CMS Physics Analysis Summary

Contact: cms-pag-conveners-b2g@cern.ch 2021/07/29

Search for heavy resonances decaying to a pair of boosted Higgs bosons in final states with leptons and a bottom quark-antiquark pair at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search for new heavy resonances decaying to a pair of Higgs bosons in proton-proton collisions at a center-of-mass energy of 13 TeV is presented. Data were collected by the CMS detector at the LHC from 2016 to 2018, corresponding to an integrated luminosity of 138 fb^{-1}. The search considers resonances with a mass between 0.8 and 4.5 TeV using events in which one Higgs boson decays into a bottom quark-antiquark pair and the other decays into final states with either one or two charged leptons. Specifically, these include the single-lepton final state of the $\text{HH} \to \text{b}\bar{\text{b}}\text{WW}^* \to \text{b}\bar{\text{b}}\ell\nu\text{q}\bar{\text{q}}$ decay and the dilepton final states of both the $\text{HH} \to \text{b}\bar{\text{b}}\text{WW}^* \to \text{b}\bar{\text{b}}\ell\nu\ell\nu$ and $\text{HH} \to \text{b}\bar{\text{b}}\tau\tau \to \text{b}\bar{\text{b}}\ell\nu\nu\ell\nu\nu$ decays, where ℓ in the final state corresponds to e or μ. The signal is extracted using a two-dimensional maximum likelihood fit of the $\text{H} \to \text{b}\bar{\text{b}}$ jet mass and HH invariant mass distributions. No significant excess above the standard model expectation is observed in data. Model-independent exclusion limits are placed on the the product of the cross section and branching fraction ($\sigma\mathcal{B}$) for spin-0 and spin-2 massive bosons decaying to HH. The results are interpreted in the context of radion and bulk graviton production in models with a warped extra spatial dimension.

1 Introduction

The discovery of a Higgs boson (H) [1–3] established the existence of at least a simple mass generation mechanism for the standard model (SM) [4, 5], the so-called "Higgs mechanism." A number of theoretical difficulties found in the simple model are ameliorated by a so-called "extended Higgs sector" [6]. Supersymmetry [7–14] requires such an extended Higgs sector, with additional spin-0 particles. Models with warped extra dimensions, proposed by Randall and Sundrum [15], postulate the existence of a compact fourth spatial dimension with a warped metric. Such compactification creates heavy resonances arising as a tower of Kaluza–Klein excitations, leading to possible spin-0 radions [16–19] or spin-2 bulk gravitons [20–22]. The ATLAS [23–38] and CMS [39–57] Collaborations have conducted a number of searches for these particles, where the new bosons decay into vector bosons and/or Higgs bosons (WW, ZZ, WZ, HH, ZH, or WH).

In this note, we present expansion of a previous search [58] for narrow resonances (X) decaying to HH. In Ref. [58], we searched in a smaller dataset of proton-proton (pp) collisions for signal in which one H decays to a bottom quark-antiquark pair ($b\bar{b}$) and the second decays to a W boson pair WW^*, with one W boson off-shell, that subsequently decays semileptonically to $\ell\nu q\bar{q}$. This dataset corresponded to collisions at $\sqrt{s} = 13\,\text{TeV}$ recorded in 2016 with $\mathcal{L} = 36\,\text{fb}^{-1}$. In this search, we include two other signal decay channels by considering dilepton decays of the H boson that does not decay to $b\bar{b}$: the $H \to WW^* \to \ell\nu\ell\nu$ and the $H \to \tau\tau \to \ell\nu\nu\ell\nu\nu$ decays. In all three cases, the ℓ denotes an electron or a muon; the analysis is sensitive to leptonically decaying τ leptons in the $b\bar{b}\ell\nu q\bar{q}$ final state. The analysis is optimized for the three $X \to HH$ channels just mentioned. Signal events from $HH \to b\bar{b}ZZ$ are included in our acceptance and constitute 1–3% of the total expected signal yield.

The search is performed on a dataset collected from 2016 to 2018 at the CERN LHC, corresponding to an integrated luminosity of $138\,\text{fb}^{-1}$, and considering resonances in the mass range $0.8 < m_X < 4.5\,\text{TeV}$. The Higgs bosons have a high Lorentz boost because of the large values of m_X considered, so the decay products of each one are contained in a collimated cone. The distinguishing characteristic of the signal is a peak in the two-dimensional plane of the $H \to b\bar{b}$ jet (hereafter called $b\bar{b}$ jet) mass $m_{b\bar{b}}$ and the reconstructed HH invariant mass m_{HH}.

