CMS PAS EXO-16-050

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 $\mathrm{b\bar{b}}$ 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 Z′ simplified models. For the baryonic Z′ model, the presented results constitute the most stringent constraints to date. The 2HDM+a model is tested experimentally for the first time.

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 that decays into a bottom quark-antiquark pair ($b\bar{b}$). As the $\mathrm{h} \to \mathrm{b\bar{b}}$ 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, 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 $A$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 $\mathrm{H}^{\pm}$ 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 $\lambda_3 = \lambda_{\mathrm{P1}} = \lambda_{\mathrm{P2}} = 3$ [467]. Masses of the

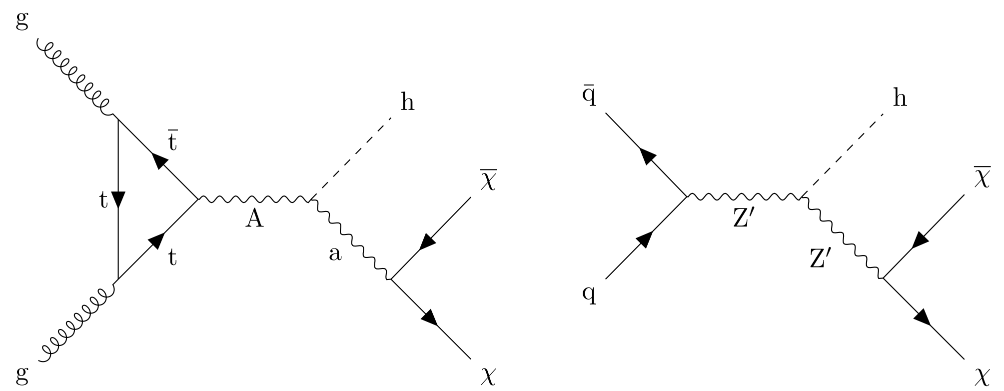

Figure 1: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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., $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,\mathrm{GeV}$.

The baryonic Z′ model [465] is an extension of the SM with an additional $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B, with a coupling of $\mathrm{h}_B\mathrm{Z'Z'}$ and vacuum expectation value of v_B. The Z′ and SM Higgs boson, h, interact with a coupling strength of $g_{\mathrm{hZ'Z'}} = m^2_{\mathrm{Z'}} \sin\theta / v_B$, 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 $g_{\mathrm{q}}, g_{\chi} \sim 4\pi$ 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 $g_{\mathrm{hZ'Z'}}/m_{\mathrm{Z'}} = 1$, which implies $v_B = m_{\mathrm{Z'}} \sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{\mathrm{Z'}}$ 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.

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 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 μs, using information from the calorimeters and muon detectors. Subsequently, the high-level trigger processor farm reduces the event rate to 1 kHz.

3 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 $t\bar{t}$ 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 $t\bar{t}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_{T}. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\mathrm{SD}}^2/p_{\mathrm{T}}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_{T}, decorrelating the N_2 variable as a function of ρ and p_{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^{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 $p_{\mathrm{T}} > 18$ 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 $p_{\mathrm{T}} > 15$ 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.

5 Event selection

Signal events are characterized by a high value of $p_{\mathrm{T}}^{\mathrm{miss}}$, 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 $\mathrm{t\bar{t}}$ production. To predict the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $\mathrm{t\bar{t}}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_{T} of the lepton(s) from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation in CRs, is used as a proxy for the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$, where $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ ($H_{\mathrm{T}}^{\mathrm{miss}}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\mathrm{T}}$ of all the particles (jets with $p_{\mathrm{T}} > 20$ GeV) at the trigger level. Muon candidates are excluded from the online $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ calculation. Thresholds on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are reffered to as $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\mathrm{SD}} < 150$ GeV and $N_2^{\mathrm{DDT}} < 0$. In the SR (CRs), $p_{\mathrm{T}}^{\mathrm{miss}}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($\vec{U}$) 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, $t\bar{t}$, W+jets	0	0	pass
Single-lepton, b-tagged	$t\bar{t}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $t\bar{t}$	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 $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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_{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 $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 40$ (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.

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

The transfer factors used to predict the W+jets and $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the $\mathrm{t\bar{t}}$ process, transfer factors to predict this contamination from $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{Z} \to \nu\nu$ and the $\mathrm{Z} \to \ell\ell$ decays and the impacts of lepton acceptances and selection efficiencies.

7 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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.* $p_{\mathrm{T}}^{\mathrm{miss}}/U$. 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 $\mathrm{t\bar{t}}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_{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 $\mathrm{g} \to \mathrm{b\bar{b}}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substruc-

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 $m_{\mathrm{A}} = 1.1\,\mathrm{TeV}$ and $m_{\mathrm{a}} = 150\,\mathrm{GeV}$ (with $\sin\theta = 0.35$ and $\tan\beta = 1$).

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

ture variable N_2^{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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 ($\mathrm{t\bar{t}}$, 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

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 $p_{T}^{miss} > 250\,\text{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{T}^{miss}$ 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_{T}^{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 $\mu =$

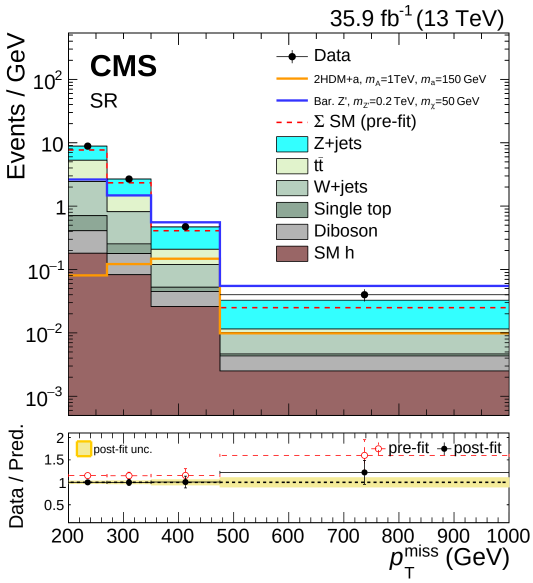

Figure 2: The $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate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

Table 3: Post-fit event yield expectations per $p_{\mathrm{T}}^{\mathrm{miss}}$ 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
$\mathrm{t\bar{t}}$	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, $m_{\mathrm{A}} = 1\,\mathrm{TeV}$, $m_{\mathrm{a}} = 150\,\mathrm{GeV}$	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z′, $m_{\mathrm{Z}'} = 0.2\,\mathrm{TeV}$, $m_{\chi} = 50\,\mathrm{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

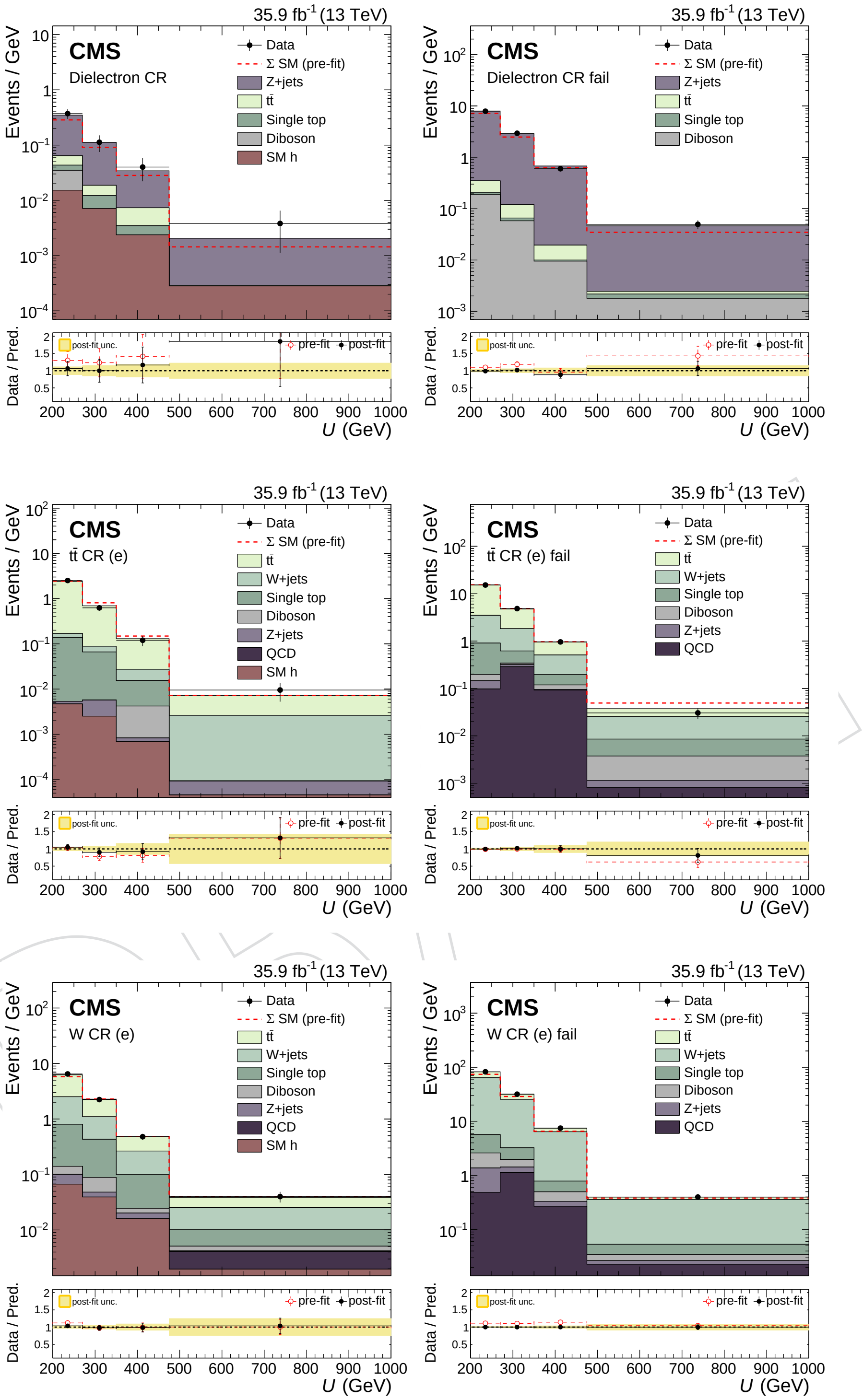

Figure 3: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

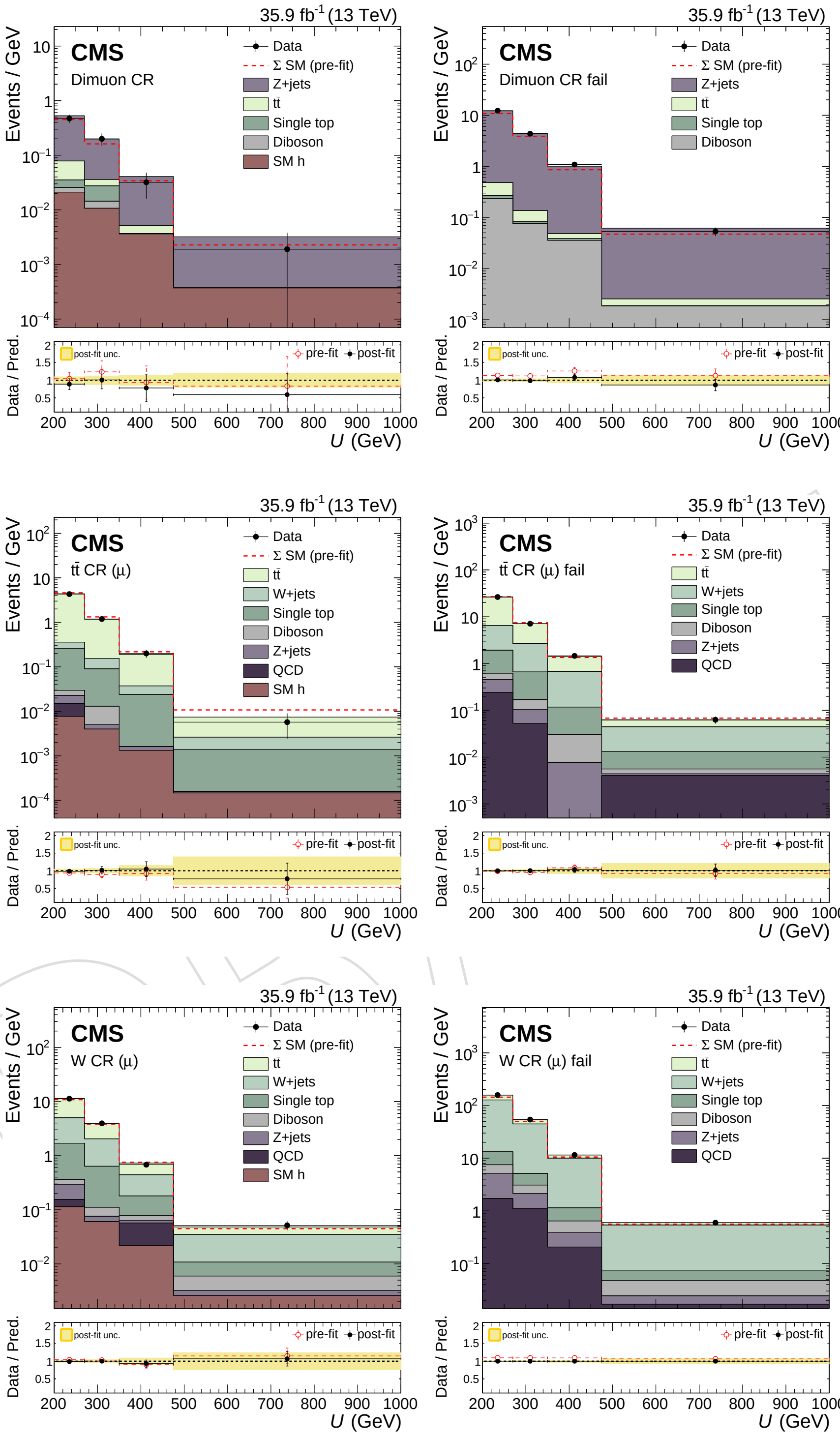

Figure 4: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

$\sigma/\sigma_{\text{theory}}$, 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 $m_{\text{a}} = 150\,\text{GeV}$, $\sin\theta = 0.35$ and $\tan\beta = 1$. Mixing angles with $0.35 < \sin\theta < 0.75$ are excluded for $m_{\text{A}} = 600\,\text{GeV}$ and $m_{\text{a}} = 200\,\text{GeV}$, assuming $\tan\beta = 1$. Also excluded are $\tan\beta$ values between 0.5 and 2.0 (1.6) for $m_{\text{a}} = 100$ (150) GeV and $m_{\text{A}} = 600\,\text{GeV}$, given $\sin\theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

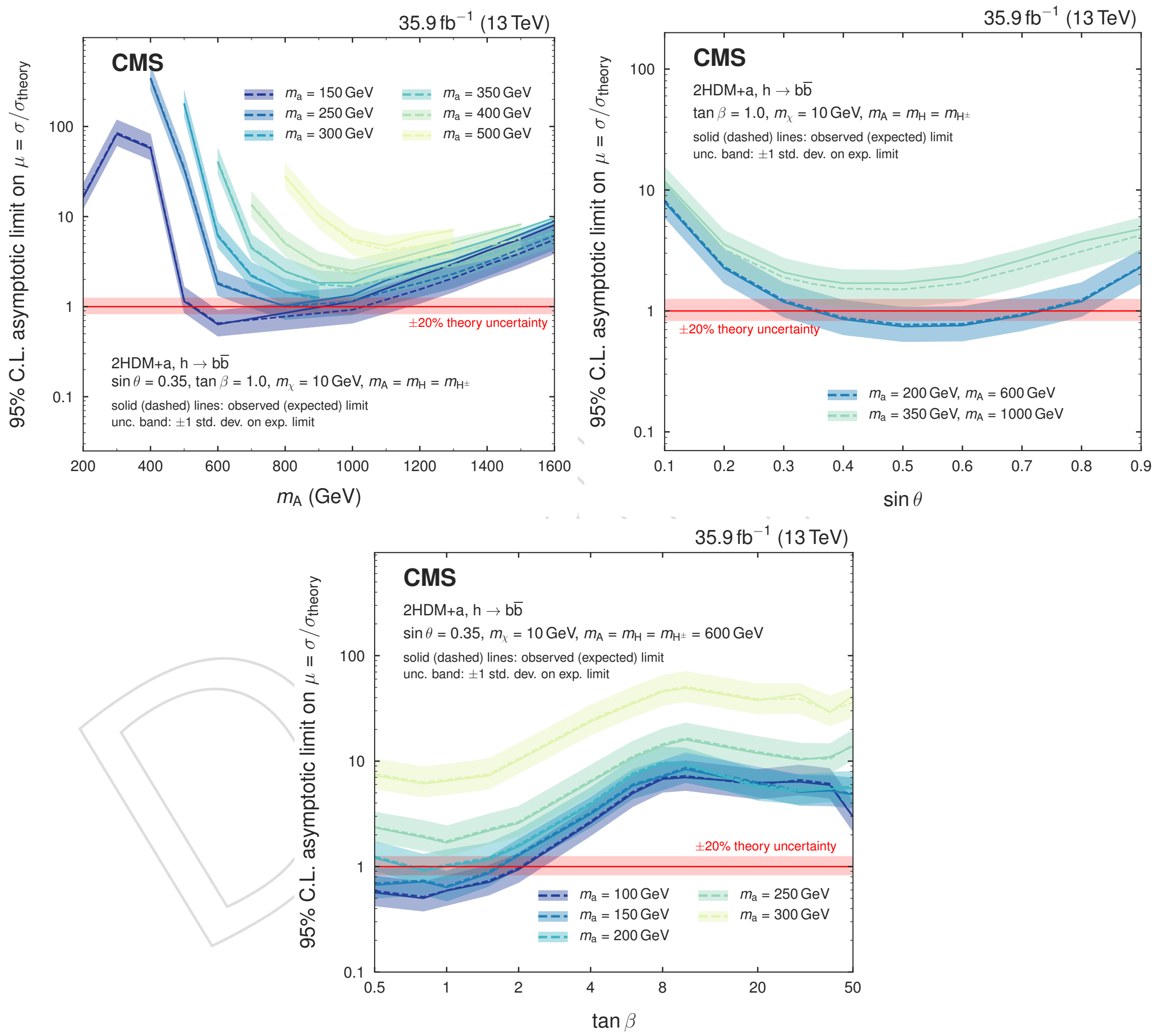

Figure 5: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_{A} and m_{a} (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

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 $m_{Z'} < 1.6\,\text{TeV}$ 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.

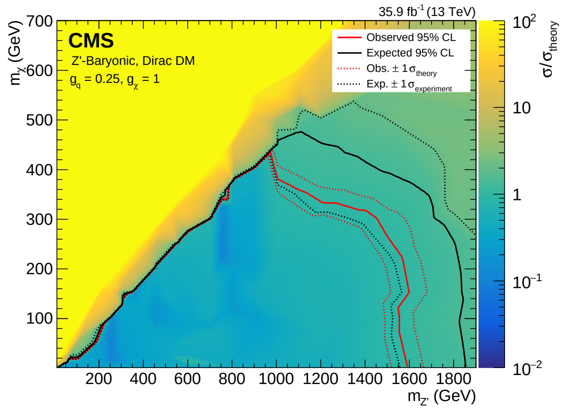

Figure 6: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

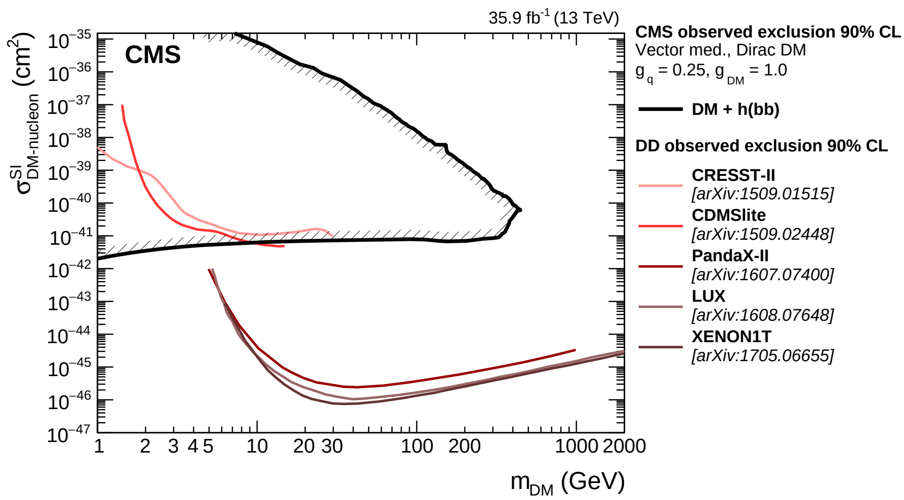

Figure 7: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from 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:

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

where $\mu_{\mathrm{n}\chi}$ is the reduced mass of the DM-nucleon system, $f(g_{\mathrm{q}})$ 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 $m_\chi < 5\,\mathrm{GeV}$.

9 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 choice of the mixing angle $\sin\theta$ and $\tan\beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900\,\mathrm{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150\,\mathrm{GeV}$. Scanning over $\sin\theta$ with $\tan\beta = 1$, we exclude $0.35 < \sin\theta < 0.75$ for $m_\mathrm{A} = 600\,\mathrm{GeV}$ and $m_\mathrm{a} = 200\,\mathrm{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_\mathrm{A} = 600\,\mathrm{GeV}$ and $m_\mathrm{a} = 100$ (150) GeV and $\sin\theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_\chi = 10\,\mathrm{GeV}$ 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 $m_\chi < 5\,\mathrm{GeV}$ 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

[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.

[5] 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.

[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.

[11] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

[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.

[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.

[15] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $\mathrm{b\bar{b}}$ 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.

[17] 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.

[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.

[23] 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.

[24] 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.

[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.

[26] 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.

[27] 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.

[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.

[31] G. Collaboration, "Geant4 – a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/10.1016/S0168-9002(03)01368-8.

[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.

[34] 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.

[35] 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.

[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.

[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.

[39] 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.

[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 $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) 10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[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.

[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 *13^{th} 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 $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) 02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

[48] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

[49] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.

[50] 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.

[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.

[52] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **06** (2014) 120, doi:10.1007/JHEP06(2014)120, arXiv:1402.1521.

[53] 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.

[54] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\to \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **763** (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 $\sqrt{s} = 13\,\mathrm{TeV}$", *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.

[60] 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.

[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.

[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.

