Search for invisible decays of the Higgs boson produced via vector boson fusion in proton-proton collisions at $\sqrt{s}=13$ TeV

A. Tumasyan et al.*
(CMS Collaboration)

(Received 27 January 2022; accepted 14 March 2022; published 20 May 2022)

A search for invisible decays of the Higgs boson produced via vector boson fusion (VBF) has been performed with 101 fb$^{-1}$ of proton-proton collisions delivered by the LHC at $\sqrt{s}=13$ TeV and collected by the CMS detector in 2017 and 2018. The sensitivity to the VBF production mechanism is enhanced by combining two analysis categories, one based on missing transverse momentum and a second based on the properties of jets. In addition to control regions with Z and W boson candidate events, a highly populated control region, based on the production of a photon in association with jets, is used to constrain the dominant irreducible background from the invisible decay of a Z boson produced in association with jets. The results of this search are combined with all previous measurements in the VBF topology, based on data collected in 2012 (at $\sqrt{s}=8$ TeV), 2015, and 2016, corresponding to integrated luminosities of 19.7, 2.3, and 36.3 fb$^{-1}$, respectively. The observed (expected) upper limit on the invisible branching fraction of the Higgs boson is found to be 0.18 (0.10) at the 95% confidence level, assuming the standard model production cross section. The results are also interpreted in the context of Higgs-portal models.

DOI: 10.1103/PhysRevD.105.092007

I. INTRODUCTION

A particle compatible with the standard model (SM) Higgs boson (H) [1–6] was discovered at the CERN LHC in 2012 [7–9]. Since then, extensive studies of this particle have been performed with data taken at $\sqrt{s} = 7, 8, \text{and } 13 \text{ TeV}$, in particular to understand how it couples to other SM particles.

In the SM, the branching fraction to invisible final states, $\mathcal{B}(H \to \text{inv})$, is only about 0.1% [10], from the decay of the Higgs boson via $ZZ^* \to 4\nu$. Several theories beyond the SM, however, predict much higher values of $\mathcal{B}(H \to \text{inv})$ (see Refs. [11–14], as well as Ref. [15] and references therein). In particular, in Higgs portal models, the Higgs boson acts as the mediator between SM particles and dark matter (DM) [16–19], strongly enhancing $\mathcal{B}(H \to \text{inv})$.

Direct searches for $H \to \text{inv}$ decays have already been performed by the ATLAS [20,21] and CMS [22–24] collaborations using data collected at $\sqrt{s} = 7, 8, \text{and } 13 \text{ TeV}$ and combining the three main Higgs boson production modes, namely gluon-gluon fusion ($ggH$), production of a Higgs boson in association with vector bosons ($VH$, with $V = W^\pm$ or $Z$), and vector boson fusion (VBF). Assuming SM production of the Higgs boson, the best observed (expected) 95% confidence level (C.L.) upper limits on $\mathcal{B}(H \to \text{inv})$ are set at 0.19 (0.15) by CMS, using data collected at $\sqrt{s} = 7, 8, \text{and } 13 \text{ TeV}$, and at 0.26 (0.17) by ATLAS, using data collected at 13 TeV. In both cases, the data at 13 TeV were collected in 2016. Combining the latest CMS constraints on both visible and invisible decays within the κ framework [25], the upper bound on $\mathcal{B}(H \to \text{inv})$ is 0.22 at the 95% C.L., using only the dataset collected at 13 TeV in 2016.

Thanks to its large production cross section [26] and distinctive event topology, the VBF production mechanism drives the overall sensitivity in the direct search for invisible decays of the Higgs boson. This paper focuses exclusively on the search for $H \to \text{inv}$ in the VBF production mode using the LHC proton-proton ($pp$) collision dataset collected during 2017–2018, corresponding to an integrated luminosity of up to 101 fb$^{-1}$, and on the combination of this search with analyses performed on previous datasets [22,27].

Employing a strategy similar to the one used in the previously published analysis [22], the invariant mass of the jet pair produced by VBF, $m_{jj}$, is used as a discriminating variable to separate the signal and the dominant backgrounds arising from vector boson production in association with two jets ($V + \text{jets}$). Representative Feynman diagrams for the signal and main background processes are shown in Fig. 1.

Control regions (CRs) enriched in $V + \text{jets}$ processes are used to constrain the associated background contributions.
FIG. 1. Leading-order Feynman diagrams for the production of the Higgs boson in association with two jets from VBF (left) and representative leading-order Feynman diagrams for the production of a Z boson in association with two jets either through VBF production (middle) or strong production (right). Diagrams for the production of a W boson in association with two jets are similar.

in the signal region. Additional sensitivity is obtained by using $\gamma + \text{jets}$ events to further constrain the $Z(\nu\bar{\nu})$ background. In the previous CMS publication, the trigger strategy was based exclusively on the invisible Higgs boson decay products, requiring a high threshold on the missing transverse momentum. With the availability of a trigger based on the jet properties from VBF production, in this analysis, additional sensitivity is achieved by including events with lower missing transverse momentum.

This article is organized as follows. Section II introduces the CMS detector. Section III summarizes the data and simulated samples. The event reconstruction is detailed in Sec. IV, followed by the analysis strategy in Sec. V. Section VI describes the systematic uncertainties. Finally, the results are presented in Sec. VII, with tabulated versions provided in HEPData [28], followed by a summary in Sec. VIII.

II. CMS DETECTOR

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two end cap sections. Hadron forward (HF) steel and quartz fiber calorimeters extend the pseudorapidity $\eta$ coverage provided by the barrel and end cap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [29].

Events of interest are selected using a two-tiered trigger system [30]. The first level (L1) is composed of custom hardware processors, which use information from the calorimeters and muon detectors to select events at a rate of about 100 kHz [31]. The second level, known as the high-level trigger (HLT), is a software-based system that runs a version of the full event reconstruction optimized for fast processing, reducing the event rate to about 1 kHz.

At the end of 2016, the first part of the CMS detector upgrade program (Phase I) was undertaken, with the replacement of the inner tracking pixel detector and the L1 trigger system. During the 2016 and 2017 data-taking periods, partial mistiming of signals in the forward region of the ECAL end caps ($2.5 < |\eta| < 3.0$) led to a large reduction in the L1 trigger efficiency [31]. A separate correction was determined using an unbiased data sample and applied to simulated events to reproduce the loss of efficiency. This problem was resolved before the 2018 data-taking period.

III. DATA AND SIMULATED SAMPLES

Data were recorded by several triggers, as detailed in Sec. VA, during 2017 and 2018, for maximum integrated luminosities corresponding to 41.5 and 59.8 fb$^{-1}$, respectively.

The signal and background processes are simulated using similar Monte Carlo (MC) generator configurations as described in detail in Ref. [22] and summarized below. Separate independent samples were produced for each data-taking year. The same generator settings were used for the 2017 and 2018 samples.

The Higgs boson signal events, produced through $ggH$, VBF, $VH$ and in association with top quarks ($t\bar{t}H$), are generated with POWHEG2.0 [32–36] at next-to-leading-order (NLO) approximation in perturbative quantum chromodynamics (pQCD). The signal yields are normalized to the inclusive Higgs boson production cross sections, calculated in Ref. [26] at approximate next-to-NLO (NNLO) in pQCD, with NLO electroweak (EW) corrections. For the VBF production process, an additional event weight is applied to the simulated events to account for EW NLO effects, dependent on the boson transverse momentum ($p_T$). The correction factor is determined using HAWK [37] and parametrized as $(1 - 0.000372 p_T/\text{GeV} - 0.0304)/0.95$, with the numerator accounting for the full leading-order (LO)-to-NLO correction, and the denominator representing the overall normalization effect of the EW correction. The latter is already included in the inclusive cross section and therefore has to be removed here to avoid double counting.
The $Z/\gamma^*(\ell^+\ell^-) + \text{jets}$, $Z(\nu\bar{\nu}) + \text{jets}$, and $W(\ell\nu) + \text{jets}$ (with $\ell = e, \mu, \tau$) processes are simulated at NLO in pQCD using MadGraph 5_aMC@NLO2.6.5 [38], with the five-flavor scheme and the FXFX [39] merging scheme, in several bins of boson $p_T$. Up to two additional partons are included in the final state in the matrix element calculations. These processes are referred to as $V + \text{jets}$ (strong) in what follows.

The $\gamma + \text{jets}$ background is simulated at LO in pQCD using MadGraph 5_aMC@NLO2.4.2 [38], with up to four partons in the final state included in the matrix element calculations [40]. We refer to this process as $\gamma + \text{jets}$ (strong) in the rest of this paper.

The MadGraph 5_aMC@NLO generator is also used for the production of a vector boson, or a photon, in association with two jets exclusively through EW interactions at LO. These are referred to as $V + \text{jets}$ (VBF) and $\gamma + \text{jets}$ (VBF), respectively, in the following.

The LO simulation for the $\gamma + \text{jets}$ (strong) process is corrected using boson $p_T$- and $m_j$-dependent NLO pQCD $K$-factors derived with MadGraph 5_aMC@NLO2.4.2 [38]. The simulations for $\gamma + \text{jets}$ (strong) and $V + \text{jets}$ (strong) processes are also corrected as a function of boson $p_T$ with NLO EW $K$-factors derived in Ref. [41]. Similarly, $V + \text{jets}$ (VBF) processes are corrected with NLO pQCD $K$-factors derived using the VBFNLO2.7.0 event generator [42,43], as functions of boson $p_T$ and $m_j$. For the $\gamma + \text{jets}$ (VBF) process, the NLO pQCD corrections, evaluated using MadGraph 5_aMC@NLO, are found to be negligible.

Samples of QCD multijet events are generated at LO using MadGraph 5_aMC@NLO. The $t\bar{t}$ and single top quark background samples are produced at NLO in pQCD using POWHEG2.0 and POWHEG1.0, respectively [44–46]. Samples of $WZ$ and $ZZ$ events are simulated at LO with PYTHIA8.205 [47], while the $V\gamma$ and $WW$ processes are simulated at NLO in pQCD using MadGraph 5_aMC@NLO and POWHEG [48], respectively.

The NNPDF3.1 [49] NNLO parton distribution functions (PDFs) are used for all the matrix element calculations. All generators are interfaced with PYTHIA8.205 for the parton shower simulation, hadronization, and fragmentation processes. The underlying event description uses the CMS Pythia8 parameter tune 5 (cp5) parameter tune [50].

Interactions of the final-state particles with the CMS detector are simulated using GEANT4 [51]. Additional $pp$ interactions (pileup) are included in the simulation, and simulated events are weighted to reproduce the pileup distribution observed in data, separately for each data-taking year. The average number of pileup vertices is 32 in the 2017 and 2018 data.

IV. EVENT RECONSTRUCTION

The event reconstruction and object definitions closely follow those of the previous publication [22]. The main aspects are summarized below.

A global event description is available using the particle-flow (PF) algorithm [52]. Using a combination of the information provided by the tracker, calorimeters, and muon systems, the PF algorithm aims to reconstruct individual particles (PF candidates), classifying them as electrons, photons, muons, or charged and neutral hadrons. The final state for this analysis is composed solely of jets from gluons or light-flavored quarks, and missing transverse momentum. We employ explicit vetoes on events containing all other identified types of objects (electrons, photons, muons, hadronically decaying $\tau$ leptons, heavy-flavored jets), which help to reject background processes with leptonic decays ($W$ and $Z$ bosons), and those containing top quarks.

