

DRAFT

CMS Paper

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

2022/02/18

Archive Hash: 45c6d29-D

Archive Date: 2022/02/18

Search for an invisibly decaying Higgs boson produced in association with a top-quark pair or a vector boson in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search for an invisibly decaying Higgs boson, produced in association with a top-quark pair or a vector boson, has been performed using 138 fb^{-1} of proton-proton collision data collected at $\sqrt{s} = 13$ TeV by the CMS experiment at the LHC. Events are categorized based on jet multiplicity, the number of jets stemming from b-quark decays as well as the number of boosted top and W candidates that have been reconstructed as a single jet. The observed (expected) limit set on the invisible branching fraction of the 125 GeV Higgs boson is XX% (36%), assuming Standard Model production cross sections. Interpretations of this analysis in the context of simplified dark matter models with a (pseudo-)scalar boson mediator are also presented. Finally, the results of this analysis are combined with previous CMS measurements of the Higgs to invisible branching fraction in complementary production and decay channels. The combined limit on the invisible branching fraction is XX% (YY%).

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

PDFAuthor: The CMS Collaboration pdftitle=Search for invisible Higgs in ttH/VjjH production
PDFTitle: "" PDFSubject: CMS
PDFKeywords: CMS, Higgs, invisible, ttH, VH

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

1 Introduction

A Higgs boson of mass 125 GeV was discovered by the CMS and ATLAS Collaborations in 2012 [1–3]. Using LHC proton-proton collision data collected at 7, 8 and 13 TeV center-of-mass energies, several of its production mechanisms and decays have since been observed, and some of its properties measured [4, 5].

In the Standard Model (SM), the Higgs boson branching fraction to an invisible final state ($\mathcal{B}(H \rightarrow \text{inv.})$) is about 0.1% [6]. Several Beyond the Standard Model (BSM) theories predict larger branching fraction to invisible final states [7–9]. For example, an hypothesis is that the Higgs boson could connect the SM sector and dark matter [10–15]. In this scenario, called the “Higgs portal”, $\mathcal{B}(H \rightarrow \text{inv.})$ is enhanced as the Higgs boson decays to a pair of dark matter particles. In other simplified models [16], dark matter is produced in the decay of a scalar or pseudoscalar boson.

The most stringent indirect constraints on invisible decays of the Higgs boson have been set by an ATLAS and CMS combined measurement of its coupling to SM particles, excluding $\mathcal{B}(H \rightarrow \text{inv.}) > 0.34$ at 95% confidence level (CL) [17]. Direct searches for invisible decays of the 125 GeV Higgs boson have also been performed by the two collaborations, in channels where the Higgs boson is produced through vector boson fusion (VBF), gluon-gluon fusion (ggF), associate production with a vector boson (VH) or with a pair of top quarks (ttH). The most recent results can be found in Refs. [18? –23] (temporary citations). For CMS, the most stringent limit is obtained with the VBF search, using 138 fb^{-1} of data collected at 13 TeV center-of-mass energy, excluding $\mathcal{B}(H \rightarrow \text{inv.}) > 0.17$ (0.11 expected) at 95% C.L. [19] (temporary citation). ATLAS sets its best limit of 0.26 (0.17 expected) through a combination of VBF and ttH searches at 7, 8 and 13 TeV [18]. A recent search in the ggF and VH channels by CMS, using 137 fb^{-1} of data collected at 13 TeV center-of-mass energy, yields an exclusion limit of $\mathcal{B}(H \rightarrow \text{inv.}) < 0.28$ (0.25 expected) [23] (temporary citation).

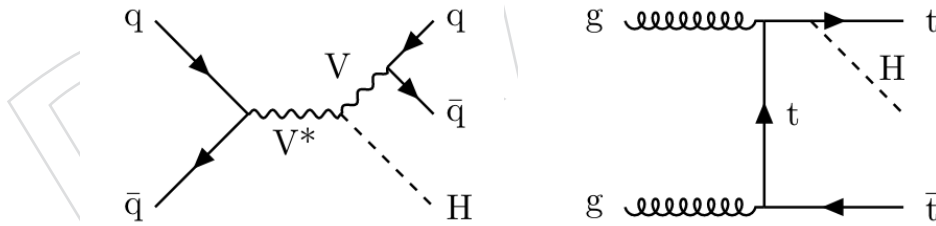


Figure 1: Feynman diagrams for the Standard Model Higgs boson production channels VH and ttH.

In this paper, a search for an invisibly decaying Higgs boson, produced in association with a top-quark pair or a vector boson is performed. The topologies for VH and ttH are presented in Figure 1. LHC proton-proton collision data from 2016, 2017 and 2018, corresponding to a total of 137 fb^{-1} collected at 13 TeV center-of-mass energy, are jointly analyzed. The search focuses on three types of hadronic final states: those with boosted top and/or boosted vector bosons reconstructed with dedicated boosted jet algorithms, accompanied by b-jet(s); those with 5 or 6 jets and b-jet(s); and those with 2 jets with a mass compatible with that of a W or Z boson. The results are also interpreted in the context of a simplified model where the dark matter is produced through the decay a scalar (ϕ) or pseudoscalar (a) boson produced in association with a pair of top quark, a vector boson, or a jet.

Missing transverse momentum is used as the discriminating variable to separate the signal from its backgrounds, which mainly consist of invisible Z-boson decays ($Z \rightarrow \nu\bar{\nu}$) and events from $t\bar{t}$ +jets and W +jets production processes where one or more leptons are mis-reconstructed or outside of the detector acceptance, thus creating fake missing transverse momentum. Control regions, in which one lepton, one photon, or two same-flavor leptons are required, are used to constrain these backgrounds. As such, the analysis is binned in “recoil”, which is defined as the magnitude of the missing transverse momentum in an event, with contributions from leptons and photons added to ensure good correspondence between signal regions and control regions.

This paper is organized as follows. Section 2 is a brief description of the CMS detector. The simulated samples used in this analysis are summarized in Section 3. Section 4 describes the event reconstruction and objects used in this analysis, while the event selection and event categorization are detailed in Section 5. The data control regions used for estimating the SM backgrounds are introduced in Section 6. Section 7 describes the statistical procedure used to constrain the backgrounds and extract the signal. The results of this search are presented in Section 8. The results of the combination with other complementary CMS searches for invisibly decaying Higgs bosons are described in Section 9, and a conclusion is proposed in Section 10.

2 The CMS detector

A detailed description of the Compact Muon Solenoid (CMS) experiment can be found in [24].

The central feature of the CMS apparatus is a superconducting solenoid, of 6 m internal diameter. Within the field volume are the silicon pixel and strip tracker, the crystal electromagnetic calorimeter (ECAL) and the brass-scintillator hadronic calorimeter (HCAL). Muons are measured in gas chambers embedded in the iron return yoke. Besides the barrel and endcap detectors, CMS has extensive forward calorimetry.

The silicon trackers consist of 1856 silicon pixel and 15148 silicon strip detector modules and are capable of measuring charged particles within the pseudorapidity acceptance range of $|\eta| < 2.5$.

The ECAL has an energy resolution of better than 0.5% above 100 GeV. The HCAL, when combined with the ECAL, measures jets with a resolution $\Delta E/E \approx 100\%/\sqrt{E} \oplus 5\%$. The calorimeter cells are grouped in projective towers, of granularity $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ at central rapidities and 0.175×0.175 at forward rapidities. Both the ECAL and HCAL are comprised of a ‘barrel’ and two ‘encaps’; there are often abbreviated as EB, EE, HB and HE respectively.

The hadron forward (HF) calorimeter uses steel as an absorber and quartz fibers as the sensitive material. The two halves of the HF are located 11.2m from the interaction region, one on each end, and together they provide coverage in the range $3.0 < |\eta| < 5.2$.

The first level (L1) of the CMS trigger system, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events, at a rate of roughly 100kHz. The High Level Trigger (HLT) processor farm performs event reconstruction similar to that of the full CMS reconstruction, but optimised for efficiency. This decreases the event rate from $\sim 100\text{kHz}$ to $\sim 1\text{kHz}$, before data storage.

The procedure for calculating the integrated luminosity recorded by the CMS detector for each year of data taking are documented in Ref. [25] (2016), Ref. [26] (2017), and Ref. [27] (2018).

3 Simulated samples

Monte Carlo (MC) simulated events are used to model signal and background contributions in all analysis regions, other than for quantum chromodynamic (QCD) multi-jet production processes, which use a dedicated data-driven estimation as detailed in Section 6.4 with MC-based transfer factors. In all cases, MC samples used are produced using either POWHEG [28], MADGRAPH5amc@NLO [29] or MADGRAPH5 [30] matrix element generators and then matched to parton showering, hadronization and underlying event properties modelled by PYTHIA version 8.202 or later [31] and event tune CUETP8M1 for 2016 samples and CP5 for 2017 and 2018 [32]. The propagation of all final states through the CMS detector is simulated using GEANT4 [33] software. The set of parton distribution functions used in the production of all samples is the: NNPDF3.1 next-to-next-to-leading order (NNLO) [34].

For simulated events describing the Higgs boson decaying invisibly POWHEG v2.0 [35–38] is used to generate separate signal samples for different Higgs boson production modes: associate production of a Higgs with a vector boson (VH) or a $t\bar{t}$ pair (ttH), vector boson fusion (VBF) and gluon-gluon fusion (ggF). These samples require the SM Higgs boson to decay to a pair of neutrinos; their cross sections are normalised to the corresponding one in the SM, computed at NNLO and NLO accuracy in QCD and electroweak (EW), respectively.

Vector boson (V) + jets processes are generated at leading order accuracy (LO) in QCD using MADGRAPH5amc@NLO version 2.6.5 generator [29] using the MLM matching scheme between hard scatter and parton shower. The γ + jets processes are generated at NLO in QCD with MADGRAPH5amc@NLO version 2.4.2 [29]. Processes consisting of top, anti-top pairs are generated at next-to-leading order (NLO) with the FxFx matching scheme with POWHEG v2.0 [39] while single-top events are generated with POWHEG v1.0 [40].

Diboson ZZ and WZ production are generated at LO using PYTHIA, while the WW process is simulated at NLO in QCD using the POWHEG generator [41]. QCD multijet samples are generated at LO using MADGRAPH5amc.

ZH processes with $H \rightarrow b\bar{b}$ and the associated Z-boson decaying to $b\bar{b}$, $\ell\ell$ and qq , are considered. These processes are generated at LO using the POWHEG generator.

Simplified Dark Matter model scenarios (DMsimp) are studied as an alternative interpretation of this analysis. Three DMsimp processes are explored in this analysis: dark matter production in association with a top quark pair ($t\bar{t}$ + DM), and dark matter production in association either with a single boosted jet or with a single boosted vector boson (collectively referred to as Mono-X + DM). Signal samples of the $t\bar{t}$ + DM model are generated using MADGRAPHv2.4.2 at LO with at most one additional parton in the matrix element. Cross-sections for the $t\bar{t}$ + DM samples are calculated at LO with MADGRAPH5. Mono-X + DM processes are generated with MADGRAPH5amc@NLO at effective LO. Cross-sections are calculated using MADGRAPH. Only processes with scalar or pseudoscalar mediators are considered. Events with a pair of DM particles and additional partons are generated at NLO in QCD using FxFx jet merging. The coupling constants between mediator and quark as well as mediator and DM particle are set to $g_q = 0.25$ and $g_\chi = 1.0$ as recommended by the LHC Dark Matter Working Group [42]. Signal samples are generated with coupling constants $g_q = 1.0$ and $g_\chi = 1.0$, and are studied over a range of parameter points for mediator mass M_ϕ and dark matter particle mass M_χ , with values and cross-sections shown in Tables 1 and 2. These processes are identical to SM $H \rightarrow inv$ processes, but with the Higgs boson replaced by a massive scalar or pseudoscalar mediator decaying to a dark matter pair.

M_χ	$M_{\phi(a)}$	σ_ϕ (pb)	σ_a
1	50	3.5080	0.3675
1	100	0.7906	0.3386
1	150	0.2561	0.1483
1	200	0.1074	0.0994
1	250	0.0563	0.0686
1	300	0.0346	0.0486
1	350	0.0235	0.0280
1	400	0.0144	0.0145
1	450	0.0093	0.0096
1	500	0.0064	0.0067

Table 1: Scalar (ϕ) and pseudoscalar (a) mediator mass parameter points and cross-sections for $t\bar{t}$ + DM samples at fixed m_χ . Masses are in units of GeV and cross sections in picobarn.

$M_{\phi(a)}$	σ_ϕ		σ_a	
	MonoV	MonoJ	MonoV	MonoJ
10	2.428 e-2	2.183	2.749e-2	4.941
50	2.393e-2	1.901	2.709e2	4.316
100	2.184e-2	1.464	2.463e-2	3.366
200	1.594e-2	0.842	1.778e-2	2.062
300	1.004e-2	0.577	1.110e-2	1.721
350	7.716e-3	0.573	6.630e-3	1.932
400	4.866e-3	0.460	3.459e-3	0.932
450	3.089e-3	0.311	2.278e-3	0.528
500	2.054e-3	0.206	1.592e-4	0.321
600	2.054e-3	0.094	8.472e-4	0.134
700	1.012e-3	0.094	8.472e-4	0.063
800	3.162e-4	0.024	2.881e-4	0.032

Table 2: Scalar (ϕ) and pseudoscalar (a) mediator mass parameter points and cross-sections for Mono-X + DM samples. Masses are in units of GeV and cross sections in picobarn.

4 Event reconstruction

4.1 Leptons and photons

Objects in this analysis are reconstructed using the *Particle-Flow* (PF) algorithm [43], which combines information from the tracker, calorimeters and muon systems to reconstruct individual objects in each event. Pileup mitigation techniques [44] are employed to filter energy deposits belonging to pileup vertices and remove objects not associated with the primary interaction vertex. For a fully hadronic analysis such as this, the final state is required to contain purely jets from quarks, gluons and recoil, with a veto on all other objects such as leptons and photons.

