended DS of particles, based on concepts such as weak-scale supersymmetry (SUSY) [26], extra dimensions [27], or extended scalar sectors [28]. An alternative approach is to hypothesize that these new particles are neutral under all the SM charges: electric, weak, and color. This new DS can have rich dynamics with previously unexplored signatures [29] that are now the target of dedicated searches by the CMS Collaboration. In this Report, we review CMS DS searches, using the Run 2 pp and HI collision data sets collected by the CMS detector from 2016–2018, or, in some cases, using data sets from Run 1 or Run 3, collected in 2010–2012 and 2022, respectively.

The relationship between theoretical models and observable final states is complex and nontrivial. We consider both perspectives to organize this Report on the overall collider effort to search for DM, as presented in this Report and depicted in Fig. 1. We begin by presenting the theoretical framework of the DM models used for CMS DM analyses in Section 2. Subsequently, we discuss the experimental apparatus and the event reconstruction in Section 3 and the experimental challenges that are common in these searches in Section 4. The data and simulation used are described in Section 5, and the final-state signatures probed by each CMS DM analysis are detailed in Section 6. Finally, we present the results and their reinterpretations in the context of the theoretical framework in Section 7, and we summarize this Report in Section 8.

2 Theoretical framework

There are numerous proposed models accessible in high-energy collisions that include new particles satisfying the cosmological and astrophysical constraints for a DM candidate. Dark matter searches at the LHC, therefore, are characterized by final states that include a DM particle or are otherwise consistent with a BSM scenario that can produce DM candidates.

In addition to the DM particle, every model includes an additional sector, called a "portal", that couples SM particles to DM particles. In most DM models, this portal consists of a new mediator particle. However, the portal can also include a Z boson or Higgs boson (H) with couplings modified to include the possibility of DM decays. Direct DM signatures, in their simplest form, consist of the production of the mediator particle, which subsequently decays into DM. Final states from such processes feature the presence of $p_{\mathrm{T}}^{\mathrm{miss}}$ because the DM particles interact

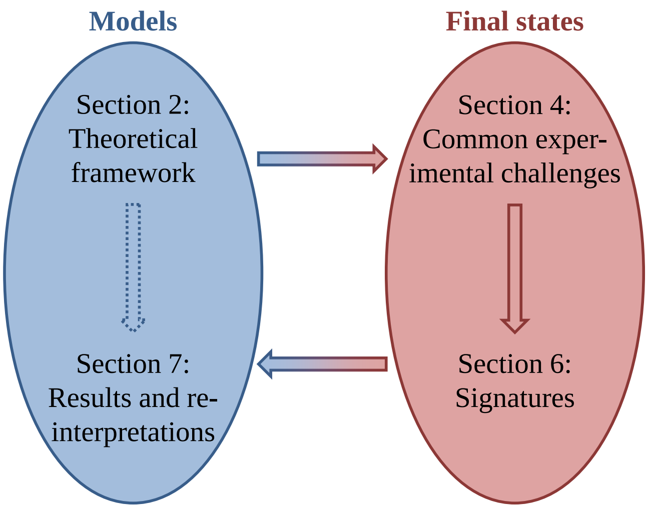

Figure 1: An outline of the paper organization in terms of theoretical models and observable final states and how the two perspectives are related.

sufficiently weakly to be invisible in the detector. To be detectable, the DM particle must be accompanied by at least one visible object, such as a jet, lepton, photon, or the decay products of a heavy SM boson, such as the Higgs, W, or Z boson. These characteristic signatures are the mainstay of "mono-X" searches, where X denotes the visible, radiated object that recoils off the system that directly produces the DM. In this Report, the DM particle is generally assumed to be a Dirac fermion, unless otherwise stated; however, this does not preclude sensitivity to other DM spin states.

Any mediator that is produced at colliders by the interaction of SM particles must also be able to decay back to those SM particles. Correspondingly, we can also search for the DM indirectly via fully visible resonances arising from the mediator production. This approach is only sensitive to the SM interactions of the mediator and therefore makes no additional assumptions about the portal. However, accessing different resonant mass ranges may require different search strategies at colliders, as discussed in subsequent sections.

Different signatures appear when the DS dynamics are modified such that there may be more mediators, additional unstable particles, or new interactions. These extended DM models give rise to a number of signatures that can be probed at the LHC. Moreover, these added signatures enhance the sensitivity of the LHC to the DS with additional visible particles and energy in the final state, compared to mono-X searches.

In this section, we present the scope of DS models probed with the CMS experiment. We classify the models into two categories: models that consist of a single mediator particle and DM are denoted *simplified DSs* and are discussed in Section 2.1, and models with more complicated DS dynamics are denoted *extended DSs* and are discussed in Section 2.2. Figures 2 and 3 give an overview of the models probed in CMS searches, which are explained in the following.

2.1 Simplified dark sectors

Originally, the exploration of the DS proceeded using an effective field theory (EFT) approach, with a single parameter Λ [30–33]. This parameter defines either the coupling strength or the interaction scale, which cannot be disentangled. Therefore, bounds on the DM production cross section are presented in terms of Λ, using a prescribed fixed coupling, when compared to noncollider experimental results. However, the higher energies at the LHC allow for exploring

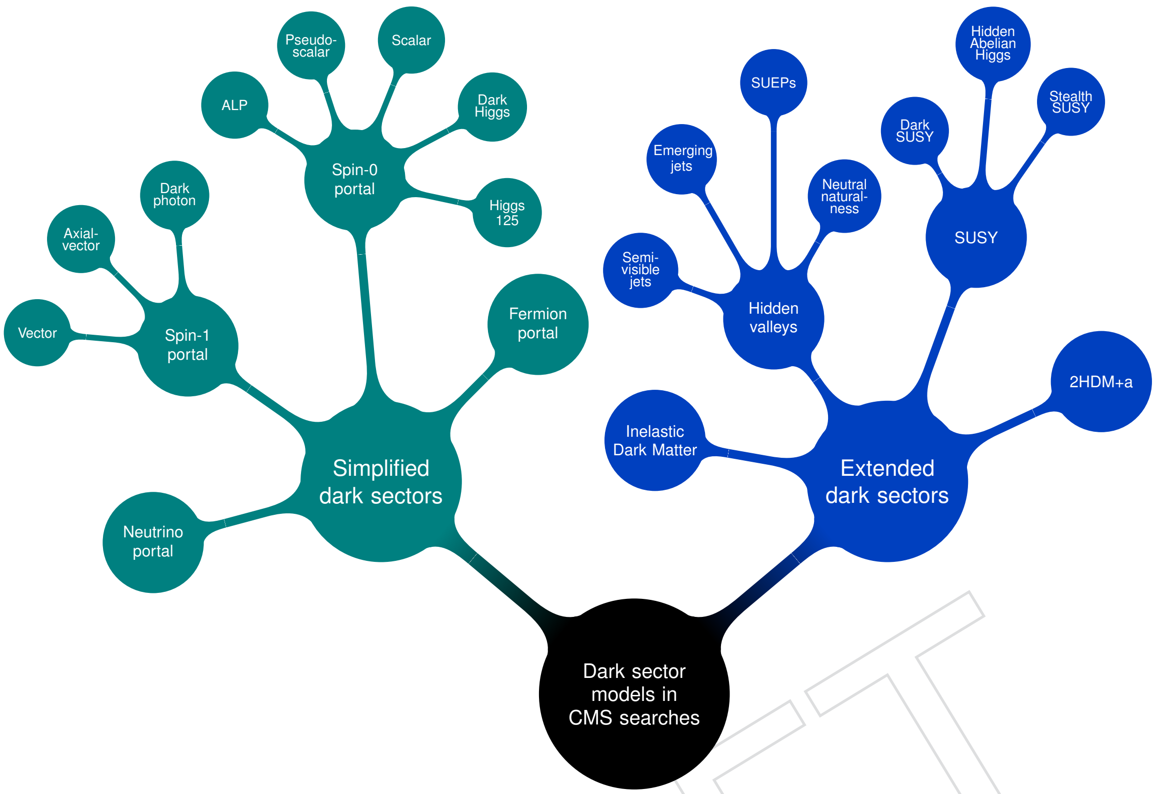

Figure 2: Map of the models probed in CMS searches for dark sectors.

more physical features that are not captured by EFT models because they are valid only for momentum transfers much smaller than the scale of the interaction. Therefore, they have largely been superseded by two classes of DM models: simplified models and DS models, the latter of which is also known as "feebly interacting particle" (FIP) models [23, 25, 34–49]. Results interpreted with EFTs will not be discussed further in this Report.

The simplified models were developed explicitly to compare LHC results with those from DD and ID searches, while the DS models were developed to facilitate comparisons with beam dump experiments targeting light DSs. These two classes largely overlap and methods exist to interpret results from one class for the other class.

For both the simplified models and the DS models, there is a framework that connects the DS with the visible sector through a mediator. The existence of a mediator resolves the limitations of EFTs, which can yield unphysical distributions because of the lack of a mediator. The mediator enables resonant production and a physical production mechanism but also adds complexity because several other parameters need to be scanned to produce interpretable results. Moreover, aspects such as renormalizability and ultraviolet completion are typically not taken into account. Despite these shortcomings, established and reliable schemes exist to present the results, and established models exist that aim to cover a variety of mediators and DM interactions.

In this Section, we present both classes together, highlighting differences when needed [17, 50]. To ensure broad coverage, four separate categories of portals are commonly utilized. These models are classified by the spin and the properties of the portal:

- **Spin-1 portal:** This category of models (Section 2.1.1) has a spin-1 mediator that couples to the SM with couplings that are uniform across flavors but deviates by

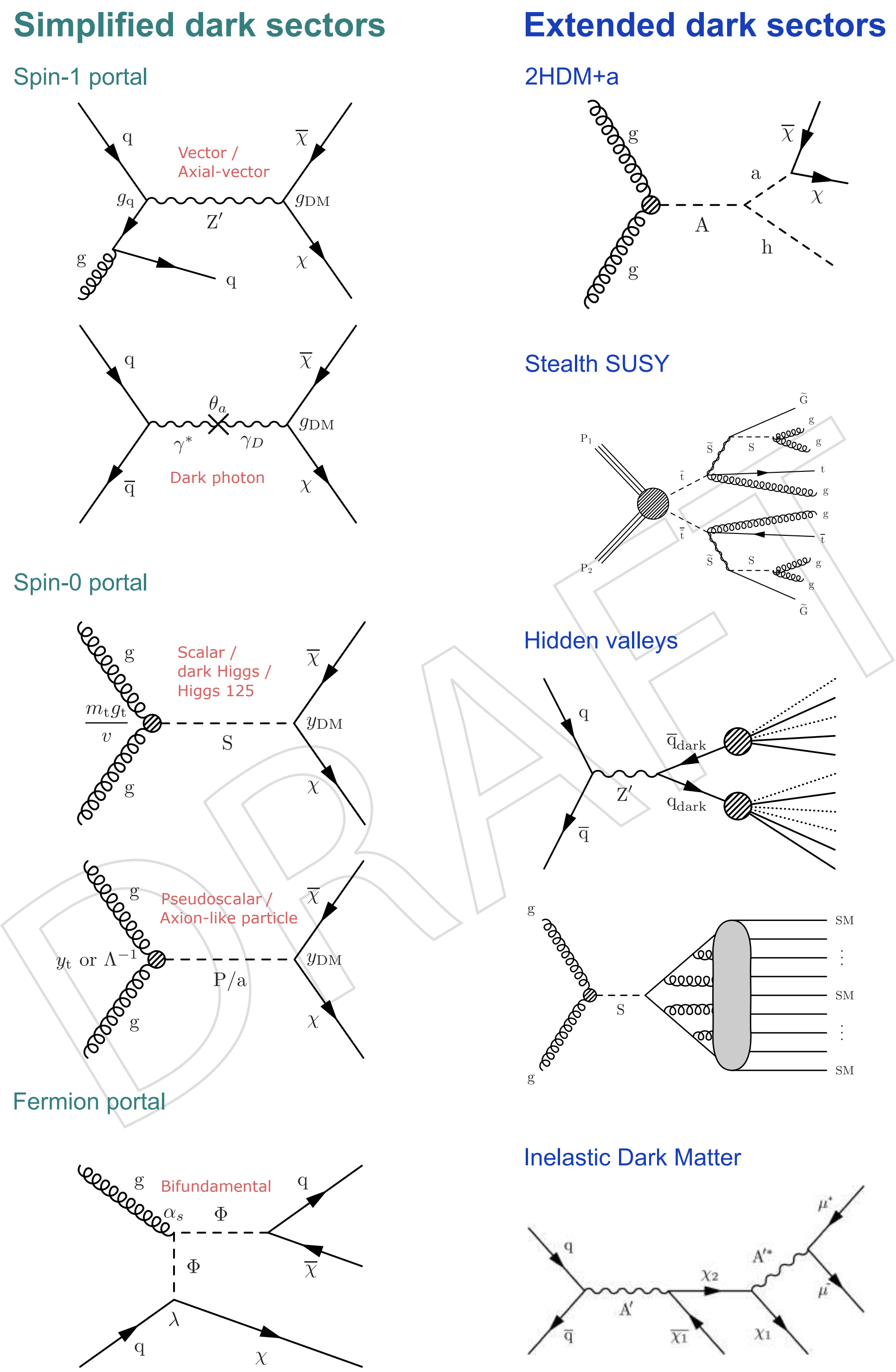

Figure 3: Example Feynman diagrams in the taxonomy of dark sector models.

particle type (leptons and quarks can have different couplings). With simplified models, a minimal model with only quark couplings is taken as the baseline, and both a pure vector and pure axial-vector coupling are allowed. In FIP models, the spin-1 mediator is assumed to mix with the Z boson, yielding a dark-photon model.

- **Spin-0 portal:** This category of models (Section 2.1.2) has a scalar or pseudoscalar particle as the mediator. The simplified model assumes the scalar particle does not mix with the Higgs boson. In the FIP models, the scalar portal mediator mixes with the Higgs boson and is often referred to as the dark-Higgs or the Higgs portal mediator (H_D). The FIP version of the mediator of the axion (a) portal is often referred to as an axion-like particle (ALP), which can be equated with the pseudoscalar mediator in the simplified model.
- **Neutrino portal:** This category of models (Section 2.1.3) includes a heavy neutral lepton (HNL), which often takes the form of a right-handed neutrino. Only one common model exists for this portal.
- **Fermion portal:** This category of models (Section 2.1.4) includes a scalar mediator Φ with a Yukawa coupling between DM and SM fermions, which allows *t*-channel interactions.

In the following subsections, we present each model from the above list. Where required, we discuss the differences between the simplified and FIP versions of the models and how to reinterpret the bounds on these models. Figure 4 shows representative diagrams for each theoretical model addressed in this Report. Note that there are no diagrams for HNL models, as these models are the subject of their own Report [51].

In order to provide constraints that are applicable to a wide range of scenarios, the analyses discussed in this Report are often interpreted using additional simplified models in which the branching fractions to exotic particles, long-lived particle (LLP) lifetimes, and final states are fixed independently of any theoretical or experimental constraints. This allows the results of these searches to be reinterpreted using both the models discussed below and complete DS models.

2.1.1 Spin-1 portal

This section discusses both commonly used spin-1 portal models, the Z′ portal and the dark photon. In addition to presenting both models, we discuss how results can be re-interpreted between the two. In both cases, the couplings are assumed to be uniform with respect to flavor. Despite that, flavor-specific spin-1 mediators do exist in the literature. These include models that motivate an explanation for the observed deviations in the anomalous magnetic moment of the muon [52]. These models are typically reinterpreted from the flavor symmetric bounds and are not extensively discussed further (more details can be found in Refs. [53–63]).

2.1.1.1 Vector and axial-vector portal

A vector mediator arises from a broken $U(1)$ symmetry with couplings to both the SM and the DS. These couplings can be strictly vector or axial-vector in nature, and they are typically assumed to be universal for each type of matter particle. The interaction terms in the Lagrangian for a vector Z′ boson are given by:

$$\mathcal{L}_{\text{vector}} \supset -g_{\text{DM}} Z'_{\mu} \bar{\chi} \gamma^{\mu} \chi - g_{\text{q}} \sum_{\text{q}} Z'_{\mu} \bar{\text{q}} \gamma^{\mu} \text{q} - g_{\ell} \sum_{\ell} Z'_{\mu} \bar{\ell} \gamma^{\mu} \ell, \tag{1}$$

where ℓ are the leptons; χ is the DM field; and g_{DM}, g_{q}, and g_{ℓ} are the couplings of the Z′ boson

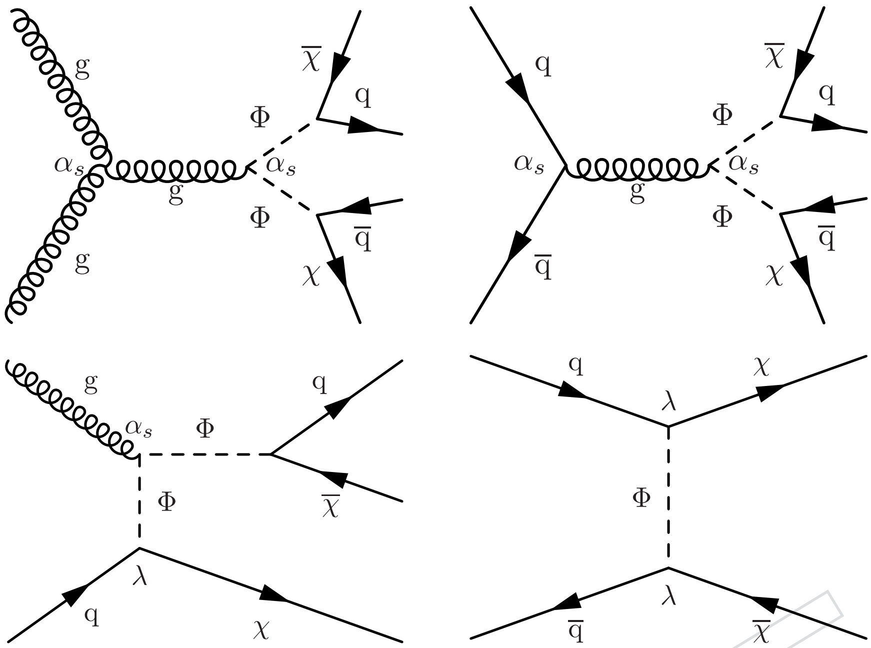

Figure 6: Feynman diagrams for production channels involving the bifundamental mediator Φ: pair production via gluon-gluon fusion (upper left), pair production via quark-antiquark annihilation (upper right), single production in association with a DM particle χ (lower left), and *t*-channel nonresonant DM production (lower right).

2.2 Extended dark sectors

Many models with complex dynamics in the DS have been theorized. They potentially communicate with the SM through any of the portals described above. Extended models of DM typically incorporate more than a single particle species in the DS, in contrast to, for example, minimal models that feature WIMPs. The additional states can give rise to enriched dynamics in the DS, with potential relevant experimental footprints in pp collision events. Specific cases motivating CMS searches are highlighted in the following sections.

2.2.1 A 2HDM-type complete model: 2HDM+a

Recently, a new class of ultraviolet-complete models has been developed with a focus on DM. One of those models is an extension of the existing two-Higgs-doublet models (2HDM) [116, 117], which adds an additional spin-0 (pseudoscalar) mediator along with a DM particle candidate. Thus, it is described as the 2HDM plus a pseudoscalar (2HDM+a). The Lagrangian of such a model is described in Ref. [118].

The interaction between the DM candidates and the SM particles is achieved by incorporating interaction terms between the 2HDM Higgs doublet fields ($h_{1,2}$) and the newly introduced pseudoscalar mediator field (P). This interaction generates a mixing between the CP-odd pseudoscalar mediator and the particles present in the 2HDM, which in turn allows for SM interactions. This yields a nondiagonal mass matrix, of which one mass eigenstate corresponds to the mediator (a), while the other eigenstates correspond to the CP-odd Higgs boson (A) and the other 2HDM fields (H, h, $H^{\pm}$). The latter fields also acquire couplings of different kinds with the mediator via the trilinear and quartic couplings introduced in the scalar potential. We follow the convention that the heaviest neutral Higgs boson in a particular model is represented by H and other neutral scalar bosons, if any, are represented by h. The DM particle nature is characterized by the Dirac fermion field χ, which couples to the two CP-odd states and whose respective coupling strengths are controlled by the mixing angle of the CP-odd sector. The Yukawa sector is taken to be the same as in the usual 2HDMs, where the structure is selected to avoid the appearance of flavor changing neutral currents. This often results in four possible configurations in terms of scalar and fermions couplings, labeled as scenarios of type: I, II, III, and IV. For the purpose of this publication, we focus our attention on the type-II scenario, where there is a differentiated interaction between the scalars and fermions for up-type and down-type quarks [116].

After electroweak symmetry breaking, the dynamics is determined by 14 parameters: v, m_{h}, m_{H}, m_{A}, $m_{\mathrm{H}^{\pm}}$, m_{a}, m_{DM}, $\cos(\beta-\alpha)$, $\tan\beta$, $\sin\theta$, y_{DM}, λ_3, λ_{P_1}, λ_{P_2}. This number is typically reduced when the existing theoretical and experimental constraints are imposed on the model. The usual theoretical constraints resulting from the unitarity and perturbativity considerations apply. Among the experimental constraints are the measurements of the properties of the SM Higgs boson, and EW precision and flavor physics observables. A more detailed description of the list of the constraints restricting the parameter phase space is given in Ref. [119], where the following benchmark parameter choices are motivated, which will be used for most results in this Report:

$$\begin{aligned} &m_{\mathrm{H}} = m_{\mathrm{A}} = m_{\mathrm{H}^{\pm}},\;\; m_{\mathrm{DM}} = 10\,\mathrm{GeV}, \\ &\cos(\beta-\alpha) = 0,\;\; \tan\beta = 1,\;\; \sin\theta = 0.35, \\ &\lambda_3 = \lambda_{\mathrm{P}_1} = \lambda_{\mathrm{P}_2} = 3,\;\; y_{\mathrm{DM}} = 1. \end{aligned} \tag{16}$$

Given the above-mentioned items and the natural complexity presented by the 2HDM+a framework, there exists very rich phenomenology in both the Higgs and dark sectors. A large number of signatures can be naturally produced in the 2HDM+a [119]. In the context of DM, the list includes resonant mono-X production, where 'X' stands for a heavy SM particle (H, Z, or W boson) recoiling against DM particles, nonresonant production of SM particles accompanied by DM such as the case of monojet and heavy-flavor quarks produced in association with DM, and many others. An example Feynman diagram is shown in the left panel of Fig. 7. Being a 2HDM-type, conventional signatures of heavy resonances decaying into SM particles can be copiously produced in the 2HDM+a context. Decays of the neutral scalar states to a pair of top quarks become important when their mass is above the $\mathrm{t\bar{t}t\bar{t}}$ threshold, which can produce signatures containing either two or four top quarks if one considers the gluon-gluon fusion and the $\mathrm{t\bar{t}}$-associated production modes of the resonance, as shown in the middle panel of Fig. 7. Other cases such as $\mathrm{A/a} \to \tau^+\tau^-$, though still present, have reduced production rates in this model because of the heavy competition with the dark channel $\mathrm{A/a} \to \chi\bar{\chi}$. In that sense, signatures involving the decay of the SM-like Higgs boson (h) are more diversified in this model. For very low m_{a}, the exotic decay $\mathrm{h} \to \mathrm{aa}$ is possible, which can involve invisible, semivisible, and visible final states; an example diagram is shown in the right panel of Fig. 7. A more comprehensive discussion of the various decay channels, covering not only the neutral scalar sector but also the charged resonances, is given in Ref. [118].

2.2.2 Supersymmetry

Many SUSY models predict a lightest supersymmetric particle that is a good candidate for DM [120–122]. These SUSY models include the minimal supersymmetric SM [123–126], gauge-mediated SUSY breaking (GMSB) [127, 128], *R*-parity violating (RPV) SUSY [129], and split SUSY [130]. In this Report, we will focus on just a few SUSY models, as described below.

2.2.2.1 Hidden Abelian Higgs model (HAHM)

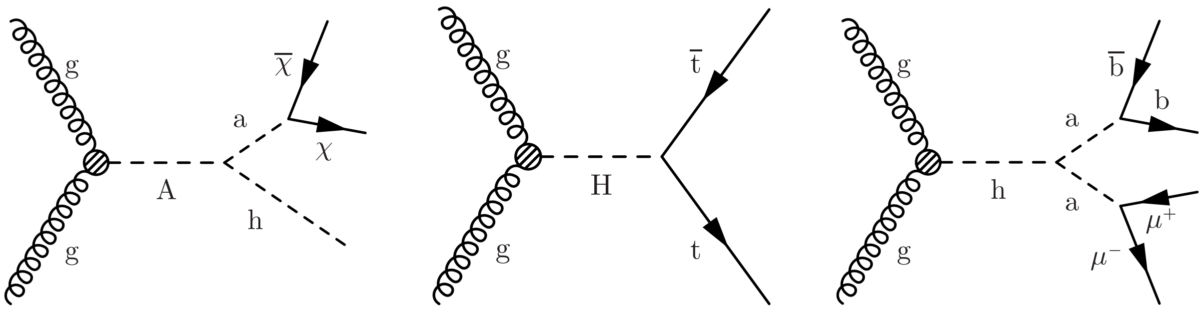

Figure 7: Feynman diagrams for 2HDM+a signatures. Left: a mono-Higgs signature, mediated by the heavy pseudoscalar A. Center: $t\bar{t}$ resonant production, mediated by the heavy scalar H. Similar processes involve $H^{\pm}$ particles, e.g. $H^{\pm} \rightarrow tb$. Right: exotic decay of the SM-like Higgs boson h.

The hidden Abelian Higgs model (HAHM) is an extension of the SM based on the group $G = SU(3)_C \times SU(2)_L \times U(1)_Y \times U(1)_\chi$. The extra $U(1)_\chi$ gauge group is added to the SM. The only coupling of this new gauge sector to the SM is through kinetic mixing with the hypercharge gauge boson. An Abelian hidden sector is coupled to the SM, and the resulting new Higgs boson and the neutral gauge boson fields are allowed to mix with the corresponding SM fields [131]. In the HAHM model, the production of long-lived dark photons is via the Higgs portal, through the mixing of the SM and dark Higgs bosons (H–H_D) via a parameter κ, with the subsequent decays via the vector portal [132].

2.2.2.2 Dark supersymmetry

Dark matter is naturally embedded in extensions of the SM motivated by solving the hierarchy problem, particularly with low-energy SUSY [133]. A hidden gauge symmetry $U(1)_D$ is broken near the GeV scale, giving rise to new dark vector bosons. A completely generic prediction is that those new bosons can be produced in cascade decays of the minimal supersymmetric SM superpartners. The lightest GeV-scale dark Higgs bosons and gauge bosons eventually decay back into light SM states, and dominantly into leptons. In this scenario, the next-to-lightest SUSY particle decays into the lightest SUSY particle in the DS, which escapes detection, plus a dark photon (A') that decays into leptons with a sizeable branching fraction. The dark-photon decay can occur promptly or after traveling some distance producing a displaced vertex (DV). Regardless of the DVs, the lepton pairs will have a small mass O(GeV), and in typical decays, will come out with small angular separation. Thus, one can produce “lepton jets”, which are boosted groups of collimated leptons with small masses. The presence of lepton jets dramatically reduces backgrounds and probes direct EW production at higher masses.

2.2.2.3 Stealth supersymmetry

Supersymmetric models with R-parity conservation often include a neutralino as the lightest SUSY particle, which makes it a candidate for weakly interacting massive particle (WIMP) DM. Searches at the LHC have placed strong constraints on these models, which has prompted interest in scenarios that could have evaded detection. One example of such a new scenario is the extension of the usual minimal supersymmetric SM particle content with a dark “stealth” sector [134–136], containing in the minimal case a scalar singlet S and its fermionic superpartner the singlino $\widetilde{S}$. There are multiple options for communication between the SM and the stealth sector, including the Higgs portal via mixing and a new vector-like $SU(5)$ messenger.

In these models, the portal between the stealth sector and the SUSY breaking sector is suppressed, such that SUSY is approximately conserved and the S and $\widetilde{S}$ are nearly mass degenerate. These stealth sector particles are not stable. Once produced, the singlino decays into the singlet and a stable DM particle. The stable DM particle is often assumed to be a gravitino ($\widetilde{\mathrm{G}}$), but it could also be an axino. In both cases, the stable DM is typically assumed to be light in these models, of order 1 GeV. Depending on the size of the mass splitting and the involved couplings, the singlino can be long lived. If it is long lived on cosmological scales, it can be a viable DM candidate and results in co-decaying DM [137, 138], which is a mechanism for thermal DM freeze-out where degenerate particles in complex DSs and out-of-equilibrium decays can both decay to obtain the observed relic density. The singlet decay depends on the assumed portal between the stealth sector and the SM. In the case of the Higgs portal and singlet masses of order 100 GeV, the decay is predominantly to two bottom quarks, whereas in the case of the vector portal, the decay is predominantly to two gluons.

At the LHC, the stealth sector particles are assumed to be produced in the decay of a SUSY particle, such as a squark. Between the many options for the production channel and possibilities for the interaction portal, the phenomenology of these models is varied. Importantly, the small assumed DM mass in combination with the small mass splitting between the singlet and singlino results in a common experimental signature with little to no $p_{\mathrm{T}}^{\mathrm{miss}}$. Searches for stealth SUSY are therefore highly complementary to traditional high-$p_{\mathrm{T}}^{\mathrm{miss}}$ SUSY searches. Feynman diagrams for two stealth SUSY models are shown in Fig. 8, where depending on the portal, additional gluons (stealth SYY) or b quarks (stealth SHH) are produced in the final state.

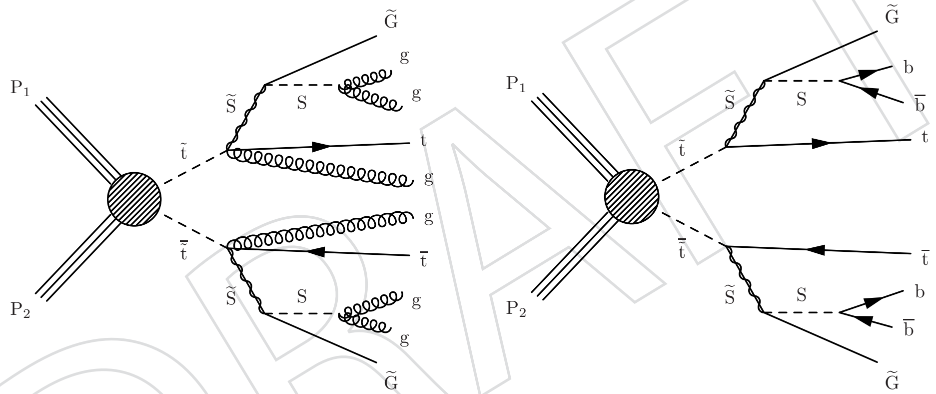

Figure 8: Feynman diagrams for pair production of top squarks under the stealth SYY (left) and stealth SHH (right) models. In these models, the signature is a pair of SM top quarks, with additional jets originating from gluons (SYY) or b quarks (SHH).

2.2.3 Inelastic dark matter

In inelastic dark matter (IDM) models [71, 139, 140], two DS states are predicted with near mass degeneracy. These states can be scalars or fermions, since this degeneracy can be induced in both cases via different mechanisms. For small mass splittings relative to the average mass, the elastic couplings between same-flavor states are suppressed compared to the inelastic ones, leading to the preferred simultaneous production of both states in pp collisions at the LHC. This production is mediated by one of the portal interactions, typically taken to be the dark-photon portal. These models can both evade increasingly stringent DM scattering constraints from DD experiments and predict the correct thermal-relic DM abundance as indicated by cosmological observations.

Focusing on the scenario with fermionic DM, a Dirac fermion can be defined as the bispinor $\psi = (\eta\ \overline{\xi})$. Assuming vector and axial-vector couplings to quarks, the interactions are de-

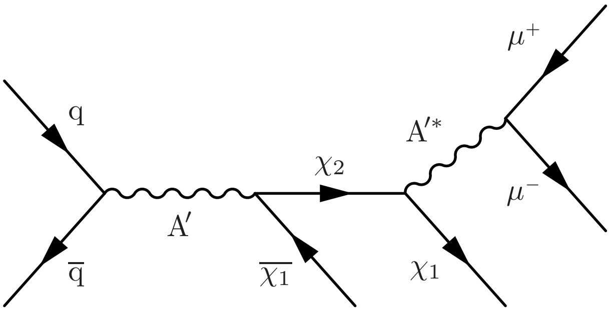

Figure 9: Feynman diagram of inelastic dark matter production and decay processes in pp collisions, for fermionic DM states. The heavier DM state χ_2 can be long-lived, and decays into χ_1 and to a muon pair via an off-shell dark photon A'.

scribed by [71]

$$\mathcal{L} \supset \overline{\psi}\, \gamma_\mu \,(g'_V + g'_A \gamma_5)\, \psi\, \overline{q}\, \gamma^\mu \,(g_V + g_A \gamma_5)\, q. \tag{17}$$

If we add also a small Majorana mass term $\frac{\Delta}{2}(\eta\eta + \overline{\xi\xi})$ to the Lagrangian, where Δ is the small mass splitting between states, the fermion mass eigenstates become

$$\begin{aligned} \chi_1 &\approx \frac{i}{\sqrt{2}}\,(\eta - \xi)\,, \\ \chi_2 &\approx \frac{1}{\sqrt{2}}\,(\eta + \xi)\,. \end{aligned} \tag{18}$$

The vector current $\overline{\psi}\, \gamma_\mu\, \psi$ in this scenario has the form

$$\overline{\psi}\, \gamma_\mu\, \psi \approx i(\overline{\chi}_1\, \overline{\sigma}_\mu\, \chi_2 - \overline{\chi}_2\, \overline{\sigma}_\mu\, \chi_1) + \frac{\Delta}{2m}(\overline{\chi}_2\, \overline{\sigma}_\mu\, \chi_2 - \overline{\chi}_1\, \overline{\sigma}_\mu\, \chi_1). \tag{19}$$

The elastic couplings in the second term are suppressed by a factor of Δ/m relative to the inelastic couplings in the first term and are negligible.

The excited state χ_2, once produced in tandem with the DM ground state χ_1 via pp $\rightarrow A' \rightarrow \chi_2\,\chi_1$, eventually decays into a χ_1 plus a pair of SM fermions by emission of an off-shell dark photon ($\chi_2 \rightarrow \chi_1 f\overline{f}$). The model is efficiently parameterized by the mass splitting Δ, the lighter state mass $m_1 = m_{\mathrm{DM}}$, and the interaction strength $y = \epsilon^2 \alpha_{\mathrm{dark}}$, where ϵ is the kinetic mixing between the dark photon and the SM hypercharge and α_{dark} is the coupling strength of the DS gauge interaction, as defined in Section 2.2.4. The small mass splitting between the states leaves only a small kinematic phase space available for the decay, leading both to a small decay width (and hence a large lifetime) of the excited state and to the production of low-energy SM fermions at the end of the decay chain. Additionally, there is near collinearity between the SM fermion pair and between the SM fermions and the χ_1 states. The displaced and low-energy SM fermion pair in the final state, combined with significant $p_{\mathrm{T}}^{\mathrm{miss}}$ from the χ_1, presents a unique and compelling experimental signature that can be searched for in pp collision events. Figure 9 shows a diagram for the displaced $\mu^+\mu^-$ + $p_{\mathrm{T}}^{\mathrm{miss}}$ signature.

2.2.4 Hidden valleys

Nonminimal DSs may include multiple new particles and potentially new interactions that are decoupled from the SM. This kind of model is often referred to as a "hidden valley" (HV) [141] because the DS may contain rich dynamics and phenomenology at relatively low energy scales while nevertheless being accessible via collider production only at high energy scales corresponding to the mass of the mediator particle. Generally, in HV models, the SM is supplemented by a non-Abelian DS $SU(N_c^{\text{dark}})$ with N_c^{dark} dark colors, gauge coupling α_{dark}, and massless dark gluons as the carriers of the new force. All SM particles are neutral under $SU(N_c^{\text{dark}})$, but there are new light particles that are charged under $SU(N_c^{\text{dark}})$ and neutral under the SM gauge groups. The basic particle content in the hidden sector comprises N_f^{dark} flavors of dark quarks (q_{dark}) charged under $SU(N_c^{\text{dark}})$ with masses $m_{q_{\text{dark}}}$.

Higher-dimensional operators, induced by a high-mass Z′ boson or a loop of heavy particles carrying both SM and hidden-sector charges, allow interactions between SM fields and the new light particles of the hidden sector. In a simple HV scenario, adding a broken $U'(1)$ gauge group introduces a heavy vector portal mediating between the two sectors. In such a scenario, the kinetic mixing between the hidden sector group $U'(1)$ and the SM group $U(1)_Y$ cannot be forbidden, implying the possible existence of an HV dark photon that may communicate with the SM via kinetic mixing. This class of models is sometimes called "dark QCD" in analogy with the SM QCD, though not all such models evince QCD-like behavior.

The confinement of this Yang–Mills theory at a scale Λ_{dark} is guaranteed only for $N_f^{\text{dark}} < 3N_c^{\text{dark}}$ [142]. Confinement and hadronization in the DS result in a spray of composite hidden-sector states, dark hadrons. This process is called a dark shower and produces dark jets. A key feature of dark shower signatures is the evolution of energy within the DS that follows the initial production at the hard process energy scale Q_{dark}. In QCD, the momentum flow from the hard scattering energy scale Q to the confinement scale is dominated by the soft and collinear singularities and can be described using perturbation theory (parton shower). This feature holds generally for theories that, like QCD, have small 't Hooft couplings $\lambda = \alpha_{\text{dark}}^2 N_c^{\text{dark}}$. In these theories, the 't Hooft coupling can become large, but only in a limited energy range near the confinement scale. The small 't Hooft coupling regime defines a QCD-like parton evolution, where well-established parton shower algorithms allow for good modeling of the partonic component of the hidden-sector evolution [143].

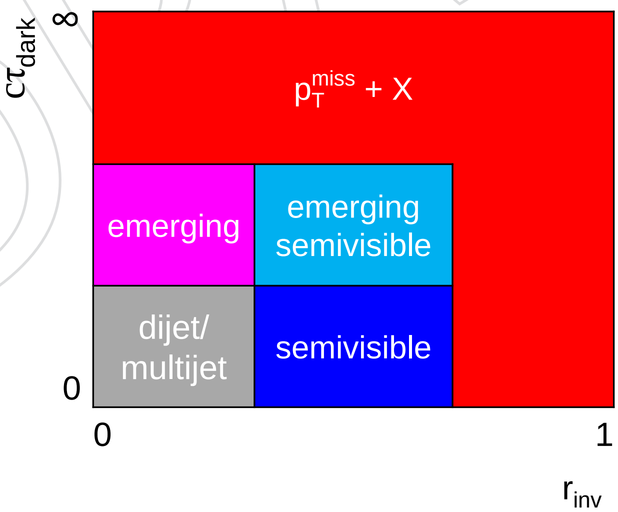

Figure 10: A qualitative depiction of the phenomenological behavior of dark QCD models depending on the fraction of invisible particles within a jet r_{inv} and the proper decay length of dark hadrons $c\tau_{\text{dark}}$. The r_{inv} parameter is defined in Section 2.2.4.1.