Electron and photon candidates [53] are selected in the range $|\eta| < 2.5$, while muon [54] candidates are selected with $|\eta| < 2.4$. When considered for event vetoes, candidates are required to satisfy loose identification and isolation criteria. These requirements ensure genuine leptons and photons are discarded with high efficiency. For electrons, the loose working point is referred as “loose” (“veto” in Ref. [53]) and has $\approx 95\%$ efficiency. For photons and muons, the loose working point corresponds to efficiencies of $\approx 90\%$ and $>99\%$, respectively. The $p_T$ threshold on loose objects is set to 10 GeV for electrons and muons and 15 GeV for photons. When leptons (photons) are explicitly selected to enhance the contributions from $V + \text{jets}$ ($\gamma + \text{jets}$) processes, which is done to populate control regions in data, “tight” identification and isolation criteria are required. These enhance the purity at the price of lower efficiency ($\approx 70\%$ for electrons and photons, $\approx 96\%$ for muons). The $p_T$ thresholds are then set to 20 GeV for the leading muon and higher values for the leading electron (40 GeV) and photon (230 GeV) because of trigger requirements. The subleading electron or muon is required to have $p_T > 10$ GeV. In 2018, a section of the hadron calorimeter end cap (HE) was not functional for part of the year, leading to the inability to properly identify electrons and photons in the region $\eta < -1.39$ and azimuthal angle $-1.6 < \phi < -0.9$. For data collected during this time, specific electron and photon selection criteria are applied. These are described in more detail in Sec. V B.

Jets are reconstructed by clustering all PF candidates associated with the primary interaction vertex using the anti-$k_T$ clustering algorithm [55], with a distance parameter of 0.4, as implemented in the FastJet package [56]. The candidate vertex with the largest value of summed physics object $p_T^V$ is taken to be the primary $pp$ interaction vertex. The physics objects used for this determination are the jets, clustered using the jet finding algorithm [55,56] with the tracks assigned to candidate vertices as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the $p_T$ of those jets. Pileup mitigation techniques [57] are used to correct the objects for energy deposits belonging to pileup vertices, as well as
to remove objects not associated with the primary interaction vertex. Loose identification criteria are applied on the jet composition to remove contributions from calorimeter noise. To correct the average measured energy of the jets to that of particle-level jets, jet energy corrections are derived using simulated events, as a function of the reconstructed jet \( p_T \) and \( \eta \). \textit{In situ} measurements of the momentum balance in dijet, \( \gamma + \text{jet}, \ Z + \text{jet}, \) and multijet events are used to determine any residual differences between the jet energy scale in data and in simulation, and appropriate corrections are made [58]. In simulated events, the jet energy is also smeared to reproduce the jet energy resolution measured in the data [58]. For jets with \( p_T < 50 \text{ GeV} \), a multivariate discriminant against pileup jets is applied, using a loose working point [59]. Jets are selected in the range \( |\eta| < 4.7 \) and with \( p_T > 30 \text{ GeV} \). Jets with an identified electron, muon, or photon within \( \Delta R < 0.4 \) are rejected, where \( \Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \).

The missing transverse momentum vector \( \vec{p}_T^{\text{miss}} \) is computed as the negative vector \( \vec{p}_T \) sum of all the PF candidates in an event, and its magnitude is denoted as \( p_T^{\text{miss}} \). Any correction applied to individual objects is propagated correspondingly to the \( p_T^{\text{miss}} \) [60]. Specific event filters have been designed to reduce the contamination arising from large misreconstructed \( p_T^{\text{miss}} \) from non-collision backgrounds [60]. For the analysis of the 2018 data, the missing HE section affects the PF reconstruction. The inability to distinguish electrons and photons from jets data, the missing HE section affects the PF reconstruction. The missing HE section affects the PF reconstruction.

The distinctive feature of VBF production is a pair of jets \( (\eta j j) \) and therefore a large \( \Delta R \) or narrow in both directions. They lead to spurious \( p_T^{\text{miss}} \) in the opposite direction in \( \phi \). Events are hence rejected if they contain any jet with \( |\eta| > 2.99 \), \( p_T > 80 \text{ GeV} \), and \( \Delta \phi(p_T^{\text{miss}}, p_T^{\text{jet}}) > 2.5 \) that does not satisfy the criteria summarized in Table I. The requirements of this selection are chosen to have a mistagging rate smaller than 10% for signal-like jets, while being more than 60%–90% efficient at rejecting noisilke jets, depending on their \( p_T \) and \( \eta \). To correct for mismodeling of these selections in simulation, the selection efficiency on signal-like jets is measured in both data and simulated \( Z(\ell\ell) + j e t \) and \( \gamma + j e t \) events, and scale factors are applied to correct the simulation. The scale factors are measured as functions of jet \( p_T \) and \( \eta \) and are consistent with unity to within 10%.

Hadronically decaying \( \tau \) leptons (\( \tau_h \)) are identified from reconstructed jets through the multivariate DeepTau algorithm [62], using a working point that has an average efficiency above 96% for a jet misidentification rate of less than 10%. The \( \tau_h \) candidates are reconstructed with \( p_T > 20 \text{ GeV} \) and \( |\eta| < 2.3 \). Jets with an identified loose electron or muon within \( \Delta R < 0.4 \) are rejected before applying the DeepTau algorithm.

The specific features of heavy-flavored jets, in particular the presence of displaced vertices, are used in a multivariate jet tagging method. The “medium” working point of the DeepCSV algorithm from Ref. [63] is used to tag b quark jets with \( p_T > 20 \text{ GeV} \) and \( |\eta| < 2.4 \) with 68% efficiency and 1.1 (12%) probability of misidentifying a light-flavor or gluon (c quark) jet as a bottom quark jet.

### V. ANALYSIS STRATEGY

The distinctive feature of VBF production is a pair of jets originating from light-flavor quarks, with a large separation in \( \eta \) (\( \Delta \eta_j \)) and therefore a large \( m_{jj} \). The signal region (SR)

| Observable | \( 2.99 < |\eta_{\text{jet}}| < 4.00 \) | \( 4.00 < |\eta_{\text{jet}}| < 5.00 \) |
|------------|---------------------------------|---------------------------------|
| \( \sigma_{\eta} - \sigma_{\phi} \) | \( <0.02 \) | \( ... \) |
| \( \sigma_{\eta} \) | \( >0.02 \) | \( <0.10 \) |
| \( \sigma_{\phi} \) | \( ... \) | \( >0.02 \) |
| \( \Delta \eta_{\text{jet}} \) | \( <3 \) | \( ... \) |
in this analysis uses selection requirements on the jet pair together with the presence of a significant amount of $p_T^{\text{miss}}$.

The shape of the $m_{jj}$ distribution is used to disentangle jet pairs produced in VBF production from other SM processes. When fitting the shape of this distribution, the strong production of the $V + \text{jets}$ processes together with the $ggH$ signal dominate at low $m_{jj}$, whereas the VBF-produced $V + \text{jets}$ processes populate the high-$m_{jj}$ tail, together with the VBF $H$ signal. The shapes of $m_{jj}$, $|\Delta \eta_{jj}|$, and the dijet separation in azimuthal angle ($|\Delta \phi_{jj}|$) predicted by the simulation are compared between strong and VBF production of both $V + \text{jets}$ and signal processes in Fig. 2. Whereas similarities are seen between the VBF production of vector bosons and Higgs bosons, there are some differences. First, the respective bosons have different coupling structures, identified as the main reason for the different behavior in the $\Delta \phi_{jj}$ distribution [64]. Second, the $V + \text{jets}$ (VBF) samples include additional diagrams compared with the VBF $H$ production, in which the vector boson is produced through a coupling to quarks and jets are produced from additional EW vertices. For the VBF $H$ production, such diagrams are strongly suppressed through

![Graph](image1.png)

**FIG. 2.** Comparison of shapes of quantities related to the dijet pair. Shown are $m_{jj}$ (upper), $|\Delta \eta_{jj}|$ (lower left), and $|\Delta \phi_{jj}|$ (lower right), as predicted by simulation, separating strong and VBF production for $V + \text{jets}$ and signal processes.
the Yukawa mechanism, also leading to differences in the kinematic behavior.

The dominant $V + $ jets backgrounds are measured using control regions ($W$, $Z$, and $\gamma$ CRs), in which one or more charged leptons (electrons and muons), or a photon, are required, but the selection on the jets and $p_T^{\text{miss}}$ is kept identical to that in the SR. The MC simulation is used to define transfer factors. These make it possible to predict the $V + $ jets yields in the SR from both the yields measured and predicted in the CRs. The full procedure is described in Sec. VII.

The $Z$ CR suffers from a lack of statistical precision, particularly in the high-$m_{jj}$ region, due to the low branching fraction of the $Z$ boson to a pair of leptons. Because of their similarities, the ratios of the yields of $W^\pm$ or $\gamma$ to $Z$ production in the SR are constrained, within theoretical uncertainties, to those predicted by the simulation.

In the following subsections, the SR selection is first presented. Then, the data-driven methods used to estimate the $V + $ jets and QCD multijet backgrounds are described. The remaining small contributions expected from diboson and top quark processes are estimated using simulation.

A. Signal region selection

Two complementary trigger strategies were used to select events online. The first category only uses the $p_T^{\text{miss}}$ information at L1 and in the HLT and is referred to as the “missing momentum triggered region” (MTR) category in the following. During 2017 and 2018, the HF region was included in the definition of $p_T^{\text{miss}}$, leading to an increase in trigger acceptance when the VBF jets are reconstructed with $3 \leq |\eta| \leq 5$, compared with 2016. The $p_T^{\text{miss}}$ thresholds varied between 65 and 90 GeV, with a $p_T^{\text{miss}}_{\text{L1,HLT}}$ threshold of 120 GeV. After correcting for the L1 mistiming inefficiency described in Sec. II, this trigger is more than 90% efficient for $p_T^{\text{miss}} > 250$ GeV. Loose muon candidates are ignored at L1 and in the HLT when calculating the $p_T^{\text{miss}}_{\text{L1}}$ and $p_T^{\text{miss}}_{\text{HLT}}$ variables, ensuring that the same trigger can be used in the muon CRs. The data collected by these triggers in 2017 and 2018 correspond to integrated luminosities of 41.5 and 59.8 fb$^{-1}$, respectively.

The second category uses a combination of $p_T^{\text{miss}}$ and jet information and is referred to as the “VBF jets triggered region” (VTR) category in the following. After the upgrade of the L1 trigger system [65], it was possible to develop an algorithm targeting jets originating from VBF production. This trigger was added after the start of data taking in 2017 and collected a dataset corresponding to an integrated luminosity of 36.7 fb$^{-1}$ during that year. The L1 VBF algorithm requires the presence of at least two jets passing $p_T^{\text{jet}}$ thresholds of 115 (110) GeV for the leading jet and 40 (35) GeV for the subleading jet in 2017 (2018). Pairs are formed from the selected jets, and the pair with the highest invariant mass $m_{jj}^{1,1}$ must satisfy $m_{jj}^{1,1} > 620$ GeV. The pair is allowed to be formed by two jets with $p_T^{\text{jet}} > 40(35)$ GeV if there is a third jet passing the leading-$p_T^{\text{jet}}$ threshold requirement, in 2017 (2018). A corresponding HLT algorithm was designed specifically for the $H \to \text{inv}$ analysis, adding a requirement on $p_T^{\text{miss}}_{\text{L1,HLT}}$ with a minimum threshold of 110 GeV. At the analysis level, after correcting for the L1 mistiming inefficiency, this trigger is more than 85% efficient for $p_T^{\text{miss}} > 160$ GeV and the leading-$m_{jj}$ jet pair passing the following requirements: $m_{jj} > 900$ GeV, with $p_T$ thresholds on the two jets forming the leading-$m_{jj}$ pair of $p_T > 140, 70$ GeV. The trigger efficiencies for the two trigger algorithms are available in the Supplemental Material, additional figures, and tables [66]. Again, loose muon candidates are ignored when calculating the $p_T^{\text{miss}}$ variables at all stages.