Candidate leptons are selected with $p_T > 10$ GeV and pseudorapidity range $|\eta| < 2.4$ for muons [45], $p_T > 10$ GeV and $|\eta| < 2.5$ for electrons [46], $|\eta| < 2.3$ and $p_T > 20$ GeV for taus, and $|\eta| < 2.5$ and $p_T > 15$ GeV for photons [47]. Object definitions are consistent across all years, and are divided into *tight*, *loose* and *veto* collections based on the working points used for identification and isolation of the candidates. For event vetoes and cleaning collections, *loose* and *veto* collections are defined using working points with efficiencies of $\simeq 90\%$, $\simeq 95\%$ and $> 99\%$ for ‘loose’ muons, ‘veto’ electrons and ‘loose’ photons, respectively. These allow for high efficiency in selecting genuine leptons and photons.

4.2 Jets

Jets are reconstructed by clustering all PF candidates with the primary interaction vertex with the *anti- k_T* jet clustering algorithm [48], using a distance parameter $\Delta R < 0.4$ (AK4), and applying Jet Energy Corrections (JEC) and Charged Hadron Subtraction (CHS) treatments [49]. JEC are derived from simulation, as functions of p_T and η , and smear the MC jet energy resolution to match that of the data. The jets are cleaned of veto electron and loose muon and photon collections, within $\Delta R < 0.4$.

For each year, a slightly different η -dependent selection is applied for data and MC. To avoid contributions from noisy jets, an additional cut is applied for 2017 only, removing jets with

$p_T < 50 \text{ GeV}$ and $2.65 < |\eta| < 3.139$. Additional analysis-specific treatment is applied to all jets. Jets must satisfy $p_T > 30$, and forward jet treatment is applied. Candidate b-jets are tagged using the *DeepCSV* algorithm with a medium working point [50] and require $p_T > 20 \text{ GeV}$.

4.3 Boosted objects

The *DeepAK8* algorithm [51] is used to classify jets that originate from heavy particles, and assists in the categorisation of boosted ttH. ‘Tight’ AK8 jets are classified as originating from either top quarks or vector bosons. Jets from top quarks are selected using: $p_T > 400 \text{ GeV}$ and $120 < m_{SD} < 210 \text{ GeV}$, with tight working point (1% mis-tag rate), and jets from vector bosons are selected with: $p_T > 200 \text{ GeV}$ and $65 < m_{SD} < 120 \text{ GeV}$ with tight working point (1% mis-tag rate). Orthogonality is ensured between the two classifications through the softdrop mass m_{SD} [52] requirements.

4.4 Recoil

Recoil is used as the binning variable for the fit to estimate the upper limit of the $H \rightarrow inv$ branching fraction. The recoil is here defined as the missing transverse energy of an event with leptonic and photonic components added, to ensure correspondence between the leptonic and hadronic regions. ‘Type-1 corrections’ based on AK4 PF CHS jets are applied to the recoil, which propagates JEC by subtracting from the MET the vector sum of jet p_T and adding the vector sum of calibrated jet p_T .

4.5 Variables

A set of cuts are applied to regions and categories as described in Section 5. Variables used include the scalar sum of all jet p_T in an event, H_T , the missing H_T in an event, H_T^{miss} , the largest jet p_T in an event p_{T,j_1} , and the azimuthal separation between the recoil and H_T^{miss} , $\Delta\phi(\text{recoil}, H_T^{miss})$.

Some angular variables are also used to suppress QCD multijet backgrounds in certain region. $\Delta\phi_{\min}(j_{1,2,3,4}, p_T^{miss})$ is defined as the minimum azimuthal angle between p_T^{miss} and any of the four jets with highest p_T . The angular variable $\tilde{\omega}_{\min}$ is designed to suppress events where missing energy is the result of a jet p_T mismeasurement, as is the case in QCD multijet events [53]. $\tilde{\omega}_{\min}$ minimises H_T^{miss} by rescaling the p_T of jet i . $\tan(\tilde{\omega})$ is defined as the ratio of the reduced value of $H_{T,\min}^{miss}$ relative to the initial jet $p_{T,i}$, i.e. $\tilde{\omega} = \arctan(H_{T,\min}^{miss}/p_{T,i})$. This angle is then minimised by testing against all jets i in the event, and goes to zero if H_T^{miss} can be made to disappear by rescaling the measured p_T of a jet.

5 Event selection and categorization

5.1 Baseline selection

This analysis makes use of seven ‘regions’. Backgrounds in the signal region (SR), detailed in Section 5.4 are weighted by events in five control regions (CRs) detailed in Sections 6.1 to 6.3, four of which are leptonic and one photonic. An additional hadronic CR for QCD multijet estimation is described in Section 6.4. All of these regions share a common baseline selection, with relevant E_T^{miss} filters applied to all data and MC where applicable. This common selection is shown in Table 3.

A ‘‘VBF veto’’ is enforced to ensure orthogonality with the VBF phase space, through an inversion of the kinematic selection employed by [54] and a veto on events with leading or sublead-

ing jets with $|\eta| > 2.4$. Meanwhile, orthogonality to leptonic ttH decays is ensured through a cut on $m_T^\ell < 110$ GeV in the single lepton CRs, and $m_{\ell\ell} < 120$ GeV in the dilepton CRs. This also suppresses ttH signal in the control regions.

Table 3: Common selection applied to all categories and regions in this analysis.

Variable	Cut Value	Purpose
recoil	≥ 200 GeV	Signal purity
H_T	≥ 200 GeV	Signal purity
H_T^{miss}	≥ 200 GeV	Signal purity
p_{T,j_1}	> 80 GeV	Signal purity
$\frac{M_{HT}}{\text{recoil}}$	< 1.2	Event quality
$\Delta\phi(\text{recoil}, H_T^{\text{miss}})$	< 0.5 rad	Event quality
$ \eta_1 , \eta_2 $	< 2.4	Analysis orthogonalization

During significant portions of data taking in 2018, areas the of the HCAL Endcap calorimeter were nonfunctional, causing no deposits to be recorded in the HCAL regions $-1.57 < \phi < -0.87$ rad and $-3.0 < \eta < -1.39$ during the affected period. This issue is known as the HE issue. Although this analysis is not especially sensitive to the HE issue due to the forward jet requirements and VBF orthogonalization, mitigation is still applied: data events from the affected period in the electron CRs with $-1.8 < E_T^{\text{miss}}(\phi) < 0.6$ are vetoed if they contain jets within the affected region, and in the hadronic regions, events are vetoed if they have $-1.8 < \phi(E_T^{\text{miss}}) < -0.6$. To ensure good correspondence between data and MC, the MC is reweighted according to the recalculated effective luminosity.

5.2 Trigger requirements

Trigger selections vary dependant on year, region, and whether an event is data or MC.

In 2016, hadronic and muonic events are required to pass one of four high-level triggers based on $PF E_T^{\text{miss}}$ and $PF H_T^{\text{miss}}$, with thresholds ranging from 90 to 120 GeV. The same triggers are applied both to data and MC. Photonic events, for both data and MC, are required to pass one of two photon triggers with a p_T requirement of roughly 170 GeV.

Electron data events in 2016 can arise from either the *SingleElectron* or *SinglePhoton* dataset. To quality as a data event containing one or more electrons, the *SingleElectron* dataset is required to pass a single-electron trigger either with a p_T threshold of 27 GeV and tight identification requirements, or a p_T threshold of 105 GeV but looser identification requirements. Meanwhile, the *SinglePhoton* dataset is required to fail both electron triggers and pass one of two photon triggers with $p_T \simeq 170$ GeV. MC events in the electron control regions meanwhile are simply required to pass any of these four triggers.

Trigger requirements in 2017 are similar to that of 2016, but with slightly different kinematic thresholds. Hadronic and muonic events are required to pass one of two triggers with $PF E_T^{\text{miss}}$ and $PF H_T^{\text{miss}}$ thresholds of 120 GeV. Single-photon data events either in the photon control region or electron control regions are required to pass one photon trigger with $p_T = 200$ GeV. *SingleElectron* data events have p_T thresholds of 35 and 115 GeV, instead of 27 and 105 GeV in 2016. Trigger requirements are otherwise the same.

in 2018, hadronic and muonic events must pass exactly one trigger with tight ID requirements and $PF MET$ and $PF H_T^{\text{miss}}$ requirements of 120 GeV. For the electron and photon control regions, a dataset which combines the *SingleElectron* and *SinglePhoton* datasets is used. Data and MC in these regions require triggers analogous to the *SingleElectron* triggers in 2016 and 2017,

but with a lower p_T threshold of 32 GeV. Photon CR triggers are the same as in 2017.

5.3 Event Categorization

Events are categorized to target the ttH and resolved VH topologies. This is achieved using three classes: ttH boosted, if at least one boosted AK8 jet is reconstructed and passes boosted-t- and boosted-W-tag requirements; ttH resolved, with high AK4 multiplicity and b-tagged jets; and VH, with 2 jets with an invariant mass compatible with a W^+ , W^- or Z_0 boson decay. In each class, events are further subcategorised according to the jet (n_j), b-jet (n_b) and boosted AK8 jet (n_t/n_W) multiplicities, as well as subleading jet p_T , p_{T,j_2} , the dijet mass m_{jj} , and ttH versus $t\bar{t}$ + jets separation variables. The subcategory definitions are summarised in Table 4. The intended outcome of this process is a set of parameter spaces with high purity of a given production mode and minimal background contamination or signal cross-contamination.

Table 4: Categorisation of the ttH and VH production modes in the analysis.

Category	Subcategory	n_{jet}	n_b	n_t	n_V	p_{T,j_2}	Other
ttH Boosted	2Boosted1b	≥ 5	1	2			
	2Boosted2b	≥ 5	≥ 2	2			
	1t1b	≥ 5	1	1	0	≥ 80	-
	1t2b	≥ 5	≥ 2	1	0		
	1W1b	≥ 5	1	0	1		
	1W2b	≥ 5	≥ 2	0	1		
ttH Resolved	5j1b	5	1	0	0		$\Delta\phi(b_1, p_T^{miss}) > 1.0$
	6j1b	≥ 6	1	0	0	≥ 80	$\& \Delta\phi(j_1, p_T^{miss}) > \pi/2$
	5j2b	5	≥ 2	0	0		$\Delta\phi(b_1, p_T^{miss}) > 1.0$
	6j2b	≥ 6	≥ 2	0	0		$\& \Delta\phi(b_2, p_T^{miss}) > \pi/2$
VH	2j0b	2	0	0	0		
	2j1b	2	1	0	0	≥ 30	$m_{jj} \in [65, 120)$
	2j2b	2	2	0	0		

ttH is characterised by high jet multiplicity and b-tagged AK4 jets. The “2Boosted” subcategories are designed to capture all ttH events with two boosted t- or W-tagged jets, and are subcategorized by b-jet multiplicity. The remaining boosted categories target events with singly-boosted t or V-tagged jets. The remaining four ($n_{jet} \geq 5$ $n_{boost} = 0$) ttH subcategories target the resolved phase space.

The VH categories are characterised by two resolved AK4 jets with dijet mass around that of the associated vector boson, and sub-categorized through b-jet multiplicity.

5.4 Signal region

The signal region (SR) is defined such that the signal-to-background ratio is enriched, QCD is suppressed, well-reconstructed events are selected and detector effects are mitigated.

To suppress events where p_T^{miss} is aligned with a jet, as in the case for QCD multijet events, we require $\Delta\phi_{min}(j_{1,2,3,4}, p_T^{miss}) > 0.5$. A cut of $\tilde{\omega}_{min} > 0.3$ is also applied in the SR to further suppress QCD multijet events.

The p_T^{miss} direction in ttH production is typically distributed about the direction of the Higgs boson. In $t\bar{t}$ events, the p_T^{miss} direction is confined parallel or anti-parallel to the leading b-jet direction, due to back-to-back top production. Therefore, the angles between the p_T^{miss} and

leading or sub-leading jet directions provide good variables for $t\bar{t}$ background suppression in ttH resolved categories. The angular variables $\Delta\phi(p_T^{miss}, b_1)$, $\Delta\phi(\text{recoil}, b_2)$ and $\Delta\phi(\text{recoil}, j_1)$ are most sensitive to ttH over $t\bar{t}$. These ttH resolved-specific cuts are summarised in Table 4.

QCD suppression cuts are applied to the SR only in ttH categories, but to both SR and CRs in VH categories in order to ensure good correspondence between regions. The cuts on $\tilde{\omega}_{\min}$ and $\Delta\phi_{\min}(j_{1,2,3,4}, p_T^{miss})$ are not applied to the CRs used for background estimation of the ttH categories, where p_T^{miss} does not stem from jet mismeasurement. This is done in order to increase event counts in the CRs, particularly in ttH boosted categories.

6 Data control regions used for SM background estimation

The analysis makes use of several data control regions to estimate the contributions of different standard model processes in the signal region. Where possible, the control regions have identical kinematic requirements to the signal region and leptons or a photon are used as a “tag” but otherwise ignored in the calculation of event observables. Single-electron and muon control regions are used to derive corrections to the MC expectation from lost lepton backgrounds, mainly consisting of $t\bar{t}$ + jets, single-top and W+jets events. Di-electron and muon as well as a photon control sample in the case of the VH categories are used to derive corrections to expected contribution of Z+jets (but also ZZ and ttZ) production, where the Z decays to a pair of neutrinos.

6.1 Single-lepton control regions

Two single-lepton CRs, the μ + jets CR and e + jets CR, are defined through the requirement of exactly one *tight* muon and one *tight* electron respectively. As detailed in Section 7, these CRs are used jointly to the SR to constrain the “lost lepton” background¹, ℓ_{lost} , which dominates ttH categories and contributes to VH categories.