The dark mesons produced in the dark shower may or may not be degenerate with mass(es) m_{dark} and proper decay length(s) $c\tau_{\text{dark}}$, while dark baryons are typically neglected, as their masses scale with N_c^{dark} and therefore their production is suppressed [144]. Alternatively, if

$N_f^{dark} = 0$ or $m_{q_{dark}} > \Lambda_{dark}$, dark glueballs form [145], along with quirks [146] in the latter case. Numerous phenomenological signatures are possible, depending on the values of these parameters that define the dark QCD model. Two major categories in the case of a small 't Hooft coupling are semivisible jets (SVJs) and emerging jets (EJs), described in Sections 2.2.4.1 and 2.2.4.2, respectively. The relationships between these two signatures are shown in Fig. 10 in terms of the novel parameters of the models, which are explained in the following sections. Here, we discuss the particular models used to motivate and design CMS searches. Comparisons of these and other models, along with other details, are detailed in Ref. [142]. Alternatively, a large 't Hooft coupling produces soft unclustered energy patterns (SUEPs), discussed in Section 2.2.4.3. Figure 11 shows examples of final states including each of the three phenomena. It is generically expected that signals of composite DM are highly suppressed at DD experiments [115], complementing other models where DD may have more sensitivity than collider production.

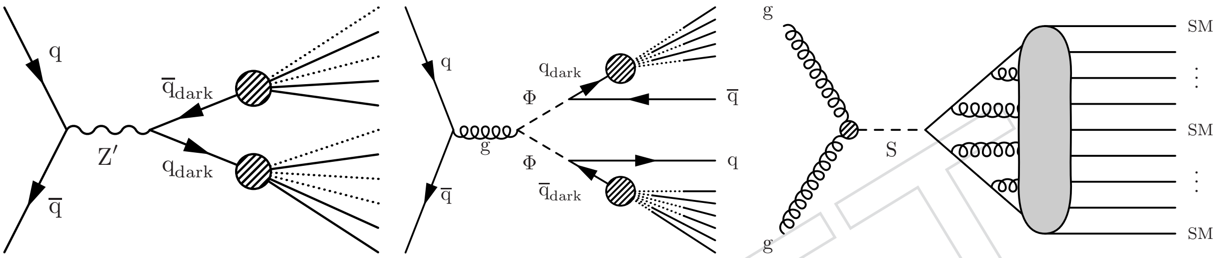

Figure 11: Illustrative Feynman diagrams showing example production modes for different hidden valley phenomena: semivisible jets (left), emerging jets (center), and soft unclustered energy patterns (right). Dotted lines indicate invisible particles.

2.2.4.1 Semivisible jets

References [115, 147] introduce a simple strongly coupled DS with $N_c^{dark} = 2$ and $N_f^{dark} = 2$, connected to the SM via a Z' mediator. This scheme produces both stable and unstable dark hadrons in varying proportions. Dark-hadron stability depends on the conservation of accidental symmetries in the DS. If the dark baryon number is conserved, then dark baryons cannot decay into SM particles. Similarly, if dark isospin is conserved, then dark vector mesons and pseudoscalars carrying nonzero dark isospin cannot decay. Therefore, combinations of different flavors of dark quarks are stable. The multiplicity of such states is proportional to $N_f^{dark}(N_f^{dark} - 1)$; however, the production of these stable hadrons may be suppressed by a mass splitting between the dark quark flavors, if $\Delta m^2_{q_{dark}} > \Lambda^2_{dark}$. This behavior is captured in an effective parameter called the invisible fraction, defined as $r_{inv} = \langle N_{stable}/(N_{stable} + N_{unstable})\rangle$, with allowed values ranging from 0 to 1. This parameter incorporates incalculable hadronic uncertainties from nonperturbative dynamics in the hidden sector. When r_{inv} assumes values different from 0 or 1, the result is a collimated mixture of visible and invisible particles, here referred to as a semivisible jet.

Both vector dark mesons ρ_{dark} and pseudoscalar dark mesons π_{dark} may form, with the former expected to occur with 75% probability, if the masses for the dark mesons are degenerate. These dark mesons are assumed to have similar mass scales, parameterized as a single value m_{dark}, and we set the constituent dark quark mass $m_{q_{dark}} = m_{dark}/2$. The unstable ρ_{dark} mesons decay democratically to any pair of SM quarks satisfying $m_{dark} \geq 2m_q$. The unstable π_{dark} mesons decay via a mass insertion, in analogy with SM pion decay, preferring the most massive species of SM quarks satisfying the above relationship. All decays of unstable dark mesons are assumed to be prompt, in accordance with theoretical predictions for this class of models [147]. The stable dark mesons traverse the detector invisibly and represent DM candidates.

The impact of the dark-coupling scale Λ_{dark} depends on m_{dark}, so its value is set by the formula $\Lambda_{\text{dark}} = 3.2(m_{\text{dark}})^{0.8}$, which is empirically found to maximize the number of dark hadrons produced in a typical dark shower [148]. The running coupling of the dark force can then be calculated as $\alpha_{\text{dark}}(\Lambda_{\text{dark}}) = \pi / (b_0 \log(Q_{\text{dark}}/\Lambda_{\text{dark}}))$, with $b_0 = (11N_c^{\text{dark}} - 2N_f^{\text{dark}})/6$ and $Q_{\text{dark}} = 1$ TeV. The mediator in this model is a leptophobic Z′ boson with universal couplings to SM quarks g_q and to dark quarks $g_{q_{\text{dark}}}$, as described in Section 2.1.1.1). To account for the multiple flavors and colors in the DS, we set $g_{q_{\text{dark}}} = 1.0/\sqrt{N_c^{\text{dark}} N_f^{\text{dark}}} = 0.5$. This produces a branching fraction to DM of 47% and a width of 5.6%, consistent with the LHC DM Working Group benchmark $g_{\text{DM}} = 1.0$ for minimal DM models [25].

2.2.4.2 Emerging jets

In some strongly coupled DS models, parton showering and fragmentation in the DS create dark mesons on a shorter time scale than that of the dark-meson decay into SM particles. Therefore, these dark mesons travel long distances before decaying into SM particles. This behavior leads to the signature of an emerging jet, a wide jet encompassing the multiple smaller displaced jets formed by the dark-meson decays. We consider models with $N_c^{\text{dark}} = 3$, such that the stable dark hadrons are dark baryons.

References [149, 150] introduce a strongly coupled DS with $N_f^{\text{dark}} = 7$ fermionic dark quarks. The dark quarks are produced via the decay of a complex scalar mediator Φ (Section 2.1.4), which is charged under both QCD and dark QCD. When produced resonantly, the mediator decays into a dark quark and SM quark: $\Phi \to q_{\text{dark}}\overline{q}$. This model assumes all dark quarks are degenerate and coupled through the mediator to SM down-type quarks, and is therefore described as "unflavored". The undetermined model parameters that influence the kinematic behavior include the mediator and dark meson masses and the dark meson lifetimes. The proper decay length can be computed as:

$$c\tau_{\text{dark}} = 80\,\text{mm} \left(\frac{1}{\kappa^4}\right) \left(\frac{2\,\text{GeV}}{f_{\pi_{\text{dark}}}}\right)^2 \left(\frac{100\,\text{MeV}}{m_{\text{d}}}\right)^2 \left(\frac{2\,\text{GeV}}{m_{\text{dark}}}\right) \left(\frac{m_\Phi}{1\,\text{TeV}}\right)^4, \tag{20}$$

where κ is the Yukawa coupling between Φ, q_{dark}, and the SM down quark; $f_{\pi_{\text{dark}}}$ is the dark pion decay constant; and m_{d} is the mass of the SM down quark.

A related model with $N_f^{\text{dark}} = 3$ [151], includes a coupling matrix $\kappa_{\alpha i}$ for the mediator Φ, where α is the dark quark flavor and i is the SM quark flavor. In particular, the "flavor-aligned" version of this model is considered, where the matrix is given by $\kappa_{\alpha i} = \kappa_0 \delta_{\alpha i}$, such that each flavor of dark quark couples to a single flavor of down-type SM quarks. Decays into the most massive allowed SM particles are preferred, leading to b quark enriched final states when the dark mesons are sufficiently massive. In this model, the proper decay length for a dark meson composed of dark quarks of flavors α and β is:

$$c\tau_{\text{dark}}^{\alpha\beta} = \frac{8\pi m_\Phi^4}{N_c m_{\text{dark}} f_{\pi_{\text{dark}}}^2 |\kappa_{\alpha i}\kappa_{\beta j}^*|^2 \left(m_i^2 + m_j^2\right) \sqrt{\left(1 - \frac{(m_i+m_j)^2}{m_{\text{dark}}^2}\right)\left(1 - \frac{(m_i-m_j)^2}{m_{\text{dark}}^2}\right)}}. \tag{21}$$

These models may also be characterized by the maximum proper decay length of any dark meson species, denoted $c\tau_{\text{dark}}^{\text{max}}$.

Reference [143] introduces a set of models with similar phenomenological behavior: the formation of long-lived dark hadrons that eventually decay into SM particles. These models fix $N_c^{\text{dark}} = 3$ and $N_f^{\text{dark}} = 1$, resulting in a spectrum with a spin-0 dark meson η_{dark} and a spin-1 dark meson ω_{dark}. Two benchmark scenarios for the dark hadron masses and dark-QCD scale are considered: $\Lambda_{\mathrm{dark}} = m_{\omega_{\mathrm{dark}}} = m_{\eta_{\mathrm{dark}}}$ and $\Lambda_{\mathrm{dark}} = m_{\omega_{\mathrm{dark}}} = 2.5 m_{\eta_{\mathrm{dark}}}$. In the first scenario, the ω_{dark} is typically stable and formed during hadronization with 75% probability (as in Section 2.2.4.1), while in the second scenario, the decay $\omega_{\mathrm{dark}} \to \eta_{\mathrm{dark}}\eta_{\mathrm{dark}}$ occurs and ω_{dark} forms with 32% probability. Typically, the η_{dark} is unstable and decays into SM particles, though there are some exceptions. The SM Higgs boson portal described in Section 2.1.2.3 is employed, producing a pair of dark quarks. However, after dark hadrons are formed, their decays into SM particles may proceed through different portals, leading to distinct phenomenology:

- gluon portal, with the decay $\eta_{\mathrm{dark}} \to \mathrm{gg}$ producing hadron-rich showers;
- photon portal, with the decay $\eta_{\mathrm{dark}} \to \gamma\gamma$ producing photon showers;
- vector portal (Section 2.1.1), in particular a heavy kinetically-mixed dark photon that allows both leptonic and hadronic decays of the vector ω_{dark} while the η_{dark} is stable, producing SVJs with a default $r_{\mathrm{inv}} = 0.25$ (Section 2.2.4.1);
- Higgs boson portal, with preferred decays $\eta_{\mathrm{dark}} \to \mathrm{b\bar{b}}$, $\eta_{\mathrm{dark}} \to \mathrm{c\bar{c}}$, and $\eta_{\mathrm{dark}} \to \tau^+\tau^-$ producing heavy flavor rich showers; and
- dark-photon portal (Section 2.1.1.2), less massive than the vector portal to allow the decay $\eta_{\mathrm{dark}} \to \mathrm{A'A'}$, with the $\mathrm{A'}$ decaying into quarks and leptons, producing lepton-rich showers.

The minimum lifetime of the unstable dark-hadron species depends on which decay portal is used. The dark-photon portal leads to a short minimum lifetime; the photon and vector portals lead to intermediate minimum lifetimes; and the gluon and Higgs boson portals lead to very long minimum lifetimes. One or more collimated decays of these particles may be observed in the tracker, calorimeter, and/or muon system of the detector, depending on the lifetime of the dark hadron.

2.2.4.3 Soft unclustered energy patterns

Dark showers produced in HV models do not necessarily result in collimated jets similar to SM QCD. In particular, SUEPs comprising a large multiplicity of spherically distributed low-momentum charged particles are also possible signatures of HV models. The underlying physics that produces such events can be varied; here, we consider quasi-conformal models in which the dark QCD force has a large 't Hooft coupling $\lambda \gg 1$ above its confinement scale [152]. When new particles shower with efficient branching over a wider energy range than in SM QCD, the initial parton momenta are not preserved, resulting in soft and isotropic emissions. In this case, the production of dark mesons proceeds similarly to hadron production in high-temperature QCD.

We focus on a benchmark model with a heavy scalar mediator S connecting the SM and DS, produced via gluon fusion. We assume the dark quark masses $m_{\mathrm{q_{dark}}}$ are less than the confinement scale Λ_{dark}, and also that $\Lambda_{\mathrm{dark}} \ll \sqrt{s}$. Therefore, the dark quarks undergo a quasi-conformal showering, forming dark pseudoscalar mesons π_{dark}. The dark-meson transverse momentum (p_{T}) spectrum follows a Boltzmann distribution that depends on the dark-meson mass m_{dark} and a temperature $T_{\mathrm{dark}} \approx \Lambda_{\mathrm{dark}}$. The pseudoscalar mesons decay into a pair of dark photons $\mathrm{A'}$. The dark photon kinetically mixes with the SM photon and decays promptly to SM particles including electrons, muons, and pions, with branching fractions ($\mathcal{B}$) that depend on its mass. Three benchmark $m_{\mathrm{A'}}$ values are considered, each with corresponding branching fractions: $m_{\mathrm{A'}} = 0.5\,\mathrm{GeV}$ ($\mathrm{A'} \to \mathrm{e^+e^-}, \mu^+\mu^-, \pi^+\pi^-$ with $\mathcal{B} = 40, 40, 20\%$), $m_{\mathrm{A'}} = 0.7\,\mathrm{GeV}$ ($\mathrm{A'} \to \mathrm{e^+e^-}, \mu^+\mu^-, \pi^+\pi^-$ with $\mathcal{B} = 15, 15, 70\%$), and $m_{\mathrm{A'}} = 1.0\,\mathrm{GeV}$ ($\mathrm{A'} \to \pi^+\pi^-$ with

$\mathcal{B} = 100\%$).

2.2.4.4 Neutral naturalness

Neutral naturalness models are motivated as a way to address the EW hierarchy problem [153]. Such scenarios include a discrete symmetry that relates SM fields to colorless counterparts. Realizations of this include the twin Higgs [154], folded SUSY [155] and quirky little Higgs [156] models. In each case, the partner particles escape LHC constraints because they are neutral under SM color charge. To address the hierarchy problem, the hidden sector must include a QCD-like gauge group with a confinement scale that is close to that of the SM. There must also be at least one more additional Higgs boson that mixes with the SM Higgs doublet and couples to particles in the hidden sector [153]. This leads to exotic decays of the Higgs boson to hidden sector particles as well as the potential production of additional Higgs bosons that decay to hidden sector particles. The lightest hidden sector particles are either effectively stable, creating $p_{\mathrm{T}}^{\mathrm{miss}}$, or undergo displaced decays to SM particles.

3 The CMS detector and event reconstruction

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. The ECAL barrel (endcap) covers the pseudorapidity range $|\eta| < 1.479$ ($1.479 < |\eta| < 3.0$), while the HCAL barrel (endcap) covers the $|\eta| < 1.3$ ($1.3 < |\eta| < 3.0$) range. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. The muon system is composed of three types of chambers: drift tubes (DTs) in the barrel ($|\eta| < 1.2$), cathode strip chambers (CSCs) in the endcaps ($0.9 < |\eta| < 2.4$), and resistive-plate chambers (RPCs) in both the barrel and the 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. [157].

Events of interest are selected using a two-tiered trigger system. The first level ("level-1"), 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 μs [158]. The second level, known as the high-level trigger (HLT), 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 [159].

The primary vertex (PV) is taken to be the vertex corresponding to the hardest scattering in the event, evaluated using tracking information alone, as described in Section 9.4.1 of Ref. [160]. The silicon tracker used in 2016 measured charged particles within the range $|\eta| < 2.5$. For nonisolated particles of $1 < p_{\mathrm{T}} < 10$ GeV and $|\eta| < 1.4$, the track resolutions were typically 1.5% in p_{T} and 25–90 (45–150) μm in the transverse (longitudinal) impact parameter [161]. At the start of 2017, a new pixel detector was installed [162]; the upgraded tracker measured particles up to $|\eta| < 3.0$ with typical resolutions of 1.5% in p_{T} and 20–75 μm in the transverse impact parameter [163] for nonisolated particles of $1 < p_{\mathrm{T}} < 10$ GeV.

The particle-flow (PF) algorithm [164] 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.

For each event, hadronic jets are clustered from these reconstructed particles using the infrared- and collinear-safe anti-k_{T} algorithm [165, 166] with a distance parameter of 0.4 (AK4 jets) or 0.8 (AK8 jets). Some analyses also use the Cambridge–Aachen algorithm [167] with a distance parameter of 1.5 (CA15 jets). 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 5–10% of the true momentum over the entire p_{T} spectrum and detector acceptance. Additional pp interactions within the same or nearby bunch crossings, known as pileup (PU), can contribute additional tracks and calorimetric energy depositions, increasing the apparent jet momentum. To mitigate this effect, tracks identified to be originating from PU vertices are discarded and an offset correction is applied to correct for remaining contributions [168]. Jet energy corrections are derived from simulation studies so that the average measured energy of jets becomes identical to that of particle-level jets. In situ measurements of the momentum balance in dijet, γ+jet, Z+jet, and multijet events are used to determine any residual differences between the jet energy scale in data and in simulation, and appropriate corrections are made [169]. Additional selection criteria are applied to each jet to remove jets potentially dominated by instrumental effects or reconstruction failures [168]. The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ [170]. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is modified to account for corrections to the energy scale of the reconstructed jets in the event.

If a resonance is much heavier than its decay products, the decay products are highly Lorentz boosted. This results in very collimated sprays of particles from those decay products, where hadronic decays cannot be reconstructed into individual small-radius jets, but are merged into one large-radius jet. In order to remove soft and wide-angle radiation in these jets, jet substructure [171] or jet grooming techniques such as trimming [172] and soft drop [173] are applied. Jet trimming is a method that removes sources of contamination by exploiting the difference in scale between the hard emissions of final state radiation and the relatively soft emissions from initial-state radiation (ISR). This algorithm begins with seed jets that are reclustered using the anti-k_{T} algorithm and then trimmed according to the subjet p_{T}. The soft-drop algorithm removes soft and wide-angle radiation from the jet by reclustering the large-radius jet with the Cambridge–Aachen algorithm and testing if $\min(p_{\mathrm{T},i}, p_{\mathrm{T},j}) > z_{\mathrm{cut}} p_{\mathrm{T},i+j} (\Delta R_{ij}/R)^{\beta}$ in each declustering step. The standard parameters used in the CMS experiment are $z_{\mathrm{cut}} = 0.1$ and $\beta = 0$. The hardest branch is followed until the soft-drop requirement is fulfilled, where the procedure stops. As a consequence, at most two soft-drop subjets are defined by this procedure. The mass is calculated as the invariant mass of the two subjets and is called the soft-drop mass m_{SD}.

4 Common experimental challenges

Searches for DS physics face common experimental challenges that are applicable to many signature types. To address these challenges, various methods are employed, shared, and continually improved across different analyses. In addition, new methods are developed to specifically address the distinctive features of DS signatures.

The design, deployment, monitoring, and characterization of trigger algorithms are fundamental components of all CMS analyses. Certain DS signatures introduce unique features that necessitate extensions to the standard trigger and data acquisition paradigm. This new data-taking paradigm is discussed in Section 4.1.

The reconstruction of $p_{\mathrm{T}}^{\mathrm{miss}}$, a key parameter in many DS searches, poses a significant challenge in the high-PU environment of the LHC. CMS has made concerted efforts to characterize the detector response and resolution to optimize the measurement of $p_{\mathrm{T}}^{\mathrm{miss}}$, as detailed in Sections 4.2 and 4.3. Additional variables that represent aspects of the event global activity are also defined and used throughout the analyses. The total hadronic transverse momentum H_{T} is defined as the scalar p_{T} sum of all jets that meet certain selection criteria. While the details of the selection may vary among different analyses, a common definition is to use all jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 3.0$. The missing hadronic transverse momentum (missing H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$) is similarly defined as the magnitude of the vector $\vec{p}_{\mathrm{T}}$ sum of all jets. In the same vein, the hadronic recoil $\vec{u}$ is defined as the vector $\vec{p}_{\mathrm{T}}$ sum of all PF candidates except for those identified with the decay products of an EW boson. It is often used as an ancillary variable to monitor the behavior of the $p_{\mathrm{T}}^{\mathrm{miss}}$.

In the context of DS searches, the reconstruction and identification of LLPs depend on their intrinsic properties, such as mass, charge, and lifetime [121, 174]. Various approaches to tackle this challenge are discussed in Section 4.4. Additionally, the particle reconstruction using the CMS-TOTEM precision proton spectrometer (PPS) is discussed in Section 4.5, and analyses of heavy ion collisions are discussed in Section 4.6.

Finally, searches for new physics must often employ methods based on control regions (CRs) in data to estimate background contributions, and DS analyses are no exception. Standard methods shared among many of the search efforts are discussed in Section 4.7.

4.1 Triggers, data scouting, and skims

Models featuring DS physics predict a wide variety of final states in pp collisions. Many triggers (as discussed in Section 4.1.1) are correspondingly developed to target these experimental signatures, which include $p_{\mathrm{T}}^{\mathrm{miss}}$ arising from stable particles that do not interact with the detector, leptons produced at the pp interaction point (prompt) or away from it (displaced), and standard or unconventional jet signatures created via enriched DS dynamics. While CMS successfully targets a range of these models, challenges arise in obtaining sensitivity to theories with exotic topologies, particularly those featuring new low-mass states in the DS. These states are generally difficult to probe because of trigger limitations. Decays of such low-mass DS states into SM particles lead to final-state particles that have either very low momentum (soft particles) or are very collinear, depending on the Lorentz boost in the laboratory frame. Both situations present triggering challenges. If these DS states are instead stable within the detector volume, they induce a soft $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum that is also difficult to use for triggering unless combined with energetic ISR jets, leading to loss of signal acceptance.

Several techniques are employed in CMS to address these challenges and improve sensitivity to DS models with exotic signatures. Here we discuss the use of data scouting (Section 4.1.2) and skims (Section 4.1.3) to expand the range of low-mass DS particles that can be probed in CMS, after describing the relevant standard triggers available in CMS during the Run 2 data-taking period.

4.1.1 Standard triggers

Event selection in CMS starts with a two-tiered trigger system, as discussed in Section 3. Standard triggers save the acquired data in a raw format that represents the complete information of the detector readout electronics. The advantage of saving the data in this format is that they can be reconstructed multiple times, profiting from more accurate calibrations that usually only become available later in the running period. The trade-off is the large size of the data volume, of order 1 MB/event. Thus the trigger system must balance the selection efficiency for signal events with the background rejection rate, which is correlated with the trigger output bandwidth. Since the HLT runs an optimized version of the full reconstruction software, a number of dedicated reconstruction techniques described later in this section are also implemented in the HLT.

The cleanest signatures for triggering are those with prompt electrons or muons in the final state. Analyses targeting these signatures usually employ general-purpose lepton triggers. For example, in 2018, the isolated single-electron trigger required $p_{\mathrm{T}} > 32\,\mathrm{GeV}$, and the dielectron trigger required $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ for both electrons. Likewise, the general-purpose isolated single-muon trigger required $p_{\mathrm{T}} > 24\,\mathrm{GeV}$, and the isolated dimuon trigger required $p_{\mathrm{T}} > 17\,(8)\,\mathrm{GeV}$ for the largest (second-largest) p_{T} muon. These algorithms are less effective for displaced leptons, for which dedicated triggers were developed. For signatures with tau leptons and b-tagged jets, the most common strategy is to use the standard reconstruction and identification techniques for the tau lepton or b jet itself and then design a dedicated trigger algorithm focusing on the final state as a whole.

The more challenging signatures are those with only photons or hadronic jets in the final state. Stringent kinematic thresholds are applied to the trigger algorithms to keep the rates within the allocated bandwidth. Dedicated triggers featuring special reconstruction algorithms for displaced or delayed objects are again deployed.

Finally, an all-purpose $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger is available to select events where a particle such as a DM candidate produced in the collision escapes the CMS detector and leaves no signal. This signature is extremely sensitive to experimental conditions such as detector calibrations and PU. The trigger requirement relies on an online calculation of $p_{\mathrm{T}}^{\mathrm{miss}}$ that is based on all PF candidates reconstructed at the HLT except for muons. It is usually combined with an $H_{\mathrm{T}}^{\mathrm{miss}}$ requirement, where jets are subjected to stringent identification requirements. The kinematic thresholds for these algorithms are $p_{\mathrm{T}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}} > 110\ (120)\,\mathrm{GeV}$ in 2016–2017 (2018) data. Unavoidable discrepancies exist between the online (trigger level) and offline reconstruction of $p_{\mathrm{T}}^{\mathrm{miss}}$, because the latter benefits from additional subdetector information and improved calibrations. The effect of those discrepancies is shown in the efficiency curve in Fig. 12. These online thresholds reach $\sim$95% efficiency for offline thresholds above 250 GeV. Table 1 displays a subset of the trigger algorithms deployed in CMS during 2018 that select events based on the presence of one or two physics objects. The complete CMS HLT event selection comprises $\sim \mathcal{O}(700)$ trigger algorithms, including those for alignment/calibration, monitoring, and backup.

4.1.2 Data scouting

The fundamental rate limitation in CMS is the total amount of data that can be transferred to storage at once, not the number of events that can be stored. A powerful technique to increase the event rate involves decreasing the information stored per event, thereby releasing some of the data bandwidth to store more events. This technique is termed “data scouting” in CMS and has been deployed since Run 1. Data scouting and “data parking,” which is another technique to save more data, are the subject of their own Report [176]. Here, we give a brief overview of

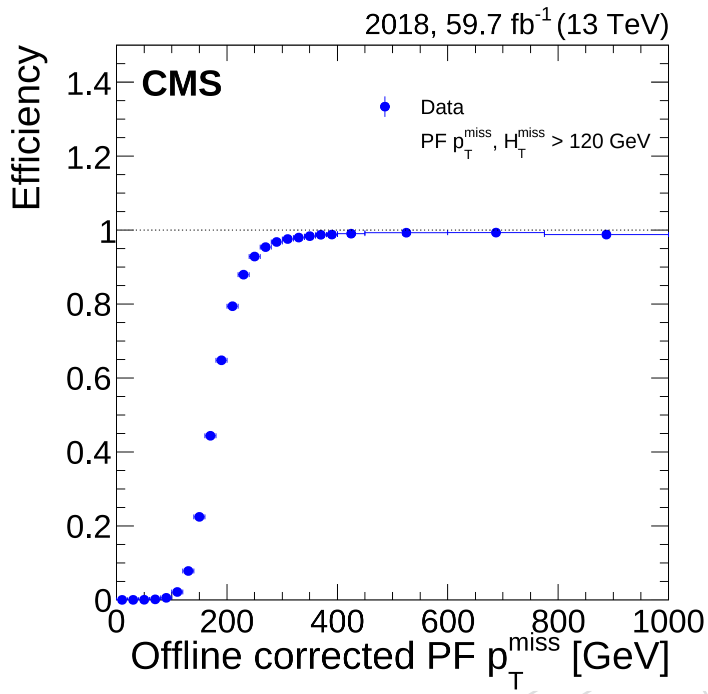

Figure 12: The event selection efficiency for requiring HLT thresholds of 120 GeV in both p_T^{miss} and H_T^{miss} as a function of the offline corrected p_T^{miss}, which takes into account jet energy scale corrections.

Table 1: Summary of p_T (or E_T) requirements (in GeV) of a subset of the HLT algorithms deployed in CMS during 2018, for trigger paths based on one or two physics objects. One p_T threshold value is given for the single-object triggers, and two p_T threshold values are given for the di-object triggers. Triggers with isolated leptons are labeled "iso.", and have generally lower kinematical thresholds than the corresponding algorithms that do not impose isolation requirements on leptons. The "1-prong" note for the tau lepton trigger refers to a selection targeting the τ decay into a single charged particle + neutrals. The "barrel" note for the photon trigger refers to a photon reconstructed solely within the barrel section of the ECAL. The "AK4" and "AK8" notes refer to jets reconstructed with the anti-k_T algorithm and a distance parameter of 0.4 and 0.8, respectively [165]; the mass threshold is applied to m_{trim}, the trimmed jet mass [172]. The "b tags" note refers to the number of jets that are b-tagged with the DEEPCSV algorithm [175].

Single-object triggers

e	μ	τ (iso.)	γ	Jet	p_T^{miss}	H_T
32 (iso.)	24 (iso.)	180	110 (iso., barrel)	500 (AK4)	120	1050
115	50		200	400, $m_{trim} > 30$ (AK8)		330 + 4 jets, 3 b tags

Di-object triggers

	e	μ	τ (iso.)	γ	Jet	p_T^{miss}	H_T
e	23, 12 (iso.) 25, 25	23, 12 (iso.) 27, 37	24 (iso.), 30		30 (iso.), 35 50, 165		28 (iso.), 150
μ	23, 12 (iso.) 27, 37	17, 8 (iso.), $m_{\mu\mu} > 3.8$ 37, 27	20 (iso.), 27	17, 30			
τ (iso.)	180		35, 35			50 (1-prong), 100	
γ				30, 18 (iso.) 70, 70			
p_T^{miss}							100, 500

data scouting, as it is relevant for some DS searches.

In Run 2, two scouting strategies are defined: one focusing on final states involving muons, and the other on hadronic final states. The “muon scouting” data set saves only muon information per event, apart from limited event-level information. This drastically reduces the event size from roughly 1 MB to about 4 (9) kB in 2017 (2018), enabling muon triggers with much lower momentum thresholds at the same instantaneous luminosity. The muon pair (dimuon) scouting trigger requires each muon to have $p_T > 3$ GeV at the HLT, compared to the standard CMS dimuon trigger requirements of $p_T > 17$ GeV for the first muon and $p_T > 8$ GeV for the second; in both cases, muons are required to be isolated. The trigger rate goes up to about 6 kHz.

Several analyses have exploited the muon scouting data set to enhance sensitivity to low-mass physics. Searches for prompt [177] and displaced [178] resonances decaying to muon pairs obtain some of the most stringent exclusion limits on dark photon production for few-GeV dark photon masses. Model-independent searches such as the one in Ref. [177] also employ muon scouting data to enable the investigation of additional DS models, such as the 2HDM+a framework.

A second scouting strategy in Run 2 collects only jet-related information per event. This data set, termed “PF scouting”, enables a considerable reduction in the jet trigger H_T thresholds, expanding the range of low-mass jet-related searches feasible in CMS. The PF scouting trigger sets a requirement of $H_T > 410$ GeV at the software level, computed by considering jets with $p_T > 40$ GeV, compared with the standard trigger requirement of $H_T > 1050$ GeV. By storing only jet-related information in the event, the event size is reduced from 1 MB to about 15 kB, and the trigger rate is increased to about 2 kHz. For comparison, the rate of the data set that comprises all standard jet triggers is close to 400 Hz.

The Run 2 jet scouting technique has been used to enhance the low-mass sensitivity to several dijet, trijet, and multijet analyses [179, 180]. For example, a search for dijet resonances [179] attained a dijet mass sensitivity as low as 350 GeV, compared to about 500 GeV when using the standard triggers. A more detailed description of DS analyses that feature scouting data sets can be found in Section 6.

4.1.3 Skims

Data skimming is a useful technique to improve the speed and robustness of analyses that are based on highly selective data sets or highly selective event content. The results of data skimming are very compact data sets here referred to as “skims”, which only contain a small subset of events and only the event content that are of interest to a particular analysis group. Skims provide a powerful and configurable way to select events for offline analysis that can significantly reduce the size of the data sets that must be processed.

Skims can be configured for several purposes: to pick specific trigger paths to accept and specific collections to save, and the level of detector reconstruction on which to operate. Additional selection requirements can also be imposed to further reduce the stored number of events. The combination of these criteria enables a data set to be distilled down to only the components (triggers and physics objects) that are relevant to an analysis or group of analyses.

Several relevant skim configurations were employed in Run 2:

- “No-BPTX” skim: Stores events collected without the beam pickup timing device (BPTX) firing, called the “No-BPTX” triggers. These triggers are active only when no proton bunches are colliding in the detector. The skim has been used to search for stopped LLPs that come to rest inside the detector before decaying, as described in Section 6.3.2.4, and also for cosmic ray muon studies.

- Displaced-jet skim: Selects events with a prescaled trigger requiring $H_{\mathrm{T}} > 400\,\mathrm{GeV}$ to monitor the performance of HLT online tracking, which is crucial for triggers targeting displaced jet signatures in CMS, as described in Section 6.3.2.1.
- High-$p_{\mathrm{T}}^{\mathrm{miss}}$ skim: Selects events acquired with $p_{\mathrm{T}}^{\mathrm{miss}}$-based triggers by requiring them to have at least $p_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$. The full event information is saved for events passing this requirement. This skim has been used by various analyses and for studies of the performance of the $p_{\mathrm{T}}^{\mathrm{miss}}$ algorithm.

Most of the skim configurations save information from the standard event content, enhanced by additional collections that are typically only available in the full event content, which is generally not stored on disk. Collections commonly saved to custom skims include the full set of calorimeter reconstructed hit information. In the standard event content, only a subset of those hits around interesting regions of the detector are made available. Access to the full hit collection is essential for several searches for DM, and skims make this possible with little additional configuration overhead.

4.2 Pileup mitigation

The CMS Collaboration has developed several widely used techniques for mitigating the impact of PU. One of these techniques, known as charged-hadron subtraction (CHS) [170], has served as the standard method for PU mitigation in jet reconstruction since the beginning of Run 2. The CHS algorithm operates by excluding charged particles associated with reconstructed vertices from PU collisions during the jet clustering process. To address the impact of neutral PU particles in jets, an event-by-event jet-area-based correction is applied to the jet four-momenta. Additionally, a technique for identifying PU-related jets (PU jet ID) is used to reject jets primarily composed of particles originating from PU interactions.

However, all these techniques have limitations when it comes to effectively removing PU contributions from neutral particles. For instance, the jet-area-based correction acts on the entire jet and is incapable of entirely eliminating PU contributions from jet shape or jet substructure observables. To address this limitation, a new PU mitigation technique, known as PU-per-particle identification (PUPPI) [168, 181], has been introduced. This algorithm works at the particle level and builds upon the preexisting CHS algorithm. The PUPPI algorithm computes the probability that a neutral particle originates from PU, based on the distribution of charged PU particles in its vicinity, and adjusts the energy of the neutral particle based on its respective probability. As a result, objects formed from hadrons, such as jets, $p_{\mathrm{T}}^{\mathrm{miss}}$, and lepton isolation, demonstrate reduced dependency on PU when PUPPI is employed [170]. The improved performance of the resolution of the PUPPI hadronic recoil in $\mathrm{Z} \to \mu\mu$ processes with respect to PU effects, represented by the number of reconstructed vertices $\mathrm{N_{vtx}}$ is shown in Fig. 13; the hadronic recoil vector is divided into components parallel ($u_{\parallel}$) and perpendicular ($u_{\perp}$) to the boson axis.

Searches for LLPs must often employ dedicated strategies for PU mitigation to avoid a significant impact on the selection efficiency. These are discussed in Section 4.4.

4.3 Filters for spurious events

Spurious events can occur because of a variety of reconstruction failures, detector malfunctions, or noncollision backgrounds and have anomalous high-$p_{\mathrm{T}}^{\mathrm{miss}}$ measurements. Such events are rejected by dedicated event filters that remove more than 85–90% of these spurious high-

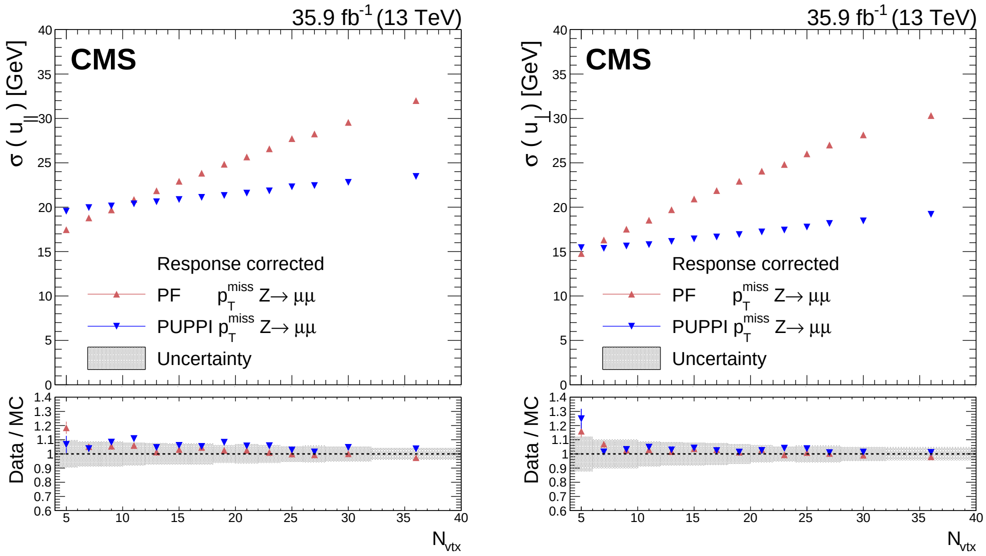

Figure 13: Upper panels: PUPPI and PF $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution of $u_{\parallel}$ (left) and $u_{\perp}$ (right) components of the hadronic recoil as a function of N_{vtx}, in $\mathrm{Z} \to \mu\mu$ data. Lower panels: data-to-simulation ratio. Systematic uncertainties are represented by the shaded band. Figure taken from Ref. [170].

$p_{\mathrm{T}}^{\mathrm{miss}}$ events with a mistagging rate of less than 0.1% [170]. These filters allow the removal of events with "artificial $p_{\mathrm{T}}^{\mathrm{miss}}$" arising from: interactions of machine-induced background particles moving along the beam direction, known as "beam halo", with the hadronic calorimeter; significant noise in the HCAL barrel and endcaps, detected by distinctive geometrical patterns of the readout electronics and by the usage of pulse shape and timing information; spurious signals in ECAL arising from sources such as anomalous large pulses in the endcap 5×5 crystal groups (supercrystals) and inoperative readout electronics; and high-p_{T} particle tracking failures leading to poorly measured PF muons and charged hadrons.

In the case that artificial $p_{\mathrm{T}}^{\mathrm{miss}}$ is the dominant source of background, custom filters optimized for a particular kinematic phase space may be needed [148]. For instance, when requiring that the jet momentum aligns with $p_{\mathrm{T}}^{\mathrm{miss}}$, over 40% of the QCD multijet background originates from events with artificial $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ caused by nonfunctional calorimeter cells. These events were not consistently detected by the dedicated filters mentioned earlier and additional analysis requirements were developed and employed.

4.4 Long-lived particle reconstruction

Particles with long lifetimes are an important possibility in the search for new phenomena, and often appear in BSM scenarios, notably in models that describe the elementary particle nature of DM. When produced at the LHC, LLPs have a distinct experimental signature: they can decay far from the primary pp interaction vertex but within the detector, or even completely pass through the detector before decaying. Some specific examples of LLP signatures include displaced and delayed leptons, photons, and jets; disappearing tracks; and nonstandard tracks produced by monopoles or heavy stable charged particles. Standard triggers, object reconstruction, and background estimation are usually inadequate for LLP searches because they are designed for promptly decaying particles, and custom techniques are often needed to analyze the data. Here we describe specific offline object reconstruction techniques that are used to identify long-lived and displaced particles in CMS.

4.4.1 Displaced tracking/vertexing

Displaced tracking and displaced vertexing are important handles to identify LLPs decaying inside the inner tracking system of CMS. The track reconstruction starts from the hit reconstruction, where the signal above specific thresholds in pixel and strip channels are clustered into hits. The initial estimation of the hit position is determined by the charge and the position of the cluster and is corrected for the Lorentz drift in the magnetic field. This initial estimation of the hit position is utilized in the following steps of seed generation and track finding.

