DRAFT CMS Physics Analysis Summary

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

2018/08/17

Head Id: 464952

Archive Id: 464955:467209MP

Archive Date: 2018/06/18

Archive Tag: trunk

Search for associated production of dark matter with a Higgs boson decaying to a pair of bottom quarks in pp collisions at sqrt(s) = 13 TeV

The CMS Collaboration

Abstract

A search for dark matter produced in association with a Higgs boson decaying to a bottom quark-antiquark pair $(b\bar{b})$ is performed in proton-proton collisions at a center-of-mass energy of 13 TeV collected with the CMS detector at the LHC. The analyzed data sample corresponds to an integrated luminosity of 35.9 fb⁻¹. The signal is characterized by a large missing transverse momentum recoiling against a bb system that has a large Lorentz boost. The number of events observed in the data is consistent with the standard model background prediction. Results are interpreted in terms of limits on parameters in the type-2 two Higgs doublet model (2HDM) extended by an additional light pseudoscalar boson a (2HDM+a) and the baryonic a2 simplified models. For the baryonic a3 model, the presented results constitute the most stringent constraints to date. The 2HDM+a4 model is tested experimentally for the first time.

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

PDFAuthor: Raman Khurana, Benedikt Maier, Shin-Shan Eiko Yu, Matteo Cremonesi, Bo

Jayatilaka

PDFTitle: Search for dark matter in H"(bb")+MET final state "(full 2016 dataset")

PDFSubject: CMS

PDFKeywords: Dark matter, LHC, CMS, boosted Higgs boson tagging, physics, software,

computing

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



1. Introduction 1

1 Introduction

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for physics beyond the standard model (SM) [456–458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consistent with the hypothesis that DM is primarily composed of weakly interacting massive particles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. Assuming the pair production of DM particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signature. If DM particles are produced in association with a detectable SM particle, which could be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding protons, or through new effective couplings between DM and SM particles, their existence could be inferred via a large transverse momentum imbalance in the collision event.

While ISR production of the SM Higgs boson (h) [460–462] is highly suppressed due to the Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs boson and DM particles can occur if the Higgs boson takes part in the interaction producing the DM particles [463–465]. Such a production mechanism would allow to directly probe the structure of the effective DM-SM coupling.

In this paper, we present a search for DM production in association with an SM Higgs boson 19 that decays into a bottom quark-antiquark pair (bb). As the $h \to bb$ decay mode has the largest 20 branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest 21 signal yield. The search is performed using the data set collected by the CMS experiment [466] 22 at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated 23 luminosity of approximately 35.9 fb⁻¹. Results are interpreted in terms of two simplified mod-24 els predicting this signature. The first one is type-2 two Higgs doublet model extended by 25 an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi \bar{\chi}$. 27 The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is 28 exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM 29 particles. Representative Feynman diagrams for the two models are presented in Fig. 50. 30

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM 31 particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 32 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no 34 significant deviation from SM predictions [468]. In contrast to previously explored 2HDM mod-35 els [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In 36 this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar h, assumed to be the observed 125 GeV Higgs boson, and a heavy neutral CP-even scalar H, that are 38 the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mix-39 ing angle α ; a heavy neutral CP-odd pseudoscalar Aand a light neutral CP-odd pseudoscalar a, that are the result of the mixing of the CP-odd mediator P with the CP-odd Higgs, with θ representing the associated mixing angle; and two heavy charged scalars H[±] with identical 43

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the

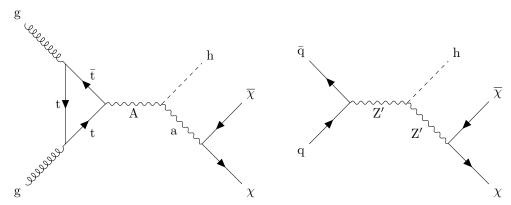


Figure 1: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z' model (right).

charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H}=m_{\rm H^\pm}=m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_\chi=10\,{\rm GeV}$.

The baryonic Z' model [465] is an extension of the SM with an additional $U(1)_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding $U(1)_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the $U(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_BZ'Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin\theta/v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_q , is 0.25 and the Z' coupling to DM, g_χ , is set to 1. This is well below the bounds g_q , $g_\chi \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin\theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'}\sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and $m_{X'}$, which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, together with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically independent data samples are used to predict the hadronic recoil distribution for these SM processes that constitute the largest sources of background. Both signal and background contributions to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

2 The CMS detector

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum (p_T) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of

3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon 81 pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta|$ < 3. The steel and quartz-fiber forward 84 Cherenkov hadron calorimeter extends the coverage to $|\eta|$ < 5. The muon system consists 85 of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and 86 covers $|\eta|$ < 2.4. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than $4 \mu s$, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces 89 the event rate to 1 kHz. 90

91 3 Simulated data samples

92

93

95

96

97

100

101

102

103

105

106

107

108

109

110

111

112

113

115

119

120

121

122

123

124

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG V2 [472-474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the p_T of the respective boson with NLO QCD corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

4 Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_{\rm T}^{i2}$, where $p_{\rm T}^i$ is the transverse momentum of the $i^{\rm th}$ track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, large-area jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central ($\eta < 2.4$). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$ is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , will affect the shape of the distribution of $m_{\rm SD}$ for the background processes. In this search, the value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore desirable to preserve a smoothly falling jet mass distribution as a function of $p_{\rm T}$. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with $p_{\rm T}$, decorrelating the N_2 variable as a function of $p_{\rm T}$ would be the most appropriate procedure. The decorrelation strategy described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494].

This search also utilizes narrow jets clustered with the anti- k_T algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta|$ < 2.4).

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in

5. Event selection 5

the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the con-172 sistency with originating from the primary vertex are imposed. The reconstructed electron is 173 required to be within $|\eta|$ < 2.5, excluding the transition region 1.44 < $|\eta|$ < 1.57 between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction 175 approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment 176 in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons 180 and poorly measured mesons tagged as muons [497]. These additional criteria include require-181 ments on the number of hits in the tracker and in the muon systems, the fit quality of the global 182 muon track, and its consistency with the primary vertex. Muon candidates with $|\eta|$ < 2.4 are 183 considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are re-185 constructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and 186 neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton 187 decays. The τ_h candidates with $p_T > 18$ GeV and $|\eta| < 2.3$ are considered [496, 498, 499]. Pho-188 ton candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15$ GeV and $|\eta| < 2.5$. 190

The missing transverse momentum $\vec{p}_{\rm T}^{\rm miss}$ is defined as the negative vectorial sum of the $p_{\rm T}$ of all the reconstructed PF candidates. Its magnitude is denoted as $p_{\rm T}^{\rm miss}$. Corrections to jet momenta are propagated to the $p_{\rm T}^{\rm miss}$ calculation as well as event filters [500] are used to remove spurious high $p_{\rm T}^{\rm miss}$ events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

5 Event selection

197

198

199

201

202

203

204

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}^{\rm miss}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T,\text{trig}}^{\text{miss}}$ or H_{T}^{miss} , where $p_{T,\text{trig}}^{\text{miss}}$ (H_{T}^{miss}) is the magnitude of the vectorial (scalar) sum of \vec{p}_{T} of all the particles (jets with $p_{T} > 20\,\text{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{T,\text{trig}}^{\text{miss}}$ calculation. Thresholds on $p_{T,\text{trig}}^{\text{miss}}$ and H_{T}^{miss} are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{T,\text{trig}}^{\text{miss}}$ and H_{T}^{miss} are reffered to as p_{T}^{miss} triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T} > 200\,{\rm GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\rm SD} < 150\,{\rm GeV}$ and $N_2^{\rm DDT} < 0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (\vec{U}) must be larger than 0.4

Table 1: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, t t , W+jets	0	0	pass
Single-lepton, b-tagged	tī, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

rad to reject multijet events that mimic signal events. Vetoes on τ_h candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from $t\bar{t}$ events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_T > 10 \,\text{GeV}$ and $|\eta| < 2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between 60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of the trigger efficiency, at least one of the two electrons must have $p_{\rm T} > 40\,{\rm GeV}$, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from W+jets and semileptonic $t\bar{t}$ events. To predict these backgrounds, four single-lepton samples are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target $t\bar{t}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the p_T^{miss} trigger triggers described above, as well as single electron events are selected using the same single-electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and isolation requirements. In addition, samples with a single electron need to have $p_T^{miss} > 50 \, \text{GeV}$ to avoid a large contamination from multijet events.

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

6 Signal extraction

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the $p_{\rm T}^{\rm miss}$ and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and

are given by the ratio of the prediction for a given bin in $p_{\rm T}^{\rm miss}$ in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the $t\bar{t}$ process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt backgrounds take into account the impact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron control samples, the additional requirement on p_T^{miss} . Since the CRs with no b-tagged AK4 jets 263 and a double-b-tagged CA15 jet also have significant contributions from the tt process, transfer 264 factors to predict this contamination from tt events are also imposed between the single-lepton 265 CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contamination from W+jets production in the tt CR with events that fail the double-b tag requirement. 267 Likewise, the Z+jets background prediction in the signal region is connected to the and dilep-268 ton CRs via transfer factors. They account for the difference in the branching fractions of the 269 $Z \rightarrow \nu \nu$ and the $Z \rightarrow \ell \ell$ decays and the impacts of lepton acceptances and selection efficiencies.

7 Systematic uncertainties

271

273

274

275

276

278

279

280

281

282

284

285

286

287

289

290

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncertainties of the search. They can either affect the rate or the shape of $p_{\rm T}^{\rm miss}$ (U) for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal distribution. The list of the systematic uncertainties considered in this search is presented in Table 23. To better estimate their impact on the results, uncertainties from a similar source (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties have been ordered according to decreasing improvement in the expected limit obtained when removing the group from the list of nuisances included in the likelihood fit. The description of each single uncertainty in the text follows the same order.

Scale factors are used to correct for the differences in the double-b tagger misidentification efficiencies between data and prediction from simulation for W/Z+jets production and for $t\bar{t}$ production. These scale factors are measured by simultaneously fitting events that pass or fail the double-b tag requirement. The correlation between the double-b tagger and the p_{T}^{miss} (or U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate independently from each other within a constraint that depends on the recoil value. Such dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. p_{T}^{miss}/U . This shape uncertainty in the double-b scale factors measurement is the one that has the largest impact on the limits on the signal cross section.

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and tt̄ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the $p_{\rm T}$ of the CA15 jet. This shape uncertainty has been derived through a measurement performed using a sample enriched in multijet events with double-muon-tagged $g \to b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substruc-

301

302

303

308

309

310

311

Table 2: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

Ctti	Т	D	Tours and any association to
Systematic uncertainty	Туре	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, tt̄	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, tt̄	1.9%
Double-b tagging	shape	SM h, signal	1.2%
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\rm T}^{\rm miss}$ magnitude	5%	all	0.7%
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%
$p_{\rm T}^{\rm miss}$ trigger muon multiplicity	shape	Z+jets, W+jets	
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	< 0.5%
single-electron trigger	1%	all	
AK4 b tagging	shape	all	< 0.5%
au lepton veto	3%	all	< 0.5%
Lepton efficiency	1% per lepton	all	< 0.5%
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%
Renorm./fact. scales	shape	SM h	
PDF	shape	SM h	< 0.59/
Multijet normalization	100%	multijet	< 0.5%
Theoretical cross section	20%	t, diboson	

ture variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes where the decay of a resonance inside the CA15 jet cone is expected. Such processes include signal production together with SM h and diboson production. The uncertainty has been derived from the efficiency measurement obtained by performing a fit in a control sample enriched in semi-leptonic $t\bar{t}$ events, where the CA15 jet originates from the W boson that comes from the hadronically decaying top quark.

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in $p_{\rm T}^{\rm miss}$ magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, obtained by varying those parameters in simulation event-by-event

The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. Differences on the order of 2% are observed between single-muon and dimuon events at lower U values and they are sources of an additional systematic uncertainty in the transfer factors for those processes whose prediction relies on data events in the single-muon and dimuon CRs (tt̄, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the identified muon, they can change the shape of the U distribution and are thus treated as shape uncertainties. The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiency is parametrized as a function of p_T^{miss} . The uncertainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies of the single-electron triggers are parametrized as a function of the electron p_T and η and an

8. Results 9

associated 1% systematic uncertainty is added into the fit.

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization scales and PDFs are included as shape variations. Being a negligible background source, an uncertainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by requiring $p_{\rm T}^{\rm miss} > 250\,{\rm GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty in QCD multijet yields in the signal region, while separate nuisance parameters are introduced for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as measured by CMS in Refs. [509, 510] and correlated across the SR and CRs.

8 Results

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficiencies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z' model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results of this search are interpreted in terms of upper limits on the signal strength modifier $\mu =$

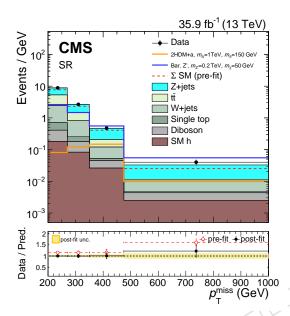


Figure 2: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

Table 3: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

$p_{\mathrm{T}}^{\mathrm{miss}}$ -bin	200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
Z+jets	248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
tŧ	199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
W+jets	121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
Single top quark	21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
Diboson	16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
SM h	12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
Σ (SM)	619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
Data	619	214	59	21
2 HDM+ a , $m_A = 1$ TeV, $m_a = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2 \text{TeV}$, $m_{\chi} = 50 \text{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

8. Results

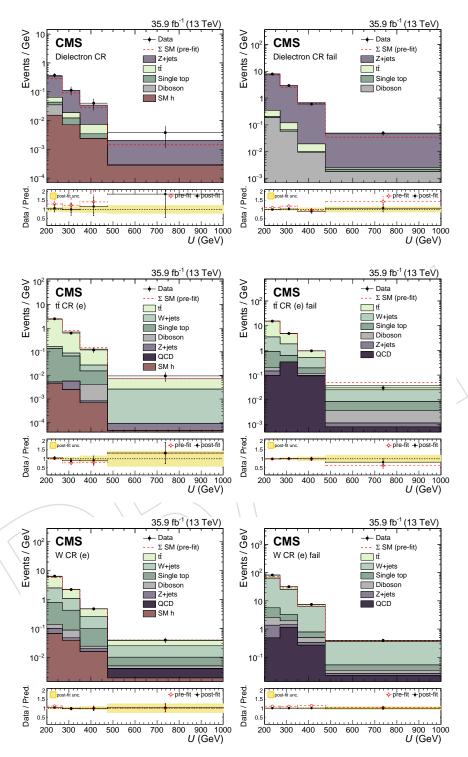


Figure 3: The *U* distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

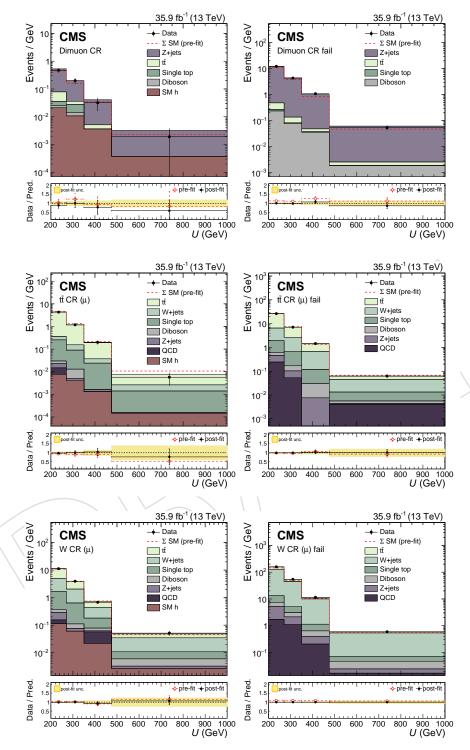


Figure 4: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

8. Results 13

 $\sigma/\sigma_{\text{theory}}$, where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

Figure 54 shows the upper limits on μ for the three scans (m_A , $\sin \theta$, and $\tan \beta$) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, $\sin \theta = 0.35$ and $\tan \beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming $\tan \beta = 1$. Also excluded are $\tan \beta$ values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

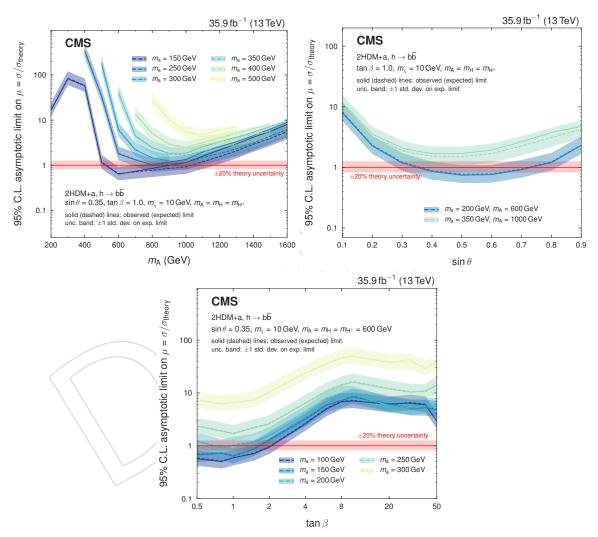


Figure 5: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus.

383

384

385

386

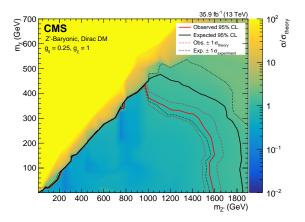


Figure 6: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

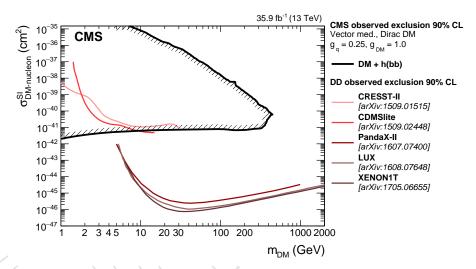


Figure 7: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2\mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},\tag{1}$$

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5\,\text{GeV}$.

9. Summary 15

9 Summary

387

388

389

390

391

392

394

395

396

397

398

399

400

401

402

403

404

405

A search for the associated production of dark matter (DM) particles with a Higgs boson decaying into a pair of bottom quarks is presented. No significant deviation from the predictions of the standard model (SM) is observed, and upper limits on the production cross section predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) and the baryonic Z' model are established. They constitute the most stringent collider exclusions placed on the parameters in these models so far. For the nominal choice of the mixing angle $\sin \theta$ and $\tan \beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900 \,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150 \,\text{GeV}$. Scanning over $\sin \theta$ with $\tan \beta = 1$, we exclude $0.35 < \sin \theta < 0.75$ for $m_A = 600$ GeV and $m_{\rm a}=200\,{\rm GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_{\rm A}=600\,{\rm GeV}$ and $m_a = 100 (150) \,\text{GeV}$ and $\sin \theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_{\chi} = 10 \,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The reinterpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross section yields a higher sensitivity for m_{χ} < 5 GeV than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed experimentally for the first time.

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent perfor-406 mance of the LHC and thank the technical and administrative staffs at CERN and at other 407 CMS institutes for their contributions to the success of the CMS effort. In addition, we grate-408 fully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Fi-410 nally, we acknowledge the enduring support for the construction and operation of the LHC 411 and the CMS detector provided by the following funding agencies: BMWFW and FWF (Aus-412 tria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Fin-415 land, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Ger-416 many); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New 419 Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, 420 RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER 421 (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and 422 NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA). 424

Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the

Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Severo Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

References

- [1] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and Constraints", Phys. Rept. 405 (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:0404175.
- [2] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",

 Ann. Rev. Astron. Astrophys. 48 (2010) 495,

 doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- [3] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", Ann. Rev. Astron. Astrophys. 49 (2011) 155,
 doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.
- [4] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron.
 Astrophys. 594 (2016) A13, doi:10.1051/0004-6361/201525830,
 arXiv:1502.01589.
- 454 [5] ATLAS Collaboration, "Observation of a new particle in the search for the standard 455 model higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, 456 doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [6] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [7] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}=7$ and 8 TeV", JHEP **06** (2013) 81, doi:10.1007/JHEP06(2013) 081, arXiv:1303.4571.
- [8] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, arXiv:1311.1511.
- [9] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", *JHEP* **06** (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- [10] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", Phys.
 Rev. D 89 (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- ⁴⁷⁰ [11] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

References 17

472 [12] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05 (2017) 138, arXiv:1701.07427.

- 475 [13] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", JHEP **08** (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
- [14] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10(2017)180, arXiv:1703.05236.
- [16] 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.
- 489 [17] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo 490 algorithms", JHEP 11 (2004) 040, doi:10.1088/1126-6708/2004/11/040, 491 arXiv:hep-ph/0409146.
- [18] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", JHEP 11 (2007) 070,
 doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [19] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* 06 (2010) 043, doi:10.1007/JHEP06 (2010) 043, arXiv:1002.2581.
- [20] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", Phys. Rev. Lett. **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [21] 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:0611129.
- [22] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", JHEP 12 (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* **03** (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- 509 [24] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet 510 production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, 511 Italy, March 21-28, 2015. 2015. arXiv:1505.05704.

- 512 [25] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", JHEP **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- ⁵¹⁵ [26] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", Comput. Phys. Commun. **191** (2015) ⁵¹⁶ 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [27] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", IHEP 07 (2011) 018, doi:10.1007/JHEP07 (2011) 018, arXiv:1105.0020.
- 519 [28] NNPDF Collaboration, "Parton distributions for the LHC Run II", JHEP **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.
- [29] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C 76 (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [30] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur. Phys. J. C 74 (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.
- 527 [31] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- 529 [32] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST 12 (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [33] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.
- 534 [34] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* 595 59 (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.
- CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST 6 (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- 539 [36] 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.
- ⁵⁴¹ [37] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", JINST **13** (2018) 05011.
- [38] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", *JHEP* 12 (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.
- J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05 (2016) 156, arXiv:1603.00027.
- [40] 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.

References 19

[41] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

- [42] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- ⁵⁵⁶ [43] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST 11 (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.
- ⁵⁵⁹ [44] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST **8** (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [45] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.
- [46] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced
 Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.
 arXiv:1009.1003.
- [47] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8 \text{ TeV}$ ", JINST 10 (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- ⁵⁷¹ [48] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- ⁵⁷³ [49] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ ", *Phys. Lett. B* ⁵⁷⁵ **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- 576 [50] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s}=7\,\text{TeV}$ ", Phys. Lett. B 735 (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- 579 [51] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", Phys. Rev. D **91** (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.
- ⁵⁸² [52] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7 \,\text{TeV}$ ", *JHEP* **06** (2014) 120, doi:10.1007/JHEP06 (2014) 120, arXiv:1402.1521.
- CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}=8\,\text{TeV}$ ", Phys. Rev. Lett. 112 (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- 588 [54] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", Phys. Lett. B 763

 590 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

- [55] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 13 \,\text{TeV}$ ", Phys. Lett. B **766** (2017) 268, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.
- [56] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs
 Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working
 Group", Technical Report CERN-2013-004, Geneva, 2013.
 doi:10.5170/CERN-2013-004, arXiv:1307.1347.
- [57] A. L. Read, "Presentation of search results: the CL_s technique", J. Phys. G 28 (2002) 2693,
 doi:10.1088/0954-3899/28/10/313.
- [58] T. Junk, "Confidence level computation for combining searches with small statistics",

 Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002(99)00498-2,

 arXiv:hep-ex/9902006.
- [59] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", Eur. Phys. J. C 71 (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3.
- 606 [60] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse energy signals using simplified *s*-channel models of dark matter", (2016).

 608 arXiv:1603.04156.
- [61] SuperCDMS Collaboration, "New results from the search for low-mass weakly interacting massive particles with the CDMS low ionization threshold experiment", Phys. Rev. Lett. 116 (2016) 071301, doi:10.1103/PhysRevLett.116.071301, arXiv:1509.02448.
- [62] LUX Collaboration, "Results from a search for dark matter in the complete LUX exposure", Phys. Rev. Lett. 118 (2017) 021303,
 doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648.
- [63] XENON Collaboration, "First dark matter search results from the XENON1T experiment", Phys. Rev. Lett. 119 (2017) 181301,
 doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655.
- 619 [64] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II experiment", *Phys. Rev. Lett.* **119** (2017) 181302, doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.
- 622 [65] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold CRESST-II detector", Eur. Phys. J. C 76 (2016) 25, doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.
- Search for associated production of dark matter with a Higgs boson decaying to a pair of bottom quarks in pp collisions at sqrt(s) = 13 TeV
- 627 [cern]The CMS Collaboration
- 628 2018/08/17

A. Introduction 21

629 A Introduction

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for physics beyond the standard model (SM) [456–458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consistent with the hypothesis that DM is primarily composed of weakly interacting massive particles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. Assuming the pair production of DM particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signature. If DM particles are produced in association with a detectable SM particle, which could be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding protons, or through new effective couplings between DM and SM particles, their existence could be inferred via a large transverse momentum imbalance in the collision event.

While ISR production of the SM Higgs boson (h) [460–462] is highly suppressed due to the Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs boson and DM particles can occur if the Higgs boson takes part in the interaction producing the DM particles [463–465]. Such a production mechanism would allow to directly probe the structure of the effective DM-SM coupling.

In this paper, we present a search for DM production in association with an SM Higgs boson that decays into a bottom quark-antiquark pair $(b\bar{b})$. As the $h\to b\bar{b}$ decay mode has the largest branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest signal yield. The search is performed using the data set collected by the CMS experiment [466] at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of approximately 35.9 fb⁻¹. Results are interpreted in terms of two simplified models predicting this signature. The first one is type-2 two Higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi\bar{\chi}$. The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM particles. Representative Feynman diagrams for the two models are presented in Fig. 50.

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no significant deviation from SM predictions [468]. In contrast to previously explored 2HDM models [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar a, assumed to be the observed 125 GeV Higgs boson, and a heavy neutral CP-even scalar a, that are the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mixing angle a; a heavy neutral CP-odd pseudoscalar a and a light neutral CP-odd pseudoscalar a, that are the result of the mixing of the CP-odd mediator a0 with the CP-odd Higgs, with a0 representing the associated mixing angle; and two heavy charged scalars a1 with identical mass.