For the μ + jets CR, one tightly-isolated muon with $p_T > 20$ GeV is required with a mass window $50 < m_T^\mu < 110$ GeV. For the e + jets CR, one tightly-isolated electron with $p_T > 40$ GeV is required with a mass window $50 < m_T^e < 110$ GeV. The full requirements for e + jets and μ + jets are shown in Table 5.

6.2 Double-lepton control regions

Two double-lepton CRs, the $\mu\mu$ + jets CR and ee + jets CR, are defined by requiring one *tight* and one *veto* (for electrons) or *loose* (for muons) lepton with opposite charge. These CRs are used jointly to the SR to constrain the $Z \rightarrow \nu\nu$ + jets background, which dominates VH categories and contributes to ttH categories especially at high recoil. The processes $Z \rightarrow \nu\nu$ and $Z \rightarrow \ell\ell$ are kinematically near-identical, largely due to lepton universality, hence the dilepton CRs are very good at constraining the $Z \rightarrow \nu\nu$ background.

For the ttH boosted categories, due to low statistics in the dilepton CRs, the individual $\mu\mu$ + jets and ee + jets CRs are summed together to form a single dilepton CR, and all ttH boosted subcategories are summed together. For the ttH resolved categories, the ee + jets and $\mu\mu$ + jets CRs remain separated, but the subcategories are again summed together for each CR. These higher statistic bins are then used as inputs for the fit to provide valuable constraints on the rate of $Z \rightarrow \nu\nu$ + jets.

¹which mainly consists in $W \rightarrow \ell\nu$ and leptonic top decays where one of the leptons fails reconstruction or is outside of acceptance.

For the $\mu\mu + \text{jets}$ CR, one tightly-isolated muon with $p_T > 20$ GeV and one loosely-isolated muon with $p_T > 10$ GeV and opposite charge is required, with a combined mass window $75 < m_{\mu\mu} < 105$ GeV for ttH and $60 < m_{\mu\mu} < 120$ GeV for VH. For the $ee + \text{jets}$ CR, one tightly-isolated electron with $p_T > 40$ GeV and one loosely-isolated electron with $p_T > 10$ GeV and opposite charge is required, with a combined mass window $75 < m_{ee} < 105$ GeV for ttH and $60 < m_{ee} < 120$ GeV for VH. A cut is required on the azimuthal angle between the recoil and $E_T^{\text{miss}_{\text{track}}}$, the missing transverse energy based on the measured track momentum. For well-defined events this should be consistent with recoil as measured in the calorimeters. Cleaning cuts of $\Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss}_{\text{track}}}) \geq \frac{\pi}{2}$ are performed to ensure good events are kept. The full requirements for $ee + \text{jets}$ and $\mu\mu + \text{jets}$ are shown in Table 5.

Table 5: A tabulated summary of all control region requirements, excluding QCD suppression cuts.

Region	Category	n_{object} reqs.	Mass requirements	p_T reqs.	Other reqs.
$\mu + \text{jets}$ control region	ttH VH	$n_\mu = 1$	$50 < m_T^\mu < 110$ GeV	$p_{T,\mu 1} > 20$	–
$\mu\mu + \text{jets}$ control region	ttH VH	$n_\mu = 2$	$75 < m_{\mu\mu} < 105$ GeV $60 < m_{\mu\mu} < 120$ GeV	$p_{T,\mu 1} > 20$ & $p_{T,\mu 2} > 10$	$\Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss}_{\text{track}}}) < \pi/2$
$e + \text{jets}$ control region	ttH VH	$n_e = 1$	$50 < m_T^e < 110$ GeV	$p_{T,e 1} > 40$	–
$ee + \text{jets}$ control region	ttH VH	$n_e = 2$	$75 < m_{ee} < 105$ GeV $60 < m_{ee} < 120$ GeV	$p_{T,e 1} > 40$ & $p_{T,e 2} > 10$	$\Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss}_{\text{track}}}) < \pi/2$
$\gamma + \text{jets}$ control region	VH	$n_\gamma = 1$	–	$p_T^\gamma > 230$	–

6.3 Photon control region

A $\gamma + \text{jets}$ CR is used for background estimation in VH categories only, and requires exactly one *loose* photon with $p_T > 230$ GeV. This CR is used to constrain the $Z \rightarrow \text{inv}$ background predictions, given it is expected that the event kinematics and topology are similar between $Z + \text{jets}$ and $\gamma + \text{jets}$ events. Candidate photon data events naturally fail the electron triggers, and require a photon trigger with threshold p_T of 200 GeV. MC candidate events for the electron or photon CRs must pass one of these three trigger requirements.

The $\gamma + \text{jets}$ requires one *loose* photon with $p_T > 230$ GeV. The requirements for $\gamma + \text{jets}$ are shown in Table 5.

Photons can usually be discriminated from other sources producing ECAL deposits due to the properties of the deposits themselves, such as ECAL and HCAL isolation of photon candidates, as well as the lack of other signatures that typically belong to other particles. However, this method is imperfect, and occasionally other particles, “fakes”, will be incorrectly identified as photons. The leading sources of fake photons is from QCD multijet events where a jet is misidentified as a photon.

In order to separate real photons from fakes in the $\gamma + \text{jets}$ CR, a purity measurement is performed. The purity is defined as the fraction of reconstructed photons that are from an isolated photon emerging from the hard scatter of the event, rather than a fake. Purity is extracted through a template fit to the data on the cluster shape parameterization variable $\sigma_{i\eta i\eta}$, which has discriminating power between real and fake photons. A real photon template is taken from data that matches the photon selection criteria, whilst the fake photon template requires at least one isolation requirement to be unfulfilled. Both templates are derived in separate bins of photon p_T , and the purity measurement is performed separately for each data taking year. This purity measurement is used to estimate the QCD multijet background in the $\gamma + \text{jets}$ CR, by replacing each event with a pseudo-event with the same recoil and p_T and weighting the new event by the purity in that p_T bin. A 25 % uncertainty is assumed, to account for effects related

to the fit procedure such as the binning of the $\sigma_{i\eta\eta}$ distribution.

6.4 Residual backgrounds from QCD multi-jet production

Searches at hadron colliders with jets in the final state are susceptible to large background contributions from QCD multijet events. While our event selection aims to minimise its impact, these processes have high cross-sections and very large numbers of events are required to accurately simulate them, which is oftentimes unfeasible. As such, in this paper, the QCD multijet background contribution is estimated using a data-driven technique.

In order to predict the expected QCD multijet background in the signal region, a sideband enriched in QCD-multijet events is defined. This sideband is defined through an inversion of the signal region requirements on $\Delta\phi_{\min}(j_{1,2,3,4}, E_T^{\text{miss}})$ and $\tilde{\omega}_{\min}$, requiring $\Delta\phi_{\min}(j_{1,2,3,4}, E_T^{\text{miss}}) < 0.5$ and $\tilde{\omega}_{\min} < 0.2$, but otherwise has an identical selection to the signal region. For the VH category, the dijet invariant mass m_{jj} requirement is also inverted in order to improve the statistics of the sideband.

The individual signal regions in both ttH and VH categories suffer from insufficient QCD MC statistics to reliably define a transfer factor. Yields are therefore integrated over all ttH subcategories and recoil bins, and separately for each VH category, in the signal (SR) and sideband (CR) regions to define MC transfer factors as follows:

$$TF_{\text{QCD}} = \frac{\sum_i \sum_j N_{i,j}^{\text{QCD,SR}}}{\sum_i \sum_j N_{i,j}^{\text{QCD,CR}}} \quad (1)$$

where indices i refer to subcategories and j to recoil bins. Within the limited QCD MC statistics available, the recoil shape as well as the relative population of the ttH subcategories do not depend on $\Delta\phi_{\min}(j_{1,2,3,4}, E_T^{\text{miss}})$ and $\tilde{\omega}_{\min}$, justifying the above choice of an “inclusive” transfer factor. As recoil shape and relative subcategory population do not depend on $\Delta\phi_{\min}(j_{1,2,3,4}, E_T^{\text{miss}})$ and $\tilde{\omega}_{\min}$ the QCD multijet MC is used to predict the fraction of QCD multijet background in each subcategory and recoil bin, defining subcategory and recoil fractions as

$$f_{c_i} = \frac{N_i^{\text{QCD,ttH}}}{N_{\text{tot}}^{\text{QCD,ttH}}} \text{ and } f_{m_j} = \frac{N_j^{\text{QCD,ttH}}}{N_{\text{tot}}^{\text{QCD,ttH}}} , \quad (2)$$

where again indices i refer to subcategories and j to recoil bins.

The estimated QCD multijet background in the ttH signal region is thus obtained as:

$$N_{i,j}^{\text{QCD,SR ttH}} = \sum_p \sum_q (N_{p,q}^{\text{data,CR ttH}} - N_{p,q}^{\text{EWK,CR ttH}}) \times TF_{\text{QCD}}^{\text{ttH}} \times f_{c_i}^{\text{ttH}} \times f_{m_j}^{\text{ttH}} , \quad (3)$$

where EWK refers to processes that are not QCD multijet and summation indices p and q are the recoil and subcategory bins respectively.

In the VH phase space, QCD multijet MC statistics allows us to define the hadronic sidebands and derive the recoil fractions f_{m_j} separately for each b-tag subcategory. The estimated QCD multijet background for the VH categories is thus obtained as:

$$N_{i,j}^{\text{QCD,SR VH}} = \sum_q (N_{i,q}^{\text{data,CR VH}} - N_{i,q}^{\text{EWK,CR VH}}) \times TF_{\text{QCD}}^{\text{VH}} \times f_{m_j}^{\text{VH}} , \quad (4)$$

where

$$f_{m_{ij}}^{\text{VH}} = \frac{N_{ij}^{\text{QCD,CR VH}}}{\sum_p N_{i,p}^{\text{QCD,CR VH}}} . \quad (5)$$

A 100% systematic uncertainty is assigned to the QCD prediction. The results of the QCD prediction aggregated over Run-2 can be found as the prefit values in Table 8.

7 Statistical interpretation

7.1 Likelihood model

The likelihood model constructed for a fit that is run simultaneously on all categories and regions (SR and CRs - μ , $\mu\mu$, e , ee and γ):

$$\mathcal{L} = \mathcal{L}_{\text{SR}} \times \mathcal{L}_{\mu\text{-CR}} \times \mathcal{L}_{\mu\mu\text{-CR}} \times \mathcal{L}_{e\text{-CR}} \times \mathcal{L}_{ee\text{-CR}} \times \mathcal{L}_{\gamma\text{-CR}} \quad (6)$$

The fit variable is $p_{\text{T}}^{\text{miss}}$. In addition to the signal strength (r), the fit implements additional free parameters, for each category i and $p_{\text{T}}^{\text{miss}}$ bin² j :

- $L^{i,j}$, which simultaneously scales the normalization of the lost lepton background in the SR and of the $t\bar{t} + W$ backgrounds in the μ and e control regions;
- $I^{i,j}$, which simultaneously scales the normalization of the $Z \rightarrow \text{invisible}$ background in the SR and of the $Z \rightarrow \mu\mu$, $Z \rightarrow ee$ and $\gamma + \text{jets}$ backgrounds in the $\mu\mu$, ee and γ CRs, respectively.

They are fit simultaneously across all regions. Due to the low statistics, in the $t\bar{t}H$ resolved and boosted class of categories, $I^{i,j}$ is shared across categories (i.e. i is fixed to “boosted” or “resolved”). In addition, in the $t\bar{t}H$ boosted categories, requiring the presence of a boosted W -tagged or top-tagged AK8 jet, the ee - and $\mu\mu$ -CR are summed into an $\ell\ell$ -CR with a likelihood similar to the one of the $\mu\mu/ee$ -CR (see below). The γ -CR is only used for the VH-targeted categories.

The likelihood for the SR reads:

$$\mathcal{L}_{\text{SR}} = \prod_{cat=i}^{n_{cat}} \prod_{\text{recoil}=j(i)}^{n_{\xi(i)}} \text{Poisson}(n_{\text{obs}}^{i,j} | n_{\text{pred}}^{i,j}), \quad (7)$$

with:

$$\begin{aligned} n_{\text{pred}}^{i,j} = & r \times s^{i,j} \times \rho_s^{i,j} \\ & + b_{Z \rightarrow \text{inv.}}^{i,j} \times I^{i,j} \times \rho_{Z \rightarrow \text{inv.}}^{i,j} \\ & + b_{\text{LL}}^{i,j} \times L^{i,j} \times \rho_{\text{LL}}^{i,j} \\ & + b_{\text{QCD}}^{i,j} \times \rho_{\text{QCD}}^{i,j} \end{aligned} \quad (8)$$

where the symbols are defined in Table 6.

The likelihood for each CR reads:

$$\mathcal{L}_{X\text{-CR}} = \prod_{cat=i}^{n_{cat}} \prod_{\text{recoil}=j(i)}^{n_{\xi(i)}} \text{Poisson}(n_{\text{obs}}^{i,j} | r \times s_{\mu}^{i,j} \times \rho_s^{i,j} + X_{\text{proc}}^{i,j} \times L^{i,j} \times \rho_{t\bar{t},W}^{i,j}), \quad (9)$$

²Since categories do not all implement the same $p_{\text{T}}^{\text{miss}}$ bins, each $p_{\text{T}}^{\text{miss}}$ bin is noted as $j(i)$ and the number of bins in each category $n_{\xi(i)}$.

Table 6: Meaning of symbols used in the likelihood.