In the seed generation, the initial possible track candidates are formed, which serve as the starting points for the propagation using the Kalman filter [182]. The CMS detector utilizes an iterative tracking process [161], with each iteration starting from a specific group of seeds. The seeds are formed using two, three, or four hits in the different layers of the pixel detector and the strip detector. The earlier iterations utilize the hits in the pixel detector to target prompt tracks, while the later iterations focus more on the tracks with larger displacements. After each iteration, hits associated with reconstructed tracks are removed. In this way, the tracking at CMS becomes efficient for reconstructing tracks with different displacements.

After the seeds belonging to a given iteration are formed, a combinatorial track finder based on the Kalman filter is applied, where the track candidates produced by the seeds are extrapolated to the next compatible layers using the Kalman filter. After the extrapolation reaches the final layer, track fitting is achieved by updating the track parameters through the smoothing step of the Kalman filter. The track candidates with too many missing hits or with p_{T} below some threshold specific to a given iteration are dropped. Since all the seeds are extrapolated at the same time, there could be some tracks with significant overlaps. When two tracks share more than 19% of the hits, the one with a smaller number of hits is removed; if both tracks have the same number of hits, the one with a larger χ^2 is discarded.

This iterative tracking approach described above is also available in the HLT system of CMS. Although HLT tracking has a degraded performance compared to offline tracking and is usually limited to some specific regions of interest, such as regions around jets, it enables us to develop and implement dedicated LLP triggers for displaced jets and displaced leptons, greatly enlarging the coverage of the LLP searches at CMS [183].

Beyond the track reconstruction, displaced vertexing using the reconstructed tracks is also a powerful tool to further discriminate exotic LLP signatures from SM background processes. The "inclusive vertex finder", which is the standard DV reconstruction algorithm at CMS [175], is tuned for reconstructing decays of heavy-flavor hadrons arising from SM processes through their secondary vertex (SV) and is not efficient in reconstructing exotic LLP decays. Therefore dedicated DV reconstruction algorithms are used in exotic LLP searches, which significantly improve the signal-to-background discrimination. Displaced vertices may also be referred to as SVs, and the vector pointing from the PV to the point of closest approach of a DV track is referred to as the impact parameter (IP) vector. Figure 14 illustrates these concepts.

In general, for vertex reconstruction tasks, it can be proven mathematically that the Kalman filter provides the optimal performance assuming Gaussian noise and no outlier tracks, which are the tracks that do not belong to the vertex but are used in the fitting. In reality, however, the presence of outlier tracks is inevitable, owing to the dense tracking environment associated with the pp collisions at CMS, especially when searching for DVs accompanied by hadronic decays. Several approaches have been adopted in CMS LLP searches to address such challenges.

One approach to filtering outlier tracks in the DV reconstruction is to start with all possible pairs of preselected tracks, which serve as the initial vertex candidates. The vertex candidates

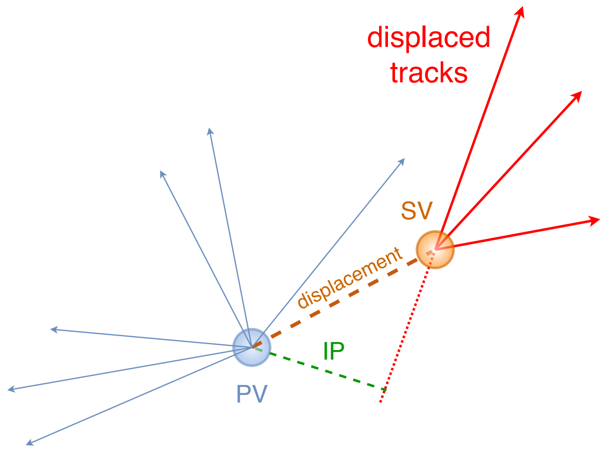

Figure 14: Illustration of the appearance of a secondary vertex (SV) from the decay of a long-lived particle resulting in charged-particle tracks that are displaced with respect to the primary interaction vertex (PV), and hence can have large impact parameter (IP) values. In BSM searches, LLPs have very long lifetimes compared to SM particles, leading to large displacements of the secondary vertices. Figure adapted from Ref. [175].

are then iteratively merged when they share tracks and have a small distance significance between the two vertices. After each merging, the new vertex candidate is refitted using the Kalman filter, and the vertex candidates with large χ^2 per degree of freedom are dropped. In this way, the input track candidates are automatically partitioned into different vertices during the vertex reconstruction process, while minimizing potential contamination from outlier tracks. This method is employed by several searches for DVs within the beam pipe [184, 185].

Another powerful technique to tackle the outlier-track contamination issue is the adaptive vertex fitter (AVF) [186], which is used in the inclusive displaced-jets search [187]. The AVF is a combination of the Kalman filter and the deterministic annealing algorithm, where, during the fitting, each track is assigned a weight according to its distance significance with respect to the vertex candidates and a given "temperature" T, which controls the shape of the weight function:

$$w_{\mathrm{track}_i} \equiv \frac{\exp(-\chi_i^2/2T)}{\exp(-\chi_i^2/2T) + \exp(-\chi_c^2/2T)}, \quad \chi_i^2 = d_i^2(\mathbf{x}_i, \mathbf{v})/\sigma_i^2 \tag{22}$$

where χ_c^2 defines a threshold such that a track with larger χ_i^2 is more likely to be an outlier than to have its position $\mathbf{x}$ associated with the vertex with position $\mathbf{v}$. The Kalman filter is then applied iteratively using the weighted track candidates. At each iteration, a specific value of T is chosen, starting at 256, and decreasing iteration by iteration until it reaches 1. The values of T are chosen such that the vertex reconstruction has good efficiency and resolution. In this way, the outlier tracks with large χ_i^2 are downweighted after each iteration, which leads to a vertex fitting that is robust against the contamination of outlier tracks.

Pileup mitigation is another important consideration that analysts must consider when vertexing displaced objects. In the case of displaced-jet searches, it has become conventional to select the vertex with the assistance of the α_{max} parameter [188], shown here for a particular vertex v_i and jet j:

$$\alpha_{\mathrm{max}}(v_i, \mathrm{j}) = \max_{v_i} \left[\frac{\sum_{\mathrm{tracks} \in v_i} p_{\mathrm{T}}}{\sum_{\mathrm{tracks}} p_{\mathrm{T}}^{\mathrm{j}}} \right]. \tag{23}$$

The parameter α_{max} takes the maximum of the ratio of the summed track p_{T} for all tracks associated with a particular vertex v_i to the total summed p_{T} for all tracks associated with the jet in consideration. Tracks are associated with the jet geometrically, e.g., by defining a ΔR requirement that is consistent with the type of jet used in the analysis. The tracks are associated with a vertex based on their weight, calculated for a given vertex as in Eq. (22). The individual values of α for a given vertex v_i range from 0 to 1, where $\alpha \approx 0$ is most consistent with displaced jets and $\alpha \approx 1$ is most consistent with prompt jets from the PV. The value of α for PU jets is within the range of 0 to 1 for a given vertex. To avoid selecting these jets, one takes the maximum of the alpha values for all vertices in the event.

4.4.2 Displaced-jet tagger

Jets displaced from the pp collision region, and arising from the decay of LLPs, are a key experimental signature for many theoretical extensions to the SM [129, 130, 189–191].

In the displaced-jets search [187], a dedicated algorithm was deployed to reconstruct the DV arising from LLP decays, using the displaced tracks associated with a dijet system. The properties of the associated tracks and DV can provide the discrimination power to distinguish LLP signals from SM backgrounds. A displaced-jets tagger is built using these properties based on a gradient-boosted decision tree (GBDT), with which the search provides world-leading sensitivities to a large number of BSM scenarios containing hadronically decaying LLPs.

A deep neural network (DNN) has also been designed to identify displaced jets [192]. The DNN architecture is inspired by the CMS DEEPJET algorithm [193, 194] that identifies jets originating from the hadronization of b quarks. The DNN provides a multiclass classification scheme similar to the DEEPJET algorithm but it also accommodates the "LLP jet" class. The network is trained using simulated events, which are typically drawn from the relevant parameter space of simplified models. Given that the experimental signature for a displaced jet depends strongly on the lifetime of the LLP, a parameterized approach [195] is adopted by using the lifetime parameter as an input to the DNN. This approach permits hypothesis testing over several orders of magnitude of lifetimes using a single network. Another key design feature of the DNN is the use of domain adaptation [196], along with the use of training examples taken from LHC data, to ensure a similar classification performance in simulation and pp collision data. The performance of the tagger is model- and lifetime-dependent, but it can typically provide a rejection factor in excess of 10 000 for jets from SM processes while maintaining a large signal efficiency (e.g., $\gtrsim$10%) for LLPs with proper decay lengths in the millimeter range. Figure 15 shows the receiver operating characteristic (ROC) curves for the DNN, for a number of SUSY models that contain an LLP and for two choices of lifetimes, $c\tau_0 = 1\,\text{mm}$ and $1\,\text{m}$.

4.4.3 Delayed calorimetry

The time resolution of the CMS calorimeter cells is around 400 ps for the ECAL [197], and a few ns for the HCAL [198]. (For Run 3, the HCAL timing resolution has been improved to around 1 ns.) This performance makes timing an excellent discriminant to identify energy deposits from slow-moving particles that arrive out of time. As shown in Fig. 16, these deposits can be delayed for two reasons: the extended path length to reach the calorimetry as compared with deposits from particles originating from the interaction point, and heavy LLPs can travel with a velocity significantly smaller than that of light. The heavier the mass and longer the lifetime of the LLP, the longer it will take to reach the detector and deposit calorimeter energy.

The CMS Collaboration has carried out two analyses that exploit the fact that LLPs decaying into hadrons nearby the calorimeter surface can be identified as out-of-time jets [199, 200]. The

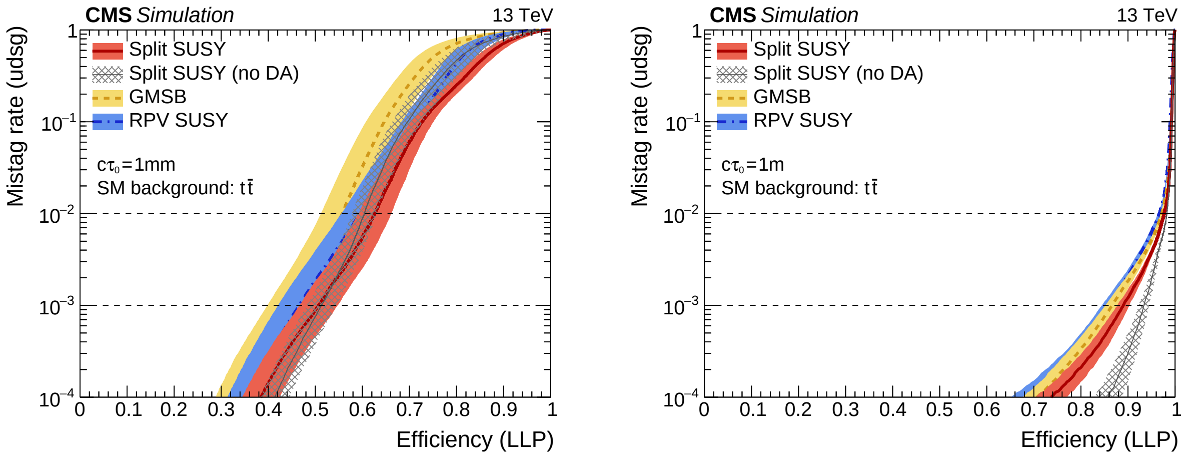

Figure 15: The ROC curves illustrating the displaced jet tagger performance for the split SUSY (solid line), GMSB SUSY (dashed line), and RPV SUSY (dot-dashed line) benchmark models, assuming $c\tau_0$ values of 1 mm (left) and 1 m (right). The thin line with hatched shading indicates the performance obtained with a DNN training using split SUSY samples but without domain adaptation (DA). Figure taken from Ref. [192].

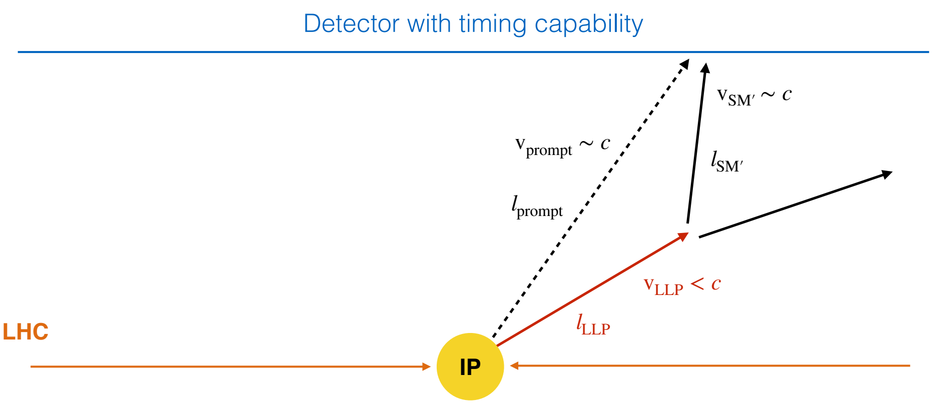

Figure 16: Illustration of contributions to the delay of particles that originate from LLP decays. For prompt decays, the path length to reach a particular location on the timing detector (l_{prompt}) is smaller than the path length for a deposit originating from an LLP decay (l_{LLP}+$l_{SM'}$). In addition, the velocity of the light SM particles (v_{prompt}) will be close to that of light while the velocity of the LLP (v_{LLP}) can be significantly lower. These factors lead to substantial delays for the decay products of LLPs, which can be exploited to improve sensitivity.

ECAL crystals associated with the jet can be used to define a new variable, the jet time, as the energy-weighted sum of the arrival times of measured pulses. The effective jet time resolution, taking into account clock jitter, size of the collision beam spot and calibration effects, ranges from 400–600 ps for jets with p_T ranging from 30–150 GeV. Any difference in the simulation of the time resolution [201] is corrected by selecting dedicated CRs in the data.

4.4.4 Displaced muons

A detailed description of the CMS muon reconstruction algorithms and their performance has been given in Refs. [202–204]. Here, we will briefly summarize how muons from pp collisions are reconstructed in CMS in general and then describe the specifics of displaced-muon reconstruction.

In general, muons from pp collisions in CMS are reconstructed using a combination of information from the tracker and the muon system. The muon system chambers are assembled into four "stations" at increasing distances from the interaction point; each station provides reconstructed hits in several detection planes, which are combined into track segments, forming the basis of muon track reconstruction in the muon system. "Standalone muon tracks" are built along the muon's trajectory using a Kalman filter technique [182] that exploits track segments from the muon subdetectors (DTs, CSCs, and RPCs). Independently, "tracker muon tracks" are built by propagating tracker tracks to the muon system with loose matching to DT or CSC segments. If at least one muon segment matches the extrapolated track, the tracker track qualifies as a tracker muon track. Finally, "global muon tracks" are built by matching standalone muon tracks with tracker tracks. In contrast to tracker muons, global muon trajectories are determined from a combined Kalman filter fit using both tracker and muon system information.

For displaced muons coming from decays of LLPs, the muon reconstruction algorithm that provides the best performance depends on how displaced the muon is from the interaction point. Muons produced relatively near the interaction point can be accurately reconstructed using the tracker muon or global muon reconstruction algorithms developed for prompt muons. The efficiency of these algorithms, however, rapidly decreases as the distance between the interaction point and the muon origin increases; the efficiency drops to zero for muons produced in the outer tracker layers and beyond. On the other hand, such muons are still efficiently reconstructed by the standalone muon reconstruction algorithms. These standalone muon algorithms reconstruct muons with displacements up to a few meters, but they have poorer spatial and momentum resolution than muons reconstructed using more precise information from the silicon tracker. In particular, a "displaced standalone" (DSA) muon track reconstruction algorithm was developed for displaced muons [204–206]. The DSA muon track algorithm uses only hits in the muon chambers and, in contrast to regular standalone muons, has the beamspot constraints removed from all stages of the muon reconstruction procedure. Thus, DSA tracks provide the largest efficiency and best resolution for displaced muons, out of all the available standalone muon track algorithms. It maintains a muon reconstruction efficiency of 0.95 up to a muon transverse production distance of 300 cm, as compared with standard algorithms, where the efficiency steeply declines after 10 cm, as shown in Fig. 17 [207].

Several analyses [132, 207] use displaced muons spanning a wide range of displacements, and take advantage of multiple muon reconstruction algorithms. For example, an attempt to match DSA tracks with global or tracker muons is made, and if such a match is found, the global or tracker muon is used for further analysis, while if not, the original DSA track is used. As a result of this matching procedure, much of the pp collision background is eliminated and the sensitivity to LLP decays in the tracker is greatly increased because tracker and global muons

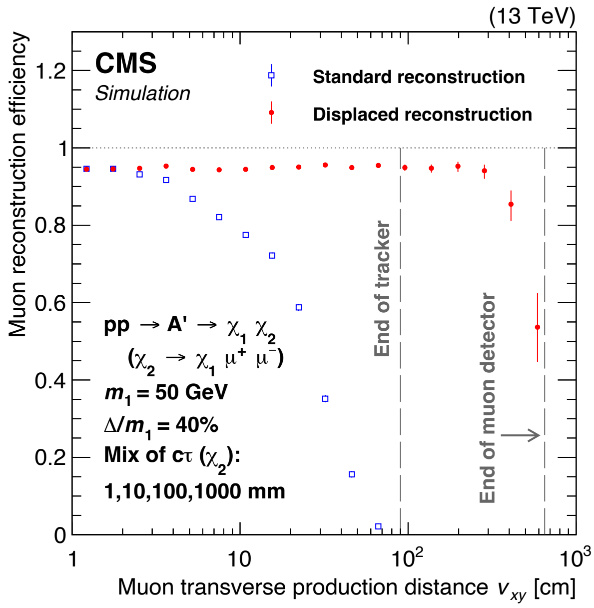

Figure 17: Simulated muon reconstruction efficiency of standard global muon (blue squares) and DSA (red circles) track reconstruction algorithms as a function of transverse vertex displacement v_{xy}, for the IDM model discussed in Section 2.2.3. The two dashed vertical gray lines denote the ends of the fiducial tracker and muon detector regions, respectively. Figure taken from Ref. [207].

have much better spatial and momentum resolution than standalone muons.

4.4.5 Muon detector showers

Long-lived particles that decay in the muon detectors could either be reconstructed as displaced muons, which are described in the previous section, or as muon detector showers. Owing to the design of the CMS muon detectors, which are composed of detector planes interleaved with the steel layers of the magnet flux-return yoke, LLPs that decay into any nonmuon particles within or just prior to the muon detectors can induce hadronic and electromagnetic showers, giving rise to a high hit multiplicity in localized detector regions. This signature uses the muon detector as a sampling calorimeter to identify displaced showers produced by LLPs that decay into hadrons, electrons, photons, or τ leptons. Additionally, due to the large amount of shielding from the calorimeters, solenoid, and steel flux-return yoke, requiring the presence of such a signature in an event reduces the otherwise large contributions from background processes.

To reconstruct the decays of LLPs in the muon detector, the muon detector hits are clustered in η and the azimuthal angle ϕ using the DBSCAN algorithm [208], which groups hits by high-density regions.

The cluster reconstruction efficiency strongly depends on the LLP decay position. The efficiency is largest when the LLP decays near the edge of the shielding material, where there is enough material to induce the shower, but not so much that it stops the shower secondaries. The cluster reconstruction efficiency also depends on whether the LLP decays hadronically or leptonically. In general, hadronic showers have larger efficiency, because they are more likely to penetrate through the steel in between stations, while showers induced from electromagnetic decays generally occupy just one station and are stopped by the steel between stations. When the LLP decays near or in the CSCs, 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, the inclusive DT cluster reconstruction efficiency is approximately 80% for fully hadronic decays, 60% for $\tau^+\tau^-$ decays, and 45% for fully leptonic decays.

4.4.6 dE/dx

Studying anomalous ionization in the tracker provides a powerful tool to search for various LLP signals. For example, heavy charged particles are characterized by low speeds, which are inferred from time-of-flight measurements in muon chambers in case of sufficiently long lifetimes and large ionization signals in the tracker. Each layer of the silicon pixel and strip trackers of CMS provides a measurement of the charge deposit, which is transformed into a $\mathrm{d}E/\mathrm{d}x$ measurement after the application of a conversion factor from charge to energy and division by the path length. Dedicated estimators and discriminators have been designed to combine the set of $\mathrm{d}E/\mathrm{d}x$ measurements in the most appropriate way.

The I_h estimator, first used in a CMS search reported in Ref. [209], is defined as

$$I_\mathrm{h} = \left(\frac{1}{N} \sum_j^N c_j^{-2} \right)^{-1/2}. \tag{24}$$

This harmonic estimator is intended to provide the most probable value for the different $\mathrm{d}E/\mathrm{d}x$ (c_j) measurements that follow a Vavilov/Landau distribution. The sum in Eq. (24) includes all of the measurements along a track that have passed a cleaning procedure to discard measurements from atypical cluster deposit distributions and deposits too close to module edges. The I_h estimator is preferred to a simple measurement average as it is very robust against upward fluctuations in c_j. It is, however, sensitive to downward fluctuations, which are unlikely to randomly occur. This I_h estimator has been used for example to search for heavy charged particles considered as stable at the scale of the CMS detector [210], and also for charged particles with much shorter lifetimes leading to disappearing track signatures [211, 212].

In addition, the I_h estimator provides an estimate of the mass of the LLP candidate under the $Q = 1$e hypothesis. It uses an approximate Bethe–Bloch parameterization in the low relativistic regime that relates the measured ionization to the particle mass m and the track momentum p:

$$I_\mathrm{h} = K\frac{m^2}{p^2} + C, \tag{25}$$

where the empirical parameters K and C extracted from low-momentum tracks in the range $0.5 < p < 5\,\mathrm{GeV}$. Figure 18 shows this parametrization for the pions, kaons, protons, and deuterons at small momenta.

In addition to the I_h estimator, two independent discriminators are defined in Eqs. (26-27): F_i^Pixels, which uses only the $\mathrm{d}E/\mathrm{d}x$ pixel detector information, and G_i^Strip based on $\mathrm{d}E/\mathrm{d}x$ measurements in the strip tracker, where the i subscript refers to ionization. Both discriminators are designed to distinguish LLP signal events (with output values close to 1) from background events (with values close to 0).

The F_i^Pixels discriminator is defined as

$$F_i^\mathrm{Pixels} = 1 - \prod_{j=1}^{n} P_j \sum_{m=0}^{n-1} \frac{[-\ln(\prod_{j=1}^{n} P_j)]^m}{m!}, \tag{26}$$

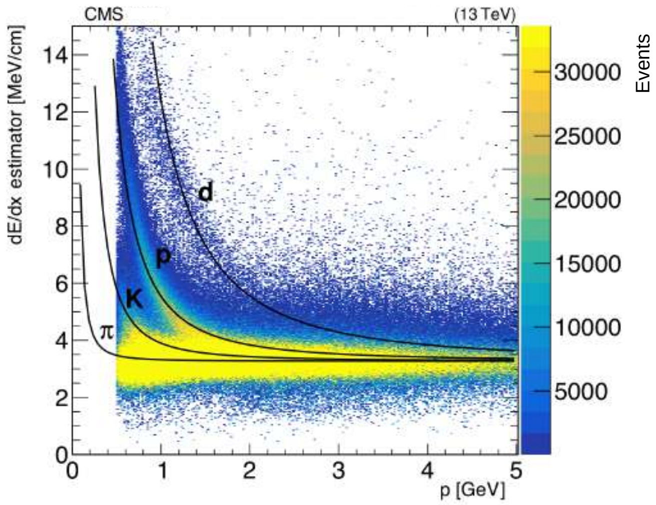

Figure 18: Distribution of the I_h estimator, computed using dE/dx measurements in the silicon strip tracker, versus the track momentum, using the data recorded in 2017 during the LHC Run 2. Expected dE/dx losses for pion, kaon, proton, and deuteron particles are shown as black lines. Tracks with $p_T < 0.5\,\text{GeV}$ are not included in this plot.

where n is the number of measurements in the silicon pixel detector, excluding the first barrel layer, and P_j is the probability that the minimum ionizing particle would produce a charge larger than or equal to the j–th measurement as predicted by a detailed simulation (called PIXELAV [213]) calibrated to data.

The G_i^{Strip} discriminator is defined as

$$G_i^{\text{Strip}} = \frac{3}{N}\left(\frac{1}{12N} + \sum_{j=1}^{N}\left[P_j\left(P_j - \frac{2j-1}{2N}\right)^2\right]\right), \qquad (27)$$

where N is the number of charge measurements in the silicon strip tracker, P_j is the probability for a minimum ionizing particle to produce a charge smaller or equal to the j–th charge measurement for the observed path length in the detector, and the sum is over the track measurements ordered in terms of increasing P_j. These P_j probabilities are determined using dE/dx templates in bins of path length values. The templates vary with detector module geometry and event PU. The probabilities are determined using data when used for data and determined using simulation when used for simulation.

These kinds of estimators can also address searches for particles with an electric charge different from unity [214]. For signals with a charge lower than unity, characterized in that case by a small dE/dx deposit, a large number of dE/dx measurements below a given threshold can be used to separate signal and background [215].

4.5 Precision proton spectrometer reconstruction

The CMS-TOTEM PPS [216] is a system of near-beam tracking and timing detectors, located in "Roman pots" at about 200 m on both sides of the CMS interaction point. The Roman pots are movable near-beam devices that allow the detectors to be moved very close (within a few mm) to the beam, directly into the beam vacuum pipe. The PPS is designed to search for the process $pp \to pp + X$ where the system X can involve SM or DS final states. It allows the measurement of the 4-momenta of scattered protons and their time-of-flight from the interaction point during standard running conditions in regular high-luminosity fills.

The proton momenta are measured by two tracking stations on each arm of the spectrometer. With the PPS setup, protons that lose approximately 3–15% of their momentum can be measured. This translates into an acceptance for the system X with a mass starting at $m_X \simeq 300\,\text{GeV}$ and up to about 2 TeV. The fractional momentum loss ξ of the protons can be measured from the proton track positions and angles (details can be found in Ref. [217]). The timing information that can be used to measure the longitudinal coordinate of the vertex via time-of-flight and suppress the background from PU is not used in the analyses discussed below.

A search using the PPS and the missing-mass technique will be described in Section 6.2.3.5.

4.6 Heavy ions

One of the main goals of the LHC as an energy-frontier pp collider is to discover new massive particles and/or FIPs. In addition to pp collisions, the LHC also provides high-energy HI collisions, and in particular lead-lead (PbPb) collisions, which are key tools to study the properties of the quark-gluon plasma.

Typically, one does not consider HI collisions as a place to look for BSM physics. They are characterized by a very large number of outgoing particles (a charged-particle multiplicity more than two orders of magnitude larger than in pp collisions [218]), which makes tracking much more challenging. Moreover, the integrated luminosity ($\mathcal{L}_{\text{int}}$) for PbPb collisions was $390\,\mu\text{b}^{-1}$ and $1650\,\mu\text{b}^{-1}$, respectively in 2015 and 2018, which is many orders of magnitude smaller than for pp collisions.

However, a fraction of HI interactions takes place with no overlap between the two nuclei. In such ultra-peripheral collisions (UPCs), the two ions only interact through the electromagnetic force, i.e., an exchange of photons, producing very low multiplicity events. Additionally, HI runs are tuned to yield no PU, which further simplifies tracking. Overall, UPCs result in extremely clean event signatures, suitable for BSM physics searches, e.g., searches for ALPs.

The CMS experiment is well equipped to record and investigate both pp and HI collisions. The main challenges in UPCs of heavy ions from the experimental perspective are related to triggering and detector noise. For instance, in light-by-light scattering (discussed in Section 6.2.3.4) the final state consists exclusively of two low-energy photons. Since there is no other activity in the detector, one cannot rely on associated tracks, muons, jets, or $p_{\text{T}}^{\text{miss}}$ to trigger the measurement. Instead, the photons themselves have to be used for triggering. For such rare processes, it is crucial to lower the photon energy requirement for both triggering and offline reconstruction as much as possible, which enters the regime where calorimeter noise becomes significant. As an example, a recent light-by-light scattering analysis, described in Section 6.2.3.4, triggers on diphotons with transverse energy $>2\,\text{GeV}$. The noise in the barrel region of the ECAL is at the level of $\approx 0.7\,\text{GeV}$. However, in the endcap region it can get as large as $\approx 6\,\text{GeV}$.

These challenges associated with UPCs are typically addressed by carefully studying triggering and reconstruction efficiencies with tag-and-probe techniques, as well as masking regions of the detector where noise levels are too large to perform the analysis. This strategy yields satisfactory results, allowing CMS to observe evidence for light-by-light scattering and derive the most stringent limits, at the time of publication, on the production of ALPs with masses between 5 and 50 GeV [219].

4.7 Background estimation strategies and statistical methods

For most of the searches presented in this Report, the CL_s method [220, 221] is used to obtain a limit at 95% confidence level (CL) using the profile likelihood test statistic [222], often in the asymptotic approximation. The CMS statistical analysis tool COMBINE [223] is used to compute these limits. The robustness and precision of the estimation of contributions from SM background processes determine the sensitivity of searches for new physics. Historically, simulated background events obtained with the Monte Carlo (MC) method have been used most of the time to seed templates for background contributions in the signal region (SR) and obtain uncertainties. These systematic uncertainties in the MC distributions are represented as nuisance parameters that are adjusted in a maximum likelihood fit, based on the observed data distribution, to obtain the final background model. In many cases, however, methods based on CRs in data or more sophisticated background estimation strategies are employed to quantify background contributions in the SRs. In the following, a few of the common methods of background estimation either fully based on CRs in data, or partially based on data and assisted through the simulation, are briefly introduced.

4.7.1 Transfer factor technique

The underlying idea of using transfer factors (TFs) to predict background contributions is to measure ratios of yields for processes across regions, rather than calibrate the absolute background shape. As a consequence, if the two samples used to build the ratios are impacted by a specific systematic uncertainty in the same or a similar way, its effect largely cancels out and does not affect the ratios. For instance, it is conceivable to assume that an event sample with a jet recoiling against a dilepton system and an event sample featuring a jet recoiling against a single photon will have the same uncertainties affecting the measurement of the jet, i.e., jet energy scale and resolution. Thus, the ratio of the (differential) yields in these two samples is largely unaffected by jet uncertainties, while being affected by lepton/photon identification and scale uncertainties.

This strategy is particularly powerful when applied in mono-X-type analyses, where the $p_{\mathrm{T}}^{\text{miss}}$ spectrum is a powerful shape discriminator between the BSM signal and the SM background and is typically used for signal extraction. Because of the symmetry of the various SM V+jets processes, the main background contribution in the SR coming from the $\mathrm{Z}(\nu\overline{\nu})$+jets process can be calibrated utilizing CRs enriched in $\mathrm{Z}(\ell\ell)$+ jets, $\mathrm{W}(\ell\nu)$+jets, and γ+jets events. By excluding the leptons and photons from the computation of $p_{\mathrm{T}}^{\text{miss}}$ in the CR, the so-called hadronic recoil becomes a proxy for the $p_{\mathrm{T}}^{\text{miss}}$ spectrum in the SR.

A binned likelihood fit to the data is performed simultaneously in different CRs and in the SRs to estimate the dominant $\mathrm{Z}(\nu\overline{\nu})$+jets and $\mathrm{W}(\ell\nu)$+jets backgrounds in each $p_{\mathrm{T}}^{\text{miss}}$ bin.

The part of the likelihood function constraining the $\mathrm{Z}(\nu\overline{\nu})$+jets background in the monojet analysis in Ref. [81], which is representative of other mono-X-type searches, is given as:

$$
\begin{aligned}
\mathcal{L}_{\mathrm{c}}(\boldsymbol{\mu}^{\mathrm{Z}(\nu\overline{\nu})},\boldsymbol{\mu},\boldsymbol{\theta}) = & \prod_i P\left(d_i^{\gamma}|B_i^{\gamma}(\boldsymbol{\theta}) + \frac{\mu_i^{\mathrm{Z}(\nu\overline{\nu})}}{R_i^{\gamma}(\boldsymbol{\theta})}\right) \\
& \times \prod_i P\left(d_i^{\mathrm{Z}}|B_i^{\mathrm{Z}}(\boldsymbol{\theta}) + \frac{\mu_i^{\mathrm{Z}(\nu\overline{\nu})}}{R_i^{\mathrm{Z}}(\boldsymbol{\theta})}\right) \\
& \times \prod_i P\left(d_i^{\mathrm{W}}|B_i^{\mathrm{W}}(\boldsymbol{\theta}) + \frac{f_i(\boldsymbol{\theta})\mu_i^{\mathrm{Z}(\nu\overline{\nu})}}{R_i^{\mathrm{W}}(\boldsymbol{\theta})}\right) \\
& \times \prod_i P\left(d_i|B_i(\boldsymbol{\theta}) + (1+f_i(\boldsymbol{\theta}))\mu_i^{\mathrm{Z}(\nu\overline{\nu})} + \mu S_i(\boldsymbol{\theta})\right).
\end{aligned}
\tag{28}
$$

In the above likelihood function, $P(n|x)$ is the Poisson probability of observing n events when

x are expected, $d_i^{\gamma/\mathrm{Z}/\mathrm{W}}$ is the observed number of events in each bin of the photon, dimuon/dielectron, and single-muon/single-electron CRs, and $B_i^{\gamma/\mathrm{Z}/\mathrm{W}}$ is the background in the respective CRs. The systematic uncertainties are modeled with nuisance parameters ($\boldsymbol{\theta}$), which enter the likelihood as additive perturbations to the TFs $R_i^{\gamma/\mathrm{Z}/\mathrm{W}}$. Each $\boldsymbol{\theta}$ parameter has an associated Gaussian constraint term in the full likelihood. The parameter $\boldsymbol{\mu}^{\mathrm{Z}(\nu\overline{\nu})}$ represents the yield of the $\mathrm{Z}(\nu\overline{\nu})$ background in the SR and is left freely floating in the maximum likelihood fit. The function $f_i(\boldsymbol{\theta})$ is the TF between the $\mathrm{Z}(\nu\overline{\nu})$+jets and $\mathrm{W}(\ell\nu)$+jets backgrounds in the SR and acts as a constraint between these backgrounds. The likelihood also includes the SR, with B_i representing all the background estimates from simulation, S representing the nominal signal prediction, and μ being the signal strength parameter also left floating in the case of an $S + B$ fit ($\mu = 0$ otherwise).

In this likelihood, the expected numbers of $\mathrm{Z}(\nu\overline{\nu})$+jets events in each bin of $p_{\mathrm{T}}^{\mathrm{miss}}$ are the free parameters of the fit. Transfer factors, derived from simulation, are used to link the yields of the $\mathrm{Z}(\ell\ell)$+ jets, $\mathrm{W}(\ell\nu)$+jets, and γ+jets processes in the CRs with the $\mathrm{Z}(\nu\overline{\nu})$+jets and $\mathrm{W}(\ell\nu)$+jets background estimates in the SR. These TFs are defined as the ratio of expected (from simulation) yields of the target process in the SR and the process being measured in the CR, e.g.:

$$R_i^{\mathrm{Z}} = \frac{N_{i,\mathrm{MC}}^{\mathrm{Z}(\mu\mu)}}{N_{i,\mathrm{MC}}^{\mathrm{Z}(\nu\overline{\nu})}}. \tag{29}$$

To estimate the $\mathrm{W}(\ell\nu)$+jets background in the SR, the TFs between the $\mathrm{W}(\ell\nu)$+jets background estimates in the SR and the $\mathrm{W}(\mu\nu_\mu)$+jets and $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}})$+jets event yields in the single-lepton CRs are constructed. These TFs take into account the impact of lepton acceptances and efficiencies, lepton veto efficiencies, and the difference in the trigger efficiencies in the case of the single-electron CR.

The $\mathrm{Z}(\nu\overline{\nu})$ background prediction in the SR is connected to the yields of $\mathrm{Z}(\mu\mu)$ and $\mathrm{Z}(\mathrm{ee})$ events in the dilepton CRs. The associated TFs account for the differences in the branching fraction of Z bosons to charged leptons relative to neutrinos and the impact of lepton acceptance and selection efficiencies. In the case of dielectron events, the TF also takes into account the difference in the trigger efficiencies. The resulting constraint on the $\mathrm{Z}(\nu\overline{\nu})$+jets process from the dilepton CRs is limited by the statistical uncertainty in the dilepton CRs because of the large difference in branching fractions between Z boson decays into neutrinos and Z boson decays into electrons and muons.

The γ+jets CR is also used to predict the $\mathrm{Z}(\nu\overline{\nu})$+jets process in the SR through a TF, which accounts for the difference in the cross sections of the γ+jets and $\mathrm{Z}(\nu\overline{\nu})$+jets processes, the effect of acceptance and efficiency of identifying photons along with the difference in the efficiencies of the photon and $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers. The addition of the γ+jets CR mitigates the impact of the limited statistical power of the dilepton constraint, because of the larger production cross section of γ+jets process compared to that of $\mathrm{Z}(\nu\overline{\nu})$+jets process.

Finally, a TF is also defined to connect the $\mathrm{Z}(\nu\overline{\nu})$+jets and $\mathrm{W}(\ell\nu)$+jets background yields in the SR, to further benefit from the larger statistical power that the $\mathrm{W}(\ell\nu)$+jets background provides, making it possible to experimentally constrain $\mathrm{Z}(\nu\overline{\nu})$+jets production at large $p_{\mathrm{T}}^{\mathrm{miss}}$.

These TFs rely on an accurate prediction of the ratio of Z+jets, W+jets, and γ+jets cross sections. Therefore, leading order (LO) simulations for these processes are corrected using boson p_{T}-dependent next-to-LO (NLO) QCD K-factors derived using MADGRAPH5_aMC@NLO. They are also corrected using p_{T}-dependent higher-order EW corrections extracted from theoretical calculations [224–229]. The higher-order corrections are found to improve the data-to-simulation agreement for both the absolute prediction of the individual Z+jets, W+jets, and γ+jets processes, and their respective ratios.

4.7.2 Bump-hunt technique

Any new mediator particle X predicted in BSM scenarios has several experimental observables, including its rest mass m_X, its decay width Γ_X, and its production cross section σ_X. If the mediator decays into SM particles or a mixture of SM and DM particles, its rest mass can be measured by determining the energy and angle of emission of all its decay products. The mass spectrum of its decay products is expected to show an increase in the number of event counts at the “resonance” m_X value because of the enhancement in the production cross section from the propagator of a massive mediator. The width of the resonance, or “bump” in the reconstructed mass spectrum, will depend on the decay interactions and the detector resolution that measures the decay products. For strong (or strong-like) interactions, with short lifetimes, the resonance shape may be wide (larger than the experimental resolution). Its shape can be approximated by a Breit–Wigner function for the intrinsic line shape, convoluted with a Gaussian function for the resolution. Parton luminosities are greater for masses below the resonance peak, such that the Breit–Wigner shape can present a significant “shoulder“ on the lower tail. This effect may be significant near the kinematic threshold of m_X production.

In some cases, a full reconstruction of m_X is impossible since the decays include invisible particles from DM candidates. In those cases, it is important to include the $\vec{p}_T^{\,\text{miss}}$ in the definition of the reconstructed m_X, such as m_{T2} [230] or the razor variable R [231, 232]. For example, in the case of SVJs $Z' \to q_{\text{dark}}\overline{q}_{\text{dark}}$, cf. Section 2.2.4.1, the invariant mass of the reconstructed (visible) jets m_{jj} is a worse proxy for $m_{Z'}$ than m_T defined to include the $\vec{p}_T^{\,\text{miss}}$ [148]:

$$m_T^2 = \left[E_{T,jj} + E_T^{\text{miss}}\right]^2 - \left[\vec{p}_{T,jj} + \vec{p}_T^{\,\text{miss}}\right]^2 = m_{jj}^2 + 2p_T^{\text{miss}}\left(\sqrt{m_{jj}^2 + p_{T,jj}^2} - p_{T,jj}\cos(\phi_{jj,\text{miss}})\right). \quad (30)$$

Here, m_{jj} is the invariant mass of the system composed of the two largest-p_T large-radius jets, and $\vec{p}_{T,jj}$ is the vector sum of their $\vec{p}_T$. The quantity $E_{T,jj}^2 = m_{jj}^2 + |\vec{p}_{T,jj}|^2$, while it is assumed that the system carrying the p_T^{miss} is massless, i.e., $E_T^{\text{miss}} = p_T^{\text{miss}}$. This enables the simplification in the second line of Eq. (30), with $\phi_{jj,\text{miss}}$ as the azimuthal angle between the dijet system and the $\vec{p}_T^{\,\text{miss}}$. In this case, m_T is much closer to $m_{Z'}$ than m_{jj}: it has better resolution and its peak reproduces $m_{Z'}$ more accurately.

