

DRAFT

CMS Paper

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

2025/01/27

Archive Hash: 32611a6

Archive Date: 2023/09/28

Search for Dark Matter Produced in Association with a Resonant Bottom-Quark Pair in Proton-Proton Collisions at $\sqrt{s} = 13 \text{ TeV}$

The CMS Collaboration

Abstract

A search for dark matter (DM) produced in association with a resonant $b\bar{b}$ pair is performed in proton-proton collisions at a center-of-mass energy of 13 TeV collected with the CMS detector during the 2016–2018 run of the Large Hadron Collider. The analyzed data sample corresponds to an integrated luminosity of 137 fb^{-1} . Results are interpreted in terms of a novel theoretical model of DM production at the LHC that predicts the presence of a Higgs-boson-like particle in the dark sector, motivated simultaneously by the need to generate the masses of the particles in the dark sector and the possibility to relax constraints from the DM relic abundance by opening up a new annihilation channel. If such a dark Higgs boson decays into standard model (SM) states via a small mixing with the SM Higgs boson, one obtains characteristic large-radius jets in association with missing transverse momentum that can be used to efficiently discriminate signal from backgrounds. Limits on the signal strength of different dark Higgs boson mass hypotheses below 160 GeV are set for the first time with CMS data.

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

PDFAuthor:	M. Cremonesi, A. Das, M. Donega, S. Eisenberger, E. Ertorer, A. Hall, M. Hildreth, B. Jayatilaka, J. Lee, N. Macilla, M. Marchegiani, C.-S. Moon, I. Pedraza, N. Smith, T. Tomei, D. Valsecchi, R. Wallny, M. Wassmer1, Z. Ye
PDFTitle:	Search for Dark Matter Produced in Association with a Resonant Bottom-Quark Pair
PDFSubject:	CMS
PDFKeywords:	CMS, physics, software, computing

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

1 Introduction

It is well established from astrophysical observations that most of the matter in the Universe is comprised of dark matter (DM) that cannot be accommodated within the standard model (SM) [1]. One of the leading hypotheses is that DM belongs to a dark sector of particles similar in structure (and possibly complexity) to that of ordinary matter. In contrast to the traditional weakly interacting massive particle (WIMP) paradigm, this dark sector is posited to be neutral under the SM forces. Similarly, all SM particles are neutral under the dark-sector forces [2]. In this hypothesis, DM is a Majorana particle, and is composed of stable, electrically neutral, massive particles which interact with baryons at least via the gravitational force. If such a DM particle also interacts non-gravitationally with SM particles, then DM could be produced in proton-proton collisions at the CERN LHC.

Since DM particles, once produced, do not leave any detectable signal in the detector, they cannot be directly observed. However, their presence can be inferred if they are produced in association with a visible object, such as a lepton, photon or hadronic jet. Such processes generate final states commonly referred to as “mono-X”, where X denotes the visible object. In the context of the LHC searches, the most sensitive analysis channel to DM production is the one that searches for DM produced in association with a gluon or a quark emitted as initial state radiation (ISR), known as the monojet channel.

The monojet DM searches at the LHC [3, 4] have strongly constrained the parameter space in which DM particles can obtain their relic abundance from direct annihilation into SM final states. This tension is relaxed if DM particles are not the lightest state particles in the dark sector, leading to new annihilation channels. If the DM mass is generated via a Higgs mechanism in the dark sector and the resulting dark Higgs boson is lighter than the DM, a new annihilation channel where DM particles annihilate into a pair of dark Higgs bosons (H_D), with subsequent decay into SM states, would be possible. This configuration would allow the model to easily match the observed relic abundance [5–7]. Assuming a small mixing with the SM Higgs boson, a light dark Higgs boson decays with the same branching fraction of the SM Higgs boson, which vary depending on its mass. If the dark sector includes an additional spin-1 Z' mediator, then the probability of the Z' boson being produced and radiating a dark Higgs boson (dark-Higgsstrahlung) can be large [8]. A representative Feynman diagram for this signal model is shown in Fig. 1.

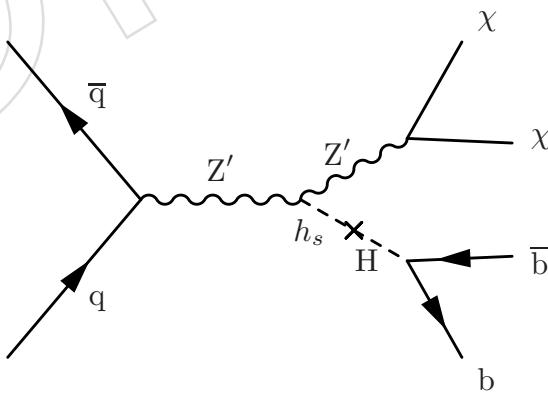


Figure 1: Feynman diagram for the production of a dark Higgs boson (H_D) in association with DM particles (χ). The interaction is mediated by a spin-1 Z' boson, and the dark Higgs boson mixes with the SM Higgs boson (H) through the θ_h mixing angle angle. In this paper we consider low-mass scenarios for the H_D particle, giving rise to a resonant $H_D \rightarrow b\bar{b}$ signal.

Searches for dark Higgs boson production in association with DM particles have already been

33 performed by the ATLAS [9] and CMS [10] collaborations. These searches focus on heavier
 34 dark Higgs boson mass hypotheses, larger than 160 GeV. For dark Higgs bosons of such mass
 35 the decay into a pair of W bosons is the dominant one. In this paper, we present the search for
 36 the associated production of DM and a light dark Higgs boson (with mass less than 160 GeV),
 37 which decays predominantly into a pair of b quarks, performed for the first time at the LHC.
 38 This final state results in a distinctive signature of large missing transverse momentum, arising
 39 from the decay of the Z' mediator into DM, and a highly-boosted large-radius jet, originated
 40 by the hadronization of two b quarks from the dark Higgs boson decay. Results are presented
 41 for the full dataset of 137 fb^{-1} collected by the CMS experiment at a center-of-mass energy of
 42 13 TeV during Run 2 of the LHC.

43 2 The CMS detector and event reconstruction

44 The CMS apparatus [11, 12] is a multipurpose, nearly hermetic detector, designed to trigger
 45 on [13–15] and identify electrons, muons, photons, and (charged and neutral) hadrons [16–18].
 46 Its central feature is a superconducting solenoid of 6 m internal diameter, providing a mag-
 47 netic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead
 48 tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron cal-
 49 orimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters
 50 extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are
 51 reconstructed using gas-ionization detectors embedded in the steel flux-return yoke outside
 52 the solenoid. More detailed descriptions of the CMS detector, together with a definition of the
 53 coordinate system used and the relevant kinematic variables, can be found in Refs. [11, 12].

54 The silicon tracker used in 2016 measured charged particles within the range $|\eta| < 2.5$. For
 55 nonisolated particles of $1 < p_T < 10\text{ GeV}$ and $|\eta| < 1.4$, the track resolutions were typically
 56 1.5% in p_T and 25–90 (45–150) μm in the transverse (longitudinal) impact parameter [18]. At the
 57 start of 2017, a new pixel detector was installed [19]; the upgraded tracker measured particles
 58 up to $|\eta| = 3.0$ with typical resolutions of 1.5% in p_T and 20–75 μm in the transverse impact
 59 parameter [20] for nonisolated particles of $1 < p_T < 10\text{ GeV}$. According to simulation stud-
 60 ies [21], similar improvements are expected in the longitudinal direction. The primary vertex
 61 (PV) is taken to be the vertex corresponding to the hardest scattering in the event, evaluated
 62 using tracking information alone, as described in Section 9.4.1 of Ref. [22].

63 In the region $|\eta| < 1.74$, the HCAL cells have widths of 0.087 in pseudorapidity and 0.087
 64 in azimuth (ϕ). In the η - ϕ plane, and for $|\eta| < 1.48$, the HCAL cells map on to 5×5 arrays
 65 of ECAL crystals to form calorimeter towers projecting radially outwards from close to the
 66 nominal interaction point. For $|\eta| > 1.74$, the coverage of the towers increases progressively
 67 to a maximum of 0.174 in $\Delta\eta$ and $\Delta\phi$. The forward hadron (HF) calorimeter uses steel as an
 68 absorber and quartz fibers as the sensitive material. The two halves of the HF are located 11.2 m
 69 from the interaction region, one on each end, and together they provide coverage in the range
 70 $3.0 < |\eta| < 5.2$. They also serve as luminosity monitors.