Symbol	Meaning
r	Signal strength
$s^{i,j}$	MC-predicted number of signal events in bin i, j of the SR
$\rho_s^{i,j}$	Systematics coming from signal MC variations in bin i, j of the SR
$b_{Z \rightarrow \text{inv.}}^{i,j}$	MC-predicted number of $Z \rightarrow \text{inv.}$ events in bin i, j of the SR
$\rho_{Z \rightarrow \text{inv.}}^{i,j}$	Systematics coming from $Z \rightarrow \text{inv.}$ MC variations in bin i, j of the SR
$b_{\text{LL}}^{i,j}$	MC-predicted number of LL (lost lepton) events in bin i, j of the SR
$\rho_{\text{LL}}^{i,j}$	Systematics affecting the LL (lost-lepton) background in bin i, j of the SR
$b_{\text{QCD}}^{i,j}$	Predicted number of QCD events in bin i, j of the SR
$\rho_{\text{QCD}}^{i,j}$	Systematics on the QCD component in bin i, j of the SR

where $X_{\text{proc}}^{i,j}$ is the yield for the background processes being constrained in CR X , and $\rho_{\text{proc}}^{i,j}$ refers to the corresponding associated systematic uncertainties.

7.2 Systematic uncertainties

Quantities used in the maximum-likelihood fit are subject to experimental and theoretical uncertainties. Theoretical uncertainties arise mostly from PDF, renormalisation and factorisation scale variations when considering higher-order effects in QCD and EW perturbative expansion. The exact recipe detailing the application of these uncertainties in V +jets and γ + jets processes is found in Lindert et. al. [55]. Experimental uncertainty is dominated by identifying and reconstructing lepton and photon candidates, followed by uncertainties on the jet energy scale, jet energy resolution, and trigger efficiencies. Data derived scale factors and uncertainties to b-tagging and DeepAk8-tagging efficiencies are applied. The full scope of uncertainties modelled as independent nuisance parameters in the Run 2 maximum-likelihood fit³ are detailed in Table 7, with the percentage range corresponding to the minimum/maximum deviation of the systematic from the nominal yield for each region.

³Systematics are decorrelated between each year in many cases, with their range tabulated. In the case of pre-firing, the range is for years 2016 and 2017 only.

Systematic Uncertainty	SR	γ + jets	e + jets	μ + jets	ee + jets	$\mu\mu$ + jets
Theoretical uncertainties						
Fact. scale V+jets (QCD)	18-23 %	0.1-0.2 %	8.6-10 %	9.4-11 %	21-26 %	21-26 %
PDF V+jets (QCD)	21-28 %	0.1 %	10-13 %	12-15 %	20-25 %	20-26 %
Ren. scale V+jets (QCD)	19-27 %	0.1-0.3 %	13-16 %	13-19 %	22-34 %	21-34 %
Ren. and Fact. scale ttH (QCD)	0.1-1.8	0.0 %	0.1-1.1 %	0.1-1.4 %	0.1-0.3 %	0.1-0.4 %
Ren. and Fact. scale $t\bar{t}$ (QCD)	7.6-14	0.1-0.4 %	8.7-17 %	7.4-13 %	1.2-2.8 %	1.4-5.8 %
Cross-section scale ttH	6.0-9.0 %	-	-	-	-	-
Cross-section PDF ttH	3.5 %	-	-	-	-	-
Experimental uncertainties						
Luminosity	1.5-2.6 %	1.5-2.6 %	1.5-2.6 %	1.5-2.6 %	1.5-2.6 %	1.5-2.6 %
Top tagging	3.5-6.5 %	-	3.7-5.6 %	3.7-5.6 %	-	-
W tagging	11-18 %	-	10-18 %	9.3-16 %	-	-
b tagging	8.9-12 %	7.0-8.9 %	8.6-11 %	8.2-10 %	9.9-14 %	9.9-16 %
Electron identification and isolation	-	-	5.6-12 %	-	11-20 %	-
Electron reconstruction	-	-	0.6-0.9 %	-	1.1-1.5 %	-
Muon identification	-	-	-	0.2-1.0 %	-	0.4-1.9 %
Muon isolation	-	-	-	0.1-0.2 %	-	0.2-0.3 %
Photon identification and isolation	-	2.4-12 %	-	-	-	-
Photon reconstruction	-	0.4-0.7 %	-	-	-	-
Pile-up	3.6-10.0 %	1.6-3.3 %	2.9-4.9 %	2.7-4.9 %	3.0-8.5 %	2.6-5.6 %
Pre-firing	1.0-1.2 %	0.7-0.9 %	1.2 %	1.0-1.2 %	0.9-1.6 %	0.7-1.3 %
Top pT reweighting	3.3-3.4	0.0 %	3.6-3.9 %	3.6-4.0 %	0.3-0.6 %	0.3-0.5 %
Trigger	2.0 %	2.0 %	2.0 %	2.0 %	2.0 %	2.0 %
Tau veto	0.0 %	0.0 %	0.1-0.2 %	0.2 %	0.1 %	0.1 %
Initial state radiation	0.5%-6.0%	8.0%	2.0%-6.0%	2.0%-6.0%	2.0%-6.0%	2.0%-6.0%
Final state radiation	0.5%-5.0%	3.0%	2.0%-5.0%	2.0%-5.0%	2.0%-5.0%	2.0%-5.0%
Jet Energy Uncertainties						
Jet Energy Resolution	2.4-3.6 %	0.9-2.9 %	1.7-2.8 %	1.7-3.0 %	1.0-3.5 %	1.1-3.2 %
Jet Energy Scale	0.2-5.8 %	0.0-3.8 %	0.2-5.0 %	0.1-5.1 %	0.1-6.7 %	0.3-5.7 %

Table 7: The percentage ranges corresponding to the pre-fit minimum/maximum deviation of a given systematic from the nominal yield for each region. Note that some systematics are decorrelated between years.

8 Results

8.1 Higgs to invisible branching fraction

The recoil distributions across all $t\bar{t}H$ and VH categories in this analysis are presented here. The “total pre-fit” yields correspond to background predictions purely from MC, while “total post-fit” is the background prediction post-minimising the likelihood presented in Section 7.1. Full Run 2 data-MC agreement plots for e +jets and μ +jets CRs are shown in Figures 2 and 3, respectively. In these distributions, the background processes $t\bar{t}$, $W \rightarrow \ell\nu$ and single-top are classified into ℓ_{lost} . The $\mu\mu$ +jets, ee +jets, $\ell\ell$ +jets (for $t\bar{t}H$) and γ +jets (for VH) CR distributions used for the prediction of backgrounds stemming from $Z \rightarrow \nu\bar{\nu}$ decays are shown in Figures 4, 5, 6 and 7 for Run 2 yields. In addition, the total SM background prediction in the SR, consisting of ℓ_{lost} , $Z \rightarrow \nu\bar{\nu}$ and QCD backgrounds, is shown in Figure 8, with the respective event yields tabulated in Table 8.

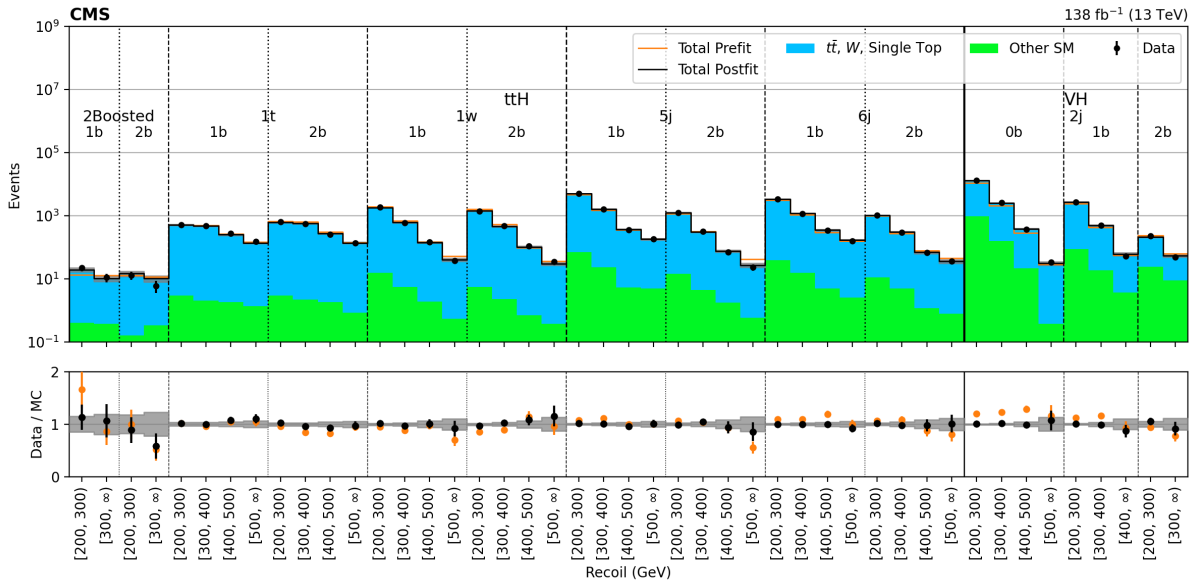
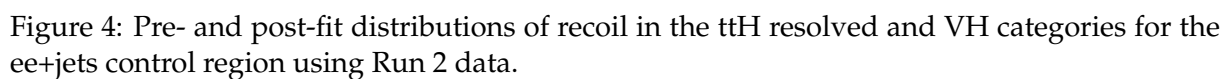
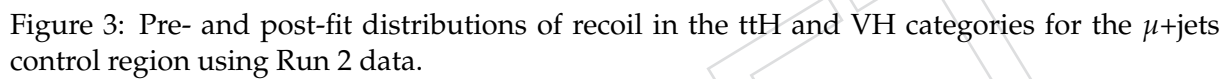


Figure 2: Pre- and post-fit distributions of recoil in the $t\bar{t}H$ and VH categories for the e +jets control region using Run 2 data.



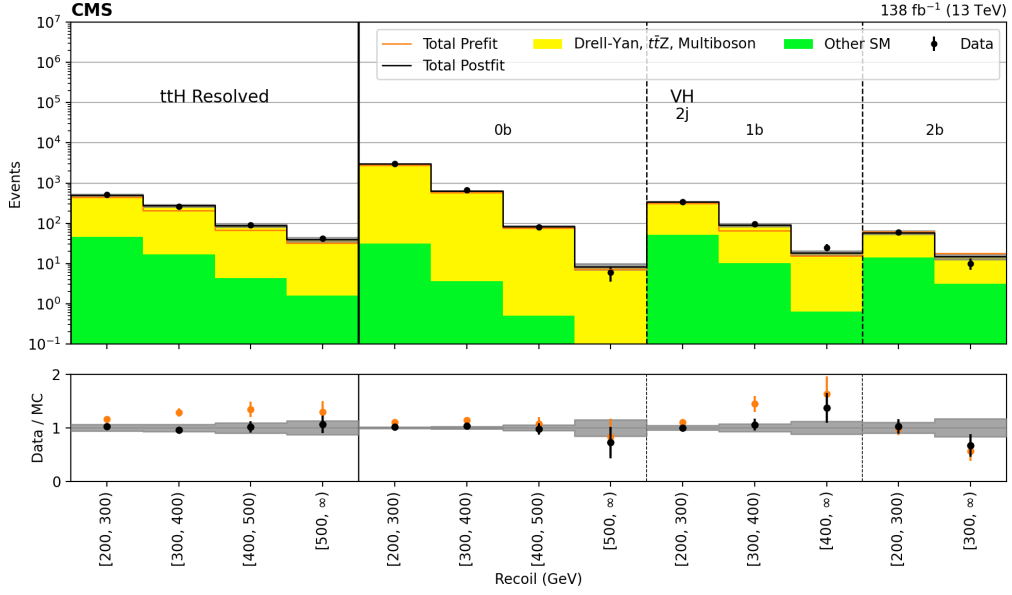


Figure 5: Pre-fit and post-fit distributions of recoil in the $t\bar{t}H$ resolved and VH categories for the $\mu\mu$ +jets control region using Run 2 data.

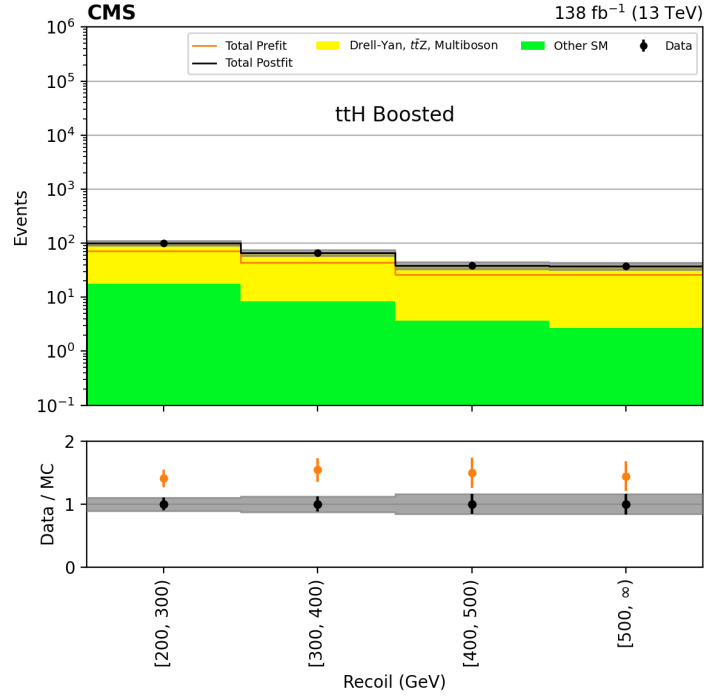


Figure 6: Pre- and post-fit distributions of recoil in the $t\bar{t}H$ boosted category for the $\ell\ell$ +jets control region using Run 2 data.