In the single-lepton (SL) channel, the quarks in the $H \to WW^* \to \ell\nu q\bar{q}$ decay are reconstructed as a single large jet, referred to as the $q\bar{q}$ jet, with a nearby lepton (e or μ). This jet is required to have a reconstructed topology consistent with a substructure arising from a decay to two quarks. This Higgs boson decay chain is reconstructed as the $q\bar{q}$ jet, the lepton, and the missing transverse momentum. In the dilepton (DL) channel, we reconstruct two leptons in close proximity to each other and with missing transverse momentum in the same azimuthal hemisphere, consistent with the neutrinos. In all channels considered, the $H \to b\bar{b}$ decay is reconstructed as a single large jet with substructure and high transverse momentum p_T.

The main standard model (SM) background to this search arises from top quark pair production ($t\bar{t}$). The top quarks in this analysis have collimated decay products because of large Lorentz boosts. In the SL channel, the largest background comes from semileptonic $t\bar{t}$ decays: one top quark decays to a charged lepton and a neutrino ($t \to Wb \to \ell\nu b$), and the other decays exclusively to quarks ($t \to Wb \to q\bar{q}b$), which can be mistakenly reconstructed as the $b\bar{b}$ jet candidate. Other significant backgrounds in this channel are production of W bosons in association with jets (with $W \to \ell\nu$, hereafter referred to as W+jets) and QCD multijet events with either a non-rpompt lepton or a hadron misidentified as a lepton. In the DL channel, the background yield is smaller than that of the SL channel by a factor of ≈60. Top quark pair pro-

duction is also the dominant background with approximately equal contributions from events where both top quarks decay leptonically and events where one decays hadronically but a lepton is misidentified. The other significant background in this channel is production of Z/γ^* bosons in association with jets (Z/γ^*+jets). These backgrounds are distinguished in data using the $m_{b\bar{b}}$ spectrum. Contributions from backgrounds with a SM Higgs boson (e.g. $t\bar{t}H$) are considered but found to be negligible in both channels.

In this analysis, the events are divided into 12 exclusive categories by the number of leptons, lepton flavor, b-tag quality, and $H \to WW^*$ decay kinematics. The signal and SM background yields are estimated using a simultaneous maximum likelihood fit to the two-dimensional $m_{b\bar{b}}$ and m_{HH} mass distributions in all 12 categories.

In addition to using a larger dataset and incorporating new DL signal modes, this analysis also employs new particle reconstruction and identification techniques in the SL channel. These include more efficient algorithms for identifying electrons and jets with b hadrons (b tagging) as well as an improved reconstruction procedure for the $H \to WW^* \to \ell\nu q\bar{q}$ decay.

2 The 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 (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gaseous detectors embedded in the steel flux-return yoke outside the solenoid. Events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of about 4 μs [59]. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage [60].

3 Simulated samples

Signal and background yields are extracted from a fit to the data in the two dimensional $m_{b\bar{b}}$ and m_{HH} mass distribution using templates obtained from samples generated from Monte Carlo simulation.

The signal processes $pp \to X \to HH \to b\bar{b}WW^*$ and $pp \to X \to HH \to b\bar{b}\tau\tau$ are simulated for both spin-0 and spin-2 X bosons. These X bosons are produced via gluon fusion and have a narrow width (1 MeV) that is small compared to the experimental resolution. The branching fractions used for H decays are the SM ones assuming $m_H = 125.09$ GeV [61]. The signal is generated at leading order (LO) using the MADGRAPH5_aMC@NLO V5 2.4.2 generator [62] with the MLM merging scheme [63] for m_X between 0.8 and 4.5 TeV.

The MADGRAPH5_aMC@NLO generator is also used to produce the W+jets, $Z \to \ell\ell$, and QCD multijet background samples at leading order (LO). The W+jets and $Z \to \ell\ell$ samples are normalized using NNLO cross sections, calculated with FEWZ V3.1 [64]. Samples of WZ diboson production and the associated production of $t\bar{t}$ with either a W or Z boson ($t\bar{t}$ + V) are also generated with MADGRAPH5_aMC@NLO but at next-to-leading-order (NLO) with the FxFx

jet merging scheme [65]. The POWHEG v2 generator is used to produce samples for $t\bar{t}$, WW, ZZ, $t\bar{t}H$, and single top production at NLO [66–73]. Furthermore, the $t\bar{t}$ process is renormalized to the NNLO cross section, computed with TOP++ v2.0 [74].