We ensure the MTR and VTR categories are orthogonal by requiring $160 < p_T^{\text{miss}} \leq 250$ GeV in the VTR category and $p_T^{\text{miss}} > 250$ GeV in the MTR category.

To enhance the selection of jets with VBF properties at the analysis level, and to reduce the contamination arising from jet pairs in QCD multijet events, the two leading-$p_T$ jets (or the jets forming the highest-$m_{jj}$ pair in the VTR category) are required to be in opposite hemispheres of the detector ($\eta_1 \eta_2 < 0$). The two selected jets are also required to have $|\Delta \eta_{jj}| > 1$ and $|\Delta \phi_{jj}| < 1.5$ ($|\Delta \phi_{jj}| < 1.8$ in the VTR category). In the MTR category, the $m_{jj}$ threshold is set to 200 GeV to use the full shape of the spectrum to better separate the signal from the background formed by strong $V + $ jets production.

In QCD multijet events, large $p_T^{\text{miss}}$ may arise from mismeasurements of the jet momenta, in which case some jets in the event could be aligned in $\phi$ with the $\vec{p}_T^{\text{miss}}$. To reduce the contamination from such events, the minimum value of the azimuthal angle between the $\vec{p}_T^{\text{miss}}$ vector and any of the first four leading jets ($p_T > 30$ GeV), $\min(\Delta \phi(\vec{p}_T^{\text{miss}}, \vec{p}_T^{\text{jet}}))$, is required to be above 0.5 (1.8) in the MTR (VTR) category. Events with possible mismeasurements due to calorimeter noise, which would lead to jets with anomalously large (small) energy fractions coming from neutral (charged) particles, are rejected. This is done by rejecting the event if either of the selected VBF jets has $|\eta| < 2.5$ and a neutral hadron energy fraction (NHEF) [charged hadron energy fraction (CHEF)] NHEF > 0.8 and CHEF < 0.1. This selection rejects at most 2% of events, independent of the process and uniformly in $m_{jj}$. A criterion is also applied on the difference between the $p_T^{\text{miss}}$ measured using the PF algorithm and that using only the calorimeters. This difference is required to be less than 50% of the $p_T^{\text{miss}}$. This selection rejects at most 2 (1)% of all events, mostly at low $m_{jj}$, for the 2017 (2018) data-taking conditions.

A second source of QCD multijet background is due to the remaining impact of jets originating from HF noise, where the $\vec{p}_T^{\text{miss}}$ is balanced in $\phi$ with such a jet, and the jet...
still passes the selection criteria from Table I. Combined with genuine jets from QCD multijet production, such events can pass the SR selection. These large energy deposits are generally close to the outer HF boundary ($|\eta| < 3.25$), where the readout boxes are located, though can extend up to $|\eta| = 5$.

Finally, a veto on all other types of loosely identified objects (electrons, muons, photons, $\tau_h$ candidates, and $b$-tagged jets), as described in Sec. IV, is applied.

The criteria for the SR selections are summarized in Table II, for the MTR and VTR categories. After these selections, the dominant backgrounds come from the $V + j$ets processes. Due to the large branching fraction of the $Z$ boson decay to neutrinos, the $Z(\nu\bar{\nu}) + j$ets process accounts for about two-thirds of the total background. After the lepton vetoes and the $p_T^{\text{miss}}$ requirement, the contributions from other decay modes are negligible. The next largest background arises from $\nu\bar{\nu}$ + jets process in which the charged lepton from the $Z^{\pm}$ boson decay is outside of the acceptance of the tracking detector, leading to additional $p_T^{\text{miss}}$. In the case of muons, which deposit very little energy in the calorimeters, the $p_T^{\text{miss}}$ is significant. The hadronic decay modes of the $W^{\pm}$ boson are rejected by the large $p_T^{\text{miss}}$ requirement. The VBF production of $V + j$ets contributes about 2% of the total $V + j$ets background for $m_{jj} \approx 1.5$ TeV, and to more than 48% for $m_{jj} > 3.5$ TeV.

### B. Lepton-based control regions

As the boson recoil properties are driven by the production mode and are independent of the boson decay mode, the dominant $Z(\nu\bar{\nu}) + j$ets background is modeled using CRs with leptonic decays of the $Z$ boson [$Z(ee)$ and $Z(\mu\mu)$]. To reduce the contribution from Drell-Yan $\gamma^*$ decays to leptons, the invariant mass of the selected leptons is required to lie in the range 60–120 GeV. The lepton selection is chosen to maximize the event yield while still ensuring leptonic $Z$ boson decays are selected with high purity.

To stay as close as possible to the SR selection, the same trigger as in the SR is used for the $Z(\mu\mu)$ CR. As a result, systematic uncertainties in the trigger efficiencies largely cancel when estimating the corresponding transfer factors. Instead of the muon veto, a pair of oppositely charged muons, consisting of a tight muon with $p_T > 20$ GeV and a loose muon with $p_T > 10$ GeV, is required. All other criteria from Table II are applied. The $p_T^{\text{miss}}$ variable is recalculated ignoring the muons, to mimic the boson recoil.

For the $Z(ee)$ CR, the triggers used in the SR are inefficient, as electrons deposit their energy in the calorimeter. Single-electron triggers are therefore used. The lowest-threshold trigger requires a minimum $p_T$ of 35 GeV and imposes isolation requirements. It is supplemented by an electron trigger with a $p_T$ threshold of 115 GeV, but no isolation requirements, as well as by a photon trigger requiring $p_T > 200$ GeV. For the last, no isolation criteria are applied, and it does not rely on track reconstruction. Taken together, this set of triggers optimizes the efficiency over the full $p_T$ range. Instead of the electron veto, a pair of oppositely charged electrons, consisting of a tight electron with $p_T > 40$ GeV and a loose electron with $p_T > 10$ GeV, is required. The $p_T^{\text{miss}}$ variable is recalculated ignoring the electrons, to mimic the boson recoil.

For the $W(e\nu)$ and $W(\mu\nu)$ CRs favor $W^{\pm}$ boson decays with a high-$p_T$ central lepton [$|\eta| < 2.5$ (electrons) or 2.4 (muons)], whereas the background expected in the SR consists of $W^{\pm}$ boson decays in which the leptons (including $\tau$ leptons) are

<table>
<thead>
<tr>
<th>Observable</th>
<th>MTR</th>
<th>VTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leading pair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading (subleading) jet</td>
<td>$p_T &gt; 80$ (40) GeV, $</td>
<td>\eta</td>
</tr>
<tr>
<td>$p_T^{\text{miss}}$</td>
<td>$&gt;250$ GeV</td>
<td>$160 &lt; p_T^{\text{miss}} &lt; 250$</td>
</tr>
<tr>
<td>$\Delta \phi (\vec{p}_T^{\text{miss}}, \vec{p}_T^{\text{jet}})$</td>
<td>$&gt;0.5$</td>
<td>$&gt;1.8$</td>
</tr>
<tr>
<td>$</td>
<td>\Delta \phi_l</td>
<td>$</td>
</tr>
<tr>
<td>$m_{jj}$</td>
<td>$&gt;200$ GeV</td>
<td>$&gt;900$ GeV</td>
</tr>
<tr>
<td>$</td>
<td>p_T^{\text{miss}} - \text{calorimeter} p_T^{\text{miss}}</td>
<td>/ p_T^{\text{miss}}$</td>
</tr>
<tr>
<td>Leading/subleading jets $</td>
<td>\eta</td>
<td>&lt; 2.5$</td>
</tr>
<tr>
<td>$\tau_h$ candidates</td>
<td>$N_{\tau_h} = 0$ with $p_T &gt; 20$ GeV, $</td>
<td>\eta</td>
</tr>
<tr>
<td>b quark jet</td>
<td>$N_{b} = 0$ with $p_T &gt; 20$ GeV, \text{deepCSV Medium}</td>
<td></td>
</tr>
<tr>
<td>$\eta_1 \eta_2$</td>
<td>$&lt; 0$</td>
<td>$&lt; 0$</td>
</tr>
<tr>
<td>$</td>
<td>\Delta \eta_{ij}</td>
<td>$</td>
</tr>
<tr>
<td>Electrotons (muons)</td>
<td>$N_{e\nu} = 0$ with $p_T &gt; 10$ GeV, $</td>
<td>\eta</td>
</tr>
<tr>
<td>Photons</td>
<td>$N_{\gamma} = 0$ with $p_T &gt; 15$ GeV, $</td>
<td>\eta</td>
</tr>
</tbody>
</table>
outside of the acceptance. This has an impact on the $p_T^{\text{miss}}$ distribution. As explained in greater detail in Sec. VII, the impact of the lepton acceptance is accounted for in the definition of the transfer factors between the CRs and the SR, using the simulation. For the $W(e\nu)$ and $W(\mu\nu)$ CRs, the lepton veto is replaced by the selection of a tight electron (muon) with $p_T > 40$ (20) GeV and a veto on any other identified loose electron or loose muon. To reduce the contribution from misidentified electrons stemming from QCD multijet production, the $p_T^{\text{miss}}$ associated with the $W(e\nu)$ decay (i.e., not ignoring the electron contribution) is required to be above 80 GeV. In the 2018 analysis, due to the missing HE sector in the data, jets in the corresponding $\eta$-$\phi$ area are often identified as electrons or photons, and hence events with an electron in this region are rejected.

C. $\gamma + \text{jets}$ control region

To further constrain the $Z(\nu\bar{\nu})$ background in the SR, a photon CR is used. At large $p_T$, the kinematic properties of photon production become similar to those of the $Z(\nu\bar{\nu})$ process [41] and can therefore be used to estimate the latter. Events are selected using a trigger requiring an online photon $p_T$ of at least 200 GeV. In the offline analysis, photons are required to be located in the central part of the detector ($|\eta| < 1.4442$), have $p_T > 230$ GeV to ensure full trigger efficiency, and pass additional identification criteria based on the properties of the associated energy deposit (supercluster) in the ECAL, as well as the isolation of the photon relative to nearby objects. Exactly one such photon is required, and all other criteria from Table II are applied. The $p_T^{\text{miss}}$ variable is recalculated ignoring the photon, to mimic the boson recoil.