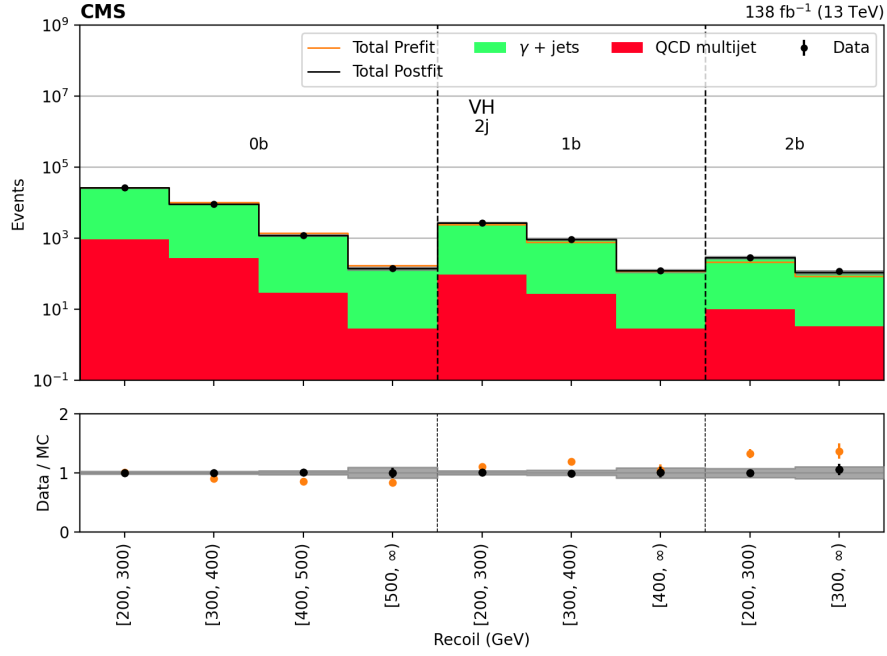


Figure 7: Pre- and post-fit distributions of recoil in the VH category for γ +jets control region using Run 2 data.

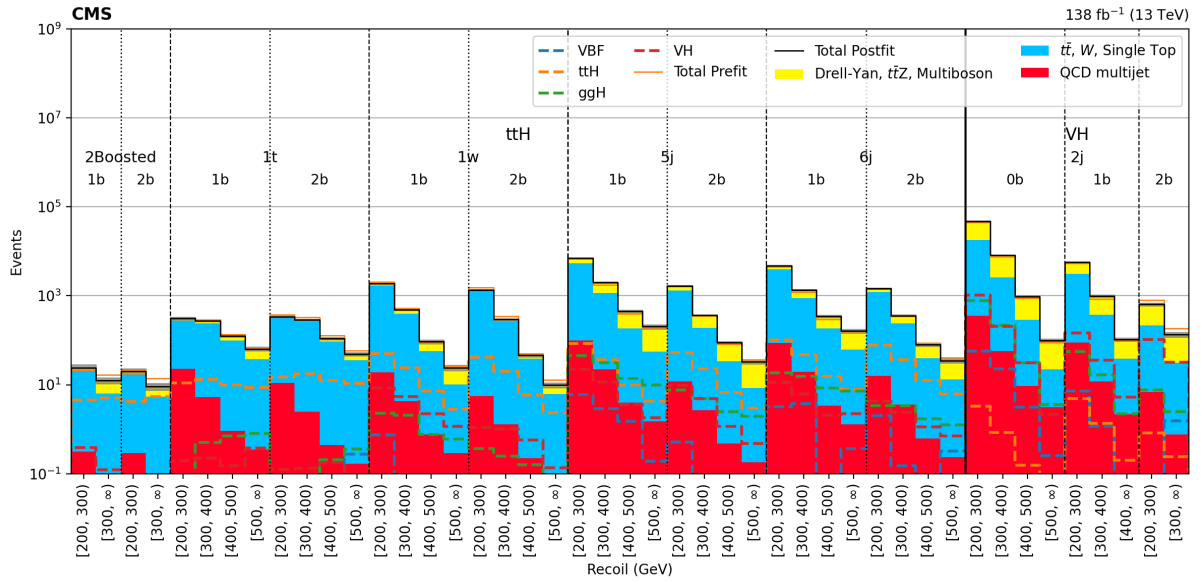
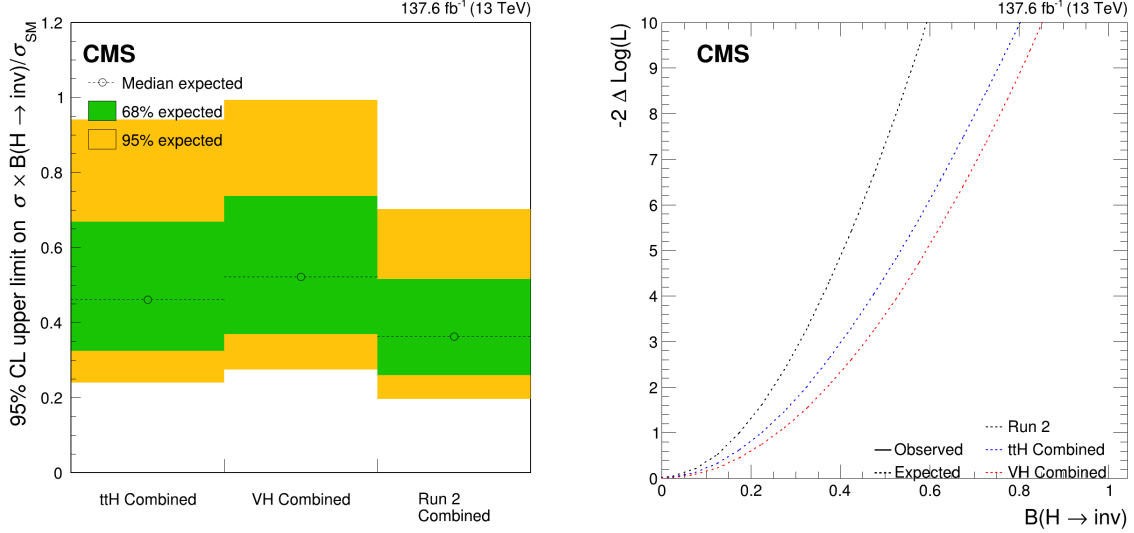


Figure 8: Pre- and post-fit distributions of recoil in the $t\bar{t}H$ and VH categories for the signal region using Run 2 data.

Table 8: Total Run 2 post-fit yields in each category in the SR, broken down into the contributing processes for pre and post-fit. Errors are inclusive of statistical and systematic contributions.