The parton showering and hadronization are simulated with PYTHIA 8.226 [75] using the CUETP8M1 [76] tune for the 2016 samples. For 2017 and 2018, PYTHIA 8.230 and the CP5 tune [77] are used to produce the samples. Additionally, some 2016 datasets were produced with the CP5 tune, namely $t\bar{t}$, $t\bar{t}H$, and the $X \to HH \to b\bar{b}VV^*$ signal. The parton distribution functions used to produce the samples are the NNPDF 3.0 [78] set for 2016 datasets and the NNPDF 3.1 [79] set for 2017 and 2018 datasets. The simulation of the CMS detector is performed with the GEANT4 [80] toolkit. Additional pp collisions in the same or nearby bunch crossings (pileup) are simulated and the samples are weighted to have the same distribution of the multiplicity of pileup interactions in data.

4 Event reconstruction

All signal events, regardless of lepton multiplicity, feature a high-p_T jet that has substructure consistent with two b quark decays. This jet is opposite in the transverse plane to a collection of other physics objects from a boosted decay. In the SL channel, signal events feature a lepton originating from a W boson decay and a nearby jet that has substructure consistent with a $W \to q\bar{q}$ decay. In the DL channel, there are two high-p_T leptons originating from the boosted decay of either a W boson pair or a τ lepton pair, but there is no jet in the vicinity of the leptons, resulting in a cleaner experimental signature.

Event reconstruction relies on the particle-flow (PF) algorithm [81], which identifies each individual particle in an event with an optimized combination of information from the various elements of the CMS detector. The missing transverse momentum vector $\vec{p}_T^{\,miss}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as p_T^{miss} [82]. The $\vec{p}_T^{\,miss}$ is modified to account for corrections to the energy scale of the reconstructed jets in the event. The candidate vertex with the largest value of summed physics-object p_T^2 is taken to be the primary pp interaction vertex. The physics objects are the jets, clustered using the anti-k_T jet finding algorithm [83, 84] 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.

4.1 Electron and muon identification

Different selection criteria are required for the SL and DL channels to identify signal-like leptons. The section describes the selection criteria for a lepton to qualify as a candidate in either the SL or DL channels, denoted as the SL or DL lepton criteria. Each lepton candidate can be a candidate in both channels. In Sec. 5.1, we describe the procedure for determining whether an event is classified in the SL or DL channels.

Any event in both channels must contain either a muon with $p_T > 27\,GeV$ or an electron with $p_T > 30\,GeV$. In the DL lepton criteria, the other lepton must have $p_T > 10\,GeV$. All muons are required to have $|\eta| < 2.4$. Electrons in the DL lepton criteria are required to have $|\eta| < 2.5$, but those in the SL lepton criteria are restricted to $|\eta| < 1.479$ to suppress a significant contribution from the QCD multijet background at low expense to signal acceptance. Leptons must satisfy reconstruction quality and identification requirements that are optimized to maintain large efficiencies and low rates misidentifying hadrons as leptons [85–87]. Additionally, the impact parameter of lepton tracks with respect to the primary vertex is required to be consistent with

those originating from this vertex. Looser constraints on the impact parameter are used in the DL channel because some of the leptons originate from $H \to \tau\tau$ decays. Leptons are required to be isolated with an isolation cone size designed for leptons from boosted decays, in which the cone size becomes smaller with larger p_T [58]. The allowed extra transverse energy in the isolation cone is smaller for the DL lepton criteria than for the SL lepton criteria.

In the SL channel, signal muons are selected with approximately 90% efficiency for $m_X = 0.8\,\text{TeV}$ down to 60% efficiency at $m_X = 4.5\,\text{TeV}$, with isolation as the leading source of inefficiency. The electron selection is less efficient, from 70% at $m_X = 0.8\,\text{TeV}$ to 7.5% at $m_X = 4.5\,\text{TeV}$. The leading source of electron inefficiency is a selection imposed at a low-level reconstruction step on the ratio of the energy deposited in the HCAL to that deposited in the ECAL. Electrons in the $H \to WW^* \to e\nu q\bar{q}$ decay often fail this selection because of the nearby energy deposits.

In the DL channel, where there is no $q\bar{q}$ jet, the lepton selection efficiency is larger for all m_X than in the SL channel. Due to the increased boost of the system, the efficiency still drops toward high m_X, though not as severely as for the SL channel. For electrons, the reconstruction efficiency is higher because there is no jet. The electron efficiency ranges from approximately 82% at $m_X = 0.8\,\text{TeV}$ down to 71% at $m_X = 4.5\,\text{TeV}$. For muons, most of the inefficiency is due to the isolation, as in the SL channel. The muon efficiency ranges from approximately 96% at $m_X = 0.8\,\text{TeV}$ down to 91% at $m_X = 4.5\,\text{TeV}$. The lepton efficiencies are measured in simulation and data in a $Z \to \ell\ell$ sample, and the simulation is corrected to match the efficiency in data. The systematic uncertainties in these measurements are applied to the normalization of the signal.