In addition to the desired $\gamma + \text{jets}$ events with a genuine well-identified and isolated (“prompt”) photon, small contributions from QCD multijet events with hadronic jets misidentified as photons are present in this region (“nonprompt”). To estimate this background contribution, a purity measurement is performed. The measurement is based on the distribution of the lateral width, $\sigma_{\text{inv}}$ [53], of the ECAL supercluster associated with the photon. For prompt photons, the distribution of $\sigma_{\text{inv}}$ peaks sharply around values of 0.01 and below, while nonprompt photons show a much smaller peak and a shoulder toward values larger than 0.01. To extract the contamination, a template fit to the $\sigma_{\text{inv}}$ distribution is performed in data collected with a looser version of the CR selection. In this looser region, instead of the usual selection that is applied to the VBF jets, we require the presence of a single jet with $p_T > 100$ GeV to enhance the available number of events. Additionally, the photon identification criteria are modified by removing a requirement on $\sigma_{\text{inv}}$ that is otherwise included. A template for prompt photons is obtained from simulated $\gamma + \text{jets}$ events, while a nonprompt template is derived from a data sample that is enriched in nonprompt events by inverting the isolation requirements that are part of the photon identification criteria. The nonprompt fraction is defined as the fraction of nonprompt photons present below the $\sigma_{\text{inv}}$ threshold set by the identification criteria. The template fit is performed separately in bins of the photon $p_T$ and yields a nonprompt fraction between around 4% at $p_T = 200$ GeV and 2%–3% at $p_T = 800$ GeV, depending on the data-taking period. The final QCD multijet contribution is then determined by weighting the events observed in the data by the nonprompt fraction. A 25% uncertainty is assigned to the normalization of the QCD multijet background to account for any mismodeling of $\sigma_{\text{inv}}$ in the simulation. The uncertainty is estimated by repeating the measurement while varying the binning of the $\sigma_{\text{inv}}$ distribution used for fitting, capturing the effect of the mismodeling. The statistical uncertainty in the determination of the differential $m_{jj}$ shape is negligible.

D. Multijet background

The QCD multijet background is estimated using events in data in which the $p_T^{\text{miss}}$ arises from mismeasured jets. Depending on the source of the mismeasurement, the jet that was mismeasured is either balanced in the case of additional HF noise or aligned with the $p_T^{\text{miss}}$ in $\phi$.

For the multijet background stemming from HF noise, an $m_{jj}$ template is extracted by inverting the requirements on the HF jet shape variables, hence requiring at least one jet in the event to fail the selection criteria given in Table I. The probability for an HF noise jet candidate to pass or fail the criteria is parametrized as a function of the jet $p_T$ and $\eta$, using events selected to have large $p_T^{\text{miss}}$ and to contain only one HF jet balanced in $\phi$ with the $p_T^{\text{miss}}$. An event-by-event weight is applied to estimate the contribution in the SR from the “failing” events. The estimated contamination from other SM processes is then removed bin by bin for the distribution under study. A closure test is performed by selecting events with the leading- $p_T$ jet within $3 < |\eta| < 3.25$. With this selection, the signal contamination is $< 2\%$ assuming $B(H \rightarrow \text{inv}) = 0.19$, as previously excluded in Ref. [22]. For events with spurious $p_T^{\text{miss}}$ from noise, one expects a full decorrelation between the $p_T^{\text{miss}}$ from the tracker acceptance only, $p_T^{\text{miss}}_{\text{track}}$, and from the full event, whereas for events with true $p_T^{\text{miss}}$, a correlation exists. Noise events are therefore expected to dominate in the region with large $\Delta\phi(p_T^{\text{miss}})$.

Data are compared with the estimated template for the $\Delta\phi(p_T^{\text{miss}})$ distribution in Fig. 3. From the agreement observed in the closure test in both years, a 20% systematic uncertainty is assigned to the template shape in the SR.

For the second category of multijet events, the requirement on $\min(\Delta\phi(p_T^{\text{miss}}_{\text{track}}, \phi_{\text{miss}}))$ is inverted to define a control region enriched in QCD multijet events (QCD CR). The $m_{jj}$ shape for the contribution of this background in the SR...
The candidates are chosen between leptons selected in the CRs and vetoed when vetoing events with identified objects. For the electrons, photons, muons, and $\tau$ candidates, b-tagged jets have been measured in both data and simulation [53,54,63,67]. The simulated events are corrected by scale factors, which are usually dependent on the $p_T$ and $\eta$ of the object, and have associated systematic uncertainties. The scale factors, and their uncertainties, are also propagated when vetoing events with identified objects. For the electrons and muons, when different working points are chosen between leptons selected in the CRs and vetoed in the SR, the corresponding uncertainties are kept higher orders in the $V +$ jets cross section calculations, and the closure of the HF noise contribution between data and simulation. The fitted values of the normalizations are used when subtracting the $W(\ell\nu)$ + jets, $Z(\ell\ell)$ + jets, and HF noise contributions from the data to obtain the $m_{jj}$ template. Their fitted uncertainties are included in the final systematic uncertainty in the QCD multijet estimate.

In both the MTR and VTR categories, the fit range is $0 < \min(\Delta\phi(p_T^{\text{miss}}, p_T^{\text{jet}})) < 1.8$, chosen to minimize the overlap of events in data with the SR. The fits are performed separately for the 2017 and 2018 datasets. Figure 4 shows the $\min(\Delta\phi(p_T^{\text{miss}}, \vec{p}_T^{\text{jet}}))$ distribution in data used in the fit and the contributions from $V +$ jets, diboson plus top quark processes, HF noise, and QCD multijet events resulting from the fit. The sums of the 2017 and 2018 datasets are shown for the MTR and VTR categories.

The function $f_{\text{QCD}}(x)$ provides an estimate of the QCD multijet template normalization, $N$, in the SR via

$$N = \int_X f_{\text{QCD}}(x)dx,$$

where $X = 0.5$ for the MTR categories and $X = 1.8$ for the VTR categories. An uncertainty in $N$ is derived by generating, and then refitting, pseudodata to extract the standard deviation of log $N$. This uncertainty includes the statistical uncertainty in the fit for the QCD multijet normalization and the uncertainty in the templates from simulation used to subtract the backgrounds due to the limited number of simulated events.

VI. SOURCES OF UNCERTAINTY

Several sources of systematic uncertainty affect the predictions of the signal and background components. They are separated into two categories, experimental and theoretical sources. Their effect is propagated either directly to the yields expected in the SR (for signal and backgrounds estimated directly from simulation) or to the transfer factors (for the $V +$ jets processes estimated from the CRs).

A. Experimental uncertainties

The reconstruction and identification efficiencies of electrons, photons, muons, $t\bar{t}$ candidates, and b-tagged jets have been measured in both data and simulation [53,54,63,67]. The simulated events are corrected by scale factors, which are usually dependent on the $p_T$ and $\eta$ of the object, and have associated systematic uncertainties. The scale factors, and their uncertainties, are also propagated when vetoing events with identified objects. For the electrons and muons, when different working points are chosen between leptons selected in the CRs and vetoed in the SR, the corresponding uncertainties are kept

FIG. 3. The $\Delta\phi(p_T^{\text{miss}}, \vec{p}_T^{\text{jet}})$ distribution in the SR with the additional requirement that the leading-$p_T$ jet passes $3 < |\eta| < 3.25$. The 2017 and 2018 data are compared with the sum of the HF noise template and other backgrounds from simulation. The uncertainty band includes only statistical uncertainties in the simulation.

derived from the yields in data in each $m_{jj}$ bin in the QCD CR, after subtracting the contributions from $V +$ jets, diboson, and top quark processes estimated from simulation, as well as HF noise contributions estimated in the data. The template is normalized as follows. The distribution of $\min(\Delta\phi(p_T^{\text{miss}}, \vec{p}_T^{\text{jet}}))$ in data is fit with the sum of templates derived from the simulated $V +$ jets, diboson, and top quark events; a HF noise template derived in data; and a functional form $f_{\text{QCD}}$ representing the QCD multijet contribution. The functional form is

$$f_{\text{QCD}}(x) = p_0 e^{-p_1 x}$$

for the MTR region and

$$f_{\text{QCD}}(x) = p_0 e^{-\frac{(x-x_0)^2}{2\sigma^2}}$$

for the VTR region. The parameters $p_i$ are allowed to float, and $x = \min(\Delta\phi(p_T^{\text{miss}}, \vec{p}_T^{\text{jet}}))$. The choices of these functions are validated by fitting this model to simulated QCD multijet events, and they are found to describe the distributions in the MTR and VTR categories well.

The normalizations of the $W(\ell\nu) +$ jets, $Z(\ell\ell) +$ jets, and HF noise contributions are allowed to vary independently in the fit. They are constrained within 20% of the prediction from simulation to account for systematic uncertainties related to jet energy calibrations, missing
fit uncertainty

VBF

The fit uncertainty is studied explicitly by considering uncorrelated between the 2017 and 2018 sample. All these uncertainties are correlated between the MTR and VTR categories. Except for the muons, where the dominant source of systematic uncertainty comes from the experimental method employed, these uncertainties are considered uncorrelated between the 2017 and 2018 datasets.

The effects of the jet energy scale (JES) and jet energy resolution (JER) uncertainties are studied explicitly by varying them by one standard deviation [58], propagating the change to the $p_T^{miss}$ accordingly, and checking the impact on the transfer factors for the $V + jets$ processes as functions of $m_{jj}$. For the signal and the minor backgrounds, the impact is studied on the simulated yields in the SR as a function of $m_{jj}$, with a fit procedure to remove the statistical contribution in the less-populated high-$m_{jj}$ bins. The JER uncertainties impact the signal yields by 4%--9% (2%--15%), increasing with $m_{jj}$, for VBF ($ggH$) production. The impacts of the JES uncertainties are 5%--25% (7%--35%), increasing with $m_{jj}$, for the VBF ($ggH$) process. Eleven independent JES sources are considered, with partial correlations between the 2017 and 2018 datasets. The dominant source is the $\eta$ dependence of the corrections. The corrections become particularly large for forward jets, which explains the large increase at high $m_{jj}$. The JER uncertainties are uncorrelated between the two data-taking years. Both the JER and JES uncertainties are correlated between the MTR and VTR categories.

Simulated events are weighted to match the distribution of reconstructed vertices to the distribution observed in the data. An uncertainty associated with this procedure is obtained by varying the total inelastic $pp$ cross section by $\pm 5\%$ [68] and repeating the background estimation procedure. This uncertainty is correlated across categories and datasets.

The trigger efficiencies are measured in both data and simulation, and a scale factor is extracted. For the signal triggers, the scale factor is parametrized as a function of the $p_T^{miss}$. The associated systematic uncertainty partially cancels in the transfer factors between the muon CRs and the SR. The associated uncertainty is uncorrelated across categories as different triggers are used but partially correlated between years. For the electron CRs, the scale factor is parametrized as a function of the lepton $p_T$ and $\eta$, and the impact of the uncertainties is propagated to the corresponding expectations in the SR. The uncertainty in the electron trigger scale factor is uncorrelated between years but correlated across categories.

The uncertainty in the integrated luminosity of the 2017 (2018) dataset is 2.3% [69] (2.5% [70]). When combined together with the 2016 dataset [71], the uncertainty is reduced to 1.6%. The improvement in the precision reflects the (uncorrelated) time evolution of some systematic effects. Eight independent sources are identified to take into account the correlations across datasets. These uncertainties are considered correlated between categories.