category	recoil	ℓ_{lost}				$Z \rightarrow \text{inv}$				QCD		TotalSM		Signal	$S/\Delta B$
		pre-fit	post-fit	pre-fit	post-fit	pre-fit	post-fit	pre-fit	post-fit	pre-fit	post-fit				
ttH 11b	[200, 300]	254.0 \pm 18.3	253.0 \pm 9.06	25.4 \pm 3.61	34.9 \pm 3.56	23.1 \pm 16.0	23.1 \pm 14.4	280.0 \pm 21.9	288.0 \pm 12.6	11.3 \pm 0.89	11.3 \pm 0.89	0.9 \pm 0.11	0.9 \pm 0.11		
	[300, 400]	243.0 \pm 16.5	235.0 \pm 8.76	23.8 \pm 2.09	35.6 \pm 4.41	5.25 \pm 3.57	5.25 \pm 3.19	267.0 \pm 18.6	270.0 \pm 13.2	14.0 \pm 1.1	14.0 \pm 1.1	1.06 \pm 0.13	1.06 \pm 0.13		
	[400, 500]	104.0 \pm 7.67	98.5 \pm 5.18	18.4 \pm 1.58	27.8 \pm 4.6	0.92 \pm 0.63	0.92 \pm 0.58	122.0 \pm 9.25	126.0 \pm 9.78	10.9 \pm 0.84	10.9 \pm 0.84	1.11 \pm 0.17	1.11 \pm 0.17		
	[500, ∞]	41.0 \pm 3.66	38.0 \pm 2.55	19.0 \pm 3.01	26.3 \pm 4.39	0.35 \pm 0.27	0.35 \pm 0.24	59.9 \pm 6.67	64.3 \pm 6.94	9.95 \pm 0.78	9.95 \pm 0.78	1.43 \pm 0.27	1.43 \pm 0.27		
	[200, 300]	339.0 \pm 25.8	315.0 \pm 10.3	13.0 \pm 1.55	18.7 \pm 1.85	10.9 \pm 6.88	10.9 \pm 6.23	352.0 \pm 27.3	334.0 \pm 12.2	15.4 \pm 1.31	15.4 \pm 1.31	1.27 \pm 0.15	1.27 \pm 0.15		
ttH 112b	[300, 400]	298.0 \pm 22.8	266.0 \pm 9.06	13.1 \pm 1.14	20.0 \pm 2.39	2.49 \pm 1.68	2.49 \pm 1.44	311.0 \pm 23.9	286.0 \pm 11.5	17.6 \pm 1.46	17.6 \pm 1.46	1.54 \pm 0.19	1.54 \pm 0.19		
	[400, 500]	105.0 \pm 8.6	94.0 \pm 4.6	10.4 \pm 0.9	15.4 \pm 2.53	0.44 \pm 0.28	0.44 \pm 0.25	116.0 \pm 9.51	109.0 \pm 7.13	12.9 \pm 1.07	12.9 \pm 1.07	1.81 \pm 0.27	1.81 \pm 0.27		
	[500, ∞]	36.8 \pm 3.64	35.7 \pm 2.54	10.3 \pm 1.29	13.9 \pm 2.32	0.17 \pm 0.11	0.17 \pm 0.12	47.1 \pm 4.94	49.6 \pm 4.86	11.3 \pm 1.01	11.3 \pm 1.01	2.33 \pm 0.44	2.33 \pm 0.44		
	[200, 300]	1839.0 \pm 140.0	1709.0 \pm 37.6	144.0 \pm 26.9	191.0 \pm 19.5	18.8 \pm 12.4	18.8 \pm 11.2	1984.0 \pm 167.0	1900.0 \pm 57.1	60.8 \pm 5.36	60.8 \pm 5.36	1.07 \pm 0.13	1.07 \pm 0.13		
	[300, 400]	432.0 \pm 33.9	395.0 \pm 12.5	62.4 \pm 7.8	90.5 \pm 10.9	4.3 \pm 2.89	4.3 \pm 2.64	495.0 \pm 41.7	485.0 \pm 23.5	31.3 \pm 2.75	31.3 \pm 2.75	1.33 \pm 0.18	1.33 \pm 0.18		
ttH 1W2b	[400, 500]	56.7 \pm 4.99	55.5 \pm 3.59	24.3 \pm 2.6	35.8 \pm 6.21	0.75 \pm 0.56	0.75 \pm 0.47	81.0 \pm 7.59	91.3 \pm 9.81	10.3 \pm 0.93	10.3 \pm 0.93	1.05 \pm 0.21	1.05 \pm 0.21		
	[500, ∞]	12.9 \pm 1.58	9.9 \pm 1.18	10.1 \pm 1.69	14.1 \pm 2.82	0.29 \pm 0.21	0.29 \pm 0.18	23.0 \pm 3.27	24.0 \pm 4.0	4.63 \pm 0.44	4.63 \pm 0.44	1.16 \pm 0.23	1.16 \pm 0.23		
	[200, 300]	1472.0 \pm 120.0	1297.0 \pm 31.0	36.9 \pm 5.19	52.5 \pm 5.21	5.55 \pm 3.7	5.55 \pm 3.53	1509.0 \pm 126.0	1349.0 \pm 36.2	42.6 \pm 3.91	42.6 \pm 3.91	1.17 \pm 0.14	1.17 \pm 0.14		
	[300, 400]	310.0 \pm 26.9	266.0 \pm 9.95	17.9 \pm 2.11	27.0 \pm 3.3	1.27 \pm 0.87	1.27 \pm 0.83	328.0 \pm 29.0	293.0 \pm 13.3	21.2 \pm 1.98	21.2 \pm 1.98	1.6 \pm 0.22	1.6 \pm 0.22		
	[400, 500]	37.2 \pm 3.96	38.2 \pm 3.16	5.32 \pm 0.5	7.97 \pm 1.31	0.22 \pm 0.16	0.22 \pm 0.14	42.5 \pm 4.47	46.2 \pm 4.48	6.55 \pm 0.66	6.55 \pm 0.66	1.46 \pm 0.29	1.46 \pm 0.29		
ttH 2Boosted1b	[500, ∞]	7.79 \pm 1.15	6.03 \pm 0.96	2.66 \pm 0.48	3.76 \pm 0.64	0.08 \pm 0.07	0.08 \pm 0.06	10.5 \pm 1.63	9.79 \pm 1.6	2.52 \pm 0.25	2.52 \pm 0.25	1.58 \pm 0.42	1.58 \pm 0.42		
	[200, 300]	13.8 \pm 1.88	19.9 \pm 3.34	2.52 \pm 0.31	3.76 \pm 0.39	0.31 \pm 0.29	0.31 \pm 0.21	16.3 \pm 2.18	23.7 \pm 3.73	4.95 \pm 0.53	4.95 \pm 0.53	1.33 \pm 0.35	1.33 \pm 0.35		
	[300, ∞]	7.76 \pm 1.35	6.34 \pm 1.4	3.79 \pm 0.38	6.0 \pm 0.78	0.09 \pm 0.08	0.09 \pm 0.07	11.5 \pm 1.73	12.3 \pm 2.19	5.23 \pm 0.51	5.23 \pm 0.51	2.39 \pm 0.66	2.39 \pm 0.66		
	[200, 300]	14.5 \pm 2.19	15.9 \pm 2.73	2.38 \pm 0.49	3.72 \pm 0.81	0.29 \pm 0.23	0.29 \pm 0.19	16.9 \pm 2.69	19.7 \pm 3.54	4.17 \pm 0.42	4.17 \pm 0.42	1.18 \pm 0.33	1.18 \pm 0.33		
	[300, ∞]	6.07 \pm 1.07	5.35 \pm 1.2	2.38 \pm 0.23	3.73 \pm 0.45	0.08 \pm 0.1	0.08 \pm 0.07	8.45 \pm 1.3	9.07 \pm 1.65	5.49 \pm 0.54	5.49 \pm 0.54	3.33 \pm 0.94	3.33 \pm 0.94		
ttH 51b	[200, 300]	5010.0 \pm 328.0	5294.0 \pm 84.5	1562.0 \pm 348.0	1705.0 \pm 65.6	99.1 \pm 73.8	99.1 \pm 68.2	6572.0 \pm 676.0	6999.0 \pm 150.0	158.0 \pm 11.9	158.0 \pm 11.9	1.05 \pm 0.1	1.05 \pm 0.1		
	[300, 400]	1039.0 \pm 66.3	1137.0 \pm 25.0	651.0 \pm 85.8	839.0 \pm 41.8	22.5 \pm 17.2	22.5 \pm 15.5	1690.0 \pm 152.0	1976.0 \pm 66.8	82.1 \pm 6.29	82.1 \pm 6.29	1.23 \pm 0.14	1.23 \pm 0.14		
	[400, 500]	174.0 \pm 13.7	183.0 \pm 7.2	209.0 \pm 18.9	267.0 \pm 21.6	3.95 \pm 3.08	3.96 \pm 2.61	383.0 \pm 32.7	450.0 \pm 28.8	28.9 \pm 2.89	28.9 \pm 2.89	1.0 \pm 0.16	1.0 \pm 0.16		
	[500, ∞]	53.4 \pm 6.6	54.5 \pm 3.29	123.0 \pm 25.6	146.0 \pm 18.1	1.5 \pm 1.1	1.5 \pm 1.01	177.0 \pm 32.2	201.0 \pm 21.4	16.2 \pm 1.35	16.2 \pm 1.35	0.76 \pm 0.14	0.76 \pm 0.14		
	[200, 300]	1236.0 \pm 91.5	1324.0 \pm 37.9	316.0 \pm 67.6	350.0 \pm 13.3	11.8 \pm 8.57	11.8 \pm 7.71	1552.0 \pm 159.0	1673.0 \pm 51.3	70.0 \pm 5.56	70.0 \pm 5.56	1.37 \pm 0.15	1.37 \pm 0.15		
ttH 52b	[300, 400]	193.0 \pm 15.3	190.0 \pm 8.46	133.0 \pm 17.1	174.0 \pm 8.9	2.68 \pm 2.03	2.69 \pm 1.82	326.0 \pm 32.4	364.0 \pm 17.4	32.6 \pm 2.62	32.6 \pm 2.62	1.88 \pm 0.24	1.88 \pm 0.24		
	[400, 500]	33.0 \pm 3.29	33.6 \pm 3.14	42.0 \pm 3.92	53.8 \pm 4.56	0.48 \pm 0.37	0.48 \pm 0.34	75.1 \pm 7.22	87.4 \pm 7.7	10.3 \pm 0.84	10.3 \pm 0.84	1.33 \pm 0.23	1.33 \pm 0.23		
	[500, ∞]	12.3 \pm 1.97	8.32 \pm 1.25	21.2 \pm 3.94	24.7 \pm 3.16	0.18 \pm 0.16	0.18 \pm 0.15	33.4 \pm 5.92	33.0 \pm 4.41	5.39 \pm 0.46	5.39 \pm 0.46	1.22 \pm 0.27	1.22 \pm 0.27		
	[200, 300]	3526.0 \pm 227.0	3840.0 \pm 65.5	726.0 \pm 149.0	801.0 \pm 30.7	85.9 \pm 65.1	86.0 \pm 58.7	4252.0 \pm 376.0	4641.0 \pm 96.2	134.0 \pm 9.95	134.0 \pm 9.95	1.4 \pm 0.13	1.4 \pm 0.13		
	[300, 400]	797.0 \pm 53.1	874.0 \pm 21.3	342.0 \pm 42.1	438.0 \pm 21.9	19.5 \pm 14.5	19.5 \pm 13.2	1139.0 \pm 95.3	1312.0 \pm 43.2	72.9 \pm 5.92	72.9 \pm 5.92	1.69 \pm 0.19	1.69 \pm 0.19		
ttH 61b	[400, 500]	147.0 \pm 11.5	179.0 \pm 7.97	127.0 \pm 11.3	162.0 \pm 13.3	3.42 \pm 2.55	3.42 \pm 2.27	274.0 \pm 22.8	341.0 \pm 21.2	26.2 \pm 2.05	26.2 \pm 2.05	1.23 \pm 0.17	1.23 \pm 0.17		
	[500, ∞]	55.1 \pm 5.95	61.1 \pm 3.86	83.2 \pm 17.4	98.2 \pm 11.9	1.29 \pm 0.93	1.29 \pm 0.88	138.0 \pm 23.4	159.0 \pm 15.8	17.7 \pm 1.46	17.7 \pm 1.46	1.13 \pm 0.2	1.13 \pm 0.2		
	[200, 300]	1155.0 \pm 89.2	1214.0 \pm 29.2	208.0 \pm 38.4	235.0 \pm 8.98	15.6 \pm 12.1	15.6 \pm 10.9	1363.0 \pm 128.0	1448.0 \pm 38.2	87.4 \pm 7.3	87.4 \pm 7.3	2.29 \pm 0.25	2.29 \pm 0.25		
	[300, 400]	213.0 \pm 18.0	239.0 \pm 10.3	90.3 \pm 10.2	118.0 \pm 6.07	3.56 \pm 2.95	3.57 \pm 2.49	304.0 \pm 28.2	357.0 \pm 16.3	42.7 \pm 3.44	42.7 \pm 3.44	2.61 \pm 0.33	2.61 \pm 0.33		
	[400, 500]	42.6 \pm 4.45	38.7 \pm 3.66	31.7 \pm 3.03	40.6 \pm 3.49	0.62 \pm 0.52	0.62 \pm 0.42	74.2 \pm 7.48	79.4 \pm 7.15	15.2 \pm 1.24	15.2 \pm 1.24	2.13 \pm 0.37	2.13 \pm 0.37		
VH 2j0b	[500, ∞]	15.9 \pm 2.1	12.9 \pm 1.71	18.8 \pm 3.43	21.5 \pm 2.63	0.24 \pm 0.23	0.24 \pm 0.17	34.7 \pm 5.54	34.4 \pm 4.34	8.41 \pm 0.84	8.41 \pm 0.84	1.94 \pm 0.44	1.94 \pm 0.44		
	[200, 300]	15754.0 \pm 1753.0	17888.0 \pm 241.0	27040.0 \pm 5739.0	29125.0 \pm 520.0	358.0 \pm 199.0	359.0 \pm 254.0	42794.0 \pm 7492.0	47013.0 \pm 761.0	1875.0 \pm 81.5	1875.0 \pm 81.5	2.46 \pm 0.15	2.46 \pm 0.15		
	[300, 400]	2167.0 \pm 164.0	2557.0 \pm 52.5	5138.0 \pm 748.0	5505.0 \pm 94.3	56.9 \pm 31.4	56.9 \pm 38.6	7304.0 \pm 912.0	8062.0 \pm 147.0	447.0 \pm 19.7	447.0 \pm 19.7	3.04 \pm 0.19	3.04 \pm 0.19		
	[400, 500]	214.0 \pm 17.7	282.0 \pm 14.3	627.0 \pm 45.3	680.0 \pm 21.8	9.47 \pm 5.18	9.47 \pm 7.15	841.0 \pm 63.1	962.0 \pm 36.1	66.0 \pm 3.37	66.0 \pm 3.37	1.83 \pm 0.16	1.83 \pm 0.16		
	[500, ∞]	17.2 \pm 2.5	19.4 \pm 2.75	68.0 \pm 7.77	76.0 \pm 6.75	3.12 \pm 2.11	3.12 \pm 2.11	85.2 \pm 10.3	95.3 \pm 9.5	5.99 \pm 0.31	5.99 \pm 0.31	0.63 \pm 0.1	0.63 \pm 0.1		
VH 2j1b	[200, 300]	2752.0 \pm 163.0	3045.0 $\pm</$												

The observed (expected) upper limits on $BR(H \rightarrow inv)$ at 95% C.L. extracted from the likelihood fit presented in Section 7.1 are shown in Figure 9 together with the observed (expected) profile likelihood distribution. Numerical values for the observed and expected limits are tabulated in Table 9.



(a) Run 2 expected limits in ttH and VH categories

(b) Profile likelihood scan of the fit model for ttH and VH

Figure 9: 95% C.L. limits for ttH, VH and the Run2 combination with a the profile likelihood scan corresponding to these limits.

Table 9: Summary of observed and expected limits for $\mathcal{B}(H \rightarrow inv.)$.

	ttH	VH	Run 2
95% C.L. limit	X.XX (0.46)	X.XX (0.52)	X.XX (0.36)

8.2 Dark matter production in association with a $t\bar{t}$ pair

Despite significant cosmological motivation for the existence of dark matter, there is currently no direct evidence for its existence. As such, there is significant interest in exploring the possibility of dark matter being produced at CMS. In addition to the Higgs portal hypothesis explored in this paper, some of the simplest forms of dark matter production, called DMSimp models, have also been studied.

One such process is the production of top quark pairs in association with dark matter ($t\bar{t} + DM$). This process is topologically similar to ttH production, with the Higgs replaced by a scalar or pseudoscalar mediator decaying into two DM particle with $m_\chi \leq m_{\phi(a)}/2$ for DM particle χ and scalar (pseudoscalar) mediator ϕ (a). Is it therefore an ideal signal to study in a Higgs to Invisible search targeting the ttH topology. This process was studied for a range of parameter points according to mediator mass m_ϕ , as well as DM particle mass M_χ , as shown in Table 1. For fixed $M_\phi(M_a) = 100$ GeV and floating M_χ , an upper limit on the signal strength of 0.46 is found, which is independent of M_χ for $M_\chi < 50$ GeV. For $M_\chi > 50$ GeV, sensitivity drops significantly, because the dark matter mass becomes off-shell (greater than half of that of the mediator mass). For fixed $M_\chi = 1$ GeV, a strong dependence of the signal strength on M_ϕ is seen. This process is the dominating source of sensitivity for Figure 10 at low mass points.

8.3 Dark matter production in association with a jet or a vector boson

Other DMSimp signals of interest are dark matter production in association with a boosted vector boson or boosted jet. These processes are akin to VH and ggF production respectively, again with the Higgs particle swapped out by a scalar or pseudoscalar mediator decaying to a DM pair. Although these processes are not directly targeted by this analysis, which focused on ttH production and VH production with the vector boson decaying to two resolved jets, reasonable signal sensitivity is found in the ttH and VH categories. For a given mediator, the two processes in this analysis are considered concurrently and jointly referred to as Mono-X + DM. The parameter points and cross-sections for these processes, over a range over mediator masses M_ϕ and for both a scalar and pseudoscalar mediator, is shown in Table 2.

Although the $t\bar{t} + \text{DM}$ signal dominates sensitivity at low M_ϕ , Mono-X + DM processes help improve the upper limit on the signal strength at higher mass points. Results for this combination are shown in Figure 10.

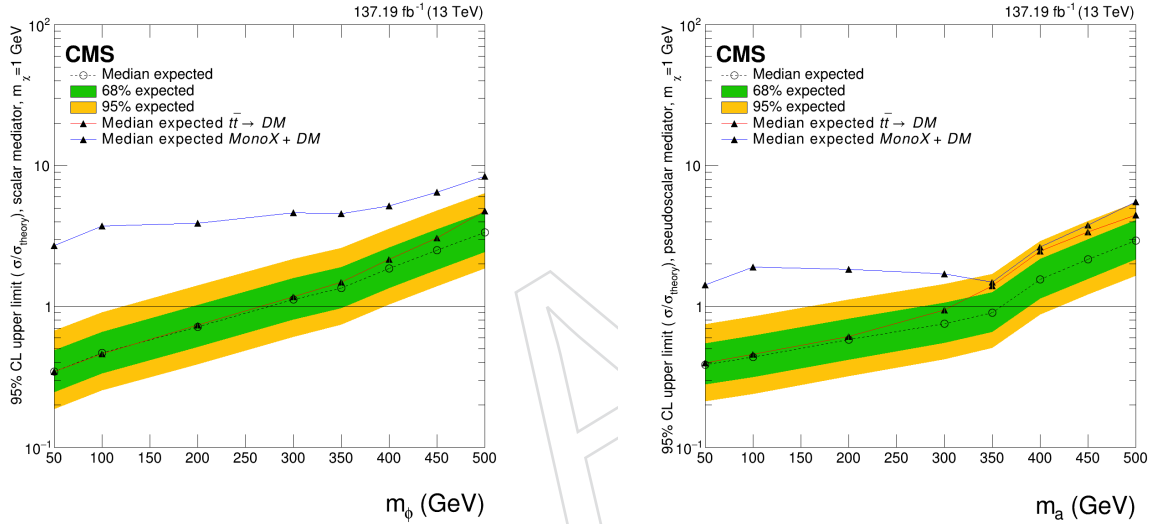


Figure 10: Upper limits on the signal strength $\mu = \sigma / \sigma_{\text{theoretical}}$ for the combination of DMSimp models $t\bar{t} + \text{DM}$ and Mono-X + DM signal samples for a scalar (left) and pseudoscalar (right) mediator. Median expected values for $t\bar{t} + \text{DM}$ and Mono-X + DM individually are also shown in red and blue respectively.

9 Combination

The results of this analysis are further combined with those of previous analyses. Different analyses are considered for the two interpretations.