The estimation of the background is critical when looking for a bump in the reconstructed mass spectrum is the estimation of the background. In contrast to the signal, the background (typically QCD multijet) spectrum is smoothly falling. Despite the progress of QCD multijet MC generators with NLO and next-to-NLO (NNLO) accuracy, the mass spectra obtained from MC generators tend not to agree very well with the data in both shape and normalization. This is caused by the large theoretical uncertainties (such as nonperturbative effects, parton distribution functions [PDFs], and the renormalization and factorization scales) and experimental uncertainties (such as the jet energy scale and resolution smearing), which can be even more pronounced in final states with large $\vec{p}_T^{\,\text{miss}}$ where misreconstructed SM jets are the dominant background. Therefore, many searches estimate the QCD multijet background parametrically, directly from data. The fit can include templates from signal (at different mass values) or a parameterized signal function, and other components for background. If no significant deviation from the background-only hypothesis is found, limits on the cross section as a function of m_X can be set. Using the data to describe the background solves the problem of poor modeling of detector effects in novel signatures, although limited event counts at large invariant mass may become a problem.

At the LHC, several families of fit functions have been used to model the QCD multijets background, which are called the "dijet function" (f_{dijet} and its enhanced version f_{dijet2}) and the "UA2 function" (f_{UA2}):

$$\begin{aligned} f_{\text{dijet}}(x) &= \frac{p_0(1-x)^{p_1}}{x^{p_2+p_3\ln x+p_4\ln^2 x}}, \\ f_{\text{dijet2}}(x) &= \frac{p_0(1-x)^{p_1+p_2\ln x+p_3\ln^2 x}}{x^{p_4+p_5\ln x+p_6\ln^2 x}}, \\ f_{\text{UA2}}(x) &= \frac{p_0 \mathrm{e}^{-p_1 x - p_2 x^2}}{x^{p_3[1+p_4\ln x+p_5\ln^2 x]}}. \end{aligned} \tag{31}$$

Here x is the reconstructed mass divided by $\sqrt{s}$. These families of functions have been found in the past to fit the observed QCD spectrum in hadron colliders [148, 233–236]. The number of parameters p_N used in each function must be optimized in each case. The Fisher test [237, 238] can determine if adding a new parameter to a function improves the fit to a given distribution. Two functions (one with fewer parameters than the other) are fit to the same distribution and the value

$$F_{\text{test}} = \frac{(q_1 - q_2)/(n_2 - n_1)}{q_2/(n_{\text{bins}} - n_2)}$$

is calculated, where q_i, n_i refer to the goodness-of-fit measurement and number of parameters in each function ($n_1 < n_2$), and n_{bins} is the number of bins in the distribution. The goodness-of-fit parameter is usually the χ^2 value, which has been observed to give more stable results than the residual sum of squares. The value of F_{test} is then compared to F_{crit}, which is defined by $\int_{F_{\text{crit}}}^{\infty} F_{\text{dist}} dx = \alpha_{\text{crit}}$, where F_{dist} is an F-distribution with $n_2 - n_1$ and $n_{\text{bins}} - n_2$ degrees of freedom and $\alpha_{\text{crit}} = 0.05$. If $F_{\text{test}} > F_{\text{crit}}$, the function with more parameters (n_2) provides a better fit than the function with fewer parameters (n_1). The value of α_{crit} may be adjusted depending on the result of the bias tests, described next, and the stability of the results.

This way of estimating the background from a fit to the data will typically be one of the largest experimental uncertainties in the statistical analysis to extract the signal. We typically assign the statistical uncertainty in the fit parameters as a background shape systematic uncertainty, and this tends to be large for large values of the reconstructed mass. It is also very important to test alternate functions to describe the QCD multijet background and check if using them introduces a bias in the results because the data in reality follows a different distribution from what was chosen for the fit. Some analyses use discrete profiling to estimate the uncertainty from different background functions and possible bias [239]. Some possible alternate functions are listed here [179, 240, 241]:

$$\begin{aligned} f_{\text{polynomial}}(x) &= \frac{p_0}{(1+p_1x+p_2x^2+p_3x^3)^{p_4}}, \\ f_{\text{extended polynomial}}(x) &= \frac{p_0(1-x)^{p_1}(1+p_2x+p_3x^2)}{x^{p_4+p_5\ln x}}, \\ f_{\text{power-law times exponential}}(x) &= \frac{p_0 e^{-p_1 x}}{x^{p_2}}, \\ f_{\text{other}}(x) &= \frac{p_0(1-x^{1/3})^{p_1}}{x^{p_2}}. \end{aligned} \tag{32}$$

A self-closure test can be performed by generating pseudo-experiments with the main background function and fitting them with the same function to extract a signal measurement.

Of course, the result here should be zero signal, but the spread in the results measures how robust the main function is to data fluctuations. This can be compared (and the corresponding uncertainty estimated) with a bias-closure test in which the main function is used to fit pseudo-experiments now generated with the alternate function. The results again should yield zero signal, and will tell us if our choice of background function has any potential to bias our results: if the self and bias-closure tests agree within their uncertainties, then no additional systematic uncertainties need to be included for this background estimation method. In addition, one can perform similar tests by injecting signal in both tests at the time of generating the pseudo-experiments and observing if the sensitivity to the signal also behaves similarly in both cases.

An alternative strategy to model the background without empirical functions is to measure the observed distribution in a CR and derive correction factors from simulation to account for differences between the CR and the SR. This method can have smaller uncertainties than methods using empirical functions, but it can only be employed when the CR is not biased by trigger requirements.

4.7.3 The "ABCD" method

Background estimations based on CRs in data are often used for more reliable descriptions of backgrounds. One of the most widely used such methods is the matrix ("ABCD") method, which was first introduced in Ref. [242]. An example of how this method is used in a CMS analysis is shown in Fig. 19. The ABCD method uses two independent variables to define four statistically independent regions, including the SR D and CRs A, B, and C. The two variables that are used to define the ABCD plane need to be statistically independent for the background process, allowing the prediction of the background yield in the SR to be constrained by the background yield in CRs A, B, and C: $N_{\mathrm{D}} = N_{\mathrm{B}} N_{\mathrm{C}} / N_{\mathrm{A}}$, where N_{X} is the number of background events in region X. Ideally, the CRs should be enriched with background events and devoid of signal events.

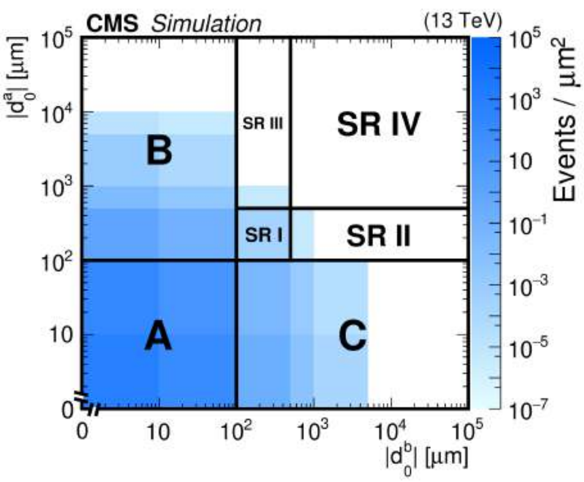

Figure 19: A diagram of the ABCD method, shown for illustration on simulated background events in a search for LLPs that decay to displaced leptons. The CRs are regions A, B, and C. There are four SRs, labeled I–IV, in this search. Figure taken from Ref. [243].

In cases where there is potential signal contamination in the CRs, a binned maximum likelihood fit is performed simultaneously in the four bins, with the signal strength included as a floating parameter. The background yields in the four regions are constrained to obey the standard ABCD relationship. This is possible because the background yields in the four regions require only three parameters to be fully described, given the independence of the two variables defining the ABCD plane. Thus one degree of freedom remains, which is used to fit the signal strength across all regions. Systematic uncertainties that impact the signal and background yields are treated as nuisance parameters with log-normal probability density functions.

Potential small correlations between the two variables defining the ABCD plane can be understood and controlled with additional validation regions adjacent to the SR [244]. These regions are located in between the corresponding CR and the SR in the ABCD plane and provide a path to estimate the correlation between the two observables.

Additionally, CMS explores the usage of machine-learning-assisted ABCD techniques to derive discriminators that are decorrelated from a variable of interest or from another discriminator following the distance correlation technique proposed in Ref. [245].

5 Data set and signal simulation

Most of the analyses presented in this Report use the Run 2 pp collision data sample, corresponding to $\mathcal{L}_{\text{int}}$ up to $140\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$, collected by the CMS detector in 2016–2018. The $\mathcal{L}_{\text{int}}$ for the 2016, 2017, and 2018 data-taking years have 1.2–2.5% individual uncertainties [246–248], while the overall uncertainty in $\mathcal{L}_{\text{int}}$ for the 2016–2018 period is 1.6%. Some analyses use Run 1 pp collision data, taken in 2010–2012 with $\sqrt{s} = 7$ and 8 TeV, or Run 3 pp collision data, taken in 2022 with $\sqrt{s} = 13.6\,\text{TeV}$. Finally, some analyses use Run 2 HI collision data, namely, PbPb collisions taken in 2015 with $\sqrt{s_{_{\text{NN}}}} = 5.02\,\text{TeV}$.

Data sets of simulated events, for both the SM background and BSM signals, are used by the searches to optimize the analysis criteria for sensitivity as well to check the agreement with data for basic kinematic variables. The simulation of collision events is implemented through a fixed-order perturbative calculation of up to four noncollinear high-p_{T} partons for the QCD terms, supplemented with a description of the underlying event, parton showering, multiparton interactions and hadronization. The perturbative calculation step is usually performed by a matrix-element calculator and event generator; versions 2.2.2 and 2.6.5 of the MADGRAPH5_aMC@NLO [249] package are used for almost all the analyses presented in this Report, and POWHEG v2 [250–252] is also used for certain processes, primarily single top, $\text{t}\bar{\text{t}}$, and Higgs boson production. The next step is, in turn, usually implemented by the PYTHIA 8 [253] generator. The combination of the two steps is a delicate procedure; a matching procedure is implemented to avoid double-counting of processes in the combination, with the exact recipe depending on the order of the perturbative calculation. The MLM matching [254] is used for LO calculations, while the FxFx [255] and POWHEG [251] methods are used for NLO. The PDFs are used to map the simulated colliding protons to the initial-state partons that are present in the matrix-element calculation; conversely, the PYTHIA parameters are adjusted to a set of values that better describe the observed dynamics of high-energy proton collisions, which is referred to as a tune. By the end of Run 2, most analyses discussed in this Report converged in the usage of the NNPDF3.1 NNLO PDFs [256] and the CP5 tune [257]. The simulation of specific new physics models may differ in particular aspects of these steps.

The detector response to simulated particles is modeled using the GEANT4 software [258]. Custom simulations of the detector electronics are used to produce readouts similar to those observed in data, in a process known as digitization. Pileup interactions are also included in the simulation. The simulated samples are corrected to make the PU distribution match the distribution in data as closely as possible. Event generation of new physics processes may need modifications to any of the steps of the simulation. The most notable case is the treatment of LLPs; the mass, charge, interactions, and lifetime of those particles are relayed to GEANT4, in a manner consistent with its treatment by the previous steps.

When simulating dark QCD models, the dedicated HV module in PYTHIA 8 is used for showering and hadronization in the DS. PYTHIA version 8.230 or higher is used to access important features, such as the running of the dark coupling. In earlier versions, these features were added by patching the source code [150]. Additional modifications to PYTHIA are required to simulate the flavored emerging jet model [151]. The SUEPs are simulated using a custom PYTHIA module that produces dark hadrons according to a Boltzmann distribution [152]. Dark-hadron properties, including branching fractions and lifetimes, are computed separately and specified in the PYTHIA configuration as needed for each signal model. In particular, r_{inv} for SVJ models is implemented by reducing the branching fractions to SM quarks for all dark hadron species; dark hadrons that do not decay into SM quarks are marked as stable. Because stable dark hadrons must be produced in pairs (in order to conserve quantum numbers), events with an odd number of stable dark hadrons are rejected. For the dark QCD signal models studied in this Report, PYTHIA is used for the LO matrix-element calculations as well as hadronization and showering. For other models, such as those requiring processes not implemented in PYTHIA or more accurate simulation of ISR, DS particles produced by MADGRAPH5_aMC@NLO can be interfaced with PYTHIA for hadronization and showering in both the DS and the SM [115].

In the following results, for some models, we present a minimum allowed coupling that will satisfy relic density constraints. Typically, there is a minimum allowed coupling between the standard model and the DS. For couplings smaller than the minimum, which would have earlier freeze-out times, the DM production in the early universe would exceed the observed DM, as measured by the Planck experiment [1]. The minimum coupling can be determined by computing the relic density for various coupling values and scanning over the range of values to yield the smallest that satisfies the observed constraint. To perform the relic density calculation, we use the MADDM 3.0 software framework [259] with the appropriate MADGRAPH5_aMC@NLO signal models for the quoted searches. For fixed DM and mediator masses and a fixed DM coupling (typically $g_{\text{DM}} = 1$), the relic density follows a parabolic form, allowing the minimum allowed coupling to be determined through a coupling scan.

6 Signatures

The CMS Collaboration has a broad program of searches for models of BSM physics that provide DM candidates; an overview of the theoretical framework for these models is provided in Section 2. In this section, we briefly discuss the details of each search and the signatures of the models targeted. The sensitivity of a broad range of signatures to DSs is probed, and no significant excess of events is observed over the background predictions. These searches are categorized by their final states: invisible, prompt final states are summarized in Section 6.1; visible, prompt final states in Section 6.2; and displaced and long-lived signatures in Section 6.3. It is notable that many general categories of theoretical models can potentially present any of these final states. For example, strongly coupled hidden sectors can produce SVJs with invisible final states, SUEPs with visible final states, EJs with displaced final states, or potentially mixtures of these novel objects. Further, there may be deep connections between different final states: any mediator produced via an SM process can decay into the same SM particles, leading to a visible final state. Therefore, investigation of the visible final state can help exclude other final states without depending on the detailed phenomenology. These considerations motivate the breadth and continued expansion of the CMS search program, as the nature of DM remains

6.1.3 Signatures from hidden valley models

As described in Section 2.2.4, some HV models predict unique signatures from a QCD-like force in the DS with corresponding dark quarks ($\mathrm{q_{dark}\overline{q}_{dark}}$). When produced at the LHC, dark quarks shower and hadronize in the DS giving rise to dark jets made of stable and unstable dark hadrons. While stable dark bound states do not interact with the detector, unstable ones decay promptly to SM quarks. This leads to an SVJ made of collimated visible and invisible particles.

6.1.3.1 Search for semivisible jets

A search is performed [148] for SVJs using data collected during Run 2 and corresponding to $\mathcal{L}_{\text{int}} = 138\,\text{fb}^{-1}$. Resonant production of a leptophobic Z′ mediator decaying into dark quarks, $\mathrm{q\overline{q}} \to \mathrm{Z'} \to \mathrm{q_{dark}\overline{q}_{dark}}$, leads to a final state with two SVJs. The $\vec{p}_{\mathrm{T}}^{\text{miss}}$ is aligned with one of the jets, as shown in Fig. 28, and has a moderate magnitude. Both jets carry a fraction of invisible momentum, leading to a partial cancellation when the jets are back-to-back. The SVJs are expected to be larger than typical SM jets, because they arise from a double parton shower and hadronization process: first in the DS and later in the SM sector. Depending on the parameter of the model, the signature can vary significantly. We assess models with $1.5 \leq m_{\mathrm{Z'}} \leq 5.1\,\text{TeV}$, $1 \leq m_{\text{dark}} \leq 100\,\text{GeV}$, and $0 \leq r_{\text{inv}} \leq 1$. Because of the invisible momentum carried by stable dark hadrons, the mass of the mediator cannot be fully reconstructed. Instead, a bump hunt is performed using the transverse mass m_{T} of the dijet system and the $p_{\mathrm{T}}^{\text{miss}}$. The SM backgrounds, dominated by QCD multijets with artificial $p_{\mathrm{T}}^{\text{miss}}$ but also including significant fractions of $\mathrm{t\overline{t}}$, W+jets, and Z+jets processes with genuine $p_{\mathrm{T}}^{\text{miss}}$ from neutrinos, are expected to have steeply falling m_{T} distributions. Two versions of the search are performed: an inclusive search using only selections on event-level kinematic variables, and a model-dependent search using a BDT trained on specific signal models to identify SVJs.

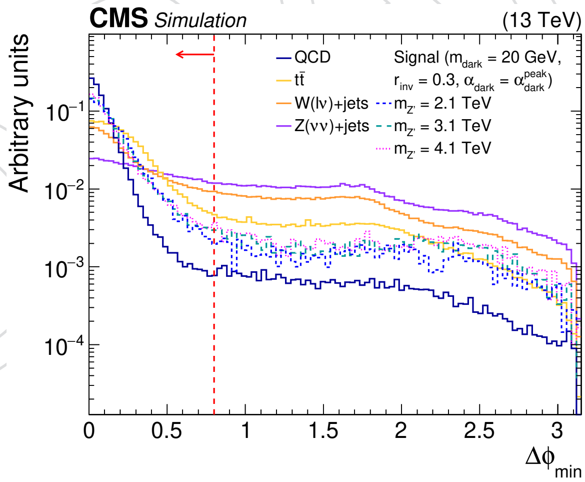

Figure 28: The normalized distribution of the minimum azimuthal angle between the $\vec{p}_{\mathrm{T}}^{\text{miss}}$ and each of the two leading jets ($\Delta\phi_{\text{min}}$) for simulated SM backgrounds and several SVJ signal models. The red vertical dotted line indicates the selection requirement on this variable. Figure taken from Ref. [148].

The SVJ models with extreme values of r_{inv}, close to 0 or 1, overlap with the phase space of dijet resonance searches (Section 6.2.2.2) and monojet DM searches (Section 6.1.1.1). Hence, we can reinterpret these two searches for the SVJ signal model. Accordingly, the DM coupling in the dark QCD model is set to $g_{\mathrm{q_{dark}}} = 0.5$ in order for the Z′ boson to have width and branching fractions consistent with the LHC DM Working Group benchmark model for simplified DM,

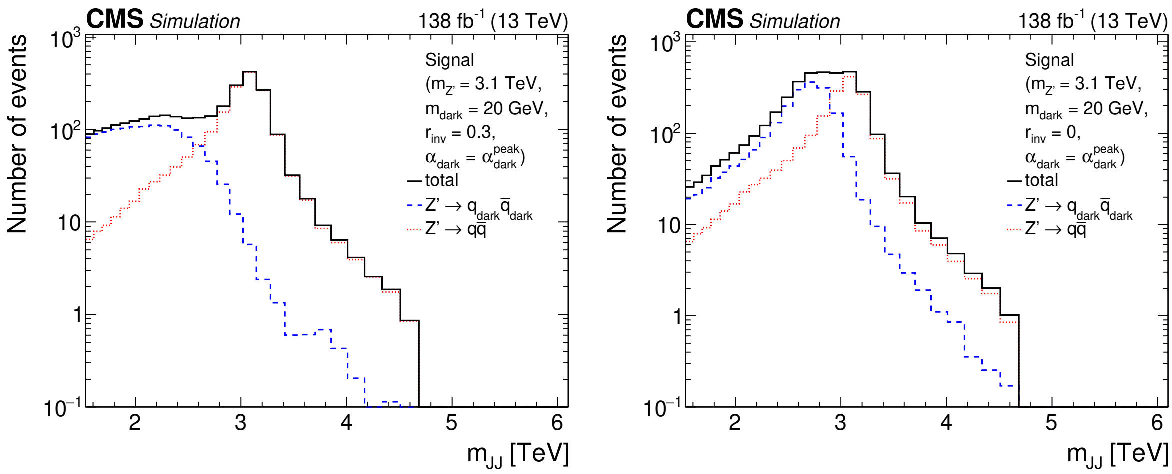

Figure 29: The dijet mass distributions for the combination of $Z' \rightarrow q_{dark}\bar{q}_{dark}$ and $Z' \rightarrow q\bar{q}$ events, for $r_{inv} = 0.3$ (left) and $r_{inv} = 0.0$ (right), in SVJ signal models.

as noted in Section 2.2.4. Because both possible final states have visible components, Fig. 29 shows the dijet mass distributions from $Z' \rightarrow q_{dark}\bar{q}_{dark}$ and $Z' \rightarrow q\bar{q}$, both individually and summed, in the correct proportions for the specified coupling values. For $r_{inv} = 0.3$, the $Z' \rightarrow q_{dark}\bar{q}_{dark}$ events have relatively lower dijet mass values, so they do not contribute substantially to the sensitivity of a dijet resonance search, which remains dominated by $Z' \rightarrow q\bar{q}$ events. However, for $r_{inv} = 0.0$, the two contributions to the dijet mass distribution are similar enough that the summed distribution is enhanced around the resonant peak, providing correspondingly greater sensitivity. The remaining minor degradation in the $Z' \rightarrow q_{dark}\bar{q}_{dark}$ dijet mass distribution primarily occurs because of the presence of neutrinos from decays of heavy-flavor hadrons, which are more likely to be produced in SVJs than in SM jets.

For a reinterpretation of the monojet DM search for the SVJ model, it is important to note that the efficiency of triggering on p_T^{miss}, which imposes an offline requirement of $p_T^{miss} > 250\,\text{GeV}$, is maximized for $r_{inv} = 0.5$, as shown in Fig. 30. At higher r_{inv} values, the majority of dark hadrons are stable and invisible, leading to increased cancellation of the invisible momenta of the two jets from the Z' boson decay, which correspondingly reduces the transverse component. However, the efficiencies of several other requirements are maximized for $r_{inv} = 1.0$: $\Delta\phi(\vec{p}_T^{\,jet}, \vec{p}_T^{\,miss}) > 0.5$ for the leading four jets with $p_T > 30\,\text{GeV}$, and $N_{b\text{-jet}} = 0$ considering all jets with $p_T > 20\,\text{GeV}$. As r_{inv} increases, fewer dark hadrons decay into visible particles, decreasing the number of possible reconstructed jets in each event; since visible and invisible dark hadrons are produced together in collimated sprays, any reconstructed jets may be aligned with p_T^{miss}. At $r_{inv} = 1.0$, the only visible particles in the signal events come from ISR. SVJs tend to be enriched in b hadrons because of the higher mass scale of dark hadrons compared to SM quarks, which enables them to decay into $b\bar{b}$ pairs. In the models considered here, $m_{dark} = 20\,\text{GeV}$, leading to $\mathcal{B}(\rho_{dark} \rightarrow b\bar{b}) = 0.2$ and $\mathcal{B}(\pi_{dark} \rightarrow b\bar{b}) = 0.94$. The signal model specifies that ρ_{dark} are produced 75% of the time, leading to an overall branching fraction of 0.385 for any unstable dark hadron to decay into b quarks. The relative efficiencies for these requirements are also presented in Fig. 30.

We present results for the two reinterpretations in Section 7.2.4.1.

6.2 Fully visible and prompt signatures

In addition to searching for decays into invisible final states as described in Section 6.1, the DS can also be probed by searching for decays of the mediator to SM particles and fully visible

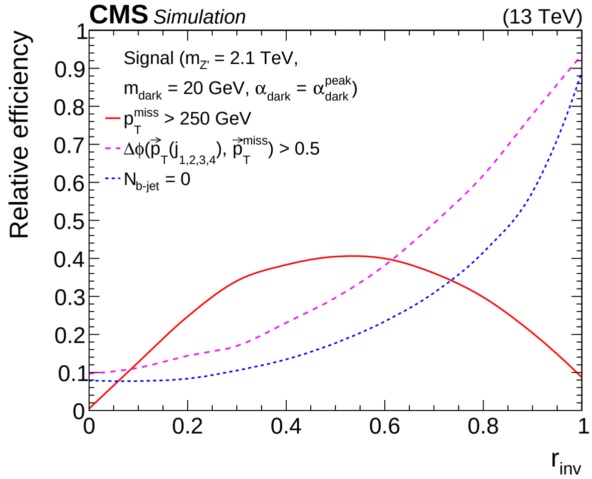

Figure 30: The relative efficiencies of several selection criteria from the monojet search for SVJ signals. The efficiencies of the $\Delta\phi(\vec{p}_T^{\,jet}, \vec{p}_T^{\,miss})$ and $N_{\text{b-jet}}$ requirements are evaluated after the $p_T^{miss} > 250$ GeV requirement. The uncertainty in the simulation is negligible.

final states. For example, we can search for mediators that decay into pairs of leptons or jets. These searches provide results that are complementary to the invisible decays described above. We organize this section into searches for low-mass resonances (Section 6.2.1), i.e., resonances below several hundreds of GeV; searches for high-mass resonances (Section 6.2.2), i.e., resonances above several hundred GeV; and searches with other prompt and visible signatures that do not easily fit into these two categories (Section 6.2.3).

6.2.1 Low-mass resonance searches

Searches for low-mass dijet resonances [277] are strongly limited by the trigger bandwidth because of overwhelming background rates. The triggers, listed in Table 3, result in a lower threshold of 1.8 TeV on the resonance masses probed by conventional dijet resonance searches.

Table 3: Trigger thresholds for various jet-based triggers in Run 2. All values are in GeV.

Trigger	2016	2017	2018
H_T	800, 900	1050	1050
AK4 PF jet p_T	450	500	500
AK8 PF jet p_T	450	500	500
AK8 PF jet p_T (m_{trim})	360 (30)	400 (30)	400 (30)
Single AK4 calo jet p_T	500	500	500

The CMS Collaboration has utilized a number of techniques to circumvent this limitation:

- Resonances with masses as small as 600 GeV can be probed with the data scouting technique [278], wherein the trigger thresholds are lower by saving to disk only high-level physics objects, i.e., jets clustered from calorimeter towers or particle flow candidates, rather than the full detector readout.
- Online b tagging has been used to allow jet energy thresholds to be reduced at the trigger level. This allows sensitivity to resonance masses as small as 325 GeV [279].
- Resonances with masses as small as 10 GeV can be probed by requiring significant ISR, either in the form of jets [280–282] or photons [283]. In this topology, acceptable trigger rates are achieved by placing selection criteria on variables that are not

strongly correlated with the resonance mass, e.g., the ISR object momenta. The dijet system itself is significantly boosted and hence is reconstructed as a single large-radius jet (AK8 or CA15) with a two-pronged substructure. Several of these searches are described below.

Additionally, several DS models predict the existence of a dark photon A′, which can decay into pairs of SM leptons. Searches for low-mass resonances decaying into a pair of muons are described in Ref. [178] and Sections 6.2.1.4 and 6.2.1.5. Related to these, a search where the dark photon is long lived is also presented in Section 6.3.1.3.

The sensitivities of the searches described in this section to a range of simplified DM models are shown in Section 7.1.

6.2.1.1 Search for low-mass vector resonances decaying into quark-antiquark pairs

The most recent search for low-mass, boosted dijet resonances, which uses data from 2016 and 2017 corresponding to $\mathcal{L}_{\text{int}} = 77\,\text{fb}^{-1}$, is described in Ref. [281]. The analysis searches for new, spin-1 Z′ bosons decaying into quark-antiquark pairs, targeting a mass range of $50 < m_{Z'} < 450\,\text{GeV}$. The Z′ bosons are assumed to couple equally to all flavors of quarks, with a universal coupling constant g_{q}. The trigger selects AK8 jets with $p_{\text{T}} > 380\,(400)\,\text{GeV}$ in 2016 (2017) and a trimmed mass greater than 30 GeV; the trigger has good efficiency for Z′ boson masses greater than 50 GeV, which sets the lower bound on the search range. The analysis uses offline AK8 and CA15 jets, depending on the signal mass considered. The p_{T} requirements for offline AK8 jets are $p_{\text{T}} > 500\,(525)\,\text{GeV}$ in 2016 (2017) data and $p_{\text{T}} > 575\,\text{GeV}$ for CA15 jets. Jet substructure techniques are used to distinguish the signal from the backgrounds, which include QCD multijets, $\text{t}\bar{\text{t}}$, and W/Z+jets. The signal resonance is identified using the soft-drop mass variable m_{SD} [173], which removes soft and wide-angle radiation from the jet. The soft drop algorithm reduces the mass of jets from QCD, where the mass arises in part from soft gluon radiation, while preserving the mass of two-pronged signal jets. Second, the variable N_2^1, defined using ratios of ECFs [265], is used to reject QCD events; two-pronged jets tend to have a lower value of N_2^1 than QCD jets.

The QCD multijet background is estimated from data, using a "fail" CR consisting of events failing a requirement on N_2^1. In simulation, the "designed decorrelated tagger" method ensures that the m_{SD} shape in the CR is the same as the one in the SR by construction. The m_{SD} distribution is shown for a representative category in Fig. 31.

6.2.1.2 Search for low-mass quark-antiquark resonances produced in association with a photon

Another strategy to extend dijet searches to small Z′ boson masses is to focus on events in which the resonance is produced in association with a high-momentum ISR photon. The analyses described previously, Refs. [281] and [282], probe resonance masses down to about 50 GeV; this bound arises from the offline lower p_{T} jet threshold of 500 GeV, which causes the lowest-mass resonances to be extremely collimated, as well as directly from HLT selections on the jet mass. Lower masses can be probed by triggering on photons. Specifically, in 2016, the CMS trigger system recorded events containing photons with $p_{\text{T}} > 175\,\text{GeV}$. A search for dijet resonances with masses from 10 to 125 GeV and produced in association with an ISR photon is described in Ref. [283], using data collected in 2016 corresponding to $\mathcal{L}_{\text{int}} = 36\,\text{fb}^{-1}$.

The offline analysis of this dijet resonance search uses events containing a photon with $p_{\text{T}} > 200\,\text{GeV}$. Events with additional photons with $p_{\text{T}} > 14\,\text{GeV}$ or leptons with $p_{\text{T}} > 10\,\text{GeV}$ are

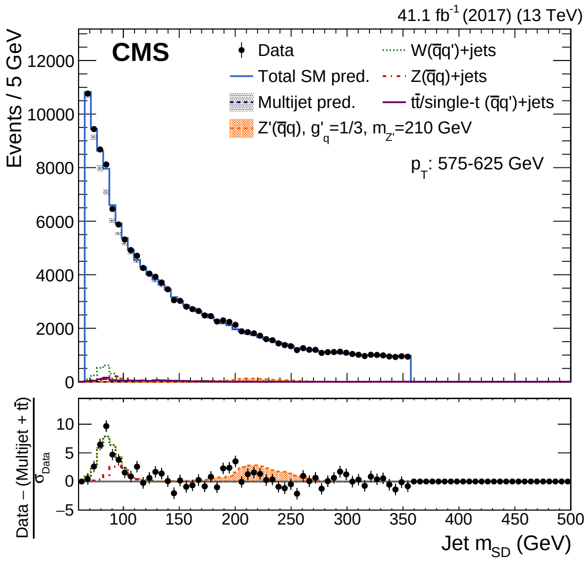

Figure 31: Jet m_{SD} distribution in data for CA15 jets for a p_{T} range of the fit from 575 to 625 GeV, in the search for low-mass vector resonances decaying into quark-antiquark pairs. Data are shown as black points. The QCD multijet background prediction, including uncertainties, is shown by the shaded bands. Smaller contributions from the W and Z bosons, and top quark background processes are shown as well. A hypothetical Z′ boson signal with a mass of 210 GeV is also indicated. In the bottom panel, the ratio of the data to its statistical uncertainty, after subtracting the nonresonant backgrounds, is shown. Figure taken from Ref. [281].

discarded to avoid overlap with other searches and to reduce backgrounds from EW sources. The analysis strategy is otherwise similar to Ref. [281], described above. The Z′ boson is reconstructed as a single AK8 jet and produces a local excess in the m_{SD} spectrum. The main background, coming from photons produced in association with jets from SM processes, is determined using a variation of the ABCD method with additional correction factors to account for the statistical dependencies of the variables. The m_{SD} distribution for the SR is shown in Fig. 32.

6.2.1.3 Search for low-mass resonances decaying into bottom quark-antiquark pairs

An analysis searching for new spin-0 resonances decaying into bottom quark-antiquark pairs, with resonance masses between 50 and 350 GeV is described in Ref. [282].

The analysis follows the general strategy of Ref. [281], a search for low-mass, boosted dijet resonances, and adapts it for new scalar resonances decaying into b$\bar{\mathrm{b}}$ pairs, using a data sample corresponding to $\mathcal{L}_{\mathrm{int}} = 36\,\mathrm{fb}^{-1}$, taken during 2016. Resonances are produced with high p_{T} because of significant ISR, ensuring events pass stringent trigger restrictions set by bandwidth limitations. In such events, the decay products of the resonance are reconstructed as a single large-radius jet with jet substructure consistent with originating from two b quarks. Both AK8 and CA15 jets are considered as candidates, with p_{T} thresholds of 450 and 500 GeV, respectively. The AK8 algorithm provides better sensitivity at signal masses less than 175 GeV, while the CA15 algorithm provides better sensitivity at higher masses. Jet substructure techniques and dedicated b-tagging algorithms are used to distinguish the signal from the QCD background. The signal is identified as a narrow resonance in the m_{SD} spectrum. The main algorithm for distinguishing signal jets from the QCD background, called the “double-b tagger,”

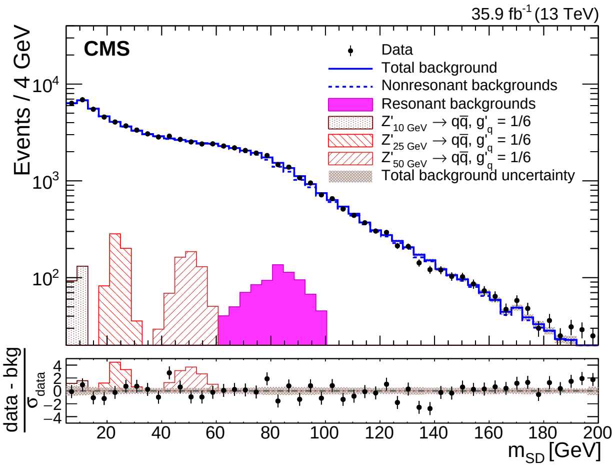

Figure 32: The soft drop jet mass distribution of the SR in the search for low-mass quark-antiquark resonances produced in association with a photon, after the main background estimation fit is performed. The nonresonant background is indicated by a dashed line, while the total background composed of the sum of this nonresonant background and the resonant backgrounds is shown by the solid line. Representative signals are plotted for comparison. The bottom panel shows the difference between the data and the final background estimate, divided by the statistical uncertainty of the data in each bin. The shaded region represents the total uncertainty in the background estimate in each bin. Figure taken from Ref. [283].

is a multivariate algorithm based on boosted decision trees, and uses kinematic information from tracks and SVs relative to two leading subjet axes. The N_2^1 [264, 265] variable is also used to further distinguish the two-pronged signal jets from QCD jets. The m_{SD} distribution is shown for a representative category in Fig. 33.

6.2.1.4 Search for a prompt dark photon resonance decaying into two muons including data scouting

Reference [178] presents a search for a narrow resonance, in the 11.5 to 200 GeV mass range, decaying into a pair of oppositely charged muons. For masses less than ≈40 GeV, a dedicated scouting trigger (as discussed in Section 4.1.2) with an exceptionally low muon p_{T} threshold was used. For higher masses, standard triggers were used. The data correspond to $\mathcal{L}_{\mathrm{int}} = 97$ and 137 fb^{-1} for the scouting and conventional triggering strategies, respectively. The dimuon mass resolution depends strongly on the pseudorapidity of the muons. Therefore, events are divided into two categories. The barrel category consists of events in which both muons are in the barrel region, and the forward category contains events in which at least one of the two muons is not in the barrel region.

In the high-mass search performed with the standard triggers, events are required to have at least one well-reconstructed PV and two oppositely charged muons. The muons are required to be isolated and to pass selection requirements based on the quality of their reconstructed tracks. In the search performed using the scouting triggers, events are required to contain two muons of opposite charge that are consistent with originating from the same vertex, with similar requirements on muon isolation and track quality as in the search using standard triggers. The dimuon invariant mass distribution is shown for a representative category in Fig. 34.

6.2.1.5 Search for prompt dimuon resonances with data scouting

An analysis [177] similar to the one described in Section 6.2.1.4 is performed to search for

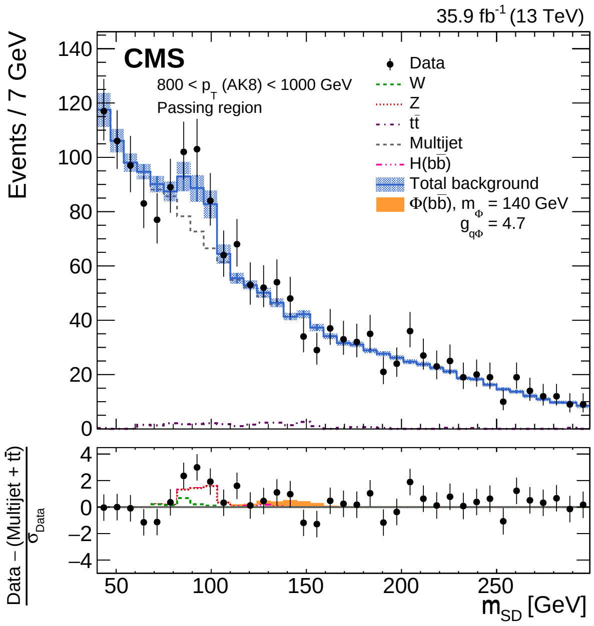

Figure 33: The observed and fitted background m_{SD} distributions in the $800 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$ category for the AK8 selection in the passing regions, in the search for low-mass resonances decaying into bottom quark-antiquark pairs. The fit is performed under the background-only hypothesis. A hypothetical signal at a mass of 140 GeV is also indicated. The shaded blue band shows the systematic uncertainty in the total background prediction. The bottom panel shows the difference between the data and the nonresonant background prediction, divided by the statistical uncertainty in the data. Figure taken from Ref. [282].

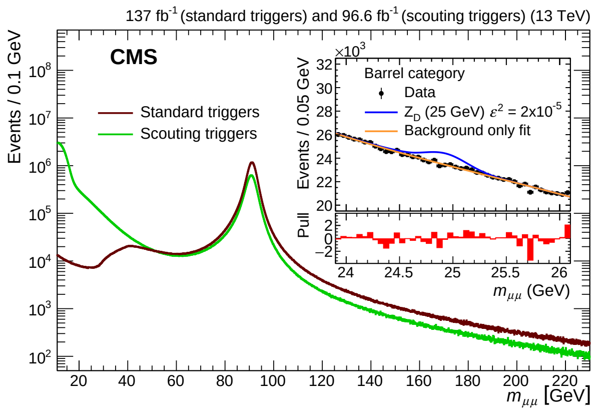

Figure 34: The dimuon invariant mass distributions of events selected with the standard muon triggers (brown, darker), and the scouting dimuon triggers (green, lighter), in the search for a prompt dark photon resonance decaying into two muons. Events are required to pass all the selection requirements. The inset shows the data (black points), the signal model (blue line), and the background-only fit (orange line), and it is restricted to events in the barrel category in the mass range 23.9–26.1 GeV. A function describing the background is fit to these data. The bottom panel of the inset shows the bin-by-bin difference between the number of events in data and the prediction from the background fit, divided by the statistical uncertainty. Figure taken from Ref. [178].