[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

[cern]The CMS Collaboration

2018/08/17

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 $\mathrm{h} \to \mathrm{b\bar{b}}$ 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, 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 $A$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 $\mathrm{H}^{\pm}$ 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 $\lambda_3 = \lambda_{\mathrm{P1}} = \lambda_{\mathrm{P2}} = 3$ [467]. Masses of the

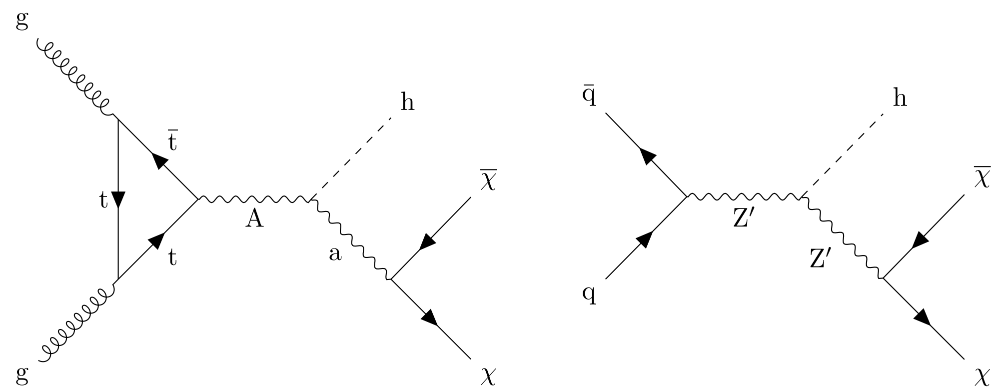

Figure 8: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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., $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,\mathrm{GeV}$.

The baryonic Z′ model [465] is an extension of the SM with an additional $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B, with a coupling of $\mathrm{h}_B\mathrm{Z'Z'}$ and vacuum expectation value of v_B. The Z′ and SM Higgs boson, h, interact with a coupling strength of $g_{\mathrm{hZ'Z'}} = m^2_{\mathrm{Z'}} \sin\theta / v_B$, 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 $g_{\mathrm{q}}, g_{\chi} \sim 4\pi$ 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 $g_{\mathrm{hZ'Z'}} / m_{\mathrm{Z'}} = 1$, which implies $v_B = m_{\mathrm{Z'}} \sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{\mathrm{Z'}}$ 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 μs, 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 $t\bar{t}$ 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 $t\bar{t}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\overline{b}}$ 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 $\mathrm{b\overline{b}}$ 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 $\mathrm{b\overline{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_{T}. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\mathrm{SD}}^2/p_{\mathrm{T}}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_{T}, decorrelating the N_2 variable as a function of ρ and p_{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^{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 $p_{\mathrm{T}} > 18$ 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 $p_{\mathrm{T}} > 15$ 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

Signal events are characterized by a high value of $p_{\mathrm{T}}^{\mathrm{miss}}$, 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 $\mathrm{t\bar{t}}$ production. To predict the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $\mathrm{t\bar{t}}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_{T} of the lepton(s) from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation in CRs, is used as a proxy for the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$, where $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ ($H_{\mathrm{T}}^{\mathrm{miss}}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\mathrm{T}}$ of all the particles (jets with $p_{\mathrm{T}} > 20$ GeV) at the trigger level. Muon candidates are excluded from the online $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ calculation. Thresholds on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are reffered to as $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\mathrm{SD}} < 150$ GeV and $N_2^{\mathrm{DDT}} < 0$. In the SR (CRs), $p_{\mathrm{T}}^{\mathrm{miss}}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($\vec{U}$) 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, $\mathrm{t\bar{t}}$, W+jets	0	0	pass
Single-lepton, b-tagged	$\mathrm{t\bar{t}}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $\mathrm{t\bar{t}}$	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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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_{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 $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 40$ (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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 bℓ denotes the $\mathrm{t\bar{t}}$ process in the b-tagged single-lepton control sample that is used to estimate the $\mathrm{t\bar{t}}$ contribution in the SR, the expected number of $\mathrm{t\bar{t}}$ events, N_i, in the i^{th} bin of the SR is then given by $N_i = \mu_i^{\mathrm{t\bar{t}}}/T_i^{\mathrm{b}\ell}$, where $\mu_i^{\mathrm{t\bar{t}}}$ is a freely floating parameter included in the likelihood to scale the $\mathrm{t\bar{t}}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the $\mathrm{t\bar{t}}$ process, transfer factors to predict this contamination from $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{Z} \to \nu\nu$ and the $\mathrm{Z} \to \ell\ell$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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.* $p_{\mathrm{T}}^{\mathrm{miss}}/U$. 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 $\mathrm{t\bar{t}}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_{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 $\mathrm{g} \to \mathrm{b\bar{b}}$ 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 $m_{\mathrm{A}} = 1.1\,\mathrm{TeV}$ and $m_{\mathrm{a}} = 150\,\mathrm{GeV}$ (with $\sin\theta = 0.35$ and $\tan\beta = 1$).

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

ture variable N_2^{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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 ($\mathrm{t\bar{t}}$, 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

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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $\mu =$

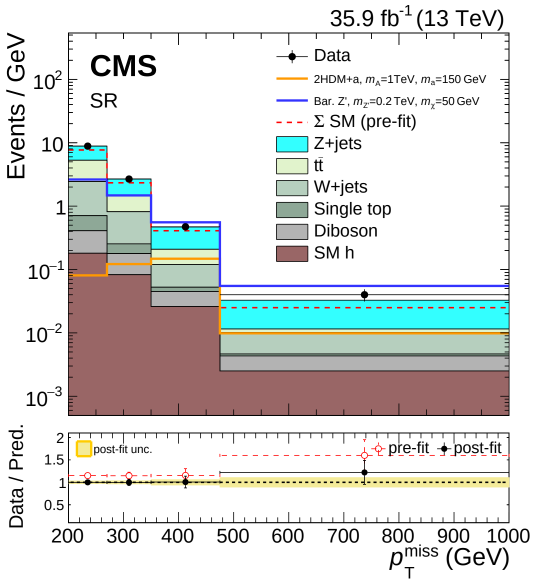

Figure 9: The $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate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

Table 6: Post-fit event yield expectations per $p_{\mathrm{T}}^{\mathrm{miss}}$ 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
$\mathrm{t\bar{t}}$	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, $m_{\mathrm{A}} = 1\,\mathrm{TeV}$, $m_{\mathrm{a}} = 150\,\mathrm{GeV}$	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2\,\mathrm{TeV}$, $m_\chi = 50\,\mathrm{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

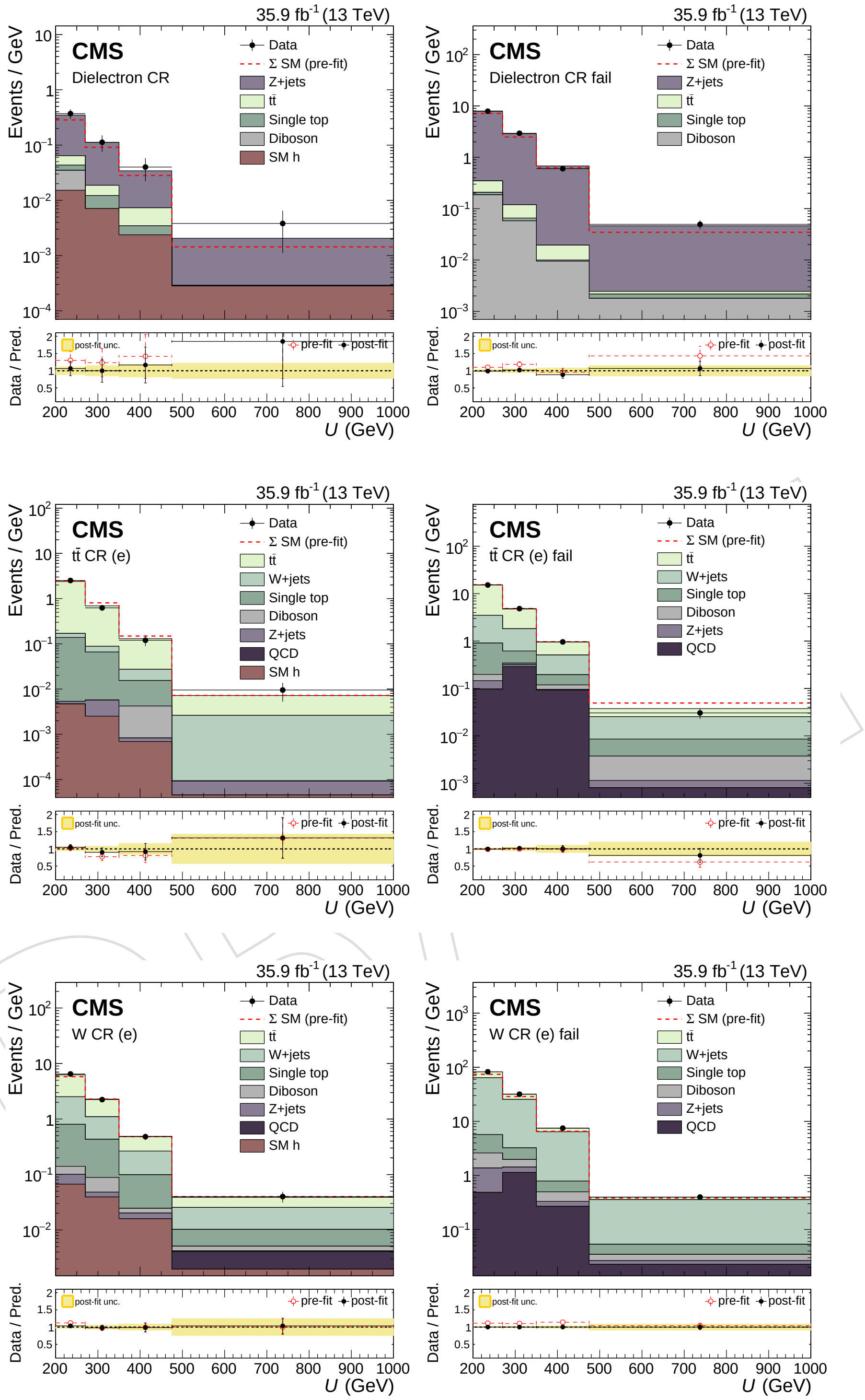

Figure 10: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

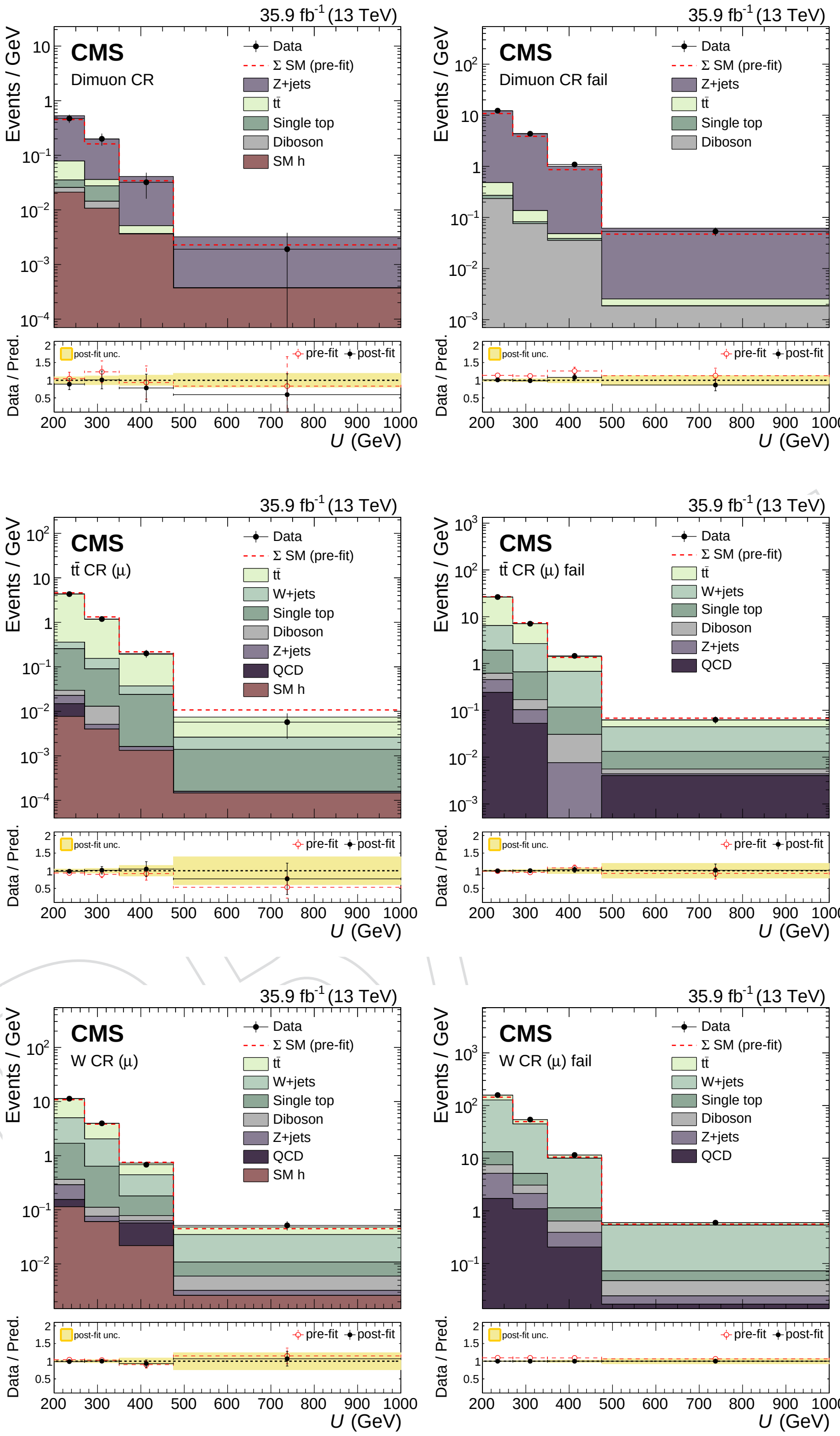

Figure 11: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

$\sigma/\sigma_{\text{theory}}$, 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 $m_{\text{a}} = 150\,\text{GeV}$, $\sin\theta = 0.35$ and $\tan\beta = 1$. Mixing angles with $0.35 < \sin\theta < 0.75$ are excluded for $m_{\text{A}} = 600\,\text{GeV}$ and $m_{\text{a}} = 200\,\text{GeV}$, assuming $\tan\beta = 1$. Also excluded are $\tan\beta$ values between 0.5 and 2.0 (1.6) for $m_{\text{a}} = 100$ (150) GeV and $m_{\text{A}} = 600\,\text{GeV}$, given $\sin\theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

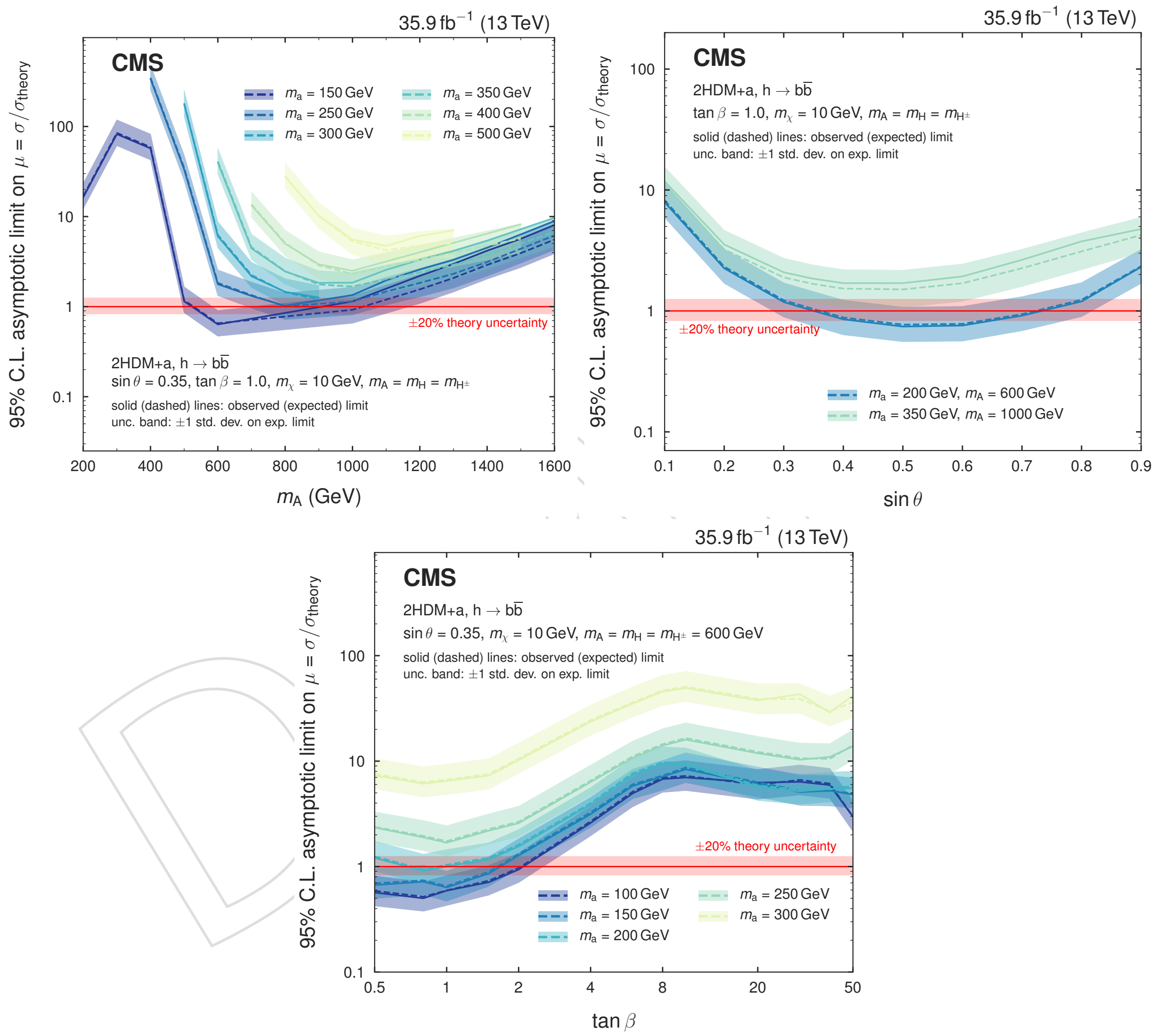

Figure 12: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_{A} and m_{a} (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

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 $m_{Z'} < 1.6\,\text{TeV}$ 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.

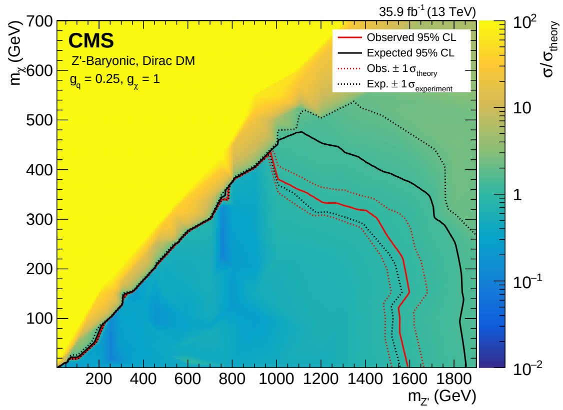

Figure 13: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

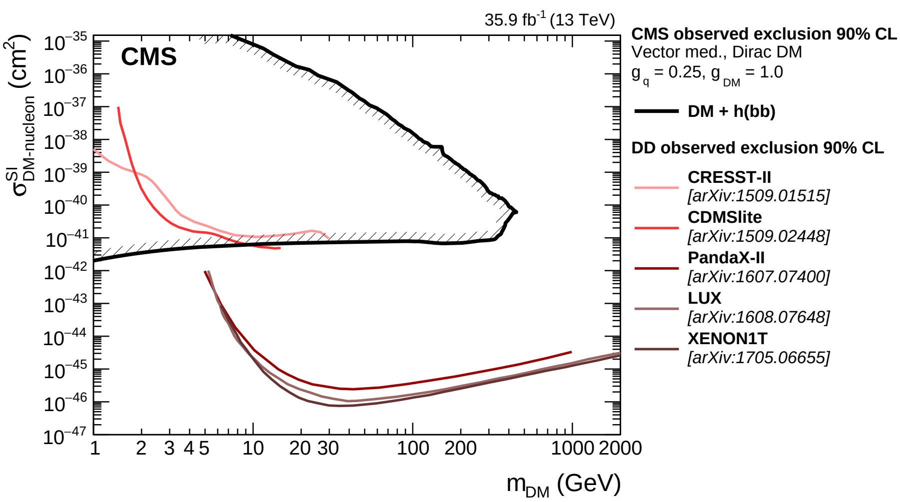

Figure 14: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits 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:

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

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 $m_\chi < 5$ 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 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 $m_\text{A} = 600\,\text{GeV}$ and $m_\text{a} = 200\,\text{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_\text{A} = 600\,\text{GeV}$ and $m_\text{a} = 100$ (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 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

[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`.

[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`.

[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`.

[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`.

[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

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 $\mathrm{b\bar{b}}$ 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.

[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.

[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.

[98] 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.

[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* **12** (2016) 153, doi:10.1007/JHEP12(2016)153, arXiv:1609.07483.

[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 $\sqrt{s} = 7$ 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.

[111] 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.

[112] 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.

[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 $\sqrt{s} = 7$ 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 $\sqrt{s} = 7$ 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 $\to \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13$ 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 $\sqrt{s} = 13$ 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.

[125] 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`, `arXiv:1509.02448`.

[127] 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`.

[128] 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`.

[129] 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`.

[130] 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

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 $\mathrm{h} \to \mathrm{b\bar{b}}$ 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, 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 $A$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 $\mathrm{H}^{\pm}$ 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 $\lambda_3 = \lambda_{\mathrm{P1}} = \lambda_{\mathrm{P2}} = 3$ [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., $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,\mathrm{GeV}$.