B. Theoretical uncertainties

The uncertainties in the $ggH$ and VBF predictions due to PDFs and renormalization and factorization scale variations are taken from Ref. [26]. For the $ggH$ process, an additional uncertainties...
uncertainty of 40% is assigned to take into account the limited knowledge of the $ggH$ production cross section in association with two or more jets, as well as the uncertainty in the prediction of the $ggH$ differential cross section for large Higgs boson $p_T$, $p_T^H > 250$ GeV, following the recipe described in Ref. [22]. The uncertainties in the signal acceptance due to the choice of the PDF set, and the renormalization and factorization scales, are evaluated independently for the different signal processes [22] and are treated as independent nuisance parameters in the fit.

Some of the theoretical uncertainties in the $V + \text{jets}$ and $\gamma + \text{jets}$ processes are expected to mostly cancel in the ratio of the $W^\pm$ ($\gamma$) to $Z$ processes. The uncertainties are estimated using the strong production and applied to both the VBF and strong production processes. As a conservative choice, for the pQCD NLO corrections, the renormalization and factorization scales are varied independently by a factor of 2. They are also varied independently for the $W(\ell\nu) + \text{jets}$ and $Z(\ell\nu) + \text{jets}$ processes. The maximum variation in the ratio of the $W(\ell\nu) + \text{jets}$ to $Z(\ell\nu) + \text{jets}$ yields, per $m_{jj}$ bin, is taken as the uncertainty. The maximum variation is generally given by that of the $W$ process. Uncertainties in the PDFs are directly propagated to the ratio of $W(\ell\nu) + \text{jets}$ to $Z(\ell\nu) + \text{jets}$ in a correlated way and are also applied per $m_{jj}$ bin. The full EW correction is taken as an additional uncertainty in the ratio of $W(\ell\nu) + \text{jets}$ ($\gamma + \text{jets}$) to $Z(\ell\nu) + \text{jets}$ for the strong production of those processes and is assumed to be uncorrelated between $m_{jj}$ bins. The theoretical uncertainties are assumed uncorrelated between the VBF- and strong-produced $V + \text{jets}$ processes, as well as between the MTR and VTR categories.

All theoretical uncertainty sources are fully correlated between years.

VII. RESULTS

A binned maximum likelihood fit is performed simultaneously across the SR and all CRs in both categories and for both datasets. In the fit, one parameter per bin $i$ of the $m_{jj}$ distribution, for each category and year, is left freely floating. This parameter represents the expected rate of the background from the strong production of $Z(\ell\nu) + \text{jets}$ events, and it is labeled $\kappa_{i}^{e\nu}$. The $m_{jj}$ distribution is binned in the same way as in Ref. [22]. This binning has been found to be optimal for the 2017 and 2018 datasets as well. Similar to the method described in Ref. [22], the likelihood function is defined as

$$
\mathcal{L}(\mu, \kappa^{e\nu}, \theta) = \prod_{i} P(d_{i}|B_{i}(\theta)) + Z_{i}(\kappa^{e\nu}, \mu, \theta) + W_{i}(\kappa^{e\nu}, \theta) + \mu S_{i}(\theta)) \times \prod_{\text{CR}} \left( \prod_{j} P(d_{i,\text{CR}}|B_{i,\text{CR}}(\theta)) + V_{i,\text{CR, strong}}(\kappa^{e\nu}, \theta) + V_{i,\text{CR, VBF}}(\kappa^{e\nu}, \theta) \right) \prod_{j} P(\theta_{j}),
$$

where $P(x|y) = y^e e^{-y}/x!$ and $d_{i}$ are the observed number of events in each bin $i$ of the $m_{jj}$ distribution in each of the CRs and in the SR, respectively. The index $i$ runs over the $m_{jj}$ bins in the two years and all categories. The symbol $\theta$ refers to constrained nuisance parameters used for the modeling of the systematic uncertainties. The signal term $S_{i}$ represents the expected signal prediction from the sum of the main Higgs boson production mechanisms ($ggH$, VBF, VH, $ttH$) assuming the cross sections predicted in the SM. The parameter $\mu = (\sigma_{H}/\sigma_{H}^{SM})B(H \rightarrow \text{inv})$ denotes the signal strength parameter and is also left freely floating in the fit.

In a given bin, the $V + \text{jets}$ background yields expected in the SR are obtained from transfer factors relating the yields in the different CRs to the yields in the SR, separately for the VBF and strong production processes. These transfer factors are denoted $R^\text{VBF}_i(\theta)$, where “proca” can be strong or VBF, and are obtained from the simulation. For the single-lepton (dilepton) CRs, the factors $R^\text{VBF}_i(\theta)$ refer to the ratio of $W + \text{jets}$ ($Z + \text{jets}$) yields from the corresponding CR to the SR. In the photon CR, which is only available in the MTR category, $R^\text{VBF}_i(\theta) = 1$.

In addition, transfer factors are defined between the $W$ ($\gamma$) and the $Z$ processes, separately for the VBF and strong production processes. These transfer factors are denoted as $f^W/\gamma_{i,res}(\theta)$ ($f^Z/\gamma_{i,res}(\theta)$), where “proca” can be strong or VBF. Finally, a transfer factor, denoted as $Z^\text{VBF}_{i,\text{MTR}}$, relates the VBF production to the strong production of $Z(\ell\nu) + \text{jets}$. The factors $C^\text{VBF, strong}_i(\theta)$ and $C^\text{VBF, VBF}_i(\theta)$ are dependent on

$$
V_{i}^{\text{CR, strong}}(\kappa^{e\nu}, \theta) = C_{i}^{\text{CR, strong}}(\theta) \kappa^{e\nu},
$$

where $P(x|y) = y^e e^{-y}/x!$ and $d_{i}$ are the observed number of events in each bin $i$ of the $m_{jj}$ distribution in each of the CRs and in the SR, respectively. The index $i$ runs over the $m_{jj}$ bins in the two years and all categories. The symbol $\theta$ refers to constrained nuisance parameters used for the modeling of the systematic uncertainties. The signal term $S_{i}$ represents the expected signal prediction from the sum of the main Higgs boson production mechanisms ($ggH$, VBF, VH, $ttH$) assuming the cross sections predicted in the SM. The parameter $\mu = (\sigma_{H}/\sigma_{H}^{SM})B(H \rightarrow \text{inv})$ denotes the signal strength parameter and is also left freely floating in the fit.
the nature of the CR, with $C_{\gamma,proc}^{(e,e,\mu)} = 1$, $C_{\gamma,proc}^{(e,\mu)} = f_{\gamma/Z,proc}(\theta)$, and $C_{\gamma,proc}^{(\gamma,\gamma)} = f_{\gamma/Z,proc}(\theta)$.

The contributions from subleading backgrounds in each region are estimated directly from simulation, and they are denoted by $B_{i}^{CR}(\theta)$ in the CRs and $B_{i}(\theta)$ in the SR.

Systematic uncertainties are modeled as constrained nuisance parameters ($\theta$), with a log-normal distribution for those that affect the overall normalization of a given process, and Gaussian priors for those that directly affect the transfer factors, indicated by $P(\theta)$ in Eq. (4). The impact

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>Ratios</th>
<th>Uncertainty vs $m_{jj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Renormalization scale $V +$ jets (VBF)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>7.5%</td>
</tr>
<tr>
<td>Renormalization scale $V +$ jets (strong)</td>
<td>$\frac{W/Z_{strong}}{W/Z_{strong}}$</td>
<td>8.2%</td>
</tr>
<tr>
<td>Factorization scale $V +$ jets (VBF)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>1.5%</td>
</tr>
<tr>
<td>Factorization scale $V +$ jets (strong)</td>
<td>$\frac{W/Z_{strong}}{W/Z_{strong}}$</td>
<td>0%</td>
</tr>
<tr>
<td>PDF $V +$ jets (VBF)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>0%</td>
</tr>
<tr>
<td>NLO EW corr. $V +$ jets (strong)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>0.5%</td>
</tr>
<tr>
<td>Renormalization scale $\gamma +$ jets (VBF)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>6%-10%</td>
</tr>
<tr>
<td>Renormalization scale $\gamma +$ jets (strong)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>6%-10%</td>
</tr>
<tr>
<td>Factorization scale $\gamma +$ jets (VBF)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>2.5%</td>
</tr>
<tr>
<td>Factorization scale $\gamma +$ jets (strong)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>2.5%</td>
</tr>
<tr>
<td>PDF $\gamma +$ jets (VBF)</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>2.5%</td>
</tr>
<tr>
<td>NLO EW corr. $\gamma +$ jets</td>
<td>$\frac{W/Z_{VBF}}{W/Z_{VBF}}$</td>
<td>3%</td>
</tr>
<tr>
<td>Electron reconstruction efficiency</td>
<td>$R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $W(\mu)$</td>
<td>$\approx0.5%$ (per lepton)</td>
</tr>
<tr>
<td>Electron identification efficiency</td>
<td>$R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $W(\mu)$</td>
<td>$\approx1%$ (per lepton)</td>
</tr>
<tr>
<td>Muon identification efficiency</td>
<td>$R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $W(\mu)$</td>
<td>$\approx0.5%$ (per lepton)</td>
</tr>
<tr>
<td>Muon isolation efficiency</td>
<td>$R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $W(\mu)$</td>
<td>$\approx0.1%$ (per lepton)</td>
</tr>
<tr>
<td>Photon identification efficiency</td>
<td>$R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $W(\mu)$</td>
<td>5%</td>
</tr>
<tr>
<td>Electron veto (reconstruction)</td>
<td>$f_{W-Z_{VBF}}$, $R_{1}^{CR,proc}$, $CR = W(\ell\nu)$</td>
<td>$\approx1.5$ (1%) for VBF (strong)</td>
</tr>
<tr>
<td>Electron veto (identification)</td>
<td>$f_{W-Z_{VBF}}$, $R_{1}^{CR,proc}$, $CR = W(\ell\nu)$</td>
<td>$\approx2.5$ (2%) for VBF (strong)</td>
</tr>
<tr>
<td>Muon veto</td>
<td>$f_{W-Z_{VBF}}$, $R_{1}^{CR,proc}$, $CR = W(\ell\nu)$</td>
<td>0.5%</td>
</tr>
<tr>
<td>$\tau_{1}$ veto</td>
<td>$f_{W-Z_{VBF}}$, $R_{1}^{CR,proc}$, $CR = W(\ell\nu)$</td>
<td>0.5%</td>
</tr>
<tr>
<td>Electron trigger</td>
<td>$R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $W(\mu)$</td>
<td>1%</td>
</tr>
<tr>
<td>$p_{T}^{miss}$ trigger</td>
<td>$R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $W(\mu)$</td>
<td>1%–2%</td>
</tr>
<tr>
<td>Photon trigger</td>
<td>$f_{W-Z_{VBF}}$, $R_{1}^{CR,proc}$, $CR = W(\mu)$</td>
<td>1%–1.5%</td>
</tr>
<tr>
<td>JES</td>
<td>$f_{W-Z_{VBF}}$, $R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $Z(\mu\mu)$</td>
<td>1%</td>
</tr>
<tr>
<td>JER</td>
<td>$f_{W-Z_{VBF}}$, $R_{1}^{CR,proc}$, $CR = Z(\mu\mu)$ or $Z(\mu\mu)$</td>
<td>1%</td>
</tr>
</tbody>
</table>

TABLE III. Experimental and theoretical sources of systematic uncertainty in the $V +$ jets transfer factors. The second column indicates which ratio a given source of uncertainty is applied to. The impact on $m_{jj}$ is given in the third column, as a single value if no dependence on $m_{jj}$ is observed, or as a range. When a range is shown, the uncertainty increases with the value of $m_{jj}$. The quoted uncertainty values represent the absolute value of the relative change in the transfer factor corresponding to a variation of $\pm 1$ standard deviation in the systematic uncertainty, or equivalently to a variation of $\pm 1$ in the corresponding nuisance parameter $\theta$ in Eq. (4).
on the transfer factors of each of the uncertainties described in Sec. VI is summarized in Table III.