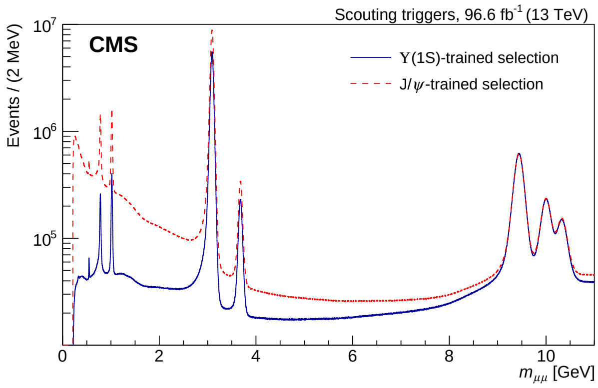

Figure 35: The dimuon invariant mass distribution obtained with the muon scouting data collected during 2017–2018 with two sets of selections: the $\Upsilon(1S)$-trained muon MVA identification (blue solid line), and the J/ψ-trained muon MVA identification (red dashed line). Figure taken from Ref. [177].

dimuon resonances with masses below the $\Upsilon(1S)$ resonance in the range of 1.1–2.6 GeV and 4.2–7.9 GeV using data collected by the dimuon scouting trigger during 2017–2018 with $\mathcal{L}_{\text{int}} = 97\,\text{fb}^{-1}$.

The event candidate is required to have at least one PV reconstructed by the HLT system and to contain a pair of oppositely charged muons originating from this vertex. To identify good-quality muon candidates, two multi-variate analysis (MVA) discriminants are used depending on the reconstructed dimuon mass, optimized for the signal kinematic properties in each mass range. The MVA identification utilizes information on the quality of the muon tracks, the relative isolation of the muon, and the vertex associated with the muons. Different vertex displacement criteria with respect to the beam spot are imposed in different mass ranges to account for the increased uncertainty in the PV position from the larger boost of the dimuon system and hence the more collinear tracks for smaller dimuon masses. The dimuon invariant mass distribution with both selections is shown in Fig. 35.

6.2.2 High-mass resonance searches

While resonances decaying into leptons have been excluded over a wide mass range and down to small couplings, resonances decaying into quarks are more challenging to detect because of the multijet background at hadronic colliders. Searches for resonances decaying into a quark pair have been performed mainly at high masses (e.g., $m > 1000$ GeV) in the dijet final state, while the low-mass range (e.g., $m < 200$ GeV) has been covered by the search for boosted resonances reconstructed as a single large-radius jet.

Three resonance searches are described in this section. We discuss a search for dijet resonances in events with three jets (which targets more mid-range masses), a search for high-mass dijet resonances, and a search for high-mass dilepton resonances.

The sensitivities of the searches in this section to a range of simplified DM models are shown in Section 7.1.

6.2.2.1 Search for dijet resonances using events with three jets

The search presented in Ref. [179] combines the data scouting technique with the requirement of an additional jet with high p_{T} to enhance signal sensitivity in the low-mass region. The analysis is performed on part of the data collected in 2016 (corresponding to $\mathcal{L}_{\text{int}} = 18.3\,\text{fb}^{-1}$)

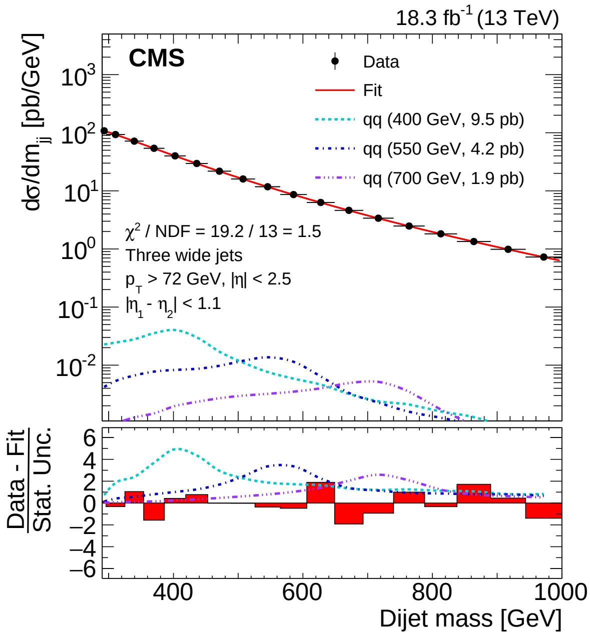

Figure 36: Dijet mass spectrum (points) compared to a fitted parameterization of the background (solid curve) in the search for dijet resonances using events with three jets, where the fit is performed in the range $290 < m_{jj} < 1000\,\text{GeV}$. The horizontal bars show the widths of each bin in dijet mass. The dashed lines represent the dijet mass distribution from 400, 550, and 700 GeV resonance signals expected to be excluded at 95% CL by this analysis. The lower panel shows the difference between the data and the fitted parametrization, divided by the statistical uncertainty of the data. Figure taken from Ref. [179].

when the trigger threshold was particularly low ($H_T > 240\,\text{GeV}$) in an attempt to extend the search to the lowest mass possible.

This analysis uses wide jets to recover the energy from final-state radiation, improving the dijet mass resolution. A selection on the η separation is used to suppress and reduce the QCD multijet background, which is dominated by t-channel production of jets. A bump search is then performed on the dijet mass spectrum, which is shown in Fig. 36.

6.2.2.2 Search for high-mass dijet resonances

There are several models [23, 24, 42, 43] in which DM mediators arise from an interaction between quarks and DM. The natural width of such mediators, which will appear as dijet resonances, increases with the coupling and may vary from narrow to broad, as defined in comparison to the experimental resolution. In Ref. [277], we describe a largely model-independent search for narrow or broad s-channel dijet resonances with masses greater than 1.8 TeV, shown in Fig. 37. We use data corresponding to $\mathcal{L}_{\text{int}} = 137\,\text{fb}^{-1}$ collected in Run 2.

Each of the two leading jets is formed into a “wide jet” using an algorithm introduced for previous CMS dijet searches in Ref. [284]. The SR is defined by vetoing events with a large η separation between the jets, which maximizes the search sensitivity for isotropic decays of dijet resonances in the presence of QCD dijet background.

The main background from QCD multijet production is predicted by fitting the m_{jj} distribution with an empirical functional form. For $m_{jj} > 2.4\,\text{TeV}$, a new background estimation method is

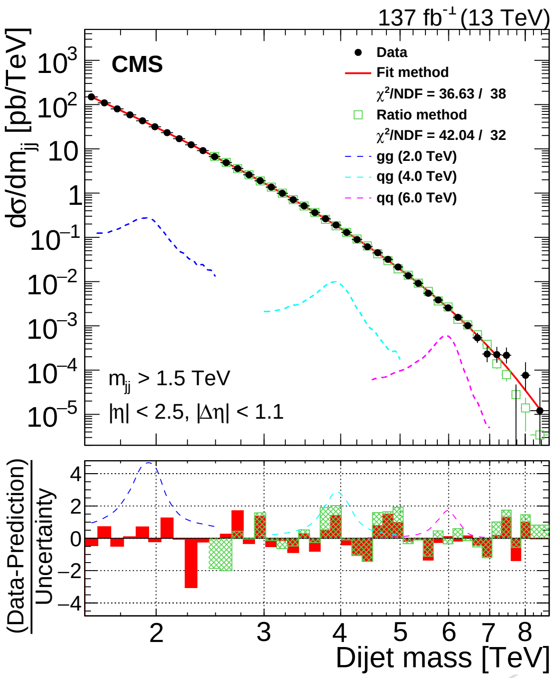

Figure 37: Dijet mass spectrum in the SR (points) compared to a fitted parameterization of the background (solid line) and the one obtained from the CR (green squares), in the search for high-mass dijet resonances. The lower panel shows the difference between the data and the fitted parametrization (red, solid), and the data and the prediction obtained from the CR (green, hatched), divided by the statistical uncertainty in the data, which for the ratio method includes the statistical uncertainty in the data in the CR. Examples of predicted signals from narrow gluon-gluon, quark-gluon, and quark-quark resonances are shown (dashed colored lines) with cross sections equal to the observed upper limits at 95% CL. Figure taken from Ref. [277].

introduced, which predicts the background from a CR where the pseudorapidity separation of the two jets, $|\Delta\eta|$, is large. This new background estimation method yields smaller systematic uncertainties.

Mediators with intrinsic widths larger than 50% have also been probed in CMS dijet events in a dedicated analysis of the dijet angular distributions [285] using a data set corresponding to $\mathcal{L}_{\text{int}} = 36\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$. While constraints on g_{q} from the dijet angular analysis are not competitive with the dijet resonance search, the dijet angular analysis allows to extend the excluded range of widths from 50 to 100% for mediator masses $<4.6\,\text{TeV}$.

6.2.2.3 Search for new physics in high-mass dilepton final state

Various theoretical models have been proposed in which DM particles interact with those of the SM via high-mass, weakly coupled mediator particles [286]. The decay of these mediator particles into SM particles could be observed through dilepton final states. A search for BSM physics using electron or muon pairs with high invariant mass [65] is sensitive to such mediator particles. Standard reconstruction techniques are used for high-p_{T} electrons and muons in this search; however, dedicated identification selection criteria are employed to ensure that high efficiency is maintained for both electrons [287] and muons [288]. The pp collision data at $\sqrt{s} = 13\,\text{TeV}$ collected in 2016–2018 are used in the search, corresponding to $\mathcal{L}_{\text{int}}$ up to $140\,\text{fb}^{-1}$.

The SM background processes are modeled with simulation (except for leptons produced inside jets or jets misidentified as leptons, which are estimated from CRs in data) and are nor-

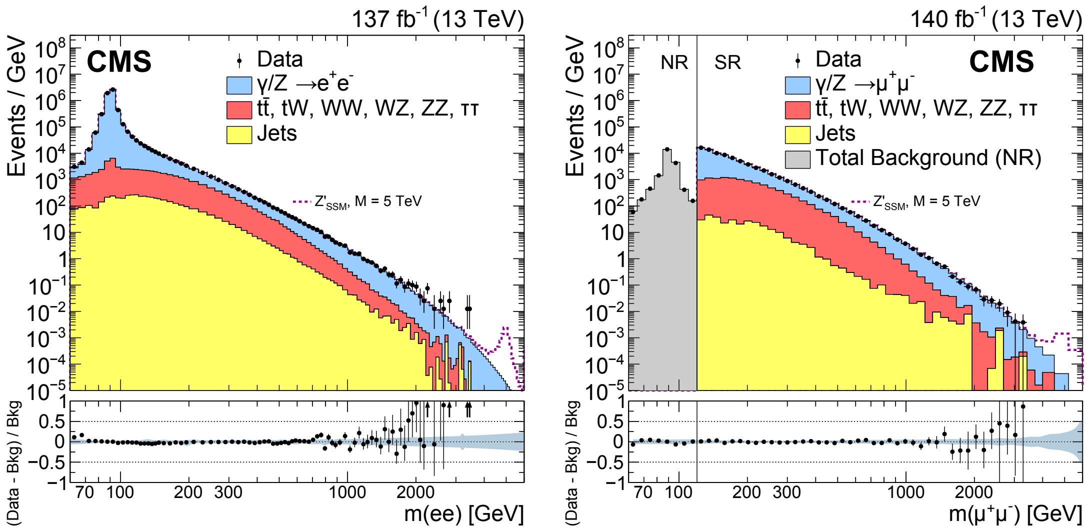

Figure 38: The invariant mass distribution of pairs of (left) electrons and (right) muons observed in data (black dots with statistical uncertainties) and expected from the SM processes (stacked histograms), in the high-mass dilepton search. For the dimuon channel, a prescaled trigger with a p_{T} threshold of 27 GeV was used to collect events in the normalization region (NR) with dimuon mass less than 120 GeV. The corresponding offline threshold is 30 GeV. Events in the SR corresponding to masses greater than 120 GeV are collected using an unprescaled single-muon trigger. The bin width gradually increases with mass. The ratios of the data yields after background subtraction to the expected background yields are shown in the lower plots. The blue shaded band represents the combined statistical and systematic uncertainties in the background. Signal contributions expected from simulated resonances are shown. Figures adapted from Ref. [65].

malized to the observed data yields in a mass window of 60–120 GeV around the Z boson peak, separately for the dielectron and dimuon channels. The search for resonant signatures is performed in a mass window around the assumed resonance mass, whose size depends on the assumed intrinsic decay width of the resonance and the mass-dependent detector resolution. A range of masses and widths is scanned to provide results covering a wide selection of signal models. Unbinned maximum likelihood fits are performed inside the mass windows, allowing the background normalization to be determined from the data. Through setting upper limits on the ratio of the product of the production cross section and the branching fraction of a new narrow dilepton resonance to that of the SM Z boson, many experimental and theoretical uncertainties common to both measurements cancel out or are reduced, leaving only uncertainties in the ratio that vary with the dilepton mass to be considered. The dielectron and dimuon invariant mass distributions are shown in Fig. 38.

6.2.3 Other signatures

In this Section, we describe searches for visible and prompt signatures that do not fall into the low- and high-mass resonance categories described in Sections 6.2.1 and 6.2.2, respectively. The searches described here include a search for fractionally charged particles, a search for SUEPs, a search for stealth or RPV top squarks, a search for ALPs in ultraperipheral PbPb collisions, and a search using the missing-mass technique in CMS and CMS-TOTEM events.

6.2.3.1 Search for fractionally charged particles (FCPs)

In the search for LLPs carrying a fraction of the electron charge, i.e., $Q_{\mathrm{FCP}} = \varepsilon e$, where ε is lower than 1, described in Ref. [215], we consider a signal generated via DY production using a data set corresponding to $\mathcal{L}_{\text{int}} = 138\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$. The experimental signature of an FCP is close to that of a muon, but with a larger mass and a lower charge. Therefore, we require events to contain exactly one or two high-p_{T} isolated muons.

The analysis strategy relies on the measurement of the ionization loss per unit length ($\text{d}E/\text{d}x$) associated with the hits in the modules of the CMS silicon tracker (described in Section 4.4.6). The energy loss process in silicon is stochastic; the most probable hit $\text{d}E/\text{d}x$ value for a muon is around 3 MeV/cm. A low-charge particle is expected to deposit lower amounts of energy, systematically across all hits. The scaling goes with the square of the FCP charge, as described by the Bethe-Bloch function. To discriminate signal from background, we build a binomial distribution by asking the following question for each hit on a track: is the $\text{d}E/\text{d}x$ less than a threshold value? The threshold is adapted layer-by-layer to take into account experimental effects such as radiation damage to a module. The variable $N_{\text{hits}}^{\text{low}\,\text{d}E/\text{d}x}$ is the total number of hits on a track that pass the requirement, shown in Fig. 39 for 2018 data. It accumulates at small values for charge e particles such as muons and extends to larger values as the charge decreases.

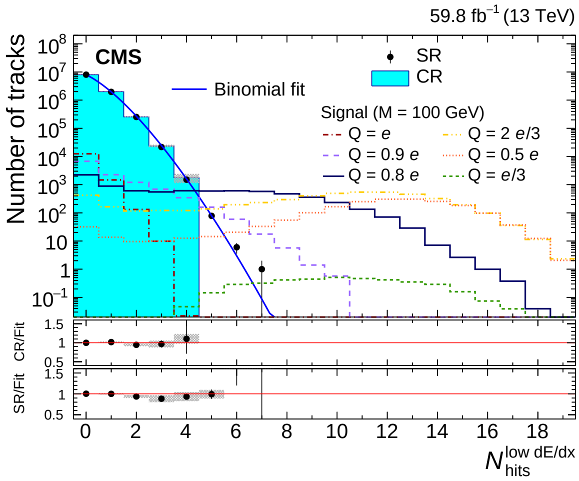

Figure 39: Distribution of $N_{\text{hits}}^{\text{low}\,\text{d}E/\text{d}x}$ in the search and CRs for the early 2018 data set, in the search for fractionally charged particles. The middle (lower) panels show the ratio of the number of tracks observed in the CR (SR) and the fit function. Figure taken from Ref. [215].

We fit the $N_{\text{hits}}^{\text{low}\,\text{d}E/\text{d}x}$ distribution in the CR to estimate our background and compare it to the observation in the SR.

6.2.3.2 Search for soft unclustered energy patterns

A search for SUEPs arising from the decay of a heavy scalar mediator is reported in Ref. [289]. Motivated by HV models with a dark-QCD sector and large ’t Hooft coupling, the signature of a SUEP is a high multiplicity of spherically distributed low-momentum charged particles in the final state. The data, which correspond to $\mathcal{L}_{\text{int}} = 138\,\text{fb}^{-1}$, were collected in 2016–2018 using traditional hadronic triggers, which often select events with high-p_{T} ISR jets. As a result, boosted topologies are favored in this analysis. The charged particle tracks in the event are clustered into wide jets and of the two leading jets, the jet with the larger number of constituent tracks is chosen to be the SUEP candidate. An example signal event is shown in Fig. 40.

The primary background in this search comes from QCD multijet events with a large number of tracks. This search utilizes a novel approach to predict the background by estimating the

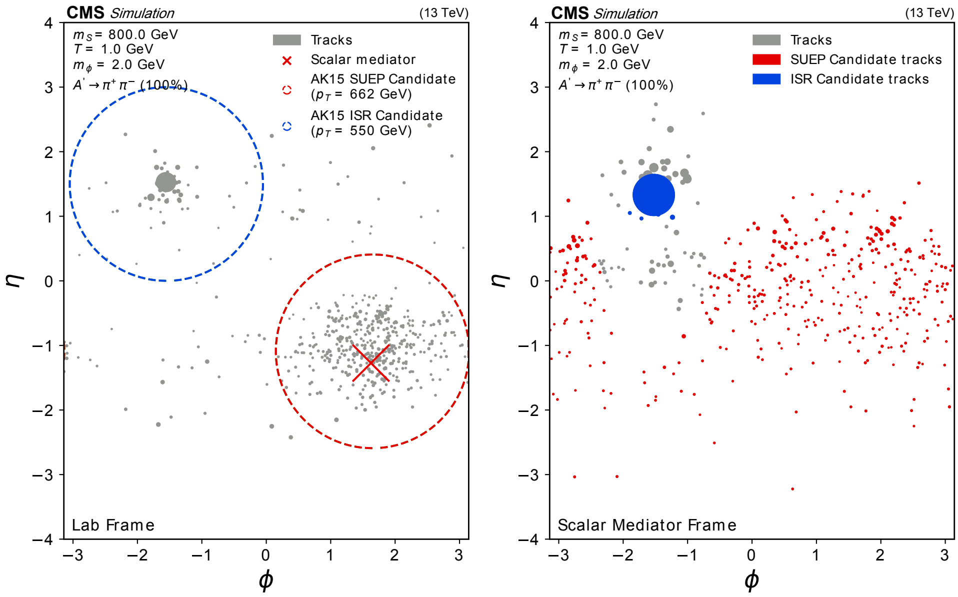

Figure 40: An example SUEP event from a representative model with a scalar mediator of mass 800 GeV shown in the lab frame (left) and the generator-level S mediator frame (right). The jets are clustered from charged particle tracks associated with the primary vertex using the anti-k_{T} algorithm with a distance parameter of 1.5. The size of each dot is scaled based on the p_{T} of the corresponding track.

contribution from traditional processes directly from data using an "extended" version of the ABCD method described in Section 4.7.3 [244].

The sensitivity of the search is shown in Section 7.2.4.3.

6.2.3.3 Search for stealth top squarks

As detailed in Section 2.2.2.3, models of stealth SUSY result in final states where the typical $p_{\mathrm{T}}^{\mathrm{miss}}$ of SUSY searches is replaced with additional visible objects, which are jets in the models considered here. The search described in Ref. [290] targets pair production of top squarks with decays via the stealth sector through the vector portal (labeled 'SYY'), resulting in a final state with two top quarks and six gluons.

The search selects events with exactly one electron or muon, at least seven jets, and at least one b-tagged jet, using data corresponding to $\mathcal{L}_{\mathrm{int}} = 137\,\mathrm{fb}^{-1}$, collected in 2016–2018. No requirement is placed on $p_{\mathrm{T}}^{\mathrm{miss}}$. The signal is distinguished from the background by means of a neural network that uses the jet kinematics as well as overall event shape variables as input features. Crucially, the network was trained to be independent of the jet multiplicity by using the gradient reversal technique. This enabled the background estimation to be done via a simultaneous fit to the jet multiplicity distribution in four bins of the neural network score. The jet multiplicity is modeled with a recursive fit function based on QCD jet scaling patterns. The distribution of the neural network score for 2017–2018 data and simulation is shown in Fig. 41. The sensitivity of the search is shown in Section 7.2.2.2.

6.2.3.4 Search for axion-like particles in ultraperipheral PbPb collisions

The CMS Collaboration has searched for ALPs (Section 2.1.2.5) that couple to photons in PbPb

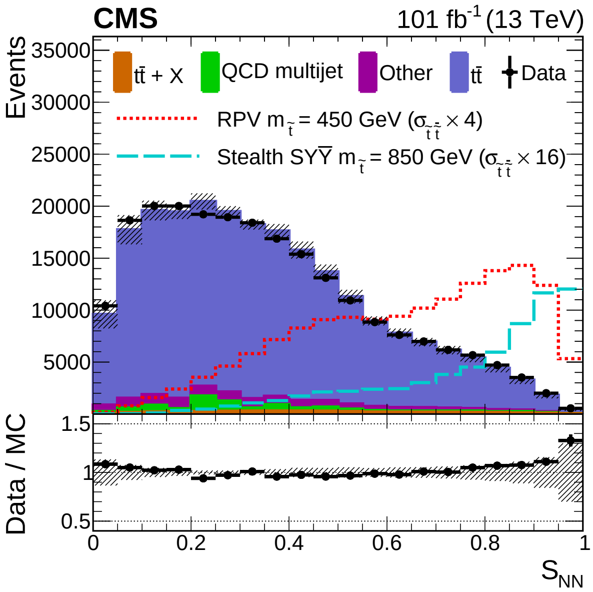

Figure 41: The neural network score (S_{NN}) distribution for 2017–2018 shows the data in the SR (black points); simulated background normalized to the number of data events (filled histograms); RPV signal model with a top squark mass of 450 GeV (red short dashed line); and stealth SYY signal model with a top squark mass of 850 GeV (cyan long dashed line), in the search for stealth top squarks. The band on the total background histogram denotes the dominant systematic uncertainties, as well as the statistical uncertainty for the non-$\mathrm{t\bar{t}}$ components. The lower panel shows the ratio of the number of data events to the number of normalized simulated events with the band representing the difference between the nominal ratio and the ratio obtained when varying the total background by its uncertainty. Figure taken from Ref. [290].

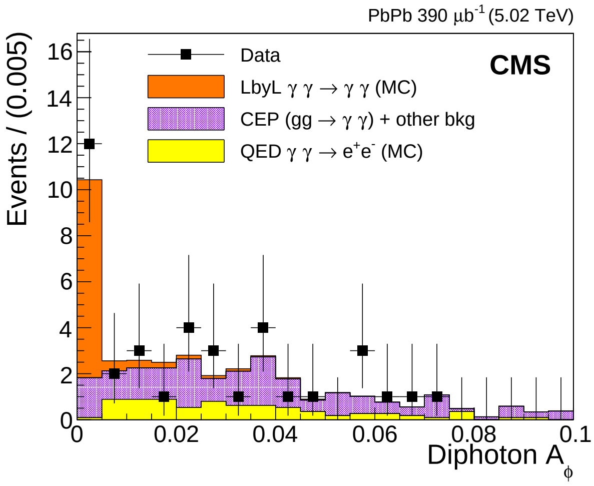

Figure 42: Diphoton acoplanarity distribution in the search for axion-like particles in ultra-peripheral PbPb collisions, for exclusive events measured in the data after selection criteria (squares), compared to the expected light-by-light scattering signal (orange histogram), quantum electrodynamics e^+e^- (yellow histogram), and the CEP+other (purple histogram) backgrounds. Signal and quantum electrodynamics e^+e^- MC samples are scaled according to their theoretical cross sections and integrated luminosity. The error bars around the data points indicate statistical uncertainties. The horizontal bars around the data symbols indicate the bin size. Figure taken from Ref. [219].

UPCs [219]. UPCs are defined as collisions in which the impact parameter is larger than twice the nucleus radius, where passing heavy ions do not break up and are so close that their electromagnetic fields are intense enough to interact as quasi-real photon beams. The PbPb collisions provide an enhancement of a factor given by the atomic number to the power of four for photon-photon scattering processes as compared to pp collisions, since the photon flux scales as the atomic number squared of the emitting ion. The production of a resonant ALP ($\gamma\gamma \to a \to \gamma\gamma$) is expected to modify the rate of the light-by-light scattering process ($\gamma\gamma \to \gamma\gamma$) that shares the same final state.

Potential backgrounds to ALPs production include the major nonresonant light-by-light process, the quantum electrodynamics $\gamma\gamma \to e^+e^-$ process where both electrons are misidentified as photons, and the central exclusive production (CEP) $gg \to \gamma\gamma$ where the exclusive diphotons are produced via strong interactions. Events with exactly two photons with $E_T > 2\,\text{GeV}$ and $|\eta| < 2.4$, no extra charged particles, and no calorimeter activity are selected. The nonexclusive diphoton background is eliminated by requiring events to have diphoton acoplanarity $A_\phi < 0.01$ and diphoton transverse momentum $p_T^{\gamma\gamma} < 1\,\text{GeV}$. The diphoton acoplanarity distribution, before the criterion on this variable is applied, is shown in Fig. 42. The measured diphoton invariant mass distribution is used to search for possible narrow diphoton resonances. The sensitivity of the search is discussed in Section 7.1.2.5.

6.2.3.5 Search for new physics in central exclusive production using the missing-mass technique with CMS and CMS-TOTEM

Studies of CEP processes in high-energy pp collisions provide a unique method to access a class of physics processes, such as new physics via anomalous production of fermions, V bosons (where V is a γ, W, or Z boson), high-p_T jet production, and possibly the production of new resonances or pair production of new particles. The addition of new detectors further extends the coverage and enhances the sensitivity of the LHC experiments thus offering a new opportunity to explore processes and final states previously not covered. The CMS-TOTEM PPS [216] allows the surviving scattered protons during standard running conditions in regular "high-luminosity" fills to be measured [217] (Section 4.5).

A generic search for a hypothetical massive particle X produced in association with one or more SM particles in CEP processes is performed [291]. In the interaction, the two colliding protons survive after exchanging two colorless particles and can be recorded in the PPS. The detection and precise measurement of both forward protons allows a full kinematic reconstruction of the event, including the four-momentum of X measured from the balance between the tagged SM particle(s) and the forward protons. This technique—the "missing-mass" technique—allows for searches for BSM particles without assumptions about their decay properties, except that the decay width can be considered narrow enough to produce a resonant mass peak, thus providing a new tool for generic BSM searches. A search for a massive particle produced in association with a Z boson or a photon in the final state is considered, using data samples corresponding to $\mathcal{L}_{\text{int}} = 37$ and $2.3\,\text{fb}^{-1}$, respectively.

The excellent proton momentum reconstruction of PPS allows us to search for missing-mass signatures at high invariant masses with unprecedented resolution. In this high-mass range, EW processes are generally enhanced relative to QCD-induced processes. The main goal is the search for a $\gamma\gamma$-induced exclusive production process in which an unspecified weakly interacting BSM particle with a narrow decay width is produced. No assumption is made on its decay properties. Leptonically decaying Z bosons or an isolated photon are selected in the central detector, and the missing mass is constructed from the kinematics of the reconstructed boson in the central detector and the final-state protons in PPS (Fig. 43). A hypothetical X resonance is searched for in the mass region between 0.6 and 1.6 TeV.

6.3 Searches for long-lived particles

As mentioned in Section 4.4, scenarios with LLPs can provide a DM candidate. Here we describe the signatures and searches for LLPs in CMS that provide sensitivity to the DS. We first describe searches for LLPs that decay into displaced leptons in Section 6.3.1, then searches for LLPs that decay hadronically in Section 6.3.2, and lastly searches for LLPs and $p_{\mathrm{T}}^{\text{miss}}$ in Section 6.3.3.

6.3.1 Displaced leptons

Displaced leptons provide a powerful handle to identify LLP decays while maintaining sensitivity to a wide range of models. Events with displaced leptons have a clean signature because of the reduced background contribution from SM processes. In this section, we describe several displaced-lepton analyses with distinct signatures. The reconstruction of displaced signatures with the tracker is described in detail in Section 4.4.1 and the reconstruction of displaced muons is described in detail in Section 4.4.4.

6.3.1.1 Search for displaced leptons in eμ, ee, and $\mu\mu$ final states

The analysis described in Ref. [243] is carried out on a pp collision data set corresponding to $\mathcal{L}_{\text{int}} = 115\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$. This analysis targets the displaced lepton signature by studying events with at least two leptons (any combination of electrons and muons) with transverse impact parameters between 0.01 and 10 cm. Requiring two such leptons with transverse momenta thresholds varying from 35 to 75 GeV, depending on lepton flavor and data-taking year, and relatively little nearby activity is sufficient to reject nearly all SM backgrounds without placing any requirements on the dilepton charge product or flavor combination, constraining

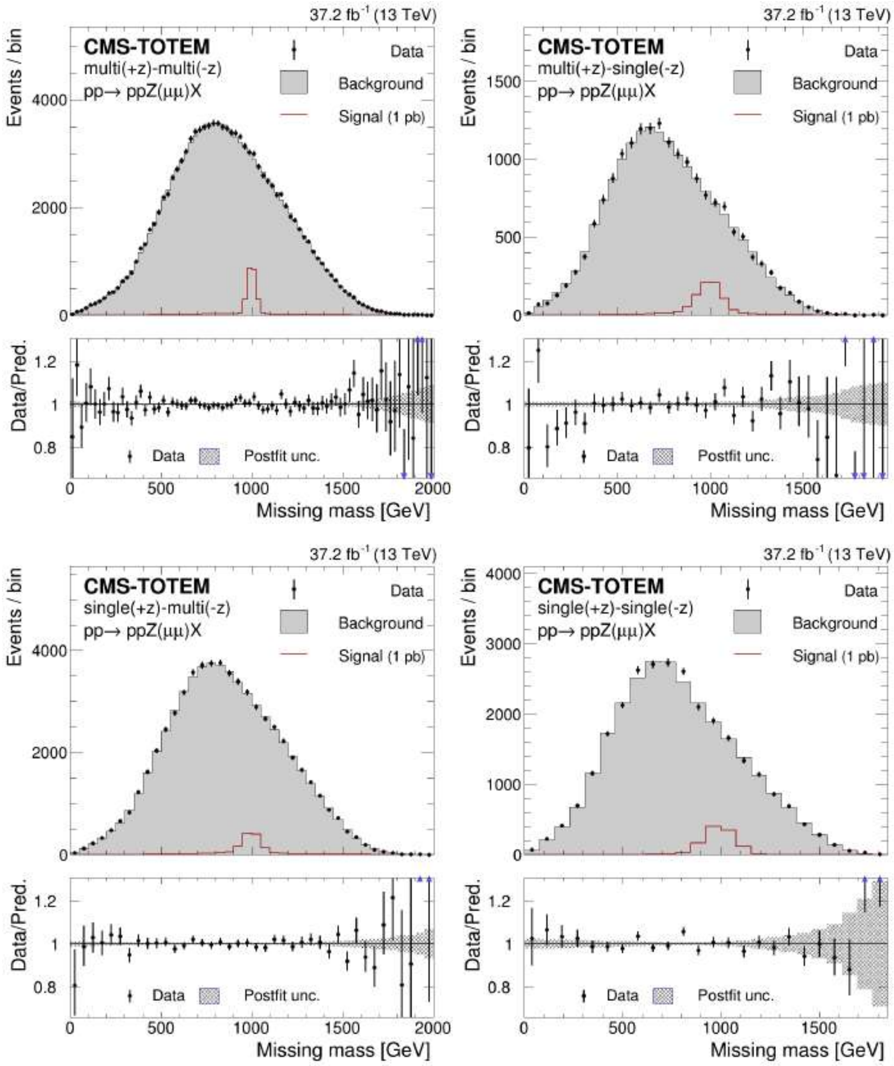

Figure 43: Missing-mass distributions in the $Z \to \mu\mu$ final state of the CMS and CMS-TOTEM search using the missing-mass technique. The distributions are shown for protons reconstructed with (from left to right) the multi-multi, multi-single, single-multi, and single-single methods, respectively. The background distributions are shown after the fit. The lower panels display the ratio between the data and the background model, with the arrows indicating values lying outside the displayed range. The expectations for a signal with $m_X = 1000\,\text{GeV}$ are superimposed and normalized to 1 pb. Figure taken from Ref. [291].

other event properties (such as hadronic activity or $p_{\mathrm{T}}^{\mathrm{miss}}$), or requiring that the leptons form a common vertex. The signature for this search is shown in Fig. 44. This approach allows the analysis to be sensitive to effectively any new physics process that involves at least one LLP whose decay includes at least two leptons or two LLPs whose decays each include at least one lepton. This is the only CMS Run 2 search for displaced leptons where the leptons are not required to come from a common SV.

This analysis uses dilepton triggers that do not require the leptons to originate from the collision point. The SM background is dominated by leptons with poorly measured displacement values, and care is taken to reject sources of genuine displaced leptons such as cosmic ray muons, displaced decays of SM mesons, and material interactions. The SM background estimate uses the ABCD method described in Section 4.7.3 within 15 orthogonal SRs that differ in lepton flavor, displacement, and momentum, an approach that maximizes the sensitivity to a range of LLP masses and lifetimes.

The sensitivity of this search to Higgs boson decays to LLPs is discussed in Section 7.2.4.4.

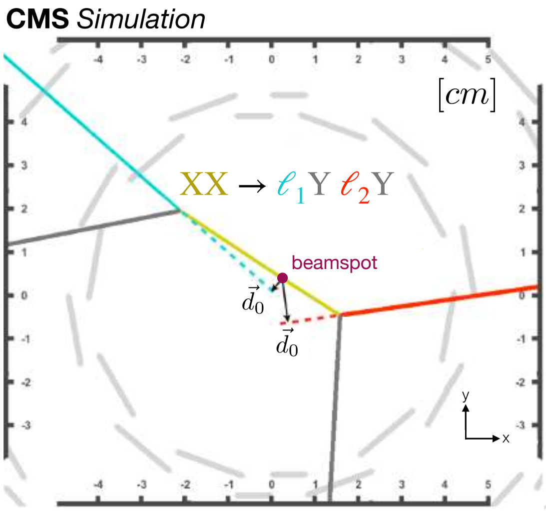

Figure 44: A diagram of a simulated signal event in the inclusive displaced-leptons search, from a transverse view of the interaction point, in the analysis presented in Ref. [243]. The black arrows indicate the lepton transverse impact parameter vectors.

6.3.1.2 Search for displaced muon pairs

Reference [132] presents an inclusive search for an exotic massive LLP decaying into a pair of oppositely charged muons ("dimuon") originating from a common SV. The SV can be spatially separated from the pp interaction point by a distance ranging from several hundred μm to several meters. The analysis uses muons produced within the silicon tracker, which can be reconstructed by both the tracker and the muon system, as well as muons produced in the outer tracker layers or beyond, which are reconstructed by only the muon system. The data sample corresponds to $\mathcal{L}_{\text{int}} = 98\,\text{fb}^{-1}$. The minimal set of requirements and loose event selection criteria used in the search allow us to be sensitive to a wide range of LLP models. Figure 45 shows the distribution of a key discriminating variable, namely, the minimum d_0 significance, for globally reconstructed dimuon pairs with 2018 data.

Reference [292] presents a continuation and extension of the search for displaced dimuons produced within and beyond the tracker described in Ref. [132]. The search is based on data collected during 2022 at $\sqrt{s} = 13.6\,\text{TeV}$, corresponding to $\mathcal{L}_{\text{int}} = 36.6\,\text{fb}^{-1}$ and recorded with an improved set of HLT and level-1 trigger algorithms [183], aimed at increasing the signal efficiency by lowering the p_{T} thresholds as much as possible without increasing the resulting trigger rate considerably. Overall, the addition of the new trigger algorithms improves the trigger efficiency for LLPs with a mass of a few tens of GeV and displacement $\gtrsim$0.1 cm by a factor of 2 to 4, depending on displacement and mass, as compared to Run 2.

The sensitivity of this search to HAHM scenarios is shown in Section 7.2.2.1 and the sensitivity to Higgs boson decays to LLPs is shown in Section 7.2.4.4.

6.3.1.3 Search for displaced dimuons in final states with 4μ+X

Reference [293] describes another analysis that uses displaced muons to search for evidence of DS particles. In this analysis, we search for the production of two LLPs per event, selecting

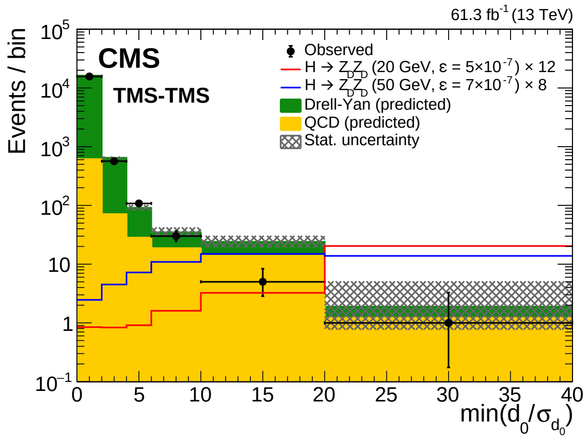

Figure 45: Comparison of the number of events observed in 2018 data with the expected number of background events, as a function of the smaller of the two d_0 significance values ($\min(d_0/\sigma_{d_0})$) for pairs of muons that are globally reconstructed in the tracker and muon system (TMS), in the search for displaced muon pairs. The black points with error bars show the number of observed events; the green and yellow components of the stacked histograms represent the estimated numbers of DY and QCD events, respectively. The last bin includes events in the overflow. The uncertainties in the total expected background (shaded area) are statistical only. Signal contributions expected from simulated decays of exotic Higgs bosons to dark Z bosons, with Z boson masses of 20 and 50 GeV are shown in red and blue, respectively. Their yields are set to the corresponding combined median expected exclusion limits at 95% CL, scaled up as indicated in the legend to improve visibility. Figure taken from Ref. [132].

pairs of displaced dimuons reconstructed in the tracker in a data sample with $\mathcal{L}_{\text{int}} = 36\,\text{fb}^{-1}$. Events that can mimic the signal come from pair-production of bottom quarks through QCD processes (QCD $\text{b}\bar{\text{b}}$), double J/ψ production, and EW processes. The 2D distribution of the invariant masses of the isolated dimuon systems is shown in Fig. 46.

In the case of the QCD $\text{b}\bar{\text{b}}$ background, CRs in data are used to estimate its contribution, while for the J/ψ and EW processes, such as $\text{ZZ} \to 4\mu$ and $\text{Z}^*/\gamma \to 2\mu$ (where a second Z boson is radiated and decays into a muon pair), the backgrounds are estimated with CRs in data and from simulation, respectively.

The sensitivity of this search to HAHM scenarios is shown in Section 7.2.2.1.