The baryonic Z′ model [465] is an extension of the SM with an additional $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B, with a coupling of $\mathrm{h}_B\mathrm{Z'Z'}$ and vacuum expectation value of v_B. The Z′ and SM Higgs boson, h, interact with a coupling strength of $g_{\mathrm{hZ'Z'}} = m_{\mathrm{Z'}}^2 \sin\theta / v_B$, 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 $g_{\mathrm{q}}, g_{\chi} \sim 4\pi$ 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 $g_{\mathrm{hZ'Z'}}/m_{\mathrm{Z'}} = 1$, which implies $v_B = m_{\mathrm{Z'}} \sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{\mathrm{Z'}}$ 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

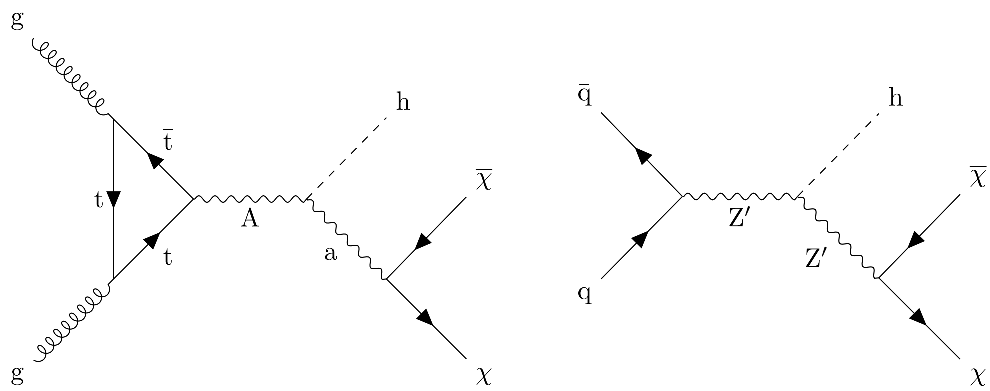

Figure 15: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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_{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 μs, 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 $t\bar{t}$ 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 $t\bar{t}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_T. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{SD}^2/p_T^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_T, decorrelating the N_2 variable as a function of ρ and p_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^{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 $\vec{p}_T^{miss}$ 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.

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, $\mathrm{t\bar{t}}$, W+jets	0	0	pass
Single-lepton, b-tagged	$\mathrm{t\bar{t}}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $\mathrm{t\bar{t}}$	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

E Event selection

Signal events are characterized by a high value of $p_{\mathrm{T}}^{\mathrm{miss}}$, 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 $\mathrm{t\bar{t}}$ production. To predict the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $\mathrm{t\bar{t}}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_{T} of the lepton(s) from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation in CRs, is used as a proxy for the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$, where $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ ($H_{\mathrm{T}}^{\mathrm{miss}}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\mathrm{T}}$ of all the particles (jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ calculation. Thresholds on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are reffered to as $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\mathrm{SD}} < 150\,\mathrm{GeV}$ and $N_2^{\mathrm{DDT}} < 0$. In the SR (CRs), $p_{\mathrm{T}}^{\mathrm{miss}}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($\vec{U}$) must be larger than 0.4 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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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_{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 $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 40$ (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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $\mathrm{b}\ell$ denotes the $\mathrm{t\bar{t}}$ process in the b-tagged single-lepton control sample that is used to estimate the $\mathrm{t\bar{t}}$ contribution in the SR, the expected number of $\mathrm{t\bar{t}}$ events, N_i, in the i^{th} bin of the SR is then given by $N_i = \mu_i^{\mathrm{t\bar{t}}} / T_i^{\mathrm{b}\ell}$, where $\mu_i^{\mathrm{t\bar{t}}}$ is a freely floating parameter included in the likelihood to scale the $\mathrm{t\bar{t}}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the $\mathrm{t\bar{t}}$ process, transfer factors to predict this contamination from $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{Z} \to \nu\nu$ and the $\mathrm{Z} \to \ell\ell$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $m_{\mathrm{A}} = 1.1\,\mathrm{TeV}$ and $m_{\mathrm{a}} = 150\,\mathrm{GeV}$ (with $\sin\theta = 0.35$ and $\tan\beta = 1$).

Systematic uncertainty	Type	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	1.9%
Double-b tagging	shape	SM h, signal	1.2%
N_2^{DDT} efficiency	7%	diboson, SM h, signal	
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\mathrm{T}}^{\mathrm{miss}}$ 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	$< 0.5\%$
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	
single-electron trigger	1%	all	
AK4 b tagging	shape	all	$< 0.5\%$
τ lepton veto	3%	all	$< 0.5\%$
Lepton efficiency	1% per lepton	all	
Heavy-flavor fraction	4-5%	Z+jets, W+jets	$< 0.5\%$
Renorm./fact. scales	shape	SM h	$< 0.5\%$
PDF	shape	SM h	
Multijet normalization	100%	multijet	
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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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.* $p_{\mathrm{T}}^{\mathrm{miss}}/U$. 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 $\mathrm{t\bar{t}}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_{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 $\mathrm{g} \to \mathrm{b\bar{b}}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substructure variable N_2^{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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 ($\mathrm{t\bar{t}}$, 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 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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ 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

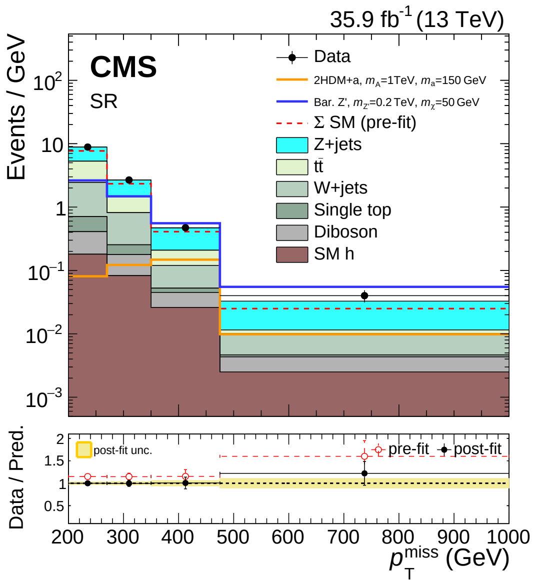

Figure 16: The p_T^{miss} distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

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_T^{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].

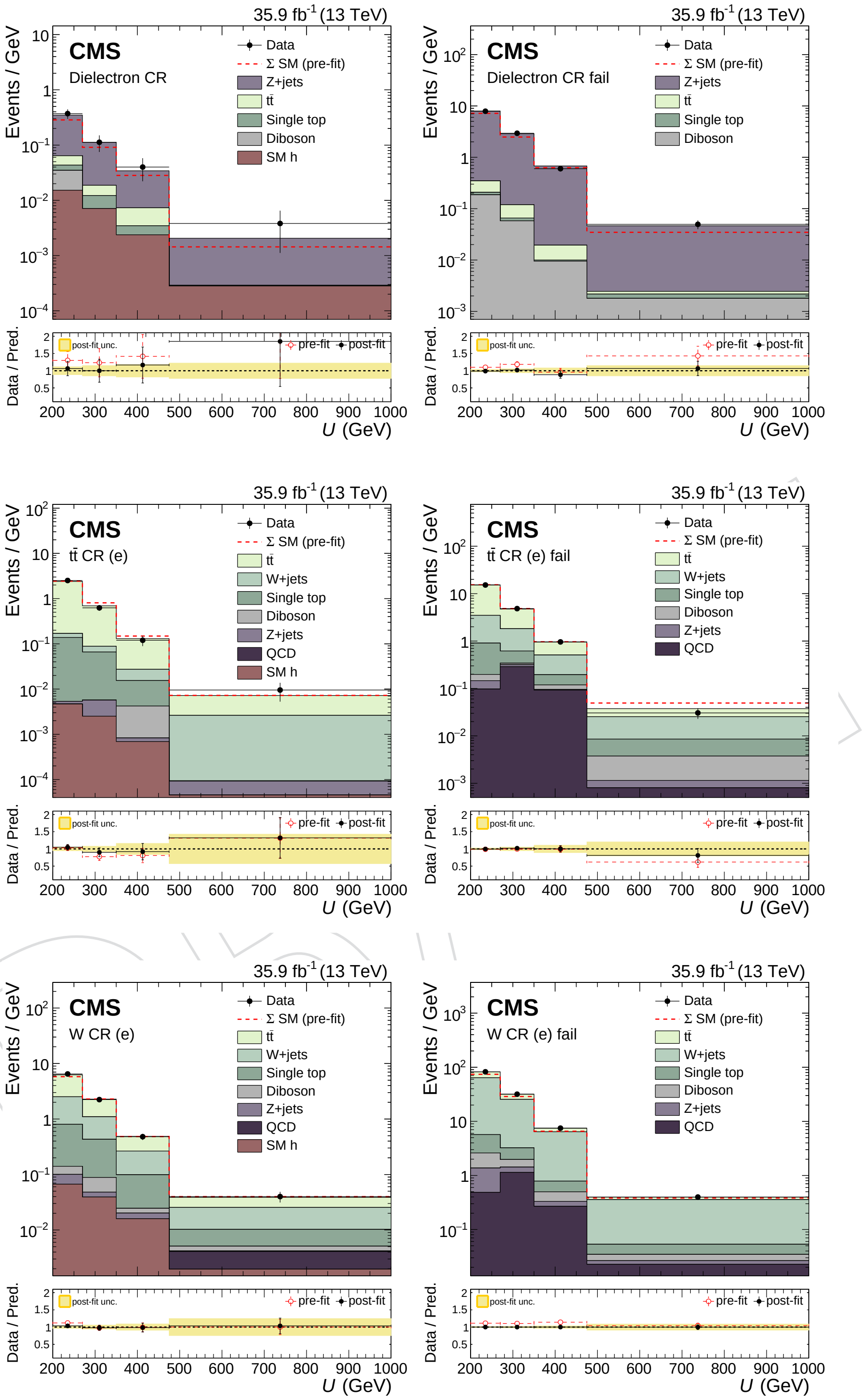

Figure 17: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

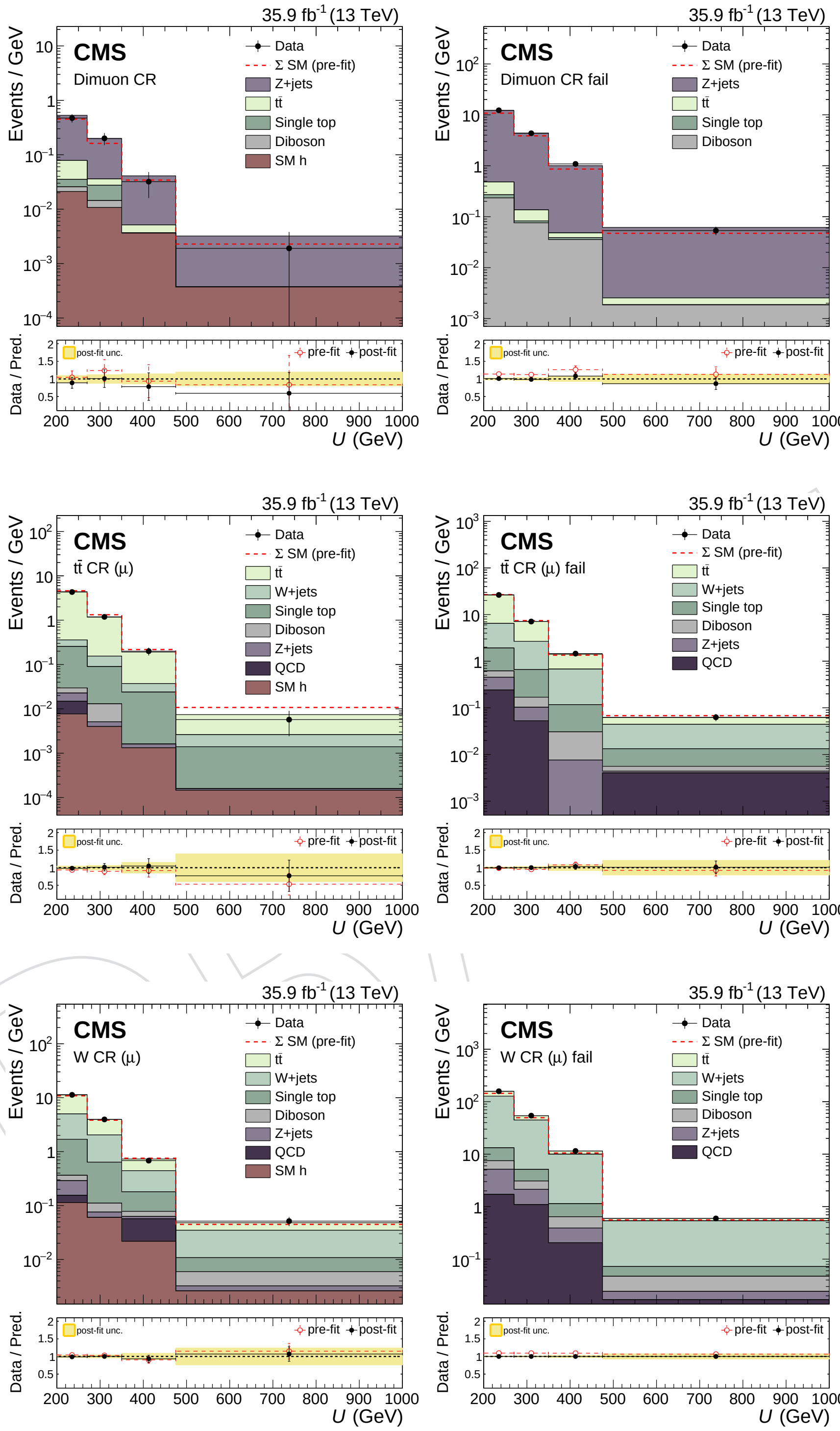

Figure 18: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

Table 9: Post-fit event yield expectations per $p_{\mathrm{T}}^{\mathrm{miss}}$ 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
$\mathrm{t\bar{t}}$	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, $m_{\mathrm{A}} = 1$ TeV, $m_{\mathrm{a}} = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z′, $m_{\mathrm{Z'}} = 0.2$ TeV, $m_\chi = 50$ 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 $m_{\mathrm{a}} = 150$ GeV, $\sin\theta = 0.35$ and $\tan\beta = 1$. Mixing angles with $0.35 < \sin\theta < 0.75$ are excluded for $m_{\mathrm{A}} = 600$ GeV and $m_{\mathrm{a}} = 200$ GeV, assuming $\tan\beta = 1$. Also excluded are $\tan\beta$ values between 0.5 and 2.0 (1.6) for $m_{\mathrm{a}} = 100$ (150) GeV and $m_{\mathrm{A}} = 600$ 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 $m_{\mathrm{Z'}}$ and m_χ for the baryonic Z′ model. For a DM mass of 1 GeV, masses $m_{\mathrm{Z'}} < 1.6$ 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:

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

where $\mu_{\mathrm{n}\chi}$ is the reduced mass of the DM-nucleon system, $f(g_{\mathrm{q}})$ 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 $m_\chi < 5$ 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

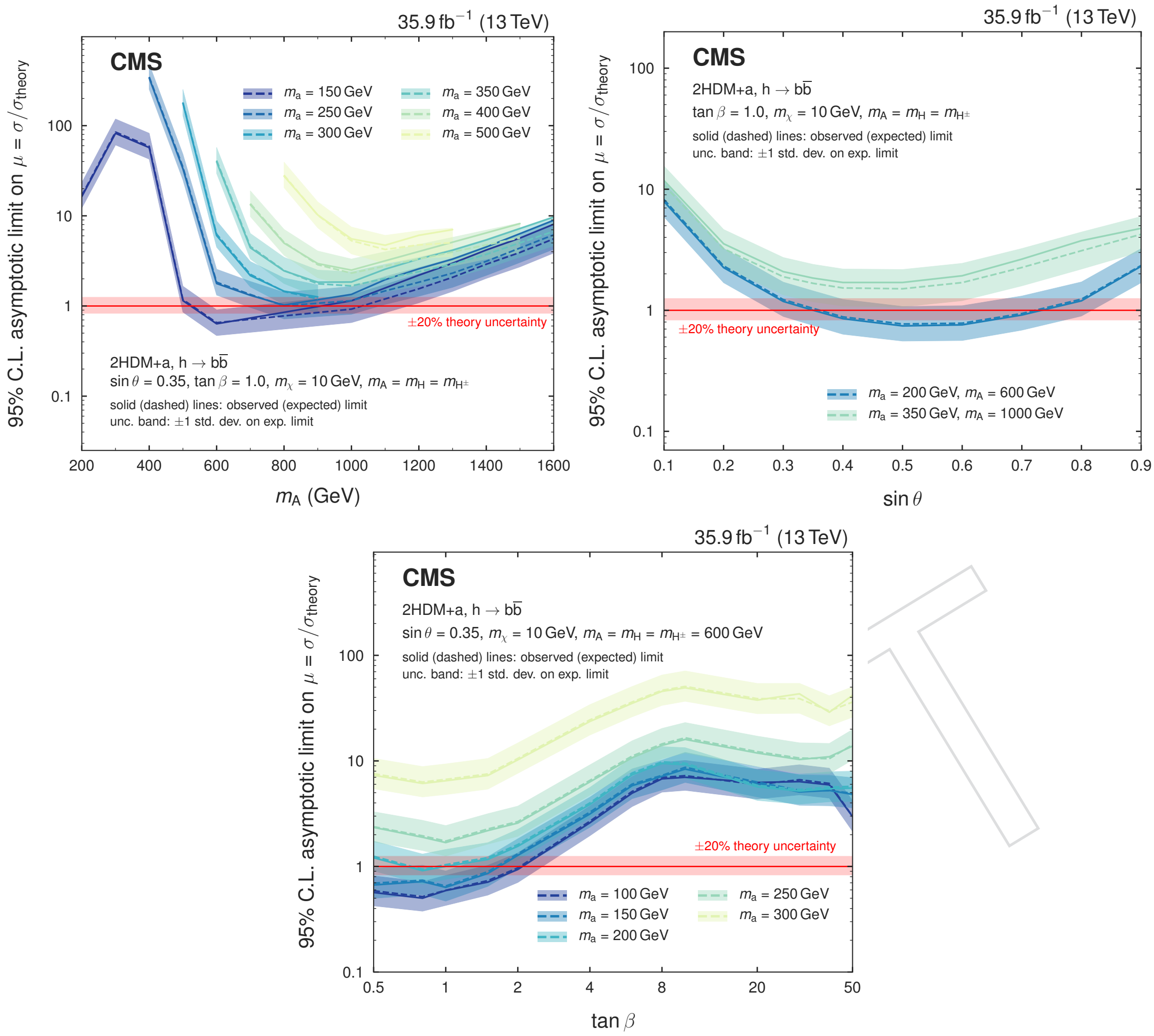

Figure 19: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_{A} and m_{a} (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

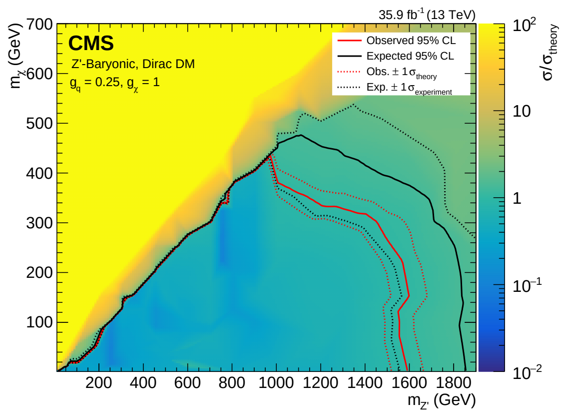

Figure 20: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

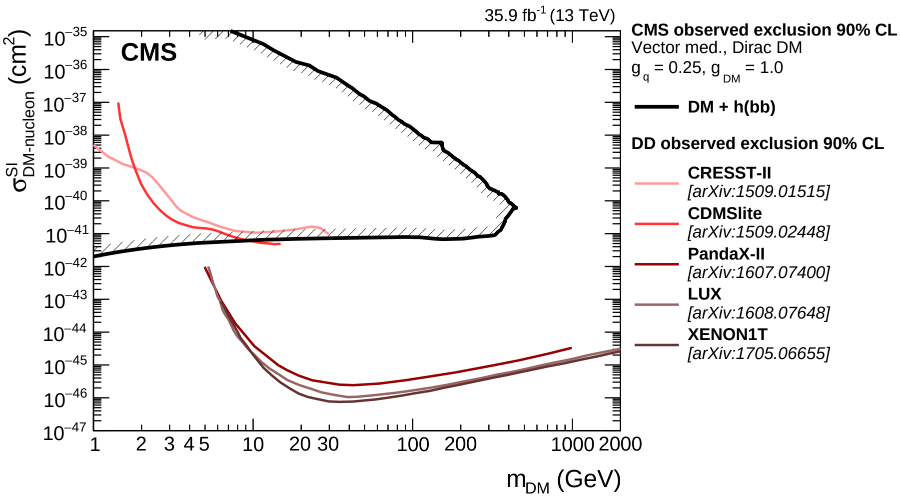

Figure 21: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits 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 $500 < m_A < 900$ GeV (where A is the heavy pseudoscalar boson) assuming $m_a = 150$ GeV. Scanning over $\sin\theta$ with $\tan\beta = 1$, we exclude $0.35 < \sin\theta < 0.75$ for $m_A = 600$ GeV and $m_a = 200$ GeV. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_A = 600$ GeV and $m_a = 100$ (150) GeV and $\sin\theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_\chi = 10$ 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 $m_\chi < 5$ 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

[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.

[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.

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

[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. 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.

[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 $\mathrm{b\bar{b}}$ 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.

[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.

[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.

[154] 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.

[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", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[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.

[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 $\sqrt{s} = 7$ 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 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 $\sqrt{s} = 7$ 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 $\sqrt{s} = 7$ 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 $\sqrt{s} = 7$ 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 $\sqrt{s} = 8$ 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 $\to \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13$ 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 $\sqrt{s} = 13$ 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.

[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.

[190] 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.

[191] 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.

[192] 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.

[193] 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.

[194] 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.

[195] 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

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 $\mathrm{h} \to \mathrm{b\bar{b}}$ 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, 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 $A$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., $m_H = m_{H^{\pm}} = m_A$. The DM particle χ is assumed to have a mass of $m_\chi = 10\,\text{GeV}$.