71 Events of interest are selected using a two-tiered trigger system. The first level (L1), composed
 72 of custom hardware processors, uses information from the calorimeters and muon detectors to
 73 select events at a rate of around 100 kHz within a fixed latency of $4\text{ }\mu\text{s}$ [13]. The second level,
 74 known as the high-level trigger (HLT), consists of a farm of processors running a version of the
 75 full event reconstruction software optimized for fast processing, and reduces the event rate to
 76 a few kHz before data storage [14, 15].

77 A particle-flow algorithm [23] aims to reconstruct and identify each individual particle in an

78 event, with an optimized combination of information from the various elements of the CMS
 79 detector. In this process, the identification of the PF candidate type (photon, electron, muon,
 80 and charged and neutral hadrons) plays an important role in the determination of the particle
 81 direction and energy. The energy of photons is obtained from the ECAL measurement. The
 82 energy of electrons is determined from a combination of the electron momentum at the primary
 83 interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster,
 84 and the energy sum of all bremsstrahlung photons spatially compatible with originating from
 85 the electron track. The energy of muons is obtained from the curvature of the corresponding
 86 track. The energy of charged hadrons is determined from a combination of their momentum
 87 measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for
 88 the response function of the calorimeters to hadronic showers. Finally, the energy of neutral
 89 hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

90 For each event, hadronic jets are clustered from the PF candidates using the infrared and
 91 collinear safe anti- k_T algorithm [24, 25] with a distance parameter of 0.4 or 1.5. Depending
 92 on the respective distance parameter, these jets are referred to as “AK4” or “AK15” jets. Jet
 93 momentum is determined as the vectorial sum of all particle momenta in the jet, and is found
 94 from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_T
 95 spectrum and detector acceptance. Additional proton-proton interactions within the same or
 96 nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depo-
 97 sitions to the jet momentum. To mitigate this effect in the AK4 jets, charged particles identified
 98 to be originating from pileup vertices are discarded and an offset correction is applied to cor-
 99 rect for remaining contributions. Jet energy corrections are derived from simulation to bring
 100 the measured response of jets to that of particle level jets on average. In situ measurements of
 101 the momentum balance in dijet, photon + jet, Z + jet, and multijet events are used to account
 102 for any residual differences in the jet energy scale between data and simulation [26]. The jet
 103 energy resolution amounts typically to 15–20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV [26].
 104 Additional selection criteria are applied to each jet to remove jets potentially dominated by
 105 anomalous contributions from various subdetector components or reconstruction failures.

106 The missing transverse momentum vector \vec{p}_T^{miss} is computed as the negative vector sum of
 107 the transverse momenta of all the PF candidates in an event, and its magnitude is denoted
 108 as p_T^{miss} [27]. The \vec{p}_T^{miss} is modified to account for corrections to the energy scale of the recon-
 109 structed jets in the event. Anomalous high- p_T^{miss} events can be due to a variety of reconstruc-
 110 tions, detector malfunctions or noncollision backgrounds. Such events are rejected by event
 111 filters that are designed to identify more than 85–90% of the spurious high- p_T^{miss} events with a
 112 mistagging rate less than 0.1% [27].

113 Large-radius AK15 jets are used for the identification of the decays of the dark Higgs boson
 114 into a b-quark pair. The pileup per particle identification algorithm (PUPPI) [28, 29] is used
 115 to mitigate the effect of pileup at the reconstructed particle level, making use of local shape
 116 information, event pileup properties, and tracking information. A local shape variable is de-
 117 fined, which distinguishes between collinear and soft diffuse distributions of other particles
 118 surrounding the particle under consideration. The former is attributed to particles originating
 119 from the hard scatter and the latter to particles originating from pileup interactions. Charged
 120 particles identified to be originating from pileup vertices are discarded. For each neutral par-
 121 ticle, a local shape variable is computed using the surrounding charged particles compatible
 122 with the primary vertex within the tracker acceptance ($|\eta| < 2.5$), and using both charged and
 123 neutral particles in the region outside of the tracker coverage. The momenta of the neutral par-
 124 ticles are then rescaled according to their probability to originate from the primary interaction
 125 vertex deduced from the local shape variable, superseding the need for jet-based pileup correc-

126 tions [28]. The modified mass drop tagger algorithm [30, 31], also known as the soft-drop (SD)
127 algorithm, with the angular exponent $\beta = 0$, soft cutoff threshold $z_{\text{cut}} < 0.1$, and characteristic
128 radius $R_0 = 1.5$ [32], is applied to remove soft, wide-angle radiation from the jet.

129 3 Simulated samples

130 Monte Carlo (MC) simulated event samples are used to model signal and background contribu-
131 tions to all the analysis regions. In all cases, parton showering, hadronization, and underlying
132 event properties are modeled using PYTHIA [33] version 8.202 or later with the underlying
133 event tune CUETP8M1 or CP5 [34], based on the year of data taking. Simulation of inter-
134 actions between particles and the CMS detector is based on GEANT4 [35]. The NNPDF 3.0
135 next-to-next-to-leading order (NNLO) [36] and the NNPDF 3.1 NNLO [37] parton distribution
136 functions (PDFs) are used for the generation of all samples based on the year of data taking.
137 The same reconstruction algorithms used for data are applied to simulated samples.

138 For the V + jets processes, predictions with up to two partons in the final state are obtained
139 at leading order (LO) in QCD using MADGRAPH5_aMC@NLO [38] with the MLM matching
140 scheme [39] between the jets from the matrix element calculations and the parton shower. Sam-
141 ples of events with top quark pairs are generated at next-to-leading (NLO) in QCD with up to
142 two additional partons in the matrix element calculations using MADGRAPH5_aMC@NLO and
143 the FxFx jet matching scheme [40]. Their cross sections are normalized to the inclusive cross
144 section of the top quark pair production at NNLO in QCD [41]. Events with single top quarks
145 are simulated using POWHEG 2.0 [42, 43] and normalized to the inclusive cross section calcu-
146 lated at NLO in QCD [44, 45]. Production of diboson events (WW, WZ, and ZZ) is simulated
147 at NLO in QCD using PYTHIA, and normalized to the cross sections at NNLO precision for
148 WW production [46] and at NLO precision for the others [47]. Several production mechanisms
149 of the SM Higgs boson decaying into a bottom-quark pair are also produced at LO with the
150 POWHEG generator. Samples of QCD multijet production events are generated at LO using
151 MADGRAPH5_aMC@NLO.

152 The simulated samples for the dark Higgs boson signal process are generated with the
153 MADGRAPH5_aMC@NLO with up to one additional parton in the matrix element calculations
154 at LO with the MLM matching scheme. In this scenario, the dark-matter (DM) particle χ is a
155 Majorana fermion that couples *axially* to the new gauge boson Z' , while the coupling to Stan-
156 dard Model (SM) quarks is vector-like. Accordingly, the relevant part of the spin-1 sector of the
157 model Lagrangian is:

$$\mathcal{L}_{\text{spin-1}} \supset -g_{\text{DM}} Z'_\mu \bar{\chi} \gamma^\mu \gamma^5 \chi - g_q Z'_\mu \sum_q \bar{q} \gamma^\mu q, \quad (1)$$

158 where g_{DM} and g_q are the couplings to the Majorana DM and to the SM quarks, respectively.

159 Following the recommendations of the LHC Dark Matter Working Group [48], we set $g_q = 0.25$
160 and $g_{\text{DM}} = 1.0$ in our model. Separate samples are generated for different mass hypotheses for
161 the mediator, DM particles, and dark Higgs boson. In all samples we set $\theta_h = 0.01$, a value that
162 is large enough to ensure prompt decay of the dark Higgs boson while being small enough to
163 have no observable effect on the the SM-like Higgs boson couplings [49].

164 4 Event selection

165 Signal region (or “SR”) events are selected using triggers that require large p_T^{miss} (at least 120 GeV)
166 and large H_T^{miss} online. They contain events with a large-radius, high- p_T jet and large p_T^{miss} . The

key feature of the analysis is the extensive use of control data samples for the purpose of precise prediction of the background contributions in the signal regions (SRs). The leading SM background contributions originate from $Z \rightarrow \nu\nu$ and $W \rightarrow \ell\nu$ production ($\ell = e, \mu, \tau$), the properties of which are constrained using control regions (CRs) with zero or one charged lepton, that are enriched in $Z \rightarrow \nu\nu$ and $W \rightarrow \ell\nu$ events, respectively. Additionally, CRs enriched in $t\bar{t}$ production events are also defined. The V + jets and $t\bar{t}$ production events in these CRs share many kinematic properties of the processes in the SRs and are used to constrain the latter. The CR and SR definitions share as many of the selection criteria as possible, in order to ensure that minimal selection biases are introduced. Seven CRs are defined: six single-electron and single-muon CRs enriched in $W \rightarrow \ell\nu$ and $t\bar{t}$ production events, and a seventh CR enriched in $Z \rightarrow \nu\nu$ production events.