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the

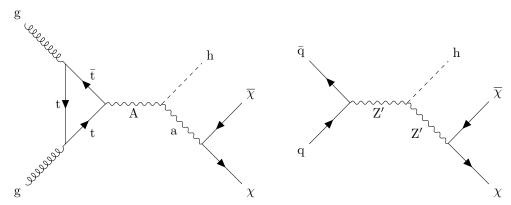


Figure 8: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z' model (right).

charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H}=m_{\rm H^\pm}=m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_\chi=10\,{\rm GeV}$.

The baryonic Z' model [465] is an extension of the SM with an additional $U(1)_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding $U(1)_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the $U(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_B Z' Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin \theta / v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_q , is 0.25 and the Z' coupling to DM, g_{χ} , is set to 1. This is well below the bounds $g_{q_1}g_{\chi} \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin \theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'} \sin \theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and m_{χ} , which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, together with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically independent data samples are used to predict the hadronic recoil distribution for these SM processes that constitute the largest sources of background. Both signal and background contributions to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

B The CMS detector

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum (p_T) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of

712

713

717

718

719

720

721

723

724

725

726

727

728

729

730

732

733

734

735

736

737

738

739

740

741

743

744

745

747

748

749

750

751

3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta| < 3$. The steel and quartz-fiber forward Cherenkov hadron calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and covers $|\eta| < 2.4$. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than $4\,\mu\text{s}$, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to 1 kHz.

C Simulated data samples

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG V2 [472-474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the p_T of the respective boson with NLO QCD corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

D Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_{\rm T}^{i2}$, where $p_{\rm T}^i$ is the transverse momentum of the $i^{\rm th}$ track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, largearea jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central ($\eta < 2.4$). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$ is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , will affect the shape of the distribution of $m_{\rm SD}$ for the background processes. In this search, the value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore desirable to preserve a smoothly falling jet mass distribution as a function of $p_{\rm T}$. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with $p_{\rm T}$, decorrelating the N_2 variable as a function of ρ and $p_{\rm T}$ would be the most appropriate procedure. The decorrelation strategy described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494].

This search also utilizes narrow jets clustered with the anti- k_T algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta|$ < 2.4).

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in

E. Event selection 25

the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the consistency with originating from the primary vertex are imposed. The reconstructed electron is required to be within $|\eta|$ < 2.5, excluding the transition region 1.44 < $|\eta|$ < 1.57 between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons and poorly measured mesons tagged as muons [497]. These additional criteria include requirements on the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates with $|\eta|$ < 2.4 are considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are reconstructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton decays. The τ_h candidates with $p_T > 18$ GeV and $|\eta| < 2.3$ are considered [496, 498, 499]. Photon candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15$ GeV and $|\eta| < 2.5$.

The missing transverse momentum \vec{p}_{T}^{miss} is defined as the negative vectorial sum of the p_{T} of all the reconstructed PF candidates. Its magnitude is denoted as p_{T}^{miss} . Corrections to jet momenta are propagated to the p_{T}^{miss} calculation as well as event filters [500] are used to remove spurious high p_{T}^{miss} events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

E Event selection

800

801

803

804

805

808

809

812

813

814

816

817

818

825

826

827

829

830

831

832

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}^{\rm miss}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T,\text{trig}}^{\text{miss}}$ or H_{T}^{miss} , where $p_{T,\text{trig}}^{\text{miss}}$ (H_{T}^{miss}) is the magnitude of the vectorial (scalar) sum of \vec{p}_{T} of all the particles (jets with $p_{T} > 20\,\text{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{T,\text{trig}}^{\text{miss}}$ calculation. Thresholds on $p_{T,\text{trig}}^{\text{miss}}$ and H_{T}^{miss} are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{T,\text{trig}}^{\text{miss}}$ and H_{T}^{miss} are reffered to as p_{T}^{miss} triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T} > 200\,{\rm GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\rm SD} < 150\,{\rm GeV}$ and $N_2^{\rm DDT} < 0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than $200\,{\rm GeV}$, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (\vec{U}) must be larger than 0.4

Table 4: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, t t , W+jets	0	0	pass
Single-lepton, b-tagged	t t , W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

rad to reject multijet events that mimic signal events. Vetoes on τ_h candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from $t\bar{t}$ events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_T > 10 \,\text{GeV}$ and $|\eta| < 2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between 60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of the trigger efficiency, at least one of the two electrons must have $p_{\rm T} > 40\,{\rm GeV}$, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from W+jets and semileptonic $t\bar{t}$ events. To predict these backgrounds, four single-lepton samples are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target $t\bar{t}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the p_T^{miss} trigger triggers described above, as well as single electron events are selected using the same single-electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and isolation requirements. In addition, samples with a single electron need to have $p_T^{miss} > 50 \, \text{GeV}$ to avoid a large contamination from multijet events.

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

F Signal extraction

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the $p_{\rm T}^{\rm miss}$ and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and

are given by the ratio of the prediction for a given bin in $p_{\rm T}^{\rm miss}$ in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the tt process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt backgrounds take into account the impact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron control samples, the additional requirement on p_T^{miss} . Since the CRs with no b-tagged AK4 jets 891 and a double-b-tagged CA15 jet also have significant contributions from the tt process, transfer 892 factors to predict this contamination from tt events are also imposed between the single-lepton 893 CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contamination from W+jets production in the tt CR with events that fail the double-b tag requirement. 895 Likewise, the Z+jets background prediction in the signal region is connected to the and dilep-896 ton CRs via transfer factors. They account for the difference in the branching fractions of the 897 $Z \rightarrow \nu \nu$ and the $Z \rightarrow \ell \ell$ decays and the impacts of lepton acceptances and selection efficiencies.

G Systematic uncertainties

899

900

902

903

904

905

906

907

908

909

910

913

914

917

918

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncertainties of the search. They can either affect the rate or the shape of $p_{\rm T}^{\rm miss}$ (U) for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal distribution. The list of the systematic uncertainties considered in this search is presented in Table 23. To better estimate their impact on the results, uncertainties from a similar source (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties have been ordered according to decreasing improvement in the expected limit obtained when removing the group from the list of nuisances included in the likelihood fit. The description of each single uncertainty in the text follows the same order.

Scale factors are used to correct for the differences in the double-b tagger misidentification efficiencies between data and prediction from simulation for W/Z+jets production and for $t\bar{t}$ production. These scale factors are measured by simultaneously fitting events that pass or fail the double-b tag requirement. The correlation between the double-b tagger and the p_{T}^{miss} (or U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate independently from each other within a constraint that depends on the recoil value. Such dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. p_{T}^{miss}/U . This shape uncertainty in the double-b scale factors measurement is the one that has the largest impact on the limits on the signal cross section.

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and tt̄ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the $p_{\rm T}$ of the CA15 jet. This shape uncertainty has been derived through a measurement performed using a sample enriched in multijet events with double-muon-tagged $g \to b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substruc-

Table 5: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

· · · · · · · · · · · · · · · · · · ·				
Systematic uncertainty	Type	Processes	Impact on sensitivity	
Double-b mistagging	shape	Z+jets, W+jets, tt̄	4.8%	
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, t t	1.9%	
Double-b tagging	shape	SM h, signal	1.2%	
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0	
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%	
$p_{\rm T}^{\rm miss}$ magnitude	5%	all	0.7%	
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%	
p _T ^{miss} trigger muon multiplicity	shape	Z+jets, W+jets		
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	< 0.5%	
single-electron trigger	1%	all		
AK4 b tagging	shape	all	< 0.5%	
au lepton veto	3%	all	< 0.5%	
Lepton efficiency	1% per lepton	all	< 0.5 /6	
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%	
Renorm./fact. scales	shape	SM h		
PDF	shape	SM h	< 0.5%	
Multijet normalization	100%	multijet	< 0.5 /6	
Theoretical cross section	20%	t, diboson		

ture variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes where the decay of a resonance inside the CA15 jet cone is expected. Such processes include signal production together with SM h and diboson production. The uncertainty has been derived from the efficiency measurement obtained by performing a fit in a control sample enriched in semi-leptonic $t\bar{t}$ events, where the CA15 jet originates from the W boson that comes from the hadronically decaying top quark.

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in $p_{\rm T}^{\rm miss}$ magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, obtained by varying those parameters in simulation event-by-event

The $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. Differences on the order of 2% are observed between single-muon and dimuon events at lower U values and they are sources of an additional systematic uncertainty in the transfer factors for those processes whose prediction relies on data events in the single-muon and dimuon CRs (t̄t̄, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the identified muon, they can change the shape of the U distribution and are thus treated as shape uncertainties. The $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ trigger efficiency is parametrized as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$. The uncertainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies of the single-electron triggers are parametrized as a function of the electron p_{T} and η and an

H. Results 29

949 associated 1% systematic uncertainty is added into the fit.

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization scales and PDFs are included as shape variations. Being a negligible background source, an uncertainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by requiring $p_{\rm T}^{\rm miss} > 250\,{\rm GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty in QCD multijet yields in the signal region, while separate nuisance parameters are introduced for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as measured by CMS in Refs. [509, 510] and correlated across the SR and CRs.

H Results

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficiencies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z' model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results of this search are interpreted in terms of upper limits on the signal strength modifier $\mu =$

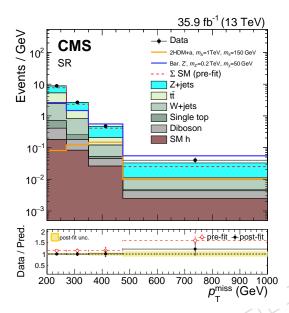


Figure 9: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

Table 6: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

p _T ^{miss} -bin	200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
Z+jets	248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
tŧ	199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
W+jets	121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
Single top quark	21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
Diboson	16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
SM h	12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
Σ (SM)	619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
Data	619	214	59	21
2 HDM+ a , $m_{\rm A} = 1$ TeV, $m_{\rm a} = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z' , $m_{Z'} = 0.2 \text{TeV}$, $m_{\chi} = 50 \text{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

H. Results 31

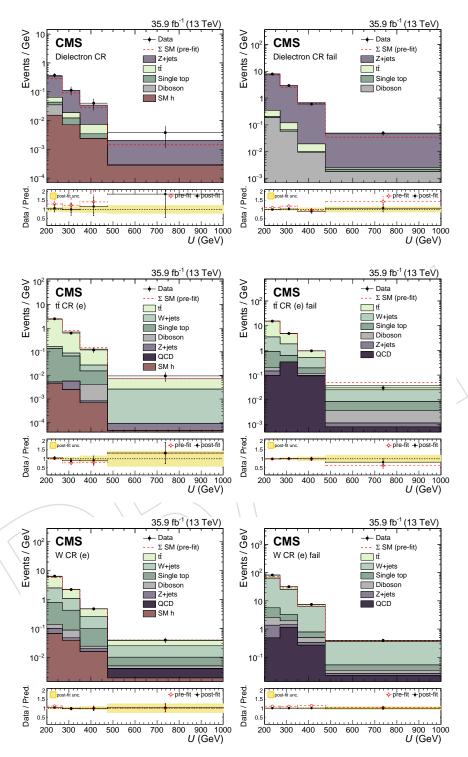


Figure 10: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

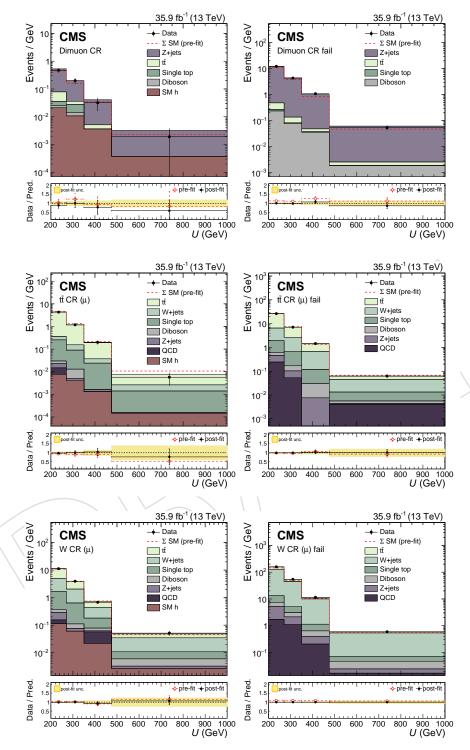


Figure 11: The *U* distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

H. Results

 $\sigma/\sigma_{\text{theory}}$, where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

Figure 54 shows the upper limits on μ for the three scans (m_A , $\sin \theta$, and $\tan \beta$) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, $\sin \theta = 0.35$ and $\tan \beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming $\tan \beta = 1$. Also excluded are $\tan \beta$ values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

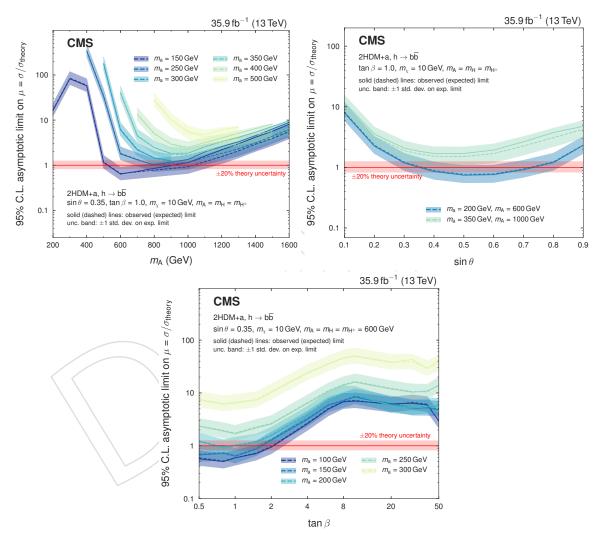


Figure 12: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus.

1011

1013

1014

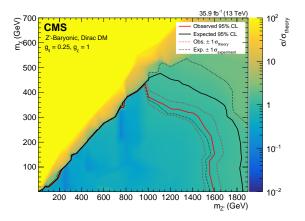


Figure 13: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

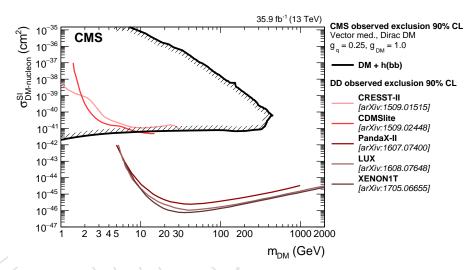


Figure 14: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_{χ} . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2\mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},$$
 (2)

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5\,\text{GeV}$.

I. Summary 35

I Summary

1015

1016

1017

1018

1019

1020

1022

1023

1024

1026

1027

1028

1029

1031

1032

1033

1034

1035

1036

1038

1039

1040

1041

1043

1044

1045

1047

1048

1049

1050

1051

1052

1053

1054

1055

1056

1058

1059

A search for the associated production of dark matter (DM) particles with a Higgs boson decaying into a pair of bottom quarks is presented. No significant deviation from the predictions of the standard model (SM) is observed, and upper limits on the production cross section predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) and the baryonic Z' model are established. They constitute the most stringent collider exclusions placed on the parameters in these models so far. For the nominal choice of the mixing angle $\sin \theta$ and $\tan \beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900 \,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150 \,\text{GeV}$. Scanning over $\sin \theta$ with $\tan \beta = 1$, we exclude $0.35 < \sin \theta < 0.75$ for $m_A = 600$ GeV and $m_{\rm a}=200\,{\rm GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_{\rm A}=600\,{\rm GeV}$ and $m_a = 100 (150) \,\text{GeV}$ and $\sin \theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_{\chi} = 10 \,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The reinterpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross section yields a higher sensitivity for m_{χ} < 5 GeV than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed experimentally for the first time.

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 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 and the CMS detector provided by the following funding agencies: BMWFW and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the

Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Severo Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

References

1069

- [66] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and
 Constraints", Phys. Rept. 405 (2005) 279, doi:10.1016/j.physrep.2004.08.031,
 arXiv:0404175.
- [67] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",
 Ann. Rev. Astron. Astrophys. 48 (2010) 495,
 doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- 1076 [68] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", Ann. Rev. Astron. Astrophys. 49 (2011) 155,

 doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.
- [69] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron.

 Astrophys. 594 (2016) A13, doi:10.1051/0004-6361/201525830,

 arXiv:1502.01589.
- [70] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- 1085 [71] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [72] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}=7$ and 8 TeV", *JHEP* **06** (2013) 81, doi:10.1007/JHEP06(2013) 081, arXiv:1303.4571.
- [73] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, arXiv:1311.1511.
- [74] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", JHEP 06 (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- [75] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", Phys.
 Rev. D 89 (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- [76] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.
- 1100 [77] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05 (2017) 138, arXiv:1701.07427.
- 1103 [78] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis

References 37

```
of the LHC pp collision data at \sqrt{s}=7 and 8 TeV", JHEP 08 (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
```

- [79] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- [80] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10 (2017) 180, arXiv:1703.05236.
- [81] 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.
- [82] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- 1120 [83] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [84] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- 1126 [85] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", Phys. Rev. Lett. 110 (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [86] 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:0611129.
- [87] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", JHEP **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [88] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* **03** (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- [89] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, Italy, March 21-28, 2015. 2015. arXiv:1505.05704.
- [90] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", JHEP **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- [91] 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.
- [92] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07 (2011) 018, arXiv:1105.0020.
- [93] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.

- [94] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [95] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur.
 Phys. J. C 74 (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y,
 arXiv:1404.5630.
- [96] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- [97] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST 12 (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- 1160 [98] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-IME-09-001, 2009.
- [99] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* 59 (2014) 059, doi:10.1007/JHEP10 (2014) 059, arXiv:1407.6013.
- [100] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST 6 (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- [101] 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.
- [102] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) 05011.
- [103] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", *JHEP* 1172 12 (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.
- 1173 [104] J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05 (2016) 156, arXiv:1603.00027.
- [105] 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.
- [106] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}=8$ TeV", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [107] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [108] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST 11 (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.
- [109] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST 8 (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [110] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.

References 39

1193 [111] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced
1194 Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.
1195 arXiv:1009.1003.

- [112] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8 \text{ TeV}$ ", JINST **10** (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [113] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [114] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* 741 (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- [115] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s} = 7 \,\text{TeV}$ ", *Phys. Lett. B* **735** (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- [116] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", Phys. Rev. D **91** (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.
- [117] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7 \, \text{TeV}$ ", JHEP **06** (2014) 120, doi:10.1007/JHEP06 (2014) 120, arXiv:1402.1521.
- [118] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}=8$ TeV", Phys. Rev. Lett. 112 (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- [119] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", Phys. Lett. B 763 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.
- [120] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 13 \,\text{TeV}$ ", *Phys. Lett. B* **766** (2017) 268, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.
- [121] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group", Technical Report CERN-2013-004, Geneva, 2013.

 doi:10.5170/CERN-2013-004, arXiv:1307.1347.
- [122] A. L. Read, "Presentation of search results: the CL_s technique", J. Phys. G **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.
- [123] T. Junk, "Confidence level computation for combining searches with small statistics",

 Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002(99)00498-2,

 arXiv:hep-ex/9902006.
- [124] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", Eur. Phys. J. C 71 (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3.
- [123] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse energy signals using simplified *s*-channel models of dark matter", (2016).

 arXiv:1603.04156.
- [126] SuperCDMS Collaboration, "New results from the search for low-mass weakly interacting massive particles with the CDMS low ionization threshold experiment",

```
Phys. Rev. Lett. 116 (2016) 071301, doi:10.1103/PhysRevLett.116.071301,
1239
         arXiv:1509.02448.
    [127] LUX Collaboration, "Results from a search for dark matter in the complete LUX
1241
         exposure", Phys. Rev. Lett. 118 (2017) 021303,
1242
         doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648.
1243
    [128] XENON Collaboration, "First dark matter search results from the XENON1T
1244
         experiment", Phys. Rev. Lett. 119 (2017) 181301,
1245
         doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655.
1246
    [129] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II
1247
         experiment", Phys. Rev. Lett. 119 (2017) 181302,
1248
         doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.
1249
    [130] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold
1250
         CRESST-II detector", Eur. Phys. J. C 76 (2016) 25,
1251
         doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.
1252
    Search for associated production of dark matter with a Higgs boson decaying to a pair of bot-
1253
    tom quarks in pp collisions at sqrt(s) = 13 \text{ TeV}
1254
    [cern]The CMS Collaboration
1255
    2018/08/17
1256
```

A Introduction

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for physics beyond the standard model (SM) [456–458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consistent with the hypothesis that DM is primarily composed of weakly interacting massive particles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. Assuming the pair production of DM particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signature. If DM particles are produced in association with a detectable SM particle, which could be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding protons, or through new effective couplings between DM and SM particles, their existence could be inferred via a large transverse momentum imbalance in the collision event.

While ISR production of the SM Higgs boson (h) [460–462] is highly suppressed due to the Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs boson and DM particles can occur if the Higgs boson takes part in the interaction producing the DM particles [463–465]. Such a production mechanism would allow to directly probe the structure of the effective DM-SM coupling.

In this paper, we present a search for DM production in association with an SM Higgs boson that decays into a bottom quark-antiquark pair $(b\bar{b})$. As the $h\to b\bar{b}$ decay mode has the largest branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest signal yield. The search is performed using the data set collected by the CMS experiment [466] at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of approximately 35.9 fb⁻¹. Results are interpreted in terms of two simplified models predicting this signature. The first one is type-2 two Higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar

A. Introduction 41

and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi\bar{\chi}$. The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM particles. Representative Feynman diagrams for the two models are presented in Fig. 50.

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no significant deviation from SM predictions [468]. In contrast to previously explored 2HDM models [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar a, that are the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mixing angle a; a heavy neutral CP-odd pseudoscalar a, and a light neutral CP-odd pseudoscalar a, that are the result of the mixing of the CP-odd mediator a0 with the CP-odd Higgs, with a1 representing the associated mixing angle; and two heavy charged scalars a2 with identical mass.

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H} = m_{\rm H^{\pm}} = m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,{\rm GeV}$.

The baryonic Z' model [465] is an extension of the SM with an additional $U(1)_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding $U(1)_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the $U(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_BZ'Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin\theta/v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_q , is 0.25 and the Z' coupling to DM, g_χ , is set to 1. This is well below the bounds g_q , $g_\chi \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin\theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'}\sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and m_χ , which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, together with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically independent data samples are used to predict the hadronic recoil distribution for these SM processes

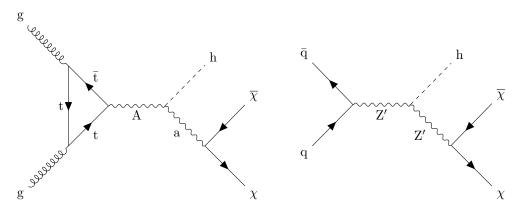


Figure 15: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z' model (right).

that constitute the largest sources of background. Both signal and background contributions to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

B The CMS detector

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum ($p_{\rm T}$) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of 3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta| < 3$. The steel and quartz-fiber forward Cherenkov hadron calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and covers $|\eta| < 2.4$. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than $4\,\mu s$, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to 1 kHz.

C Simulated data samples

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG v2 [472–474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the p_T of the respective boson with NLO QCD

D. Event reconstruction 43

corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

D Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_T^{i2}$, where p_T^i is the transverse momentum of the i^{th} track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, large-area jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central (η < 2.4). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$

1429

1430

1431

1432

1433

1434

1435

1436

1437

1438

1439

1440

1442

1443

1444

1445

is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , 1410 will affect the shape of the distribution of m_{SD} for the background processes. In this search, the value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection 1412 of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore 1413 desirable to preserve a smoothly falling jet mass distribution as a function of p_T . As explained 1414 in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_T , decorrelating the N_2 variable 1416 as a function of ρ and p_T would be the most appropriate procedure. The decorrelation strategy 1417 described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds 1418 to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494]. 1420

This search also utilizes narrow jets clustered with the anti- $k_{\rm T}$ algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta|$ < 2.4).

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the consistency with originating from the primary vertex are imposed. The reconstructed electron is required to be within $|\eta| < 2.5$, excluding the transition region 1.44 $< |\eta| < 1.57$ between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons and poorly measured mesons tagged as muons [497]. These additional criteria include requirements on the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates with $|\eta|$ < 2.4 are considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are reconstructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton decays. The τ_h candidates with $p_T > 18 \,\text{GeV}$ and $|\eta| < 2.3$ are considered [496, 498, 499]. Photon candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15$ GeV and $|\eta| < 2.5$.

The missing transverse momentum $\vec{p}_{\rm T}^{\rm miss}$ is defined as the negative vectorial sum of the $p_{\rm T}$ of all the reconstructed PF candidates. Its magnitude is denoted as $p_{\rm T}^{\rm miss}$. Corrections to jet momenta are propagated to the $p_{\rm T}^{\rm miss}$ calculation as well as event filters [500] are used to remove spurious high $p_{\rm T}^{\rm miss}$ events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

E. Event selection 45

Table 7: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, t t , W+jets	0	0	pass
Single-lepton, b-tagged	tī, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

E Event selection

1452

1453

1454

1455

1457

1458

1459

1460

1461

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T, \rm trig}^{\rm miss}$ or $H_{\rm T}^{\rm miss}$, where $p_{\rm T, trig}^{\rm miss}$ ($H_{\rm T}^{\rm miss}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\rm T}$ of all the particles (jets with $p_{\rm T} > 20\,{\rm GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\rm T, trig}^{\rm miss}$ calculation. Thresholds on $p_{\rm T, trig}^{\rm miss}$ and $H_{\rm T}^{\rm miss}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\rm T, trig}^{\rm miss}$ and $H_{\rm T}^{\rm miss}$ are reffered to as $p_{\rm T}^{\rm miss}$ triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T} > 200\,{\rm GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\rm SD} < 150\,{\rm GeV}$ and $N_2^{\rm DDT} < 0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than $200\,{\rm GeV}$, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (U) must be larger than 0.4 rad to reject multijet events that mimic signal events. Vetoes on $\tau_{\rm h}$ candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from $t\bar{t}$ events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_{\rm T}>10\,{\rm GeV}$ and $|\eta|<2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between 60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of

the trigger efficiency, at least one of the two electrons must have $p_T > 40 \,\text{GeV}$, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from W+jets and semileptonic tt events. To predict these backgrounds, four single-lepton samples are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target tt events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the $p_{\rm T}^{\rm miss}$ trigger triggers described above, as well as single electron events are selected using the same single-electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and isolation requirements. In addition, samples with a single electron need to have $p_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ to avoid a large contamination from multijet events.