6.3.1.4 Search for displaced dimuon resonances with data scouting

Scouting triggers such as those described in Section 4.1.2 also provide opportunities for DS searches with displaced leptons. A search for narrow, long-lived dimuon resonances [294] is performed based on data collected during Run 2 in 2017 and 2018, using a dedicated dimuon scouting trigger stream. The selected data correspond to $\mathcal{L}_{\text{int}} = 101\,\text{fb}^{-1}$. The rate of scouting triggers is higher than that of the standard triggers allowing less stringent requirements on the muon p_{T}. This enables dimuon resonance searches across mass and lifetime ranges that are otherwise inaccessible; in particular, the search described here has sensitivity to masses in the 1–3 GeV range. The scouting trigger algorithms used in this search select events containing muons with $p_{\text{T}} > 3\,\text{GeV}$. The search targets narrow, low-mass, long-lived resonances decaying into a pair of oppositely charged muons, where the lifetime of the LLP is such that the transverse displacement (l_{xy}) of its decay vertex is within 11 cm of the PV. Muon tracks are used in pairs to form dimuon vertices, considering all possible pairings. These vertices are considered to be candidate SVs, and they may be displaced from the PV or not. The dimuon invariant

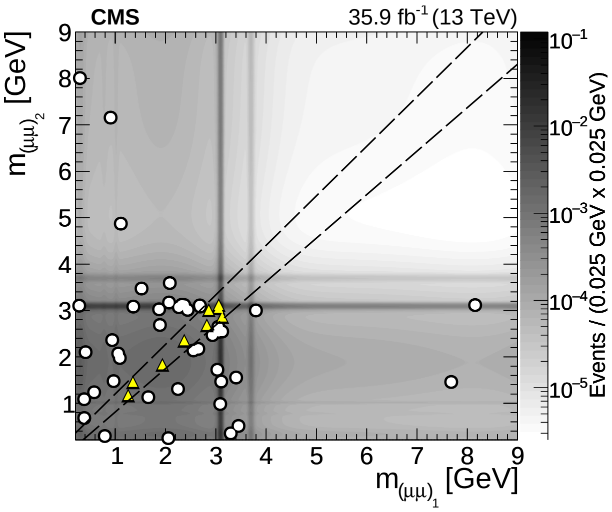

Figure 46: Distribution of the invariant masses $m_{(\mu\mu)1}$ vs. $m_{(\mu\mu)2}$ of the isolated dimuon systems, in the search for displaced dimuons in final states with 4μ+X. Triangles represent data events passing all the selection criteria and falling in the SR $m_{(\mu\mu)1} \approx m_{(\mu\mu)2}$ (outlined by dashed lines), and white bullets represent data events that pass all selection criteria but fall outside the SR. Figure taken from Ref. [293].

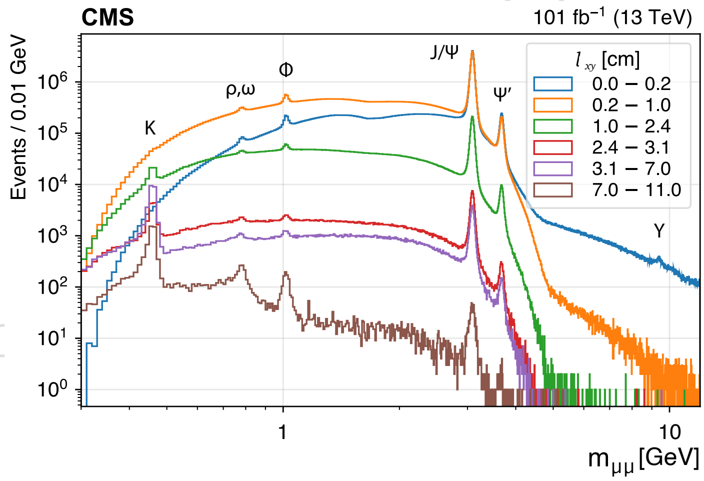

Figure 47: The dimuon invariant mass distribution from the search for displaced dimuon resonances with data scouting, shown in bins of l_{xy} as obtained from all selected dimuon events. Figure taken from Ref. [176].

mass distribution in bins of l_{xy} is shown in Fig. 47. The signal is expected to appear as a narrow peak on the dimuon mass continuum, with a resonance width smaller than the experimental mass resolution. Events are required to contain at least one pair of oppositely charged muons associated with a selected SV, and those that contain a single muon pair are then categorized according to transverse displacement and the p_T and isolation of the muon pair. In each category, we define mass windows sliding along the dimuon invariant mass spectrum, and we perform a search for a resonant peak in each mass window.

The sensitivity of the search to Higgs boson decays to LLPs is discussed in Section 7.2.4.4.

6.3.2 Hadronic LLP decays

Hadronic decays of LLPs can provide sensitivity to a large variety of DS models. Here we describe several CMS searches that utilize hadronic LLP decays. The decay positions of the LLPs targeted in these searches span a wide range, including decays in the tracker, calorimeters, and even in the muon system.

6.3.2.1 Search for LLPs decaying into displaced jets

In Ref. [187], we present a model-independent search for LLPs decaying into jets, with at least one LLP having a decay vertex within the tracker acceptance, which goes up to ≈550 mm in the plane transverse to the beam direction. The data sample corresponds to $\mathcal{L}_{\text{int}} = 132\,\text{fb}^{-1}$. Events were collected with dedicated displaced-jets triggers, which select jets with small numbers of prompt tracks or with displaced tracks. With these tracking requirements, the H_{T} trigger threshold has been lowered from 1000 to 430 GeV, which significantly increases the trigger efficiencies for a large variety of models with LLPs.

After the trigger selections, we look for all possible pairs of jets in a given event. For each jet pair (dijet), we attempt to reconstruct one DV using the displaced tracks associated with the two jets.

The vertex reconstruction is performed using the adaptive vertex fitter described in 4.4.1. The properties of the DV, such as the number of tracks and the transverse displacement significance, provide discrimination power to distinguish LLP signatures from SM backgrounds. The distribution of the vertex track multiplicity is shown in Fig. 48. The relations among the DV, displaced tracks, and the dijet are also examined to construct more discriminating variables. Using these variables, a multivariate classifier based on a GBDT is developed to further improve the signal-to-background discrimination. The use of displaced jet tagging is described in detail in Section 4.4.2.

The sensitivity of the search to Higgs boson decays to LLPs is presented in Section 7.2.4.4. The sensitivities to models containing heavy Z′ and heavy H_{D} bosons are described in Section 7.2.4.5.

6.3.2.2 Search for new physics with displaced vertices

This inclusive and largely model-independent search for pair-produced LLPs that decay hadronically focuses on LLPs with mean proper decay lengths less than 100 mm [184]. The reconstruction of DVs is detailed in Section 4.4.1. To perform the search, the LLP decay positions are reconstructed as DVs, which are formed from charged particle tracks using a custom vertex reconstruction algorithm. The search is performed using data corresponding to $\mathcal{L}_{\text{int}} = 140\,\text{fb}^{-1}$ from 2015–2018, collected at $\sqrt{s} = 13\,\text{TeV}$, and relies on events collected with H_{T} triggers that require large jet activity. After forming the DVs, a series of selection criteria are used to suppress backgrounds. For instance, to eliminate backgrounds originating from material interactions, the DVs are required to be located within the radius of the beam pipe. Several other criteria are additionally used in the search to distinguish signal from background, including requirements on the uncertainty in the beamspot-to-vertex distance, which is crucial for mitigating backgrounds from genuine b quark decay vertices, as well as a requirement that each signal-like vertex be formed from at least five charged particle tracks, to reduce combinatorial backgrounds. The primary search variable is the distance between two signal-like vertices in the x-y plane (d_{VV}), shown in Fig. 49, as the LLPs considered are often expected to be produced back-to-back and to each have large x-y displacement, while the separation between background vertices tends to be smaller.

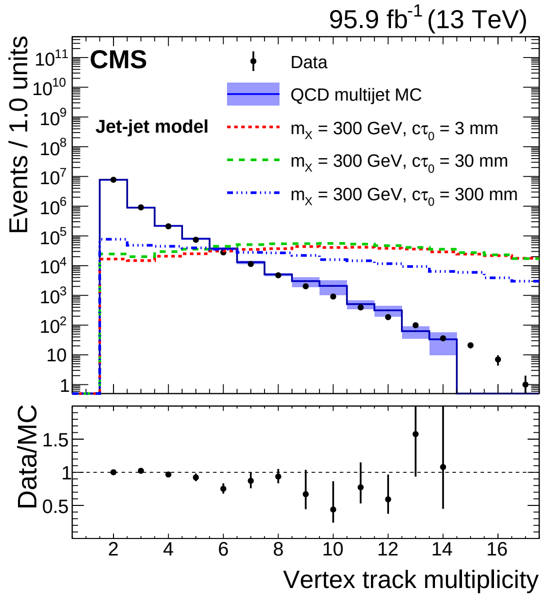

Figure 48: Distribution of the vertex track multiplicity, for data, simulated QCD multijet events, and simulated signal events, in the displaced-jets search. For a given event, if there is more than one SV candidate being reconstructed, the one with the largest vertex track multiplicity is chosen. If the track multiplicities are the same, the one with the smallest χ^2/ndof is chosen, where ndof is the number of degrees of freedom. The lower panel shows the ratios between the data and the simulated QCD multijet events. The blue shaded error bands and vertical bars represent the statistical uncertainties. Three benchmark signal distributions are shown (dashed lines). For visualization purposes, each signal process is given a cross section that yields 106 events produced in the analyzed data sample. Figure taken from Ref. [187].

The sensitivities of the search to models containing heavy Z' and heavy H_D bosons are shown in Section 7.2.4.5.

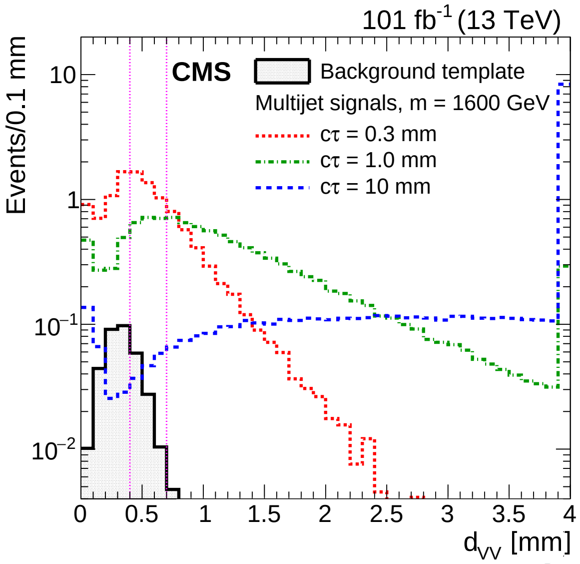

Figure 49: The distribution of distances between vertices in the x-y plane, d_{VV}, for the displaced-vertices search, for three simulated multijet signals each with a mass of 1600 GeV, with the background template distribution overlaid. The production cross section for each signal model is assumed to be the lower limit excluded by Ref. [295], corresponding to values of 0.8, 0.25, and 0.15 fb for the samples with $c\tau_0 = 0.3$, 1.0, and 10 mm, respectively. The last bin includes the overflow events. The two vertical pink dashed lines separate the regions used in the fit. Figure taken from Ref. [184].

6.3.2.3 Searches for emerging jets

Emerging jet phenomena may be observable at the LHC detectors when the DS is strongly coupled and the composite dark mesons have a finite lifetime comparable to the detector size, as described in the HV description in Section 2.2.4. The signature of an EJ differs from that of an SM jet in that the associated tracks will originate from many vertices, which can appear at various distances from the collision point depending on the dark meson lifetimes. The axis of each vertex within the jet points radially from the collision point. Dark quark production occurs via the decay of a complex scalar mediator Φ, which is charged under both SM QCD and dark QCD. The mediator is produced in pairs at the LHC primarily through gluon-gluon fusion, and it decays into a dark quark and SM quark: $\Phi\Phi^\dagger \rightarrow q_{dark}\overline{q}q'\overline{q}'_{dark}$.

The displacement features from tracks associated with a jet are used to tag the EJ signal. Because there are no dedicated triggers for this signature, an H_T-based trigger is used, as the signal includes multiple hard jets.

The first iteration of the search [296] uses a set of requirements on several jet- and track-based variables to tag the EJs. This search uses a data set corresponding to $\mathcal{L}_{int} = 16\,fb^{-1}$, which is approximately half of the 2016 data [296].

The second iteration of the search [297] employs both a model-agnostic EJ tagger, similar to the first search, and a more powerful, but model-dependent, graph neural network (GNN) EJ tagger. Distributions of the output score of the GNN are shown in Fig. 50. The Run 2 data set corresponding to $\mathcal{L}_{int} = 138\,fb^{-1}$ is analyzed.

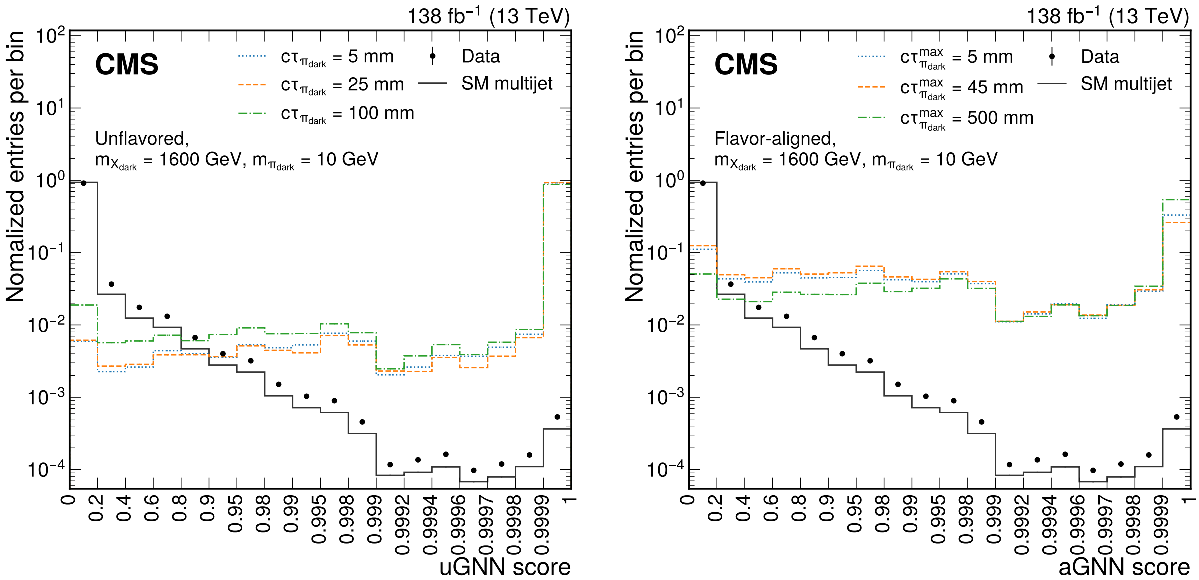

Figure 50: Distributions of the GNN output score for the data (points with error bars), SM multijet simulation (dark gray line), and signal simulation (colored lines), for the search for emerging jets. Separate GNNs are trained for the unflavored model (uGNN, left) and the flavor-aligned model (aGNN, right). Bins are chosen to correspond to the jet selection criteria applied in the analysis. The sums of the entries are normalized to unity. Figure taken from Ref. [297].

The sensitivity of the search is shown in Section 7.2.4.2.

6.3.2.4 Search for decays of stopped LLPs

No particles with masses of the order of 100 GeV and significant lifetimes are present in the SM. Therefore, any sign of them would be an indication of new physics. At the LHC, the LLPs could stop inside the detector material if they lose all of their kinetic energy while traversing the detector, which will typically occur for particles with initial velocities $\beta < 0.5$ [298]. This energy loss can occur via nuclear processes if they are strongly interacting and/or through ionization if they are charged. The observation of a stopped particle decay signature would not only indicate new physics but also help measure the lifetime of LLPs, giving insights into various BSM scenarios.

If these stopped LLPs have lifetimes longer than tens of nanoseconds, most of their decays would be reconstructed as separate events unrelated to their production [299]. Owing to the difficulty of differentiating between the LLP decay products and SM particles from LHC pp collisions, these subsequent decays are most easily identified when there are no proton bunches in the detector. The detector is quiet during these out-of-collision time periods with the exception of rare noncollision backgrounds, such as cosmic rays, beam halo particles, and detector noise. If LLPs come to a stop in the detector, they are most likely to do so in the densest detector materials, which in the CMS detector are the ECAL, the HCAL, and the steel yoke in the muon system. If the stopped LLPs decay in the calorimeters, relatively large energy deposits occurring in the intervals between collisions could be observed. Furthermore, if the stopped LLPs decay into muons, displaced muon tracks out of time with the collisions could be detected. Both signatures require dedicated triggers to select events in between bunch crossings.

Two searches are performed for stopped LLPs that decay out of time with respect to the presence of proton bunches in the detector [206]. One search targets hadronic decays detected in

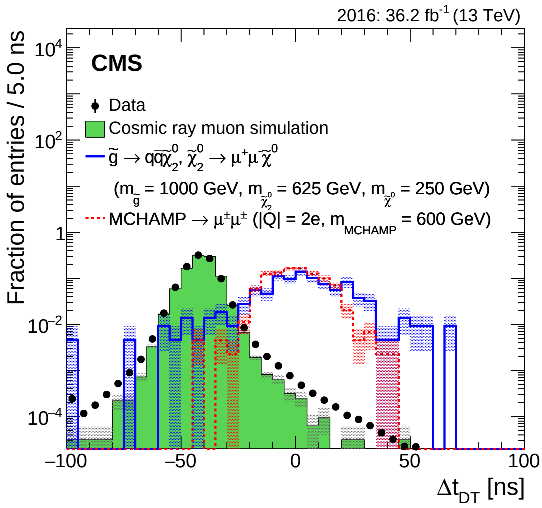

Figure 51: The muon timing distribution in the DTs for 2016 data, simulated cosmic ray muon events, and simulated signal events, for the muon channel of the stopped-LLPs search. The gray bands indicate the statistical uncertainty in the simulation. The histograms are normalized to unit area. Figure taken from Ref. [206].

the calorimeters and the other looks for decays into muon pairs in the muon system. These two search channels are analyzed independently using data collected in 2015 and 2016 with separate dedicated triggers. The triggers select calorimeter deposits or muons during gaps between proton bunches in the LHC beams. The calorimeter (muon) search uses $\sqrt{s} = 13$ TeV data corresponding to $\mathcal{L}_{\text{int}} = 38.6\,(39.0)\,\text{fb}^{-1}$ collected with LHC pp collisions separated by 25 ns during a search interval totaling 721 (744) hours. Figure 51 shows the muon timing distribution used in the muon search.

6.3.3 Signatures with LLPs and p_T^{miss}

Some DS models lead to striking signatures with both displaced particles and significant p_T^{miss}. This p_T^{miss} can arise from either stable particles, which could be a DM candidate, or from an LLP that escapes the detector before decaying. These signatures often have very low levels of SM backgrounds and can be sensitive to unique DS interpretations. Four of these searches are described below.

6.3.3.1 Searches for neutral LLPs decaying in the muon system

Reference [300] describes the first search at the LHC that uses a muon detector as a sampling calorimeter to identify showers produced by decays of LLPs. The analysis uses a data set corresponding to $\mathcal{L}_{\text{int}} = 137\,\text{fb}^{-1}$ collected during 2016–2018 with p_T^{miss} triggers. Based on a unique detector signature, the search is largely model-independent, with sensitivity to a broad range of LLP decay modes and to LLP masses as small as a few GeV. Decays of LLPs in the muon detectors induce hadronic and electromagnetic showers, giving rise to a high hit multiplicity in localized detector regions. The use of muon detector showers is described in detail in Section 4.4.5.

This first search effort used the CSC endcap muon detectors. To identify displaced showers, the CSC hits are clustered to form CSC clusters with a large hit multiplicity, which has a high efficiency of about 80% for $d\bar{d}$ and $b\bar{b}$ decays and 65% for $\tau^+\tau^-$ decays. A number of selections are applied to suppress SM background clusters from punch-through jets, muons that undergo bremsstrahlung, and decays of SM LLPs, such as the neutral kaon K_L^0.

A second analysis, presented in Ref. [301], is an extension of the muon endcap search described above and in Ref. [300]. This second analysis is the first search at the LHC that uses both the barrel and endcap muon detectors as a sampling calorimeter to identify showers produced by decays of LLPs. As in the previous search, the CSC/DT hits are clustered to form muon detector showers with a large hit multiplicity to identify displaced showers in the muon detector. The efficiency for this clustering is shown in Fig. 52.

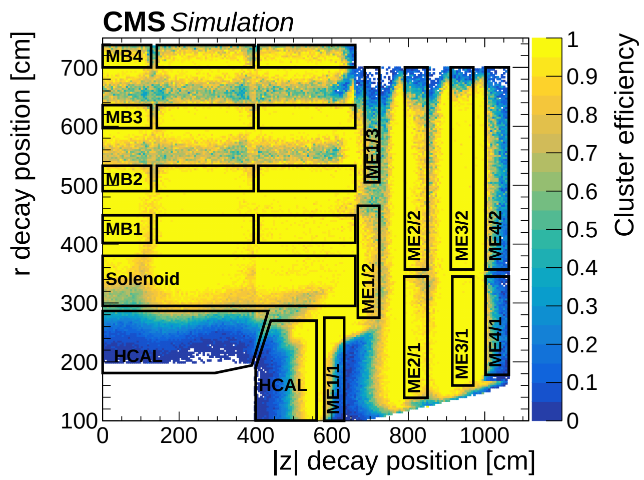

Figure 52: The cluster reconstruction efficiency as a function of the simulated r and $|z|$ decay positions of an LLP with a mass of 40 GeV and a range of $c\tau_0$ values between 1 and 10 m, for the search for neutral LLPs decaying in the muon system. Figure taken from Ref. [301].

The sensitivity of the search to EJ signatures is presented in 7.2.4.2. The sensitivity of the search to Higgs boson decays to LLPs is given in Section 7.2.4.4. The sensitivities to models containing heavy Z' and heavy H_D bosons are provided in Section 7.2.4.5.

6.3.3.2 Search for inelastic dark matter

The traditional "mono-X" approach can be combined with searches for LLPs to probe new models and new signatures. In this analysis [207], the final state of interest includes two displaced, nonresonant muons that are produced collinearly with the $\vec{p}_T^{miss}$ arising from the DM production. The DM and the muons also recoil against an ISR jet. The muons are too soft to be used for triggering, but by requiring the presence of a hard ISR jet in the final state, the use of data recorded with p_T^{miss} triggers is possible. The data sample corresponds to $\mathcal{L}_{int} = 138\,fb^{-1}$. The results are interpreted in the context of an IDM model [71, 139, 140], described in Section 2.2.3.

The event selection requires significant p_T^{miss} and hadronic activity. Two muons reconstructed with the DSA muon reconstruction algorithm [204–206] are required. The DSA muon reconstruction algorithm only uses information from the muon spectrometer system, but similar to the approach developed in Ref. [132], different categories of events are defined depending on whether the DSA muons can be matched to muons reconstructed using both the tracker and muon spectrometer. The minimum displacement min-d_{xy} distribution is shown in Fig. 53 for the most sensitive category.

6.3.3.3 Search for new physics with delayed jets

This search [199] presents the first use of timing signatures with the ECAL to identify delayed jets from the decays of heavy LLPs [302], using a data sample corresponding to $\mathcal{L}_{int} = 137\,fb^{-1}$.

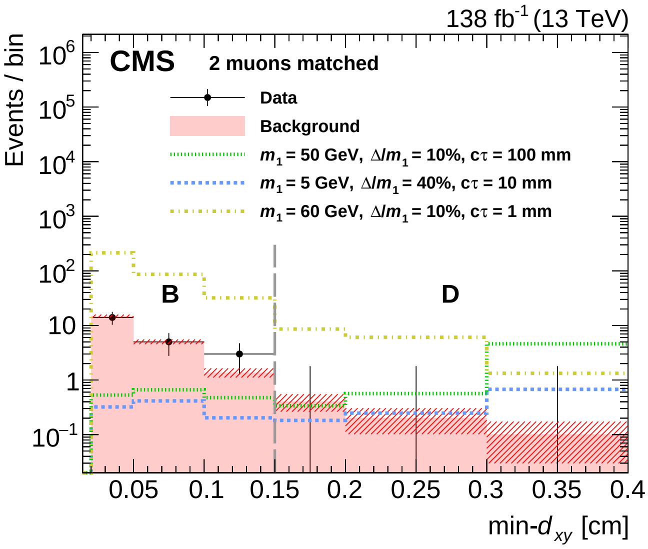

Figure 53: Measured min-d_{xy} distribution in the 2-match category of the IDM search, after requiring the min-d_{xy} muon to pass the isolation requirement $I_{\mathrm{PF}}^{\mathrm{rel}} < 0.25$. Overlaid with a red histogram is the background predicted from the region of the ABCD plane failing the same requirement, as well as three signal benchmark hypotheses (as defined in the legends), assuming $\alpha_{\mathrm{D}} = \alpha_{\mathrm{EM}}$ (the fine-structure constant). The red hatched bands correspond to the background prediction uncertainty. The last bin includes the overflow. Figure taken from Ref. [207].

The use of timing to provide sensitivity to LLPs is discussed in detail in Section 4.4.3. There are two effects that contribute to the time delay of jets from the decay of heavy LLPs relative to deposits from jets originating at the interaction point. First, the total path, composed of the initial LLP trajectory and the subsequent jet trajectories, will be longer, and second, the LLP will move with a lower velocity owing to its high mass, as was shown earlier in Fig. 16. The two contributions are shown in Fig. 54 for a representative LLP signal model. The use of this technique provides sensitivity to models with displacements significantly larger than those allowed by tracker-based searches.

This search for heavy BSM LLPs also requires that the events contain significant $p_{\mathrm{T}}^{\mathrm{miss}}$. The $p_{\mathrm{T}}^{\mathrm{miss}}$ can originate from invisible particles in the final state or from decays occurring beyond the detector acceptance. The $p_{\mathrm{T}}^{\mathrm{miss}}$ is used as a trigger requirement as it allows substantially lower thresholds than H_{T} triggers. A series of selections is performed to reject backgrounds from both prompt collisions and noncollision processes, such as cosmic ray muons and beam halo. Example selections include using the tracker to veto deposits originating from the interaction point and using the muon systems to reject beam halo and cosmic ray muon deposits. The remaining background components are individually characterized and their residual contributions are predicted using CRs in data.

The sensitivities to models containing a heavy Z′ boson are discussed in Section 7.2.4.5.

6.3.3.4 Search for LLPs with trackless and out-of-time jets and $p_{\mathrm{T}}^{\mathrm{miss}}$

Another search [200], which uses timing information, targets events with LLP decays into hadronically decaying Higgs or Z bosons with $p_{\mathrm{T}}^{\mathrm{miss}}$. Signal events are characterized by large $p_{\mathrm{T}}^{\mathrm{miss}}$, either because of the production of particles that do not interact with the detector material, or because of the LLP decaying at a macroscopic distance, outside of the calorimeters, and by the presence of trackless and out-of-time (OOT) jets. A hadronic LLP decay in the outer regions of the tracker or within the calorimeter volume will result in jets with a low track mul-

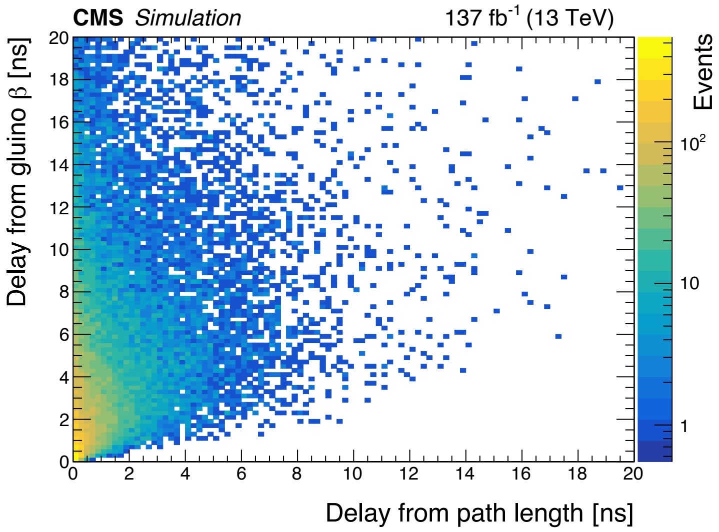

Figure 54: The contributions to the delay of the LLP from the path length and the lower velocity of the parent particle, in the delayed-jets search [199]. For this model, which features LLPs with proper decay lengths of 10 m and masses of 3 TeV, the lower velocity dominates the contribution to the delay.

tiplicity (nearly trackless) and OOT with regard to the LHC collisions. The time delay is due to the low speed of massive LLPs, heavier than 600 GeV, and the large flight distance to the outer parts of the detector.

The search uses $p_{\mathrm{T}}^{\mathrm{miss}}$ as a trigger selection and is performed on a data sample corresponding to $\mathcal{L}_{\mathrm{int}} = 138\,\mathrm{fb}^{-1}$. The features of the tracks and the electromagnetic calorimeter crystal hits associated with the jets induced by the LLP decays are the inputs of a DNN that tags trackless and OOT jets. The efficiency of the jet tagger as a function of LLP transverse decay length is shown in Fig. 55.

The sensitivities of the search to models containing heavy Z' and heavy $\mathrm{H_D}$ bosons are provided in Section 7.2.4.5.

6.3.3.5 Search for new physics with at least one displaced vertex and $p_{\mathrm{T}}^{\mathrm{miss}}$

This search [185] targets LLPs in signatures with at least one DV and $p_{\mathrm{T}}^{\mathrm{miss}}$ using pp collision events taken during 2016–2018 at $\sqrt{s} = 13\,\mathrm{TeV}$. The reconstruction of DVs is detailed in Section 4.4.1. This search expands on Ref. [184], which targets a pair of DVs and triggers on H_{T}. Compared to the search described in Section 6.3.2.2, this search aims to target DVs with low H_{T} and a broader range of displacement. A $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger is used to record events. A customized vertex reconstruction algorithm, which takes displaced tracks and iteratively creates vertices from them, is used to reconstruct DVs. A set of vertex selections is applied to avoid background vertices from material interactions and SM backgrounds originating from decays of particles with nonnegligible lifetimes, such as b hadrons. For LLP events with low H_{T}, fewer displaced tracks are available to be used for vertex reconstruction, and thus the vertex reconstruction efficiency is smaller. To overcome this difficulty, this search only requires one DV, which improves the search sensitivity to signal events with low H_{T} and longer LLP lifetime. After the vertex selections, the dominant source of background stems from the accidental crossing of tracks originating from the pp collision, which are fit to a spurious vertex. To further mitigate such background vertices, an interaction network, a machine-learning algorithm based on a GNN, is used as an event classifier. The distribution of the output score of the interaction network is

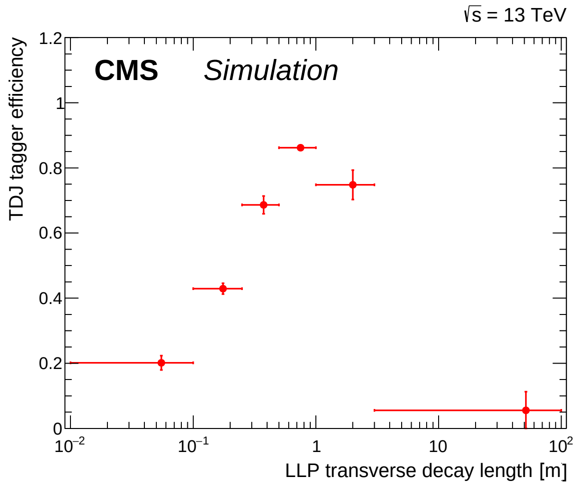

Figure 55: The efficiency of the jet tagger working point used in the trackless and OOT jets and $p_{\mathrm{T}}^{\mathrm{miss}}$ analysis shown as a function of the lab frame LLP transverse decay length. The uncertainties shown account for lifetime dependence and statistical uncertainty. Figure taken from Ref. [200].

shown in Fig. 56.

7 Results and reinterpretations

This section summarizes the results of the previously described DS searches performed by the CMS experiment. None of the searches produce evidence for the existence of new physics. Accordingly, limits on model parameters are presented in the following. The results also include reinterpretations of certain analyses in terms of DS models that are presented for the first time. The results are organized in terms of the DS models introduced in Section 2.

7.1 Simplified dark sectors

For simplified models of DSs, limits are presented as a function of the essential parameters of such models, which are the masses of the mediator (i.e., the portal) states, of the DM, as well as couplings and, for FIP models, mixing strengths.

7.1.1 Spin-1 portal

7.1.1.1 Vector and axial-vector portal

Summaries of the 95% CL observed exclusion limits in the plane of the mediator mass and the DM mass (the m_{med}-m_{DM} plane) for different $p_{\mathrm{T}}^{\mathrm{miss}}$-based DM searches in the leptophobic vector and axial-vector models are presented in Table 4 and Fig. 57. Summaries of the 95% CL observed exclusion limits for a nonleptophobic vector and axial-vector mediator are presented in Fig. 58. In order to compare with various DD experiments, the 90% CL observed exclusion limits from the vector (axial-vector) model are converted to upper limits on the spin-independent (-dependent) DM-nucleon scattering cross section [24] and shown in Fig. 59, where σ_{SI} (σ_{SD}) is the spin-independent (-dependent) DM-nucleon scattering cross section.

Cross section exclusions can be converted to limits on g_{q} assuming the benchmark values for the DM coupling $g_{\mathrm{DM}} = 1.0$ and DM mass $m_{\mathrm{DM}} = m_{\mathrm{Z'}}/3$ [312], following the procedure outlined in Ref. [313]. Briefly, in the narrow-width approximation, the dependence of the cross

7.1.3 Fermion portal

Figure 69 presents 95% CL limits for the fermion portal model, obtained from the monojet search [81]. In the specific model probed, the mediator Φ couples to DM particles and right-handed u quarks with coupling strength $\lambda = 1$. Exclusions are presented in terms of the DM mass and the mass of the mediator.

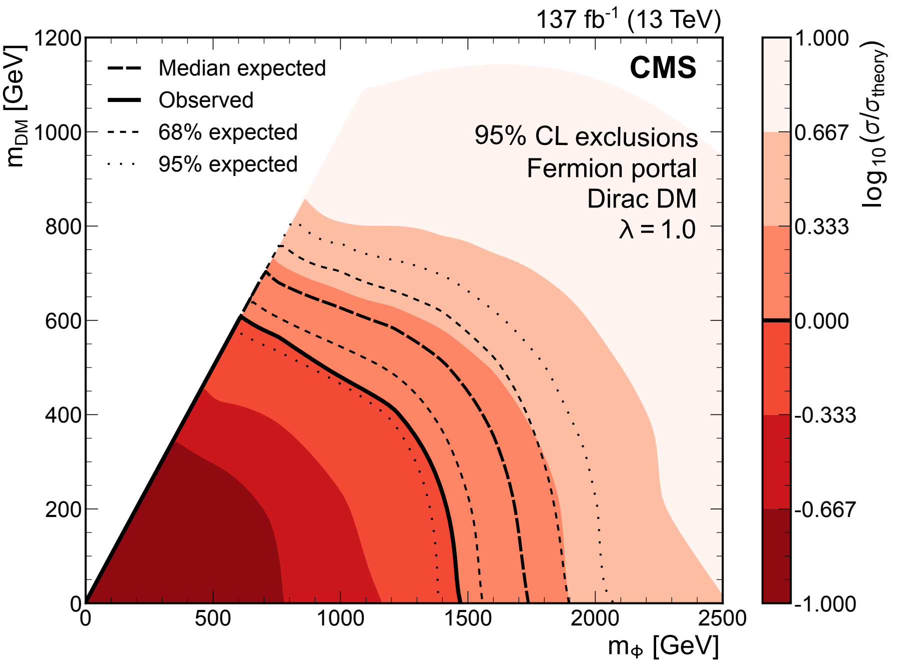

Figure 69: Observed (solid line) and expected (dashed lines) exclusions at 95% CL in the m_Φ-m_{DM} plane for the fermion portal model scenario obtained from the monojet search performed using data collected in 2016–2018. Figure adapted from Ref. [81].

7.2 Extended dark sectors

7.2.1 The 2HDM+a scenario

This section presents results interpreted in the 2HDM+a, as described in Section 2.2.1. A summary table and a plot for the 95% CL observed exclusion limits in the m_{a}-m_{A} plane for different $p_{\mathrm{T}}^{\mathrm{miss}}$-based DM searches from CMS are presented in Table 8 and Fig. 70, respectively. From the figure it can be seen that the mono-Z analysis sets exclusion limits that depend on the ratio of the pseudoscalar masses $m_{\mathrm{A}}/m_{\mathrm{a}}$; this is because the process is dominated by resonant production of the heavy scalar H and subsequent decay $\mathrm{H} \to \mathrm{Za}$; an analogous situation occurs in the mono-Higgs analysis, with the $\mathrm{A} \to \mathrm{Ha}$ channel being dominant instead. On the other hand, the exclusion limit set by the monojet analysis is almost independent of m_{A}; this is because in this case the process reduces to the simplified model case with $p_{\mathrm{T}}^{\mathrm{miss}}$ +ISR, and the heavier pseudoscalar plays essentially no role.

Figure 71 summarizes searches for the 2HDM+a scenario that approach the problem from the viewpoint of exotic decays of the 125 GeV Higgs boson instead. If the $\mathrm{a} \to \chi\chi$ decay is not kinematically allowed, searches for the visible products of the $\mathrm{H} \to \mathrm{aa}$ process are the most stringent. Otherwise, the interpretation of the Higgs boson invisible decay limits in terms of the 2HDM+a scenario gives the strongest limits.

Table 8: Summary of 95% CL observed exclusion limits in the heavy pseudoscalar mass m_A for p_T^{miss}-based DM searches from CMS in the 2HDM+a scenario. Each search listed here used data corresponding to $\mathcal{L}_{int}$=137 fb^{-1}.

Reference	Channel	95% CL lower limit on m_A [TeV]
[86]	Mono-Z	1.2
[81]	Monojet	0.39
[270]	Mono-Higgs	1.0

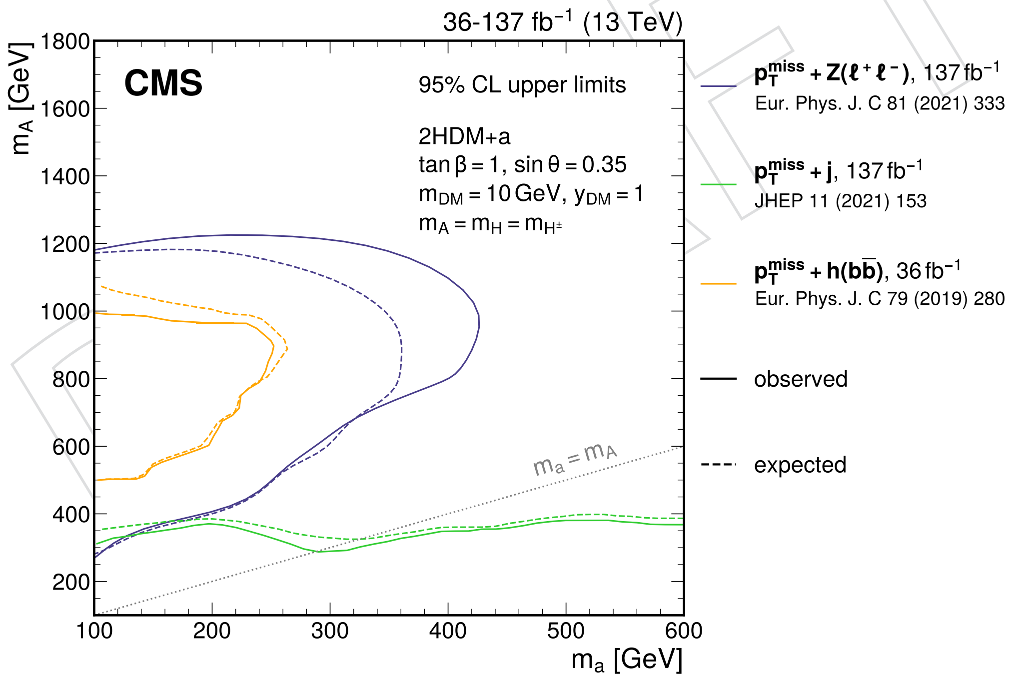

Figure 70: Observed (solid lines) and expected (dashed lines) exclusion regions at 95% CL in the m_a-m_A plane for the 2HDM+a scenario arising from various "mono-X" searches performed using data collected in 2016–2018 [81, 86, 270]. Following the recommendation of the LHC DM Working Group [24, 25], the projection is performed for values of the other parameters as follows: $m_H = m_A = m_{H^\pm}$, $\sin\theta = 0.35$, $\tan\beta = 1$, $m_{DM} = 10\,\text{GeV}$, and $y_{DM} = 1$.