The trigger requirement for the SRs is based on an online calculation of p_T^{miss} based on all PF candidates reconstructed at the HLT, except for muons. Events with high- p_T muons are therefore also assigned large online p_T^{miss} , and the same trigger is used to collect data populating the single-muon and CRs. The control samples with electrons are selected using two single-electron triggers: one requiring $p_T > 27$ (2016), 35 (2017), 32 (2018) GeV, while the other requiring $p_T > 105$ (2016), 115 (2017–2018) GeV. Additionally, a single-photon trigger with $p_T > 200$ GeV is used in 2017 and 2018. The single-electron triggers differ in their usage of isolation requirements: while the lower threshold trigger requires electrons to be well isolated, the higher-threshold trigger does not, which gives an improved efficiency at high p_T . Similarly, the single-photon trigger avoids the reliance on the online track reconstruction and increases the overall efficiency for electrons with $p_T > 200$ GeV. During the 2016 and 2017 data taking, a gradual shift in the timing of the inputs of the ECAL L1 trigger in the region at $|\eta| > 2.0$ caused a specific trigger inefficiency. For events containing an electron or a photon (a jet) with $p_T \gtrsim 50$ (100) GeV in this region, the efficiency loss is up to $\approx 10\text{--}20\%$, depending on p_T , η , and time. This issue is known as the L1 pre-firing. Correction factors are computed from data and applied to the acceptance evaluated by simulation for the 2016 and 2017 samples.

At the analysis level, a requirement of $p_T^{\text{miss}} > 250$ GeV is applied to the SR events in order to ensure a p_T^{miss} trigger efficiency of at least 95%. The leading AK15 jet in p_T is required to have $p_T > 160$ GeV, $|\eta| < 2.4$, and SD-corrected mass (m_{SD}) of $40 < m_{\text{SD}} < 300$ GeV. In order to preferentially select events where the leading AK15 jet originates from a hadronic decay of a dark Higgs boson, the jet is further required to be double-b tagged with the DEEPAK15 algorithm [50]. The DEEPAK15 algorithm employs a deep neural network to differentiate between jets from vector boson, top quark, and Higgs boson decays, as well as jets originating from QCD radiation. The inputs to the neural network are features of up to 100 jet constituent PF candidates of a given jet and features related to up to seven secondary vertices reconstructed in a given collision event. For each jet, the output of the neural network is one numerical score for each of the jet classes, representing the likelihood that the jet originates from that class. In this analysis, separation between dark Higgs bosons and QCD jets is sought, and a binary score is constructed by taking the ratio of the sum of the SM $Z \rightarrow bb$ and $H \rightarrow bb$ scores to the sum of the SM $Z \rightarrow bb$, $H \rightarrow bb$, and QCD scores.

Further requirements are imposed in order to suppress reducible background processes. Events are rejected if they contain a well-reconstructed and isolated electron (photon) with $p_T > 10$ (15) GeV and $|\eta| < 2.5$ or a muon with $p_T > 10$ GeV and $|\eta| < 2.4$ [16, 51]. Hadronically decaying τ leptons are identified using the DEEPTAU algorithm [52]. Events with a hadronically decaying τ lepton candidate with $p_T > 20$ GeV and $|\eta| < 2.3$ are removed. These requirements efficiently reject events with leptonic decays of the V bosons and top quarks, as well as backgrounds with photons. Contributions from top quark processes are further suppressed by

215 rejecting events with AK4 jets that do not overlap with the leading AK15 jet, have $p_T > 20 \text{ GeV}$
 216 and $|\eta| < 2.4$, and are identified to have originated from the hadronization of a bottom quark
 217 (“b-tagged jets”). The b-tagging of AK4 jets is performed using the DEEPJET algorithm [53]
 218 with a “loose” working point, corresponding to a b-tagging efficiency of 93% with a probabili-
 219 ty of 10% of misidentifying a light-flavor quark or gluon jet. Finally, topological requirements
 220 are applied in order to reject contributions from QCD multijet events. These events do not
 221 have p_T^{miss} from genuine sources and require a p_T^{miss} mismeasurement in order to pass the SR
 222 selections, which can happen in two main ways. In the first case, the energy of a jet in the
 223 event could be misreconstructed either as a result of an interaction between the jet with poorly
 224 instrumented or inactive parts of the detector, or because of failures in the readout of other-
 225 wise functioning detector modules. In these cases, artificial p_T^{miss} is generated with a charac-
 226 teristically small azimuthal angle difference between the misreconstructed jet \vec{p}_T and the \vec{p}_T^{miss}
 227 vectors. Such events are rejected by requiring the minimum azimuthal angle between the \vec{p}_T^{miss}
 228 direction and each AK4 jet in the event to be larger than 0.5 radians. With the same goal, the
 229 azimuthal angle between the \vec{p}_T^{miss} direction and each AK15 jet in the event must be larger than
 230 1.5 radians. In the second case, large p_T^{miss} is generated due to failures of the PF reconstruction,
 231 which are suppressed by considering an alternative calculation of p_T^{miss} based on calorimeter
 232 energy clusters and muon candidates, rather than the full set of all PF candidates. While the
 233 calorimeter-based p_T^{miss} has significantly worse resolution than PF p_T^{miss} , it is much simpler and
 234 more robust. To reduce the multijet background caused by PF reconstruction failures, events
 235 are required to have $\Delta p_T^{\text{miss}}(\text{PF-calorimeter}) = |p_T^{\text{miss}}(\text{PF})/p_T^{\text{miss}}(\text{calorimeter}) - 1| < 0.5$. Fi-
 236 nally, a section of the HCAL was not functioning during a part of the 2018 data taking pe-
 237 riod corresponding to 65% of the total integrated luminosity recorded in that year, leading
 238 to irrecoverable mismeasurement in a localized region of the detector ($-1.57 < \phi < -0.87$,
 239 $-3.0 < \eta < -1.3$). To avoid contamination from such mismeasurement, events where any
 240 jet with $p_T > 30 \text{ GeV}$ is found in the corresponding η - ϕ region are rejected in the analysis of
 241 the 2018 data set. Events where the mismeasurement is so severe that a jet is fully lost in this
 242 region are found to contribute at low values of $p_T^{\text{miss}} < 470 \text{ GeV}$ and to have a characteristic
 243 signature in $\phi(\vec{p}_T^{\text{miss}})$. Such events are rejected by requiring that $\phi(\vec{p}_T^{\text{miss}}) \notin [-1.62, -0.62]$ if
 244 $p_T^{\text{miss}} < 470 \text{ GeV}$. Expected yields from different processes in SR are reported in Table 1.

245 A control region (labelled as “ZCR”) composed of those events that satisfy all the SR require-
 246 ments, but have the leading AK15 jet failing the DEEPAK15 selection, is used to constrain
 247 $Z(\nu\nu) + \text{jets}$ production in SR. Single-lepton CRs are used to constrain $W(\ell\nu) + \text{jets}$ events in
 248 SR. These CRs are labeled “WEPCR”, which stands for W+jets single Electron Pass Control
 249 Region, and “WMPCR”, which stands for W+jets single Muon Pass Control Region. The same
 250 selection criteria are applied to these CRs as for the SR, with the exception of the charged-lepton
 251 rejection criteria being inverted to allow for exactly one muon or one electron. The \vec{p}_T^{miss} vec-
 252 tor used in the SR definition is replaced by the hadronic recoil vector \vec{U} . The hadronic recoil
 253 is defined as the vectorial sum of the \vec{p}_T^{miss} vector and the transverse momentum vectors of
 254 the selected charged lepton in each event. The hadronic recoil therefore acts as a proxy of the
 255 momentum of the W boson in each CR, convolved with the p_T^{miss} resolution, which is equiva-
 256 lent to the role of p_T^{miss} in the SR. In order to enhance the purity of the CRs, specific additional
 257 selection criteria are applied. For the charged-lepton CRs, at least one of the leptons is re-
 258 quired to pass a more strict set of quality criteria and have $p_T > 40$ (20) GeV electrons (muons).
 259 Additionally, in the single-electron CR are required to have $p_T^{\text{miss}} > 100 \text{ GeV}$ in order to re-
 260 ject contributions from QCD multijet events. Additional single-lepton CRs composed of those
 261 single-lepton events with the leading AK15 jet failing the DEEPAK15 selection are also used to
 262 constrain $W(\ell\nu) + \text{jets}$ events in SR. These CRs are labeled “WEFCR”, which stands for W+jets
 263 single electron Fail Control Region, and “WMFCR”, which stands for W+jets single Muon Fail

264 Control Region. Finally, in order to select single-lepton events enriched in $t\bar{t}$ production, additional
 265 CRs (labelled as “TECR” and “TMCR”) are identified by inverting the veto on b-tagged
 266 AK4 jets outside the leading AK15 jet cone.