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

F Signal extraction

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the p_T^{miss} and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and are given by the ratio of the prediction for a given bin in p_T^{miss} in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the $t\bar{t}$ process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt̄ backgrounds take into account the impact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron control samples, the additional requirement on p_T^{miss} . Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the tt̄ process, transfer factors to predict this contamination from tt̄ events are also imposed between the single-lepton CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contamination from W+jets production in the tt̄ CR with events that fail the double-b tag requirement. Likewise, the Z+jets background prediction in the signal region is connected to the and dilepton CRs via transfer factors. They account for the difference in the branching fractions of the $Z \rightarrow \nu \nu$ and the $Z \rightarrow \ell \ell$ decays and the impacts of lepton acceptances and selection efficiencies.

G Systematic uncertainties

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncertainties of the search. They can either affect the rate or the shape of $p_T^{miss}(U)$ for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal

Table 8: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

Systematic uncertainty	Туре	Processes	Impact on sensitivity	
Double-b mistagging	shape	Z+jets, W+jets, tt	4.8%	
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, tt̄	1.9%	
Double-b tagging	shape	SM h, signal	1.2%	
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0	
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%	
$p_{\rm T}^{\rm miss}$ magnitude	5%	all	0.7%	
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%	
p _T ^{miss} trigger muon multiplicity	shape	Z+jets, W+jets		
$p_{\mathrm{T}}^{\mathrm{\hat{m}iss}}$ trigger efficiency	1%	all	< 0.5%	
single-electron trigger	1%	all		
AK4 b tagging	shape	all	< 0.5%	
au lepton veto	3%	all	< 0.5%	
Lepton efficiency	1% per lepton	all	V 0.570	
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%	
Renorm./fact. scales	shape	SM h	()	
PDF	shape	SM h	< 0.5%	
Multijet normalization	100%	multijet	0.5%	
Theoretical cross section	20%	t, diboson		

distribution. The list of the systematic uncertainties considered in this search is presented in Table 23. To better estimate their impact on the results, uncertainties from a similar source (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties have been ordered according to decreasing improvement in the expected limit obtained when removing the group from the list of nuisances included in the likelihood fit. The description of each single uncertainty in the text follows the same order.

Scale factors are used to correct for the differences in the double-b tagger misidentification efficiencies between data and prediction from simulation for W/Z+jets production and for $t\bar{t}$ production. These scale factors are measured by simultaneously fitting events that pass or fail the double-b tag requirement. The correlation between the double-b tagger and the p_{T}^{miss} (or U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate independently from each other within a constraint that depends on the recoil value. Such dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. p_{T}^{miss}/U . This shape uncertainty in the double-b scale factors measurement is the one that has the largest impact on the limits on the signal cross section.

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and tt̄ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the $p_{\rm T}$ of the CA15 jet. This shape uncertainty has been derived through a measurement performed using a sample enriched in multijet events with double-muon-tagged $g \to b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substructure variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes where the decay of a resonance inside the CA15 jet cone is expected. Such processes include sig-

nal production together with SM h and diboson production. The uncertainty has been derived from the efficiency measurement obtained by performing a fit in a control sample enriched in semi-leptonic tt events, where the CA15 jet originates from the W boson that comes from the hadronically decaying top quark.

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in p_T^{miss} magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, obtained by varying those parameters in simulation event-by-event

The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. 1568 Differences on the order of 2% are observed between single-muon and dimuon events at lower 1569 U values and they are sources of an additional systematic uncertainty in the transfer factors 1570 for those processes whose prediction relies on data events in the single-muon and dimuon CRs 1571 (tt, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the 1572 identified muon, they can change the shape of the U distribution and are thus treated as shape 1573 uncertainties. The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiency is parametrized as a function of p_{T}^{miss} . The uncer-1574 tainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies 1575 of the single-electron triggers are parametrized as a function of the electron p_T and η and an 1576 associated 1% systematic uncertainty is added into the fit. 1577

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization scales and PDFs are included as shape variations. Being a negligible background source, an uncertainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by requiring $p_{\rm T}^{\rm miss} > 250$ GeV, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty in QCD multijet yields in the signal region, while separate nuisance parameters are introduced for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single

H. Results 49

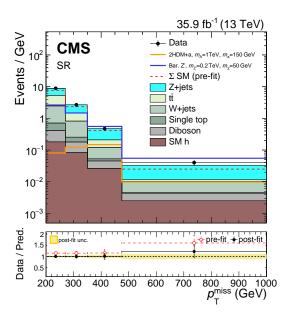


Figure 16: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as measured by CMS in Refs. [509, 510] and correlated across the SR and CRs.

H Results

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficiencies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z' model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results of this search are interpreted in terms of upper limits on the signal strength modifier $\mu = \sigma/\sigma_{\text{theory}}$, where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

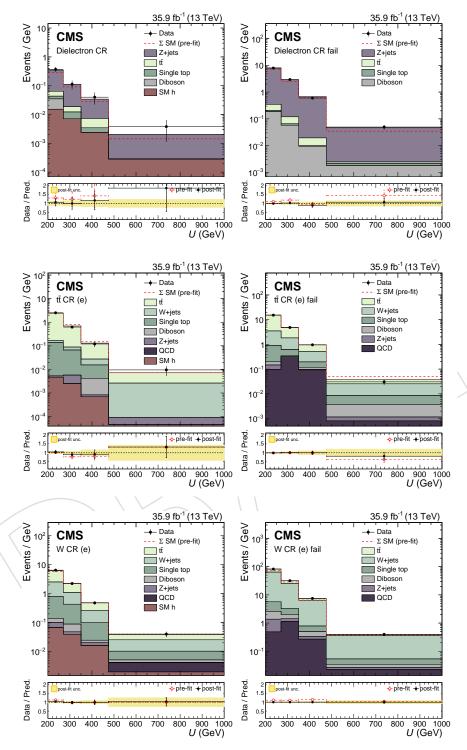


Figure 17: The *U* distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

H. Results 51

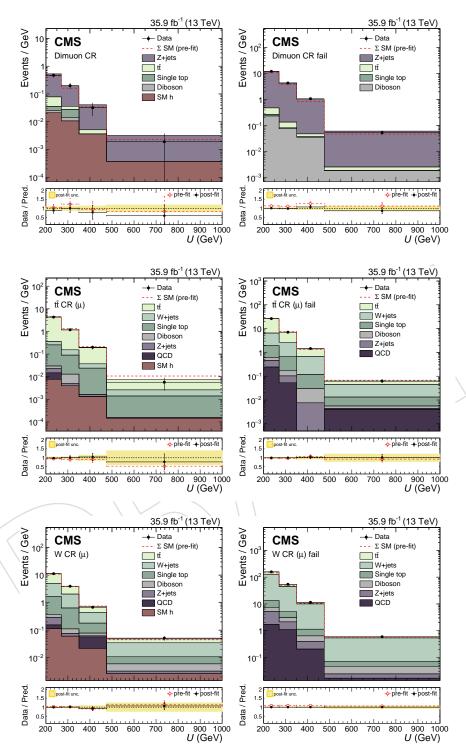


Figure 18: The *U* distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

Table 9: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

$p_{\mathrm{T}}^{\mathrm{miss}}$ -bin	200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
Z+jets	248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
tŧ	199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
W+jets	121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
Single top quark	21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
Diboson	16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
SM h	12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
Σ (SM)	619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
Data	619	214	59	21
2HDM+ a , $m_{\rm A} = 1 \text{TeV}$, $m_{\rm a} = 150 \text{GeV}$	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2 \text{TeV}$, $m_{\chi} = 50 \text{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

Figure 54 shows the upper limits on μ for the three scans (m_A , $\sin \theta$, and $\tan \beta$) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, $\sin \theta = 0.35$ and $\tan \beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming $\tan \beta = 1$. Also excluded are $\tan \beta$ values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus. Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2 \mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},$$
(3)

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5\,\text{GeV}$.

I Summary

A search for the associated production of dark matter (DM) particles with a Higgs boson decaying into a pair of bottom quarks is presented. No significant deviation from the predictions of the standard model (SM) is observed, and upper limits on the production cross section predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar boson *a* (2HDM+*a*) and the baryonic Z' model are established. They constitute the most stringent collider exclusions placed on the parameters in these models so far. For the nominal

I. Summary 53

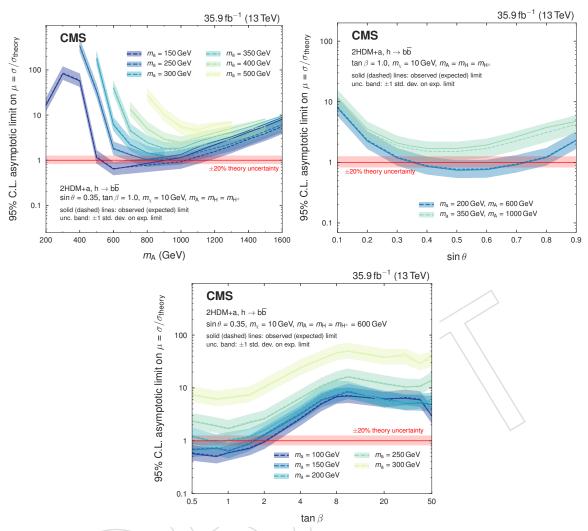


Figure 19: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

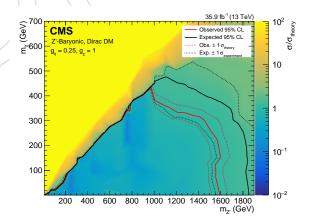


Figure 20: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

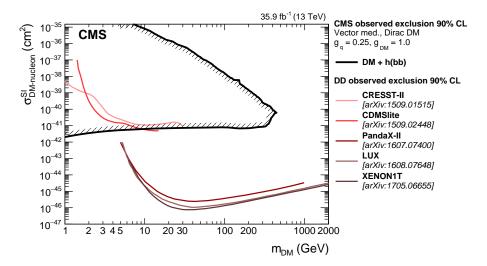


Figure 21: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_{χ} . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

choice of the mixing angle $\sin\theta$ and $\tan\beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900\,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150\,\text{GeV}$. Scanning over $\sin\theta$ with $\tan\beta = 1$, we exclude $0.35 < \sin\theta < 0.75$ for $m_A = 600\,\text{GeV}$ and $m_a = 200\,\text{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_A = 600\,\text{GeV}$ and $m_a = 100\,$ (150) GeV and $\sin\theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_\chi = 10\,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The reinterpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross section yields a higher sensitivity for $m_\chi < 5\,\text{GeV}$ than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed experimentally for the first time.

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 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 and the CMS detector provided by the following funding agencies: BMWFW and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New

References 55

Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie program and the European Research 1681 Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foun-1682 dation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Fed-1683 eral Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans 1684 l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie 1685 (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the 1686 Council of Science and Industrial Research, India; the HOMING PLUS program of the Foun-1687 dation for Polish Science, cofinanced from European Union, Regional Development Fund, the Mobility Plus program of the Ministry of Science and Higher Education, the National Science 1689 Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 1690 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the 1691 National Priorities Research Program by Qatar National Research Fund; the Programa Severo 1692 Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn 1694 University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project 1695 (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA). 1696

References

1697

- [131] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and Constraints", *Phys. Rept.* **405** (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:0404175.
- [132] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",
 Ann. Rev. Astron. Astrophys. 48 (2010) 495,
 doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- 1704 [133] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", *Ann. Rev. Astron. Astrophys.* **49** (2011) 155, doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.
- 1707 [134] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron.
 1708 Astrophys. 594 (2016) A13, doi:10.1051/0004-6361/201525830,
 1709 arXiv:1502.01589.
- 1710 [135] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [136] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [137] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", JHEP **06** (2013) 81, doi:10.1007/JHEP06 (2013) 081, arXiv:1303.4571.
- [138] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, arXiv:1311.1511.

- [139] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", *JHEP* **06** (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- [140] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", *Phys.*1725 *Rev. D* **89** (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- [141] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.
- [142] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05 (2017) 138, arXiv:1701.07427.
- [143] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}=7$ and 8 TeV", JHEP **08** (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
- 1735 [144] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- [145] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to bb or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10 (2017) 180, arXiv:1703.05236.
- [146] 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.
- 1745 [147] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* 11 (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- 1748 [148] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [149] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [150] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", Phys. Rev. Lett. 110 (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [151] 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:0611129.
- [152] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12 (2012) 061, arXiv:1209.6215.
- [153] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* **03** (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- 1765 [154] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet 1766 production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, 1767 Italy, March 21-28, 2015. 2015. arXiv:1505.05704.

References 57

- [155] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", *JHEP* **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- [156] 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.
- [157] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [1774] *JHEP* **07** (2011) 018, doi:10.1007/JHEP07 (2011) 018, arXiv:1105.0020.
- 1775 [158] NNPDF Collaboration, "Parton distributions for the LHC Run II", JHEP **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.
- [159] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C 76 (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [160] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur. Phys. J. C 74 (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.
- 1783 [161] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A 506 (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- [162] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST 12 (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [163] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.
- [164] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* 59 (2014) 059, doi:10.1007/JHEP10 (2014) 059, arXiv:1407.6013.
- [165] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST 6 (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- [166] 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.
- [167] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", JINST 13 (2018) 05011.
- [168] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", *JHEP* **12** (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.
- [169] J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05 (2016) 156, arXiv:1603.00027.
- [170] 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.
- [171] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [172] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [173] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST 11 (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.

- [174] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST 8 (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [175] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.
- [176] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.

 arXiv:1009.1003.
- [177] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8 \text{ TeV}$ ", JINST 10 (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [178] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [179] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ ", *Phys. Lett. B* **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- [180] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ ", Phys. Lett. B **735** (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- [181] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", Phys. Rev. D **91** (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.
- [182] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s}=7\,\text{TeV}$ ", JHEP **06** (2014) 120, doi:10.1007/JHEP06 (2014) 120, arXiv:1402.1521.
- [183] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ ", Phys. Rev. Lett. 112 (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- [184] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", Phys. Lett. B 763 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.
- [185] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", *Phys. Lett. B* **766** (2017) 268, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.
- [186] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs
 Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working
 Group", Technical Report CERN-2013-004, Geneva, 2013.
 doi:10.5170/CERN-2013-004, arXiv:1307.1347.
- 1854 [187] A. L. Read, "Presentation of search results: the CL_s technique", J. Phys. G **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.
- [188] T. Junk, "Confidence level computation for combining searches with small statistics",

 Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002(99)00498-2,

 arXiv:hep-ex/9902006.
- [189] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", Eur. Phys. J. C **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3.

A. Introduction 59

```
[190] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse
         energy signals using simplified s-channel models of dark matter", (2016).
1863
         arXiv:1603.04156.
1864
    [191] SuperCDMS Collaboration, "New results from the search for low-mass weakly
1865
         interacting massive particles with the CDMS low ionization threshold experiment",
1866
         Phys. Rev. Lett. 116 (2016) 071301, doi:10.1103/PhysRevLett.116.071301,
1867
         arXiv:1509.02448.
1868
    [192] LUX Collaboration, "Results from a search for dark matter in the complete LUX
1869
         exposure", Phys. Rev. Lett. 118 (2017) 021303,
         doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648.
1871
    [193] XENON Collaboration, "First dark matter search results from the XENON1T
1872
         experiment", Phys. Rev. Lett. 119 (2017) 181301,
1873
         doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655.
1874
    [194] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II
1875
         experiment", Phys. Rev. Lett. 119 (2017) 181302,
1876
         doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.
1877
    [195] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold
1878
         CRESST-II detector", Eur. Phys. J. C 76 (2016) 25,
1879
         doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.
1880
    Search for associated production of dark matter with a Higgs boson decaying to a pair of bot-
1881
    tom quarks in pp collisions at sqrt(s) = 13 \text{ TeV}
    [cern]The CMS Collaboration
1883
    2018/08/17
1884
```

Introduction Α

1885

1886

1887

1888

1889

1890

1891

1892

1893

1894

1895

1896

1862

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for physics beyond the standard model (SM) [456-458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consistent with the hypothesis that DM is primarily composed of weakly interacting massive particles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. Assuming the pair production of DM particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signature. If DM particles are produced in association with a detectable SM particle, which could be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding protons, or through new effective couplings between DM and SM particles, their existence could be inferred via a large transverse momentum imbalance in the collision event.

While ISR production of the SM Higgs boson (h) [460-462] is highly suppressed due to the 1898 Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs 1899 boson and DM particles can occur if the Higgs boson takes part in the interaction producing 1900 the DM particles [463–465]. Such a production mechanism would allow to directly probe the 1901 structure of the effective DM-SM coupling. 1902

In this paper, we present a search for DM production in association with an SM Higgs boson 1903 that decays into a bottom quark-antiquark pair (bb). As the h \rightarrow bb decay mode has the largest 1904 branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest 1905

signal yield. The search is performed using the data set collected by the CMS experiment [466] at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of approximately $35.9 \, \mathrm{fb}^{-1}$. Results are interpreted in terms of two simplified models predicting this signature. The first one is type-2 two Higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi\bar{\chi}$. The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM particles. Representative Feynman diagrams for the two models are presented in Fig. 50.

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no significant deviation from SM predictions [468]. In contrast to previously explored 2HDM models [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar a, assumed to be the observed 125 GeV Higgs boson, and a heavy neutral CP-even scalar a, that are the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mixing angle a; a heavy neutral CP-odd pseudoscalar a and a light neutral CP-odd pseudoscalar a, that are the result of the mixing of the CP-odd mediator a0 with the CP-odd Higgs, with a1 representing the associated mixing angle; and two heavy charged scalars a2 with identical mass.

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H} = m_{\rm H^{\pm}} = m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,{\rm GeV}$.

The baryonic Z' model [465] is an extension of the SM with an additional U(1) $_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding U(1) $_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the U(1) $_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_BZ'Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin\theta/v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_q , is 0.25 and the Z' coupling to DM, g_χ , is set to 1. This is well below the bounds g_q , $g_\chi \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin\theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'}\sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and m_χ , which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, to-

B. The CMS detector 61

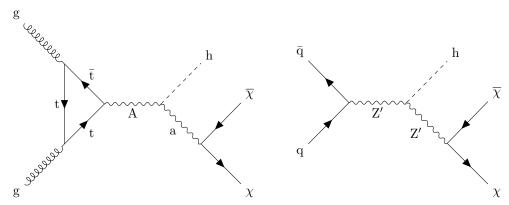


Figure 22: Feynman diagrams for the 2HDM+*a* model (left) and the baryonic Z' model (right).

gether with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically independent data samples are used to predict the hadronic recoil distribution for these SM processes that constitute the largest sources of background. Both signal and background contributions to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

B The CMS detector

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum ($p_{\rm T}$) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of 3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta| < 3$. The steel and quartz-fiber forward Cherenkov hadron calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and covers $|\eta| < 2.4$. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than $4\,\mu\rm s$, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to $1\,\rm kHz$.

C Simulated data samples

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG V2 [472–474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four

partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the p_T of the respective boson with NLO QCD corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

D Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_T^{i2}$, where p_T^i is the transverse momentum of the i^{th} track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, largearea jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central ($\eta < 2.4$). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

D. Event reconstruction 63

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$ is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , will affect the shape of the distribution of $m_{\rm SD}$ for the background processes. In this search, the value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore desirable to preserve a smoothly falling jet mass distribution as a function of $p_{\rm T}$. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with $p_{\rm T}$, decorrelating the N_2 variable as a function of $p_{\rm T}$ and $p_{\rm T}$ would be the most appropriate procedure. The decorrelation strategy described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494].

This search also utilizes narrow jets clustered with the anti- $k_{\rm T}$ algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta| < 2.4$).

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the consistency with originating from the primary vertex are imposed. The reconstructed electron is required to be within $|\eta|$ < 2.5, excluding the transition region 1.44 < $|\eta|$ < 1.57 between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons and poorly measured mesons tagged as muons [497]. These additional criteria include requirements on the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates with $|\eta|$ < 2.4 are considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are reconstructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton decays. The τ_h candidates with $p_T > 18$ GeV and $|\eta| < 2.3$ are considered [496, 498, 499]. Photon candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15$ GeV and $|\eta| < 2.5$.

The missing transverse momentum \vec{p}_T^{miss} is defined as the negative vectorial sum of the p_T of all the reconstructed PF candidates. Its magnitude is denoted as p_T^{miss} . Corrections to jet momenta

Table 10: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, t t , W+jets	0	0	pass
Single-lepton, b-tagged	t t , W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

are propagated to the $p_{\rm T}^{\rm miss}$ calculation as well as event filters [500] are used to remove spurious high $p_{\rm T}^{\rm miss}$ events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

E Event selection

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T, \rm trig}^{\rm miss}$ or $H_{\rm T}^{\rm miss}$, where $p_{\rm T, trig}^{\rm miss}$ ($H_{\rm T}^{\rm miss}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\rm T}$ of all the particles (jets with $p_{\rm T} > 20\,{\rm GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\rm T, trig}^{\rm miss}$ calculation. Thresholds on $p_{\rm T, trig}^{\rm miss}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\rm T, trig}^{\rm miss}$ and $H_{\rm T}^{\rm miss}$ are reffered to as $p_{\rm T}^{\rm miss}$ triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T}>200\,{\rm GeV}$ and $|\eta|<2.4$ is required. It is also required that $100< m_{\rm SD}<150\,{\rm GeV}$ and $N_2^{\rm DDT}<0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than $200\,{\rm GeV}$, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (\vec{U}) must be larger than 0.4 rad to reject multijet events that mimic signal events. Vetoes on $\tau_{\rm h}$ candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from $t\bar{t}$ events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_{\rm T}>10\,{\rm GeV}$ and $|\eta|<2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between

F. Signal extraction 65

60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of the trigger efficiency, at least one of the two electrons must have $p_{\rm T} > 40$ GeV, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from 2119 W+jets and semileptonic tt events. To predict these backgrounds, four single-lepton samples 2120 are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target tt events, while the other two 2122 single-lepton CRs target W+jets events. Single-muon events are selected using the p_T^{miss} trigger 2123 triggers described above, as well as single electron events are selected using the same single-2124 electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and 2126 isolation requirements. In addition, samples with a single electron need to have $p_{\rm T}^{\rm miss} > 50\,{\rm GeV}$ 2127 to avoid a large contamination from multijet events. 2128

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

F Signal extraction

2133

2134

2135

2136

2137

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the $p_{\rm T}^{\rm miss}$ and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and are given by the ratio of the prediction for a given bin in $p_{\rm T}^{\rm miss}$ in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the $t\bar{t}$ process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt backgrounds take into account the im-2145 pact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron 2146 control samples, the additional requirement on p_T^{miss} . Since the CRs with no b-tagged AK4 jets 2147 and a double-b-tagged CA15 jet also have significant contributions from the tt process, transfer 2148 factors to predict this contamination from tt events are also imposed between the single-lepton CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contam-2150 ination from W+jets production in the tt CR with events that fail the double-b tag requirement. 2151 Likewise, the Z+jets background prediction in the signal region is connected to the and dilep-2152 ton CRs via transfer factors. They account for the difference in the branching fractions of the $Z \rightarrow \nu \nu$ and the $Z \rightarrow \ell \ell$ decays and the impacts of lepton acceptances and selection efficiencies.

Table 11: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

Systematic uncertainty	Туре	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, tt̄	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, tt̄	1.9%
Double-b tagging	shape	SM h, signal	1.2%
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\rm T}^{\rm miss}$ magnitude	5%	all	0.7%
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger muon multiplicity	shape	Z+jets, W+jets	
$p_{\mathrm{T}}^{\mathrm{\hat{m}iss}}$ trigger efficiency	1%	all	< 0.5%
single-electron trigger	1%	all	
AK4 b tagging	shape	all	< 0.5%
au lepton veto	3%	all	< 0.5%
Lepton efficiency	1% per lepton	all	< 0.5 /6
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%
Renorm./fact. scales	shape	SM h	. \
PDF	shape	SM h	< 0.5%
Multijet normalization	100%	multijet	< 0.5%
Theoretical cross section	20%	t, diboson	

G Systematic uncertainties

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncertainties of the search. They can either affect the rate or the shape of $p_{\rm T}^{\rm miss}$ (U) for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal distribution. The list of the systematic uncertainties considered in this search is presented in Table 23. To better estimate their impact on the results, uncertainties from a similar source (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties have been ordered according to decreasing improvement in the expected limit obtained when removing the group from the list of nuisances included in the likelihood fit. The description of each single uncertainty in the text follows the same order.

Scale factors are used to correct for the differences in the double-b tagger misidentification efficiencies between data and prediction from simulation for W/Z+jets production and for $t\bar{t}$ production. These scale factors are measured by simultaneously fitting events that pass or fail the double-b tag requirement. The correlation between the double-b tagger and the p_{T}^{miss} (or U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate independently from each other within a constraint that depends on the recoil value. Such dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. p_{T}^{miss}/U . This shape uncertainty in the double-b scale factors measurement is the one that has the largest impact on the limits on the signal cross section.

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and tt̄ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_T of the CA15 jet. This shape uncertainty has been derived through a 2179 measurement performed using a sample enriched in multijet events with double-muon-tagged $g \rightarrow b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substruc-2181 ture variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes 2182 where the decay of a resonance inside the CA15 jet cone is expected. Such processes include sig-2183 nal production together with SM h and diboson production. The uncertainty has been derived 2184 from the efficiency measurement obtained by performing a fit in a control sample enriched in 2185 semi-leptonic tt events, where the CA15 jet originates from the W boson that comes from the 2186 hadronically decaying top quark. 2187

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in $p_{\rm T}^{\rm miss}$ magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, obtained by varying those parameters in simulation event-by-event

The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. 2196 Differences on the order of 2% are observed between single-muon and dimuon events at lower 2197 U values and they are sources of an additional systematic uncertainty in the transfer factors for those processes whose prediction relies on data events in the single-muon and dimuon CRs 2199 (tt, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the 2200 identified muon, they can change the shape of the *U* distribution and are thus treated as shape 2201 uncertainties. The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiency is parametrized as a function of p_{T}^{miss} . The uncer-2202 tainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies 2203 of the single-electron triggers are parametrized as a function of the electron p_T and η and an 2204 associated 1% systematic uncertainty is added into the fit. 2205

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization scales and PDFs are included as shape variations. Being a negligible background source, an un-

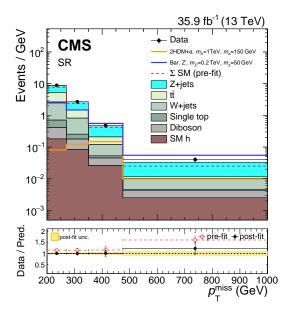


Figure 23: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

certainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by requiring $p_{\rm T}^{\rm miss} > 250\,{\rm GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty in QCD multijet yields in the signal region, while separate nuisance parameters are introduced for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as measured by CMS in Refs. [509, 510] and correlated across the SR and CRs.