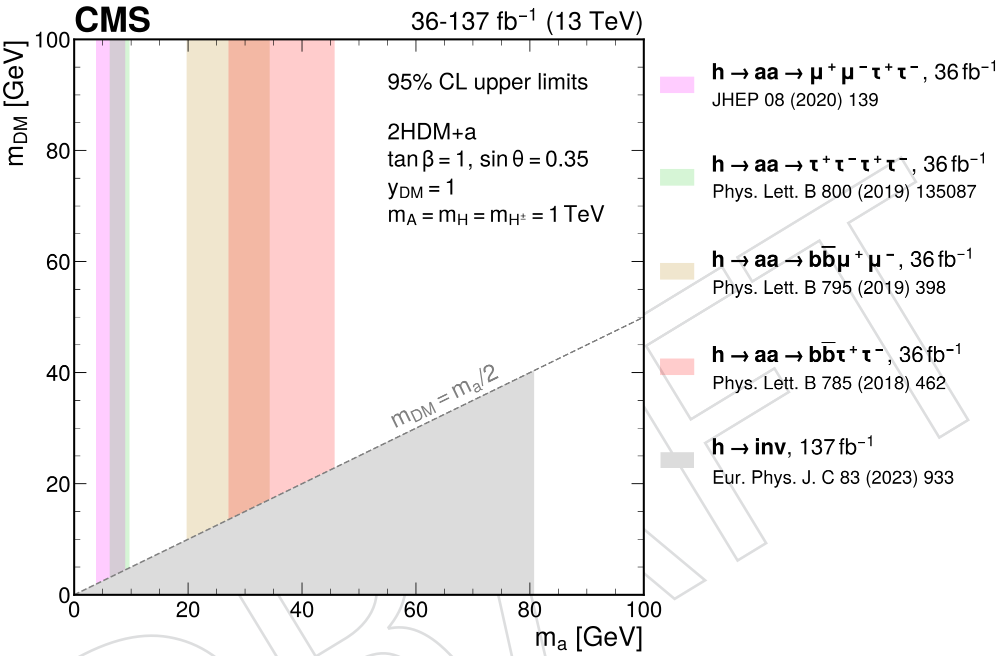

Figure 71: Exclusion regions at 95% CL in the m_a-m_{DM} plane for the 2HDM+a scenario arising from searches for exotic and invisible decays of the 125 GeV Higgs boson performed using data collected in 2016–2018 [85, 318–321]. Following the recommendation of the LHC DM Working Group [24, 25], the projection is performed for values of the other parameters as follows: $m_H = m_A = m_{H^\pm} = 1$ TeV, $\sin\theta = 0.35$, $\tan\beta = 1$, and $y_{DM} = 1$. The branching fractions of the pseudoscalar boson to SM and DM particles are computed using the MADWIDTH [322] functionality within MADGRAPH5_aMC@NLO.

7.2.2 Supersymmetry

7.2.2.1 Dark supersymmetry and Hidden Abelian Higgs model

Results interpreted in a dark SUSY scenario and in the HAHM, as described in Sections 2.2.2.2 and 2.2.2.1, respectively, are presented in this section. Figure 72 shows a summary of LLP results for dark bosons, in contrast to the dark photon summary with prompt analyses shown in Section 7.1.1.2. Three analyses are covered in this figure. The first is a search for displaced dimuons [132] with a HAHM signal benchmark (Section 2.2.2.1). The second analysis, which uses the same benchmark model, is a search for displaced dimuon resonances with data scouting [294]. The third search evaluates the CMS sensitivity to displaced dimuons in final states with $4\mu + X$ in the context of a dark SUSY signal scenario (Section 2.2.2.2) [293]. For all three searches, $\mathcal{B}(\mathrm{h} \to 2\mathrm{A}') = 1\%$ is assumed. The $\mathcal{L}_{\mathrm{int}}$ used for each analysis varies depending on the available triggers and data sets at that time.

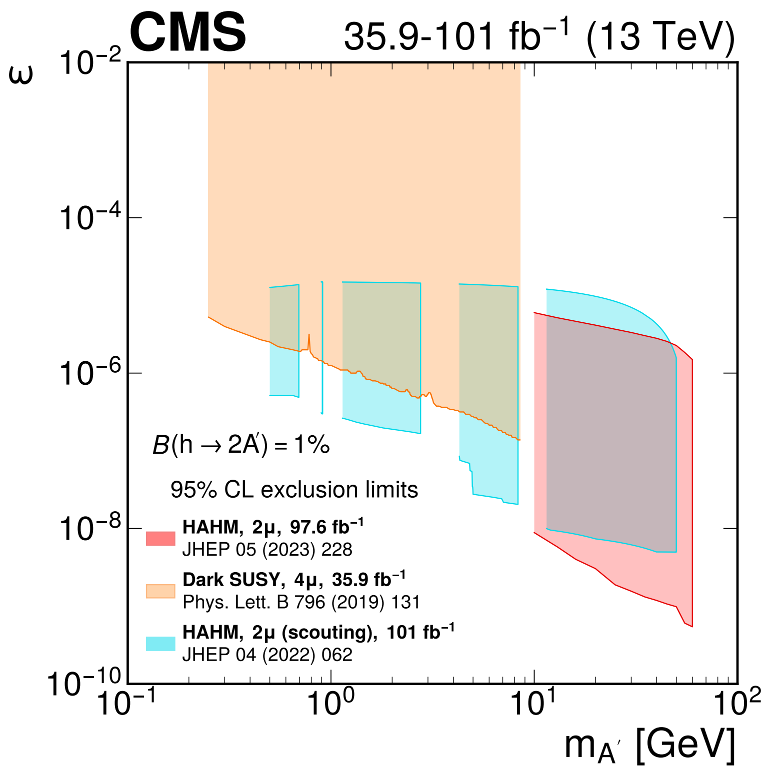

Figure 72: Observed 95% CL exclusion contours in the plane defined by the kinetic mixing parameter (ϵ) and the mass of the new dark boson. A summary of Run 2 CMS searches focusing on displaced signatures is presented. Two of those searches, namely Refs. [132] (red) and [294] (blue), consider the HAHM signal and use a final state with at least two muons $(2\mu + X)$, and the latter one uses data scouting. The third search (orange) [293] uses a final state with at least four muons $(4\mu + X)$ and a dark SUSY signal scenario.

7.2.2.2 Stealth supersymmetry

Stealth SUSY models are detailed in Section 2.2.2.3. The stealth SUSY search described in Section 6.2.3.3 targeted top squark pair production with decays via the stealth vector portal, and the limits on this model are shown in the upper plot in Fig. 73. However, other portals such as a Higgs portal are possible. The main difference between the two scenarios is that the six gluons in the event are replaced by four b quarks, resulting in a reduction of the number of jets in the event. However, the signal still features many more jets, as shown in Fig. 8, than the dominant $\mathrm{t\bar{t}}$ background, and thus, sensitivity to this model is still expected for this search. The full analysis chain was used to interpret the results in the context of the stealth Higgs portal. No changes were made to the original analysis.

The lower plot in Fig. 73 shows the expected and observed 95% CL upper limit on the product of the top squark pair production cross section and branching fraction via the Higgs portal in terms of the top squark mass. The branching fractions are assumed to be 100% for the chosen decay chain: $\widetilde{\mathrm{t}} \to \mathrm{t}\widetilde{\mathrm{S}}$, $\widetilde{\mathrm{S}} \to \mathrm{S}\widetilde{\mathrm{G}}$, and $\mathrm{S} \to \mathrm{b\bar{b}}$. The observed (expected) mass exclusion is found to be 570 (670) GeV, compared to 870 (920) GeV for the vector portal and 670 (720) GeV for the RPV model. The sensitivity can be improved by explicitly taking advantage of the additional b quarks expected from decays via the Higgs portal.

Considering the $\mathrm{SY\overline{Y}}$ and Higgs portal stealth SUSY models discussed above, if the singlino is long lived, then dedicated LLP searches could be sensitive to these SUSY models. In addition to the stealth SUSY search [290], four LLP-style searches, including the displaced-jets search [187], the DVs search [184], the trackless- and OOT-jets search [200], and the muon system showers search (MS clusters) [301] reinterpret their analyses for these stealth SUSY models, where the proper decay length of the singlino ($c\tau_{\widetilde{\mathrm{S}}}$) ranges from 0.01 mm to 1000 mm. Figure 74 shows observed exclusions on the product of the top squark pair production cross section and branching fraction in terms of the top squark mass and proper decay length of the singlino for the $\mathrm{SY\overline{Y}}$ and Higgs portal versions of the stealth SUSY model. Two singlino mass scenarios are considered: where $m_{\widetilde{\mathrm{S}}} = 100\,\mathrm{GeV}$ and where $m_{\widetilde{\mathrm{S}}} = m_{\widetilde{\mathrm{t}}} - 225\,\mathrm{GeV}$. The branching fractions are assumed to be 100% for the decay chain for either the $\mathrm{SY\overline{Y}}$ ($\widetilde{\mathrm{t}} \to \mathrm{t}\widetilde{\mathrm{S}}\mathrm{g}$, $\widetilde{\mathrm{S}} \to \mathrm{S}\widetilde{\mathrm{G}}$, and $\mathrm{S} \to \mathrm{gg}$) or Higgs portal ($\widetilde{\mathrm{t}} \to \mathrm{t}\widetilde{\mathrm{S}}$, $\widetilde{\mathrm{S}} \to \mathrm{S}\widetilde{\mathrm{G}}$, and $\mathrm{S} \to \mathrm{b\bar{b}}$). Each exclusion contour bounds (bounding direction denoted by hatching) the 2D parameter space that is excluded by the respective search.

7.2.3 Inelastic dark matter

The first dedicated collider search for IDM has been conducted by the CMS Collaboration [207] and is described in Section 6.3.3.2. No evidence for the signal is observed. Limits at 95% CL are set on the product of the DM production cross section and decay branching fraction of the excited state $\sigma(\mathrm{pp} \to \mathrm{A}' \to \chi_2\chi_1)\,\mathcal{B}(\chi_2 \to \chi_1\mu^+\mu^-)$. These limits can be translated into limits on the interaction strength y and the DM particle mass m_{DM}, in terms of the mass split Δ between the DM states and the coupling strength α_{dark} of the DS gauge interaction. That translation has a strong dependency both on α_{dark} itself and on the dark photon (mediator) mass m_{med}, therefore the results, shown in Fig. 75 for the 10% mass-split scenario, are presented for the recommended $m_{\mathrm{DM}} = m_{\mathrm{med}}/3$ choice and for two α_{dark} hypotheses. For $\Delta = 0.1\,m_{\mathrm{DM}}$, at $m_{\mathrm{DM}} = 3$ and 80 GeV respectively, values of y greater than $\approx 10^{-7}$–10^{-6} are excluded for the $\alpha_{\mathrm{dark}} = 0.1$ hypothesis. Conversely, for the $\alpha_{\mathrm{dark}} = \alpha_{\mathrm{EM}}$ hypothesis, values of y greater than $\approx 10^{-8}$–10^{-7} are excluded for the same m_{DM} values. The A'-Z resonance effect greatly improves the limits when $m_{\mathrm{DM}} \simeq 30\,\mathrm{GeV}$.

7.2.4 Hidden valleys

This section presents results interpreted in dark QCD models, as described in Section 2.2.4.

7.2.4.1 Semivisible jets

As explained in Section 6.1.3.1, we reinterpret the dijet resonance search and monojet DM

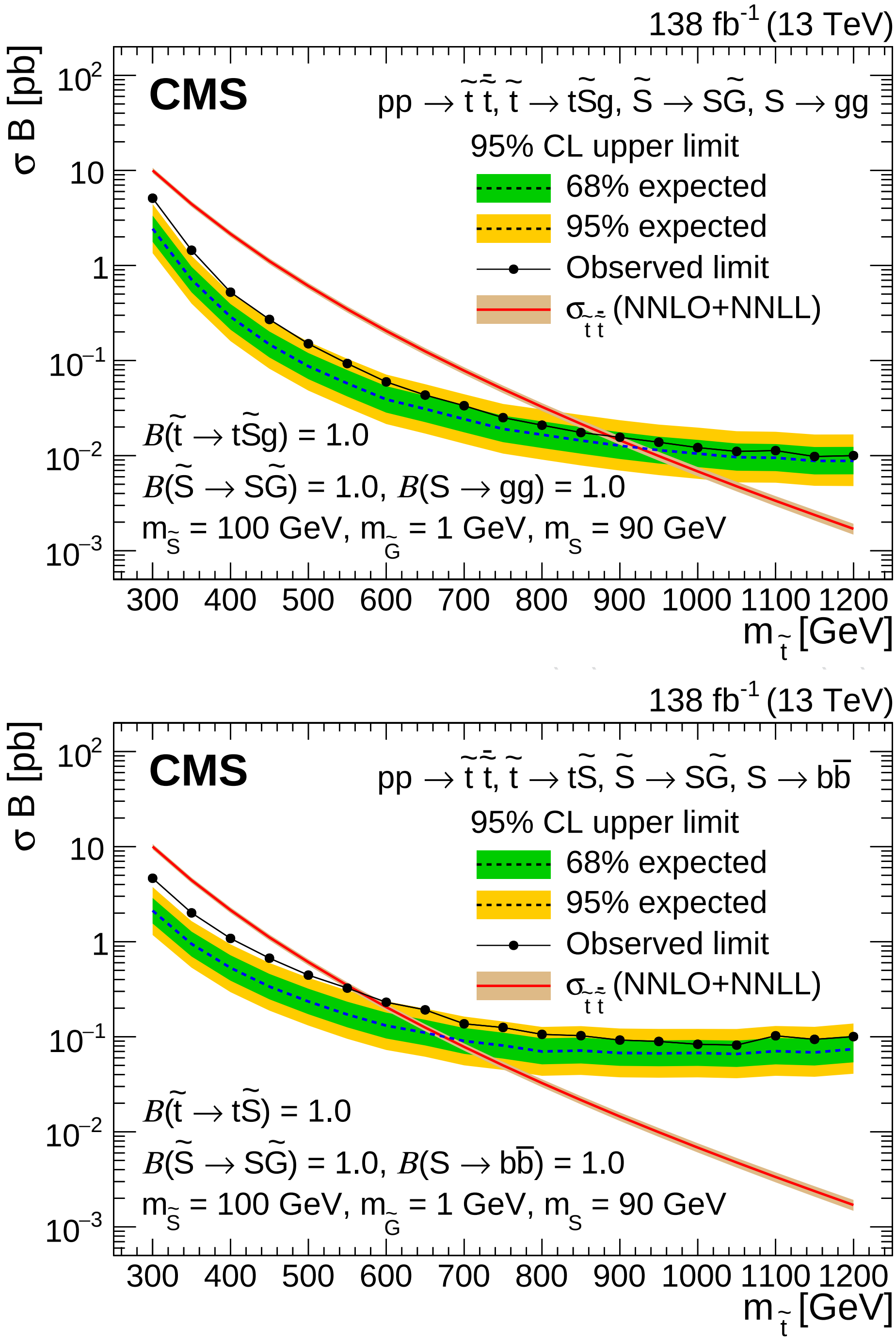

Figure 73: Expected and observed 95% CL upper limit on the product of the top squark pair production cross section and branching fraction in terms of the top squark mass for the stealth SYY SUSY model (upper) and stealth SHH SUSY model (lower). Particle masses and branching fractions assumed for the model are included. The expected cross section is computed at NNLO accuracy, improved by using the summation of soft gluons at next-to-next-to-leading logarithmic (NNLL) order, and is shown in the red curve. Upper figure adapted from Ref. [290].

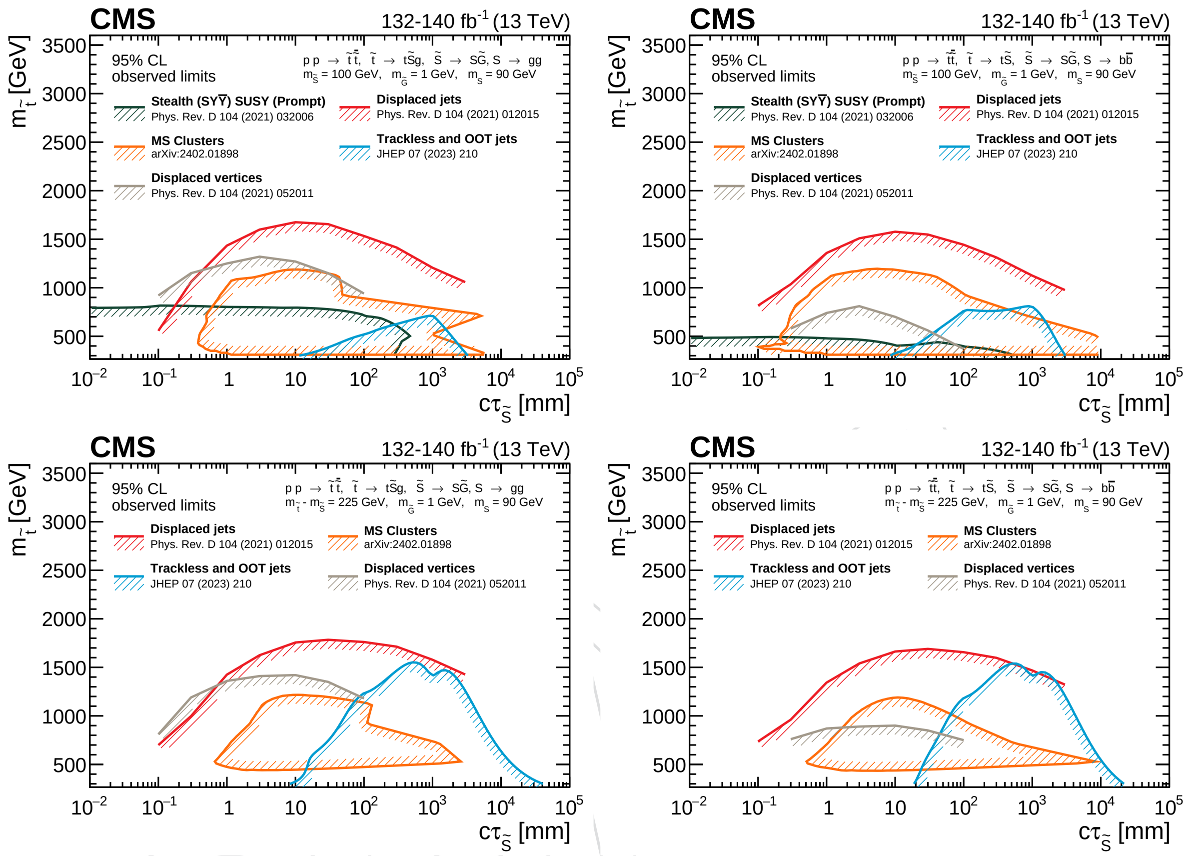

Figure 74: Observed 95% CL exclusions of the product of the top squark pair production cross section and branching fraction as functions of the top squark mass and proper decay length of the singlino for the stealth SY$\overline{\text{Y}}$ (left) and stealth SHH (right) SUSY model where the mass of the singlino is 100 GeV (upper) and $m_{\tilde{t}} - 225$ GeV (lower). Exclusions are for the stealth SUSY search [290] (dark green), the displaced vertices search [184] (gray), the displaced-jets search [187] (red), the trackless- and OOT-jets search [200] (blue), and muon system showers search (MS clusters) [301] (orange). The hatching direction on each contour denotes the region of excluded 2D phase space that is bounded by the respective contour. Note that the displaced-jets search has no sensitivity less than $c\tau_{\tilde{S}} = 0.1$ mm, the DVs search has no sensitivity less than $c\tau_{\tilde{S}} = 0.1$ (0.3) mm for the SY$\overline{\text{Y}}$ (SHH) model, and the stealth SUSY search has no sensitivity to either stealth SUSY model when $m_{\tilde{S}} - m_{\tilde{t}} = 225$ GeV. Additionally, for the specific result here, the muon system showers search only uses the CSCs component of the muon system.

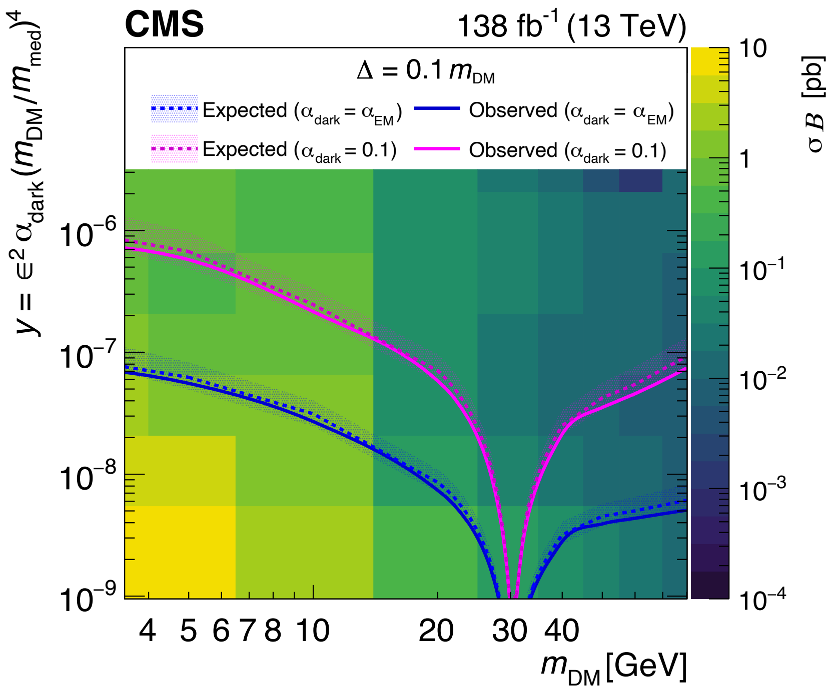

Figure 75: Two-dimensional exclusion surface in the search for IDM, assuming $\Delta = 0.1\,m_{DM}$, in terms of the DM mass m_{DM} and the signal strength y, with $m_{med} = 3\,m_{DM}$. Filled histograms denote observed limits on $\sigma(pp \to A' \to \chi_2\chi_1)\,\mathcal{B}(\chi_2 \to \chi_1\mu^+\mu^-)$. Solid (dashed) curves denote the observed (expected) exclusion limits at 95% CL, with 68% CL uncertainty bands around the expectation. Regions above the curves are excluded, depending on the α_{dark} hypothesis: $\alpha_{dark} = \alpha_{EM}$ (dark blue) or 0.1 (light magenta). The sensitivity is higher in the region near $m_{DM} \approx 30$ GeV or $m_{med} \approx 90$ GeV because of the A′ mixing with the Z boson in that mass range. Figure adapted from Ref. [207].

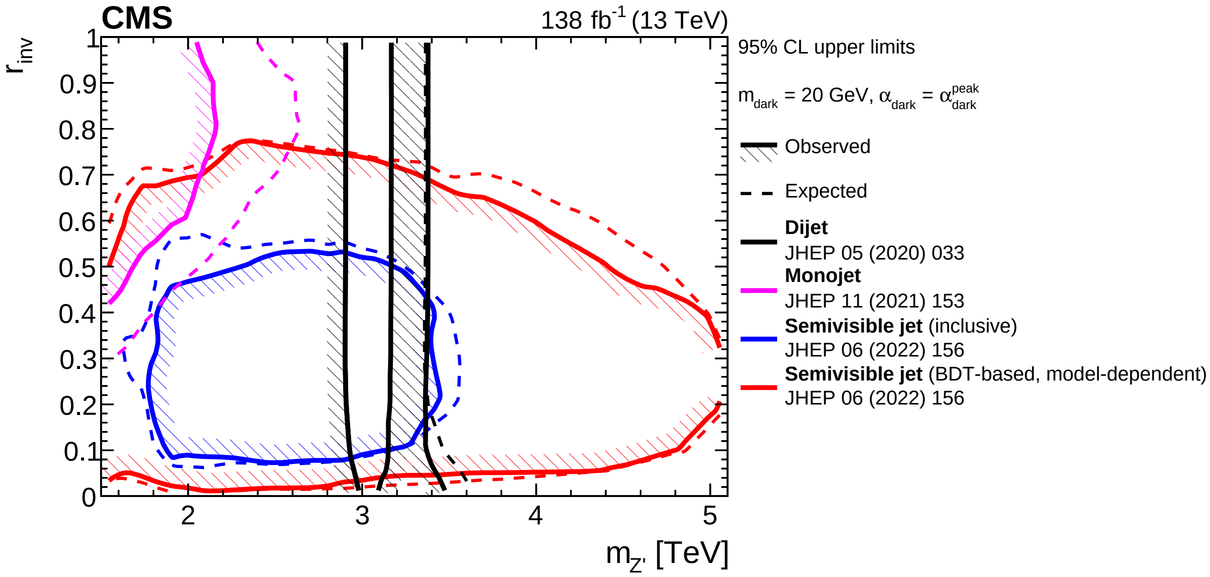

Figure 76: Observed and expected 95% CL excluded regions of the $m_{Z'}$-r_{inv} plane from the dedicated SVJ search [148], the dijet search [277] (Section 6.2.2.2), and the monojet search [81] (Section 6.1.1.1). The hashed areas indicate the direction of the excluded area from the observed limits.

search for the SVJ model. For the dijet resonance search, following Ref. [277], the background estimation from CRs in data is used for signals with $m_{Z'} \geq 3\,\text{TeV}$, while the analytic fit-based background estimation is used for lower $m_{Z'}$. For the reinterpretation of the monojet search, we use the MADANALYSIS implementation [317].

The results from both reinterpretations are compared to the results from the dedicated SVJ search, with and without the BDT tagger, in Fig. 76. The complementary sensitivities of each strategy are clearly visible. The monojet search is more sensitive for large r_{inv} values, and the standard DM reinterpretation of the dijet search, effectively considering only $Z' \to q\bar{q}$ events, also provides good sensitivity in this region. Accounting for the combination of effects of SVJ model parameters on observables used in the monojet search, the most stringent exclusion is found for $r_{inv} = 0.8$, as this maximizes the overall selection efficiency for SVJ signals. For very small r_{inv} values, the reinterpreted dijet search provides the best sensitivity. At intermediate r_{inv} values, the dedicated SVJ search is the most sensitive, especially when the BDT is used to identify SVJs, though the latter strategy introduces more model dependence. The advantage of the dedicated strategy would increase with the branching fraction for $Z' \to q_{dark}\bar{q}_{dark}$, which grows for larger $g_{q_{dark}}$ or smaller g_q values.

These cross section limits can be interpreted as limits on g_q for fixed parameter values $g_{q_{dark}} = 0.5$ and $m_{dark} = 20\,\text{GeV}$, following the procedure described in Section 7.1.1.1. For both the SVJ search and the reinterpretation of the monojet search, the initial and final states for the procedure are $q\bar{q}$ and $q_{dark}\bar{q}_{dark}$, respectively. Those searches do not depend strongly on the Z' boson width within the narrow-width regime, because the resolutions of the search variables are intrinsically limited by the information lost in p_T^{miss}. In contrast, the resolution of the dijet mass used in the dijet search is small enough that even minor increases in the mediator width are visible [278]. Therefore, the existing g_q exclusion from the dijet search is used directly; though this underestimates the exclusion at small r_{inv}, SVJ events do not contribute to the dijet search limit for $r_{inv} \gtrsim 0.1$, so this is a reasonable approximation in the majority of the signal model parameter space. Figure 77 shows the excluded values of g_q for SVJ signals from all searches for two representative values $r_{\mathrm{inv}} = 0.3$ and 0.6. Only values that satisfy the narrow-width approximation $\Gamma_{Z'}/m_{Z'} < 10\%$ are shown. For $r_{\mathrm{inv}} = 0.3$, the acceptance of the SVJ search is maximized, and even without the BDT tagger, it provides the strongest exclusions for a wide range of Z' boson masses. For $r_{\mathrm{inv}} = 0.6$, the BDT-based SVJ search still provides a strong exclusion even as the search acceptance decreases, while the monojet search has the best exclusion at small $m_{Z'}$. The $r_{\mathrm{inv}} = 1$ case is equivalent to the vector DM simplified model, so the coupling exclusion from the dijet and monojet searches can be seen in Fig. 60.

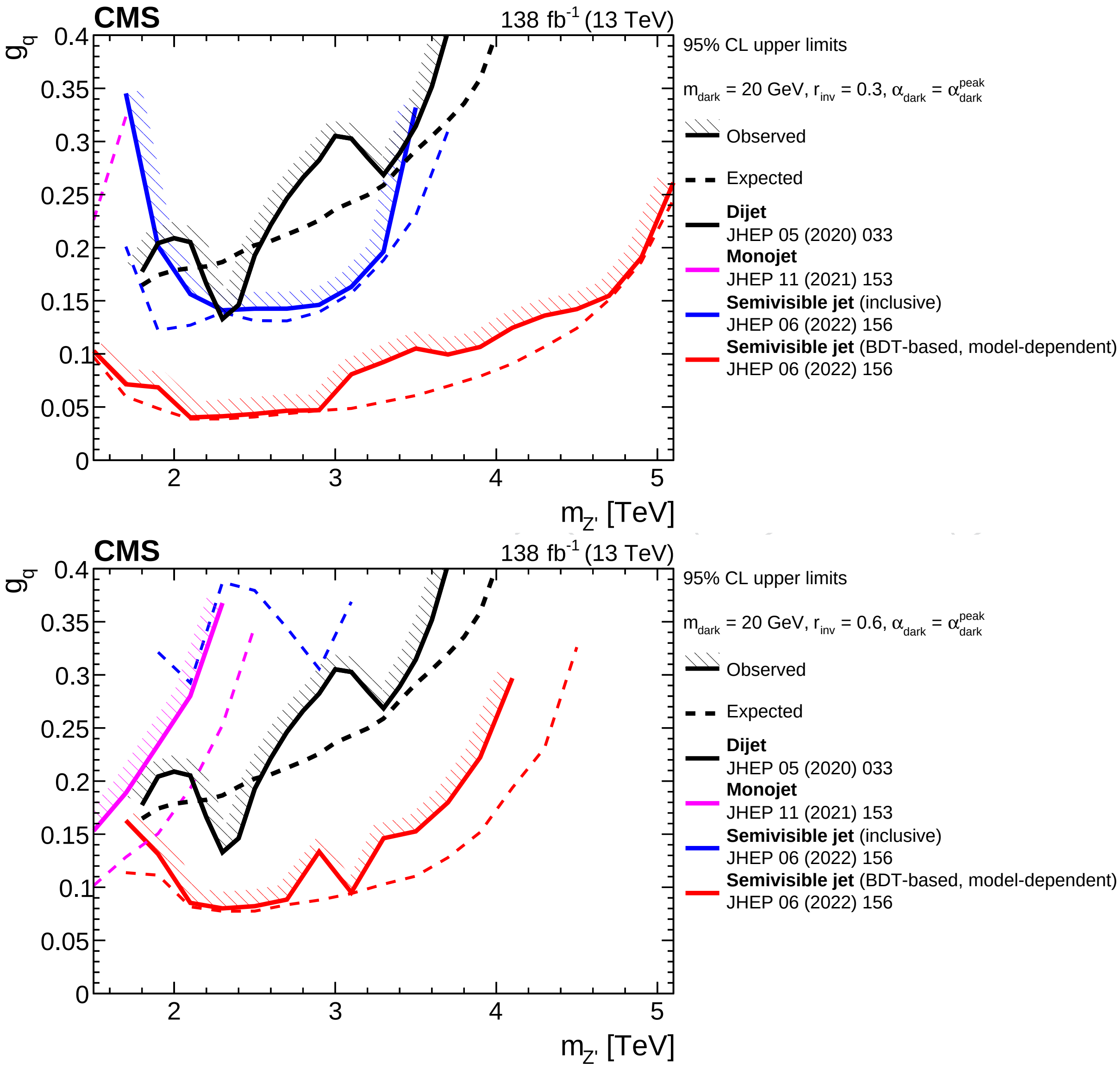

Figure 77: Observed and expected 95% CL exclusion limits on g_{q} for SVJ signals from the dedicated SVJ search [148], the dijet search [277], and the monojet search [81], for $r_{\mathrm{inv}} = 0.3$ (upper) and $r_{\mathrm{inv}} = 0.6$ (lower). The hashed areas indicate the direction of the excluded area from the observed limits. The observed limits from the monojet search in the upper plot and the inclusive SVJ search in the lower plot are outside the range of validity of the narrow-width approximation, so they are not shown.

7.2.4.2 Emerging jets

The track-based EJ search and the muon detector shower search (Sections 6.3.2.3 and 6.3.3.1) have complementary sensitivity to EJ signatures, targeting smaller and larger lifetimes, respectively. The exclusion limits for unflavored and flavor-aligned EJ models from both searches are shown in Fig. 78 for signals with $m_{\text{dark}} = 10\,\text{GeV}$. For the dedicated EJ search, the results from both the model-agnostic EJ tagger and the model-dependent GNN tagger are shown. For the muon detector shower search, results are obtained by clustering CSC hits. The sensitivity of the muon detector shower search to the flavor-aligned model is reduced because this model has a broader spread of lifetimes and therefore fewer particles reach the muon detectors.

Other LLP searches are not sensitive to EJ models, for various reasons. The searches for delayed jets (Section 6.3.3.3) and trackless jets (Section 6.3.3.4) use timing measurements that rely on the exotic particles being sufficiently delayed, and EJs do not satisfy this requirement. The displaced-jet search (Section 6.3.2.1) uses triggers that require at most two prompt tracks to be associated with the jets, which rejects most EJs because they contain tracks with a broader mix of displacements. The DV search (Section 6.3.2.2) relies on reconstructing DVs, which is inefficient for EJs, as each vertex tends to have only a few tracks associated with it.

As detailed in Section 2.2.4.2, dark QCD signatures may be produced through decays of various mediators, such as the SM Higgs boson, to dark hadrons. The search for neutral decays in the muon system (Section 6.3.3.1) is also interpreted using a set of perturbative benchmark dark QCD models [143]. The decays back to the SM can proceed via multiple portals, comprehensively considered in Ref. [301]. The representative exclusion limits for two decay portals are shown in Fig. 79.

7.2.4.3 Soft unclustered energy patterns

The SUEP search (Section 6.2.3.2) is interpreted in terms of limits on the production cross section for different values of the signal model parameters m_{S}, m_{dark}, and T_{dark}. The excluded ranges in the m_{S}–m_{dark}–$m_{\text{A}'}$–T_{dark} parameter space are obtained by comparing the expected and observed cross section limits to the theoretical signal cross section. Figure 80 shows the exclusions for all m_{S} values in the plane of m_{dark} and T_{dark} with $m_{\text{A}'} = 1.0\,\text{GeV}$. Similar exclusions are obtained for other $m_{\text{A}'}$ values and their corresponding decay patterns. In the signal models with the highest track multiplicity, corresponding to the most SUEP-like signatures and arising when $m_{\text{S}}/T_{\text{dark}} \approx m_{\text{S}}/m_{\text{dark}} \approx 100$, the most stringent limits are set.

7.2.4.4 Higgs boson decays into long-lived particles

Exotic decays of the Higgs boson into LLPs are well motivated in a variety of models, such as those motivated by neutral naturalness, as described in Section 2.2.4.4. Several CMS searches have been reinterpreted in a scenario in which an exotic Higgs boson is produced in pp collisions and then decays into two LLPs, here denoted X (as shown in the right diagram in Fig. 7). These reinterpretations are shown in Figs. 81, 82, and 83. Figure 81 shows the upper limits on the branching fraction of Higgs bosons decaying into LLPs with masses between 40 and 55 GeV, as functions of the LLP proper decay length. Figure 82 shows the same but for masses between 15 and 30 GeV, and Fig. 83 shows the same but for masses between 0.4 and 7 GeV.

7.2.4.5 Heavy long-lived particles

Dark sectors may have complex constituents including TeV scale scalar and vector bosons that decay into LLPs in the DS as well as to DM candidate particles [142]. This can include scenarios motivated by neutral naturalness, as described in Section 2.2.4.4. The LLPs may be boosted if their mass is significantly less than the parent particle. These particles can typically decay both to displaced leptonic and hadronic final states. The displaced signatures that can be recon-

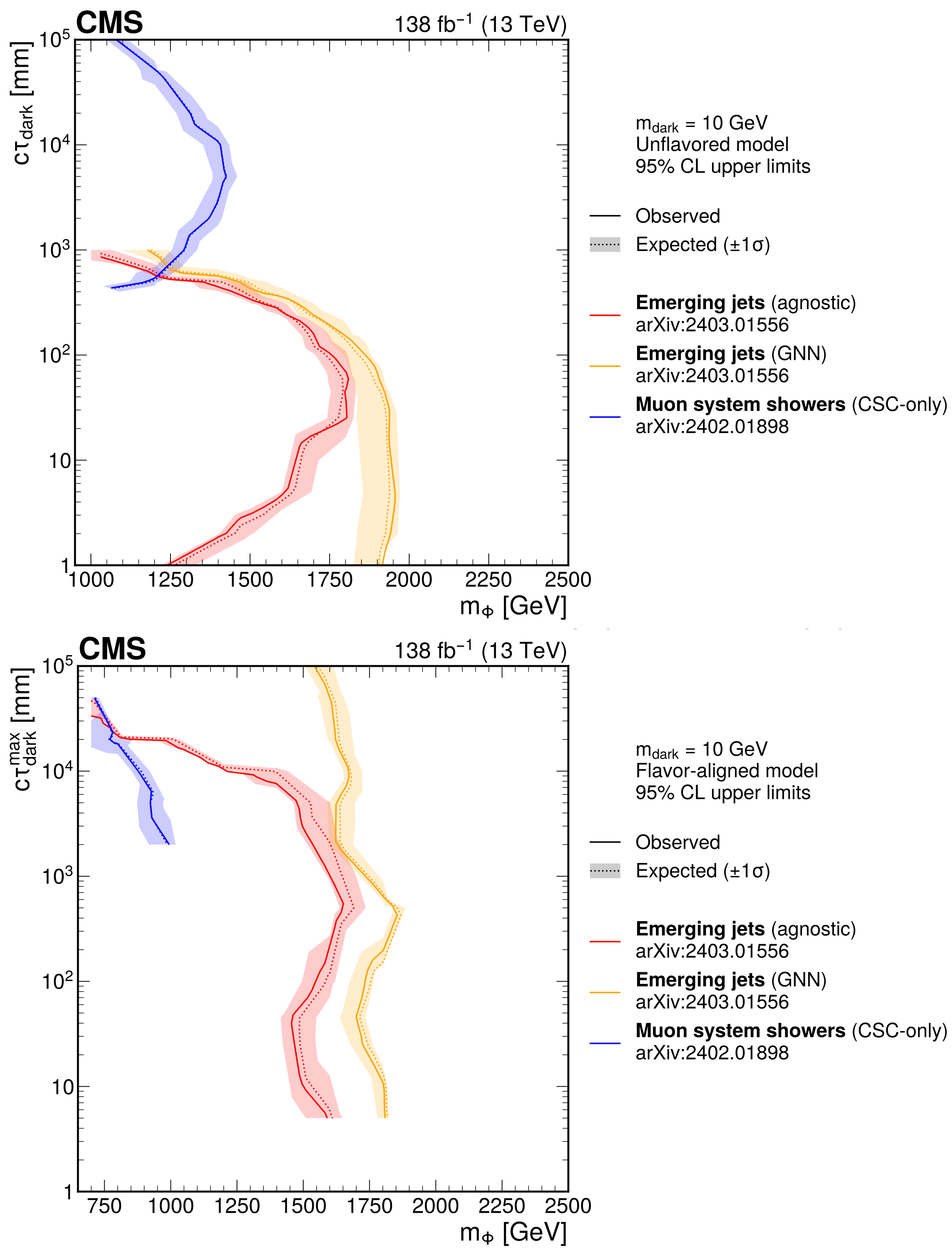

Figure 78: Observed and expected 95% CL exclusion limits from the track-based [297] and muon detector shower-based [301] searches for pair production of a bifundamental mediator that decays into a jet and an emerging jet, for $m_{\text{dark}} = 10\,\text{GeV}$ and various choices of Φ masses and π_{dark} proper decay lengths, in the unflavored model (upper) and the flavor-aligned model (lower).