267 The selection criteria that define the different SR and CRs are summarized in Table 2.

268 5 Background estimation

269 Background estimation and signal extraction are performed simultaneously, using a joint maximum
 270 likelihood (ML) fit across all SR and CRs. A likelihood function is constructed to model
 271 the expected background contributions in each bin of the two-dimensional recoil-vs- m_{SD} variable
 272 of the SR and CRs, as well as the expected signal yield in each bin of the SR. The best
 273 fit background model, as well as the best fit signal strength, are obtained by maximizing the
 274 joint likelihood function of all categories, where the signal strength (μ) is defined as the ratio
 275 of the observed signal cross section to the theoretical cross section predicted by the dark Higgs
 276 model.

$$\mu = \frac{\sigma}{\sigma_{\text{theo}}} \quad (2)$$

277 5.1 Likelihood function

278 The likelihood function maximized by the fit is:

$$\begin{aligned} \mathcal{L}_c(\mu_{Z\text{CR}}^Z, \mu_{SR}^{t\bar{t}}, \mu, \theta) = & \prod_{i,j} \text{Poisson} \left(d_{i,j}^{Z\text{CR}} | B_{i,j}^{Z\text{CR}}(\theta) + (1 + R^{W-Z}_{i,j}(\theta)) \mu_{Z\text{CR}i,j}^Z \right) \\ & \times \prod_{i,j} \text{Poisson} \left(d_{i,j}^{\text{TTECR}} | B_{i,j}^{\text{TTECR}}(\theta) + \frac{\mu_{SRi,j}^{t\bar{t}}}{R_{\text{TTECR}i,j}^{t\bar{t}}(\theta)} \right) \\ & \times \prod_{i,j} \text{Poisson} \left(d_{i,j}^{\text{TTMCR}} | B_{i,j}^{\text{TTMCR}}(\theta) + \frac{\mu_{SRi,j}^{t\bar{t}}}{R_{\text{TTMCR}i,j}^{t\bar{t}}(\theta)} \right) \\ & \times \prod_{i,j} \text{Poisson} \left(d_{i,j}^{\text{WEFCR}} | B_{i,j}^{\text{WEFCR}}(\theta) + \frac{R^{W-Z}_{i,j}(\theta) \mu_{Z\text{CR}i,j}^Z}{R_{\text{WEFCR}i,j}^W(\theta)} \right) \\ & \times \prod_{i,j} \text{Poisson} \left(d_{i,j}^{\text{WMFCR}} | B_{i,j}^{\text{WMFCR}}(\theta) + \frac{R^{W-Z}_{i,j}(\theta) \mu_{Z\text{CR}i,j}^Z}{R_{\text{WMFCR}i,j}^W(\theta)} \right) \\ & \times \prod_{i,j} \text{Poisson} \left(d_{i,j}^{\text{WEPCR}} | B_{i,j}^{\text{WEPCR}}(\theta) + \frac{R_p^W(\theta) R^{W-Z}_{i,j}(\theta) \mu_{Z\text{CR}i,j}^Z}{R_{\text{WEPCR}i,j}^W(\theta)} + \frac{\mu_{SRi,j}^{t\bar{t}}}{R_{\text{WEPCR}i,j}^{t\bar{t}}(\theta)} \right) \\ & \times \prod_{i,j} \text{Poisson} \left(d_{i,j}^{\text{WMPCR}} | B_{i,j}^{\text{WMPCR}}(\theta) + \frac{R_p^W(\theta) R^{W-Z}_{i,j}(\theta) \mu_{Z\text{CR}i,j}^Z}{R_{\text{WMPCR}i,j}^W(\theta)} + \frac{\mu_{SRi,j}^{t\bar{t}}}{R_{\text{WMPCR}i,j}^{t\bar{t}}(\theta)} \right) \\ & \times \prod_i \text{Poisson} \left(d_{i,j}^{\text{SR}} | B_{i,j}^{\text{SR}}(\theta) + (R_{p/f}^Z(\theta) + R_{p/f}^W(\theta) R_{i,j}^{W-Z}(\theta)) \mu_{Z\text{CR}i,j}^Z + \mu_{SRi,j}^{t\bar{t}} + \mu S_{i,j}(\theta) \right) \end{aligned} \quad (3)$$

Table 1: Expected yields from background processes and a subset of signal hypotheses in the signal region, for the three data sets considered. Values are before the maximum likelihood fit is performed (pre-fit) and uncertainties are statistical-only.

	2016	2017	2018
Z($\rightarrow \nu\nu$) + jets	7514.8 ± 29.2	7035.2 ± 33.3	6978.5 ± 38.8
W($\rightarrow \ell\nu$) + jets	3997.8 ± 38.5	2991.0 ± 40.2	2826.6 ± 50.5
t \bar{t}	5486.5 ± 199.7	5810.7 ± 60.0	6784.2 ± 133.7
Single t	646.0 ± 10.9	567.4 ± 12.5	614.6 ± 12.8
Diboson	718.0 ± 17.5	623.4 ± 17.8	606.4 ± 20.8
DY($\rightarrow \ell\ell$)+jets	56.8 ± 2.2	43.3 ± 2.0	37.1 ± 3.0
H $\rightarrow b\bar{b}$	57.6 ± 0.3	72.0 ± 0.3	83.8 ± 0.3
QCD multijet	93.3 ± 25.8	154.9 ± 41.7	163.2 ± 64.6
Total expected background	18570.7 ± 208.1	17297.9 ± 92.4	18094.2 ± 163.4
Signal hypotheses with $m_{\text{med}} = 1000 \text{ GeV}$ and $m_{H_D} = 130 \text{ GeV}$			
$m_{\text{DM}} = 150 \text{ GeV}$	684.8 ± 4.1	626.7 ± 3.9	687.2 ± 4.6
$m_{\text{DM}} = 500 \text{ GeV}$	$(381.6 \pm 2.1) \times 10^{-4}$	$(357.6 \pm 2.0) \times 10^{-4}$	$(399.6 \pm 2.4) \times 10^{-4}$
$m_{\text{DM}} = 1000 \text{ GeV}$	$(1341.2 \pm 6.7) \times 10^{-8}$	$(1005.9 \pm 6.7) \times 10^{-8}$	$(1341.2 \pm 10.1) \times 10^{-8}$

Table 2: Summary of selection criteria that define the different analysis regions.

Selection	SR	ZCR	WMPCR	WEPCR	WMFCR	WEFCR	TTMCR	TTECR
$U > 250 \text{ GeV}$	✓	✓	✓	✓	✓	✓	✓	✓
$\Delta p_T^{\text{miss}}(\text{PF-calorimeter}) > 0.5$	✓	✓	✓	✓	✓	✓	✓	✓
Leading AK15 $p_T > 160 \text{ GeV}$	✓	✓	✓	✓	✓	✓	✓	✓
Leading AK15 $m_{SD} \in [40, 300] \text{ GeV}$	✓	✓	✓	✓	✓	✓	✓	✓
$\min \Delta\phi(\vec{U}, A\vec{K}4s) > 0.5$	✓	✓	✓	✓	✓	✓	✓	✓
$\min \Delta\phi(\vec{U}, A\vec{K}15s) > 1.5$	✓	✓	✓	✓	✓	✓	✓	✓
$p_T^{\text{miss}} > 100 \text{ GeV}$	✗	✗	✗	✓	✗	✓	✗	✓
# of muons	0	0	1	0	1	0	1	0
# of electrons	0	0	0	1	0	1	0	1
# of photons	0	0	0	0	0	0	0	0
# of taus	0	0	0	0	0	0	0	0
# of non-overlapping AK4s	0	0	0	0	0	0	≥ 1	≥ 1
DeepAK15 $> \text{wp(90\% signal eff.)}$	pass	fail	pass	pass	fail	fail	pass	pass

In the above likelihood, $d_{i,j}^{*R}$ are the observed number of events in each (i,j) bin of the two-dimensional recoil-vs- m_{SD} distribution in the SR and CRs, while $B_{i,j}^{*R}$ is the number of background events. The parameter μ_{ZCR}^Z represents the yield of the $Z \rightarrow \nu\nu + \text{jets}$ background in the ZCR, and is left freely floating in the fit. The parameter $\mu_{SR}^{t\bar{t}}$ represents the yield of the $t\bar{t}$ background in the SR, that is left freely floating in the fit as well. The likelihood also includes the SR with μ being the signal strength parameter also left floating in the fit. The systematic uncertainties (θ) enter the likelihood as additive perturbations to the transfer factors R used in the modeling of the main backgrounds, as well as to the minor background and signal expectations, and are modeled as Gaussians.