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 $g_q, g_\chi \sim 4\pi$ 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-

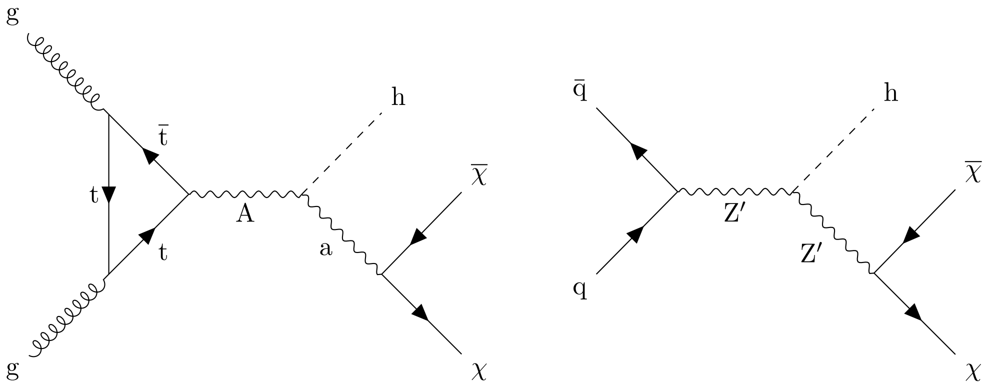

Figure 22: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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_{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 μs, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_{T}. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\mathrm{SD}}^2/p_{\mathrm{T}}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_{T}, decorrelating the N_2 variable as a function of ρ and p_{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^{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 $p_{\mathrm{T}} > 18$ 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 $p_{\mathrm{T}} > 15$ 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

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, $t\bar{t}$, W+jets	0	0	pass
Single-lepton, b-tagged	$t\bar{t}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $t\bar{t}$	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

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

Signal events are characterized by a high value of $p_{\mathrm{T}}^{\mathrm{miss}}$, 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 $t\bar{t}$ production. To predict the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $t\bar{t}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_{T} of the lepton(s) from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation in CRs, is used as a proxy for the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$, where $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ ($H_{\mathrm{T}}^{\mathrm{miss}}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\mathrm{T}}$ of all the particles (jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ calculation. Thresholds on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are reffered to as $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\mathrm{SD}} < 150\,\mathrm{GeV}$ and $N_2^{\mathrm{DDT}} < 0$. In the SR (CRs), $p_{\mathrm{T}}^{\mathrm{miss}}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($\vec{U}$) must be larger than 0.4 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 $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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_{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 $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 40$ (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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $\mathrm{b}\ell$ denotes the $\mathrm{t\bar{t}}$ process in the b-tagged single-lepton control sample that is used to estimate the $\mathrm{t\bar{t}}$ contribution in the SR, the expected number of $\mathrm{t\bar{t}}$ events, N_i, in the i^{th} bin of the SR is then given by $N_i = \mu_i^{\mathrm{t\bar{t}}}/T_i^{\mathrm{b}\ell}$, where $\mu_i^{\mathrm{t\bar{t}}}$ is a freely floating parameter included in the likelihood to scale the $\mathrm{t\bar{t}}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the $\mathrm{t\bar{t}}$ process, transfer factors to predict this contamination from $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{Z} \to \nu\nu$ and the $\mathrm{Z} \to \ell\ell$ 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 $m_{\mathrm{A}} = 1.1\,\mathrm{TeV}$ and $m_{\mathrm{a}} = 150\,\mathrm{GeV}$ (with $\sin\theta = 0.35$ and $\tan\beta = 1$).

Systematic uncertainty	Type	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	1.9%
Double-b tagging	shape	SM h, signal	1.2%
N_2^{DDT} efficiency	7%	diboson, SM h, signal	
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\mathrm{T}}^{\mathrm{miss}}$ 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	$< 0.5\%$
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	
single-electron trigger	1%	all	
AK4 b tagging	shape	all	$< 0.5\%$
τ lepton veto	3%	all	$< 0.5\%$
Lepton efficiency	1% per lepton	all	
Heavy-flavor fraction	4-5%	Z+jets, W+jets	$< 0.5\%$
Renorm./fact. scales	shape	SM h	$< 0.5\%$
PDF	shape	SM h	
Multijet normalization	100%	multijet	
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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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.* $p_{\mathrm{T}}^{\mathrm{miss}}/U$. 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 $\mathrm{t\bar{t}}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_{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 $\mathrm{g} \to \mathrm{b\bar{b}}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substructure variable N_2^{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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 ($\mathrm{t\bar{t}}$, 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 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 $\mathrm{t\bar{t}}$ 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-

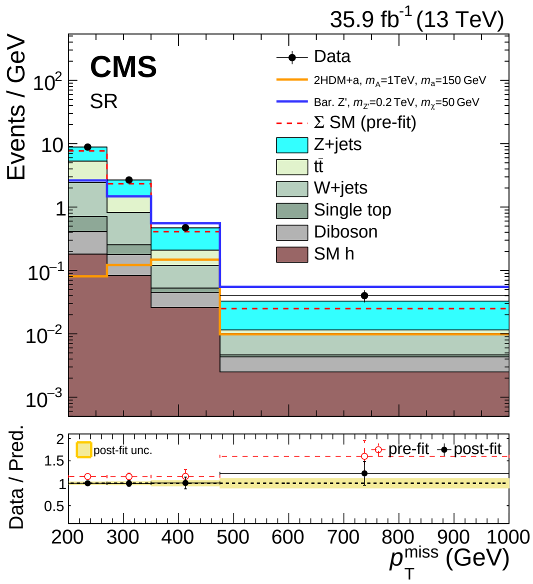

Figure 23: The $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate 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 $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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

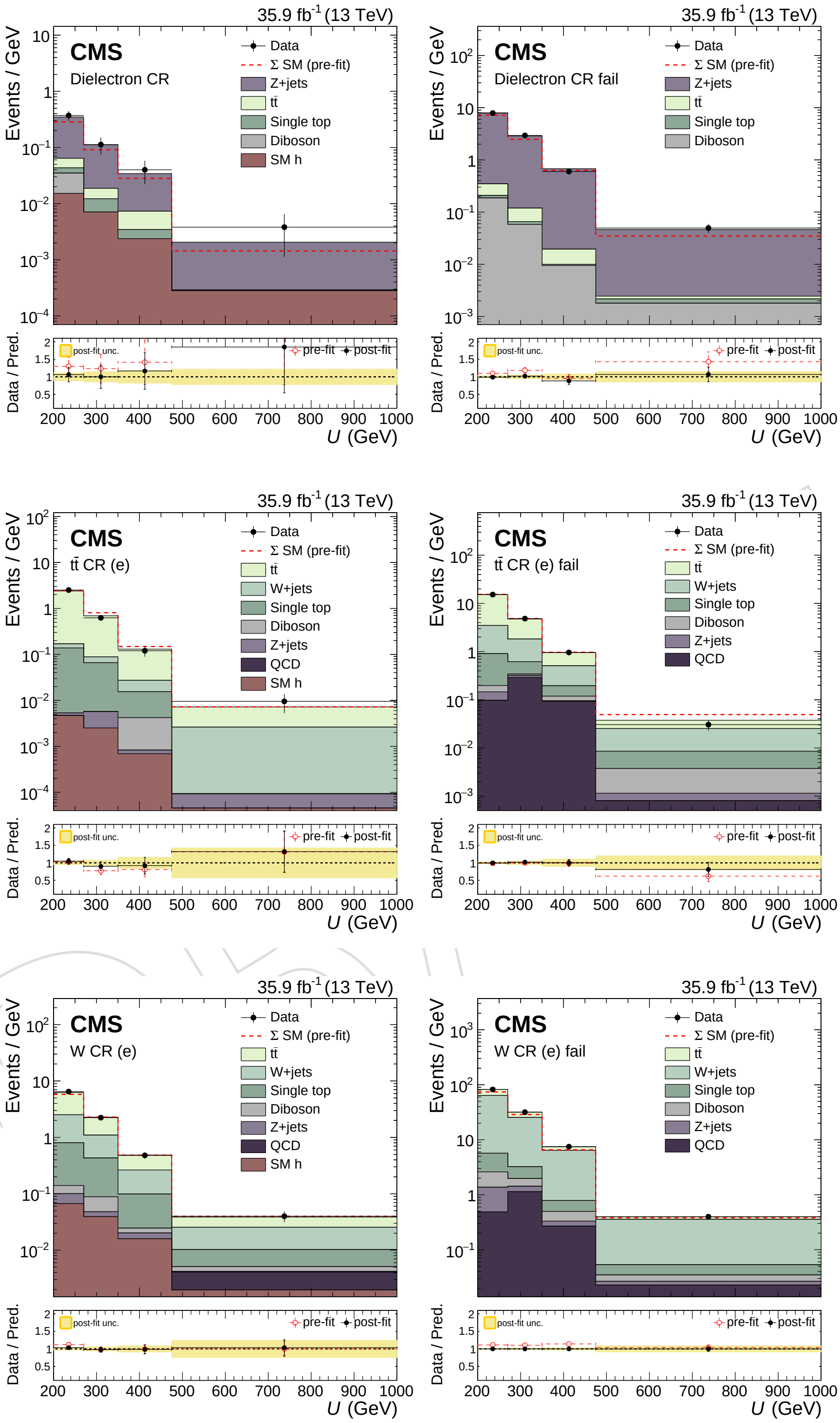

Figure 24: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

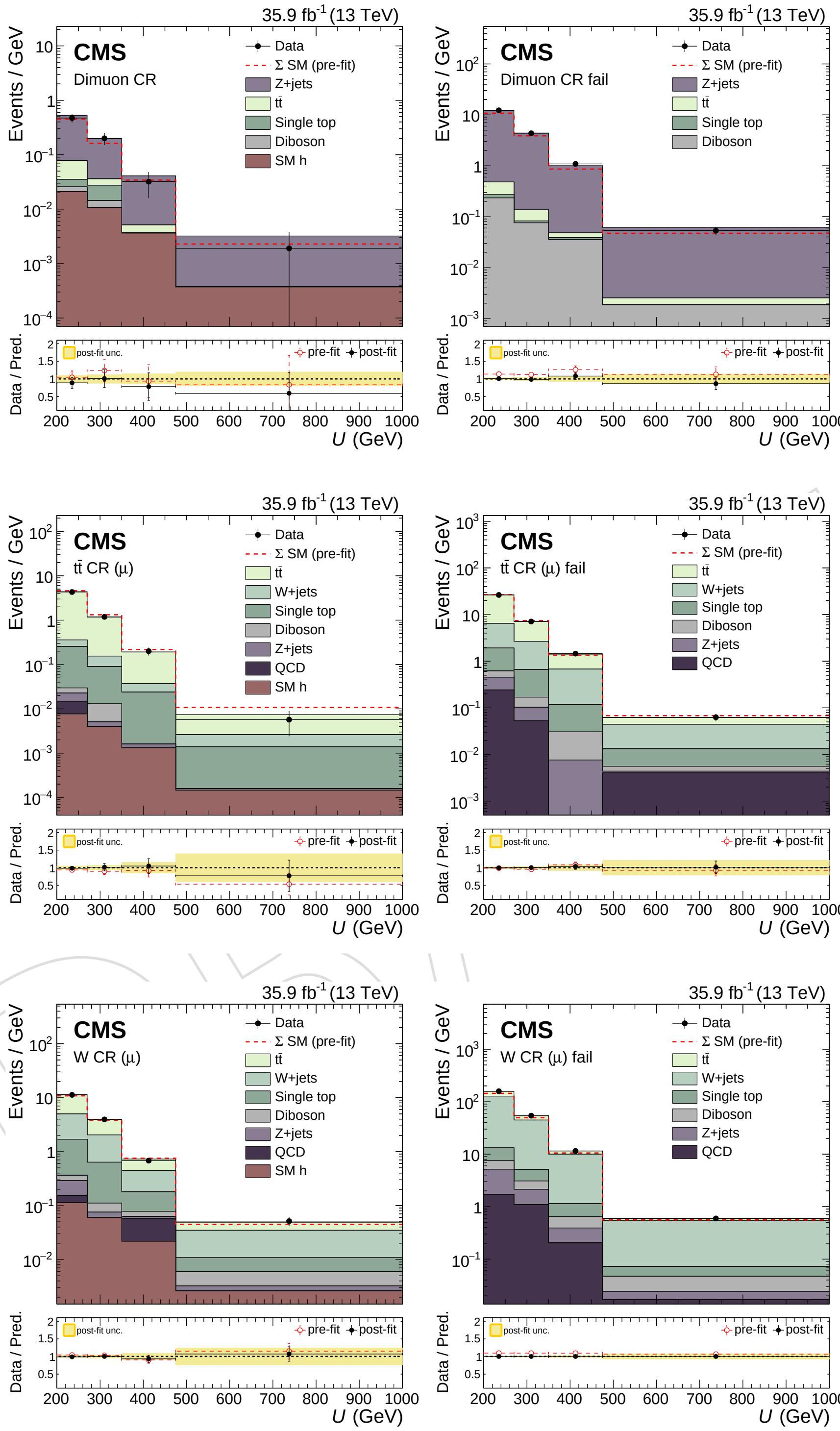

Figure 25: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

Table 12: Post-fit event yield expectations per $p_{\mathrm{T}}^{\mathrm{miss}}$ 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
$\mathrm{t\bar{t}}$	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, $m_{\mathrm{A}} = 1$ TeV, $m_{\mathrm{a}} = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z′, $m_{Z'} = 0.2$ TeV, $m_\chi = 50$ 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 $\mu = \sigma/\sigma_{\mathrm{theory}}$, 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 $m_{\mathrm{a}} = 150$ GeV, $\sin\theta = 0.35$ and $\tan\beta = 1$. Mixing angles with $0.35 < \sin\theta < 0.75$ are excluded for $m_{\mathrm{A}} = 600$ GeV and $m_{\mathrm{a}} = 200$ GeV, assuming $\tan\beta = 1$. Also excluded are $\tan\beta$ values between 0.5 and 2.0 (1.6) for $m_{\mathrm{a}} = 100$ (150) GeV and $m_{\mathrm{A}} = 600$ 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:

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

where $\mu_{\mathrm{n}\chi}$ is the reduced mass of the DM-nucleon system, $f(g_{\mathrm{q}})$ 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 $m_\chi < 5$ GeV.

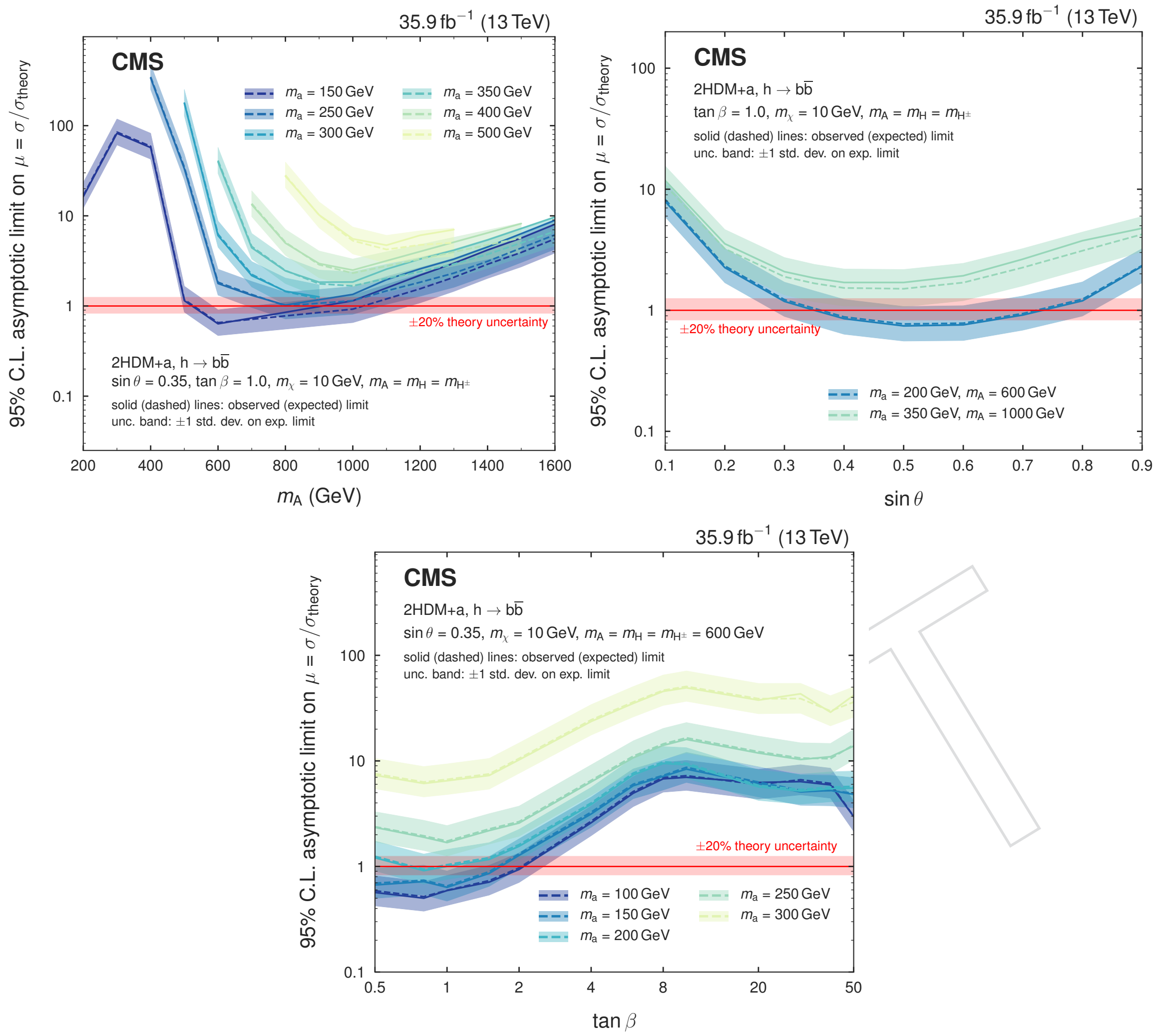

Figure 26: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_{A} and m_{a} (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

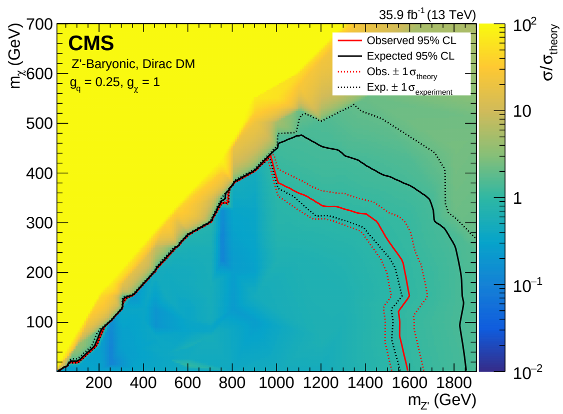

Figure 27: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

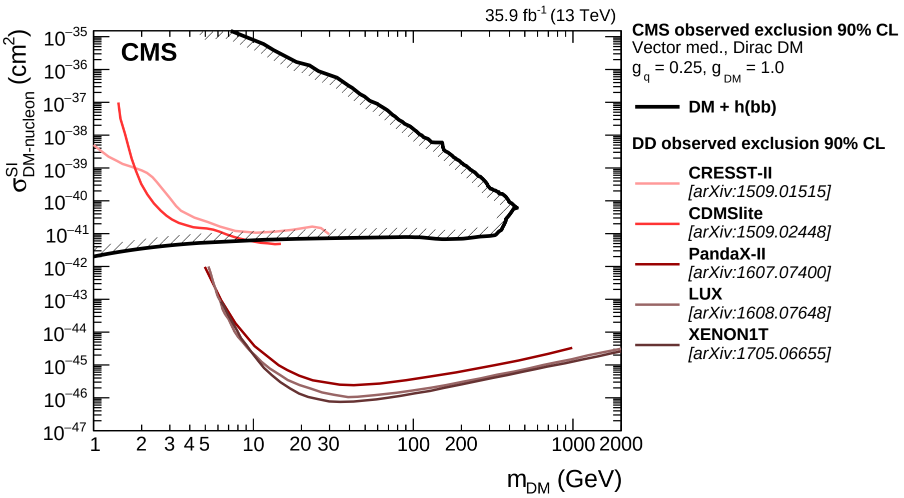

Figure 28: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits from CDMSLite [516], LUX [517], XENON-1T [518], PandaX-II [519], and CRESST-II [520] are shown.

I Summary

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 $m_\text{A} = 600\,\text{GeV}$ and $m_\text{a} = 200\,\text{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_\text{A} = 600\,\text{GeV}$ and $m_\text{a} = 100$ (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 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

[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.

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

[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`.

[204] 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`.

[205] 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`.

[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 $\mathrm{b\bar{b}}$ 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`.

[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.

[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.

[219] 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.

[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.

[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", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[223] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[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.

[225] 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.

[226] G. Collaboration, "Geant4 – a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/10.1016/S0168-9002(03)01368-8.

[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.

[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.

[229] 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.

[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.

[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.

[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 $\sqrt{s} = 8$ 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.

[244] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.

[245] 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.

[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.

[247] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7$ 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 $\sqrt{s} = 8$ 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 $\to \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13$ 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.

[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.

[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.

[257] 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.

[258] 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.

[259] 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.

[260] 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

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 $\mathrm{h} \to \mathrm{b\bar{b}}$ 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, 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 $A$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 $\mathrm{H}^{\pm}$ 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 $\lambda_3 = \lambda_{\mathrm{P1}} = \lambda_{\mathrm{P2}} = 3$ [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., $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$. The DM particle χ is assumed to have a mass of $m_\chi = 10\,\mathrm{GeV}$.

The baryonic Z′ model [465] is an extension of the SM with an additional $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B, with a coupling of $\mathrm{h}_B\mathrm{Z'Z'}$ and vacuum expectation value of v_B. The Z′ and SM Higgs boson, h, interact with a coupling strength of $g_{\mathrm{hZ'Z'}} = m^2_{\mathrm{Z'}} \sin\theta / v_B$, 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 $g_{\mathrm{q}}, g_\chi \sim 4\pi$ 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 $g_{\mathrm{hZ'Z'}}/m_{\mathrm{Z'}} = 1$, which implies $v_B = m_{\mathrm{Z'}} \sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{\mathrm{Z'}}$ 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

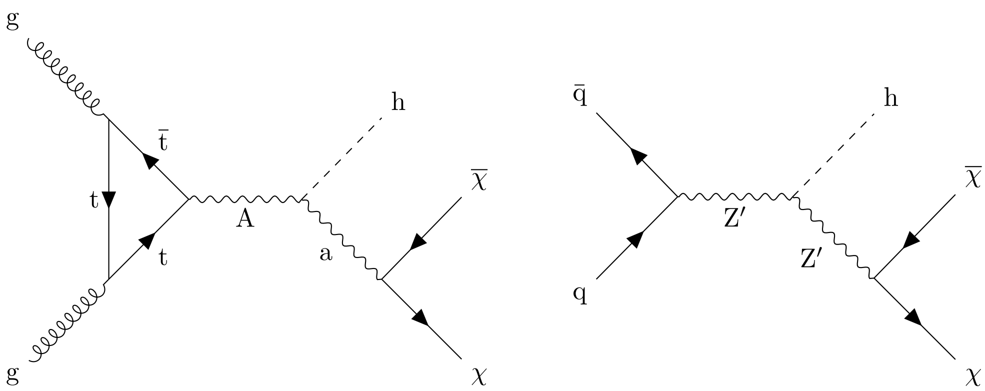

Figure 29: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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 μs, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_{T}. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\mathrm{SD}}^2/p_{\mathrm{T}}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_{T}, decorrelating the N_2 variable as a function of ρ and p_{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^{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 $p_{\mathrm{T}} > 18$ 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 $p_{\mathrm{T}} > 15$ 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.