In the following, the expected background yields used as input to the fit procedure are denoted as “prefit,” while the yields after a fit to the CRs, or the CRs and SR, are denoted as “CR-postfit,” or “postfit,” respectively.

The results are presented for the MTR and VTR categories separately. They are shown separately for the 2017 and 2018 datasets when presenting the transfer factors, and for the datasets added together otherwise. The corresponding figures and tables split into individual control regions and data-taking periods are available in the Supplemental Material [66].

All results are obtained from a combined fit across all categories and datasets.

For the maximum likelihood fit, the different datasets and categories are treated separately. The final results are obtained from a fit using the combined likelihood function

![Graphs showing data and simulation comparison for MTR and VTR categories.](image)

FIG. 5. Comparison between data and simulation for the Z(ℓℓ) + jets/W(ℓν) + jets (upper row) and γ + jets/Z(ℓℓ) + jets (lower row) prefit and CR-postfit ratios, as functions of \( m_{jj} \), for the MTR category using the 2017 (left) and 2018 (right) event samples. The minor backgrounds (bkg.) in each CR are subtracted from the data using estimates from simulation. The gray bands include the theoretical and experimental systematic uncertainties listed in Table III, as well as the statistical uncertainty in the simulation.
including all categories and datasets, which takes into account nuisance correlations.

A. Control regions for the MTR category

The ratio of dilepton to single-lepton events is studied to validate the predictions from simulation and uncertainties used to model the ratio of $Z + \text{jets}$ to $W + \text{jets}$ events in the SR [parameters $f_i^{W/Z,\text{proc}}(\theta)$ in Eq. (4)]. The prefit ratio between the number of $Z + \text{jets}$ and $W + \text{jets}$ events in the CRs in bins of $m_{jj}$ is shown in Fig. 5 (upper row) for the 2017 (left) and 2018 (right) datasets. The prediction is compared to the ratio of observed data yields, from which the estimates of minor background processes have been subtracted. The CR-postfit results are shown together with the prefit ratio. A reasonable agreement is observed between data and simulation, with differences in most bins covered by the systematic uncertainties listed in Table III.

FIG. 5. The postfit $m_{jj}$ distributions in the dielectron (upper left), dimuon (upper right), single-electron (lower left), and single-muon (lower right) CR for the MTR category, showing the summed 2017 and 2018 data samples and the background processes. The background contributions are estimated from the fit to the data described in the text (S+B fit). The total background (bkg.) estimated from a fit assuming $B(H \rightarrow \text{inv}) = 0$ (B-only fit) is also shown. The yields from the 2017 and 2018 samples are summed and the correlations between their uncertainties are neglected. The last bin of each distribution integrates events above the bin threshold divided by the bin width.

FIG. 6. The postfit $m_{jj}$ distributions in the dielectron (upper left), dimuon (upper right), single-electron (lower left), and single-muon (lower right) CR for the MTR category, showing the summed 2017 and 2018 data samples and the background processes. The background contributions are estimated from the fit to the data described in the text (S+B fit). The total background (bkg.) estimated from a fit assuming $B(H \rightarrow \text{inv}) = 0$ (B-only fit) is also shown. The yields from the 2017 and 2018 samples are summed and the correlations between their uncertainties are neglected. The last bin of each distribution integrates events above the bin threshold divided by the bin width.
In 2017, the simulation predicts a lower ratio than observed in data. The compatibility of the data with the prefit prediction is measured, in this particular category and year only, using a $\chi^2$ test accounting for correlations between the 2017 MTR CRs and each $m_{jj}$ bin. This test indicates that there is a local discrepancy of approximately two standard deviations. The disagreement is attributed to the $Z +$ jets regions with low event yields and is partially compensated in the fit through the movement of nuisance parameters representing uncertainties such as the renormalization and factorization scales. The significance of the discrepancy is low, and none of the nuisance parameters moves by more than one standard deviation from their prefit value in the combined fit across all categories and years. The $p$-value for the 2017 dataset in the MTR category after the combined fit is 38.4%, and it is 37.0% for all categories when combining the 2017 and 2018 datasets. Closure tests have also been performed comparing the decays to electrons or muons separately for the $W$ and $Z$ CRs, again showing reasonable agreement between data and simulation.

The ratio of photon to dilepton events is also studied to validate the predictions from simulation and uncertainties used to model the ratio of $\gamma +$ jets to $Z +$ jets events in the SR [parameters $f^i_{\gamma/Z, \text{proc}}(\theta)$ in Eq. (4)]. The prefit and CR-postfit ratios between the number of events in the $\gamma$ + jets over $Z$ + jets CRs in bins of $m_{jj}$ are shown in Fig. 5 (lower row) for the 2017 (left) and 2018 (right) datasets. The prediction is compared to the ratio of observed data yields, from which the estimates of minor background processes have been subtracted. A similar observation can be made for the 2017 MTR category as for the $Z$ + jets over

FIG. 7. The postfit $m_{jj}$ distribution in the photon CR for the MTR category, showing the summed 2017 and 2018 data samples and the background processes. The background contributions are estimated from the fit to the data described in the text (S + B fit). The total background (bkg.) estimated from a fit assuming $B(H \to \text{inv}) = 0$ (B-only fit) is also shown. The yields from the 2017 and 2018 samples are summed, and the correlations between their uncertainties are neglected. The last bin of each distribution integrates events above the bin threshold divided by the bin width.

FIG. 8. Comparison between data and simulation for the $Z(\ell\ell) +$ jets/$W(\ell\nu) +$ jets prefit and CR-postfit ratios, as functions of $m_{jj}$, for the VTR category in the 2017 (left) and 2018 (right) data samples. The minor backgrounds (bkg.) in each CR are subtracted from the data using estimates from simulation. The gray bands include the theoretical and experimental systematic uncertainties listed in Table III, as well as the statistical uncertainty in the simulation.
This effect is again attributed to the $Z + \text{jets}$ CRs with low event yields.

The $m_{jj}$ distributions in data in the dilepton and single-lepton CRs, along with the postfit estimates of the background contributions, are shown in Fig. 6. Similar distributions in the photon CR are shown in Fig. 7. The total background estimated from a fit assuming $B(H \to \text{inv}) = 0$ is also shown. The distributions show the sum of the 2017 and 2018 datasets in each region. The postfit predictions are in good agreement with the data within one standard deviation for most of the bins, with discrepancies of just over one standard deviation in the dielectron MTR CR for two bins at $m_{jj} \approx 2 \text{ TeV}$. As discussed previously in the context of the validation of the

![Image of CMS plots](image-url)

**FIG. 9.** The postfit $m_{jj}$ distributions in the dielectron (upper left), dimuon (upper right), single-electron (lower left), and single-muon (lower right) CRs for the VTR category, showing the summed 2017 and 2018 data samples and the SM background processes. The background contributions are estimated from the fit to the data described in the text (S+B fit). The total background (bkg.) estimated from a fit assuming $B(H \to \text{inv}) = 0$ (B-only fit) is also shown. The yields from the 2017 and 2018 samples are summed, and the correlations between their uncertainties are neglected. The last bin of each distribution integrates events above the bin threshold divided by the bin width.
Z + jets over W + jets and the $\gamma +$ jets over Z + jets ratios in
the CRs, the prefit disagreement observed in the 2017 MTR
is partially compensated by the nuisance parameters in the fit,
resulting in an overall $p$-value of 37.0%.

B. Control regions for the VTR category

The prefit and CR-postfit ratios between the number of $Z$ + jets and $W$ + jets events in the CRs in bins of $m_{jj}$ are
shown in Fig. 8. The predictions from simulated events are
found to model the data in all bins, within the quoted uncertainties.

The $m_{jj}$ distributions in the dilepton and single-lepton CRs are shown in Fig. 9, along with the CR-
postfit and postfit estimates. Again, a good agreement
between the data and the predictions is shown, within the uncertainties.

C. Signal region fits

The background estimates in the SR are reported for each $m_{jj}$ bin of the MTR category in Table IV and for each $m_{jj}$ bin of the VTR category in Table V. The observed and expected $m_{jj}$ distributions in the SR are shown in Fig. 10 for the MTR (left) and VTR (right) categories.

D. Combination of results

No significant deviations from the SM expectations are observed. The results of this search are interpreted in terms of
an upper limit on the product of the Higgs boson production
cross section and its branching fraction to invisible particles,$\sigma_H B(H \rightarrow \text{inv})$, relative to the predicted cross section
assuming SM interactions, $\sigma_{SM}^H$. Observed and expected 95% C.L. upper limits are computed using an asymptotic
approximation of the CL$_S$ method detailed in Refs. [73,74].

---

**TABLE IV.** Expected event yields in each $m_{jj}$ bin for the different background processes in the SR of the MTR category, summing the 2017 and 2018 samples. The background yields and the corresponding uncertainties are obtained after performing a combined fit across all of the CRs and the SR. The expected signal contributions for the Higgs boson, produced in the non-VBF and VBF modes, decaying to invisible particles with a branching fraction of $B(H \rightarrow \text{inv}) = 1$, and the observed event yields are also reported. The yields from the 2017 and 2018 samples are summed, and the correlations between their uncertainties are neglected.