4.2 Reconstruction and flavor identification of jets

Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. Additional pp interactions within the same or nearby bunch crossings can contribute additional tracks and calorimetric energy depositions, increasing the apparent jet momentum. The pileup per particle identification algorithm [88, 89] is used to mitigate the effect of pileup at the reconstructed particle level, making use of local shape information, event pileup properties, and tracking information. Charged particles identified to be originating from pileup vertices are discarded. For each neutral particle, a local shape variable is computed using the surrounding charged particles compatible with the primary vertex within the tracker acceptance ($|\eta| < 2.5$), and using both charged and neutral particles in the region outside of the tracker coverage. The momenta of the neutral particles are then rescaled according to their probability to originate from the primary interaction vertex deduced from the local shape variable, superseding the need for jet-based pileup corrections [88]. Jet energy corrections are derived from simulation studies so that the average measured energy of jets becomes identical to that of particle level jets.

Two types of jets, both clustered with the anti-k_T algorithm [83, 84], are used in this analysis according to their cluster radius R: those clustered with a distance parameter $R = 0.8$ (AK8 jets) and $R = 0.4$ (AK4 jets). Because of the boost imparted to the Higgs bosons by the much more massive X boson, both the $H \to b\bar{b}$ and $W \to q\bar{q}$ decays are each reconstructed as a single, merged AK8 jet with two-prong substructure. In the SL channel, the PF particles associated with the lepton are not included in the clustering of the set of jets from which the $q\bar{q}$ jet is selected in order to prevent the $q\bar{q}$ jet from containing the lepton's momentum. The AK4 jets are used primarily to suppress background from $t\bar{t}$ production by identifying additional jets from b quarks and vetoing events that contain these jets. Jets of both types are required to have

$|\eta| < 2.4$ so that a majority of the jet area is within the acceptance of the tracker.

Because of the large mass of the X boson, the Higgs bosons are boosted, with collimated decays, and typically produced opposite each other in the transverse plane, i.e., $\Delta\phi \approx \pi$, where ϕ is the azimuthal angle. In the SL channel, the $\mathrm{b\overline{b}}$ jet candidate is required to be $\Delta\phi > 2.0$ from the lepton and $\Delta R > 1.6$ from the $\mathrm{q\overline{q}}$ jet. The $\mathrm{q\overline{q}}$ jet is chosen as the closest AK8 jet in ΔR to the lepton, provided that it is found within $\Delta R < 1.2$ from the lepton and has $p_\mathrm{T} > 50\,\mathrm{GeV}$. In the DL channel, we require the $\mathrm{b\overline{b}}$ jet candidate to be $\Delta\phi > 2.0$ from the dilepton momentum and that neither lepton overlap with the $\mathrm{b\overline{b}}$ jet cone. Within both the $\mathrm{b\overline{b}}$ jet and the $\mathrm{q\overline{q}}$ jet, two subjets are reconstructed that must each have $p_\mathrm{T} > 20\,\mathrm{GeV}$. Constituents of the AK8 jets are first reclustered using the Cambridge–Aachen algorithm [90, 91]. The "modified mass drop tagger" algorithm [92, 93], also known as the "soft drop" (SD) algorithm, with angular exponent $\beta = 0$, soft cutoff threshold $z_\mathrm{cut} < 0.1$, and characteristic radius $R_0 = 0.8$ [94], is applied to remove soft, wide-angle radiation from the jet. The subjets used are those remaining after the algorithm has removed all recognized soft radiation. The jet mass is the invariant mass of these two subjets. The SD mass of the $\mathrm{b\overline{b}}$ jet is used to obtain the search variable $m_{\mathrm{b\overline{b}}}$, after applying p_T-dependent corrections so that $m_{\mathrm{b\overline{b}}}$ in simulation is on average equal to the Higgs boson mass 125 GeV.

Jets are identified as likely to have originated from b hadron decays with different strategies depending on the size of the jet. The AK4 jets considered for b tagging are required to have $p_\mathrm{T} > 30\,\mathrm{GeV}$. These jets are identified as b jets using the DeepJet tagger [95–97] at a working point that has an efficiency of ≈80% for selecting b jets and a misidentification rate of ≈1% for light jets.

The AK8 jets are identified as having substructure consistent with a $\mathrm{b\overline{b}}$ decay using the DeepAK8 mass-decorrelated $\mathrm{Z/H} \to \mathrm{bb}$ tagger [98] designed for large, high-p_T jets. The b tagging efficiencies are measured in data, and the simulation is corrected for any discrepancies.