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, $\mathrm{t\bar{t}}$, W+jets	0	0	pass
Single-lepton, b-tagged	$\mathrm{t\bar{t}}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $\mathrm{t\bar{t}}$	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

E Event selection

Signal events are characterized by a high value of $p_{\mathrm{T}}^{\mathrm{miss}}$, 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 $\mathrm{t\bar{t}}$ production. To predict the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $\mathrm{t\bar{t}}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_{T} of the lepton(s) from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation in CRs, is used as a proxy for the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$, where $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ ($H_{\mathrm{T}}^{\mathrm{miss}}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\mathrm{T}}$ of all the particles (jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ calculation. Thresholds on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are reffered to as $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\mathrm{SD}} < 150\,\mathrm{GeV}$ and $N_2^{\mathrm{DDT}} < 0$. In the SR (CRs), $p_{\mathrm{T}}^{\mathrm{miss}}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($\vec{U}$) must be larger than 0.4 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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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_{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 $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ events, while the other two single-lepton CRs target W+jets events. Single-muon events are selected using the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 40$ (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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $\mathrm{b}\ell$ denotes the $\mathrm{t\bar{t}}$ process in the b-tagged single-lepton control sample that is used to estimate the $\mathrm{t\bar{t}}$ contribution in the SR, the expected number of $\mathrm{t\bar{t}}$ events, N_i, in the i^{th} bin of the SR is then given by $N_i = \mu_i^{\mathrm{t\bar{t}}}/T_i^{\mathrm{b}\ell}$, where $\mu_i^{\mathrm{t\bar{t}}}$ is a freely floating parameter included in the likelihood to scale the $\mathrm{t\bar{t}}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the $\mathrm{t\bar{t}}$ process, transfer factors to predict this contamination from $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{Z} \to \nu\nu$ and the $\mathrm{Z} \to \ell\ell$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $m_{\mathrm{A}} = 1.1\,\mathrm{TeV}$ and $m_{\mathrm{a}} = 150\,\mathrm{GeV}$ (with $\sin\theta = 0.35$ and $\tan\beta = 1$).

Systematic uncertainty	Type	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	1.9%
Double-b tagging	shape	SM h, signal	1.2%
N_2^{DDT} efficiency	7%	diboson, SM h, signal	
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\mathrm{T}}^{\mathrm{miss}}$ 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	$< 0.5\%$
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	
single-electron trigger	1%	all	
AK4 b tagging	shape	all	$< 0.5\%$
τ lepton veto	3%	all	$< 0.5\%$
Lepton efficiency	1% per lepton	all	
Heavy-flavor fraction	4-5%	Z+jets, W+jets	$< 0.5\%$
Renorm./fact. scales	shape	SM h	$< 0.5\%$
PDF	shape	SM h	
Multijet normalization	100%	multijet	
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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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.* $p_{\mathrm{T}}^{\mathrm{miss}}/U$. 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 $\mathrm{t\bar{t}}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_{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 $\mathrm{g} \to \mathrm{b\bar{b}}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substructure variable N_2^{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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 ($t\bar{t}$, 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 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 $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ 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

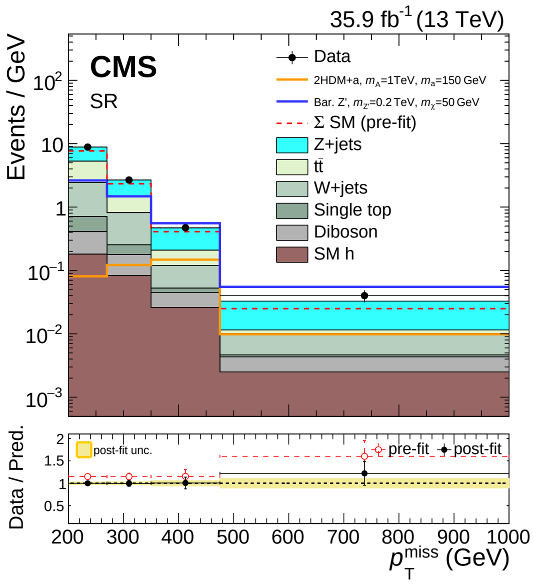

Figure 30: The $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate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

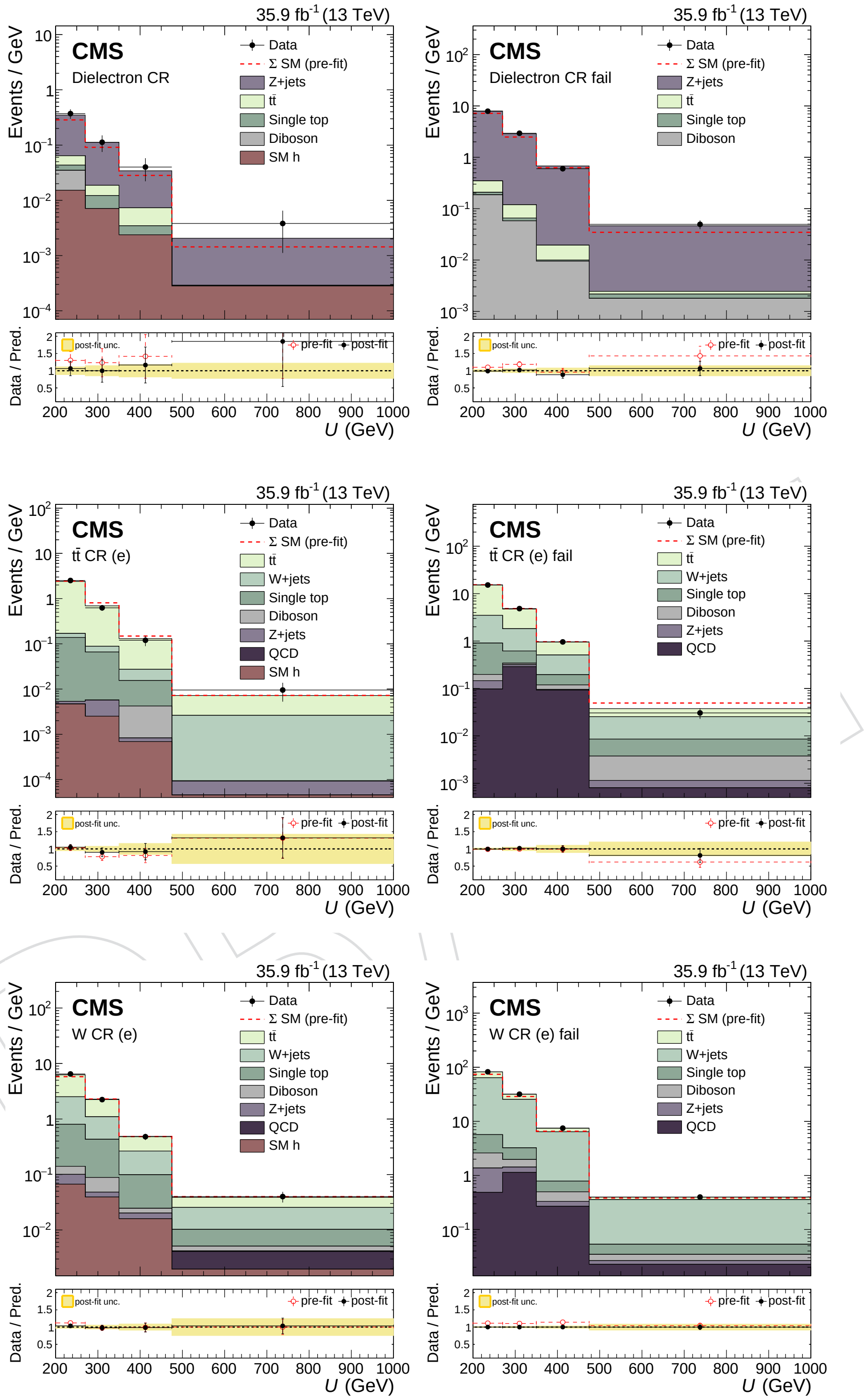

Figure 31: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

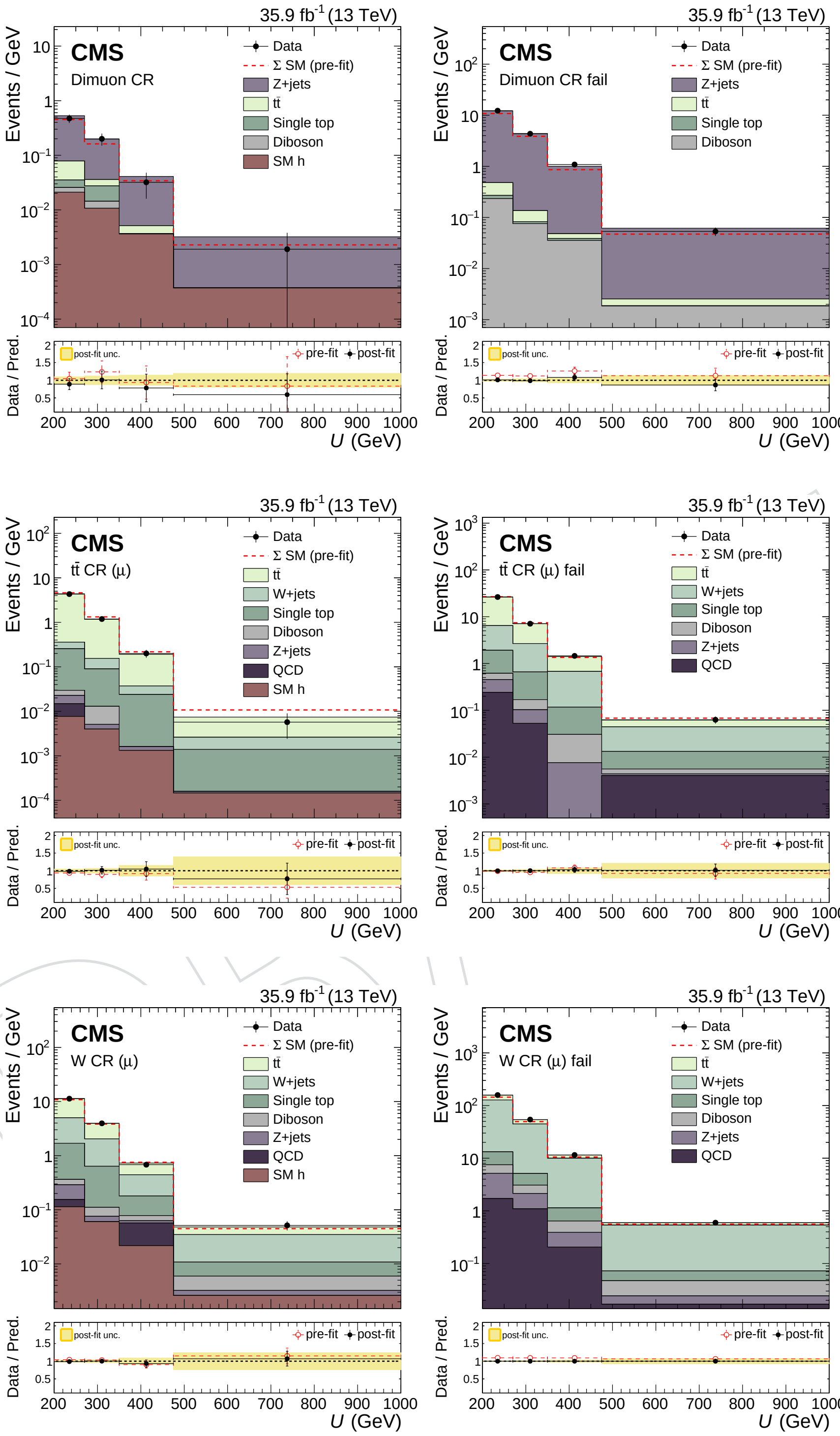

Figure 32: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

Table 15: Post-fit event yield expectations per $p_{\mathrm{T}}^{\mathrm{miss}}$ 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
$\mathrm{t\bar{t}}$	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, $m_{\mathrm{A}} = 1$ TeV, $m_{\mathrm{a}} = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z′, $m_{\mathrm{Z'}} = 0.2$ TeV, $m_{\chi} = 50$ 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 $m_{\mathrm{a}} = 150$ GeV, $\sin\theta = 0.35$ and $\tan\beta = 1$. Mixing angles with $0.35 < \sin\theta < 0.75$ are excluded for $m_{\mathrm{A}} = 600$ GeV and $m_{\mathrm{a}} = 200$ GeV, assuming $\tan\beta = 1$. Also excluded are $\tan\beta$ values between 0.5 and 2.0 (1.6) for $m_{\mathrm{a}} = 100$ (150) GeV and $m_{\mathrm{A}} = 600$ 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 $m_{\mathrm{Z'}}$ and m_{χ} for the baryonic Z′ model. For a DM mass of 1 GeV, masses $m_{\mathrm{Z'}} < 1.6$ 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:

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

where $\mu_{\mathrm{n}\chi}$ is the reduced mass of the DM-nucleon system, $f(g_{\mathrm{q}})$ 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 $m_{\chi} < 5$ 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

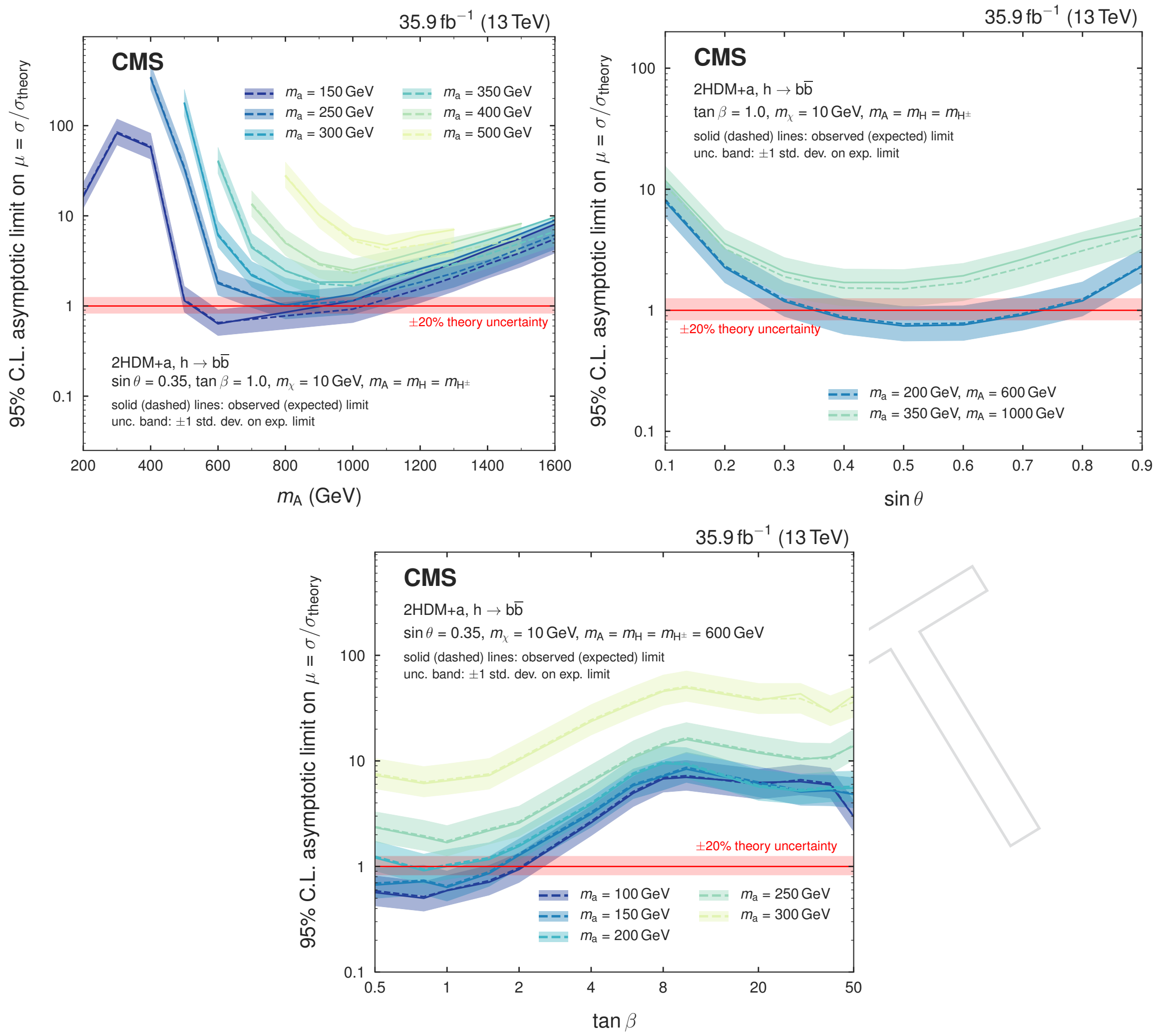

Figure 33: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_{A} and m_{a} (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

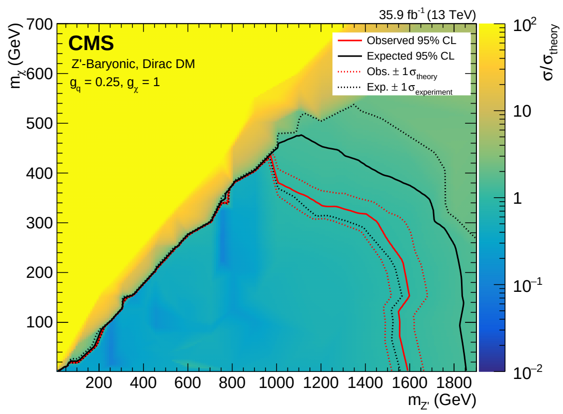

Figure 34: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

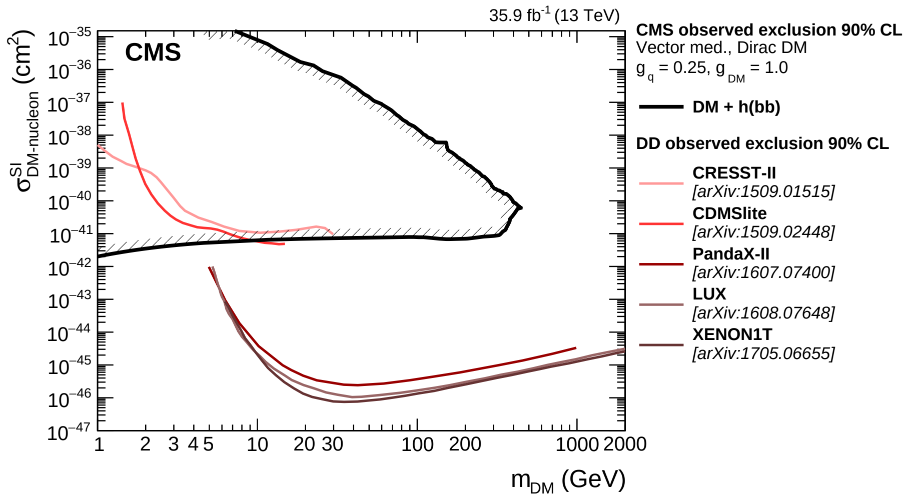

Figure 35: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits 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 $500 < m_A < 900$ GeV (where A is the heavy pseudoscalar boson) assuming $m_a = 150$ GeV. Scanning over $\sin\theta$ with $\tan\beta = 1$, we exclude $0.35 < \sin\theta < 0.75$ for $m_\mathrm{A} = 600$ GeV and $m_\mathrm{a} = 200$ GeV. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_\mathrm{A} = 600$ GeV and $m_\mathrm{a} = 100$ (150) GeV and $\sin\theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_\chi = 10$ 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 $m_\chi < 5$ 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

[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`.

[262] 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`.

[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`.

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

[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`.

[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`.

[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`.

[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.

[269] 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.

[270] 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.

[271] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

[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.

[275] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $\mathrm{b\bar{b}}$ or $\gamma\gamma$ at $\sqrt{s} = 13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10(2017)180, arXiv:1703.05236.

[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.

[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.

[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.

[286] 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.

[287] 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.

[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.

[290] 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.

[291] G. Collaboration, "Geant4 – a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/10.1016/S0168-9002(03)01368-8.

[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.

[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.

[294] 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.

[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.

[296] 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.

[297] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) 05011.

[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.

[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 TeV", *JINST* **10** (2015) 06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[302] 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.

[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 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.

[309] 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 $\sqrt{s} = 7$ 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.

[312] CMS Collaboration, "Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$", *JHEP* **06** (2014) 120, doi:10.1007/JHEP06(2014)120, arXiv:1402.1521.

[313] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$", *Phys. Rev. Lett.* **112** (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.

[314] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\rightarrow \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Phys. Lett. B* **763** (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

[315] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Phys. Lett. B* **766** (2017) 268, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.

[316] 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.

[317] 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.

[318] 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.

[319] 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.

[320] 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.

[321] 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.

[322] 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.

[323] 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.

[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.

[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

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 $\mathrm{h} \to \mathrm{b\bar{b}}$ 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 $35.9\,\mathrm{fb}^{-1}$. 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 h, as-

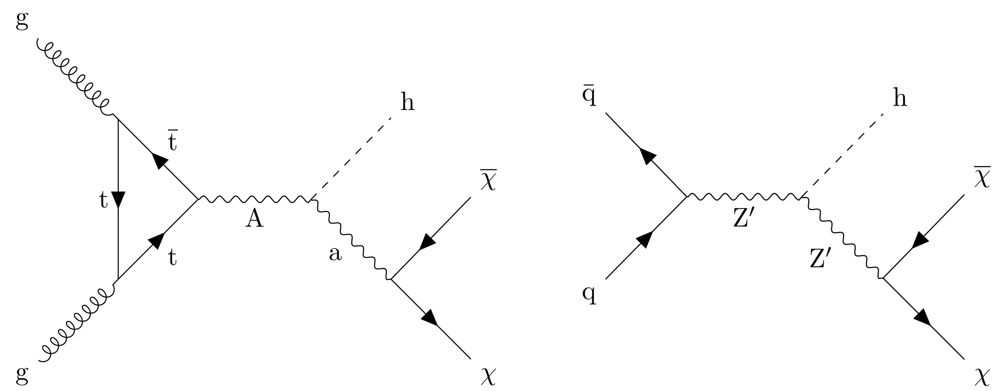

Figure 36: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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 $\mathrm{H}^{\pm}$ 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 $\lambda_3 = \lambda_{\mathrm{P1}} = \lambda_{\mathrm{P2}} = 3$ [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., $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$. The DM particle χ is assumed to have a mass of $m_\chi = 10\,\mathrm{GeV}$.