H Results

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficiencies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z' model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results

H. Results 69

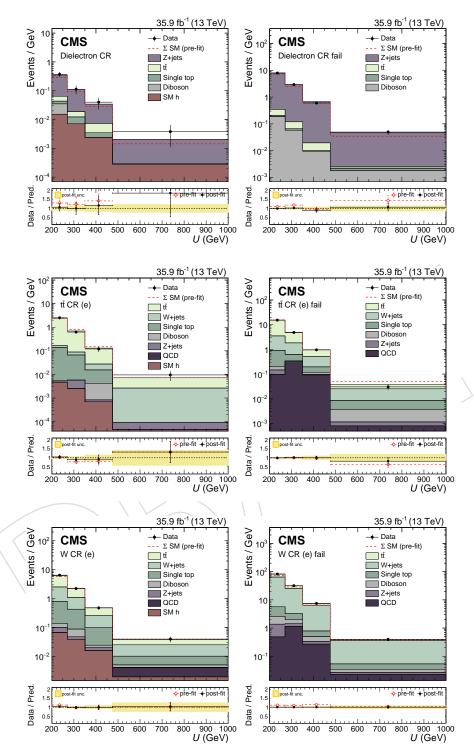


Figure 24: The *U* distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

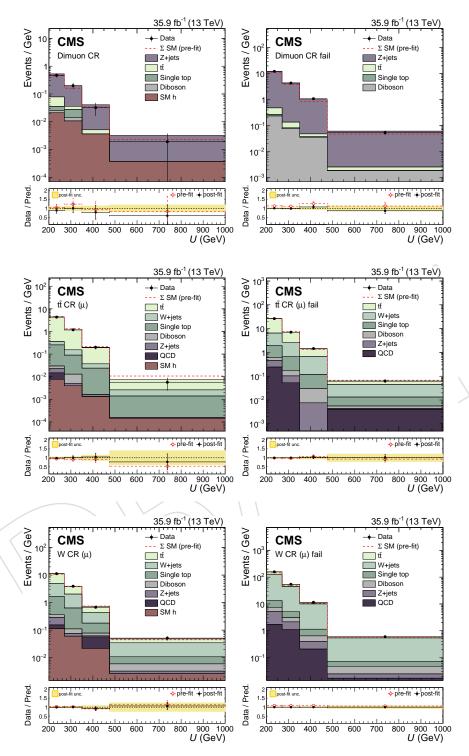


Figure 25: The *U* distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

I. Summary 71

Table 12: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

$p_{\mathrm{T}}^{\mathrm{miss}}$ -bin	200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
Z+jets	248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
tŧ	199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
W+jets	121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
Single top quark	21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
Diboson	16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
SM h	12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
Σ (SM)	619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
Data	619	214	59	21
2 HDM+ a , $m_A = 1$ TeV, $m_a = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2 \text{TeV}$, $m_{\chi} = 50 \text{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

of this search are interpreted in terms of upper limits on the signal strength modifier $\mu = \sigma/\sigma_{\text{theory}}$, where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

Figure 54 shows the upper limits on μ for the three scans (m_A , $\sin \theta$, and $\tan \beta$) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, $\sin \theta = 0.35$ and $\tan \beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming $\tan \beta = 1$. Also excluded are $\tan \beta$ values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus. Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2\mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},\tag{4}$$

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5\,\text{GeV}$.

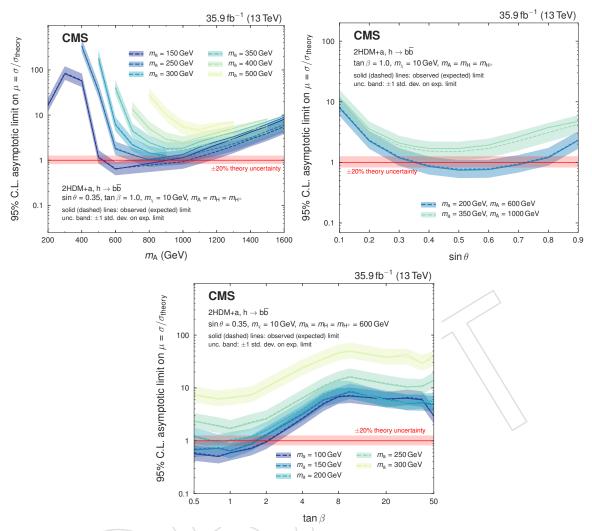


Figure 26: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

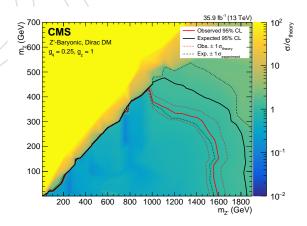


Figure 27: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

I. Summary 73

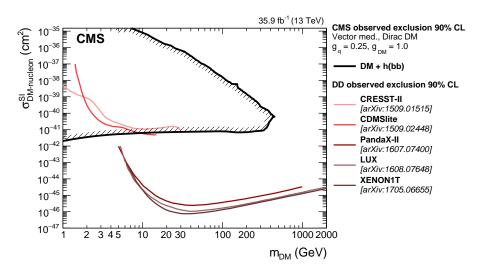


Figure 28: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_{χ} . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

I Summary

2271

2272

2273

2274

2275

2277

2278

2279

2280

2282

2283

2284

2285

2287

2288

2289

2290

2291

2293

2294

2295

A search for the associated production of dark matter (DM) particles with a Higgs boson decaying into a pair of bottom quarks is presented. No significant deviation from the predictions of the standard model (SM) is observed, and upper limits on the production cross section predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) and the baryonic Z' model are established. They constitute the most stringent collider exclusions placed on the parameters in these models so far. For the nominal choice of the mixing angle $\sin \theta$ and $\tan \beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900 \,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150 \,\text{GeV}$. Scanning over $\sin \theta$ with $\tan \beta = 1$, we exclude $0.35 < \sin \theta < 0.75$ for $m_A = 600$ GeV and $m_a = 200 \,\text{GeV}$. Finally, $\tan \beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_A = 600 \,\text{GeV}$ and $m_a = 100 (150) \,\text{GeV}$ and $\sin \theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_{\chi} = 10 \,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The reinterpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross section yields a higher sensitivity for m_{χ} < 5 GeV than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed experimentally for the first time.

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 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

and the CMS detector provided by the following funding agencies: BMWFW and FWF (Aus-2296 tria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); 2297 CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Fin-2299 land, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Ger-2300 many); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI 2301 (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM 2302 (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New 2303 Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, 2304 RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER 2305 (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United 2307 Kingdom); DOE and NSF (USA). 2308

Individuals have received support from the Marie-Curie program and the European Research 2309 Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foun-2310 dation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Fed-2311 eral Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans 2312 l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the 2314 Council of Science and Industrial Research, India; the HOMING PLUS program of the Foun-2315 dation for Polish Science, cofinanced from European Union, Regional Development Fund, the 2316 Mobility Plus program of the Ministry of Science and Higher Education, the National Science 2317 Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the 2319 National Priorities Research Program by Qatar National Research Fund; the Programa Severo 2320 Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and 2321 the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn 2322 University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA). 2324

References

```
[196] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and Constraints", Phys. Rept. 405 (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:0404175.
```

- [197] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",

 Ann. Rev. Astron. Astrophys. 48 (2010) 495,

 doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- [198] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", Ann. Rev. Astron. Astrophys. 49 (2011) 155,

doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.

2335 [199] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron. 2336 Astrophys. 594 (2016) A13, doi:10.1051/0004-6361/201525830, arXiv:1502.01589.

References 75

2338 [200] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

- [201] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [202] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", JHEP **06** (2013) 81, doi:10.1007/JHEP06(2013) 081, arXiv:1303.4571.
- [203] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, arXiv:1311.1511.
- ²³⁵⁰ [204] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", ²³⁵¹ *JHEP* **06** (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- ²³⁵² [205] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", *Phys.*²³⁵³ *Rev. D* **89** (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- ²³⁵⁴ [206] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.
- [207] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05 (2017) 138, arXiv:1701.07427.
- [208] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}=7$ and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
- [209] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- [210] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10 (2017) 180, arXiv:1703.05236.
- [211] 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.
- 2373 [212] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", JHEP 11 (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [213] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* 11 (2007) 070,
- doi:10.1088/1126-6708/2007/11/070,arXiv:0709.2092.

- [214] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [215] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [216] 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:0611129.
- 2388 [217] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12 (2012) 061, arXiv:1209.6215.
- [218] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* **03** (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- 2393 [219] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet 2394 production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, 2395 Italy, March 21-28, 2015. 2015. arXiv:1505.05704.
- 2396 [220] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", *JHEP* **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- 2399 [221] 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.
- [222] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2402] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2402] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2402] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2403] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2404] J. W. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2405] J. W. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2406] J. W. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2407] J. W. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2408] J. W. Campbell, R. W. Ellis, and C. Williams, "Vector boson pair production at the LHC", [2408] J. W. Campbell, R. W
- ²⁴⁰³ [223] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.
- ²⁴⁰⁵ [224] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- ²⁴⁰⁸ [225] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur. ²⁴⁰⁹ Phys. J. C **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, ²⁴¹⁰ arXiv:1404.5630.
- [226] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- ²⁴¹³ [227] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST **12** (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- ²⁴¹⁶ [228] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.

References 77

²⁴¹⁸ [229] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* ²⁴¹⁹ **59** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

- [230] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST 6 (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- ²⁴²³ [231] 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.
- ²⁴²⁵ [232] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) 05011.
- [233] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", *JHEP* **12** (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.
- [234] J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05 (2016) 156, arXiv:1603.00027.
- [235] 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.
- [236] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [237] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at \sqrt{s} = 7 TeV", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [238] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST 11 (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.
- [239] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST 8 (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [240] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.
- [241] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.

 arXiv:1009.1003.
- [242] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8 \text{ TeV}$ ", JINST 10 (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [243] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period",
 CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

- ²⁴⁵⁷ [244] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ ", *Phys. Lett. B* **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- ²⁴⁶⁰ [245] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ ", *Phys. Lett. B* **735** (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- [246] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", Phys. Rev. D **91** (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.
- ²⁴⁶⁶ [247] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7 \, \text{TeV}$ ", *JHEP* **06** (2014) 120, doi:10.1007/JHEP06 (2014) 120, arXiv:1402.1521.
- [248] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}=8\,\text{TeV}$ ", Phys.~Rev.~Lett.~ 112 (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- [249] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", Phys. Lett. B 763 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.
- [250] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **766** (2017) 268, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.
- [251] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs
 Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working
 Group", Technical Report CERN-2013-004, Geneva, 2013.
 doi:10.5170/CERN-2013-004, arXiv:1307.1347.
- ²⁴⁸² [252] A. L. Read, "Presentation of search results: the CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.
- ²⁴⁸⁴ [253] T. Junk, "Confidence level computation for combining searches with small statistics",

 Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002(99)00498-2,

 arXiv:hep-ex/9902006.
- [254] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3.
- [255] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse energy signals using simplified *s*-channel models of dark matter", (2016).

 arXiv:1603.04156.
- [256] SuperCDMS Collaboration, "New results from the search for low-mass weakly interacting massive particles with the CDMS low ionization threshold experiment", *Phys. Rev. Lett.* **116** (2016) 071301, doi:10.1103/PhysRevLett.116.071301, arXiv:1509.02448.

A. Introduction 79

```
[257] LUX Collaboration, "Results from a search for dark matter in the complete LUX
2497
         exposure", Phys. Rev. Lett. 118 (2017) 021303,
2498
         doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648.
    [258] XENON Collaboration, "First dark matter search results from the XENON1T
2500
         experiment", Phys. Rev. Lett. 119 (2017) 181301,
2501
         doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655.
2502
    [259] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II
2503
         experiment", Phys. Rev. Lett. 119 (2017) 181302,
2504
         doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.
2505
    [260] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold
2506
         CRESST-II detector", Eur. Phys. J. C 76 (2016) 25,
2507
         doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.
2508
    Search for associated production of dark matter with a Higgs boson decaying to a pair of bot-
2509
    tom quarks in pp collisions at sqrt(s) = 13 \text{ TeV}
2510
    [cern]The CMS Collaboration
2511
    2018/08/17
```

A Introduction

2513

2514

2515

2516

2517

2519

2521

2522

2524

2525

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for physics beyond the standard model (SM) [456–458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consistent with the hypothesis that DM is primarily composed of weakly interacting massive particles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. Assuming the pair production of DM particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signature. If DM particles are produced in association with a detectable SM particle, which could be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding protons, or through new effective couplings between DM and SM particles, their existence could be inferred via a large transverse momentum imbalance in the collision event.

While ISR production of the SM Higgs boson (h) [460–462] is highly suppressed due to the Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs boson and DM particles can occur if the Higgs boson takes part in the interaction producing the DM particles [463–465]. Such a production mechanism would allow to directly probe the structure of the effective DM-SM coupling.

In this paper, we present a search for DM production in association with an SM Higgs boson 2531 that decays into a bottom quark-antiquark pair $(b\bar{b})$. As the h \rightarrow bb decay mode has the largest branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest 2533 signal yield. The search is performed using the data set collected by the CMS experiment [466] 2534 at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated 2535 luminosity of approximately 35.9 fb⁻¹. Results are interpreted in terms of two simplified mod-2536 els predicting this signature. The first one is type-2 two Higgs doublet model extended by 2537 an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar 2538 and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi \bar{\chi}$.

The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM particles. Representative Feynman diagrams for the two models are presented in Fig. 50.

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no significant deviation from SM predictions [468]. In contrast to previously explored 2HDM models [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar h, assumed to be the observed 125 GeV Higgs boson, and a heavy neutral CP-even scalar h, that are the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mixing angle α ; a heavy neutral CP-odd pseudoscalar h and a light neutral CP-odd pseudoscalar h, that are the result of the mixing of the CP-odd mediator h with the CP-odd Higgs, with h representing the associated mixing angle; and two heavy charged scalars h with identical mass.

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H} = m_{\rm H^{\pm}} = m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,{\rm GeV}$.

The baryonic Z' model [465] is an extension of the SM with an additional $U(1)_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding $U(1)_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the $U(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_B Z' Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin\theta/v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_q , is 0.25 and the Z' coupling to DM, g_χ , is set to 1. This is well below the bounds $g_{q_1}, g_\chi \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin \theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'} \sin \theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and m_{χ} , which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, together with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically independent data samples are used to predict the hadronic recoil distribution for these SM processes that constitute the largest sources of background. Both signal and background contributions

B. The CMS detector 81

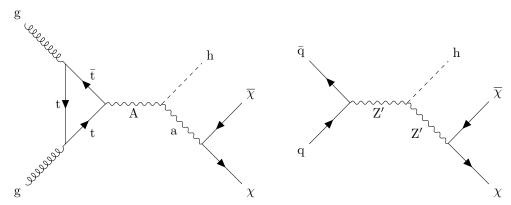


Figure 29: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z' model (right).

to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

B The CMS detector

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum (p_T) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of 3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta| < 3$. The steel and quartz-fiber forward Cherenkov hadron calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and covers $|\eta| < 2.4$. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than 4 μ s, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to 1 kHz.

C Simulated data samples

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG V2 [472–474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the *p*_T of the respective boson with NLO QCD corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO

and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

D Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_T^{i2}$, where p_T^i is the transverse momentum of the i^{th} track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, large-area jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central ($\eta < 2.4$). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$ is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs

D. Event reconstruction 83

boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , 2666 will affect the shape of the distribution of m_{SD} for the background processes. In this search, the 2667 value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore 2669 desirable to preserve a smoothly falling jet mass distribution as a function of p_T . As explained 2670 in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass 2671 distribution for QCD multijet events is expected to scale with p_T , decorrelating the N_2 variable 2672 as a function of ρ and p_T would be the most appropriate procedure. The decorrelation strategy 2673 described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds 2674 to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we 2675 denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494]. 2676

This search also utilizes narrow jets clustered with the anti- $k_{\rm T}$ algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta|$ < 2.4).

2683

2684

2685

2686

2687

2688

2689

2691

2692

2693

2694

2695

2696

2697

2698

2699

2700

2701

2702

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the consistency with originating from the primary vertex are imposed. The reconstructed electron is required to be within $|\eta| < 2.5$, excluding the transition region 1.44 $< |\eta| < 1.57$ between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons and poorly measured mesons tagged as muons [497]. These additional criteria include requirements on the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates with $|\eta| < 2.4$ are considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are reconstructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton decays. The τ_h candidates with $p_T > 18 \,\text{GeV}$ and $|\eta| < 2.3$ are considered [496, 498, 499]. Photon candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15$ GeV and $|\eta| < 2.5$.

The missing transverse momentum $\vec{p}_{\rm T}^{\rm miss}$ is defined as the negative vectorial sum of the $p_{\rm T}$ of all the reconstructed PF candidates. Its magnitude is denoted as $p_{\rm T}^{\rm miss}$. Corrections to jet momenta are propagated to the $p_{\rm T}^{\rm miss}$ calculation as well as event filters [500] are used to remove spurious high $p_{\rm T}^{\rm miss}$ events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

Table 13: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, t t , W+jets	0	0	pass
Single-lepton, b-tagged	tī, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

E Event selection

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T, trig}^{miss}$ or H_{T}^{miss} , where $p_{T, trig}^{miss}$ (H_{T}^{miss}) is the magnitude of the vectorial (scalar) sum of \vec{p}_{T} of all the particles (jets with $p_{T} > 20\,\text{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{T, trig}^{miss}$ calculation. Thresholds on $p_{T, trig}^{miss}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{T, trig}^{miss}$ and H_{T}^{miss} are reffered to as p_{T}^{miss} triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T} > 200\,{\rm GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\rm SD} < 150\,{\rm GeV}$ and $N_2^{\rm DDT} < 0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (U) must be larger than 0.4 rad to reject multijet events that mimic signal events. Vetoes on $\tau_{\rm h}$ candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from tt events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_{\rm T}>10\,{\rm GeV}$ and $|\eta|<2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between 60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of

F. Signal extraction 85

the trigger efficiency, at least one of the two electrons must have $p_{\rm T} > 40\,{\rm GeV}$, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from W+jets and semileptonic tt events. To predict these backgrounds, four single-lepton samples are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target tt events, while the other two 2750 single-lepton CRs target W+jets events. Single-muon events are selected using the $p_{\rm T}^{\rm miss}$ trigger 2751 triggers described above, as well as single electron events are selected using the same single-2752 electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and 2754 isolation requirements. In addition, samples with a single electron need to have $p_T^{\text{miss}} > 50 \,\text{GeV}$ 2755 to avoid a large contamination from multijet events. 2756

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

F Signal extraction

2761

2762

2763

2764

2765

2768

2769

2770

2771

2772

2773

2774

2777

2778

2782

2783

2784

2785

2786

2787

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the $p_{\rm T}^{\rm miss}$ and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and are given by the ratio of the prediction for a given bin in $p_{\rm T}^{\rm miss}$ in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the $t\bar{t}$ process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt̄ backgrounds take into account the impact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron control samples, the additional requirement on $p_{\rm T}^{\rm miss}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the tt̄ process, transfer factors to predict this contamination from tt̄ events are also imposed between the single-lepton CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contamination from W+jets production in the tt̄ CR with events that fail the double-b tag requirement. Likewise, the Z+jets background prediction in the signal region is connected to the and dilepton CRs via transfer factors. They account for the difference in the branching fractions of the Z $\rightarrow \nu\nu$ and the Z $\rightarrow \ell\ell$ decays and the impacts of lepton acceptances and selection efficiencies.

G Systematic uncertainties

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncertainties of the search. They can either affect the rate or the shape of $p_T^{miss}(U)$ for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal

Table 14: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

Systematic uncertainty	Туре	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, tt̄	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, tt̄	1.9%
Double-b tagging	shape	SM h, signal	1.2%
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
p _T ^{miss} magnitude	5%	all	0.7%
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%
$p_{\rm T}^{\rm miss}$ trigger muon multiplicity	shape	Z+jets, W+jets	
$p_{\rm T}^{\rm miss}$ trigger efficiency	1%	all	< 0.5%
single-electron trigger	1%	all	
AK4 b tagging	shape	all	< 0.5%
au lepton veto	3%	all	< 0.5%
Lepton efficiency	1% per lepton	all	< 0.5 /6
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%
Renorm./fact. scales	shape	SM h	
PDF	shape	SM h	< 0.5%
Multijet normalization	100%	multijet	< 0.5%
Theoretical cross section	20%	t, diboson	

distribution. The list of the systematic uncertainties considered in this search is presented in Table 23. To better estimate their impact on the results, uncertainties from a similar source (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties have been ordered according to decreasing improvement in the expected limit obtained when removing the group from the list of nuisances included in the likelihood fit. The description of each single uncertainty in the text follows the same order.

Scale factors are used to correct for the differences in the double-b tagger misidentification efficiencies between data and prediction from simulation for W/Z+jets production and for $t\bar{t}$ production. These scale factors are measured by simultaneously fitting events that pass or fail the double-b tag requirement. The correlation between the double-b tagger and the p_{T}^{miss} (or U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate independently from each other within a constraint that depends on the recoil value. Such dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. p_{T}^{miss}/U . This shape uncertainty in the double-b scale factors measurement is the one that has the largest impact on the limits on the signal cross section.

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and tt̄ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the $p_{\rm T}$ of the CA15 jet. This shape uncertainty has been derived through a measurement performed using a sample enriched in multijet events with double-muon-tagged $g \to b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substructure variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes where the decay of a resonance inside the CA15 jet cone is expected. Such processes include sig-

nal production together with SM h and diboson production. The uncertainty has been derived
 from the efficiency measurement obtained by performing a fit in a control sample enriched in
 semi-leptonic tt events, where the CA15 jet originates from the W boson that comes from the
 hadronically decaying top quark.

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in p_T^{miss} magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, obtained by varying those parameters in simulation event-by-event

The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. 2824 Differences on the order of 2% are observed between single-muon and dimuon events at lower 2825 U values and they are sources of an additional systematic uncertainty in the transfer factors 2826 for those processes whose prediction relies on data events in the single-muon and dimuon CRs 2827 (tt, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the 2828 identified muon, they can change the shape of the U distribution and are thus treated as shape 2829 uncertainties. The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiency is parametrized as a function of p_{T}^{miss} . The uncer-2830 tainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies 2831 of the single-electron triggers are parametrized as a function of the electron p_T and η and an 2832 associated 1% systematic uncertainty is added into the fit. 2833

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization scales and PDFs are included as shape variations. Being a negligible background source, an uncertainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by requiring $p_{\rm T}^{\rm miss} > 250\,{\rm GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty in QCD multijet yields in the signal region, while separate nuisance parameters are introduced for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single

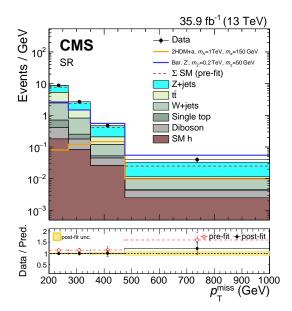


Figure 30: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as measured by CMS in Refs. [509, 510] and correlated across the SR and CRs.

H Results

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficiencies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z' model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results of this search are interpreted in terms of upper limits on the signal strength modifier $\mu = \sigma/\sigma_{\text{theory}}$, where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

H. Results 89

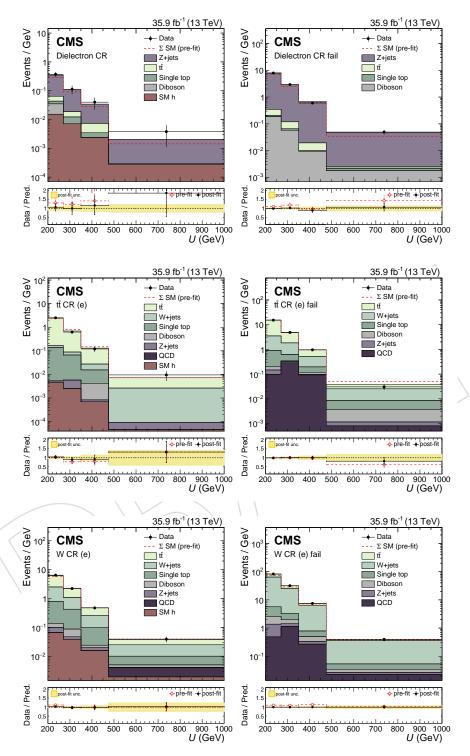


Figure 31: The *U* distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

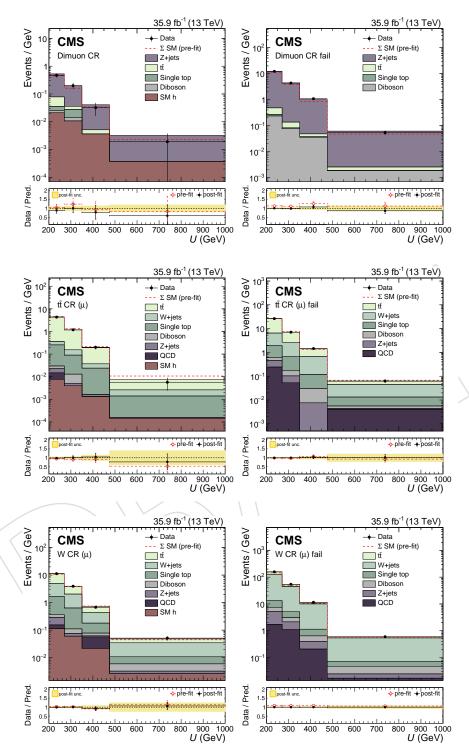


Figure 32: The *U* distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

I. Summary 91

Table 15: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

$p_{\mathrm{T}}^{\mathrm{miss}}$ -bin	200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
Z+jets	248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
tŧ	199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
W+jets	121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
Single top quark	21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
Diboson	16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
SM h	12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
Σ (SM)	619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
Data	619	214	59	21
2 HDM+ a , $m_A = 1$ TeV, $m_a = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2 \text{TeV}$, $m_{\chi} = 50 \text{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

Figure 54 shows the upper limits on μ for the three scans (m_A , $\sin \theta$, and $\tan \beta$) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, $\sin \theta = 0.35$ and $\tan \beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming $\tan \beta = 1$. Also excluded are $\tan \beta$ values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus. Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2\mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},\tag{5}$$

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5\,\text{GeV}$.