4.3 Reconstructing the Di-Higgs mass

Depending on whether the final state has one or two leptons, different strategies are employed to reconstruct the momentum of the Higgs boson that does not decay to $\mathrm{b\overline{b}}$. The mass m_HH is then the invariant mass of this momentum and the $\mathrm{b\overline{b}}$ jet momentum. In Sec. 4.3.1 and 4.3.2 respectively, the reconstruction strategies are described for the SL and DL channel.

4.3.1 Single-lepton channel

To reconstruct the Higgs boson four-momentum in the $\mathrm{H} \to \mathrm{WW}^* \to \ell\nu\mathrm{q\overline{q}}$ decay chain from the visible and invisible decay products, a likelihood-based technique is employed. For each event, values for the following five parameters are extracted by maximizing a likelihood function:

- $\vec{p}_\nu$: the three components of the neutrino momentum.
- $V_{\mathrm{q\overline{q}}}$: an indicator of whether the $\mathrm{q\overline{q}}$ jet favors a larger or smaller mass than the leptonic W boson decay.
- $R_{\mathrm{q\overline{q}}}$: the jet response, a multiplicative scale factor on the p_T of the $\mathrm{q\overline{q}}$ jet. We allow the jet p_T to vary because the $\mathrm{q\overline{q}}$ jet has large energy resolution.

With these parameters, the $\mathrm{H} \to \mathrm{WW}^*$ momentum can be fully determined. The $\mathrm{H} \to \mathrm{WW}^*$ momentum is then the summed momentum of the neutrino momentum p_ν, the $\mathrm{q\overline{q}}$ jet momen-

tum (modified by $R_{q\bar{q}}$), and the momentum of the lepton.

The likelihood function is constructed with six probability density functions (pdfs) $P(x|\vec{y})$ estimated from signal simulation, where x is the corresponding observable for the pdf (such as m_{jet}). The symbol $\vec{y}$ represents the set of free parameters associated to that pdf, such as $V_{q\bar{q}}$. These pdfs are represented as one-dimensional histograms. The full likelihood function is:

$$\mathcal{L} = P(m_{\text{jet}}|V_{q\bar{q}})P(R_{\text{jet}}|R_{q\bar{q}}, V_{q\bar{q}})P(m_{\ell\nu q\bar{q}}|\vec{p}_\nu, R_{q\bar{q}}, V_{q\bar{q}})P(m_{\ell\nu}|\vec{p}_\nu, V_{q\bar{q}})P(\vec{p}_{\text{T}}^{\,\text{miss}}|\vec{p}_\nu, V_{q\bar{q}}). \quad (1)$$

The observable m_{jet} is the SD mass of the $q\bar{q}$ jet and its corresponding pdf is coarsely binned to remain insensitive to the precise modeling of the soft-drop algorithm. The observable R_{jet} is the jet p_{T} response. Another two observables $m_{\ell\nu}$ and $m_{\ell\nu q\bar{q}}$ are masses of their corresponding composite physics objects.

The last term in Eq. 1 represents the product of two pdfs, each corresponding to a single component of the missing transverse momentum $\vec{p}_{\text{T}}^{\,\text{miss}}$:

$$P(\vec{p}_{\text{T}}^{\,\text{miss}}|\vec{p}_\nu, V_{q\bar{q}}) = P(p_{T,\parallel}^{\text{miss}}|\vec{p}_\nu, V_{q\bar{q}})P(p_{T,\perp}^{\text{miss}}|\vec{p}_\nu, V_{q\bar{q}}). \quad (2)$$

The two observables $p_{T,\parallel}^{\text{miss}}$ and $p_{T,\perp}^{\text{miss}}$ are defined with respect to the reference frame of the $H \to WW^*$ decay, along the direction of $\vec{p}_{\text{T}}^{\,\text{reco}}$:

$$\vec{p}_{\text{T}}^{\,\text{reco}} = \vec{p}_{\text{T}}^{\,\text{miss}} + (\vec{p}_\ell + \vec{p}_{q\bar{q}\text{ jet}})_{\text{T}}. \quad (3)$$

The two $\vec{p}_{\text{T}}^{\,\text{miss}}$ pdf terms are parameterized as the extra $p_{\text{T}}^{\text{miss}}$ (relative to the neutrino momentum) that is parallel and perpendicular to this vector $\vec{p}_{\text{T}}^{\,\text{reco}}$. The extra $p_{\text{T}}^{\text{miss}}$ along this direction arises from mismeasurement of the $b\bar{b}$ jet, while the orthogonal component arises from pileup and the underlying event.