The baryonic Z′ model [465] is an extension of the SM with an additional $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B, with a coupling of $\mathrm{h}_B\mathrm{Z'Z'}$ and vacuum expectation value of v_B. The Z′ and SM Higgs boson, h, interact with a coupling strength of $g_{\mathrm{hZ'Z'}} = m_{\mathrm{Z'}}^2 \sin\theta / v_B$, 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 $g_{\mathrm{q}}, g_\chi \sim 4\pi$ 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 $g_{\mathrm{hZ'Z'}}/m_{\mathrm{Z'}} = 1$, which implies $v_B = m_{\mathrm{Z'}} \sin\theta$. 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 μs, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_{T}. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\mathrm{SD}}^2/p_{\mathrm{T}}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_{T}, decorrelating the N_2 variable as a function of ρ and p_{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^{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 $p_{\mathrm{T}} > 18$ 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 $p_{\mathrm{T}} > 15$ 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

Signal events are characterized by a high value of $p_{\mathrm{T}}^{\mathrm{miss}}$, 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 $\mathrm{t\bar{t}}$ production. To predict the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $\mathrm{t\bar{t}}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_{T} of the lepton(s) from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation in CRs, is used as a proxy for the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$, where $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ ($H_{\mathrm{T}}^{\mathrm{miss}}$) 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, $t\bar{t}$, W+jets	0	0	pass
Single-lepton, b-tagged	$t\bar{t}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $t\bar{t}$	0	1	pass/fail
Dilepton	Z+jets	0	2	pass/fail

magnitude of the vectorial (scalar) sum of $\vec{p}_{\mathrm{T}}$ of all the particles (jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$) at the trigger level. Muon candidates are excluded from the online $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ calculation. Thresholds on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are reffered to as $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\mathrm{SD}} < 150\,\mathrm{GeV}$ and $N_2^{\mathrm{DDT}} < 0$. In the SR (CRs), $p_{\mathrm{T}}^{\mathrm{miss}}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($\vec{U}$) must be larger than 0.4 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 $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of the Z+jets process in the SR, dimuon and dielectron CRs are used. Dimuon events are selected online employing the same $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ 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_{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 $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 40$ (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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $\mathrm{b}\ell$ denotes the $\mathrm{t\bar{t}}$ process in the b-tagged single-lepton control sample that is used to estimate the $\mathrm{t\bar{t}}$ contribution in the SR, the expected number of $\mathrm{t\bar{t}}$ events, N_i, in the i^{th} bin of the SR is then given by $N_i = \mu_i^{\mathrm{t\bar{t}}}/T_i^{\mathrm{b}\ell}$, where $\mu_i^{\mathrm{t\bar{t}}}$ is a freely floating parameter included in the likelihood to scale the $\mathrm{t\bar{t}}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the $\mathrm{t\bar{t}}$ process, transfer factors to predict this contamination from $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{Z} \to \nu\nu$ and the $\mathrm{Z} \to \ell\ell$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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

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 $m_{\mathrm{A}} = 1.1\,\mathrm{TeV}$ and $m_{\mathrm{a}} = 150\,\mathrm{GeV}$ (with $\sin\theta = 0.35$ and $\tan\beta = 1$).

Systematic uncertainty	Type	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, $\mathrm{t\bar{t}}$	1.9%
Double-b tagging	shape	SM h, signal	1.2%
N_2^{DDT} efficiency	7%	diboson, SM h, signal	
CA15 jet energy	4%	t, diboson, multijet, SM h, signal	0.8%
$p_{\mathrm{T}}^{\mathrm{miss}}$ 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	$< 0.5\%$
$p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	1%	all	
single-electron trigger	1%	all	
AK4 b tagging	shape	all	$< 0.5\%$
τ lepton veto	3%	all	$< 0.5\%$
Lepton efficiency	1% per lepton	all	
Heavy-flavor fraction	4-5%	Z+jets, W+jets	$< 0.5\%$
Renorm./fact. scales	shape	SM h	$< 0.5\%$
PDF	shape	SM h	
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.* $p_{\mathrm{T}}^{\mathrm{miss}}/U$. 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 $\mathrm{t\bar{t}}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_{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 $\mathrm{g} \to \mathrm{b\bar{b}}$ splittings. A 7% rate uncertainty in the efficiency of the requirement on the substructure variable N_2^{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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 ($\mathrm{t\bar{t}}$, 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 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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ 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 efficien-

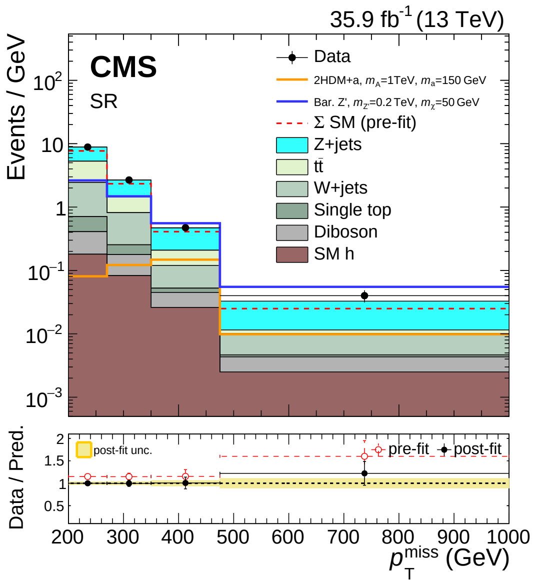

Figure 37: The p_T^{miss} distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

Table 18: Post-fit event yield expectations per p_T^{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
tt̄	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, $m_A = 1$ TeV, $m_a = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2$ TeV, $m_\chi = 50$ 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_T^{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 $\mu =$

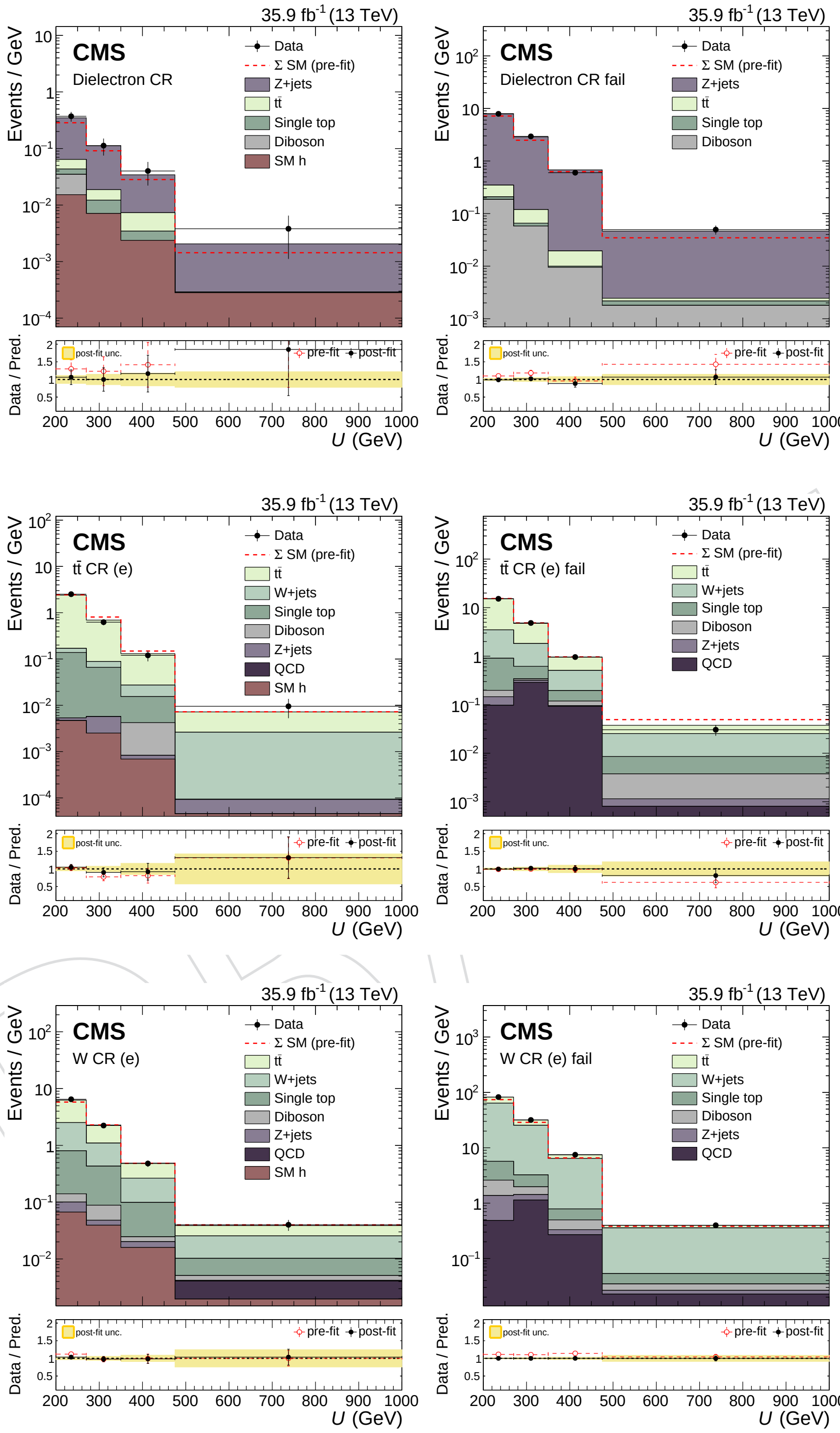

Figure 38: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

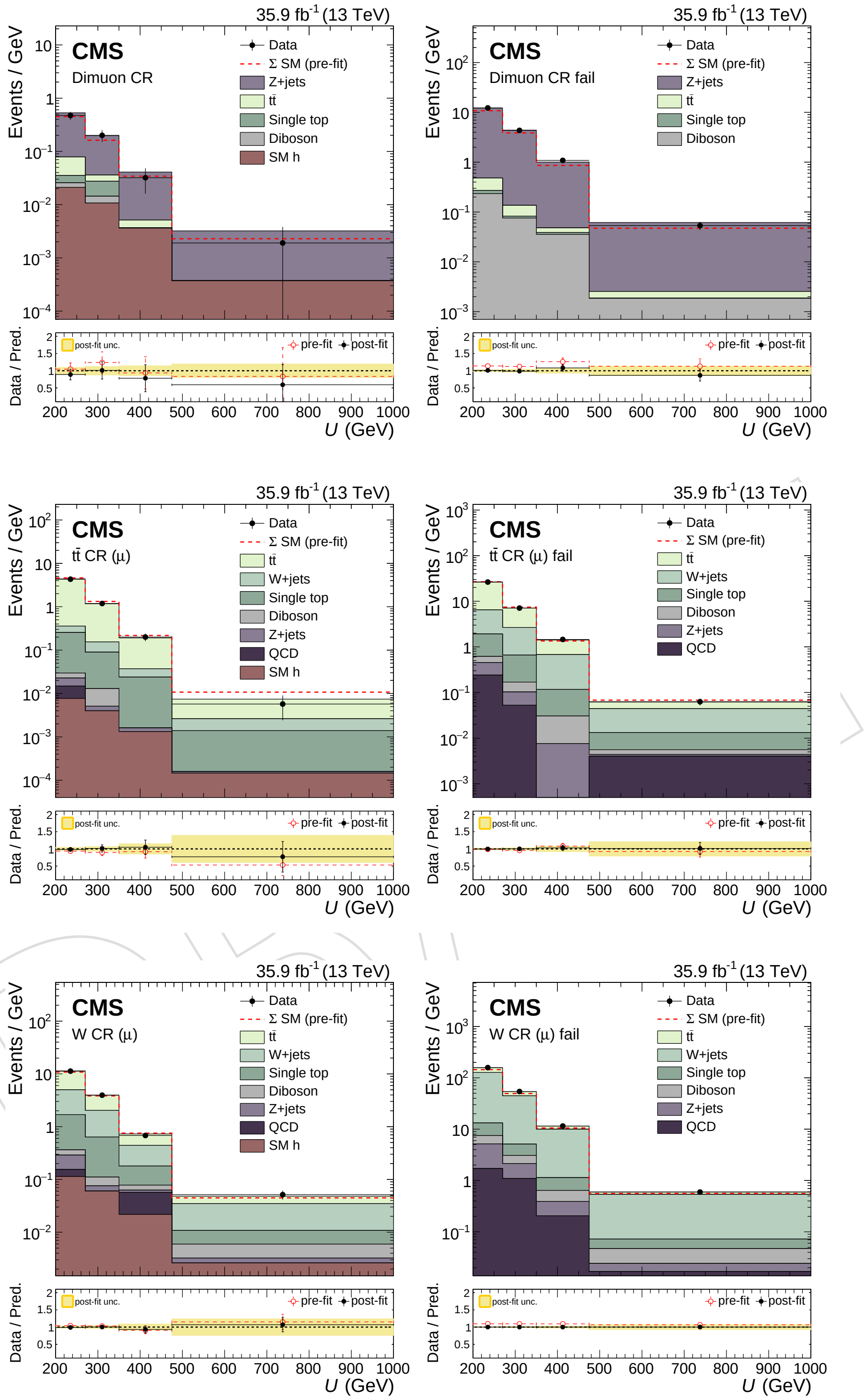

Figure 39: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

$\sigma/\sigma_{\text{theory}}$, 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.

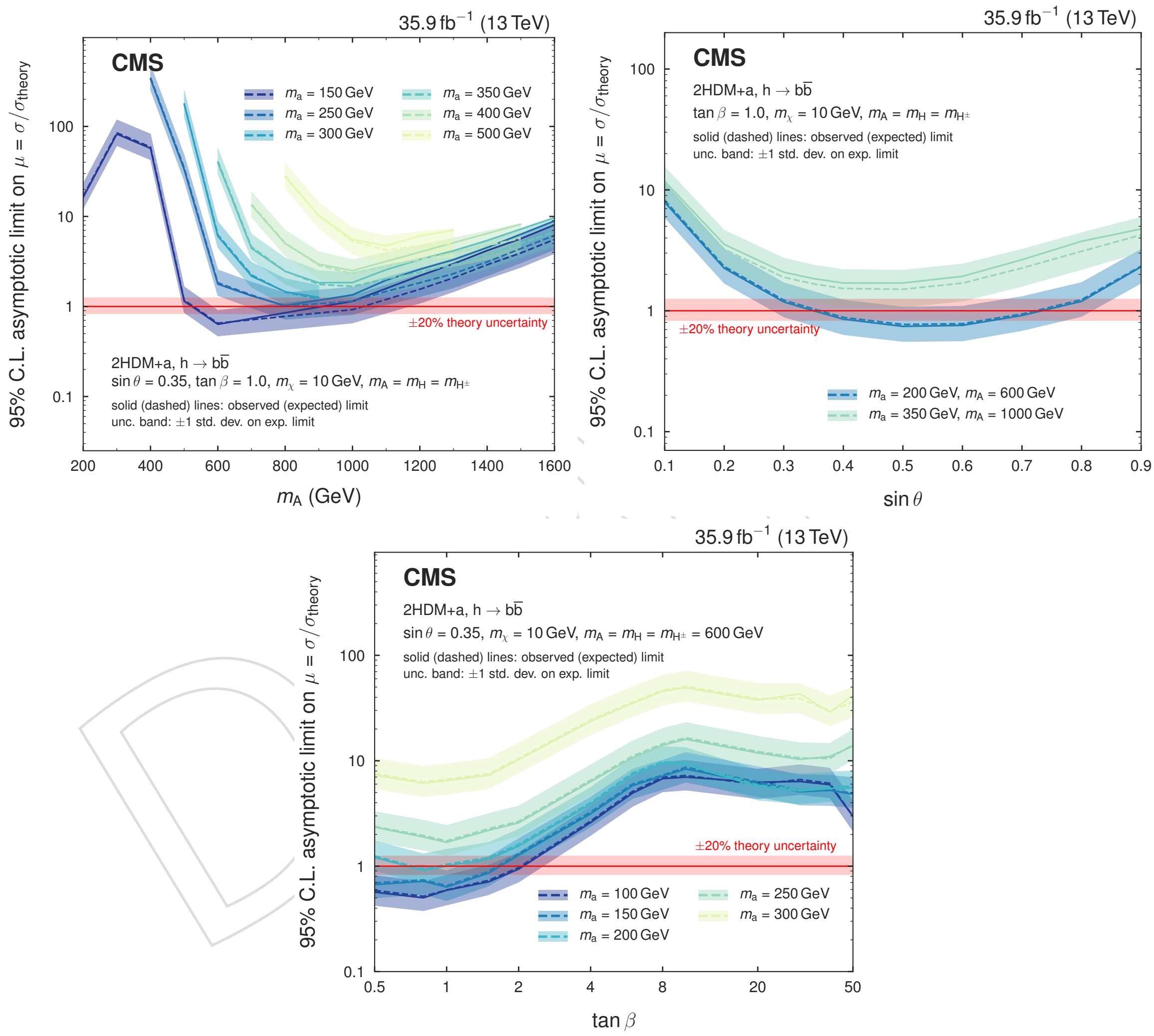

Figure 40: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

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.

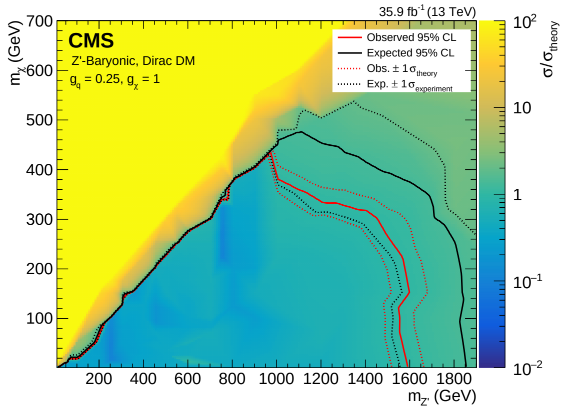

Figure 41: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

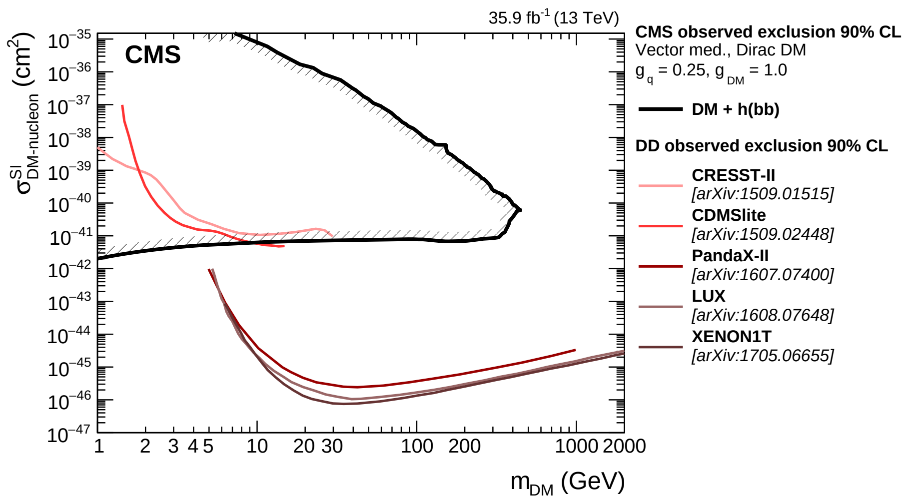

Figure 42: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits 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:

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

where $\mu_{\mathrm{n}\chi}$ is the reduced mass of the DM-nucleon system, $f(g_{\mathrm{q}})$ 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 $m_\chi < 5\,\mathrm{GeV}$.

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 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 $m_\text{A} = 600\,\text{GeV}$ and $m_\text{a} = 200\,\text{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_\text{A} = 600\,\text{GeV}$ and $m_\text{a} = 100$ (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 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

[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.

[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.

[334] 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.

[335] 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.

[336] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

[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 $\mathrm{b\bar{b}}$ 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.

[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.

[347] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 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.

[351] 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.

[352] 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.

[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.

[355] 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.

[356] G. Collaboration, "Geant4 – a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/10.1016/S0168-9002(03)01368-8.

[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.

[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.

[359] 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.

[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.

[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.

[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.

[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.

[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.

[367] 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.

[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. arXiv:1009.1003.

[372] 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.

[373] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

[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.

[375] 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.

[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 $\sqrt{s} = 7$ 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 $\sqrt{s} = 8$ 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 $\to \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13$ 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 at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **766** (2017) 268, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.

[381] 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.

[382] 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.

[383] 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.

[384] 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.

[385] 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.

[386] 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.

[387] 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.

[388] 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.

[389] 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.

[390] 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

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 $\mathrm{h} \to \mathrm{b\bar{b}}$ 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, 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 $A$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 $\mathrm{H}^{\pm}$ 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 $\lambda_3 = \lambda_{\mathrm{P1}} = \lambda_{\mathrm{P2}} = 3$ [467]. Masses of the

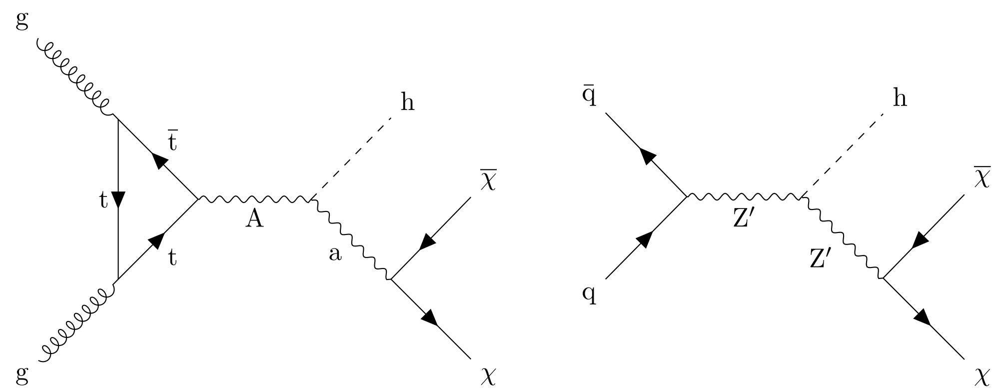

Figure 43: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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., $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,\mathrm{GeV}$.