I Summary

A search for the associated production of dark matter (DM) particles with a Higgs boson decaying into a pair of bottom quarks is presented. No significant deviation from the predictions of the standard model (SM) is observed, and upper limits on the production cross section predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar boson *a* (2HDM+*a*) and the baryonic Z' model are established. They constitute the most stringent collider exclusions placed on the parameters in these models so far. For the nominal

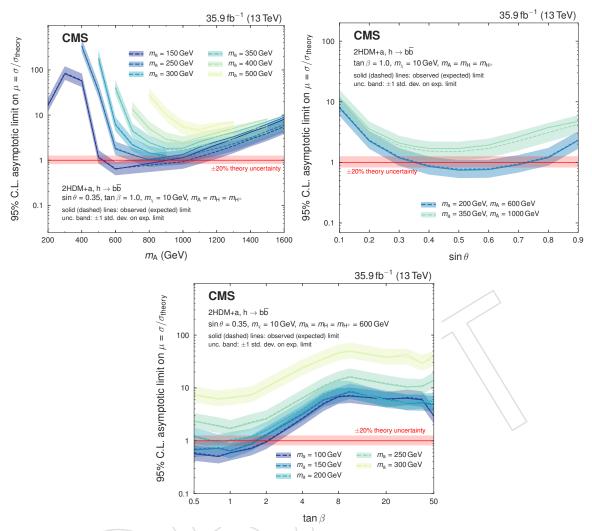


Figure 33: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

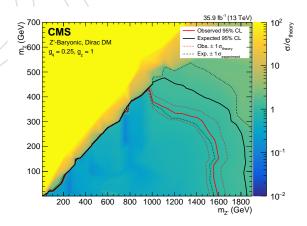


Figure 34: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

I. Summary 93

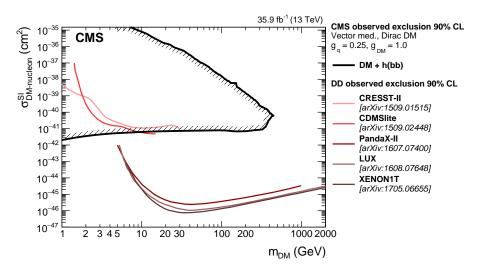


Figure 35: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_{χ} . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

choice of the mixing angle $\sin\theta$ and $\tan\beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900\,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150\,\text{GeV}$. Scanning over $\sin\theta$ with $\tan\beta = 1$, we exclude $0.35 < \sin\theta < 0.75$ for $m_A = 600\,\text{GeV}$ and $m_a = 200\,\text{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_A = 600\,\text{GeV}$ and $m_a = 100\,$ (150) GeV and $\sin\theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_\chi = 10\,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The reinterpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross section yields a higher sensitivity for $m_\chi < 5\,\text{GeV}$ than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed experimentally for the first time.

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 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 and the CMS detector provided by the following funding agencies: BMWFW and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New

Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie program and the European Research 2937 Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foun-2938 dation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Fed-2939 eral Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie 2941 (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the 2942 Council of Science and Industrial Research, India; the HOMING PLUS program of the Foun-2943 dation for Polish Science, cofinanced from European Union, Regional Development Fund, the Mobility Plus program of the Ministry of Science and Higher Education, the National Science 2945 Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2946 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the 2947 National Priorities Research Program by Qatar National Research Fund; the Programa Severo 2948 Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn 2950 University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project 2951 (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA). 2952

References

2953

- ²⁹⁵⁴ [261] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and Constraints", *Phys. Rept.* **405** (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:0404175.
- ²⁹⁵⁷ [262] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",
 ²⁹⁵⁸ Ann. Rev. Astron. Astrophys. 48 (2010) 495,
 ²⁹⁵⁹ doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- ²⁹⁶⁰ [263] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", Ann. Rev. Astron. Astrophys. **49** (2011) 155, doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.
- ²⁹⁶³ [264] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron. ²⁹⁶⁴ Astrophys. **594** (2016) A13, doi:10.1051/0004-6361/201525830, ²⁹⁶⁵ arXiv:1502.01589.
- 2966 [265] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- ²⁹⁶⁹ [266] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- ²⁹⁷² [267] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **06** (2013) 81, doi:10.1007/JHEP06 (2013) 081, arXiv:1303.4571.

References 95

²⁹⁷⁵ [268] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, arXiv:1311.1511.

- ²⁹⁷⁸ [269] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", ²⁹⁷⁹ JHEP **06** (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- ²⁹⁸⁰ [270] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", *Phys.*²⁹⁸¹ *Rev. D* **89** (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- ²⁹⁸² [271] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.
- ²⁹⁸⁴ [272] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05 (2017) 138, arXiv:1701.07427.
- [273] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}=7$ and 8 TeV", JHEP **08** (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
- [274] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- ²⁹⁹⁵ [275] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10 (2017) 180, arXiv:1703.05236.
- ²⁹⁹⁸ [276] 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.
- ³⁰⁰¹ [277] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [278] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* 11 (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [279] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [280] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [281] 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:0611129.

- 3016 [282] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", JHEP 12 (2012) 061, doi:10.1007/JHEP12 (2012) 061, arXiv:1209.6215.
- [283] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* 03 (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- [284] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, Italy, March 21-28, 2015. 2015. arXiv:1505.05704.
- [285] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", *JHEP* **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- ³⁰²⁷ [286] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", Comput. Phys. Commun. **191** (2015) ³⁰²⁸ 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- ³⁰²⁹ [287] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", ³⁰³⁰ *JHEP* **07** (2011) 018, doi:10.1007/JHEP07 (2011) 018, arXiv:1105.0020.
- ³⁰³¹ [288] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.
- [289] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C 76 (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- ³⁰³⁶ [290] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur. ³⁰³⁷ Phys. J. C **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, ³⁰³⁸ arXiv:1404.5630.
- 3039 [291] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- ³⁰⁴¹ [292] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST **12** (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- ³⁰⁴⁴ [293] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.
- ³⁰⁴⁶ [294] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* ³⁰⁴⁷ **59** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.
- ³⁰⁴⁸ [295] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST **6** (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- 3051 [296] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", JHEP **05** (2014) 146, 3052 doi:10.1007/JHEP05 (2014) 146, arXiv:1402.2657.
- ³⁰⁵³ [297] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) 05011.

References 97

[298] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", JHEP
 12 (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.

- [299] J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05 (2016) 156, arXiv:1603.00027.
- ³⁰⁶⁰ [300] 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.
- [301] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8 \text{ TeV}$ ", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- 3065 [302] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- 3068 [303] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST 11 (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.
- [304] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST 8 (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [305] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.
- [306] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.

 arXiv:1009.1003.
- [307] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8 \text{ TeV}$ ", JINST 10 (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [308] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* 741 (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- [310] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ ", Phys. Lett. B 735 (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- [311] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", Phys. Rev. D **91** (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.

3129

3130

- [312] CMS Collaboration, "Measurement of the production cross sections for a Z boson and 3094 one or more b jets in pp collisions at $\sqrt{s} = 7 \text{ TeV}''$, JHEP **06** (2014) 120, 3095 doi:10.1007/JHEP06(2014)120, arXiv:1402.1521. 3096 [313] CMS Collaboration, "Observation of the associated production of a single top quark and 3097 a W boson in pp collisions at $\sqrt{s} = 8 \text{ TeV}''$, Phys. Rev. Lett. 112 (2014) 231802, 3098 doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942. 3099 [314] CMS Collaboration, "Measurement of the ZZ production cross section and Z 3100 $\rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13 \text{ TeV}''$, Phys. Lett. B 763 3101 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834. 3102 [315] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions 3103 at $\sqrt{s} = 13 \text{ TeV}''$, Phys. Lett. B **766** (2017) 268, 3104 doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943. 3105 [316] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs 3106 Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working 3107 Group", Technical Report CERN-2013-004, Geneva, 2013. 3108 doi:10.5170/CERN-2013-004, arXiv:1307.1347. 3109 [317] A. L. Read, "Presentation of search results: the CL_s technique", J. Phys. G 28 (2002) 3110 2693, doi:10.1088/0954-3899/28/10/313. 3111 [318] T. Junk, "Confidence level computation for combining searches with small statistics", 3112 Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002 (99) 00498-2, 3113 arXiv:hep-ex/9902006. 3114 [319] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for 3115 likelihood-based tests of new physics", Eur. Phys. J. C 71 (2011) 1554, 3116 doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3. 3117 [320] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse 3118 energy signals using simplified s-channel models of dark matter", (2016). 3119 arXiv:1603.04156. 3120 [321] SuperCDMS Collaboration, "New results from the search for low-mass weakly 3121 interacting massive particles with the CDMS low ionization threshold experiment", 3122 Phys. Rev. Lett. 116 (2016) 071301, doi:10.1103/PhysRevLett.116.071301, 3123 arXiv:1509.02448. 3124 [322] LUX Collaboration, "Results from a search for dark matter in the complete LUX 3125 exposure", Phys. Rev. Lett. 118 (2017) 021303, 3126 doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648. 3127
- [324] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II experiment", *Phys. Rev. Lett.* **119** (2017) 181302, doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.

[323] XENON Collaboration, "First dark matter search results from the XENON1T

doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655.

experiment", Phys. Rev. Lett. 119 (2017) 181301,

A. Introduction 99

```
[325] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold CRESST-II detector", Eur. Phys. J. C 76 (2016) 25, doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.

Search for associated production of dark matter with a Higgs boson decaying to a pair of bottom quarks in pp collisions at sqrt(s) = 13 TeV

[cern]The CMS Collaboration

2018/08/17
```

A Introduction

3141

3159

3160

3161

3162

3163

3164

3165

3166

3168

3169

3170

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for 3142 physics beyond the standard model (SM) [456–458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consis-3144 tent with the hypothesis that DM is primarily composed of weakly interacting massive parti-3145 cles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could 3146 be produced by colliding SM particles at high energy. Assuming the pair production of DM 3147 particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signa-3149 ture. If DM particles are produced in association with a detectable SM particle, which could 3150 be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding pro-3151 tons, or through new effective couplings between DM and SM particles, their existence could 3152 be inferred via a large transverse momentum imbalance in the collision event. 3153

While ISR production of the SM Higgs boson (h) [460–462] is highly suppressed due to the Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs boson and DM particles can occur if the Higgs boson takes part in the interaction producing the DM particles [463–465]. Such a production mechanism would allow to directly probe the structure of the effective DM-SM coupling.

In this paper, we present a search for DM production in association with an SM Higgs boson that decays into a bottom quark-antiquark pair $(b\bar{b})$. As the $h\to b\bar{b}$ decay mode has the largest branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest signal yield. The search is performed using the data set collected by the CMS experiment [466] at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of approximately 35.9 fb⁻¹. Results are interpreted in terms of two simplified models predicting this signature. The first one is type-2 two Higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi\bar{\chi}$. The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM particles. Representative Feynman diagrams for the two models are presented in Fig. 50.

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no significant deviation from SM predictions [468]. In contrast to previously explored 2HDM models [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar h, as-

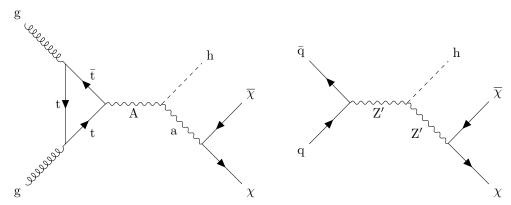


Figure 36: Feynman diagrams for the 2HDM+*a* model (left) and the baryonic Z' model (right).

sumed to be the observed 125 GeV Higgs boson, and a heavy neutral CP-even scalar H, that are the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mixing angle α ; a heavy neutral CP-odd pseudoscalar A and a light neutral CP-odd pseudoscalar a, that are the result of the mixing of the CP-odd mediator P with the CP-odd Higgs, with θ representing the associated mixing angle; and two heavy charged scalars H^{\pm} with identical mass.

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H} = m_{\rm H^{\pm}} = m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10$ GeV.

The baryonic Z' model [465] is an extension of the SM with an additional $U(1)_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding $U(1)_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the $U(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_BZ'Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin\theta/v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_q , is 0.25 and the Z' coupling to DM, g_χ , is set to 1. This is well below the bounds g_q , $g_\chi \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin\theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'}\sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and m_χ , which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, together with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically indepen-

B. The CMS detector 101

dent data samples are used to predict the hadronic recoil distribution for these SM processes that constitute the largest sources of background. Both signal and background contributions to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

B The CMS detector

3213

3214

3216

3217

3218

3219

3220

3222

3223

3224

3225

3227

3228

3229

3230

3231

3233

3234

3235

3236

3238

3239

3240

3241

3243

3244

3245

3248

3249

3251

3253

3254

3255

3256

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum ($p_{\rm T}$) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of 3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta| < 3$. The steel and quartz-fiber forward Cherenkov hadron calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and covers $|\eta| < 2.4$. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than 4 μ s, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to 1 kHz.

C Simulated data samples

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG V2 [472-474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the p_T of the respective boson with NLO QCD corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or

a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

D Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_{\rm T}^{i2}$, where $p_{\rm T}^i$ is the transverse momentum of the $i^{\rm th}$ track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, large-area jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central ($\eta < 2.4$). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$ is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , will affect the shape of the distribution of $m_{\rm SD}$ for the background processes. In this search, the value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore desirable to preserve a smoothly falling jet mass distribution as a function of $p_{\rm T}$. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with $p_{\rm T}$, decorrelating the N_2 variable as a function of ρ and $p_{\rm T}$ would be the most appropriate procedure. The decorrelation strategy

E. Event selection 103

described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494].

This search also utilizes narrow jets clustered with the anti- k_T algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta|$ < 2.4).

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the consistency with originating from the primary vertex are imposed. The reconstructed electron is required to be within $|\eta| < 2.5$, excluding the transition region 1.44 $< |\eta| < 1.57$ between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons and poorly measured mesons tagged as muons [497]. These additional criteria include requirements on the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates with $|\eta| < 2.4$ are considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are reconstructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton decays. The τ_h candidates with $p_T > 18$ GeV and $|\eta| < 2.3$ are considered [496, 498, 499]. Photon candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15$ GeV and $|\eta| < 2.5$.

The missing transverse momentum $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the negative vectorial sum of the p_{T} of all the reconstructed PF candidates. Its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$. Corrections to jet momenta are propagated to the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation as well as event filters [500] are used to remove spurious high $p_{\mathrm{T}}^{\mathrm{miss}}$ events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

E Event selection

3305

3307

3308

3309

3310

3311

3312

3313

3314

3315

3317

3318

3319

3320

3322

3323

3324

3325

3326

3327

3328

3329

3330

3336

3337

3338

3340

3341

3344

3345

3346

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T,\text{trig}}^{\text{miss}}$ or H_{T}^{miss} , where $p_{T,\text{trig}}^{\text{miss}}$ (H_{T}^{miss}) is the

Table 16: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, tt, W+jets	0	0	pass
Single-lepton, b-tagged	tī, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

magnitude of the vectorial (scalar) sum of \vec{p}_T of all the particles (jets with $p_T > 20\,\text{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{T,\text{trig}}^{\text{miss}}$ calculation. Thresholds on $p_{T,\text{trig}}^{\text{miss}}$ and H_T^{miss} are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{T,\text{trig}}^{\text{miss}}$ and H_T^{miss} are reffered to as p_T^{miss} triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T} > 200\,{\rm GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\rm SD} < 150\,{\rm GeV}$ and $N_2^{\rm DDT} < 0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than $200\,{\rm GeV}$, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (U) must be larger than 0.4 rad to reject multijet events that mimic signal events. Vetoes on $\tau_{\rm h}$ candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from $t\bar{t}$ events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_{\rm T} > 10\,{\rm GeV}$ and $|\eta| < 2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between 60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of the trigger efficiency, at least one of the two electrons must have $p_{\rm T} > 40\,{\rm GeV}$, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from W+jets and semileptonic $t\bar{t}$ events. To predict these backgrounds, four single-lepton samples are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target $t\bar{t}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the p_T^{miss} trigger triggers described above, as well as single electron events are selected using the same single-electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and isolation requirements. In addition, samples with a single electron need to have $p_T^{miss} > 50 \, \text{GeV}$ to avoid a large contamination from multijet events.

F. Signal extraction 105

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

F Signal extraction

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the $p_{\rm T}^{\rm miss}$ and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and are given by the ratio of the prediction for a given bin in $p_{\rm T}^{\rm miss}$ in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the $t\bar{t}$ process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt backgrounds take into account the impact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron control samples, the additional requirement on $p_{\rm T}^{\rm miss}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the tt process, transfer factors to predict this contamination from tt events are also imposed between the single-lepton CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contamination from W+jets production in the tt CR with events that fail the double-b tag requirement. Likewise, the Z+jets background prediction in the signal region is connected to the and dilepton CRs via transfer factors. They account for the difference in the branching fractions of the Z $\rightarrow \nu\nu$ and the Z $\rightarrow \ell\ell$ decays and the impacts of lepton acceptances and selection efficiencies.

G Systematic uncertainties

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncertainties of the search. They can either affect the rate or the shape of $p_{\rm T}^{\rm miss}$ (U) for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal distribution. The list of the systematic uncertainties considered in this search is presented in Table 23. To better estimate their impact on the results, uncertainties from a similar source (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties have been ordered according to decreasing improvement in the expected limit obtained when removing the group from the list of nuisances included in the likelihood fit. The description of each single uncertainty in the text follows the same order.

Scale factors are used to correct for the differences in the double-b tagger misidentification efficiencies between data and prediction from simulation for W/Z+jets production and for $t\bar{t}$ production. These scale factors are measured by simultaneously fitting events that pass or fail the double-b tag requirement. The correlation between the double-b tagger and the p_{T}^{miss} (or 3426 U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate independently from each other within a constraint that depends on the recoil value. Such

3434

3435

3436

3438

3439

3440

3441

3443

Table 17: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

Systematic uncertainty	Туре	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, tt	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, tt̄	1.9%
Double-b tagging	shape	SM h, signal	1.2%
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\rm T}^{\rm miss}$ magnitude	5%	all	0.7%
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%
p _T ^{miss} trigger muon multiplicity	shape	Z+jets, W+jets	
$p_{\mathrm{T}}^{\mathrm{\hat{m}iss}}$ trigger efficiency	1%	all	< 0.5%
single-electron trigger	1%	all	
AK4 b tagging	shape	all	< 0.5%
au lepton veto	3%	all	< 0.5%
Lepton efficiency	1% per lepton	all	V 0.570
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%
Renorm./fact. scales	shape	SM h	()
PDF	shape	SM h	< 0.5%
Multijet normalization	100%	multijet	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
Theoretical cross section	20%	t, diboson	

dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. p_T^{miss}/U . This shape uncertainty in the double-b scale factors measurement is the one that has the largest impact on the limits on the signal cross section.

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and tt̄ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the $p_{\rm T}$ of the CA15 jet. This shape uncertainty has been derived through a measurement performed using a sample enriched in multijet events with double-muon-tagged $g \to b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substructure variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes where the decay of a resonance inside the CA15 jet cone is expected. Such processes include signal production together with SM h and diboson production. The uncertainty has been derived from the efficiency measurement obtained by performing a fit in a control sample enriched in semi-leptonic $t\bar{t}$ events, where the CA15 jet originates from the W boson that comes from the hadronically decaying top quark.

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in p_T^{miss} magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, ob-

tained by varying those parameters in simulation event-by-event

The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. 3452 Differences on the order of 2% are observed between single-muon and dimuon events at lower 3453 U values and they are sources of an additional systematic uncertainty in the transfer factors 3454 for those processes whose prediction relies on data events in the single-muon and dimuon CRs 3455 (tt, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the 3456 identified muon, they can change the shape of the *U* distribution and are thus treated as shape 3457 uncertainties. The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiency is parametrized as a function of p_{T}^{miss} . The uncer-3458 tainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies 3459 of the single-electron triggers are parametrized as a function of the electron p_T and η and an 3460 associated 1% systematic uncertainty is added into the fit. 3461

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization 3477 scales and PDFs are included as shape variations. Being a negligible background source, an un-3478 certainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a 3479 sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by requiring $p_{\rm T}^{\rm miss} > 250$ GeV, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ 3481 and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty 3482 in QCD multijet yields in the signal region, while separate nuisance parameters are introduced 3483 for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single 3484 top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR 3485 and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as 3486 measured by CMS in Refs. [509, 510] and correlated across the SR and CRs. 3487

H Results

3488

3489

3490

3492

3493

3451

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficien-

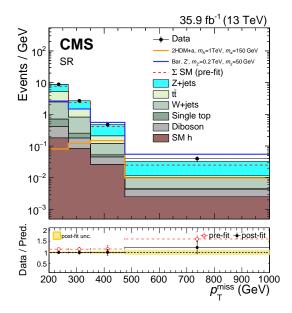


Figure 37: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

Table 18: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

p _T ^{miss} -bin	200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
Z+jets	248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
tŧ \	199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
W+jets	121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
Single top quark	21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
Diboson	16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
SM h	12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
Σ (SM)	619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
Data	619	214	59	21
2 HDM+ a , $m_A = 1$ TeV, $m_a = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z' , $m_{Z'} = 0.2 \text{ TeV}$, $m_{\chi} = 50 \text{ GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

cies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z′ model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results of this search are interpreted in terms of upper limits on the signal strength modifier $\mu =$

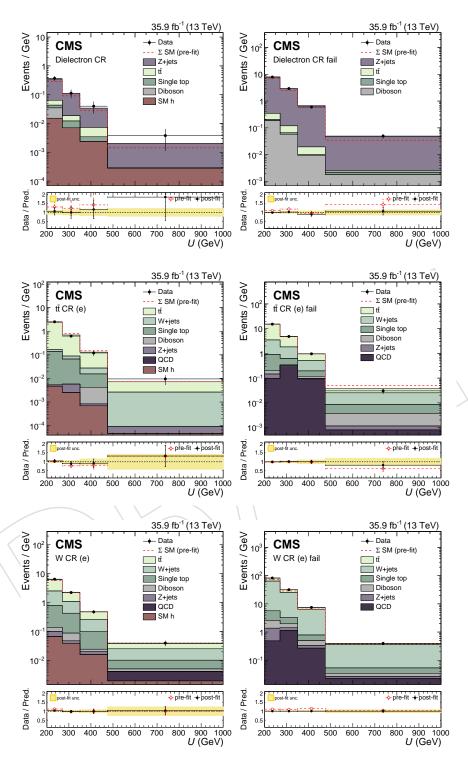


Figure 38: The *U* distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

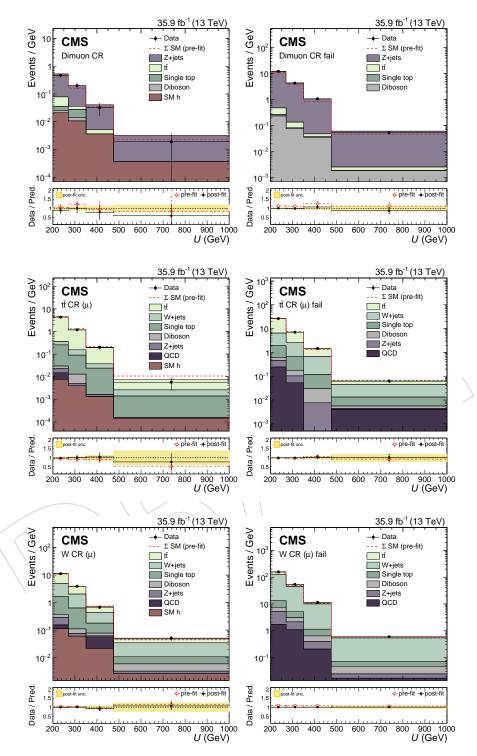


Figure 39: The *U* distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

 σ/σ_{theory} , where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

Figure 54 shows the upper limits on μ for the three scans (m_A , sin θ , and tan β) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, sin $\theta = 0.35$ and tan $\beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming tan $\beta = 1$. Also excluded are tan β values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

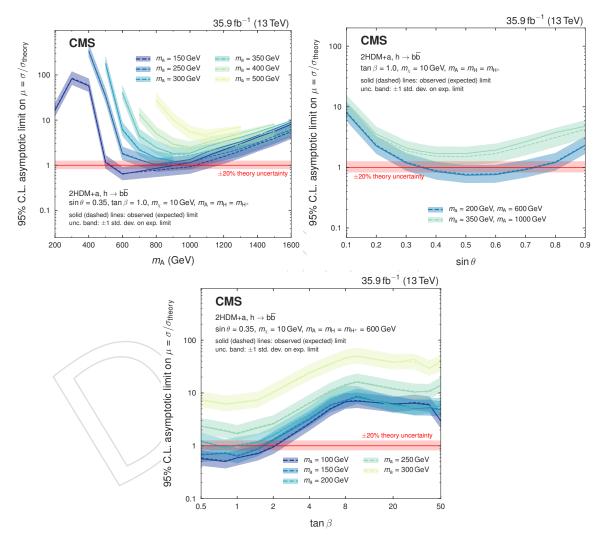


Figure 40: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus.

3523

3524

3525

3526

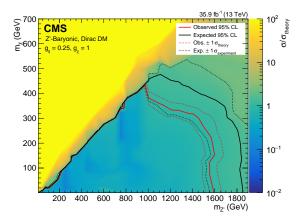


Figure 41: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

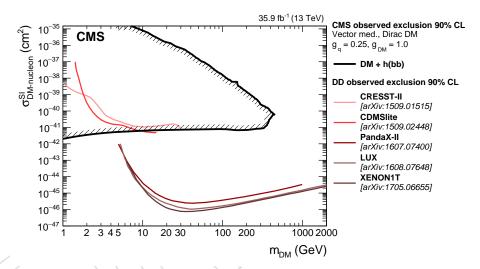


Figure 42: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_{χ} . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2\mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},$$
 (6)

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5\,\text{GeV}$.