The pdfs P of the observables are generally independent of m_X, but there is still some residual dependence. We account for this by producing two sets of pdfs, one at low $p_{\text{T}}^{\text{reco}}$ and one at high $p_{\text{T}}^{\text{reco}}$. Then, event-by-event, the histogram of the pdf is obtained by interpolating between the two histograms at the two regimes of $p_{\text{T}}^{\text{reco}}$. This interpolation is performed linearly as a function of the event $p_{\text{T}}^{\text{reco}}$. The P are all dependent on whether the hadronically decaying W boson is heavier than the leptonically decaying W, so each term is dependent on the free parameter $V_{q\bar{q}}$. Otherwise, the P are largely uncorrelated.

In Ref. [58], this reconstruction was performed by enforcing a 125 GeV mass value on $H \to WW^*$ and then solving for $\vec{p}_\nu$ with a simple second-order equation, using the $\vec{p}_{\text{T}}^{\,\text{miss}}$ directly for the transverse components of $\vec{p}_\nu$. The signal m_{HH} resolution is very similar between the previous and the current method, and the background events tend to move to lower masses than before. We then take advantage of this fact with an alternative likelihood $\mathcal{L}_{\text{alt}}$, which is less constrained by the intermediate masses. Instead of fitting for the neutrino p_z, we include the masses $m_{\ell\nu}$ and $m_{\ell\nu q\bar{q}}$ as free parameters. We then use both likelihoods to construct a discriminating variable between signal and background:

$$D_{\ell\nu q\bar{q}} = -2\log \mathcal{L}/\mathcal{L}_{\text{alt}}, \quad (4)$$

where $\mathcal{L}$ is the likelihood described in Eq. 1. We discuss how $D_{\ell\nu q\bar{q}}$ is used in Sec. 5.

4.3.2 Dilepton channel

Because of the absence of a $q\bar{q}$ jet, the presence of larger p_T^{miss}, and much smaller backgrounds, there is no need for a likelihood-based technique to separate signal and background in the DL channel. Instead, we make simple assumptions of the decay kinematics. First, because the decay of the boosted Higgs boson is collimated, we assume the polar angle θ of the total invisible momentum due to neutrinos to be identical to the that of the dilepton momentum: $\theta_{inv} = \theta_{\ell\ell}$. With this constraint, the z-component of the invisible four-momentum p_{inv} is obtained. The invisible invariant mass m_{inv} due to neutrinos is estimated to be 55 GeV, the average of the true distribution from signal simulation. The invisible four-momentum is reconstructed by directly using $\vec{p}_T^{miss}$ as the transverse components of p_{inv}. The corresponding Higgs boson four-momentum is the summed four-momentum of p_{inv} and the dilepton four-momentum $p_{\ell\ell}$.

5 Event selection and categorization

Events are selected in this search if they pass the following criteria that indicate they could originate from a X boson decay and are then divided into 12 distinct categories (eight SL and four DL). A separate set of criteria is used to define control regions, which are used to validate the modeling of background processes.

5.1 Event preselection

We define the event preselection as the criteria imposed for all events before an event is considered for placement into one of the 12 SL or DL categories. Events are selected by the trigger system with small year-to-year differences. Events are triggered if they contain one of the following: an isolated muon with $p_T > 24$ GeV (27 GeV in 2017), an isolated electron with $p_T > 32$ GeV (27 GeV in 2016), or $H_T > 1050$ GeV (900 in 2016), where H_T is the scalar sum of jet p_T for all online AK4 jets with $p_T > 30$ GeV. A combination (inclusive OR) of lepton and H_T triggers is used because the high-m_X SL signal does not have leptons that are sufficiently isolated to pass the online lepton isolation selection, as the decay products $WW^* \to \ell\nu q\bar{q}$ are highly collimated. Additional multi-object triggers that select events with at least one lepton and $H_T > 450$ GeV (400 GeV in 2016) supplement these two single-object triggers, thereby maintaining high trigger efficiency for signal over the entire range of m_X. The trigger efficiency is measured for $t\bar{t}$ events in data for events passing offline selection criteria for H_T in the SL channel and both H_T and lepton p_T in the DL channel. We use $t\bar{t}$ events because the lepton and jet multiplicities resemble those in signal events. Simulation is corrected such that the trigger efficiency matches that in the data. In the SL channel, the trigger efficiency for signal events at over 96% at $m_X = 0.8$ TeV and increases to $> 99\%$ above $m_X = 1.0$ TeV. In the DL channel, the trigger efficiency is $> 99\%$ over the full range of m_X.