Separate approaches are adopted to estimate the dominant ($Z + \text{jets}$, $W + \text{jets}$, $t\bar{t}$) and subdominant (single top, diboson, Higgs, and QCD multijet) backgrounds.

The predictions for the dominant $Z + \text{jets}$ and $W + \text{jets}$ backgrounds are based on the yield of $Z \rightarrow \nu\nu$ events in each bin of the ZCR. The per-bin yields for this process are defined as free parameters of the likelihood function. The yields for the $Z + \text{jets}$ and $W + \text{jets}$ contribution to the SR, as well as the yields of the $W + \text{jets}$ process in the CRs are defined relative to the $Z \rightarrow \nu\nu$ yields by introducing a set of per-bin transfer factors. The yields of $t\bar{t}$ events in the single-lepton CRs are similarly related via transfer factors to the $t\bar{t}$ event yields in the SRs. This choice of transfer factors takes into account the correlations between the $V + \text{jets}$ background contributions in all regions. In all cases, the central values of the transfer factors are obtained from the ratios of the simulated recoil-vs- m_{SD} spectra of the respective processes in the SRs to those in CRs. For the minor backgrounds the nominal expected yield per region is obtained directly from simulation.

Systematic uncertainties are incorporated in the likelihood function as nuisance parameters, as described in more detail below. In the case of the $V + \text{jets}$ and $t\bar{t}$ processes, the nuisance parameters affect the values of the transfer factors in each recoil-vs- m_{SD} variable bin and thus control the ratios of the contributions from different processes, as well as the ratios of the yields in the SRs to those in various CRs. For the subdominant background processes, the yields in each bin are directly parameterized in terms of the nuisance parameters. The final free parameter of the likelihood function is the signal strength modifier μ , which—for a given signal hypothesis—controls the signal normalization relative to the theoretical cross section.

The likelihood method relies on the accurate predictions of the ratios between the dominant backgrounds in the SRs and CRs, as well as on the absolute normalization and shape of the recoil-vs- m_{SD} distributions for the subdominant backgrounds. To achieve the most accurate possible predictions for these quantities, weights are applied to each simulated event to take into account both experimental and theoretical effects not present in the MC simulated samples. The experimental corrections are related to the trigger efficiencies, identification and reconstruction efficiencies of charged leptons and of b-tagged and doubleb-tagged jets, and the pileup distribution in simulation. Theoretical corrections are applied to the $V + \text{jets}$ processes in order to model the effects of NLO terms in the perturbative EW corrections [54]. The corrections are parameterized as functions of the generator-level boson p_T and are evaluated separately for the $W(\ell\nu) + \text{jets}$ and $Z(\ell\ell) + \text{jets}$ processes.

5.2 Systematic uncertainties

The inputs to the ML fit are subject to various experimental and theoretical uncertainties. Uncertainties in the measurement of the integrated luminosity in each year of data taking are 0.6–2% [55–57]. The uncertainties in the corrections for the L1 pre-firing effect in 2016 and

324 2017, as well as the uncertainties in the pile up correction are of the order of 1%. The uncertain-
 325 ties in the efficiencies of reconstructing and identifying electron candidates are 1% and 2–3%,
 326 respectively. For muons, the uncertainties in the identification efficiency are 1%, with an ad-
 327 dditional 1% uncertainty in the efficiency of the isolation criteria. A systematic uncertainty for
 328 each lepton/photon veto selection has been obtained by propagating the overall uncertainties
 329 in the identification of muons, electrons, photons, and taus, into the vetoed regions. While
 330 the uncertainties are found to be negligible for photon, muon, and electron vetoes, a 3% un-
 331 certainty in the tau veto is included. The uncertainties in the trigger efficiency are 1% for the
 332 single electron trigger and 1–2% for the p_T^{miss} trigger. The uncertainty in the modeling of p_T^{miss}
 333 in simulation [58] is dominated by the uncertainty in the jet energy corrections. The resulting
 334 bin migration affects the acceptance of the minimum requirement in p_T^{miss} . The change in rate
 335 is estimated to be 5% and it is included as a systematic uncertainty. An additional systematic
 336 uncertainty is included to cover the effect of the uncertainties in the AK15 jet energy corrections
 337 on the AK15 jet p_T . Also in this case, the resulting bin migration affects the acceptance of the
 338 minimum requirements in AK15 jet p_T . This introduces an effect on the rate of the order of 4%.
 339 The uncertainty in the b-tagging efficiency leads to a shape uncertainty applied to all processes
 340 in all regions. The uncertainty in the doubleb-tagging efficiency results in a shape uncertainty
 341 applied to the signal processes in SR. Uncertainties of 100% are assigned to the normalization
 342 of the QCD multijet background contributions in all the regions. These uncertainties are corre-
 343 lated between regions with the same source of fake: one uncertainty is applied to QCD multijet
 344 events in the SR and in the ZCR, a separate uncertainty is applied to QCD multijet events in the
 345 single-muon CRs, and similarly for the single-electron CRs. Additionally, uncertainties of 20%
 346 are assigned to the cross section of diboson, SM Higgs boson, and $Z(\rightarrow \ell\ell)$ +jets productions.
 347 Similarly, 10% uncertainties in the single top quark and $t\bar{t}$ production cross sections are also
 348 assigned. The theoretical uncertainties in the transfer factors related to higher-order effects in
 349 the QCD and EW perturbative expansions are calculated according to the prescription given
 350 in Ref. [54] and implemented, as described in Ref. [59]. Bin-by-bin statistical uncertainties are
 351 incorporated following the Barlow-Beeston-lite approach [60].
 352 The likelihood functions obtained for the three data taking years are combined in order to max-
 353 imize the statistical power of the analysis. The combination is performed by defining a com-
 354 bined likelihood describing all the analysis regions in all data sets. For this purpose, the effects
 355 of all theoretical uncertainties are assumed to be correlated. Most experimental uncertainties
 356 are dominated by the inherent precision of auxiliary measurements specific to each data set
 357 and are thus assumed to be uncorrelated between different data taking years. The experimen-
 358 tal uncertainties related to the determination of the integrated luminosity and to the b-tagging
 359 efficiency are partially correlated between the data taking years, which is taken into account by
 360 splitting the total uncertainty into its correlated and uncorrelated components. A summary of
 361 all the uncertainties considered for this analysis is reported in Table 3.

362 6 Results and interpretation

363 The ML fit is performed by combining the analysis categories as well as the data sets corre-
 364 sponding to the different years of data taking. The recoil-vs- m_{SD} distributions in SR before
 365 (“pre-fit”) and after (“post-fit”) the fit are shown in Figs. 2–4. In all cases, good agreement is
 366 observed between the background-only post-fit result and the data.
 367 Signal strength exclusion limits are presented for different signal hypotheses. All data sets and
 368 categories are included. The exclusion limits are calculated using the CL_s criterion [61–63], and
 369 an asymptotic approximation to the distribution of the profiled likelihood ratio test statistic. In

Table 3: Summary of statistical and systematic uncertainties included in the analysis.