9.1 Combination of H(inv) results

Invisible decays of the Higgs boson can be sought after in a variety of production modes of the Higgs boson. The results of this analysis are combined with CMS results based on different analysis channels using the same data sets, as well as results from earlier data sets. The combination is performed by defining a new likelihood function, which is the product of the likelihood functions of all individual channels. Unless explicitly specified below, parameters of the individual likelihood functions are treated as independent.

This section is organized as follows. First, the individual results included in the combination are described. Then, the treatment of event selection overlaps between different analyses in the same data sets, as well as the treatment of uncertainty correlations are discussed. Finally, the results of the combination are presented.

9.1.1 Description of analyses

The VBF analysis [54] targets production of the Higgs boson in association with a dijet system that exhibits large invariant mass, a large rapidity gap between the jets, and small angular separation of the jets in the transverse plane. The analysis is performed in two categories, dubbed “VTR” and “MTR”, which target the p_T^{miss} ranges of 160-250 GeV and >250 GeV, respectively. In both cases, a possible signal is extracted from the distribution of the dijet invariant mass variable after applying a selection focused on the dijet properties. A maximum-likelihood fit is performed for background estimation and signal extraction, and control regions with leptons or a photon are included in order to constrain the leading backgrounds from V+jets production. A similar fitting scheme to this analysis is used, in which the shape and normalization of the leading background contributions are directly obtained from data, with predictions for transfer factors between processes and regions being obtained from simulation. Leading uncertainty contributions arise from the limited theoretical knowledge of the transfer factors between the Z, W and photon production processes.

In the case of $Z(\ell\ell)H$ production [56], the presence of charged leptons in the final state enables the use of lepton triggers, and thus allows access to a significantly lower range of $p_T^{\text{miss}} > 100$ GeV. Events are selected by requiring two leptons consistent with the decay of a Z boson, and strict requirements are imposed on the relationship between the transverse component of the Z boson candidate momentum, and the \vec{p}_T^{miss} vector, focusing on a topology where the two are of similar magnitude and oriented in opposite directions. The leading background contributions from ZZ and WZ production are constrained using control regions with three or four charged leptons, which are included in an ML fit together with the signal region. Similarly to the VBF analysis, the leading systematic uncertainties relate to the theoretical prediction of the transfer factors between the ZZ and WZ processes.

Further hadronic final states are explored in an analysis targeting ggH and boosted $V(qq)$ production [23]. At large $p_T^{\text{miss}} > 250$ GeV, the decay products of Z or W bosons are often close-by, and can be reconstructed as a single wide jet. In this case, the ggH and VH modes produce a similar topology with a high- p_T leading jet recoiling against large p_T^{miss} . The analysis is categorized based on the properties of the leading jet, which is a narrow light jet in the case of ggH, or a wide jet with a substructure consistent with a Z or W decay, as identified by the DeepAK8 tagger, in the case of VH. The VH category is further split into two low-purity and high-purity categories based on the quality of the V tag. In each category, a signal is sought in the p_T^{miss} distribution, which is analyzed using the same fit method as in this analysis as well as the VBF analysis. The leading backgrounds from V+jets production are again constrained using control regions with leptons or a photon. In contrast to the other topologies, high-precision predictions for the transfer factors between the different processes are available, and the effect of theory uncertainties is therefore reduced significantly. Leading uncertainty contributions arise from the modeling of photon and electron identification and reconstruction, which affect the relationship between control and signal regions.

Finally, constraints on $t\bar{t}H$ (inv) production in final states with one or two leptons are obtained from analyses originally designed for the detection of supersymmetric partners of the top quark [57, 58]. A combination of these analyses has previously been presented in Ref. [59].

Results are further combined with analyses targeting all production modes based on data sets collected by CMS between 2011 and 2015 at center-of-mass energies of 7, 8, and 13 TeV. A combination of these analysis has previously been presented in Ref. [60].

For presentational purposes, the set of analyses using the data sets collected between 2016 and 2018 will be referred to as “Run-2” analyses. Earlier results are identified as “2015” if they are based on the 13 TeV data set collected in 2015. Results based on the 2011–2012 data taking period are referred to as “Run-1”.

9.1.2 Treatment of selection overlap

The individual analyses employed in the combination select partially overlapping analysis regions. In order for the statistical combination to be feasible, such overlap must be removed. Initially, significant overlap is present between the VBF and ggH/boosted VH analyses, as well as between the ggH/boosted VH analysis and the VH2j0b categories of this analysis. In both cases, the overlap is treated by modifying the event selection criteria in the ggH/boosted VH analysis to remove the overlap.

For the case of the overlap with the VBF analysis, events are considered for rejection if they have with at least two AK4 jets with $|\eta| < 4.7$, where the leading (trailing) jet has $p_T > 80$ GeV ($p_T > 40$ GeV). An event is removed if the two jets are from opposite detector hemispheres ($\eta_1 \times \eta_2 < 0$), have large transverse mass ($m_{jj} > 200$ GeV), small angular separation in the transverse plane $\Delta\Phi_{jj} < 1.5$ and a large rapidity gap ($|\eta_1 - \eta_2| > 1.0$). These requirements mirror the selection used to enhance the characteristic VBF phase space in Ref. [54]. This additional selection is applied in all categories of the ggH/boosted VH analysis.

The overlap between the ggH/boosted VH analysis and the VH2j0b category of this analysis is driven by the low-purity VH category of the boosted analysis. Overlap with the other categories is negligible. The overlap is removed by explicitly removing events from the low-purity VH category if they contain exactly two AK4 jets forming a dijet candidate with $65 < m_{jj} < 120$ GeV.

Overall, the modifications of the ggH/boosted VH selection result in a relative reduction of the exclusion sensitivity of that analysis of X%.

9.1.3 Treatment of uncertainty correlations

When combining analyses, correlations between source of uncertainties in different analysis channels and data sets are taken into account whenever possible. In this section, the correlations between the analyses performed on the data sets taken between 2016 and 2018 are described. For the earlier analyses, the correlation scheme established in Ref. [60] is used.

Uncertainties in the overall cross section of the signal processes are treated as correlated among analysis channels, and among data sets with the center-of-mass energy. The uncertainties relating to missing higher-order corrections, as well as PDF variations are obtained from Ref. [61]. In some of the channels, additional uncertainty contributions relating to signal acceptance modeling are considered. These are treated as uncorrelated between different analysis channels.

The main sources of theoretical modeling uncertainties in the background estimate vary for the different analysis channels. The different analyses preferentially select different phase space regions, and employ different assumptions for the modeling of theoretical uncertainties on transfer factors between different analysis regions. The resulting uncertainties are therefore treated as uncorrelated.

Significant correlations appear in the treatment of experimental uncertainties. The determination of the luminosity estimate is affected by a number of sources of uncertainty, which are assumed to be correlated among all channels, and partially correlated among data sets. Some of the analysis channels share trigger requirements, and the uncertainties in the efficiency of these common triggers are assumed to be correlated among channels and uncorrelated among data sets. Furthermore, analysis channels often share criteria used for identifying b tagged jets, as well as the hadronic decay products of tau leptons. The uncertainties in the efficiencies of such identification criteria are assumed to be correlated between channels using the same criteria in the same data set. Finally, uncertainties in the calibration of the jet energy scale and resolution are treated as correlated between this analysis and the VBF and ggH/boosted VH channels. All other experimental uncertainties are assumed to be uncorrelated between channels.

9.1.4 Results

Maximum-likelihood fits are performed in the individual channels, as well as in the combination of channels. The dependence of the profiled negative log-likelihood functions on the signal strength parameter μ is shown in Fig. 11. In all channels, consistent best-fit values of the signal strengths are found. The best-fit signal strength is found to be $0.07^{+0.04}_{-0.04}$. This value is dominated by the Run-2 results, and specifically the VBF channel. A saturated goodness-of-fit test is performed using the final combined likelihood function, yielding a p value of , indicating that the best-fit signal model is consistent with the data.

Exclusion limits are calculated on the invisible branching fraction of the Higgs boson assuming SM production cross sections. The results are shown in Fig. 12. Values larger than 0.146 (0.084 expected) are excluded at 95% CL. The final combination thus represents a improvement in sensitivity of approximately 20% relative to the most sensitive single channel (Run-2 VBF), and a improvement of approximately 5% relative to the combination of all Run-2 results.

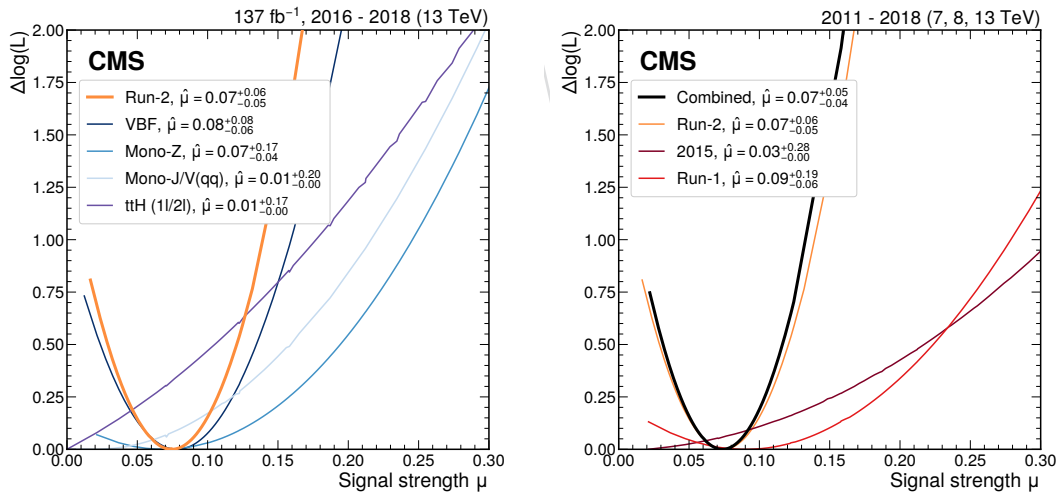


Figure 11: Profiled negative log-likelihood as a function of the signal strength. For each separate curve, the minimum has been subtracted. The left panel shows the scans for the analyses using the 2016–2018 data sets, as well as their combination. In the right panel, scans are shown for the individual data-taking periods, as well as the final combination.

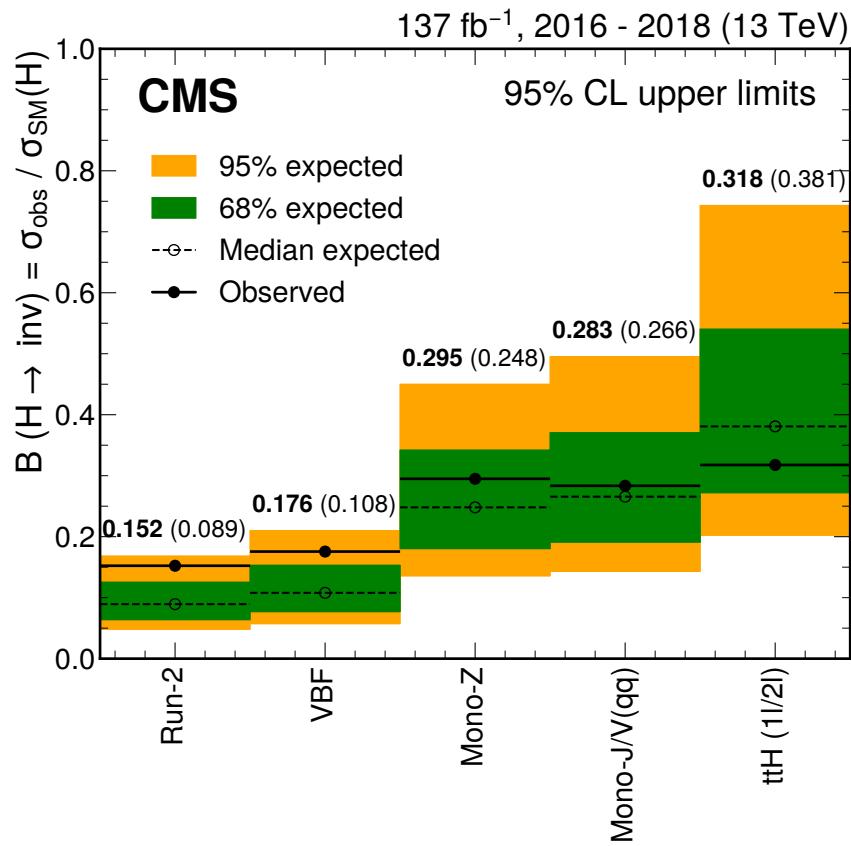


Figure 12: Exclusion limits at 95% CL on the branching fraction of the Higgs boson to invisible final states from the 2016–2018 data sets. The results are shown separately for each channel, as well as combined across channels.