Offline, all events are required to have $H_T > 400$ GeV and either one electron with $p_T > 30$ GeV or one muon with $p_T > 27$ GeV. Events are also required to have a single $b\bar{b}$ jet. Background from $t\bar{t}$ production is reduced by vetoing events with an AK4 jet that is $\Delta R > 1.2$ from the $b\bar{b}$ jet and is identified as a b jet as in Sec. 4.2.

We ensure that the sets of events belonging to the SL and DL channels are disjoint. To accomplish this, we first impose that any event with exactly two oppositely charged lepton candidates passing the DL lepton criteria be assigned as a dilepton event. Otherwise, if the event has a lepton candidate passing the SL lepton criteria and also has fewer than two lepton candidates passing the DL lepton criteria, it is classified as a SL event, and the lepton candidate with the largest p_T is selected as the event lepton. Otherwise, the event is not used for analysis.

Distributions of important variables are shown in Fig. 1 for the SL channel and in Fig. 2 for the DL channel. In each plot, the full selection to enter any of the relevant search region categories is applied, except any selection on the displayed variable. We describe each of these variables in the following sections.

5.2 Single-lepton channel event selection

In addition to the lepton, events in this channel must have a $\mathrm{q\bar{q}}$ jet, chosen only as the closest two-prong jet to the lepton and required to have $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $\Delta R < 1.2$ from the lepton. Jets in background events tend to be produced at higher $|\eta|$ than those produced in signal events, which contain jets from the decay of a heavy particle. To exploit this property, the ratio of the $\mathrm{H} \to \mathrm{WW}^*$ p_{T} divided by m_{HH}, denoted as p_{T}/m, is required to be > 0.3 and is shown in Fig. 1, .

5.3 Dilepton channel event selection

Events in the DL channel must pass additional criteria. In signal events, the invariant mass of the two leptons is kinematically constrained by the mass of the boosted Higgs boson from which they originate, as shown in Fig. 2, peaking near 30 GeV. Background in the $m_{\ell\ell}$ spectrum from Z/γ^*+jets populates predominantly lower masses from the continuum and higher masses from the Z boson. Background from $\mathrm{t\bar{t}}$ mostly populates higher masses since the leptons are typically opposite each other in the transverse plane. Requiring the dilepton invariant mass to satisfy $6 < m_{\ell\ell} < 75\,\mathrm{GeV}$ suppresses these backgrounds while preserving the signal. Enforcing that the leptons be close together in η-ϕ space with $\Delta R_{\ell\ell} < 1.0$ further helps to suppress the $\mathrm{t\bar{t}}$ background. In Z/γ^*+jets, the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is typically in the direction of the $\mathrm{b\bar{b}}$ jet, away from the leptons, due to jet mismeasurements, while in signal the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is close to the leptons because of the boosted Higgs decay. Thus, we also require that $|\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\ell\ell})| < \pi/2$ to discriminate against Z/γ^*+jets. Background is further separated from signal by requiring that there be a significant amount of missing transverse momentum in the event with $p_{\mathrm{T}}^{\mathrm{miss}} > 85\,\mathrm{GeV}$.

5.4 Event categorization

Events are categorized by event properties that reflect the signal purity, and the categorization is the same over the full range of m_{X}. In the SL channel, electron and muon events are separated because their reconstruction efficiencies for background and signal are different, resulting in different signal purities. The electron and muon categories are labeled "e" and "μ," respectively, in the figures. Likewise, in the DL channel, same-flavor and opposite-flavor lepton events are separated because the background composition is different between these two cases. These are labeled "SF" and "OF," respectively, in the figures. We do not separate ee from $\mu\mu$ in the DL channel because these events have similar ratios of signal to background. For all events, there are two categories for $\mathrm{b\bar{b}}$ jet tagging, constructed from different subsets of the distribution of the DeepAK8 mass-decorrelated $\mathrm{Z/H} \to \mathrm{b\bar{b}}$ discriminator $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}}$, introduced in Sec. 4.2. The discriminator value ranges from 0.0 to 1.0, where larger values indicate that the jet is more consistent with $\mathrm{b\bar{b}}$ substructure. We use working points designed to maintain large signal efficiency for decays from massive particles; a loose category defined by $0.8 \leq D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} < 0.97$ is labelled as "bL," and a tight category defined by $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} \geq 0.97$ is labelled as "bT."