The baryonic Z′ model [465] is an extension of the SM with an additional $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B, with a coupling of $\mathrm{h}_B\mathrm{Z'Z'}$ and vacuum expectation value of v_B. The Z′ and SM Higgs boson, h, interact with a coupling strength of $g_{\mathrm{hZ'Z'}} = m^2_{\mathrm{Z'}} \sin\theta / v_B$, 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 $g_{\mathrm{q}}, g_{\chi} \sim 4\pi$ 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 $g_{\mathrm{hZ'Z'}}/m_{\mathrm{Z'}} = 1$, which implies $v_B = m_{\mathrm{Z'}} \sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{\mathrm{Z'}}$ 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 μs, 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 $t\bar{t}$ 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 $t\bar{t}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_{T}. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\mathrm{SD}}^2/p_{\mathrm{T}}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_{T}, decorrelating the N_2 variable as a function of ρ and p_{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^{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 $\vec{p}_T^{\text{miss}}$ 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

Signal events are characterized by a high value of p_T^{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 $t\bar{t}$ production. To predict the p_T^{miss} spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $t\bar{t}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_T of the lepton(s) from the p_T^{miss} computation in CRs, is used as a proxy for the p_T^{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 $p_T > 20$ 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 $p_T > 200$ GeV and $|\eta| < 2.4$ is required. It is also required that $100 < m_{SD} < 150$ GeV and $N_2^{\text{DDT}} < 0$. In the SR (CRs), p_T^{miss} (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_T^{\text{miss}}$ ($\vec{U}$) 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, $t\bar{t}$, W+jets	0	0	pass
Single-lepton, b-tagged	$t\bar{t}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $t\bar{t}$	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_T^{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_T^{miss} triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_T > 20\,\text{GeV}$ and $p_T > 10\,\text{GeV}$ 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_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 $t\bar{t}$ 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_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 bℓ 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 $t\bar{t}$ 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 $t\bar{t}$ process, transfer factors to predict this contamination from $t\bar{t}$ 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 $t\bar{t}$ 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 $Z \to \nu\nu$ and the $Z \to \ell\ell$ 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 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 $t\bar{t}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_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 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 $\tan\beta = 1$).

Systematic uncertainty	Type	Processes	Impact on sensitivity
Double-b mistagging	shape	Z+jets, W+jets, $t\bar{t}$	4.8%
Transfer factor stat. uncertainties	shape	Z+jets, W+jets, $t\bar{t}$	1.9%
Double-b tagging	shape	SM h, signal	1.2%
N_2^{DDT} efficiency	7%	diboson, SM h, signal	
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_T^{miss} trigger muon multiplicity	shape	Z+jets, W+jets	$< 0.5\%$
p_T^{miss} trigger efficiency	1%	all	
single-electron trigger	1%	all	
AK4 b tagging	shape	all	$< 0.5\%$
τ lepton veto	3%	all	$< 0.5\%$
Lepton efficiency	1% per lepton	all	
Heavy-flavor fraction	4-5%	Z+jets, W+jets	$< 0.5\%$
Renorm./fact. scales	shape	SM h	$< 0.5\%$
PDF	shape	SM h	
Multijet normalization	100%	multijet	
Theoretical cross section	20%	t, diboson	

ture variable N_2^{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 $p_{T,trig}^{miss}$ 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 ($t\bar{t}$, 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 $p_{T,trig}^{miss}$ 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

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 $p_T^{miss} > 250\,\text{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_T^{miss}$ 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_T^{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 $\mu =$

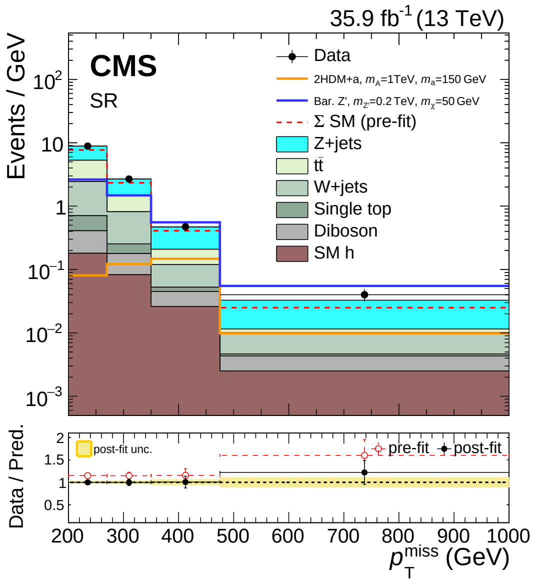

Figure 44: The $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate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

Table 21: Post-fit event yield expectations per $p_{\mathrm{T}}^{\mathrm{miss}}$ 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
$\mathrm{t\bar{t}}$	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, $m_{\mathrm{A}} = 1\,\mathrm{TeV}$, $m_{\mathrm{a}} = 150\,\mathrm{GeV}$	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2\,\mathrm{TeV}$, $m_\chi = 50\,\mathrm{GeV}$	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

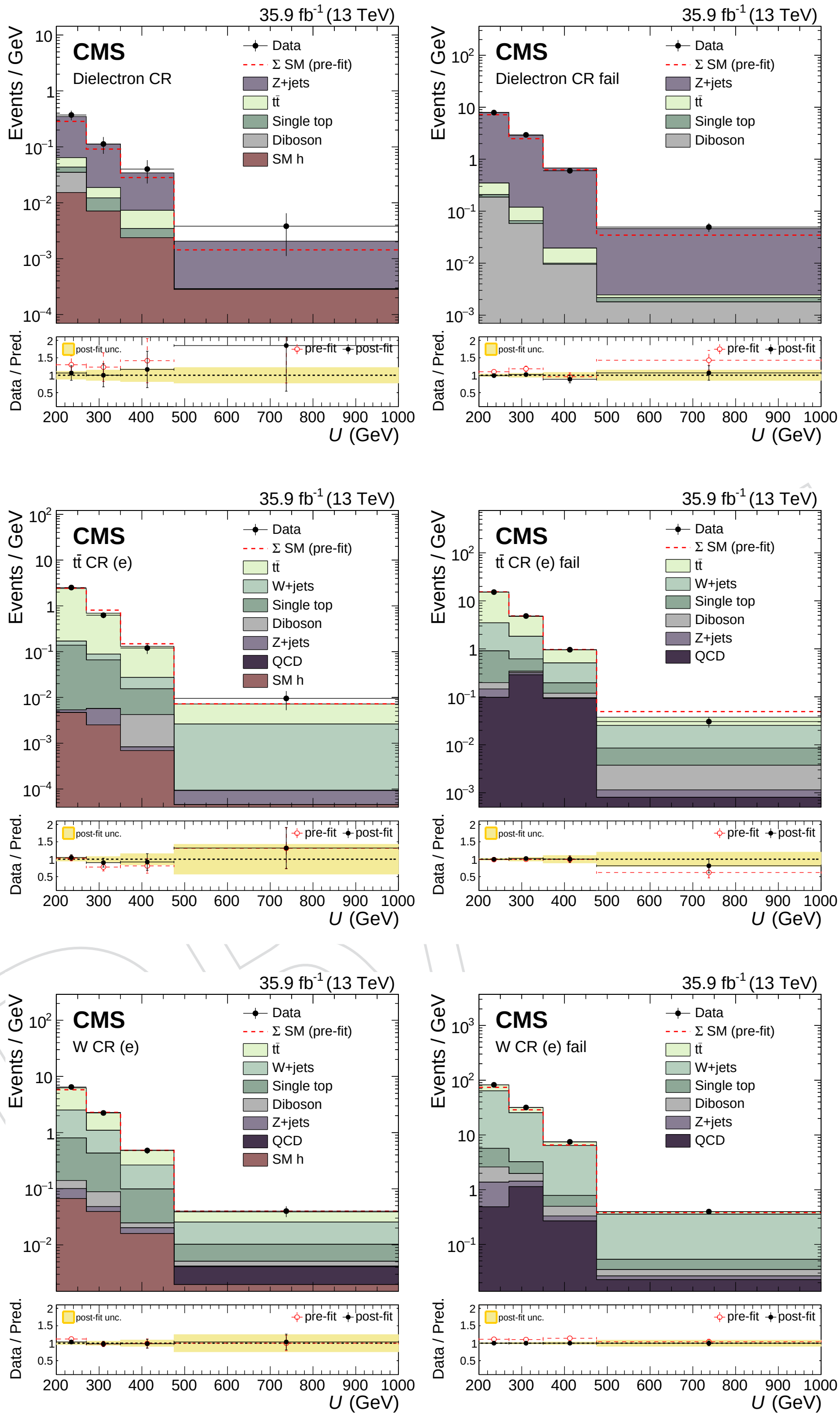

Figure 45: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

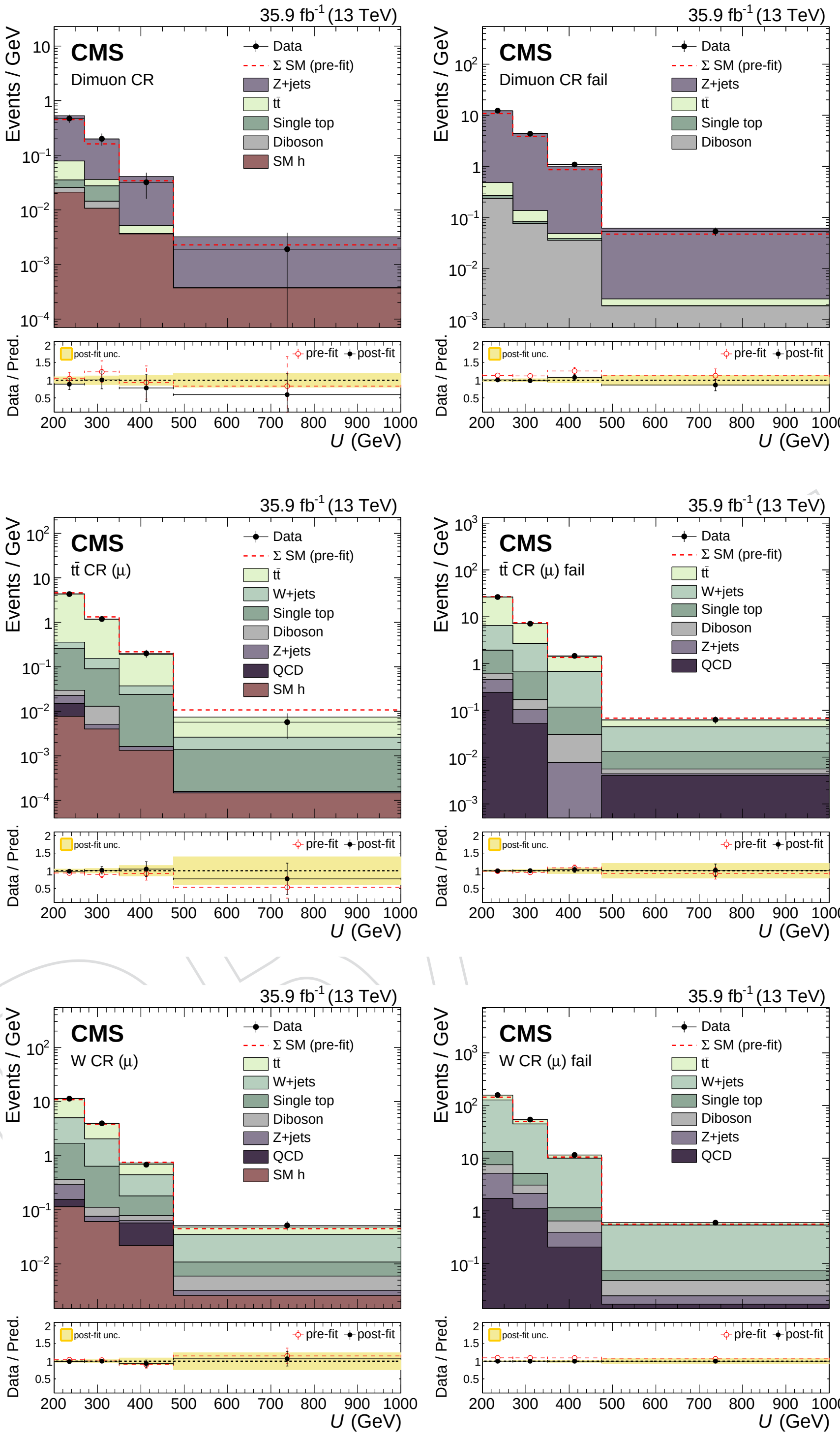

Figure 46: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

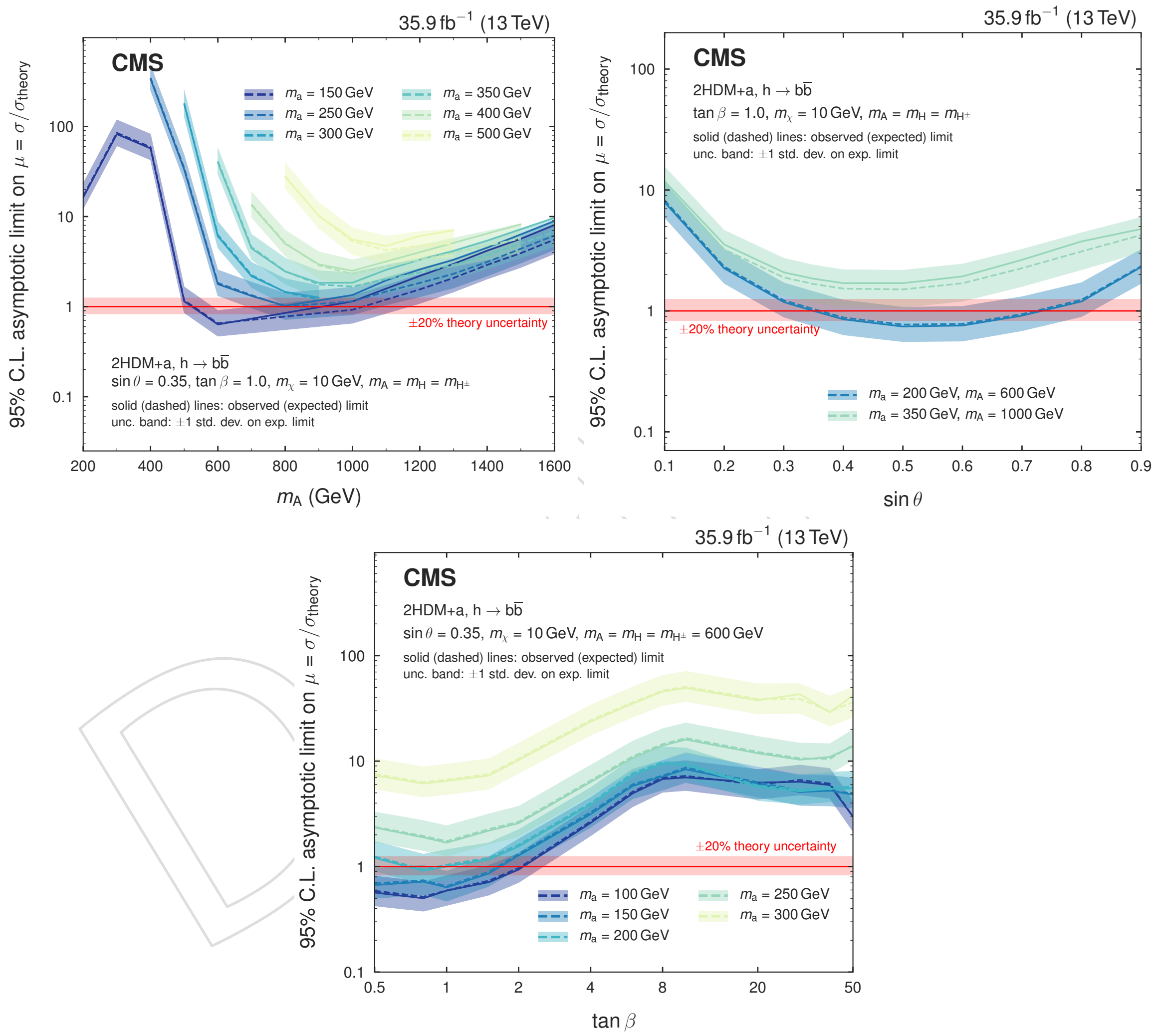

Figure 47: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

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.

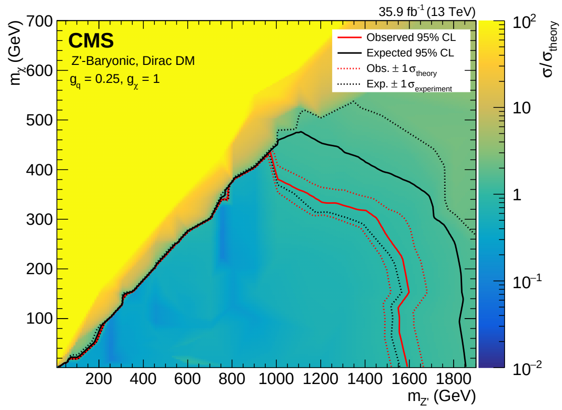

Figure 48: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

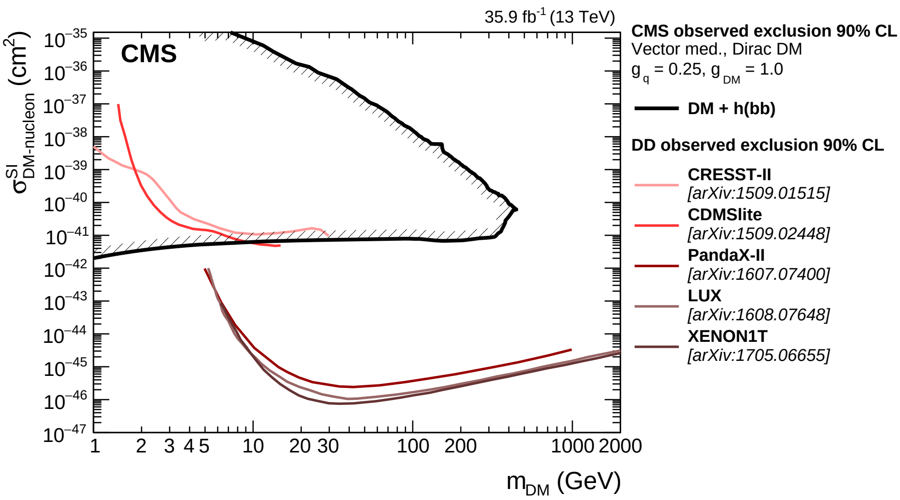

Figure 49: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits 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:

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

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 $m_\chi < 5$ 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 choice of the mixing angle $\sin\theta$ and $\tan\beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900\,\mathrm{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150\,\mathrm{GeV}$. Scanning over $\sin\theta$ with $\tan\beta = 1$, we exclude $0.35 < \sin\theta < 0.75$ for $m_\mathrm{A} = 600\,\mathrm{GeV}$ and $m_\mathrm{a} = 200\,\mathrm{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_\mathrm{A} = 600\,\mathrm{GeV}$ and $m_\mathrm{a} = 100$ (150) GeV and $\sin\theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_\chi = 10\,\mathrm{GeV}$ 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 $m_\chi < 5\,\mathrm{GeV}$ 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

[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.

[393] 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.

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

[395] 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.

[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.

[401] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) 08004, doi:10.1088/1748-0221/3/08/S08004.

[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 $\mathrm{b\bar{b}}$ 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.

[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.

[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.

[424] 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.

[425] 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.

[426] 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.

[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.

[430] 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.

[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 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.

[436] 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.

[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 $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **741** (2015) 12, doi:10.1016/j.physletb.2014.12.003, arXiv:1406.7533.

[440] 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.

[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 $\sqrt{s} = 7$ TeV", *JHEP* **06** (2014) 120, doi:10.1007/JHEP06(2014)120, arXiv:1402.1521.

[443] 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.

[444] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\to \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **763** (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

[445] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Phys. Lett. B* **766** (2017) 268, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.

[446] 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.

[447] 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.

[448] 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.

[449] 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.

[450] 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.

[451] 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.

[452] 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.

[453] 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.

[454] 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.

[455] 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

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 $\mathrm{h} \to \mathrm{b\bar{b}}$ 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, 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 $A$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 $\mathrm{H}^{\pm}$ 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 $\lambda_3 = \lambda_{\mathrm{P1}} = \lambda_{\mathrm{P2}} = 3$ [467]. Masses of the

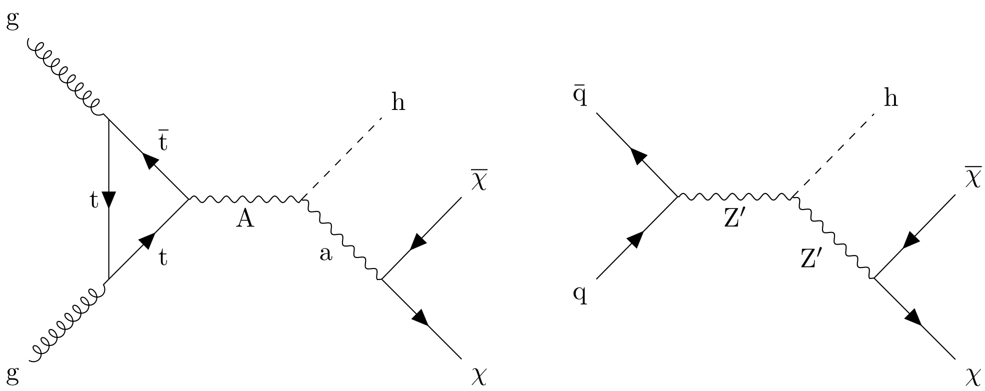

Figure 50: Feynman diagrams for the 2HDM+a model (left) and the baryonic Z′ model (right).

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., $m_{\mathrm{H}} = m_{\mathrm{H}^{\pm}} = m_{\mathrm{A}}$. The DM particle χ is assumed to have a mass of $m_{\chi} = 10\,\mathrm{GeV}$.