I. Summary 113

I Summary

3527

3528

3529

3530

3531

3532

3534

3535

3536

3538

3539

3540

3541

3543

3544

3545

3565

3566

3567 3568

3570

3571

3572

A search for the associated production of dark matter (DM) particles with a Higgs boson decaying into a pair of bottom quarks is presented. No significant deviation from the predictions of the standard model (SM) is observed, and upper limits on the production cross section predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) and the baryonic Z' model are established. They constitute the most stringent collider exclusions placed on the parameters in these models so far. For the nominal choice of the mixing angle $\sin \theta$ and $\tan \beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900 \,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150 \,\text{GeV}$. Scanning over $\sin \theta$ with $\tan \beta = 1$, we exclude $0.35 < \sin \theta < 0.75$ for $m_A = 600$ GeV and $m_{\rm a}=200\,{\rm GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_{\rm A}=600\,{\rm GeV}$ and $m_a = 100 (150) \,\text{GeV}$ and $\sin \theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_{\chi} = 10 \,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The reinterpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross section yields a higher sensitivity for m_{χ} < 5 GeV than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed experimentally for the first time.

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent perfor-3546 mance of the LHC and thank the technical and administrative staffs at CERN and at other 3547 CMS institutes for their contributions to the success of the CMS effort. In addition, we grate-3548 fully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Fi-3550 nally, we acknowledge the enduring support for the construction and operation of the LHC 3551 and the CMS detector provided by the following funding agencies: BMWFW and FWF (Aus-3552 tria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); 3553 CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Fin-3555 land, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Ger-3556 many); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM 3558 (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New 3559 Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, 3560 RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER 3561 (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United 3563 Kingdom); DOE and NSF (USA). 3564

Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the

Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Severo Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

References

3581

- [326] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and Constraints", *Phys. Rept.* **405** (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:0404175.
- [327] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",

 Ann. Rev. Astron. Astrophys. 48 (2010) 495,

 doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- [328] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", Ann. Rev. Astron. Astrophys. 49 (2011) 155,
 doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.
- [329] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron.

 Astrophys. 594 (2016) A13, doi:10.1051/0004-6361/201525830,

 arXiv:1502.01589.
- [330] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [331] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [332] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}=7$ and 8 TeV", JHEP **06** (2013) 81, doi:10.1007/JHEP06 (2013) 081, arXiv:1303.4571.
- 3603 [333] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, arXiv:1311.1511.
- 3606 [334] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", 3607 *JHEP* **06** (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- ³⁶⁰⁸ [335] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", *Phys.*³⁶⁰⁹ *Rev. D* **89** (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- 3610 [336] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

References 115

[337] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05 (2017) 138, arXiv:1701.07427.

- [338] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}=7$ and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
- [339] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- [340] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10 (2017) 180, arXiv:1703.05236.
- [341] 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.
- 3629 [342] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* 11 (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [343] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* 11 (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [344] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [345] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [346] 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:0611129.
- 3644 [347] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 3645 061, doi:10.1007/JHEP12 (2012) 061, arXiv:1209.6215.
- [348] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* **03** (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- [349] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet
 production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile,
 Italy, March 21-28, 2015. 2015. arXiv:1505.05704.

- [350] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", *JHEP* **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- ³⁶⁵⁵ [351] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", Comput. Phys. Commun. **191** (2015) ³⁶⁵⁶ 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- ³⁶⁵⁷ [352] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", ³⁶⁵⁸ JHEP **07** (2011) 018, doi:10.1007/JHEP07 (2011) 018, arXiv:1105.0020.
- ³⁶⁵⁹ [353] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.
- [354] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C 76 (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- 3664 [355] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur. 3665 Phys. J. C 74 (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, 3666 arXiv:1404.5630.
- 3667 [356] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- ³⁶⁶⁹ [357] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST **12** (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- ³⁶⁷² [358] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.
- 3674 [359] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* 3675 **59** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.
- 3676 [360] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST 6 (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- ³⁶⁷⁹ [361] 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.
- ³⁶⁸¹ [362] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) 05011.
- [363] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", JHEP
 12 (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.
- [364] J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05(2016) 156, arXiv:1603.00027.
- ³⁶⁸⁸ [365] 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.

References 117

[366] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

- [3693] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- ³⁶⁹⁶ [368] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST **11** (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.
- [369] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST 8 (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [370] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.
- [371] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.

 3707 arXiv:1009.1003.
- [372] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8 \text{ TeV}$ ", JINST 10 (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [373] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- ³⁷¹³ [374] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- In [375] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ ", Phys. Lett. B 735 (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- [376] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", *Phys. Rev. D* **91** (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.
- [377] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s}=7\,\text{TeV}$ ", JHEP **06** (2014) 120, doi:10.1007/JHEP06 (2014) 120, arXiv:1402.1521.
- [378] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}=8\,\text{TeV}$ ", Phys.~Rev.~Lett.~ 112 (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- [379] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", Phys. Lett. B 763 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

```
[380] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions
3731
         at \sqrt{s} = 13 \text{ TeV}'', Phys. Lett. B 766 (2017) 268,
3732
         doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.
3733
    [381] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs
3734
         Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working
3735
         Group", Technical Report CERN-2013-004, Geneva, 2013.
3736
         doi:10.5170/CERN-2013-004, arXiv:1307.1347.
3737
    [382] A. L. Read, "Presentation of search results: the CL<sub>s</sub> technique", J. Phys. G 28 (2002)
3738
         2693, doi:10.1088/0954-3899/28/10/313.
3739
    [383] T. Junk, "Confidence level computation for combining searches with small statistics",
3740
         Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002 (99) 00498-2,
3741
         arXiv:hep-ex/9902006.
3742
    [384] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for
3743
         likelihood-based tests of new physics", Eur. Phys. J. C 71 (2011) 1554,
3744
         doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3.
3745
    [385] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse
         energy signals using simplified s-channel models of dark matter", (2016).
3747
         arXiv:1603.04156.
3748
    [386] SuperCDMS Collaboration, "New results from the search for low-mass weakly
3749
         interacting massive particles with the CDMS low ionization threshold experiment",
3750
         Phys. Rev. Lett. 116 (2016) 071301, doi:10.1103/PhysRevLett.116.071301,
3751
         arXiv:1509.02448.
3752
    [387] LUX Collaboration, "Results from a search for dark matter in the complete LUX
3753
         exposure", Phys. Rev. Lett. 118 (2017) 021303,
3754
         doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648.
3755
    [388] XENON Collaboration, "First dark matter search results from the XENON1T
3756
         experiment", Phys. Rev. Lett. 119 (2017) 181301,
3757
         doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655.
3758
    [389] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II
3759
         experiment", Phys. Rev. Lett. 119 (2017) 181302,
3760
         doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.
3761
    [390] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold
3762
         CRESST-II detector", Eur. Phys. J. C 76 (2016) 25,
3763
         doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.
3764
3765
    Search for associated production of dark matter with a Higgs boson decaying to a pair of bot-
    tom quarks in pp collisions at sqrt(s) = 13 \text{ TeV}
3767
    [cern]The CMS Collaboration
    2018/08/17
3768
```

A. Introduction 119

A Introduction

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for physics beyond the standard model (SM) [456–458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consistent with the hypothesis that DM is primarily composed of weakly interacting massive particles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. Assuming the pair production of DM particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signature. If DM particles are produced in association with a detectable SM particle, which could be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding protons, or through new effective couplings between DM and SM particles, their existence could be inferred via a large transverse momentum imbalance in the collision event.

While ISR production of the SM Higgs boson (h) [460–462] is highly suppressed due to the Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs boson and DM particles can occur if the Higgs boson takes part in the interaction producing the DM particles [463–465]. Such a production mechanism would allow to directly probe the structure of the effective DM-SM coupling.

In this paper, we present a search for DM production in association with an SM Higgs boson that decays into a bottom quark-antiquark pair $(b\bar{b})$. As the $h\to b\bar{b}$ decay mode has the largest branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest signal yield. The search is performed using the data set collected by the CMS experiment [466] at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of approximately 35.9 fb⁻¹. Results are interpreted in terms of two simplified models predicting this signature. The first one is type-2 two Higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi\bar{\chi}$. The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM particles. Representative Feynman diagrams for the two models are presented in Fig. 50.

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no significant deviation from SM predictions [468]. In contrast to previously explored 2HDM models [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar a, assumed to be the observed 125 GeV Higgs boson, and a heavy neutral CP-even scalar a, that are the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mixing angle a; a heavy neutral CP-odd pseudoscalar a and a light neutral CP-odd pseudoscalar a, that are the result of the mixing of the CP-odd mediator a0 with the CP-odd Higgs, with a1 representing the associated mixing angle; and two heavy charged scalars a2 with identical mass.

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the

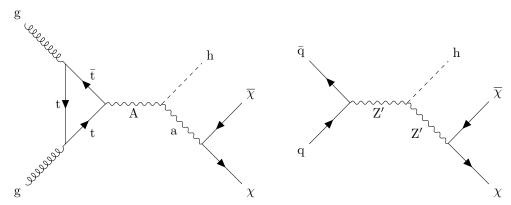


Figure 43: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z' model (right).

charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H}=m_{\rm H^\pm}=m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_\chi=10\,{\rm GeV}$.

The baryonic Z' model [465] is an extension of the SM with an additional $U(1)_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding $U(1)_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the $U(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_BZ'Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin\theta/v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_q , is 0.25 and the Z' coupling to DM, g_{χ} , is set to 1. This is well below the bounds g_q , $g_{\chi} \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin \theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'}\sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and m_{χ} , which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, together with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically independent data samples are used to predict the hadronic recoil distribution for these SM processes that constitute the largest sources of background. Both signal and background contributions to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

B The CMS detector

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum (p_T) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of

3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta| < 3$. The steel and quartz-fiber forward Cherenkov hadron calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and covers $|\eta| < 2.4$. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than $4\,\mu s$, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to 1 kHz.

C Simulated data samples

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG V2 [472-474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the p_T of the respective boson with NLO QCD corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

D Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_{\rm T}^{i2}$, where $p_{\rm T}^i$ is the transverse momentum of the $i^{\rm th}$ track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, large-area jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central ($\eta < 2.4$). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$ is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , will affect the shape of the distribution of $m_{\rm SD}$ for the background processes. In this search, the value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore desirable to preserve a smoothly falling jet mass distribution as a function of $p_{\rm T}$. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with $p_{\rm T}$, decorrelating the N_2 variable as a function of ρ and $p_{\rm T}$ would be the most appropriate procedure. The decorrelation strategy described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494].

This search also utilizes narrow jets clustered with the anti- $k_{\rm T}$ algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta|$ < 2.4).

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in

E. Event selection 123

the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the consistency with originating from the primary vertex are imposed. The reconstructed electron is required to be within $|\eta| < 2.5$, excluding the transition region 1.44 $< |\eta| < 1.57$ between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons and poorly measured mesons tagged as muons [497]. These additional criteria include requirements on the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates with $|\eta|$ < 2.4 are considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are reconstructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton decays. The τ_h candidates with $p_T > 18$ GeV and $|\eta| < 2.3$ are considered [496, 498, 499]. Photon candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15 \,\text{GeV}$ and $|\eta| < 2.5$.

The missing transverse momentum \vec{p}_{T}^{miss} is defined as the negative vectorial sum of the p_{T} of all the reconstructed PF candidates. Its magnitude is denoted as p_{T}^{miss} . Corrections to jet momenta are propagated to the p_{T}^{miss} calculation as well as event filters [500] are used to remove spurious high p_{T}^{miss} events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

E Event selection

3940

3941

3943

3944

3945

3948

3949

3950

3951

3952

3953

3954

3956

3957

3958

3959

3960

3961

3963

3965

3966

3967

3969

3970

3971

3972

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}^{\rm miss}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T, \rm trig}^{\rm miss}$ or $H_{\rm T}^{\rm miss}$, where $p_{\rm T, trig}^{\rm miss}$ ($H_{\rm T}^{\rm miss}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\rm T}$ of all the particles (jets with $p_{\rm T} > 20\,{\rm GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\rm T, trig}^{\rm miss}$ calculation. Thresholds on $p_{\rm T, trig}^{\rm miss}$ and $H_{\rm T}^{\rm miss}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\rm T, trig}^{\rm miss}$ and $H_{\rm T}^{\rm miss}$ are reffered to as $p_{\rm T}^{\rm miss}$ triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T} > 200\,{\rm GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\rm SD} < 150\,{\rm GeV}$ and $N_2^{\rm DDT} < 0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than $200\,{\rm GeV}$, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (\vec{U}) must be larger than 0.4

Table 19: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, t t , W+jets	0	0	pass
Single-lepton, b-tagged	tī, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

rad to reject multijet events that mimic signal events. Vetoes on τ_h candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from $t\bar{t}$ events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_T > 10 \,\text{GeV}$ and $|\eta| < 2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between 60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of the trigger efficiency, at least one of the two electrons must have $p_{\rm T} > 40\,{\rm GeV}$, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from W+jets and semileptonic $t\bar{t}$ events. To predict these backgrounds, four single-lepton samples are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target $t\bar{t}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the p_T^{miss} trigger triggers described above, as well as single electron events are selected using the same single-electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and isolation requirements. In addition, samples with a single electron need to have $p_T^{miss} > 50 \, \text{GeV}$ to avoid a large contamination from multijet events.

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

F Signal extraction

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the $p_{\rm T}^{\rm miss}$ and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and

are given by the ratio of the prediction for a given bin in $p_{\rm T}^{\rm miss}$ in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the $t\bar{t}$ process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt backgrounds take into account the im-4029 pact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron 4030 control samples, the additional requirement on p_T^{miss} . Since the CRs with no b-tagged AK4 jets 4031 and a double-b-tagged CA15 jet also have significant contributions from the tt process, transfer 4032 factors to predict this contamination from tt events are also imposed between the single-lepton 4033 CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contamination from W+jets production in the tt CR with events that fail the double-b tag requirement. 4035 Likewise, the Z+jets background prediction in the signal region is connected to the and dilep-4036 ton CRs via transfer factors. They account for the difference in the branching fractions of the 4037 $Z \rightarrow \nu \nu$ and the $Z \rightarrow \ell \ell$ decays and the impacts of lepton acceptances and selection efficiencies.

G Systematic uncertainties

4039

4042

4043

4044

4045

4046

4047

4048

4050

4052

4053

4054

4055

4057

4058

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncertainties of the search. They can either affect the rate or the shape of $p_{\rm T}^{\rm miss}$ (U) for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal distribution. The list of the systematic uncertainties considered in this search is presented in Table 23. To better estimate their impact on the results, uncertainties from a similar source (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties have been ordered according to decreasing improvement in the expected limit obtained when removing the group from the list of nuisances included in the likelihood fit. The description of each single uncertainty in the text follows the same order.

Scale factors are used to correct for the differences in the double-b tagger misidentification efficiencies between data and prediction from simulation for W/Z+jets production and for the production. These scale factors are measured by simultaneously fitting events that pass or fail the double-b tag requirement. The correlation between the double-b tagger and the $p_{\rm T}^{\rm miss}$ (or U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate independently from each other within a constraint that depends on the recoil value. Such dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. $p_{\rm T}^{\rm miss}/U$. This shape uncertainty in the double-b scale factors measurement is the one that has the largest impact on the limits on the signal cross section.

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and $t\bar{t}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the $p_{\rm T}$ of the CA15 jet. This shape uncertainty has been derived through a measurement performed using a sample enriched in multijet events with double-muon-tagged $g \rightarrow b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substruc-

Table 20: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

Ctti	Т	D	Tours and any association to
Systematic uncertainty	Туре	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, tt̄	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, tt̄	1.9%
Double-b tagging	shape	SM h, signal	1.2%
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\rm T}^{\rm miss}$ magnitude	5%	all	0.7%
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%
$p_{\rm T}^{\rm miss}$ trigger muon multiplicity	shape	Z+jets, W+jets	
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	< 0.5%
single-electron trigger	1%	all	
AK4 b tagging	shape	all	< 0.5%
au lepton veto	3%	all	< 0.5%
Lepton efficiency	1% per lepton	all	< 0.5%
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%
Renorm./fact. scales	shape	SM h	
PDF	shape	SM h	< 0.59/
Multijet normalization	100%	multijet	< 0.5%
Theoretical cross section	20%	t, diboson	

ture variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes where the decay of a resonance inside the CA15 jet cone is expected. Such processes include signal production together with SM h and diboson production. The uncertainty has been derived from the efficiency measurement obtained by performing a fit in a control sample enriched in semi-leptonic $t\bar{t}$ events, where the CA15 jet originates from the W boson that comes from the hadronically decaying top quark.

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in p_T^{miss} magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, obtained by varying those parameters in simulation event-by-event

The $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. Differences on the order of 2% are observed between single-muon and dimuon events at lower U values and they are sources of an additional systematic uncertainty in the transfer factors for those processes whose prediction relies on data events in the single-muon and dimuon CRs (tt̄, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the identified muon, they can change the shape of the U distribution and are thus treated as shape uncertainties. The $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ trigger efficiency is parametrized as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$. The uncertainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies of the single-electron triggers are parametrized as a function of the electron p_{T} and η and an

4089 associated 1% systematic uncertainty is added into the fit.

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization 4105 scales and PDFs are included as shape variations. Being a negligible background source, an un-4106 certainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a 4107 sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by 4108 requiring $p_{\rm T}^{\rm miss} > 250$ GeV, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ 4109 and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty 4110 in QCD multijet yields in the signal region, while separate nuisance parameters are introduced 4111 for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR 4113 and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as 4114 measured by CMS in Refs. [509, 510] and correlated across the SR and CRs.

H Results

4116

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficiencies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z' model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results of this search are interpreted in terms of upper limits on the signal strength modifier $\mu =$

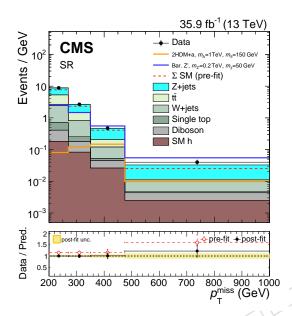


Figure 44: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

Table 21: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
619	214	59	21
5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3
184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3
	248.9 ± 22.2 199.2 ± 13.5 121.6 ± 21.6 21.0 ± 4.2 16.0 ± 3.1 12.6 ± 1.4 619.3 ± 20.1 619 5.7 ± 0.6	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$

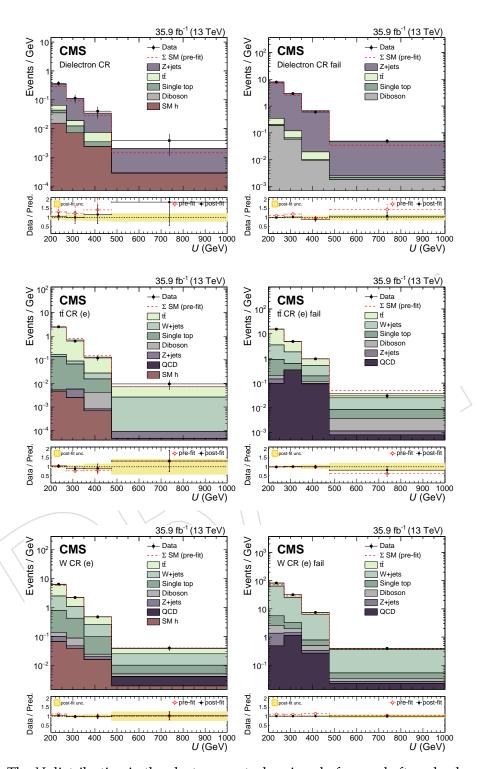


Figure 45: The *U* distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

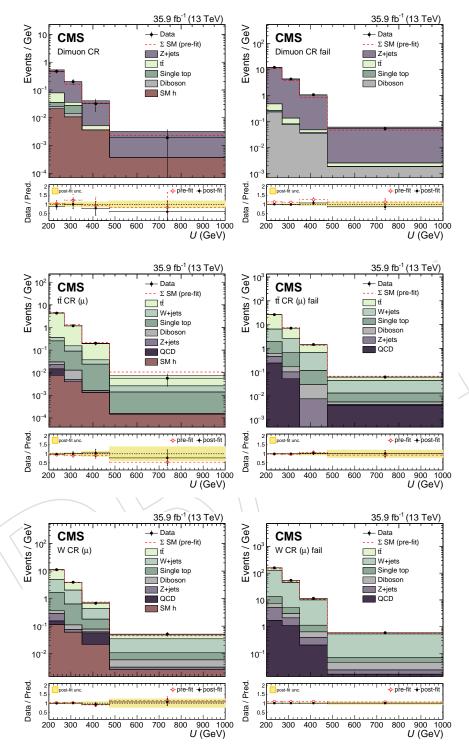


Figure 46: The *U* distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

 $\sigma/\sigma_{\text{theory}}$, where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

Figure 54 shows the upper limits on μ for the three scans (m_A , $\sin \theta$, and $\tan \beta$) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, $\sin \theta = 0.35$ and $\tan \beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming $\tan \beta = 1$. Also excluded are $\tan \beta$ values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

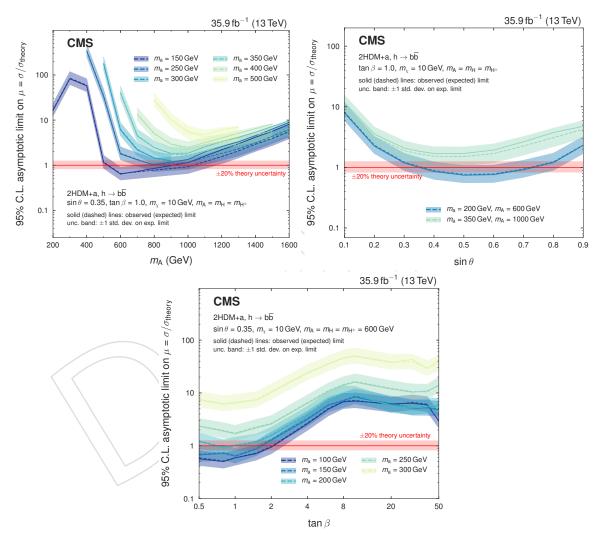


Figure 47: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus.

4151

4153

4154

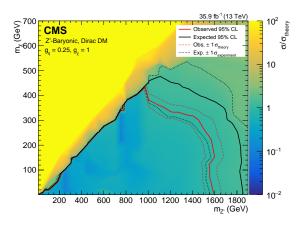


Figure 48: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

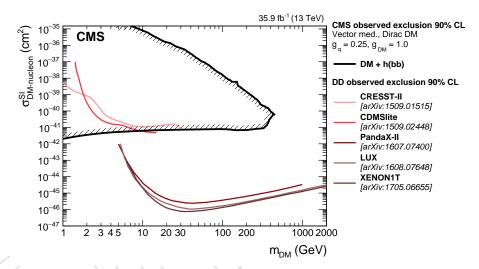


Figure 49: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_{χ} . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2\mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},$$
 (7)

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5\,\text{GeV}$.

I. Summary 133

I Summary

4155

4173

A search for the associated production of dark matter (DM) particles with a Higgs boson de-4156 caying into a pair of bottom quarks is presented. No significant deviation from the predic-4157 tions of the standard model (SM) is observed, and upper limits on the production cross section 4158 predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar 4159 boson a (2HDM+a) and the baryonic Z' model are established. They constitute the most strin-4160 gent collider exclusions placed on the parameters in these models so far. For the nominal choice of the mixing angle $\sin \theta$ and $\tan \beta$ in the 2HDM+a model, the search excludes masses 4162 $500 < m_A < 900 \,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150 \,\text{GeV}$. 4163 Scanning over $\sin \theta$ with $\tan \beta = 1$, we exclude $0.35 < \sin \theta < 0.75$ for $m_A = 600$ GeV and 4164 $m_{\rm a}=200\,{\rm GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_{\rm A}=600\,{\rm GeV}$ and $m_a = 100 (150) \,\text{GeV}$ and $\sin \theta > 0.35$. In all 2HDM+a interpretations, a DM mass of 4166 $m_{\chi} = 10 \,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV 4167 for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The rein-4168 terpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross 4169 section yields a higher sensitivity for m_{χ} < 5 GeV than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed 4171 experimentally for the first time. 4172

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent perfor-4174 mance of the LHC and thank the technical and administrative staffs at CERN and at other 4175 CMS institutes for their contributions to the success of the CMS effort. In addition, we grate-4176 fully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Fi-4178 nally, we acknowledge the enduring support for the construction and operation of the LHC 4179 and the CMS detector provided by the following funding agencies: BMWFW and FWF (Aus-4180 tria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); 4181 CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); 4182 RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Fin-4183 land, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Ger-4184 many); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI 4185 (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New 4187 Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, 4188 RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER 4189 (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United 4191 Kingdom); DOE and NSF (USA). 4192

Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the

Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Severo Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

References

4209

- [391] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and Constraints", *Phys. Rept.* **405** (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:0404175.
- [392] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",

 Ann. Rev. Astron. Astrophys. 48 (2010) 495,

 doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- 4216 [393] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", Ann. Rev. Astron. Astrophys. 49 (2011) 155,
 4218 doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.
- [394] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron.

 Astrophys. 594 (2016) A13, doi:10.1051/0004-6361/201525830,

 arXiv:1502.01589.
- 4222 [395] ATLAS Collaboration, "Observation of a new particle in the search for the standard 4223 model higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, 4224 doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [396] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [397] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}=7$ and 8 TeV", JHEP **06** (2013) 81, doi:10.1007/JHEP06(2013) 081, arXiv:1303.4571.
- [398] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, arXiv:1311.1511.
- [399] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", JHEP 06 (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- [400] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", *Phys.***Rev. D **89** (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- 4238 [401] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

References 135

[402] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05 (2017) 138, arXiv:1701.07427.

- [403] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}=7$ and 8 TeV", JHEP **08** (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
- [404] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- [405] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10 (2017) 180, arXiv:1703.05236.
- [406] 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.
- [407] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", JHEP 11 (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [408] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* 11 (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [409] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [410] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", Phys. Rev. Lett. 110 (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [411] 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:0611129.
- [412] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12 (2012) 061, arXiv:1209.6215.
- [413] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* **03** (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- [414] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, Italy, March 21-28, 2015. 2015. arXiv:1505.05704.