<table>
<thead>
<tr>
<th>$m_{jj}$ bin (GeV)</th>
<th>200–400</th>
<th>400–600</th>
<th>600–900</th>
<th>900–1200</th>
<th>1200–1500</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z(\ell\ell) +$ jets (strong)</td>
<td>$26107.6 \pm 82.7$</td>
<td>$15521.0 \pm 62.1$</td>
<td>$10747.3 \pm 48.8$</td>
<td>$4040.4 \pm 25.3$</td>
<td>$1923.4 \pm 16.7$</td>
</tr>
<tr>
<td>$Z(\ell\ell) +$ jets (VBF)</td>
<td>$431.3 \pm 6.1$</td>
<td>$498.2 \pm 6.6$</td>
<td>$620.6 \pm 7.0$</td>
<td>$452.0 \pm 6.2$</td>
<td>$294.9 \pm 5.3$</td>
</tr>
<tr>
<td>$W(\ell\nu) +$ jets (strong)</td>
<td>$13571.4 \pm 76.8$</td>
<td>$8293.4 \pm 53.0$</td>
<td>$5868.4 \pm 43.2$</td>
<td>$2409.6 \pm 23.4$</td>
<td>$1053.6 \pm 16.1$</td>
</tr>
<tr>
<td>$W(\ell\nu) +$ jets (VBF)</td>
<td>$268.0 \pm 10.5$</td>
<td>$301.5 \pm 11.3$</td>
<td>$353.5 \pm 12.7$</td>
<td>$242.8 \pm 8.5$</td>
<td>$163.0 \pm 5.9$</td>
</tr>
<tr>
<td>$t\bar{t} +$ single $t$ quark</td>
<td>$498.8 \pm 21.2$</td>
<td>$370.6 \pm 15.6$</td>
<td>$275.5 \pm 11.8$</td>
<td>$115.3 \pm 5.1$</td>
<td>$59.6 \pm 2.8$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$464.9 \pm 40.0$</td>
<td>$305.1 \pm 26.2$</td>
<td>$246.3 \pm 21.3$</td>
<td>$85.4 \pm 7.5$</td>
<td>$39.4 \pm 3.5$</td>
</tr>
<tr>
<td>$Z/\gamma^* (\ell^+\ell^-) +$ jets</td>
<td>$192.3 \pm 4.4$</td>
<td>$126.3 \pm 2.9$</td>
<td>$102.0 \pm 2.5$</td>
<td>$38.2 \pm 1.0$</td>
<td>$16.1 \pm 0.5$</td>
</tr>
<tr>
<td>Multijet</td>
<td>$10.9 \pm 2.0$</td>
<td>$10.6 \pm 1.9$</td>
<td>$10.4 \pm 1.8$</td>
<td>$4.8 \pm 0.9$</td>
<td>$2.3 \pm 0.4$</td>
</tr>
<tr>
<td>HF noise</td>
<td>$0.8 \pm 0.1$</td>
<td>$35.1 \pm 3.0$</td>
<td>$82.7 \pm 7.3$</td>
<td>$70.3 \pm 6.2$</td>
<td>$28.1 \pm 2.5$</td>
</tr>
<tr>
<td>$qqH$$\rightarrow$inv</td>
<td>$130.5$</td>
<td>$297.0$</td>
<td>$586.1$</td>
<td>$571.7$</td>
<td>$460.5$</td>
</tr>
<tr>
<td>Other $H$$\rightarrow$inv signals</td>
<td>$1430.9$</td>
<td>$1027.1$</td>
<td>$848.7$</td>
<td>$414.3$</td>
<td>$209.5$</td>
</tr>
<tr>
<td>Total background</td>
<td>$41546.0 \pm 122.3$</td>
<td>$25467.1 \pm 88.2$</td>
<td>$18306.6 \pm 71.5$</td>
<td>$7822.8 \pm 37.6$</td>
<td>$3580.2 \pm 25.1$</td>
</tr>
<tr>
<td>Observed</td>
<td>$41450$</td>
<td>$25356$</td>
<td>$18438$</td>
<td>$7793$</td>
<td>$3629$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m_{jj}$ bin (GeV)</th>
<th>1500–2000</th>
<th>2000–2750</th>
<th>2750–3500</th>
<th>&gt;3500</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z(\ell\ell) +$ jets (strong)</td>
<td>$1261.7 \pm 12.7$</td>
<td>$462.4 \pm 7.4$</td>
<td>$95.6 \pm 4.7$</td>
<td>$28.8 \pm 1.4$</td>
</tr>
<tr>
<td>$Z(\ell\ell) +$ jets (VBF)</td>
<td>$317.8 \pm 6.0$</td>
<td>$197.3 \pm 5.3$</td>
<td>$62.1 \pm 3.6$</td>
<td>$35.8 \pm 2.3$</td>
</tr>
<tr>
<td>$W(\ell\nu) +$ jets (strong)</td>
<td>$704.3 \pm 10.9$</td>
<td>$276.7 \pm 7.6$</td>
<td>$65.4 \pm 4.1$</td>
<td>$23.5 \pm 2.4$</td>
</tr>
<tr>
<td>$W(\ell\nu) +$ jets (VBF)</td>
<td>$163.9 \pm 6.1$</td>
<td>$111.9 \pm 4.6$</td>
<td>$49.4 \pm 3.2$</td>
<td>$19.2 \pm 1.6$</td>
</tr>
<tr>
<td>$t\bar{t} +$ single $t$ quark</td>
<td>$38.7 \pm 2.2$</td>
<td>$14.9 \pm 1.2$</td>
<td>$5.3 \pm 0.5$</td>
<td>$1.8 \pm 0.2$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$27.4 \pm 2.7$</td>
<td>$7.9 \pm 0.8$</td>
<td>$0.6 \pm 0.1$</td>
<td>$0.0 \pm 0.1$</td>
</tr>
<tr>
<td>$Z/\gamma^* (\ell^+\ell^-) +$ jets</td>
<td>$11.9 \pm 0.6$</td>
<td>$4.9 \pm 0.3$</td>
<td>$1.5 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
</tr>
<tr>
<td>Multijet</td>
<td>$2.1 \pm 0.4$</td>
<td>$1.0 \pm 0.2$</td>
<td>$0.4 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
</tr>
<tr>
<td>HF noise</td>
<td>$56.4 \pm 5.0$</td>
<td>$62.2 \pm 5.6$</td>
<td>$30.5 \pm 2.7$</td>
<td>$20.9 \pm 1.8$</td>
</tr>
<tr>
<td>$qqH$$\rightarrow$inv</td>
<td>$539.6$</td>
<td>$427.2$</td>
<td>$177.9$</td>
<td>$118.0$</td>
</tr>
<tr>
<td>Other $H$$\rightarrow$inv signals</td>
<td>$161.8$</td>
<td>$84.7$</td>
<td>$24.3$</td>
<td>$11.0$</td>
</tr>
<tr>
<td>Total background</td>
<td>$2584.4 \pm 19.8$</td>
<td>$1139.2 \pm 14.0$</td>
<td>$310.7 \pm 8.3$</td>
<td>$130.6 \pm 4.4$</td>
</tr>
<tr>
<td>Observed</td>
<td>$2623$</td>
<td>$1142$</td>
<td>$279$</td>
<td>$136$</td>
</tr>
</tbody>
</table>
with a profile likelihood ratio test statistic [75] in which systematic uncertainties are modeled as nuisance parameters following a frequentist approach [76].

Both VBF and non-VBF signal production modes are included, with their relative contributions fixed to the SM prediction within their uncertainties.

Between the 2017 and 2018 datasets, and the two analysis categories (MTR and VTR), the uncertainties are correlated according to the description given in Sec. VI. To combine with the data taken in 2016, the same correlation scheme as between 2017 and 2018 is used, except for the jet energy calibration uncertainties (JES and JER), which are kept fully uncorrelated. The integrated luminosity of the 2016 dataset was updated to 36.3 fb$^{-1}$ to reflect the latest improvements in the luminosity measurement [71]. To be consistent with the treatment of the VBF signal in the 2017 and 2018 analyses, the Higgs boson $p_T$-dependent EW NLO corrections are also applied to the 2016 signal shape. The VBF results obtained with the earlier datasets, namely the dataset from 2012 (2015), at $\sqrt{s} = 8$ (13) TeV, with 19.7 (2.3) fb$^{-1}$, from Ref. [27], are combined, taking into account uncertainty correlations where appropriate. Theoretical uncertainties related to signal modeling are correlated for data-taking periods with the same center-of-mass energy. Partial correlations between datasets exist for the uncertainty in the luminosity measurements. All other experimental uncertainties are decorrelated between the run periods before and after 2015. The results of the fit to the data across all data-taking periods are available in Supplemental Material [66].

### 1. Constraints on an SM-like Higgs boson

Observed and expected upper limits on $(\sigma_{\text{fit}}/\sigma_{\text{SM}}) \times B(H \to \text{inv})$ at 95% C.L. are presented in Fig. 11 and Table VI. The limits are computed for the combination of all datasets, as well as for individual categories and data-taking periods. By itself, the addition of the $\gamma + \text{jets}$ category (the addition of the VTR category) improves the expected limits by about 11% (8%) compared with the 2016-like analysis selections. Considered together, the improvements reach about 17% in both years. The upper limits for the individual categories entering the combination are available in Supplemental Material [66].

The combination of the 2017 and 2018 results yields an observed (expected) upper limit of $B(H \to \text{inv}) < 0.18 (0.12)$ at the 95% C.L., assuming an SM Higgs boson with a mass of 125.38 GeV [77]. Figure 11 additionally shows a combination with data collected in 2012 and 2015 [27] and in 2016 [22]. This combination yields an observed (expected) upper limit of $B(H \to \text{inv}) < 0.18 (0.10)$ at the 95% C.L., which is currently the stringent limit on $B(H \to \text{inv})$.

Figure 12 shows the profile likelihood ratio ($q$) as a function of $B(H \to \text{inv})$, for the individual data-taking periods and for their combination. The observed (expected) combined 2012–2018 best fit signal strength is found to be $0.086^{+0.054}_{-0.052}$ ($0.00^{+0.051}_{-0.052}$). Table VII summarizes the uncertainties in the measured $B(H \to \text{inv})$, separating the contributions from different groups of uncertainties. The systematic uncertainties with the largest impacts in the $B(H \to \text{inv})$ measurement are the theoretical uncertainties affecting the $f_W^W/Z_{\text{proc}}$ ratio, followed by the statistical uncertainties in the simulated samples, the trigger uncertainties, jet calibration effects, and the uncertainties in the QCD multijet modeling.

The upper limit on $B(H \to \text{inv})$, obtained from the combination of 2012–2018 data, is interpreted in the context of Higgs-portal models of DM interactions, in which a stable
DM particle couples to the SM Higgs boson. The interaction between a DM particle and an atomic nucleus may be mediated by the exchange of a Higgs boson, producing nuclear recoil signatures, such as those investigated by direct detection experiments. The sensitivity of these experiments depends mainly on the DM particle mass ($m_{DM}$). If $m_{DM}$ is smaller than half of the Higgs boson mass, the partial width of the invisible Higgs boson decay ($\Gamma_{inv}$) can be translated, within an effective field theory approach, into a spin-independent DM-nucleon elastic scattering cross section, as outlined in Ref. [16]. This translation is performed assuming that the DM candidate is either a scalar or a Majorana fermion, and both the central value and the uncertainty in the dimensionless nuclear form factor $f_N$ are taken from the recommendations of Ref. [79]. We do not perform the translation under the assumption of a vector DM candidate in this paper, since it requires an extended dark Higgs sector, which may lead to modifications of kinematic distributions assumed for the invisibly decaying Higgs boson signal. Figure 13 shows the

**TABLE VI.** The 95% C.L. upper limits on $(\sigma_{H}/\sigma_{H}^{SM})B(\text{inv} \rightarrow \text{inv})$ for all data-taking years considered, as well as their combination, assuming an SM Higgs boson with a mass of 125.38 GeV. The observed and median expected results are shown, along with the 68% and 95% interquartile ranges for each category and for the combinations.