Source	MC-Driven Processes	Transfer Factors	Uncertainty
Luminosity	All MC processes in all regions	\times	0.6 to 2%
Pile-up	All processes in all regions	\times	$\sigma(1\%)$
L1 prefireing	All processes in all regions	\times	$\sigma(1\%)$
p_T^{miss} trigger eff.	All processes in SR, ZCR, and single muon regions	$R_{\text{TTMCR}}^{t\bar{t}}, R_{\text{WMPCR}}^{t\bar{t}}, R_{\text{WMPCR}}^W, R_{\text{WMFCR}}^W$	1/2%
Single electron trigger eff.	All MC processes in single electron regions	$R_{\text{TTECR}}^{t\bar{t}}, R_{\text{WEPCR}}^{t\bar{t}}, R_{\text{WEPCR}}^W, R_{\text{WEFCR}}^W$	1%
Muon iso. eff.	All MC processes in single muon regions	$R_{\text{TTMCR}}^{t\bar{t}}, R_{\text{WMPCR}}^{t\bar{t}}, R_{\text{WMPCR}}^W, R_{\text{WMFCR}}^W$	1%
Muon ID eff.	All MC processes in single muon regions	$R_{\text{TTMCR}}^{t\bar{t}}, R_{\text{WMPCR}}^{t\bar{t}}, R_{\text{WMPCR}}^W, R_{\text{WMFCR}}^W$	1%
Electron reco. eff.	All MC processes in single electron regions	$R_{\text{TTECR}}^{t\bar{t}}, R_{\text{WEPCR}}^{t\bar{t}}, R_{\text{WEPCR}}^W, R_{\text{WEFCR}}^W$	1%
Electron ID eff.	All MC processes in single electron regions	$R_{\text{TTECR}}^{t\bar{t}}, R_{\text{WEPCR}}^{t\bar{t}}, R_{\text{WEPCR}}^W, R_{\text{WEFCR}}^W$	2/3%
p_T^{miss}	All MC processes in SR and ZCR	\times	5%
Jet energy corr.	All MC processes in all regions	\times	4%
AK4 b-tagging eff.	All MC processes in all regions	$R_{\text{TTECR}}^{t\bar{t}}, R_{\text{TTMCR}}^{t\bar{t}}$	shape
AK15 double b-tagging eff.	Signal MC processes in SR	\times	shape
Bernstein coefficients	\times	$R_{\text{p/f}}^{Z,\text{data}}, R_{\text{p/f}}^{W,\text{data}}$	Unconstrained nuisances
Diboson norm.	Diboson MC process in all regions	\times	20%
SM Higgs boson norm.	SM Higgs boson MC processes in all regions	\times	20%
$Z(\rightarrow \ell\ell) + \text{jets}$ normalization	$Z(\rightarrow \ell\ell) + \text{jets}$ processes in all regions	\times	20%
Single top quark norm.	Single top quark MC processes in all regions	\times	10%
$t\bar{t}$ norm.	$t\bar{t}$ MC processes in all “fail” regions	\times	10%
QCD- p_T^{miss} norm.	QCD MC processes in SR and ZCR	\times	100%
QCD-electron norm.	QCD MC processes in single electron regions	\times	100%
QCD-muon norm.	QCD MC processes in single muon regions	\times	100%
W+jets norm.	\times	$R_{\text{W-Z}}$	40%
Higher-order corrections	\times	All transfer factors but $R_{\text{p/f}}^{Z,\text{data}}, R_{\text{p/f}}^{W,\text{data}}$	shape
MC statistics	All MC processes in all regions		shape

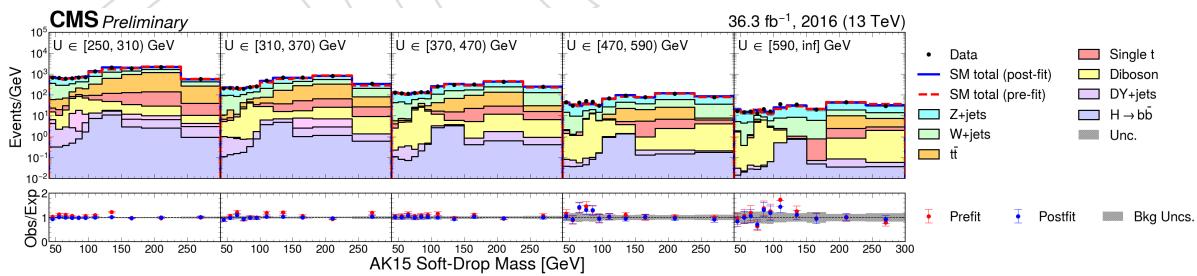


Figure 2: Post-fit m_{SD} distributions in bins of U for 2016. Distributions in signal region are shown. The top plots present stacked post-fit predictions for the backgrounds superimposed to the data. The blue curve represents the post-fit total background prediction, while the red curve represent the pre-fit one. The bottom plots present the ratio between the data (labeled as “Observed”) and the background predictions (labeled as “Expected”). The ratio between the data and the post-fit prediction is represented by the blue dots, while the ratio between the data and the pre-fit prediction is represented by the red ones.

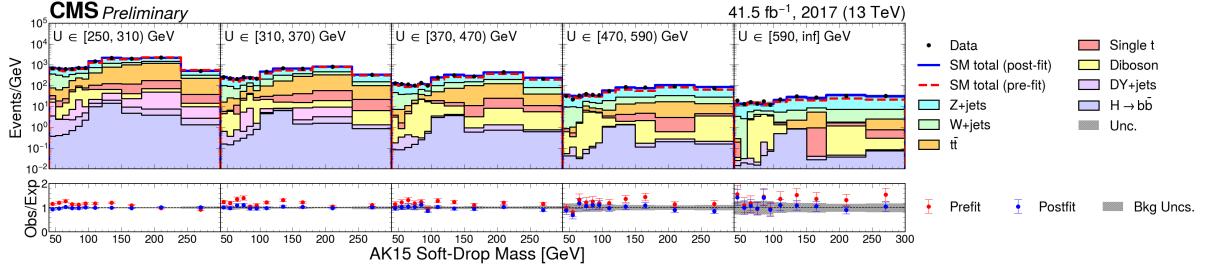


Figure 3: Post-fit m_{SD} distributions in bins of U for 2017. Distributions in signal region are shown. The top plots present stacked post-fit predictions for the backgrounds superimposed to the data. The blue curve represents the post-fit total background prediction, while the red curve represent the pre-fit one. The bottom plots present the ratio between the data (labeled as "Observed") and the background predictions (labeled as "Expected"). The ratio between the data and the post-fit prediction is represented by the blue dots, while the ratio between the data and the pre-fit prediction is represented by the red ones.

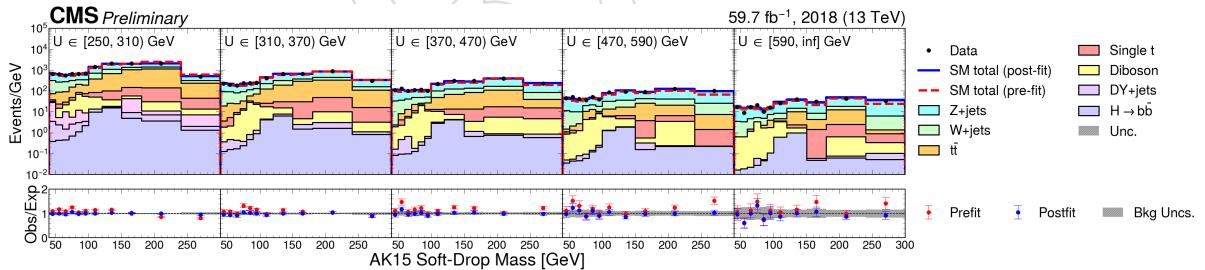


Figure 4: Post-fit m_{SD} distributions in bins of U for 2018. Distributions in signal region are shown. The top plots present stacked post-fit predictions for the backgrounds superimposed to the data. The blue curve represents the post-fit total background prediction, while the red curve represent the pre-fit one. The bottom plots present the ratio between the data (labeled as "Observed") and the background predictions (labeled as "Expected"). The ratio between the data and the post-fit prediction is represented by the blue dots, while the ratio between the data and the pre-fit prediction is represented by the red ones.

370 this method, a signal-plus-background fit is performed for each signal hypothesis in addition
 371 to the background-only fit. In the signal fits, the nuisance parameters are profiled, and the
 372 resulting best fit nuisance parameters vary for the different signal hypotheses. Consequently,
 373 different nonzero best fit values for the signal strength can be obtained for different signals
 374 even if the background-only fit succeeds in modeling the data. In the exclusion limits, this
 375 feature is represented by differences between the observed and expected limits.