- [415] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", JHEP **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- [416] 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.
- [417] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", ### JHEP 07 (2011) 018, doi:10.1007/JHEP07 (2011) 018, arXiv:1105.0020.
- [418] NNPDF Collaboration, "Parton distributions for the LHC Run II", JHEP **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.
- [419] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C 76 (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [420] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur. Phys. J. C 74 (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.
- 4295 [421] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- [422] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST 12 (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- 4300 [423] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.
- 4302 [424] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* 4303 **59** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.
- 4304 [425] CMS Collaboration, "Determination of jet energy calibration and transverse momentum 4305 resolution in CMS", JINST 6 (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, 4306 arXiv:1107.4277.
- 4307 [426] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", *JHEP* **05** (2014) 146, 4308 doi:10.1007/JHEP05 (2014) 146, arXiv:1402.2657.
- 4309 [427] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", JINST 13 (2018) 05011.
- [428] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", *JHEP* 12 (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.
- [429] J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05(2016) 156, arXiv:1603.00027.
- 4316 [430] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k_t jet clustering algorithm", *JHEP* **04** 4317 (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

References 137

[431] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

- [432] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [433] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST 11 (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.
- [434] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST 8 (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [435] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.
- 4333 [436] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced
 4334 Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.
 4335 arXiv:1009.1003.
- [437] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV", JINST 10 (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [438] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period",
 CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [439] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ ", *Phys. Lett. B* **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- 4344 [440] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ ", Phys. Lett. B 735 (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- [441] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", Phys. Rev. D 91 (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.
- [442] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s}=7\,\text{TeV}$ ", JHEP **06** (2014) 120, doi:10.1007/JHEP06 (2014) 120, arXiv:1402.1521.
- 4353 [443] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}=8\,\text{TeV}''$, Phys.~Rev.~Lett.~ 112 (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- [444] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", Phys. Lett. B 763 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

2018/08/17

4396

```
[445] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions
4359
         at \sqrt{s} = 13 \text{ TeV}'', Phys. Lett. B 766 (2017) 268,
4360
         doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.
4361
    [446] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs
4362
         Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working
4363
         Group", Technical Report CERN-2013-004, Geneva, 2013.
4364
         doi:10.5170/CERN-2013-004, arXiv:1307.1347.
4365
    [447] A. L. Read, "Presentation of search results: the CL<sub>s</sub> technique", J. Phys. G 28 (2002)
4366
         2693, doi:10.1088/0954-3899/28/10/313.
4367
    [448] T. Junk, "Confidence level computation for combining searches with small statistics",
4368
         Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002 (99) 00498-2,
4369
         arXiv:hep-ex/9902006.
4370
    [449] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for
4371
         likelihood-based tests of new physics", Eur. Phys. J. C 71 (2011) 1554,
4372
         doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3.
4373
    [450] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse
         energy signals using simplified s-channel models of dark matter", (2016).
4375
         arXiv:1603.04156.
4376
    [451] SuperCDMS Collaboration, "New results from the search for low-mass weakly
4377
         interacting massive particles with the CDMS low ionization threshold experiment",
4378
         Phys. Rev. Lett. 116 (2016) 071301, doi:10.1103/PhysRevLett.116.071301,
4379
         arXiv:1509.02448.
4380
    [452] LUX Collaboration, "Results from a search for dark matter in the complete LUX
4381
         exposure", Phys. Rev. Lett. 118 (2017) 021303,
4382
         doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648.
4383
    [453] XENON Collaboration, "First dark matter search results from the XENON1T
4384
         experiment", Phys. Rev. Lett. 119 (2017) 181301,
4385
         doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655.
4386
    [454] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II
4387
         experiment", Phys. Rev. Lett. 119 (2017) 181302,
4388
         doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.
4389
    [455] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold
4390
         CRESST-II detector", Eur. Phys. J. C 76 (2016) 25,
4391
         doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.
4392
4393
    Search for associated production of dark matter with a Higgs boson decaying to a pair of bot-
    tom quarks in pp collisions at sqrt(s) = 13 \text{ TeV}
4395
    [cern]The CMS Collaboration
```

A. Introduction 139

A Introduction

Astrophysical evidence for dark matter (DM) is one of the most compelling motivations for physics beyond the standard model (SM) [456-458]. Cosmological observations demonstrate that around 85% of the matter in the universe is comprised of DM [459] and are largely consis-tent with the hypothesis that DM is primarily composed of weakly interacting massive parti-cles (WIMPs). If nongravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. Assuming the pair production of DM particles in hadron collisions happens through a spin-0 or spin-1 bosonic mediator coupled to the initial-state particles, the DM particles leave the detector without a measurable signa-ture. If DM particles are produced in association with a detectable SM particle, which could be emitted as initial-state radiation (ISR) from the interacting constituents of the colliding protons, or through new effective couplings between DM and SM particles, their existence could be inferred via a large transverse momentum imbalance in the collision event.

While ISR production of the SM Higgs boson (h) [460–462] is highly suppressed due to the Yukawa-like nature of its coupling strength to fermions, the associated production of a Higgs boson and DM particles can occur if the Higgs boson takes part in the interaction producing the DM particles [463–465]. Such a production mechanism would allow to directly probe the structure of the effective DM-SM coupling.

In this paper, we present a search for DM production in association with an SM Higgs boson that decays into a bottom quark-antiquark pair $(b\bar{b})$. As the $h\to b\bar{b}$ decay mode has the largest branching fraction of all Higgs boson decay modes allowed in the SM, it provides the largest signal yield. The search is performed using the data set collected by the CMS experiment [466] at the CERN LHC at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of approximately 35.9 fb⁻¹. Results are interpreted in terms of two simplified models predicting this signature. The first one is type-2 two Higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) [467]. The a boson mixes with the scalar and pseudoscalar partners of the SM Higgs boson, and decays into a pair of DM particles, $\chi\bar{\chi}$. The second model is a baryonic Z' model (baryonic Z') [465] where a vector mediator Z' is exchanged in the s-channel, radiates a Higgs boson, and subsequently decays into two DM particles. Representative Feynman diagrams for the two models are presented in Fig. 50.

In the 2HDM+a model, the DM particle candidate χ is a Dirac fermion that can couple to SM particles only through a spin-0, pseudoscalar mediator. Since the couplings of the new spin-0 mediator to SM gauge bosons are strongly suppressed, the 2HDM+a model is consistent with the measurements of the SM Higgs boson production and decay modes, which so far show no significant deviation from SM predictions [468]. In contrast to previously explored 2HDM models [464, 469, 470], the 2HDM+a framework ensures gauge invariance and renormalizability. In this model, there are six mass eigenstates: a light neutral charge-parity (CP)-even scalar a, assumed to be the observed 125 GeV Higgs boson, and a heavy neutral CP-even scalar a, that are the result of the mixing of the neutral CP-even weak eigenstates with the corresponding mixing angle a; a heavy neutral CP-odd pseudoscalar a, and a light neutral CP-odd pseudoscalar a, that are the result of the mixing of the CP-odd mediator a0 with the CP-odd Higgs, with a1 representing the associated mixing angle; and two heavy charged scalars a2 with identical mass.

The masses of the two CP-odd Higgs bosons, the angle θ , and the ratio of vacuum expectation values of the two CP-even Higgs bosons $\tan \beta$ are varied in this search. Perturbativity and unitarity put restrictions on the magnitudes and the signs of the three quartic couplings λ_3 , λ_{P1} , λ_{P2} , and we therefore set their values to $\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3$ [467]. Masses of the

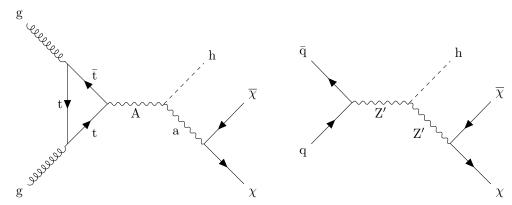


Figure 50: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z' model (right).

charged Higgs bosons and of the heavy CP-even Higgs boson are assumed to be the same as the mass of the heavy pseudoscalar, i.e., $m_{\rm H}=m_{\rm H^\pm}=m_{\rm A}$. The DM particle χ is assumed to have a mass of $m_\chi=10\,{\rm GeV}$.

The baryonic Z' model [465] is an extension of the SM with an additional $U(1)_B$ Z' gauge boson that couples to the baryon number B. The model predicts the existence of a new baryonic fermion (or scalar) that is neutral under SM gauge symmetries and stable due to the corresponding $U(1)_B$ symmetry. The state therefore serves as a good DM candidate. To generate the Z' mass, a "baryonic Higgs" scalar is introduced to spontaneously break the $U(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B , with a coupling of $h_B Z' Z'$ and vacuum expectation value of v_B . The Z' and SM Higgs boson, h, interact with a coupling strength of $g_{hZ'Z'} = m_{Z'}^2 \sin \theta / v_B$, where θ is the h- h_B mixing angle. The chosen value for the Z' coupling to quarks, g_{q_1} is 0.25 and the Z' coupling to DM, g_{χ_1} is set to 1. This is well below the bounds $g_{q_1}g_{\chi} \sim 4\pi$ where perturbativity and the validity of the effective field theory break down [465]. The mixing angle θ is assumed to have $\sin \theta = 0.3$. It is also assumed that $g_{hZ'Z'}/m_{Z'} = 1$, which implies $v_B = m_{Z'} \sin \theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{Z'}$ and m_{χ} , which are varied in this search.

Signal events are characterized by a large imbalance in the transverse momentum (or hadronic recoil), which indicates the presence of invisible DM particles, and by hadronic activity consistent with the production of an SM Higgs boson that decays into a $b\bar{b}$ pair. Thus, our search strategy is to impose requirements on the mass of the reconstructed Higgs boson candidate, together with the identification of the products of hadronization of the two b quarks produced in the Higgs boson decay, to define a data sample that is expected to be enriched in signal events. Several different SM processes can mimic this topology, such as top quark pair production and the production of a vector boson (V) in association with multiple jets. Statistically independent data samples are used to predict the hadronic recoil distribution for these SM processes that constitute the largest sources of background. Both signal and background contributions to the data are extracted with a likelihood fit to the hadronic recoil distribution, performed simultaneously in all the different analysis subsamples.

B The CMS detector

The CMS detector, described in detail in Ref. [466], is a multipurpose apparatus designed to study high-transverse momentum (p_T) processes in proton-proton (pp) and heavy ion collisions. A superconducting solenoid occupies its central region, providing a magnetic field of

4481

4482

4483

4485

4486

4487

4488

4489

4491

4492

4493

4494

4495

4496

4497

4498

4500

4501

4502

4503

4504

4505

4506

4507

4515

4517

4518

4519

4520

3.8 T parallel to the beam direction. Charged-particle trajectories are measured using silicon pixel and strip trackers that cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and extend to $|\eta| < 3$. The steel and quartz-fiber forward Cherenkov hadron calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid and covers $|\eta| < 2.4$. Online event selection is accomplished via the two-tiered CMS trigger system. The first level is designed to select events in less than $4\,\mu\text{s}$, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to 1 kHz.

C Simulated data samples

The signal processes are simulated at leading order (LO) accuracy in quantum chromodynamics (QCD) perturbation theory using the MADGRAPH5_aMC@NLO v2.4.2 [471] program. To model the contributions from SM Higgs boson processes as well as from the tt and single top quark backgrounds, we use the POWHEG V2 [472-474] generator. These processes are generated at the next-to-leading order (NLO) in QCD. The tt production cross section is further corrected using calculations at the next-to-next-to-leading order (NNLO) in QCD including corrections for soft-gluon radiation estimated with next-to-next-to-leading logarithmic accuracy [475]. Events with multiple jets produced via the strong interaction (referred to as QCD multijet events) are generated at LO using MADGRAPH5_aMC@NLO v2.2.2 with up to four partons in the matrix element calculations. The MLM prescription [476] is used for matching these partons to parton shower jets. Simulated samples of Z+jets and W+jets processes are generated at LO using MADGRAPH5_aMC@NLO v2.3.3. Up to four additional partons are considered in the matrix element and matched to their parton showers using the MLM technique. The V+jets samples are corrected by weighting the p_T of the respective boson with NLO QCD corrections obtained from large samples of events generated with MADGRAPH5_aMC@NLO and the FxFx merging technique [477] with up to two additional jets stemming from the matrix element calculations. These samples are further corrected by applying NLO electroweak corrections [478–480] that depend on the boson p_T . Predictions for the SM diboson production modes WW, WZ, and ZZ are obtained at LO with the PYTHIA 8.205 [481] generator and normalized to NLO accuracy using MCFM [482].

The LO or NLO NNPDF 3.0 parton distribution functions (PDFs) [483] are used, depending on the QCD order of the generator used for each physics process. Parton showering, fragmentation, and hadronization are simulated with PYTHIA 8.212 using the CUETP8M1 underlying event tune [484, 485]. Interactions of the resulting final state particles with the CMS detector are simulated using the GEANT4 program [486]. Additional inelastic pp interactions in the same or a neighboring bunch crossing (pileup) are included in the simulation. The pileup distribution is adjusted to match the corresponding distribution observed in data.

D Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_{\rm T}^{i2}$, where $p_{\rm T}^i$ is the transverse momentum of the $i^{\rm th}$ track associated with the vertex, is selected as the primary event vertex. The offline selection requires all events to have at least one primary vertex reconstructed within a 24 cm window along the z-axis around the nominal interaction point, and a transverse distance from the nominal interaction region less than 2 cm.

The particle-flow (PF) [487] algorithm aims to reconstruct all the physics objects described in this section. At large Lorentz boosts, the two b quarks from the Higgs boson decay may produce jets that overlap and make their individual reconstruction difficult. In this search, large-area jets clustered from PF candidates using the Cambridge–Aachen algorithm [488] with a distance parameter of 1.5 (CA15 jets) are utilized to identify the Higgs boson candidate. The large cone size is chosen in order to include events characterized by the presence of Higgs bosons with medium boost (p_T of the order of 200 GeV). To reduce the impact of particles arising from pileup interactions, the four-vector of each PF candidate is scaled with a weight calculated with the pileup per particle identification (PUPPI) algorithm [489] prior to the clustering. The absolute jet energy scale is corrected using calibrations derived from data [490]. The CA15 jets are also required to be central ($\eta < 2.4$). The "soft-drop" (SM) jet grooming algorithm [491] is applied to the jets to remove the wide-angle ISR and soft radiation emerging from the underlying event. We refer to the groomed mass of the CA15 jet as m_{SD} .

The ability to identify two b quarks inside a single CA15 jet is crucial for this search. A likelihood for the CA15 jet to contain two b quarks is derived by combining the information from primary and secondary vertices and tracks in a multivariate discriminant optimized to distinguish the Higgs to $b\bar{b}$ decays from energetic quarks or gluons [492] that appear inside the CA15 jet cone. The working point chosen for this algorithm (the "double-b tagger") corresponds to an identification efficiency of 50% for a $b\bar{b}$ system with a p_T of 200 GeV, and a probability of about 10-13% for misidentifying CA15 jets originating from other combinations of quarks or gluons. The efficiency of the algorithm increases with the p_T of the $b\bar{b}$ system.

Energy correlation functions are used to identify the two-prong structure in the CA15 jet expected from a Higgs boson decay to two b quarks. The energy correlation functions are sensitive to correlations among the constituents of CA15 jets (the PF candidates) [493]. They are defined as N-point correlation functions (e_N) of the constituents' momenta, weighted by the angular separation of the constituents. As motivated in Ref. [493], the ratio $N_2 = e_3^{(\beta)}/(e_2^{(\beta)})^2$ is proposed as a two-prong tagger for the identification of the CA15 jet containing the Higgs boson decay products; the parameter β , which controls the weighting of the angles between constituent pairs in the computation of the N_2 variable, is chosen to be 1.0.

It is noted that requiring a jet to be two-pronged by using a jet substructure variable, such as N_2 , will affect the shape of the distribution of $m_{\rm SD}$ for the background processes. In this search, the value of $m_{\rm SD}$ is required to be consistent with the Higgs boson mass. To improve the rejection of QCD-like jets (i.e., jets that do not originate from a heavy resonance decay), it is therefore desirable to preserve a smoothly falling jet mass distribution as a function of $p_{\rm T}$. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\rm SD}^2/p_{\rm T}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with $p_{\rm T}$, decorrelating the N_2 variable as a function of ρ and $p_{\rm T}$ would be the most appropriate procedure. The decorrelation strategy described in Ref. [494] is applied, choosing a background efficiency of 20%, which corresponds to a signal efficiency of roughly 50%. This results in a modified tagging variable, which we denote as $N_2^{\rm DDT}$, where DDT is designing decorrelated taggers [494].

This search also utilizes narrow jets clustered with the anti- $k_{\rm T}$ algorithm [495], with a distance parameter of 0.4 (AK4 jets). Narrow jets originating from b quarks are identified using the combined secondary vertex (CSVv2) algorithm [492]. The working point used in this search has a b jet identification efficiency of 81%, a charm jet selection efficiency of 37%, and a 9% probability of misidentifying light-flavor jets [492]. Jets that are b tagged are required to be central ($|\eta|$ < 2.4).

Electron reconstruction requires the matching of a supercluster in the ECAL with a track in

E. Event selection 143

the silicon tracker. Identification criteria [496] based on the ECAL shower shape and the consistency with originating from the primary vertex are imposed. The reconstructed electron is required to be within $|\eta|$ < 2.5, excluding the transition region 1.44 < $|\eta|$ < 1.57 between the ECAL barrel and endcap. Muons candidates are selected by two different reconstruction approaches [497]: the one in which tracks in the silicon tracker are matched to a track segment in the muon detector, and the other one in which a track fit spanning the silicon tracker and muon detector is performed starting with track segments in the muon detector. Candidates that are found by both the approaches are considered as single candidates. Further identification criteria are imposed on muon candidates to reduce the number of misidentified hadrons and poorly measured mesons tagged as muons [497]. These additional criteria include requirements on the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates with $|\eta|$ < 2.4 are considered in this analysis. With electron and muon candidates, the minimum p_T is required to be 10 GeV. Isolation is required for both the objects. Hadronically decaying τ leptons, τ_h , are reconstructed using the hadron-plus-strips algorithm [498], which uses the charged hadron and neutral electromagnetic objects to reconstruct intermediate resonances into which the τ lepton decays. The τ_h candidates with $p_T > 18$ GeV and $|\eta| < 2.3$ are considered [496, 498, 499]. Photon candidates, identified by means of requirements on the ECAL energy distribution and its distance to the closest track, must have $p_T > 15 \,\text{GeV}$ and $|\eta| < 2.5$.

The missing transverse momentum $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the negative vectorial sum of the p_{T} of all the reconstructed PF candidates. Its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$. Corrections to jet momenta are propagated to the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation as well as event filters [500] are used to remove spurious high $p_{\mathrm{T}}^{\mathrm{miss}}$ events caused by instrumental noise in the calorimeters or beam halo muons [500]. The filters remove about 1% of signal events.

E Event selection

4568

4569

4571

4572

4573

4576

4577

4578

4579

4580

4581

4582

4584

4585

4586

4592

4593

4594

4595

4597

4598

4599

4600

Signal events are characterized by a high value of $p_{\rm T}^{\rm miss}$, no isolated leptons or photons, and a CA15 jet identified as a Higgs boson candidate. In the signal region (SR) described below, the dominant background contributions arise from Z+jets, W+jets, and tt̄ production. To predict the $p_{\rm T}^{\rm miss}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the tt̄ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the $p_{\rm T}^{\rm miss}$ of the lepton(s) from the $p_{\rm T}^{\rm miss}$ computation in CRs, is used as a proxy for the $p_{\rm T}^{\rm miss}$ distribution of the main background processes in the SR. Predictions for other backgrounds are obtained from simulation.

Events are selected online by requiring large values of $p_{T, \rm trig}^{\rm miss}$ or $H_{\rm T}^{\rm miss}$, where $p_{\rm T, trig}^{\rm miss}$ ($H_{\rm T}^{\rm miss}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\rm T}$ of all the particles (jets with $p_{\rm T} > 20\,{\rm GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\rm T, trig}^{\rm miss}$ calculation. Thresholds on $p_{\rm T, trig}^{\rm miss}$ and $H_{\rm T}^{\rm miss}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\rm T, trig}^{\rm miss}$ and $H_{\rm T}^{\rm miss}$ are reffered to as $p_{\rm T}^{\rm miss}$ triggers. For CRs that require the presence of electrons, at least one electron is required by the online selections. These set of requirements are referred to as single-electron triggers.

A common set of preselection criteria is used for all regions. The presence of exactly one CA15 jet with $p_{\rm T} > 200\,{\rm GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\rm SD} < 150\,{\rm GeV}$ and $N_2^{\rm DDT} < 0$. In the SR (CRs), $p_{\rm T}^{\rm miss}$ (U) has to be larger than $200\,{\rm GeV}$, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $p_{\rm T}^{\rm miss}$ (\vec{U}) must be larger than 0.4

Table 22: Event selection criteria defining the signal and the control regions. These criteria are applied in addition to the preselection that is common to all regions, as described in the text.

Region	Main background process	Additional AK4 b tag	Leptons	Double-b tag
Signal	Z+jets, t t , W+jets	0	0	pass
Single-lepton, b-tagged	t t , W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, t t	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

rad to reject multijet events that mimic signal events. Vetoes on τ_h candidates and photon candidates are applied, and the number of AK4 jets that do not overlap with the CA15 jet must be smaller than two. This significantly reduces the contribution from $t\bar{t}$ events.

Events that meet the preselection criteria described above are split into the SR and the different CRs based on their lepton multiplicity and the presence of a b-tagged AK4 jet not overlapping with the CA15 jet, as summarized in Table 22. For the SR, events are selected if they have no isolated electrons (muons) with $p_{\rm T} > 10\,{\rm GeV}$ and $|\eta| < 2.5$ (2.4). The previously described double-b tag requirement on the Higgs boson candidate CA15 jet is imposed.

To predict the $p_{\rm T}^{\rm miss}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\rm T}^{\rm miss}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\rm T} > 20\,{\rm GeV}$ and $p_{\rm T} > 10\,{\rm GeV}$ for the leading and trailing muon, respectively) with an invariant mass between 60 and 120 GeV. The leading muon has to satisfy tight identification and isolation requirements that result in an efficiency of 95%. Dielectron events are selected online using single-electron triggers. Two oppositely charged electrons with $p_{\rm T}$ greater than 10 GeV are required offline, and they must form an invariant mass between 60 GeV and 120 GeV. To be on the plateau of the trigger efficiency, at least one of the two electrons must have $p_{\rm T} > 40\,{\rm GeV}$, and it has to satisfy tight identification and isolation requirements that correspond to an efficiency of 70%.

Events that satisfy the SR selection due to the loss of a single lepton primarily originate from W+jets and semileptonic $t\bar{t}$ events. To predict these backgrounds, four single-lepton samples are used: single-electron and single-muon, with or without a b-tagged AK4 jet outside the CA15 jet. The single-lepton CRs with a b-tagged AK4 jet target $t\bar{t}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the p_T^{miss} trigger triggers described above, as well as single electron events are selected using the same single-electron triggers used for the dielectron events online selection. The electron (muon) candidate in these events is required to have $p_T > 40$ (20) GeV and has to satisfy tight identification and isolation requirements. In addition, samples with a single electron need to have $p_T^{miss} > 50$ GeV to avoid a large contamination from multijet events.

Each CR is further split into two subsamples depending on whether or not the CA15 jet satisfies the double-b tag requirement. This allows for an in situ calibration of the scale factor that corrects the simulated misidentification probability of the double-b tagger for the three main backgrounds to the one observed in data.

F Signal extraction

As mentioned in Section A, signal and background contributions to the data are extracted with a simultaneous binned likelihood fit (using the ROOSTAT package [501]) to the $p_{\rm T}^{\rm miss}$ and U distributions in the SR and the CRs. The dominant SM process in each CR is used to predict the respective background in the SR via transfer factors T. They are determined in simulation and

are given by the ratio of the prediction for a given bin in $p_{\rm T}^{\rm miss}$ in the SR and the corresponding bin in U in the CR, for the given process. This ratio is determined independently for each bin of the corresponding distribution.

For example, if $b\ell$ denotes the $t\bar{t}$ process in the b-tagged single-lepton control sample that is used to estimate the $t\bar{t}$ contribution in the SR, the expected number of $t\bar{t}$ events, N_i , in the i^{th} bin of the SR is then given by $N_i = \mu_i^{t\bar{t}}/T_i^{b\ell}$, where $\mu_i^{t\bar{t}}$ is a freely floating parameter included in the likelihood to scale the $t\bar{t}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and tt backgrounds take into account the im-4657 pact of lepton acceptances and efficiencies, the b tagging efficiency, and, for the single-electron 4658 control samples, the additional requirement on p_T^{miss} . Since the CRs with no b-tagged AK4 jets 4659 and a double-b-tagged CA15 jet also have significant contributions from the tt process, transfer 4660 factors to predict this contamination from tt events are also imposed between the single-lepton 4661 CRs with and without b-tagged AK4 jets. A similar approach is applied to estimate the contamination from W+jets production in the tt CR with events that fail the double-b tag requirement. 4663 Likewise, the Z+jets background prediction in the signal region is connected to the and dilep-4664 ton CRs via transfer factors. They account for the difference in the branching fractions of the 4665 $Z \rightarrow \nu \nu$ and the $Z \rightarrow \ell \ell$ decays and the impacts of lepton acceptances and selection efficiencies.