<table>
<thead>
<tr>
<th>Category</th>
<th>Observed Median</th>
<th>68% expected</th>
<th>95% expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012–2016</td>
<td>0.33</td>
<td>0.21</td>
<td>[0.15, 0.29]</td>
</tr>
<tr>
<td>VTR 2017</td>
<td>0.57</td>
<td>0.45</td>
<td>[0.32, 0.66]</td>
</tr>
<tr>
<td>VTR 2018</td>
<td>0.44</td>
<td>0.34</td>
<td>[0.24, 0.49]</td>
</tr>
<tr>
<td>VTR 2017+2018</td>
<td>0.70</td>
<td>0.48</td>
<td>[0.40, 0.76]</td>
</tr>
<tr>
<td>MTR 2017</td>
<td>0.25</td>
<td>0.19</td>
<td>[0.14, 0.28]</td>
</tr>
<tr>
<td>MTR 2018</td>
<td>0.63</td>
<td>0.47</td>
<td>[0.32, 0.66]</td>
</tr>
<tr>
<td>MTR 2017+2018</td>
<td>0.26</td>
<td>0.19</td>
<td>[0.14, 0.28]</td>
</tr>
<tr>
<td>all 2017</td>
<td>0.24</td>
<td>0.18</td>
<td>[0.13, 0.26]</td>
</tr>
<tr>
<td>all 2018</td>
<td>0.25</td>
<td>0.15</td>
<td>[0.10, 0.21]</td>
</tr>
<tr>
<td>all 2017+2018</td>
<td>0.18</td>
<td>0.12</td>
<td>[0.08, 0.17]</td>
</tr>
<tr>
<td>2012–2018</td>
<td>0.18</td>
<td>0.10</td>
<td>[0.07, 0.14]</td>
</tr>
</tbody>
</table>
90% C.L. upper limits on the spin-independent DM-nucleon scattering cross section as a function of $m_{\text{DM}}$, for both the scalar and the fermion DM scenarios. The corresponding 90% C.L. upper limit on $B(H \rightarrow \text{inv})$ is 0.16. These limits are computed at the 90% C.L. so that they can be compared with those from direct detection experiments such as XENON1T [80], CRESST-II [81], CDMSlite [82], LUX [83], Panda-X 4T [84], and DarkSide-50 [85], which provide the strongest constraints in the DM range probed by this search. The collider-based results complement the direct-detection experiments in the range $m_{\text{DM}}$ smaller than 12 (6) GeV, assuming a fermion (scalar) DM candidate.

### VIII. SUMMARY

A search for the Higgs boson (H) decaying invisibly, produced in the vector boson fusion mode, is performed with 101 fb$^{-1}$ of proton-proton collisions delivered by the LHC at $\sqrt{s} = 13$ TeV and collected by the CMS detector during 2017–2018. Building upon the previously published results, an additional category targeting events at lower Higgs boson transverse momentum is added. An additional highly populated control region, based on production of a photon associated with jets, is used to constrain the dominant irreducible background from invisible decays of a Z boson produced in association with jets. Compared with the strategy of the previously published analysis, these additions improve the expected limits by approximately 17%. The observed (expected) upper limit on the invisible branching fraction of the Higgs boson, $B(H \rightarrow \text{inv})$, is found to be 0.18 (0.12) at the 95% C.L., assuming the standard model production cross section. The results are combined with previous measurements in the vector boson fusion topology, for total integrated luminosities of 19.7 fb$^{-1}$ at $\sqrt{s} = 8$ TeV and 140 fb$^{-1}$ at $\sqrt{s} = 13$ TeV, yielding an observed (expected) upper limit of 0.18 (0.10) at the 95% C.L. This is currently the most stringent limit on $B(H \rightarrow \text{inv})$. Finally, the results are interpreted in the context of Higgs-portal models. The 90% C.L. upper limits

---

**TABLE VII.** Uncertainty breakdown in $B(H \rightarrow \text{inv})$. The sources of uncertainty are separated into different groups. Observed and expected results are quoted for the full combination of 2012–2018 data. The expected results are obtained using an Asimov dataset [75] with $B(H \rightarrow \text{inv}) = 0$.

<table>
<thead>
<tr>
<th>Group of systematic uncertainties</th>
<th>Impact on $B(H \rightarrow \text{inv})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory</td>
<td>+0.026 ±0.024</td>
</tr>
<tr>
<td>Simulated event count</td>
<td>±0.022 ±0.021</td>
</tr>
<tr>
<td>Triggers</td>
<td>0.018 ±0.018</td>
</tr>
<tr>
<td>Jet calibration</td>
<td>0.014 ±0.011</td>
</tr>
<tr>
<td>QCD multijet mismodeling</td>
<td>±0.012 ±0.013</td>
</tr>
<tr>
<td>Leptons/photons/b-tagged jets</td>
<td>±0.010 ±0.010</td>
</tr>
<tr>
<td>Integrated luminosity/pileup</td>
<td>±0.004 ±0.004</td>
</tr>
<tr>
<td>Other systematic uncertainties</td>
<td>±0.009 ±0.009</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>±0.028 ±0.028</td>
</tr>
</tbody>
</table>
on the spin-independent dark-matter-nucleon scattering cross section obtained from the observed LHC data collected during 2012–2018 complement the direct detection experiments in the range of dark matter particle masses smaller than 12 (6) GeV, assuming a fermion (scalar) dark matter candidate.

ACKNOWLEDGMENTS

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid and other centers for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC, the CMS detector, and the supporting computing infrastructure provided by the following funding agencies: BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES and BNSF (Bulgaria); CERN; CAS, MoST, and NSFC (China); MINCIENCIAS (Colombia); MSES and CSF (Croatia); RIF (Cyprus); SENSCYT (Ecuador); MoER, ERC PUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRI (Greece); NKFIA (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFFR, and NRC KI (Russia); MESTD (Serbia); MCIN/AEI and PCTI (Spain); MOSTR (Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU (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, Grants No. 675440, No. 724704, No. 752730, No. 758316, No. 765710, No. 824093, and No. 884104, and COST Action CA16108 (European Union); the Leventis Foundation; the Alfred 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 F.R.S.-FNRS and FWO (Belgium) under the “Excellence of Science—EOS”—be.h Project No. 30820817; the Beijing Municipal Science & Technology Commission, Grant No. Z191100007219010; the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Deutsche Forschungsgemeinschaft (DFG), under Germany’s Excellence Strategy—EXC 2121 “Quantum Universe”—Grant No. 390833306, and under Project No. 400140256—GRK2497; the Lendület (“Momentum”) Program and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the New National Excellence Program ÚNKP, the NKFIA research Grants No. 123842, No. 123959, No. 124845, No. 124850, No. 125105, No. 128713, No. 128786, and No. 129058 (Hungary); the Council of Science and Industrial Research, India; the Latvian Council of Science; the Ministry of Science and Higher Education and the National Science Center, Contracts Opus 2014/15/B/ST2/03998 and 2015/19/B/ST2/02861 (Poland); the Fundação para a Ciência e a Tecnologia, Grant No. CEECIND/01334/2018 (Portugal); the National Priorities Research Program by Qatar National Research Fund; the Ministry of Science and Higher Education, Projects No. 0723-2020-0041 and No. FSWW-2020-0008, and the Russian Foundation for Basic Research, Project No. 19-42-703014 (Russia); Grant No. MCIN/AEI/10.13039/501100011033, ERDF “a way of making Europe,” and the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia María de Maeztu, Grant No. MDM-2017-0765, and Programa Severo Ochoa del Principado de Asturias (Spain); the Stavros Niarchos Foundation (Greece); the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University, and the Chulalongkorn Academic into its 2nd Century Project Advancement Project (Thailand); the Kavli Foundation; the Nvidia Corporation; the SuperMicro Corporation; the Welch Foundation, Contract No. C-1845; and the Weston Havens Foundation (USA).


[28] HEPData record for this analysis, 10.17182/hepdata.114357.


SEARCH FOR INVISIBLE DECAYS OF THE HIGGS BOSON … PHYS. REV. D 105, 092007 (2022)

17 State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
18 Sun Yat-Sen University, Guangzhou, China
19 Institute of Modern Physics and Key Laboratory of Nuclear Physics and Ion-beam Application (MOE)—Fudan University, Shanghai, China
20 Zhejiang University, Hangzhou, China, Zhejiang, China
21 Universidad de Los Andes, Bogota, Colombia
22 Universidad de Antioquia, Medellin, Colombia
23 University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia
24 University of Split, Faculty of Science, Split, Croatia
25 Institute Rudjer Boskovic, Zagreb, Croatia
26 University of Cyprus, Nicosia, Cyprus
27 Charles University, Prague, Czech Republic
28 Escuela Politecnica Nacional, Quito, Ecuador
29 Universidad San Francisco de Quito, Quito, Ecuador
30 Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt
31 Center for High Energy Physics (CHEP-FU), Fayoum University, El-Fayoum, Egypt
32 National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
33 Department of Physics, University of Helsinki, Helsinki, Finland
34 Helsinki Institute of Physics, Helsinki, Finland
35 Lappeenranta University of Technology, Lappeenranta, Finland
36 IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
37 Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
38 Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, France
39 Institut de Physique des 2 Infinis de Lyon (IP2I), Villeurbanne, France
40 Georgian Technical University, Tbilisi, Georgia
41 RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany
42 RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
43 RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany
44 Deutsches Elektronen-Synchrotron, Hamburg, Germany
45 University of Hamburg, Hamburg, Germany
46 Karlsruher Institut fuer Technologie, Karlsruhe, Germany
47 Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece
48 National and Kapodistrian University of Athens, Athens, Greece
49 National Technical University of Athens, Athens, Greece
50 University of Ioannina, Ioannina, Greece
51 MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary
52 Wigner Research Centre for Physics, Budapest, Hungary
53 Institute of Nuclear Research ATOMKI, Debrecen, Hungary
54 Institute of Physics, University of Debrecen, Debrecen, Hungary
55 Karoly Robert Campus, MATE Institute of Technology, Gyöngyös, Hungary
56 National Institute of Science Education and Research, HBNI, Bhubaneswar, India
57 Panjab University, Chandigarh, India
58 University of Delhi, Delhi, India
59 Saha Institute of Nuclear Physics, HBNI, Kolkata, India
60 Indian Institute of Technology Madras, Madras, India
61 Bhabha Atomic Research Centre, Mumbai, India
62 Tata Institute of Fundamental Research-A, Mumbai, India
63 Tata Institute of Fundamental Research-B, Mumbai, India
64 Indian Institute of Science Education and Research (IISER), Pune, India
65 Isfahan University of Technology, Isfahan, Iran
66 Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
67 University College Dublin, Dublin, Ireland
68 INFN Sezione di Bari, Bari, Italy
69 INFN Sezione di Bologna, Università di Bologna, Bologna, Italy
Also at Kafkas University, Kars, Turkey.
Also at Istanbul Bilgi University, Istanbul, Turkey.
Also at Hacettepe University, Ankara, Turkey.
Also at Istanbul University—Cerrahpasa, Faculty of Engineering, Istanbul, Turkey.
Also at Ozyegin University, Istanbul, Turkey.
Also at Vrije Universiteit Brussel, Brussel, Belgium.
Also at School of Physics and Astronomy, University of Southampton, Southampton, United Kingdom.
Also at Rutherford Appleton Laboratory, Didcot, United Kingdom.
Also at IPPP Durham University, Durham, United Kingdom.
Also at Monash University, Faculty of Science, Clayton, Australia.
Also at Università di Torino, Torino, Italy.
Also at Bethel University, St. Paul, Minneapolis, USA.
Also at Karamanoğlu Mehmetbey University, Karaman, Turkey.
Also at United States Naval Academy, Annapolis, Maryland, USA.
Also at Ain Shams University, Cairo, Egypt.
Also at Bingol University, Bingol, Turkey.
Also at Georgian Technical University, Tbilisi, Georgia.
Also at Sinop University, Sinop, Turkey.
Also at Erciyes University, Kayseri, Turkey.
Also at Institute of Modern Physics and Key Laboratory of Nuclear Physics and Ion-beam Application (MOE)—Fudan University, Shanghai, China.
Also at Texas A&M University at Qatar, Doha, Qatar.
Also at Kyungpook National University, Daegu, Korea.