376 Exclusion limits are calculated in the two-dimensional parameter space of the DM and me-
 377 diator masses, m_{DM} and m_{med} , constrained by the fact that only scenarios in which the DM
 378 particle is more massive than the dark Higgs boson are considered. The coupling between the
 379 mediator and SM quarks is set to a constant value of $g_q = 0.25$, and the mediator-DM coupling
 380 is set to $g_\chi = 1.0$. The resulting exclusion limits at 95% confidence level (CL) on the signal
 381 strength μ are shown in Figs. 5-6 for different hypotheses of the dark Higgs boson mass. For
 382 the smaller value allowed by a specific choice of the dark Higgs boson mass, different values
 383 of the mediator mass m_{med} are expected to be excluded. The excluded value of m_{med} reduces
 384 with increasing values of m_{DM} , as the branching fraction for decays of the mediator into dark
 385 matter candidates is reduced.

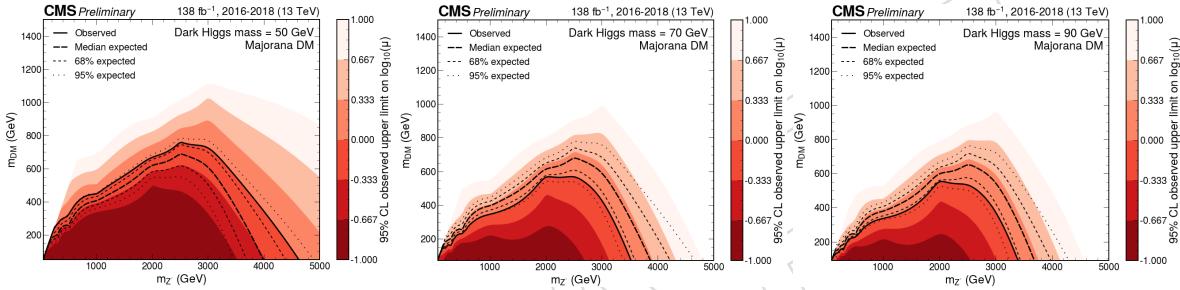


Figure 5: Expected and observed exclusion limits at 95% CL on the signal strength $\mu = \sigma/\sigma_{\text{theo}}$ as a function of m_{med} for a dark Higgs boson mass of 50 GeV (left), 70 GeV (middle), and 90 GeV (right). Only scenarios where the DM particle is more massive than the dark Higgs boson are considered. The black solid line indicates the exclusion boundary $\mu = 1$. Parameter combinations with larger values of μ are excluded.

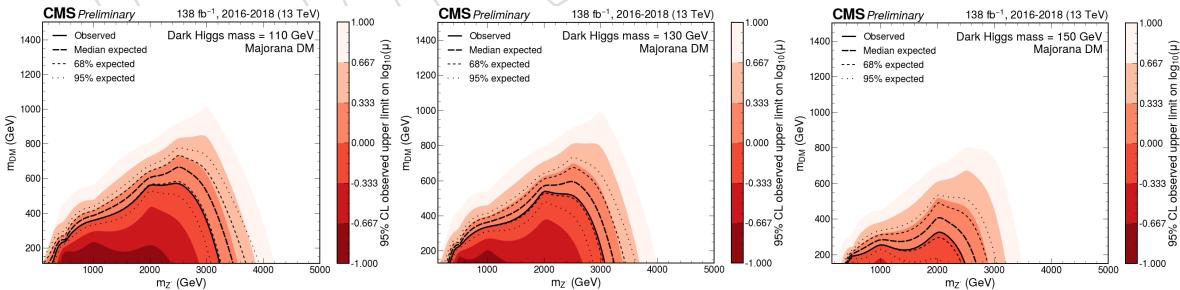


Figure 6: Expected and observed exclusion limits at 95% CL on the signal strength $\mu = \sigma/\sigma_{\text{theo}}$ as a function of m_{med} for a dark Higgs boson mass of 110 GeV (left), 130 GeV (middle), and 150 GeV (right). Only scenarios where the DM particle is more massive than the dark Higgs boson are considered. The black solid line indicates the exclusion boundary $\mu = 1$. Parameter combinations with larger values of μ are excluded.

386 7 Summary

387 A search for physics beyond the standard model in events with a large-cone energetic jet consist-
 388 ent with the hadronization of a resonant b-quark pair and large missing transverse momentum
 389 has been presented. A data set of proton-proton collisions at a center-of-mass energy of 13 TeV,
 390 corresponding to an integrated luminosity of 137 fb^{-1} is analyzed. A joint maximum likelihood
 391 fit over a combination of signal and control regions is used to constrain standard model back-
 392 ground processes and to extract a possible signal. The result is interpreted in terms of exclusion
 393 limits at 95% confidence level on the parameters of a model of production of a dark Higgs bo-
 394 son in association with dark matter particles. Values of the mediator mass of up to 2.5–4.5 TeV
 395 are excluded, depending on the mass of the dark Higgs boson and assuming the couplings of
 396 $g_q = 0.25$ between the mediator and quarks, and $g_{\text{DM}} = 1.0$ between the mediator and the DM
 397 particles.

398 Acknowledgments

399 References

- 400 [1] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: Evidence, candidates and
 401 constraints”, *Phys. Rept.* **405** (2005) 279–390,
 402 doi:[10.1016/j.physrep.2004.08.031](https://doi.org/10.1016/j.physrep.2004.08.031), arXiv:[hep-ph/0404175](https://arxiv.org/abs/hep-ph/0404175).
- 403 [2] S. Gori et al., “Dark sector physics at high-intensity experiments”, Technical Report
 404 FERMILAB-PUB-22-672-SCD-T, 2022. arXiv:[2209.04671](https://arxiv.org/abs/2209.04671).
- 405 [3] CMS Collaboration, “Search for new particles in events with energetic jets and large
 406 missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ ”, *JHEP* **11**
 407 (2021) 153, doi:[10.1007/JHEP11\(2021\)153](https://doi.org/10.1007/JHEP11(2021)153), arXiv:[2107.13021](https://arxiv.org/abs/2107.13021).
- 408 [4] ATLAS Collaboration, “Search for new phenomena in events with an energetic jet and
 409 missing transverse momentum in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector”,
 410 *Phys. Rev. D* **103** (2021), no. 11, 112006, doi:[10.1103/PhysRevD.103.112006](https://doi.org/10.1103/PhysRevD.103.112006),
 411 arXiv:[2102.10874](https://arxiv.org/abs/2102.10874).
- 412 [5] N. F. Bell, Y. Cai, and R. K. Leane, “Dark Forces in the Sky: Signals from Z' and the Dark
 413 Higgs”, *JCAP* **08** (2016) 001, doi:[10.1088/1475-7516/2016/08/001](https://doi.org/10.1088/1475-7516/2016/08/001),
 414 arXiv:[1605.09382](https://arxiv.org/abs/1605.09382).
- 415 [6] F. Kahlhoefer, K. Schmidt-Hoberg, T. Schwetz, and S. Vogl, “Implications of unitarity and
 416 gauge invariance for simplified dark matter models”, *JHEP* **02** (2016) 016,
 417 doi:[10.1007/JHEP02\(2016\)016](https://doi.org/10.1007/JHEP02(2016)016), arXiv:[1510.02110](https://arxiv.org/abs/1510.02110).
- 418 [7] N. F. Bell, Y. Cai, and R. K. Leane, “Impact of mass generation for spin-1 mediator
 419 simplified models”, *JCAP* **01** (2017) 039, doi:[10.1088/1475-7516/2017/01/039](https://doi.org/10.1088/1475-7516/2017/01/039),
 420 arXiv:[1610.03063](https://arxiv.org/abs/1610.03063).
- 421 [8] M. Duerr et al., “Hunting the dark Higgs”, *JHEP* **04** (2017) 143,
 422 doi:[10.1007/JHEP04\(2017\)143](https://doi.org/10.1007/JHEP04(2017)143), arXiv:[1701.08780](https://arxiv.org/abs/1701.08780).
- 423 [9] ATLAS Collaboration, “Search for dark matter produced in association with a dark Higgs
 424 boson decaying into W^+W^- in the one-lepton final state at $\sqrt{s}=13 \text{ TeV}$ using 139 fb^{-1} of
 425 pp collisions recorded with the ATLAS detector”, *JHEP* **07** (2023) 116,
 426 doi:[10.1007/JHEP07\(2023\)116](https://doi.org/10.1007/JHEP07(2023)116), arXiv:[2211.07175](https://arxiv.org/abs/2211.07175).