G Systematic uncertainties

4667

Nuisance parameters are introduced into the likelihood fit to represent the systematic uncer-4668 tainties of the search. They can either affect the rate or the shape of $p_T^{\text{miss}}(U)$ for a given process in the SR (CRs) and can be constrained in the fit. The shape uncertainties are incorporated by 4670 means of a prior Gaussian distribution, while the rate uncertainties are given a prior log-normal 4671 distribution. The list of the systematic uncertainties considered in this search is presented in 4672 Table 23. To better estimate their impact on the results, uncertainties from a similar source 4673 (i.e., uncertainties in the trigger efficiencies) have been grouped. The groups of uncertainties 4674 have been ordered according to decreasing improvement in the expected limit obtained when 4675 removing the group from the list of nuisances included in the likelihood fit. The description of 4676 each single uncertainty in the text follows the same order. 4677

Scale factors are used to correct for the differences in the double-b tagger misidentification 4678 efficiencies between data and prediction from simulation for W/Z+jets production and for tt production. These scale factors are measured by simultaneously fitting events that pass or fail 4680 the double-b tag requirement. The correlation between the double-b tagger and the p_T^{miss} (or 4681 U) is taken into account in the scale factor measurement by allowing recoil bins to fluctuate 4682 independently from each other within a constraint that depends on the recoil value. Such 4683 dependence is estimated from the profile of the two-dimensional distribution double-b tag vs. $p_{\rm T}^{\rm miss}/U$. This shape uncertainty in the double-b scale factors measurement is the one that has 4685 the largest impact on the limits on the signal cross section. 4686

A shape uncertainty due to bin-by-bin statistical uncertainties in the transfer factors, which are used to derive the predictions for the main backgrounds from data in CRs, is considered for the Z+jets, W+jets, and tt̄ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the $p_{\rm T}$ of the CA15 jet. This shape uncertainty has been derived through a measurement performed using a sample enriched in multijet events with double-muon-tagged $g \rightarrow b\bar{b}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substruc-

Table 23: Sources of systematic uncertainty, along with the type (rate/shape) of uncertainty and the affected processes. For the rate uncertainties, the percentage value of the prior is quoted. The last column denotes the improvement in the expected limit when removing the uncertainty group from the list of nuisances included in the likelihood fit. Such improvement is estimated considering as signal process the 2HDM+a model with $m_A = 1.1$ TeV and $m_a = 150$ GeV (with $\sin \theta = 0.35$ and $\tan \beta = 1$).

Ctti	Т	Processes	Tours and any association to
Systematic uncertainty	<i>7</i> 1		Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, tt̄	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, tt̄	1.9%
Double-b tagging	shape	SM h, signal	1.2%
$N_2^{\rm DDT}$ efficiency	7%	diboson, SM h, signal	1.2/0
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\rm T}^{\rm miss}$ magnitude	5%	all	0.7%
Integrated luminosity	2.5%	t, diboson, multijet, SM h, signal	< 0.5%
$p_{\rm T}^{\rm miss}$ trigger muon multiplicity	shape	Z+jets, W+jets	
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	< 0.5%
single-electron trigger	1%	all	
AK4 b tagging	shape	all	< 0.5%
au lepton veto	3%	all	< 0.5%
Lepton efficiency	1% per lepton	all	< 0.5%
Heavy-flavor fraction	4-5%	Z+jets, W+jets	< 0.5%
Renorm./fact. scales	shape	SM h	
PDF	shape	SM h	< 0.59/
Multijet normalization	100%	multijet	< 0.5%
Theoretical cross section	20%	t, diboson	

ture variable $N_2^{\rm DDT}$, which is used to identify two-prong CA15 jets, is assigned to all processes where the decay of a resonance inside the CA15 jet cone is expected. Such processes include signal production together with SM h and diboson production. The uncertainty has been derived from the efficiency measurement obtained by performing a fit in a control sample enriched in semi-leptonic $t\bar{t}$ events, where the CA15 jet originates from the W boson that comes from the hadronically decaying top quark.

A 4% rate uncertainty due to the imperfect knowledge of the CA15 jet energy scale [490] is assigned to all the processes obtained from simulation.

Similarly, a 5% rate uncertainty in p_T^{miss} magnitude, as measured by CMS in Ref. [502], is assigned to each processes estimated from simulation.

A rate uncertainty of 2.5% in the in the integrated luminosity measurement [503] is included and assigned to processes determined from simulation. In these cases, QCD renormalisation and factorization scales scale and PDF uncertainties are included as shape uncertainties, obtained by varying those parameters in simulation event-by-event

The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiencies are affected by uncertainties in the muon multiplicity in the event. Differences on the order of 2% are observed between single-muon and dimuon events at lower U values and they are sources of an additional systematic uncertainty in the transfer factors for those processes whose prediction relies on data events in the single-muon and dimuon CRs (tt̄, W+jets, and Z+jets production). As these uncertainties depend on the momentum of the identified muon, they can change the shape of the U distribution and are thus treated as shape uncertainties. The $p_{T,\text{trig}}^{\text{miss}}$ trigger efficiency is parametrized as a function of p_{T}^{miss} . The uncertainty in its measurement is 1% and is included in the fit as a rate uncertainty. The efficiencies of the single-electron triggers are parametrized as a function of the electron p_{T} and η and an

H. Results

associated 1% systematic uncertainty is added into the fit.

An uncertainty on the efficiency of the CSV b-tagging algorithm applied to isolated AK4 jets is assigned to the transfer factors used to predict the $t\bar{t}$ background. The scale factors that correct this efficiency are measured with standard CMS methods [492]. They depend on the p_T and η of the b-tagged (or mistagged) jet and therefore their uncertainties are included in the fit as shape uncertainties.

The uncertainty in the τ lepton veto amounts to 3%, correlated across all U bins. Also correlated across all U bins are the uncertainties in the selection efficiencies per selected electron or muon, that amount to 1%.

An uncertainty of 21% in the heavy-flavor fraction of W+jets is reported in previous CMS measurements [504, 505]. The uncertainty in the heavy-flavor fraction of the Z+jets process is measured to be 22% [506, 507]. To take into account the variation of the double-b tagging efficiency introduced by such uncertainties, the efficiencies for the W+jets and Z+jets processes are reevaluated after varying the heavy-flavor component in the simulation. The difference in the efficiency with respect to the nominal efficiency value is taken as a systematic uncertainty, and amounts to 4% in the rate of the W+jets process and of 5% in the rate of the Z+jets process.

Uncertainties in the SM h production due to variations of the of the renormalization/factorization 4733 scales and PDFs are included as shape variations. Being a negligible background source, an un-4734 certainty of 100% is assigned to the QCD multijet yield. This uncertainty is estimated using a 4735 sample enriched in multijet events. The sample is obtained by vetoing leptons and photons, by 4736 requiring $p_{\rm T}^{\rm miss} > 250$ GeV, and by requiring that the minimum azimuthal angle between $\vec{p}_{\rm T}^{\rm miss}$ 4737 and the jet directions be less than 0.1 rad. One nuisance parameter represents the uncertainty 4738 in QCD multijet yields in the signal region, while separate nuisance parameters are introduced 4739 for the muon CRs and electron CRs. A systematic uncertainty of 20% is assigned to the single top quark background yields as reported by CMS in Ref. [508] and is correlated between the SR 4741 and the CRs. An uncertainty of 20% is also assigned to the diboson production cross section as 4742 measured by CMS in Refs. [509, 510] and correlated across the SR and CRs. 4743

H Results

4744

4745

4747

4748

4749

4750

4751

4752

The expected yields for each background in the SR and their uncertainties, as determined in the likelihood fit under the background-only assumption, are presented in Table 24, along with the observed data yields. Good agreement is observed between data and the predictions. Due to anticorrelations between background processes, in some bins the uncertainty in the background sum is smaller than the one in the individual contributions, such as, for example, the Z+jets yields. Expected yields are also presented for two signal models. The selection efficiencies for the chosen points correspond to 5% for the 2HDM+*a* model and 1% for the baryonic Z' model.

Figure 51 shows the pre-fit and post-fit $p_{\rm T}^{\rm miss}$ distribution in the SR for signal and for all SM backgrounds, as well as the observed data distribution. The likelihood fit has been performed simultaneously in all analysis regions. The data agree with the background predictions at the one standard deviation level, and the post-fit estimate of the SM background is slightly larger than the pre-fit one. The distributions for U in the muon and electron CRs, after a fit to the data, are presented in Figs. 52 and 53.

No significant excess over the SM background expectation is observed in the SR. The results of this search are interpreted in terms of upper limits on the signal strength modifier $\mu =$

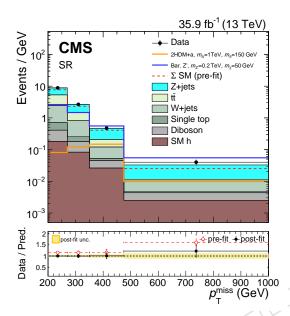


Figure 51: The $p_{\rm T}^{\rm miss}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

Table 24: Post-fit event yield expectations per $p_{\rm T}^{\rm miss}$ bin for the SM backgrounds in the signal region when including the signal region data in the likelihood fit, under the background-only assumption. Quoted are also the expected yields for two signal models. For the 2HDM+a model, we choose $\sin\theta=0.35$ and $\tan\beta=1$. Uncertainties quoted in the predictions include both the systematic and statistical components.

$p_{\mathrm{T}}^{\mathrm{miss}}$ -bin	200-270 GeV	270-350 GeV	350-475 GeV	> 475 GeV
Z+jets	248.9 ± 22.2	97.2 ± 8.5	32.6 ± 3.6	11.1 ± 1.9
tŧ	199.2 ± 13.5	52.1 ± 5.2	11.1 ± 2.0	0.7 ± 0.4
W+jets	121.6 ± 21.6	45.0 ± 8.7	8.4 ± 1.9	2.9 ± 0.9
Single top quark	21.0 ± 4.2	6.1 ± 1.2	0.9 ± 0.2	0.2 ± 0.1
Diboson	16.0 ± 3.1	7.6 ± 1.5	2.4 ± 0.5	1.0 ± 0.2
SM h	12.6 ± 1.4	6.6 ± 0.7	3.3 ± 0.3	1.3 ± 0.1
Σ (SM)	619.3 ± 20.1	214.6 ± 8.1	58.7 ± 3.7	17.2 ± 2.0
Data	619	214	59	21
2 HDM+ a , $m_A = 1$ TeV, $m_a = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2 \text{TeV}$, $m_{\chi} = 50 \text{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

H. Results 149

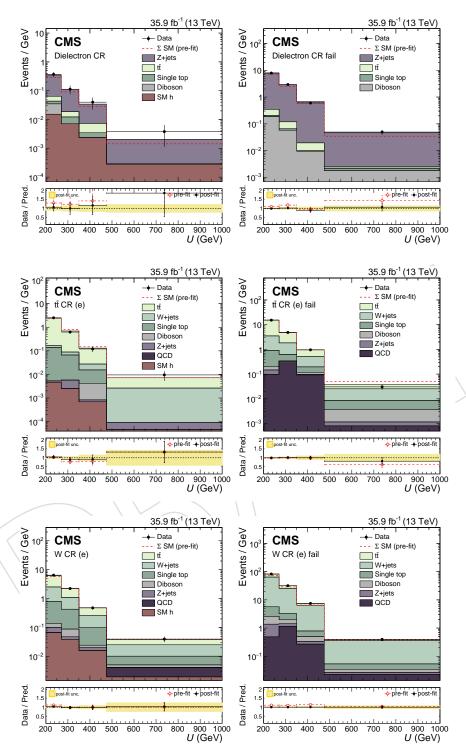


Figure 52: The *U* distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

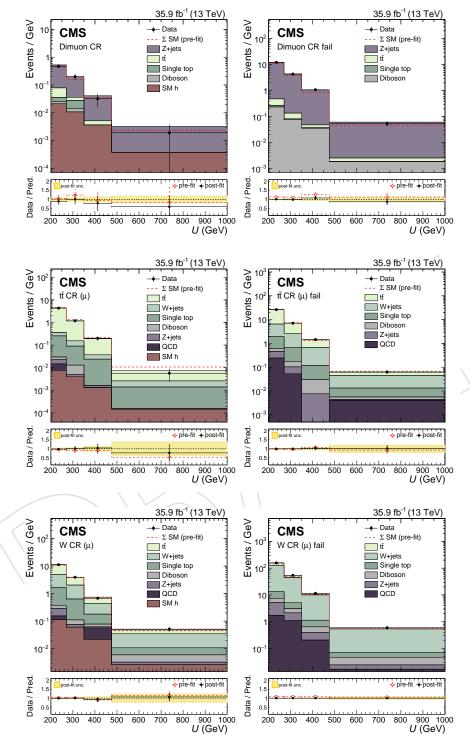


Figure 53: The *U* distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b tag requirement.

H. Results

 σ/σ_{theory} , where σ_{theory} is the predicted production cross section of DM candidates in association with a Higgs boson and σ is the upper limit on the observed cross section. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method (CL_s) [511–513] computed with an asymptotic approximation [514].

Figure 54 shows the upper limits on μ for the three scans (m_A , $\sin \theta$, and $\tan \beta$) performed. For the 2HDM+a model, m_A masses are excluded between 500 and 900 GeV for $m_a = 150$ GeV, $\sin \theta = 0.35$ and $\tan \beta = 1$. Mixing angles with $0.35 < \sin \theta < 0.75$ are excluded for $m_A = 600$ GeV and $m_a = 200$ GeV, assuming $\tan \beta = 1$. Also excluded are $\tan \beta$ values between 0.5 and 2.0 (1.6) for $m_a = 100$ (150) GeV and $m_A = 600$ GeV, given $\sin \theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

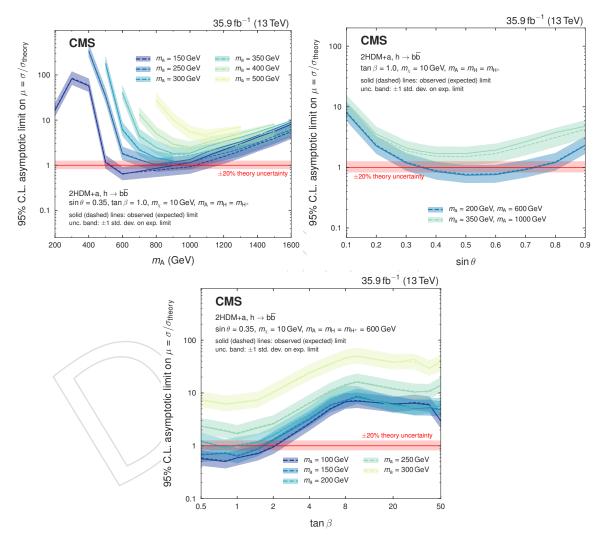


Figure 54: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or tan β (lower).

Figure 55 shows the expected and observed exclusion range as a function of $m_{Z'}$ and m_{χ} for the baryonic Z' model. For a DM mass of 1 GeV, masses $m_{Z'} < 1.6$ TeV are excluded. The expected exclusion boundary is 1.85 TeV. Masses for the DM particles of up to 430 GeV are excluded for a 1.1 TeV Z' mass. These are the most stringent limits on this model so far.

To compare results with DM direct detection experiments, limits from the baryonic Z' model are presented in terms of a spin-independent (SI) cross section for DM scattering off a nucleus.

4779

4781

4782

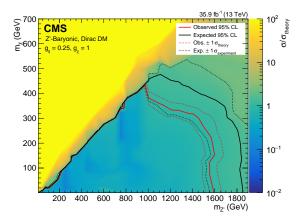


Figure 55: Upper limits on the signal strength modifier for the baryonic Z' model as a function of $m_{Z'}$ and m_{χ} . Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z' mass of 1.25 TeV.

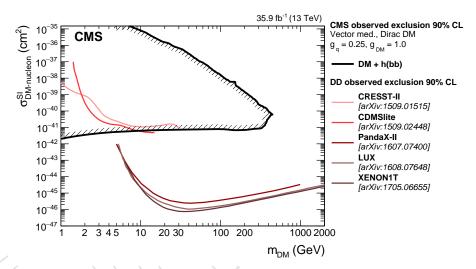


Figure 56: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_{χ} . Results on the baryonic Z' model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

Following the recommendation of Ref. [515], the value of σ_{SI} is determined by the equation:

$$\sigma_{\rm SI} = \frac{f^2(g_{\rm q})g_{\rm DM}^2\mu_{\rm n\chi}^2}{\pi m_{\rm med}^4},\tag{8}$$

where $\mu_{n\chi}$ is the reduced mass of the DM-nucleon system, $f(g_q)$ is the mediator-nucleon coupling, which is dependent on the mediator coupling to SM quarks g_q , g_{DM} is the mediator coupling to SM particles, and m_{med} is the mass of the mediator. The resulting limits as a function of DM the mass are shown in Fig. 56. Under the assumptions made for the baryonic Z' model, these limits are the most stringent to date for $m_{\chi} < 5 \, \text{GeV}$.

I. Summary 153

I Summary

4783

4784

4785

4786

4787

4788

4790

4791

4792

4794

4795

4796

4797

4799

4800

A search for the associated production of dark matter (DM) particles with a Higgs boson decaying into a pair of bottom quarks is presented. No significant deviation from the predictions of the standard model (SM) is observed, and upper limits on the production cross section predicted by a type-2 two higgs doublet model extended by an additional light pseudoscalar boson a (2HDM+a) and the baryonic Z' model are established. They constitute the most stringent collider exclusions placed on the parameters in these models so far. For the nominal choice of the mixing angle $\sin \theta$ and $\tan \beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900 \,\text{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150 \,\text{GeV}$. Scanning over $\sin \theta$ with $\tan \beta = 1$, we exclude $0.35 < \sin \theta < 0.75$ for $m_A = 600$ GeV and $m_{\rm a}=200\,{\rm GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_{\rm A}=600\,{\rm GeV}$ and $m_a = 100 (150) \,\text{GeV}$ and $\sin \theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_{\chi} = 10 \,\text{GeV}$ is assumed. For the baryonic Z' model, we exclude Z' boson masses up to 1.6 TeV for a DM mass of 1 GeV, and DM masses up to 430 GeV for a Z' boson mass of 1.1 TeV. The reinterpretation of the results for the baryonic Z' model in terms of an SI nucleon scattering cross section yields a higher sensitivity for m_{χ} < 5 GeV than existing results from direct detection experiments, under the assumptions imposed by the model. The 2HDM+a model is probed experimentally for the first time.

4801 Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent perfor-4802 mance of the LHC and thank the technical and administrative staffs at CERN and at other 4803 CMS institutes for their contributions to the success of the CMS effort. In addition, we grate-4804 fully acknowledge the computing centers and personnel of the Worldwide LHC Computing 4805 Grid for delivering so effectively the computing infrastructure essential to our analyses. Fi-4806 nally, we acknowledge the enduring support for the construction and operation of the LHC 4807 and the CMS detector provided by the following funding agencies: BMWFW and FWF (Aus-4808 tria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); 4809 CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); 4810 RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Fin-4811 land, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Ger-4812 many); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM 4814 (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New 4815 Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, 4816 RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER 4817 (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United 4819 Kingdom); DOE and NSF (USA). 4820

Individuals have received support from the Marie-Curie program and the European Research 4821 Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foun-4822 dation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Fed-4823 4824 eral Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie 4825 (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the 4826 Council of Science and Industrial Research, India; the HOMING PLUS program of the Foun-4827 dation for Polish Science, cofinanced from European Union, Regional Development Fund, the 4828

Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Severo Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

References

4837

- 4838 [456] G. Bertone, D. Hooper, and J. Silk, "Particle Dark Matter: Evidence, Candidates and Constraints", *Phys. Rept.* **405** (2005) 279, doi:10.1016/j.physrep.2004.08.031, arXiv:0404175.
- [457] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection",

 Ann. Rev. Astron. Astrophys. 48 (2010) 495,

 doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- [458] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", Ann. Rev. Astron. Astrophys. 49 (2011) 155,

 doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.
- 4847 [459] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", Astron.
 4848 Astrophys. 594 (2016) A13, doi:10.1051/0004-6361/201525830,
 4849 arXiv:1502.01589.
- 4850 [460] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC", Phys. Lett. B **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [461] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS
 experiment at the LHC", Phys. Lett. B 716 (2012) 30,
 doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- 4856 [462] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}=7$ and 8 TeV", JHEP **06** (2013) 81, doi:10.1007/JHEP06 (2013) 081, arXiv:1303.4571.
- 4859 [463] A. A. Petrov and W. Shepherd, "Searching for dark matter at LHC with Mono-Higgs 4860 production", *Phys. Lett. B* **730** (2014) 178, doi:10.1016/j.physletb.2014.01.051, 4861 arXiv:1311.1511.
- [464] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs detection of dark matter at the LHC", JHEP 06 (2014) 078, doi:10.1007/JHEP06 (2014) 078, arXiv:1402.7074.
- 4864 [465] L. Carpenter et al., "Mono-Higgs-boson: A new collider probe of dark matter", *Phys.*4865 *Rev. D* **89** (2014) 075017, doi:10.1103/PhysRevD.89.075017, arXiv:1312.2592.
- 4866 [466] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

References 155

4868 [467] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators", *JHEP* **05** (2017) 138, doi:10.1007/JHEP05(2017)138, arXiv:1701.07427.

- [468] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}=7$ and 8 TeV", JHEP **08** (2016) 045, doi:10.1007/JHEP08 (2016) 045, arXiv:1606.02266.
- [469] ATLAS Collaboration, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb⁻¹ of pp collisions at $\sqrt{s}=13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **119** (2017) 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- [470] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s}=13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10 (2017) 180, arXiv:1703.05236.
- 4882 [471] 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.
- 4885 [472] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", JHEP 11 (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [473] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* 11 (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [474] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [475] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$ ", Phys. Rev. Lett. 110 (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [476] 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:0611129.
- 4900 [477] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12 (2012) 061, arXiv:1209.6215.
- [478] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic photon production at large transverse momenta", *JHEP* **03** (2006) 059, doi:10.1088/1126-6708/2006/03/059, arXiv:hep-ph/0508253.
- [479] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet production", in 50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, Italy, March 21-28, 2015. 2015. arXiv:1505.05704.

- 4908 [480] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", JHEP **04** (2016) 021, doi:10.1007/JHEP04 (2016) 021, arXiv:1511.08692.
- [481] 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.
- 4915 [483] NNPDF Collaboration, "Parton distributions for the LHC Run II", JHEP **04** (2015) 040, doi:10.1007/JHEP04 (2015) 040, arXiv:1410.8849.
- 4917 [484] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C 76 (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [485] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", Eur.
 Phys. J. C 74 (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y,
 arXiv:1404.5630.
- 4923 [486] G. Collaboration, "Geant4 a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250, doi:10.1016/10.1016/S0168-9002 (03) 01368-8.
- 4925 [487] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST 12 (2017) 10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [488] CMS Collaboration, "A Cambridge-Aachen (C-A) based jet algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.
- [489] D. Berteloni, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* 59 (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.
- 4932 [490] CMS Collaboration, "Determination of jet energy calibration and transverse momentum 4933 resolution in CMS", JINST 6 (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, 4934 arXiv:1107.4277.
- 4935 [491] 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.
- [492] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", JINST 13 (2018) 05011.
- [493] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions", JHEP
 12 (2016) 153, doi:10.1007/JHEP12 (2016) 153, arXiv:1609.07483.
- [494] J. Dolen et al., "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure", JHEP 05 (2016) 156, doi:10.1007/JHEP05(2016) 156, arXiv:1603.00027.
- [495] 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.

References 157

```
[496] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at \sqrt{s} = 8 TeV", JINST 10 (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
```

- [497] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [498] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_{τ} at CMS", JINST 11 (2016) 01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.
- [499] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", JINST 8 (2013) 11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.
- [500] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp
 collision data using the CMS detector", CMS Physics Analysis Summary
 CMS-PAS-JME-16-004, 2016.
- [501] L. Moneta et al., "The RooStats Project", in 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010.

 arXiv:1009.1003.
- 4964 [502] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8 \text{ TeV}$ ", JINST 10 (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [503] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period",
 CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- 4969 [504] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* 4971 **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.
- [505] CMS Collaboration, "Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s}=7$ TeV", *Phys. Lett. B* **735** (2014) 204, doi:10.1016/j.physletb.2014.06.041, arXiv:1312.6608.
- [506] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s}=7$ TeV", *Phys. Rev. D* **91** (2015) 052008, doi:10.1103/PhysRevD.91.052008, arXiv:1408.3104.
- 4978 [507] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s}=7\,\text{TeV}$ ", JHEP **06** (2014) 120, doi:10.1007/JHEP06 (2014) 120, arXiv:1402.1521.
- [508] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}=8\,\text{TeV}$ ", Phys.~Rev.~Lett.~ 112 (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.
- [509] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}=13\,\text{TeV}$ ", Phys. Lett. B 763 (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

5017

- [510] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions 4987 at $\sqrt{s} = 13 \text{ TeV}''$, Phys. Lett. B **766** (2017) 268, 4988 doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943. [511] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs 4990 Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working 4991 Group", Technical Report CERN-2013-004, Geneva, 2013. 4992 doi:10.5170/CERN-2013-004, arXiv:1307.1347. 4993 [512] A. L. Read, "Presentation of search results: the CL_s technique", J. Phys. G 28 (2002) 4994 2693, doi:10.1088/0954-3899/28/10/313. 4995 [513] T. Junk, "Confidence level computation for combining searches with small statistics", 4996 Nucl. Instrum. Meth. A 434 (1999) 435, doi:10.1016/S0168-9002 (99) 00498-2, 4997 arXiv:hep-ex/9902006. 4998 [514] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for 4999 likelihood-based tests of new physics", Eur. Phys. J. C 71 (2011) 1554, 5000 doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727v3. 5001 [515] A. Boveia et al., "Recommendations on presenting LHC searches for missing transverse 5002 energy signals using simplified s-channel models of dark matter", (2016). 5003 arXiv:1603.04156. 5004 [516] SuperCDMS Collaboration, "New results from the search for low-mass weakly 5005 interacting massive particles with the CDMS low ionization threshold experiment", 5006 Phys. Rev. Lett. 116 (2016) 071301, doi:10.1103/PhysRevLett.116.071301, 5007 arXiv:1509.02448. 5008 [517] LUX Collaboration, "Results from a search for dark matter in the complete LUX 5009 exposure", Phys. Rev. Lett. 118 (2017) 021303, 5010 doi:10.1103/PhysRevLett.118.021303, arXiv:1608.07648. 5011 [518] XENON Collaboration, "First dark matter search results from the XENON1T 5012 experiment", Phys. Rev. Lett. 119 (2017) 181301, 5013 doi:10.1103/PhysRevLett.119.181301, arXiv:1705.06655. 5014 [519] PandaX-II Collaboration, "Dark matter results from 54-ton-day exposure of PandaX-II 5015
- 5018 [520] CRESST-II Collaboration, "Results on light dark matter particles with a low-threshold CRESST-II detector", Eur. Phys. J. C 76 (2016) 25, doi:10.1140/epjc/s10052-016-3877-3, arXiv:1509.01515.

doi:10.1103/PhysRevLett.119.181302, arXiv:1708.06917.

experiment", Phys. Rev. Lett. 119 (2017) 181302,