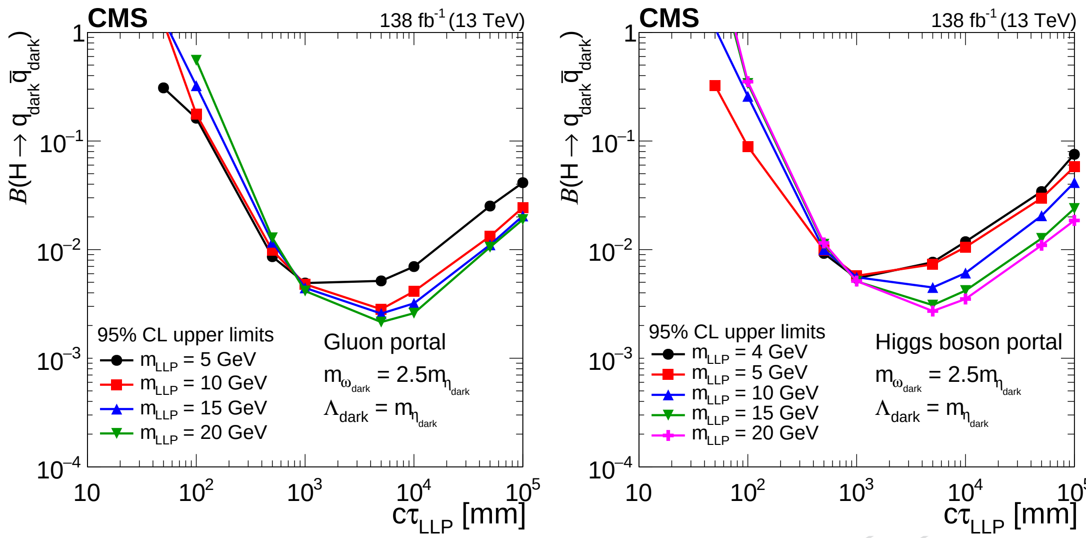

Figure 79: Observed 95% CL exclusion limits on the branching fraction of the Higgs boson decay into DS hadrons, Ψ, for the search for neutral decays in the muon system (Section 6.3.3.1). Sensitivity for the gluon (left) and Higgs boson (right) DS decay portals are shown. The model parameters considered here are $m_{\omega_{\text{dark}}} = 2.5 m_{\eta_{\text{dark}}}$, $\Lambda_{\text{dark}} = m_{\eta_{\text{dark}}}$. Figure adapted from Ref. [301].

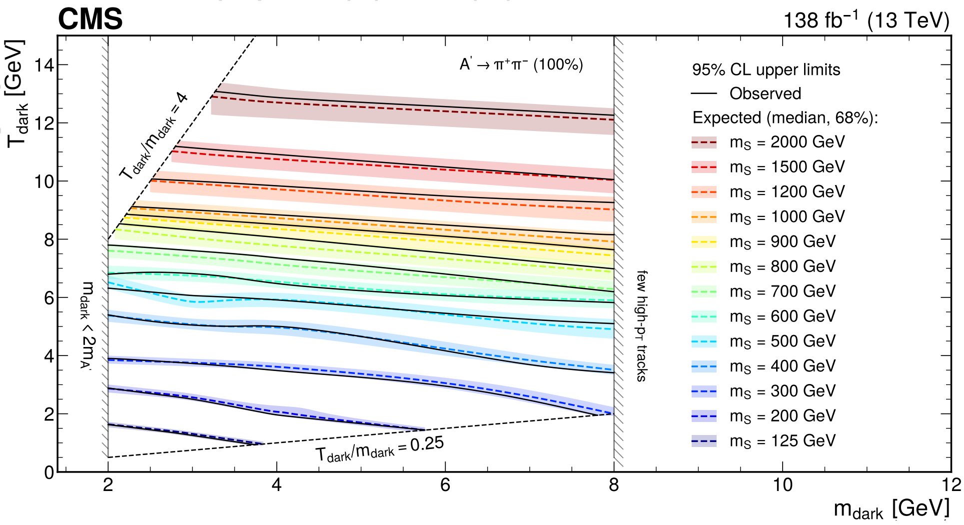

Figure 80: Observed and expected 95% CL excluded regions in the SUEP search (Section 6.2.3.2) in m_{dark}–T_{dark} for each m_S value, considering the case with $m_{A'} = 1.0\,\text{GeV}$ ($A' \to \pi^+\pi^-$ with $\mathcal{B} = 100\%$). The regions below the lines are excluded. Figure taken from Ref. [289].

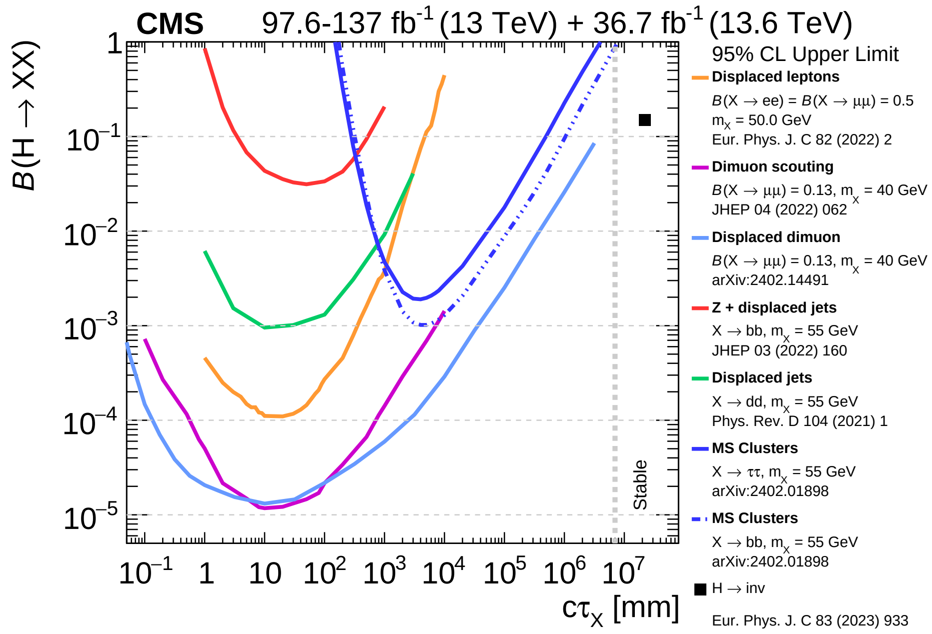

Figure 81: Observed 95% CL upper limits on the branching fraction of Higgs bosons decaying into LLPs with masses between 40 and 55 GeV [85, 187, 243, 292, 294, 301, 323].

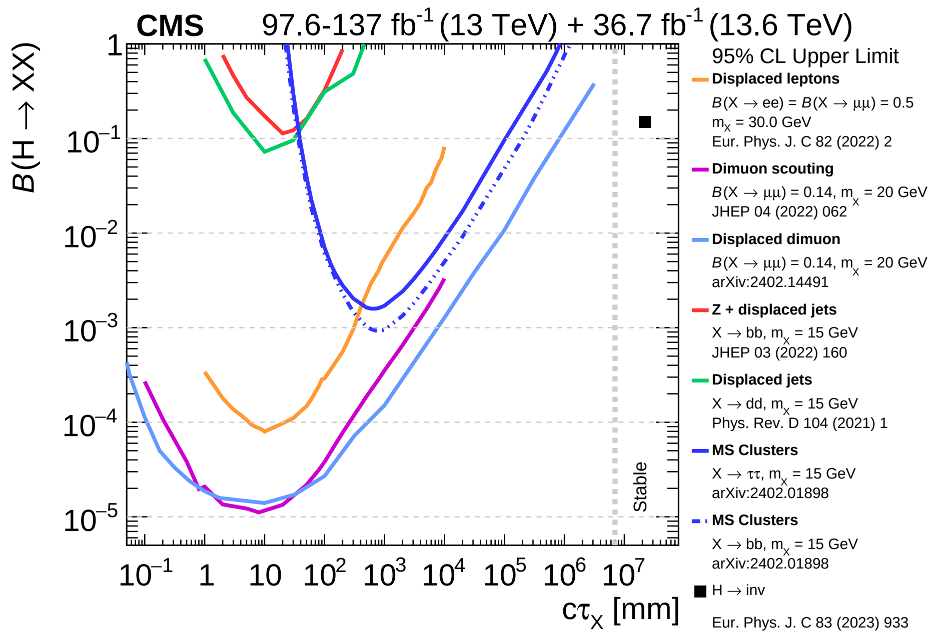

Figure 82: Observed 95% CL upper limits on the branching fraction of Higgs bosons decaying into LLPs with masses between 15 and 30 GeV [85, 187, 243, 292, 294, 301, 323].

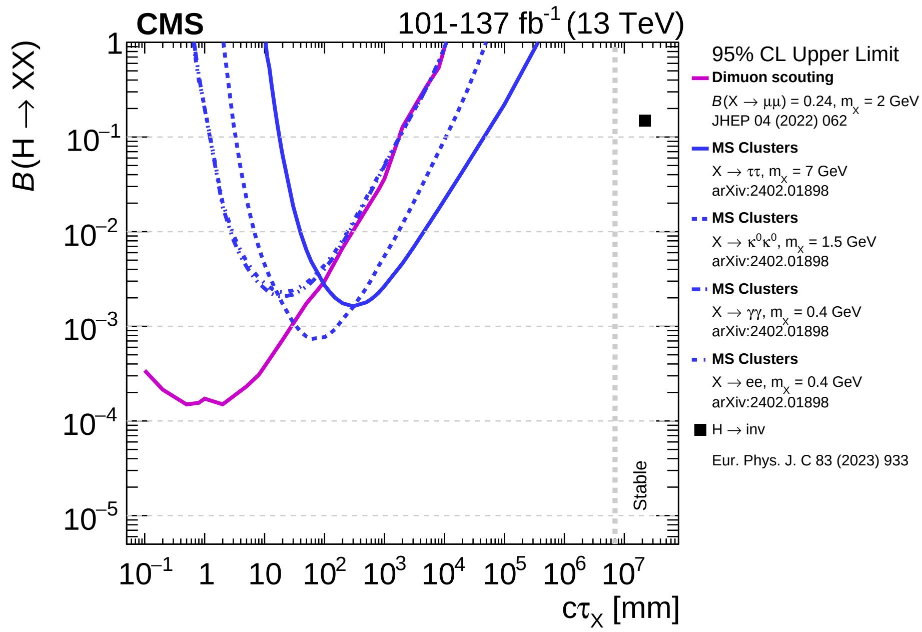

Figure 83: Observed 95% CL upper limits on the branching fraction of Higgs bosons decaying into LLPs with masses between 0.4 and 7 GeV [294, 301].

structed range from a few microns to several meters. In addition, the final state may include significant p_T^{miss} from either decays of LLPs outside of acceptance or from invisible particles produced in the decays.

Given the wide range of potential signatures, multiple search strategies have been employed to provide sensitivity, as detailed in Section 6.3. Many of these searches were originally designed to achieve sensitivity to supersymmetric models or lower energy signatures, such as decays of the SM Higgs boson. However, excellent sensitivity is achieved by these searches for DS models, such as decays of heavy Z′ and heavy H_D bosons to LLPs. Sensitivities for leptonic final states are shown in Refs. [132, 243]. Hadronic final states are considered below. The Z′ model is used to probe the sensitivity to DSs with TeV-scale production of LLPs while the H_D model is used to probe sensitivity to DSs with masses ≈100 GeV.

The exclusion limits for several CMS LLP searches to Z′ bosons decaying into a pair of LLPs are shown in Fig. 84 for Z′ boson masses of 3000 and 4500 GeV. The use of multiple search techniques provides extensive lifetime coverage. The DV search has the best sensitivity for lower lifetimes as it uses the tracker while the calorimetry and muon system based searches have optimal sensitivity for longer lifetimes. To probe spectra with DM candidates, models in which the Z′ boson decays into an LLP and a DM candidate are considered in Fig. 85. As p_T^{miss} is significantly increased, searches using p_T^{miss} show substantially improved reach compared to signal models in which the Z′ boson decays into two LLPs.

The exclusion limits for several CMS LLP searches for H_D decaying into a pair of LLPs are shown in Fig. 86 for H_D masses of 400 and 800 GeV, respectively. The use of multiple search techniques is again shown to provide extensive lifetime coverage. It can also be seen that the large energy thresholds used for the DV search cause a stronger dependence of the sensitivity on the mass of H_D compared to the muon system search. To probe spectra with DM candidates, models in which the H_D decays into an LLP and a DM candidate are considered in Fig. 87. As p_T^{miss} is significantly increased for such signatures, searches using p_T^{miss} show substantially improved reach.

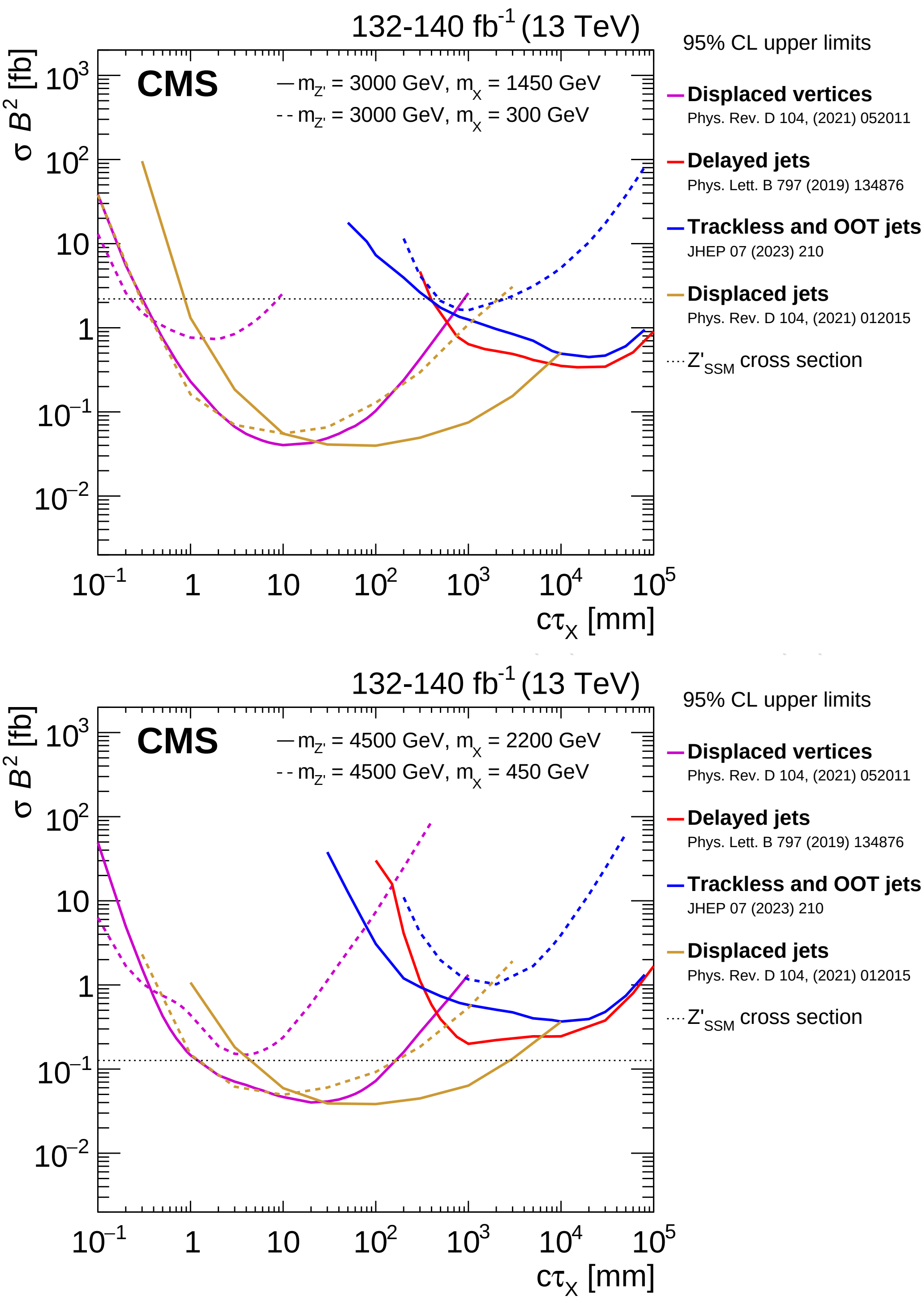

Figure 84: Observed 95% CL exclusion limits for Z' bosons decaying into LLPs with fully hadronic final states, for a Z' boson mass of 3000 GeV (upper) and 4500 GeV (lower). Analyses employing different strategies are shown to have complementary lifetime sensitivity [184, 187, 199, 200]. The theoretical cross section assumes the Z' has SM-like couplings to SM quarks [66].

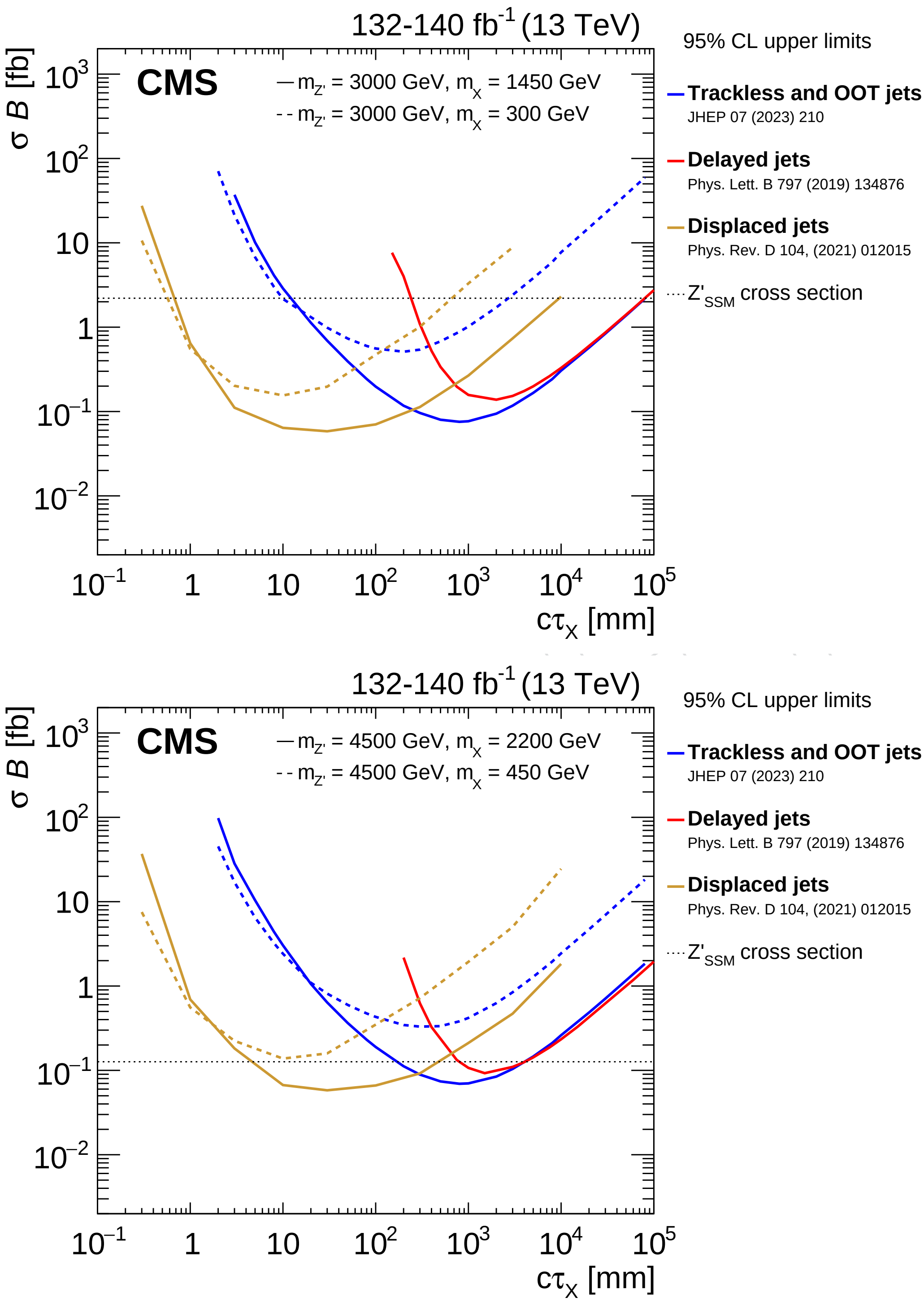

Figure 85: Observed 95% CL exclusion limits for Z′ bosons decaying into LLPs with hadronic plus p_T^{miss} final states, for a Z′ boson mass of 3000 GeV (upper) and 4500 GeV (lower). Analyses employing different strategies are shown to have complementary lifetime sensitivity [187, 199, 200]. The theoretical cross section assumes the Z′ has SM-like couplings to SM quarks [66].

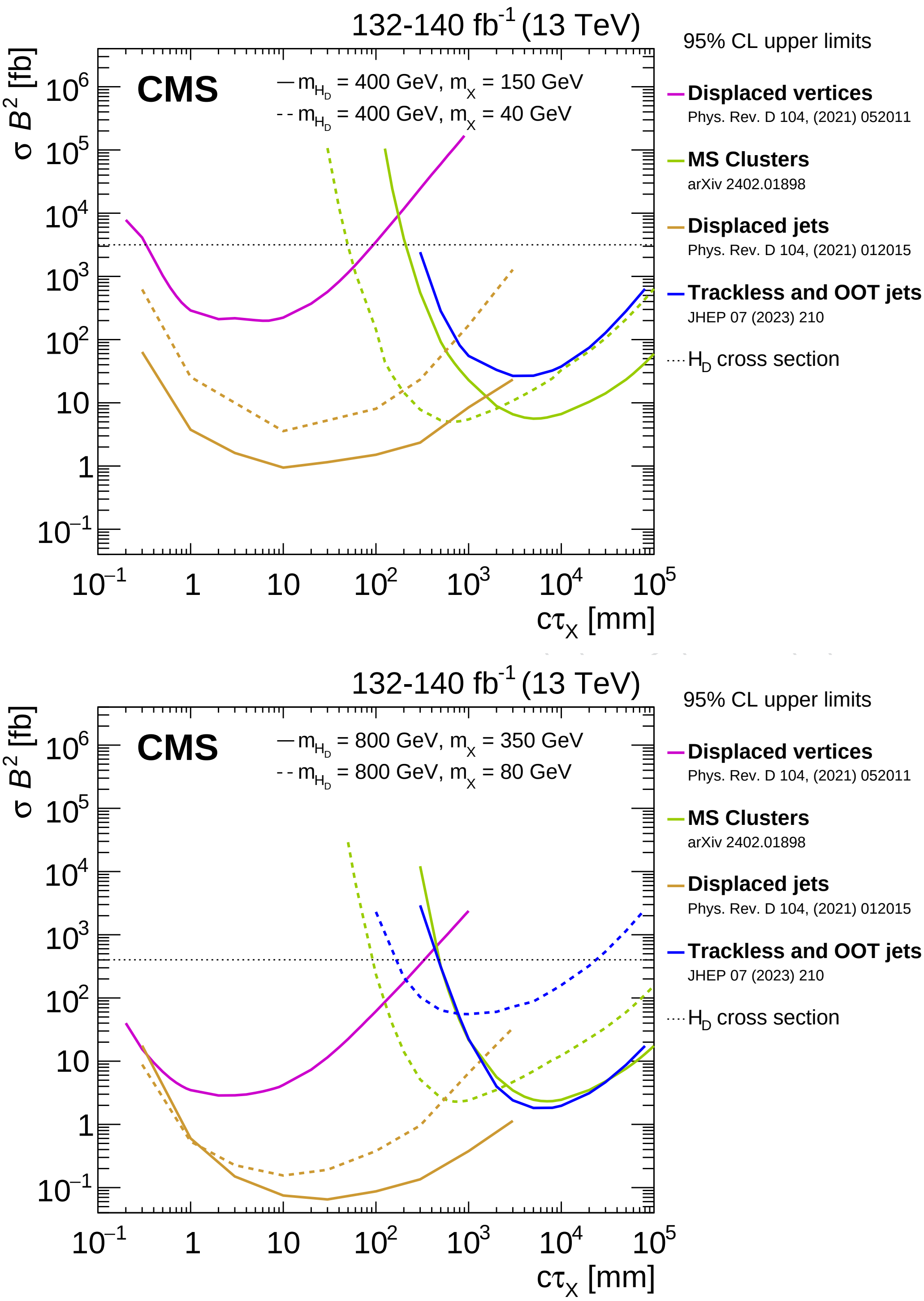

Figure 86: Observed 95% CL exclusion limits for H_D decaying into LLPs with a fully hadronic final state, for a H_D mass of 400 GeV (upper) and 800 GeV (lower). The H_D production cross section assumes point-like effective theory [274]. Analyses employing different strategies are shown to have complementary lifetime sensitivity [184, 187, 200, 301].

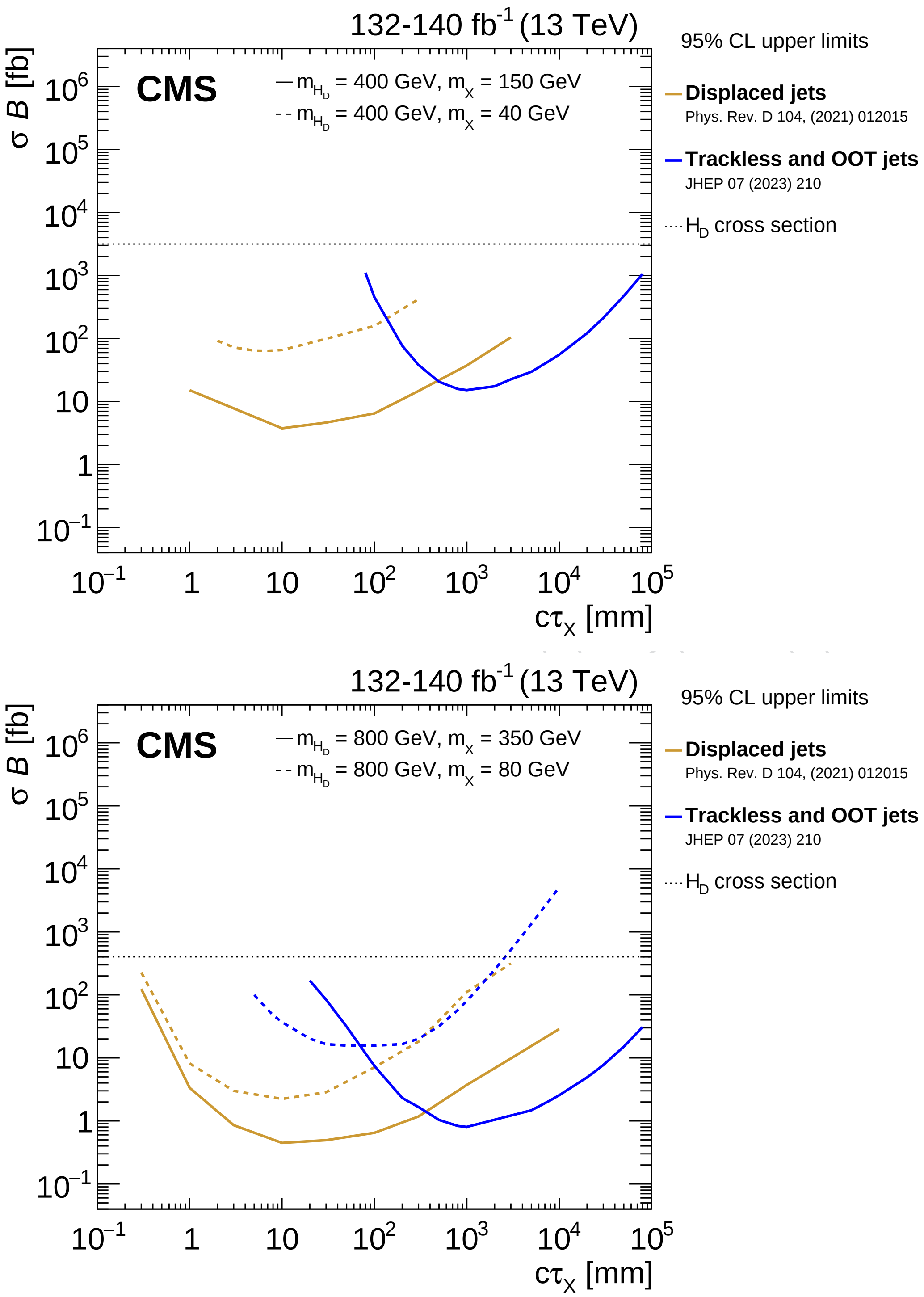

Figure 87: Observed 95% CL exclusion limits for H_D decaying into LLPs with a hadronic plus p_T^{miss} final state, for a H_D mass of 400 GeV (upper) and 800 GeV (lower). The H_D production cross section assumes point-like effective theory [274]. Analyses employing different strategies are shown to have complementary lifetime sensitivity [187, 200].

8 Summary

A comprehensive review of dark sector (DS) searches with the CMS experiment at the LHC has been presented, using proton-proton and heavy ion collision data collected in Run 2, from 2016 to 2018, or, in some cases, from Run 1 (2011–2012) or Run 3 (2022). These searches have been interpreted in simplified and extended DS models. Figure 88 qualitatively illustrates how the results map into this theoretical framework. The broad DS search program spans many different signatures, including those with invisible particles, those with particles promptly decaying into fully visible final states, and those with long-lived particles (LLPs). A number of searches have been newly reinterpreted with DS benchmark scenarios for this Report. In order to perform these searches, several unique techniques of data collection and reconstruction were employed, and they are also described in this Report. The broad variety of searches provides sensitivity across a wide range of models and parameter space, and the results represent the most complete set of constraints on DS models obtained by the CMS Collaboration to date.

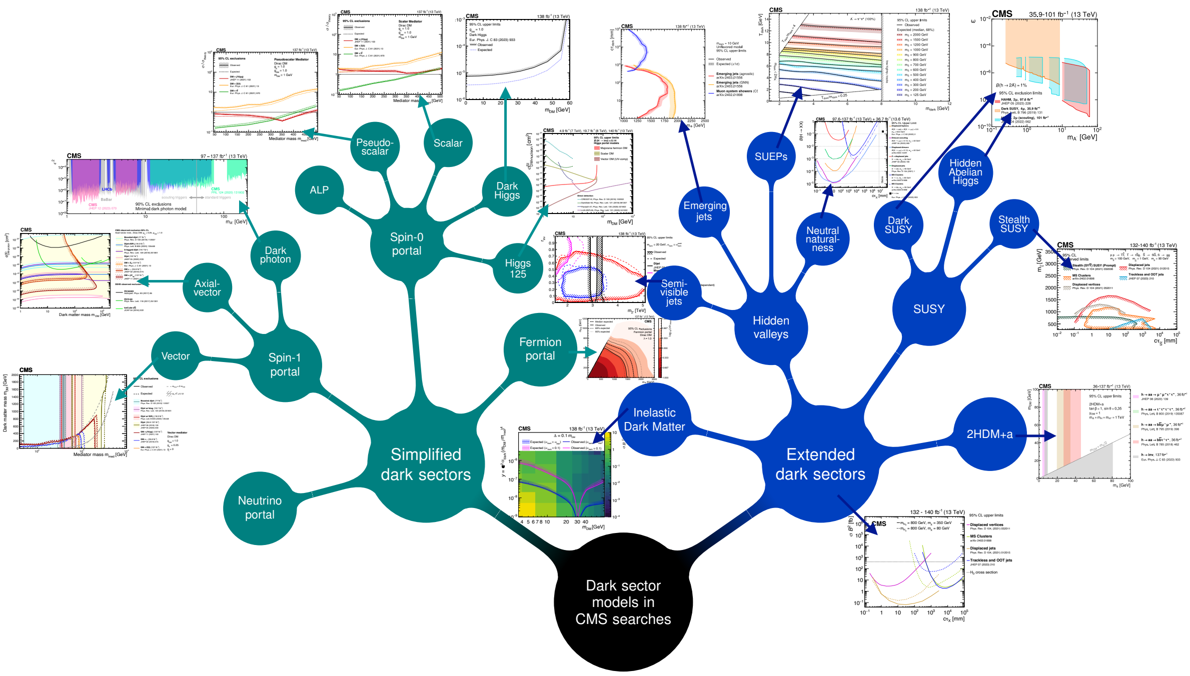

Figure 88: A qualitative depiction of how the results in this Report map onto the models probed in CMS searches for dark sectors.

In particular, this Report has presented the latest constraints from the CMS experiment on a comprehensive set of simplified dark matter models, and it has compared these constraints with those from direct-detection experiments. New reinterpretations have been shown for extended DS scenarios, including semivisible jets, emerging jets, dark supersymmetry, hidden Abelian Higgs models, and two-Higgs-doublet plus a pseudoscalar models. Several scenarios involving LLPs have been presented, including models with heavy LLPs, stealth supersymmetry, and Higgs boson decays to LLPs.

In addition, future improvements will provide increased DS sensitivity. For Run 3 of the LHC [324], new triggers are available [183], as well as improvements to unique data-collection strategies, such as data scouting and data parking [176]. These strategies have already been exploited by some of the searches presented in this Report, and more analyses in the future will also benefit from them.

Finally, the High-Luminosity LHC will provide even further DS sensitivity, owing to both the increased performance of the accelerator and the substantial upgrades of the CMS detector [160, 325–331]. The impressive extension in sensitivity that will be achieved for DS models has been shown in several studies of the physics performance at the High-Luminosity LHC [332–334].

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid and other centers for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC, the CMS detector, and the supporting computing infrastructure provided by the following funding agencies: the Armenian Science Committee, project no. 22rl-037; the Austrian Federal Ministry of Education, Science and Research and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP); the Bulgarian Ministry of Education and Science, and the Bulgarian National Science Fund; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, the National Natural Science Foundation of China, and Fundamental Research Funds for the Central Universities; the Ministerio de Ciencia Tecnología e Innovación (MINCIENCIAS), Colombia; the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research and Innovation Foundation, Cyprus; the Secretariat for Higher Education, Science, Technology and Innovation, Ecuador; the Estonian Research Council via PRG780, PRG803, RVTT3 and the Ministry of Education and Research TK202; the Academy of Finland, Finnish Ministry of Education and Culture, and Helsinki Institute of Physics; the Institut National de Physique Nucléaire et de Physique des Particules / CNRS, and Commissariat à l'Énergie Atomique et aux Énergies Alternatives / CEA, France; the Shota Rustaveli National Science Foundation, Georgia; the Bundesministerium für Bildung und Forschung, the Deutsche Forschungsgemeinschaft (DFG), under Germany's Excellence Strategy – EXC 2121 "Quantum Universe" – 390833306, and under project number 400140256 - GRK2497, and Helmholtz-Gemeinschaft Deutscher Forschungszentren, Germany; the General Secretariat for Research and Innovation and the Hellenic Foundation for Research and Innovation (HFRI), Project Number 2288, Greece; the National Research, Development and Innovation Office (NKFIH), Hungary; the Department of Atomic Energy and the Department of Science and Technology, India; the Institute for Studies in Theoretical Physics and Mathematics, Iran; the Science Foundation, Ireland; the Istituto Nazionale di Fisica Nucleare, Italy; the Ministry of Science, ICT and Future Planning, and National Research Foundation (NRF), Republic of Korea; the Ministry of Education and Science of the Republic of Latvia; the Research Council of Lithuania, agreement No. VS-19 (LMTLT); the Ministry of Education, and University of Malaya (Malaysia); the Ministry of Science of Montenegro; the Mexican Funding Agencies (BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI); the Ministry of Business, Innovation and Employment, New Zealand; the Pakistan Atomic Energy Commission; the Ministry of Education and Science and the National Science Center, Poland; the Fundação para a Ciência e a Tecnologia, grants CERN/FIS-PAR/0025/2019 and CERN/FIS-INS/0032/2019, Portugal; the Ministry of Education, Science and Technological Development of Serbia; MCIN/AEI/10.13039/501100011033, ERDF "a way of making Europe", Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia María de Maeztu, grant MDM-2017-0765, projects PID2020-113705RB, PID2020-113304RB, PID2020-116262RB and PID2020-113341RB-I00, and Plan de Ciencia, Tecnología e

A Glossary of acronyms

ALP	Axion-like particle
BDT	Boosted decision tree
BPTX	Beam pickup timing device
BSM	Beyond the standard model
CA	Cambridge–Aachen
CEP	Central exclusive production
CHS	Charged-hadron subtraction
CL	Confidence level
CMS	Compact Muon Solenoid
CP	Charge conjugation parity
CSC	Cathode strip chamber
CR	Control region
DA	Domain adaptation
DM	Dark matter
DD	Direct detection
DNN	Deep neural network
DS	Dark sector
DSA	Displaced standalone
DT	Drift tube
DV	Displaced vertex
ECAL	Electromagnetic calorimeter
EFT	Effective field theory
EJ	Emerging jet
EW	Electroweak
FCNC	Flavor-changing neutral currents
FIP	Feebly interacting particle
GBDT	Gradient-boosted decision tree
GMSB	Gauge-mediated SUSY breaking
GNN	Graph neural network
HAHM	Hidden Abelian Higgs model
HCAL	Hadronic calorimeter
HI	Heavy ion
HLT	High level trigger
HNL	Heavy neutral lepton
HV	Hidden valley
ID	Indirect detection
IDM	Inelastic dark matter
IP	Impact parameter
ISR	Initial-state radiation
LHC	Large Hadron Collider
LLP	Long-lived particle
LO	Leading order
MC	Monte Carlo
MVA	Multi-variate analysis
NLO	Next-to-leading order
NNLL	Next-to-next-to-leading logarithm
NNLO	Next-to-next-to-leading order
OOT	Out of time

PDF	Parton distribution function
PF	Particle flow
PPS	Precision proton spectrometer
PU	Pileup
PUPPI	Pileup-per-particle identification
PV	Primary vertex
QCD	Quantum chromodynamics
ROC	Receiver operating characteristic
RPC	Resistive-plate chamber
RPV	*R*-parity violating
SD	Spin dependent
SI	Spin independent
SM	Standard model
SR	Signal region
SUEP	Soft unclustered energy patterns
SUSY	Supersymmetry
SV	Secondary vertex
SVJ	Semivisible jet
TF	Transfer factor
TMS	Tracker and muon spectrometer
UPC	Ultra-peripheral collision
VBF	Vector-boson fusion
WIMP	Weakly interacting massive particle
2D	Two-dimensional
3D	Three-dimensional
2HDM	Two-Higgs-doublet model
2HDM+a	Two-Higgs-doublet model plus pseudoscalar