- [427] [10] CMS Collaboration, “Search for dark matter particles in W^+W^- events with transverse momentum imbalance in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **03** (2024) 134, doi:10.1007/JHEP03(2024)134, arXiv:2310.12229.
- [430] [11] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [432] [12] CMS Collaboration, “Development of the CMS detector for the CERN LHC Run 3”, *JINST* **19** (2024) P05064, doi:10.1088/1748-0221/19/05/P05064.
- [434] [13] CMS Collaboration, “Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **15** (2020), no. 10, P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.
- [437] [14] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017), no. 01, P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.
- [439] [15] CMS Collaboration, “Performance of the CMS high-level trigger during LHC Run 2”, *JINST* **19** (2024), no. 11, P11021, doi:10.1088/1748-0221/19/11/P11021, arXiv:2410.17038.
- [442] [16] CMS Collaboration, “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”, *JINST* **16** (2021), no. 05, P05014, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.
- [445] [17] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.
- [448] [18] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [451] [19] CMS Tracker Group Collaboration, “The CMS phase-1 pixel detector upgrade”, *JINST* **16** (2021) P02027, doi:10.1088/1748-0221/16/02/P02027, arXiv:2012.14304.
- [453] [20] CMS Collaboration, “Track impact parameter resolution for the full pseudo rapidity coverage in the 2017 dataset with the CMS phase-1 pixel detector”, CMS Detector Performance Summary CMS-DP-2020-049, 2020.
- [456] [21] CMS Collaboration, “2017 tracking performance plots”, CMS Detector Performance Summary CMS-DP-2017-015, 2017.
- [458] [22] CMS Collaboration, “Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid”, CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015.
- [460] [23] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017), no. 10, P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [463] [24] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_T jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [465] [25] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

- [26] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- [27] CMS Collaboration, “Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector”, *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.
- [28] CMS Collaboration, “Pileup mitigation at CMS in 13 TeV data”, *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/p09018, arXiv:2003.00503.
- [29] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup per particle identification”, *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.
- [30] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, “Towards an understanding of jet substructure”, *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.
- [31] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, “Jet substructure as a new Higgs search channel at the LHC”, *Phys. Rev. Lett.* **100** (2008) 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.
- [32] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft drop”, *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.
- [33] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [34] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements”, *Eur. Phys. J. C* **80** (2020), no. 1, 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.
- [35] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [36] NNPDF Collaboration, “Parton distributions for the LHC run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [37] NNPDF Collaboration, “Parton distributions from high-precision collider data”, *Eur. Phys. J. C* **77** (2017), no. 10, 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.
- [38] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [39] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, “Matching matrix elements and shower evolution for top-quark production in hadronic collisions”, *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.
- [40] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [41] M. Czakon, P. Fiedler, and A. Mitov, “Total top-quark pair-production cross section at hadron colliders through $o(\alpha_s^4)$ ”, *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

- 508 [42] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with
509 shower in POWHEG: s - and t -channel contributions”, *JHEP* **09** (2009) 111,
510 doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum:
511 doi:10.1007/JHEP02(2010)011].
- 512 [43] E. Re, “Single-top Wt-channel production matched with parton showers using the
513 POWHEG method”, *Eur. Phys. J. C* **71** (2011) 1547,
514 doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- 515 [44] M. Aliev et al., “HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR”,
516 *Comput. Phys. Commun.* **182** (2011) 1034, doi:10.1016/j.cpc.2010.12.040,
517 arXiv:1007.1327.
- 518 [45] P. Kant et al., “HATHOR for single top-quark production: Updated predictions and
519 uncertainty estimates for single top-quark production in hadronic collisions”, *Comput.*
520 *Phys. Commun.* **191** (2015) 74, doi:10.1016/j.cpc.2015.02.001,
521 arXiv:1406.4403.
- 522 [46] T. Gehrmann et al., “ W^+W^- production at hadron colliders in next to next to leading
523 order QCD”, *Phys. Rev. Lett.* **113** (2014) 212001,
524 doi:10.1103/PhysRevLett.113.212001, arXiv:1408.5243.
- 525 [47] J. M. Campbell and R. K. Ellis, “An update on vector boson pair production at hadron
526 colliders”, *Phys. Rev. D* **60** (1999) 113006, doi:10.1103/PhysRevD.60.113006,
527 arXiv:hep-ph/9905386.
- 528 [48] LHC Dark Matter Working Group, “Recommendations of the LHC dark matter working
529 group: Comparing LHC searches for dark matter mediators in visible and invisible decay
530 channels and calculations of the thermal relic density”, *Phys. Dark Univ.* **26** (2019)
531 100377, doi:10.1016/j.dark.2019.100377, arXiv:1703.05703.
- 532 [49] CMS Collaboration, “A portrait of the Higgs boson by the CMS experiment ten years
533 after the discovery.”, *Nature* **607** (2022), no. 7917, 60,
534 doi:10.1038/s41586-022-04892-x, arXiv:2207.00043. [Erratum:
535 doi:10.1038/s41586-023-06164-8].
- 536 [50] CMS Collaboration, “Identification of heavy, energetic, hadronically decaying particles
537 using machine-learning techniques”, *JINST* **15** (2020), no. 06, P06005,
538 doi:10.1088/1748-0221/15/06/P06005, arXiv:2004.08262.
- 539 [51] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction
540 with proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **13** (2018) P06015,
541 doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.
- 542 [52] C. Collaboration, “Identification of hadronic tau lepton decays using a deep neural
543 network”, *JINST* **17** (2022) P07023, doi:10.1088/1748-0221/17/07/P07023,
544 arXiv:2201.08458.
- 545 [53] E. Bols et al., “Jet Flavour Classification Using DeepJet”, *JINST* **15** (2020), no. 12, P12012,
546 doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.
- 547 [54] J. M. Lindert et al., “Precise predictions for V+jets dark matter backgrounds”, *Eur. Phys.*
548 *J. C* **77** (2017) 829, doi:10.1140/epjc/s10052-017-5389-1, arXiv:1705.04664.

- 549 [55] CMS Collaboration, “Precision luminosity measurement in proton-proton collisions at
550 $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS”, *Eur. Phys. J. C* **81** (2021), no. 9, 800,
551 doi:[10.1140/epjc/s10052-021-09538-2](https://doi.org/10.1140/epjc/s10052-021-09538-2), arXiv:[2104.01927](https://arxiv.org/abs/2104.01927).
- 552 [56] CMS Collaboration, “CMS luminosity measurement for the 2017 data-taking period at
553 $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.
- 554 [57] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at
555 $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.
- 556 [58] CMS Collaboration, “Performance of the CMS missing transverse momentum
557 reconstruction in pp data at $\sqrt{s} = 8$ TeV”, *JINST* **10** (2015) P02006,
558 doi:[10.1088/1748-0221/10/02/P02006](https://doi.org/10.1088/1748-0221/10/02/P02006), arXiv:[1411.0511](https://arxiv.org/abs/1411.0511).
- 559 [59] CMS Collaboration, “Search for new physics in final states with an energetic jet or a
560 hadronically decaying W or Z boson and transverse momentum imbalance at
561 $\sqrt{s} = 13$ TeV”, *Phys. Rev. D* **97** (2018) 092005, doi:[10.1103/PhysRevD.97.092005](https://doi.org/10.1103/PhysRevD.97.092005),
562 arXiv:[1712.02345](https://arxiv.org/abs/1712.02345).
- 563 [60] J. S. Conway, “Incorporating Nuisance Parameters in Likelihoods for Multisource
564 Spectra”, in *PHYSTAT 2011*, pp. 115–120. 2011. arXiv:[1103.0354](https://arxiv.org/abs/1103.0354).
565 doi:[10.5170/CERN-2011-006.115](https://doi.org/10.5170/CERN-2011-006.115).
- 566 [61] A. L. Read, “Presentation of search results: The CL_s technique”, *J. Phys. G* **28** (2002) 2693,
567 doi:[10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- 568 [62] T. Junk, “Confidence level computation for combining searches with small statistics”,
569 *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:[10.1016/S0168-9002\(99\)00498-2](https://doi.org/10.1016/S0168-9002(99)00498-2),
570 arXiv:[hep-ex/9902006](https://arxiv.org/abs/hep-ex/9902006).
- 571 [63] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for
572 likelihood-based tests of new physics”, *Eur. Phys. J. C* **71** (2011) 1554,
573 doi:[10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0), arXiv:[1007.1727](https://arxiv.org/abs/1007.1727). [Erratum:
574 doi:[10.1140/epjc/s10052-013-2501-z](https://doi.org/10.1140/epjc/s10052-013-2501-z)].