The baryonic Z′ model [465] is an extension of the SM with an additional $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ 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 $\mathrm{U}(1)_B$ symmetry. Analogous to the SM, there remains a physical baryonic Higgs particle, h_B, with a coupling of $\mathrm{h}_B\mathrm{Z'Z'}$ and vacuum expectation value of v_B. The Z′ and SM Higgs boson, h, interact with a coupling strength of $g_{\mathrm{hZ'Z'}} = m_{\mathrm{Z'}}^2 \sin\theta / v_B$, 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 $g_{\mathrm{q}}, g_{\chi} \sim 4\pi$ 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 $g_{\mathrm{hZ'Z'}}/m_{\mathrm{Z'}} = 1$, which implies $v_B = m_{\mathrm{Z'}} \sin\theta$. This choice maximizes the cross section without violating the bounds. The free parameters in the model under these assumptions are thus $m_{\mathrm{Z'}}$ 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 μs, 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\sum_i p_{\mathrm{T}}^{i2}$, 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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 $\mathrm{b\bar{b}}$ 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_{SD} for the background processes. In this search, the value of m_{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_{T}. As explained in Ref. [494], the stability of N_2 is tested against the variable $\rho = \ln(m_{\mathrm{SD}}^2/p_{\mathrm{T}}^2)$: since the jet mass distribution for QCD multijet events is expected to scale with p_{T}, decorrelating the N_2 variable as a function of ρ and p_{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^{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 $p_{\mathrm{T}} > 18$ 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 $p_{\mathrm{T}} > 15$ 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

Signal events are characterized by a high value of $p_{\mathrm{T}}^{\mathrm{miss}}$, 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 $\mathrm{t\bar{t}}$ production. To predict the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra of these processes in the SR, data from different control regions (CRs) are used. Single-lepton CRs are designed to predict the $\mathrm{t\bar{t}}$ and W+jets backgrounds, while dilepton CRs predict the Z+jets background contribution. The hadronic recoil, U, defined by removing the p_{T} of the lepton(s) from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation in CRs, is used as a proxy for the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$, where $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ ($H_{\mathrm{T}}^{\mathrm{miss}}$) is the magnitude of the vectorial (scalar) sum of $\vec{p}_{\mathrm{T}}$ of all the particles (jets with $p_{\mathrm{T}} > 20$ GeV) at the trigger level. Muon candidates are excluded from the online $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ calculation. Thresholds on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are between 90 and 120 GeV, depending on the data-taking period. Collectively, online requirements on $p_{\mathrm{T,trig}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ are reffered to as $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.4$ is required. It is also required that $100 < m_{\mathrm{SD}} < 150$ GeV and $N_2^{\mathrm{DDT}} < 0$. In the SR (CRs), $p_{\mathrm{T}}^{\mathrm{miss}}$ (U) has to be larger than 200 GeV, and the minimum azimuthal angle ϕ between any AK4 jet and the direction of $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($\vec{U}$) 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, $t\bar{t}$, W+jets	0	0	pass
Single-lepton, b-tagged	$t\bar{t}$, W+jets	1	1	pass/fail
Single-lepton, anti-b-tagged	W+jets, $t\bar{t}$	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_T^{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_T^{miss} triggers that are used in the SR. These events are required to have two oppositely charged muons (having $p_T > 20\,\text{GeV}$ and $p_T > 10\,\text{GeV}$ 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_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 $t\bar{t}$ 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_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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 bℓ denotes the $\mathrm{t\bar{t}}$ process in the b-tagged single-lepton control sample that is used to estimate the $\mathrm{t\bar{t}}$ contribution in the SR, the expected number of $\mathrm{t\bar{t}}$ events, N_i, in the i^{th} bin of the SR is then given by $N_i = \mu_i^{\mathrm{t\bar{t}}}/T_i^{\mathrm{b}\ell}$, where $\mu_i^{\mathrm{t\bar{t}}}$ is a freely floating parameter included in the likelihood to scale the $\mathrm{t\bar{t}}$ contribution in bin i of U in the CR.

The transfer factors used to predict the W+jets and $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the CRs with no b-tagged AK4 jets and a double-b-tagged CA15 jet also have significant contributions from the $\mathrm{t\bar{t}}$ process, transfer factors to predict this contamination from $\mathrm{t\bar{t}}$ 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 $\mathrm{t\bar{t}}$ 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 $\mathrm{Z} \to \nu\nu$ and the $\mathrm{Z} \to \ell\ell$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ (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.* $p_{\mathrm{T}}^{\mathrm{miss}}/U$. 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 $\mathrm{t\bar{t}}$ processes.

For the signal and the SM h processes, an uncertainty in the double-b tagging efficiency is applied that depends on the p_{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 $\mathrm{g} \to \mathrm{b\bar{b}}$ 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 $m_{\mathrm{A}} = 1.1\,\mathrm{TeV}$ and $m_{\mathrm{a}} = 150\,\mathrm{GeV}$ (with $\sin\theta = 0.35$ and $\tan\beta = 1$).

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

ture variable N_2^{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 $\mathrm{t\bar{t}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 ($\mathrm{t\bar{t}}$, 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

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 $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, and by requiring that the minimum azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $\mu =$

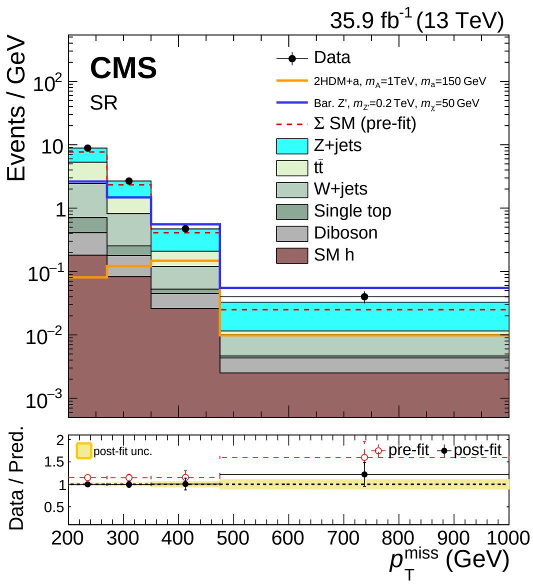

Figure 51: The p_T^{miss} distribution in the signal region before and after a likelihood fit. The data are in agreement with post-fit background predictions for the SM backgrounds, and no significant excess is observed. The dashed red histogram corresponds to the pre-fit estimate for the SM backgrounds.

Table 24: Post-fit event yield expectations per p_T^{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
tt̄	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, $m_A = 1$ TeV, $m_a = 150$ GeV	5.7 ± 0.6	9.8 ± 1.1	18.5 ± 2.1	5.2 ± 0.6
Bar. Z', $m_{Z'} = 0.2$ TeV, $m_\chi = 50$ GeV	184.2 ± 20.0	118.1 ± 12.8	69.5 ± 7.7	28.9 ± 3.3

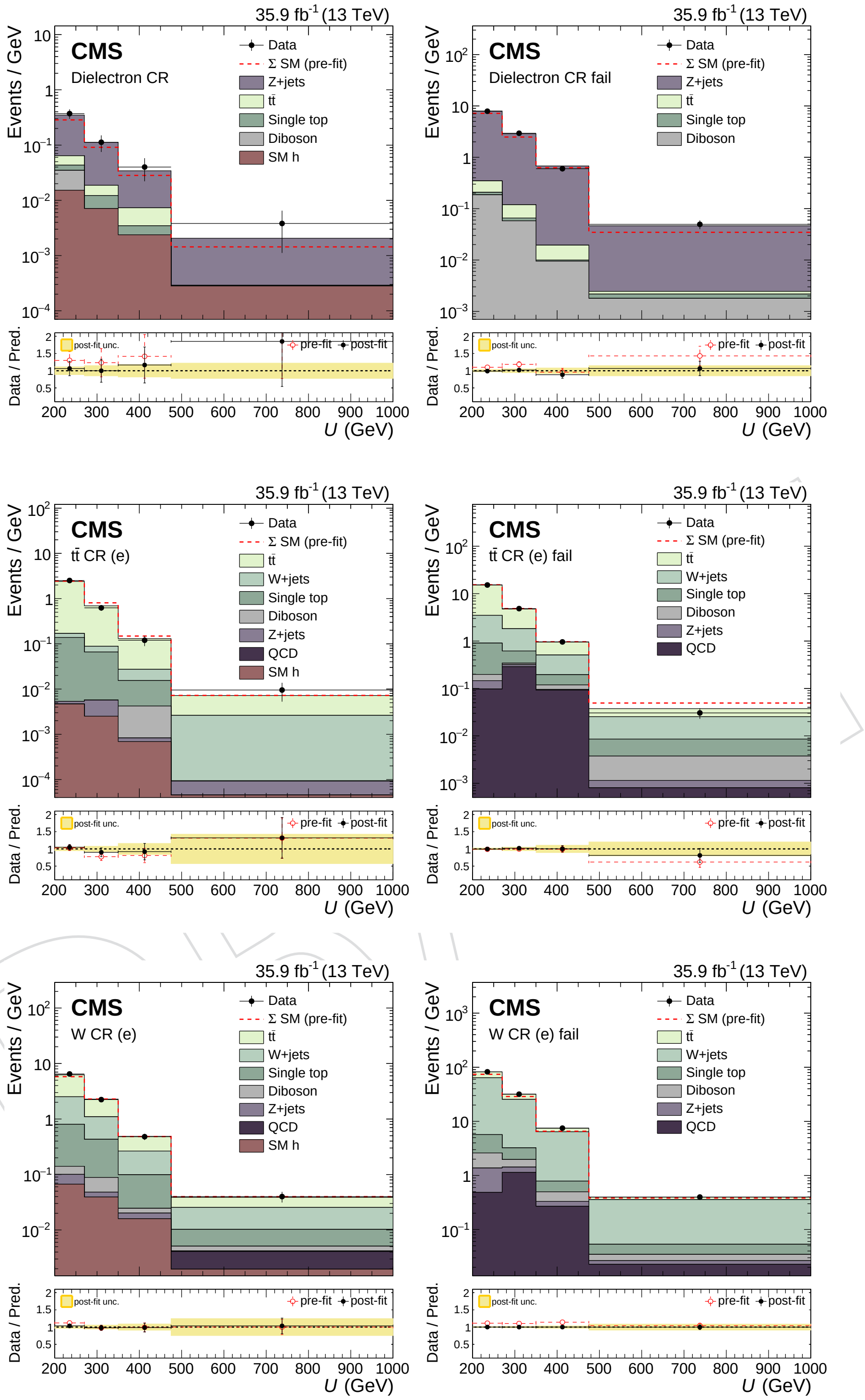

Figure 52: The U distribution in the electron control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the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

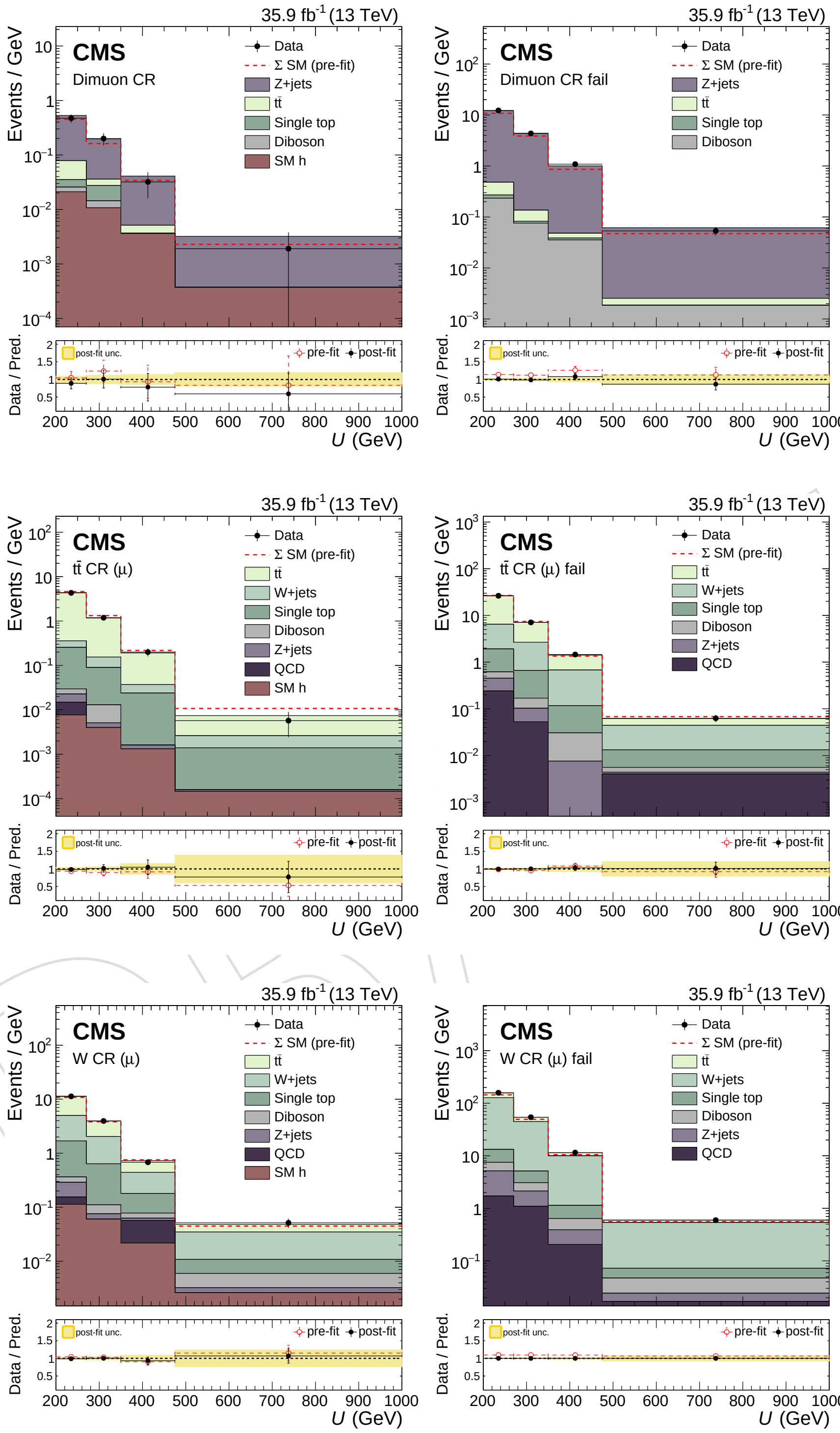

Figure 53: The U distribution in the muon control regions before and after a background-only fit to data, including the data in the signal region in the likelihood. For the distributions on the left the CA15 jet passes the double-b tag requirement and for the distributions on the right it fails the double-b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

$\sigma/\sigma_{\text{theory}}$, 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 $m_\text{a} = 150\,\text{GeV}$, $\sin\theta = 0.35$ and $\tan\beta = 1$. Mixing angles with $0.35 < \sin\theta < 0.75$ are excluded for $m_\text{A} = 600\,\text{GeV}$ and $m_\text{a} = 200\,\text{GeV}$, assuming $\tan\beta = 1$. Also excluded are $\tan\beta$ values between 0.5 and 2.0 (1.6) for $m_\text{a} = 100$ (150) GeV and $m_\text{A} = 600\,\text{GeV}$, given $\sin\theta = 0.35$. These are the first experimental limits on the 2HDM+a model.

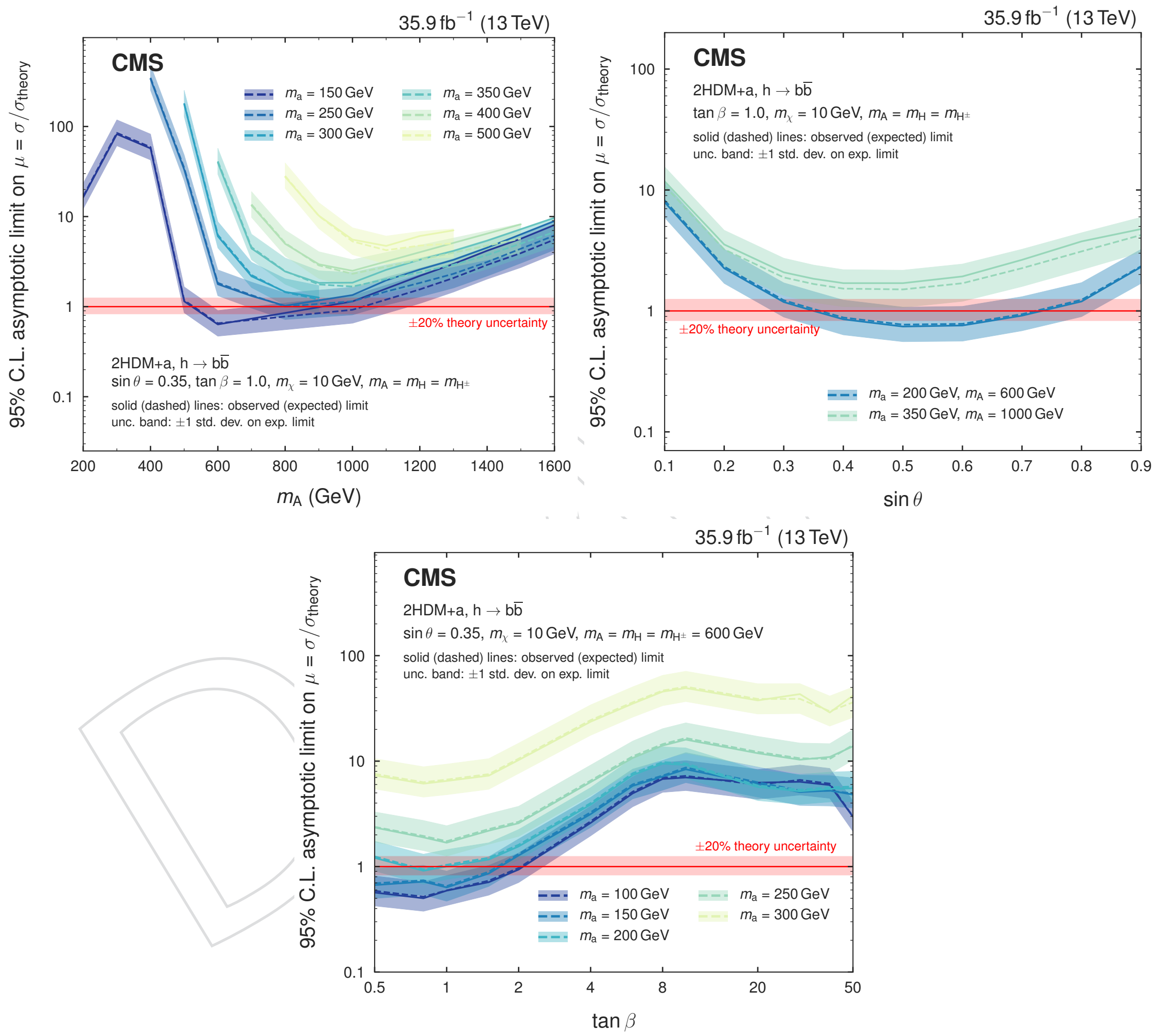

Figure 54: Upper limits on the signal strength modifier for the 2HDM+a model when scanning m_A and m_a (upper left), the mixing angle θ (upper right), or $\tan\beta$ (lower).

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 $m_{Z'} < 1.6\,\text{TeV}$ 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.

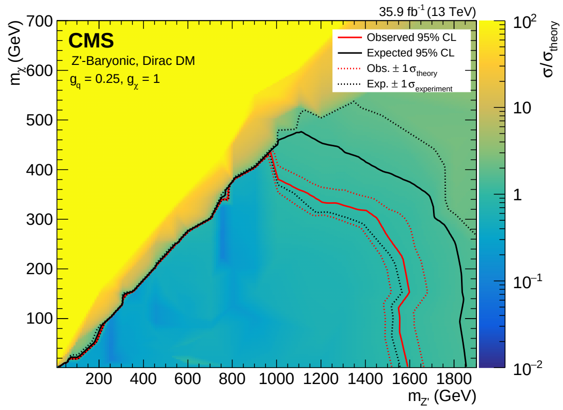

Figure 55: Upper limits on the signal strength modifier for the baryonic Z′ model as a function of $m_{Z'}$ and m_χ. Mediators of up to 1.6 TeV are excluded for a DM mass of 1 GeV. Masses of the DM particle itself are excluded up to 430 GeV for a Z′ mass of 1.25 TeV.

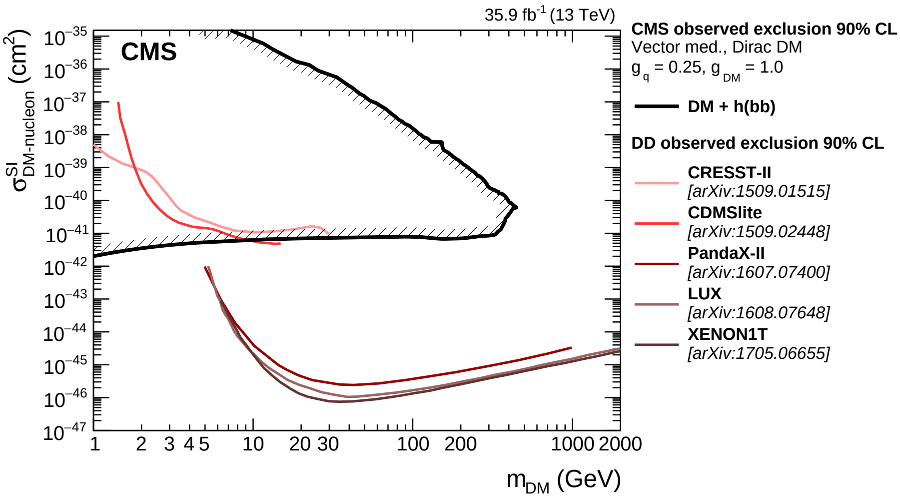

Figure 56: The 90% CL exclusion limits on the DM-nucleon SI scattering cross section as a function of m_χ. Results on the baryonic Z′ model obtained in this analysis are compared with those from a selection of direct detection (DD) experiments. The latter exclude the regions above the curves. Limits 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:

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

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 $m_\chi < 5$ 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 choice of the mixing angle $\sin\theta$ and $\tan\beta$ in the 2HDM+a model, the search excludes masses $500 < m_A < 900\,\mathrm{GeV}$ (where A is the heavy pseudoscalar boson) assuming $m_a = 150\,\mathrm{GeV}$. Scanning over $\sin\theta$ with $\tan\beta = 1$, we exclude $0.35 < \sin\theta < 0.75$ for $m_\mathrm{A} = 600\,\mathrm{GeV}$ and $m_\mathrm{a} = 200\,\mathrm{GeV}$. Finally, $\tan\beta$ values between 0.5 and 2.0 (1.6) are excluded for $m_\mathrm{A} = 600\,\mathrm{GeV}$ and $m_\mathrm{a} = 100$ (150) GeV and $\sin\theta > 0.35$. In all 2HDM+a interpretations, a DM mass of $m_\chi = 10\,\mathrm{GeV}$ 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 $m_\chi < 5\,\mathrm{GeV}$ 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

[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`.

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

[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`.

[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`.

[463] 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`.

[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`.

[465] 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`.

[466] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) 08004, `doi:10.1088/1748-0221/3/08/S08004`.

[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 $\mathrm{b\bar{b}}$ or $\gamma\gamma$ at $\sqrt{s} = 13$ TeV", *JHEP* **10** (2017) 180, doi:10.1007/JHEP10(2017)180, arXiv:1703.05236.

[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.

[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.

[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.

[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.

[482] 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.

[483] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[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.

[486] G. Collaboration, "Geant4 – a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/10.1016/S0168-9002(03)01368-8.

[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.

[490] 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.

[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.

[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 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.

[502] 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.

[503] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period”, CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

[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* **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 $\sqrt{s} = 7\,\mathrm{TeV}$”, *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.

[507] CMS Collaboration, “Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$”, *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 $\sqrt{s} = 8\,\mathrm{TeV}$”, *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 $\sqrt{s} = 13\,\mathrm{TeV}$”, *Phys. Lett. B* **763** (2016) 280, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

[510] 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.

[511] 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.

[512] 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.

[513] 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.

[514] 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.

[515] 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.

[516] 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.

[517] 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.

[518] 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.

[519] 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.

[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.