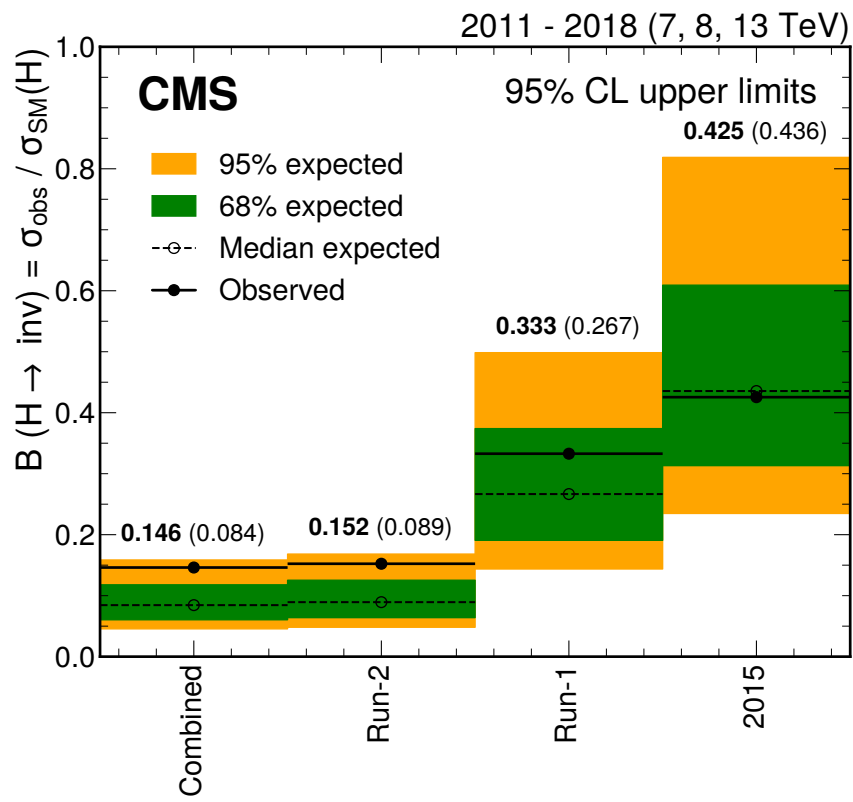


Figure 13: Exclusion limits at 95% CL on the branching fraction of the Higgs boson to invisible final states. The results are shown separately for each data-taking period, as well as combined across periods.

10 Conclusion

A search for invisible decays of a Higgs boson produced through associated production with a top-quark pair or a vector boson is presented. The analysis is based on 138 fb^{-1} of proton-proton data collected at 13 TeV in 2016, 2017 and 2018 by the CMS experiment at the LHC. $t\bar{t}H$ production is investigated using fully hadronic final states with b-jets, including event categories with boosted top quark or vector boson candidates. VH production focusses on a di-jet pair with an invariant mass that is compatible with that of a W or Z boson. Three types of control regions of data, with a photon or one or two leptons, are used directly in the fit to constrain the main SM backgrounds. In the absence of any significant excess in data, a 95% C.L. limit of 0.XX (0.36 expected) is set on the Higgs boson to invisible branching fraction, assuming SM production cross-sections.

These results are combined with those obtained from Higgs production in vector boson fusion and gluon-gluon fusion as well as leptonic $t\bar{t}H$ and ZH decays to obtain a combined limit on the Higgs to invisible branching fraction of 0.XX (0.YY), setting the strongest limit on this quantity to date.

In addition, interpretations of the result in simplified dark matter models with a scalar or pseudo-scalar mediator, accompanied by a pair of top quarks, a jet or a vector boson, are presented and limits set on the cross section of those processes, for a range of scalar and pseudo-scalar mediator masses from 50 to 500 GeV.

Acknowledgments

References

- [1] 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) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [2] 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) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [3] 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) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.
- [4] CMS Collaboration, “Combined measurements of Higgs boson couplings in proton–proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **79** (2019), no. 5, 421, doi:10.1140/epjc/s10052-019-6909-y, arXiv:1809.10733.
- [5] ATLAS Collaboration, “Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment”, *Phys. Rev. D* **101** (2020), no. 1, 012002, doi:10.1103/PhysRevD.101.012002, arXiv:1909.02845.
- [6] Particle Data Group Collaboration, “Review of Particle Physics”, *PTEP* **2020** (2020), no. 8, 083C01, doi:10.1093/ptep/ptaa104.

- [7] G. Bélanger et al., “The MSSM invisible Higgs in the light of dark matter and $g-2$ ”, *Phys. Lett. B* **519** (2001) 93, doi:10.1016/S0370-2693(01)00976-5, arXiv:hep-ph/0106275.
- [8] A. Datta, K. Huitu, J. Laamanen, and B. Mukhopadhyaya, “Linear collider signals of an invisible Higgs boson in theories of large extra dimensions”, *Phys. Rev. D* **70** (2004) 075003, doi:10.1103/PhysRevD.70.075003, arXiv:hep-ph/0404056.
- [9] D. Dominici and J. F. Gunion, “Invisible Higgs decays from Higgs-graviscalar mixing”, *Phys. Rev. D* **80** (2009) 115006, doi:10.1103/PhysRevD.80.115006, arXiv:0902.1512.
- [10] S. Baek, P. Ko, W.-I. Park, and E. Senaha, “Higgs portal vector dark matter: revisited”, *JHEP* **05** (2013) 036, doi:10.1007/JHEP05(2013)036, arXiv:1212.2131.
- [11] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, “Implications of LHC searches for Higgs–portal dark matter”, *Phys. Lett. B* **709** (2012) 65, doi:10.1016/j.physletb.2012.01.062, arXiv:1112.3299.
- [12] A. Djouadi, A. Falkowski, Y. Mambrini, and J. Quevillon, “Direct detection of Higgs–portal dark matter at the LHC”, *Eur. Phys. J. C* **73** (2013) 2455, doi:10.1140/epjc/s10052-013-2455-1, arXiv:1205.3169.
- [13] A. Beniwal et al., “Combined analysis of effective Higgs–portal dark matter models”, *Phys. Rev. D* **93** (2016) 115016, doi:10.1103/PhysRevD.93.115016, arXiv:1512.06458.
- [14] R. E. Shrock and M. Suzuki, “Invisible decays of Higgs bosons”, *Phys. Lett. B* **110** (1982) 250, doi:10.1016/0370-2693(82)91247-3.
- [15] S. Kanemura, S. Matsumoto, T. Nabeshima, and N. Okada, “Can WIMP Dark Matter overcome the Nightmare Scenario?”, *Phys. Rev. D* **82** (2010) 055026, doi:10.1103/PhysRevD.82.055026, arXiv:1005.5651.
- [16] D. Abercrombie et al., “Dark Matter benchmark models for early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum”, *Phys. Dark Univ.* **27** (2020) 100371, doi:10.1016/j.dark.2019.100371, arXiv:1507.00966.
- [17] ATLAS, CMS Collaboration, “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.
- [18] ATLAS Collaboration, “Combination of searches for invisible Higgs boson decays with the ATLAS experiment”, *Phys. Rev. Lett.* **122** (2019), no. 23, 231801, doi:10.1103/PhysRevLett.122.231801, arXiv:1904.05105.
- [19] CMS Collaboration, “Search for invisible decays of a Higgs boson produced via vector boson fusion with 138 fb⁻¹ of proton-proton collisions at $\sqrt{s} = 13$ TeV”, technical report, CERN, Geneva, 2021.
- [20] ATLAS Collaboration, “Search for an invisibly decaying Higgs boson or dark matter candidates produced in association with a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Lett. B* **776** (2018) 318–337, doi:10.1016/j.physletb.2017.11.049, arXiv:1708.09624.

- [21] CMS Collaboration, “Search for dark matter produced in association with a leptonically decaying Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **81** (2021), no. 1, 13, doi:10.1140/epjc/s10052-020-08739-5, arXiv:2008.04735.
[Erratum: *Eur.Phys.J.C* 81, 333 (2021)].
- [22] ATLAS Collaboration, “Search for new phenomena in events with an energetic jet and missing transverse momentum in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector”, *Phys. Rev. D* **103** (2021), no. 11, 112006, doi:10.1103/PhysRevD.103.112006, arXiv:2102.10874.
- [23] CMS Collaboration, “Search for new particles in events with energetic jets and large missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **11** (2021) 153, doi:10.1007/JHEP11(2021)153, arXiv:2107.13021.
- [24] .
- [25] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period”, Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
- [26] CMS Collaboration Collaboration, “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV”, Technical Report CMS-PAS-LUM-17-004, CERN, Geneva, 2018.
- [27] CMS Collaboration Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, Technical Report CMS-PAS-LUM-18-002, CERN, Geneva, 2019.
- [28] C. Oleari, “The powheg box”, *Nuclear Physics B - Proceedings Supplements* **205-206** (Aug, 2010) 3641, doi:10.1016/j.nuclphysbps.2010.08.016.
- [29] 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”, *Journal of High Energy Physics* **2014** (Jul, 2014) doi:10.1007/jhep07(2014)079.
- [30] J. Alwall et al., “Madgraph 5: going beyond”, *Journal of High Energy Physics* **2011** (Jun, 2011) doi:10.1007/jhep06(2011)128.
- [31] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG: s- and t-channel contributions”, arXiv:0907.4076.
- [32] A. M. Sirunyan et al., “Extraction and validation of a new set of CMS pythia8 tunes from underlying-event measurements”, *The European Physical Journal C* **80** (Jan, 2020) doi:10.1140/epjc/s10052-019-7499-4.
- [33] S. Agostinelli et al., “Geant4 a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250–303, doi:https://doi.org/10.1016/S0168-9002(03)01368-8.
- [34] R. D. Ball et al., “Parton distributions from high-precision collider data”, *The European Physical Journal C* **77** (Oct, 2017) doi:10.1140/epjc/s10052-017-5199-5.
- [35] H. Hartanto, B. Jger, L. Reina, and D. Wackeroth, “Higgs boson production in association with top quarks in the powheg box”, *Physical Review D* **91** (May, 2015) doi:10.1103/physrevd.91.094003.

- [36] P. Nason and C. Oleari, “Nlo higgs boson production via vector-boson fusion matched with shower in powheg”, *Journal of High Energy Physics* **2010** (Feb, 2010) doi:10.1007/jhep02(2010)037.
- [37] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, “Hw /hz + 0 and 1 jet at nlo with the powheg box interfaced to gosam and their merging within minlo”, *Journal of High Energy Physics* **2013** (Oct, 2013) doi:10.1007/jhep10(2013)083.
- [38] E. Bagnaschi, G. Degrandi, P. Slavich, and A. Vicini, “Higgs production via gluon fusion in the powheg approach in the sm and in the mssm”, *Journal of High Energy Physics* **2012** (Feb, 2012) doi:10.1007/jhep02(2012)088.
- [39] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, “Top-pair production and decay at nlo matched with parton showers”, *Journal of High Energy Physics* **2015** (Apr, 2015) doi:10.1007/jhep04(2015)114.
- [40] E. Re, “Single-top wt-channel production matched with parton showers using the powheg method”, *The European Physical Journal C* **71** (Feb, 2011) doi:10.1140/epjc/s10052-011-1547-z.
- [41] T. Melia, P. Nason, R. Rentsch, and G. Zanderighi, “W+z-, wz and zz production in the powheg box”, *Journal of High Energy Physics* **2011** (Nov, 2011) doi:10.1007/jhep11(2011)078.
- [42] A. Albert et al., “Recommendations of the LHC dark matter working group: Comparing LHC searches for heavy mediators of dark matter production in visible and invisible decay channels”, 2017.
- [43] J. Nash and CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *Journal of Instrumentation* **12** (2017), no. 10, doi:10.1088/1748-0221/12/10/P10003.
- [44] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Physics Letters B* **659** (2008), no. 1, 119–126, doi:https://doi.org/10.1016/j.physletb.2007.09.077.
- [45] T. C. collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7\text{TeV}$ ”, doi:10.1088/1748-0221/7/10/p10002.
- [46] “Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\text{ TeV}$ ”, doi:10.1088/1748-0221/10/08/p08010.
- [47] “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\text{ TeV}$ ”, doi:10.1088/1748-0221/10/06/p06005.
- [48] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-kt jet clustering algorithm”, doi:10.1088/1126-6708/2008/04/063.
- [49] V. Khachatryan et al., “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 tev”, *Journal of Instrumentation* **12** (Feb, 2017) P02014P02014, doi:10.1088/1748-0221/12/02/p02014.
- [50] CMS Collaboration Collaboration, “Heavy flavor identification at CMS with deep neural networks”,.

- [51] CMS Collaboration Collaboration, “Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques”, Technical Report CMS-PAS-JME-18-002, Geneva, 2020.
- [52] 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.
- [53] T. Sakuma, H. Flaecher, and D. Smith, “Alternative angular variables for suppression of qcd multijet events in new physics searches with missing transverse momentum at the lhc”, 2019.
- [54] CMS Collaboration, “Search for invisible decays of a Higgs boson produced via vector boson fusion with $138fb^{-1}$ of proton-proton collisions at $\sqrt{s} = 13$ TeV”, Technical Report CMS-PAS-HIG-20-003, CERN, Geneva.
- [55] J. M. Lindert et al., “Precise predictions for $v + \text{jets}$ dark matter backgrounds”, *The European Physical Journal C* **77** (Dec, 2017) doi:10.1140/epjc/s10052-017-5389-1.
- [56] CMS Collaboration, “Search for dark matter produced in association with a leptonically decaying Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **81** (2021), no. 1, 13, doi:10.1140/epjc/s10052-020-08739-5, arXiv:2008.04735. [Erratum: Eur.Phys.J.C 81, 333 (2021)].
- [57] CMS Collaboration, “Search for direct top squark pair production in events with one lepton, jets, and missing transverse momentum at 13 tev with the cms experiment”, *JHEP* **05** (2020) 032, doi:10.1007/JHEP05(2020)032, arXiv:1912.08887.
- [58] CMS Collaboration, “Search for top squark pair production using dilepton final states in pp collision data collected at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **81** (2021) 3, doi:10.1140/epjc/s10052-020-08701-5, arXiv:2008.05936.
- [59] CMS Collaboration, “Combined searches for the production of supersymmetric top quark partners in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **81** (2021) 970, doi:10.1140/epjc/s10052-021-09721-5, arXiv:2107.10892.
- [60] CMS Collaboration, “Searches for invisible decays of the higgs boson in pp collisions at $\sqrt{s} = 7, 8,$ and 13 TeV”, *JHEP* **02** (2017) 135, doi:10.1007/JHEP02(2017)135, arXiv:1610.09218.
- [61] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector”, CERN Report CERN-2017-002-M, 2016. doi:10.23731/CYRM-2017-002, arXiv:1610.07922.