One more criterion for categorization, related to the $\mathrm{H} \to \mathrm{WW}^* \to \ell\nu\mathrm{q\bar{q}}$ decay, is implemented for the SL channel but not the DL channel. This categorization relies on both the τ_2/τ_1 N-subjettiness ratio [99] of the $\mathrm{q\bar{q}}$ jet (denoted now as τ_2/τ_1) and the $\mathrm{H} \to \mathrm{WW}^*$ likelihood

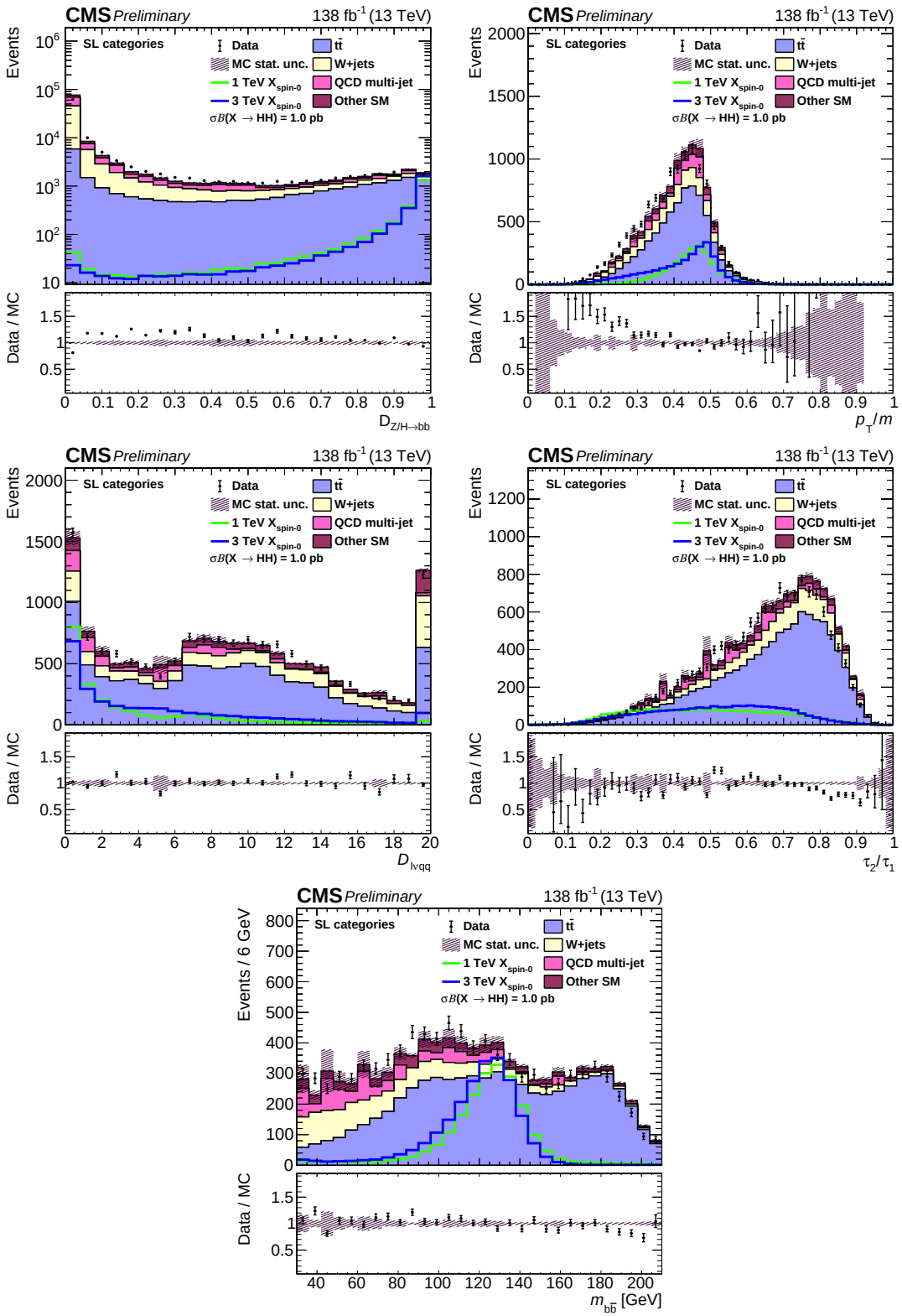

Figure 1: Single-lepton channel variables: distributions of important variables are shown for data (points), simulated SM processes (filled histograms), and simulated signal (solid lines). The statistical uncertainty of the simulated sample is shown as the hatched band. Spin-0 signals for m_X of 1.0 and 3.0 TeV are displayed. For both signal models, $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH})$ is set to 1.0 pb. The bottom panes of each plot show the ratio of the data to the sum of all background processes.

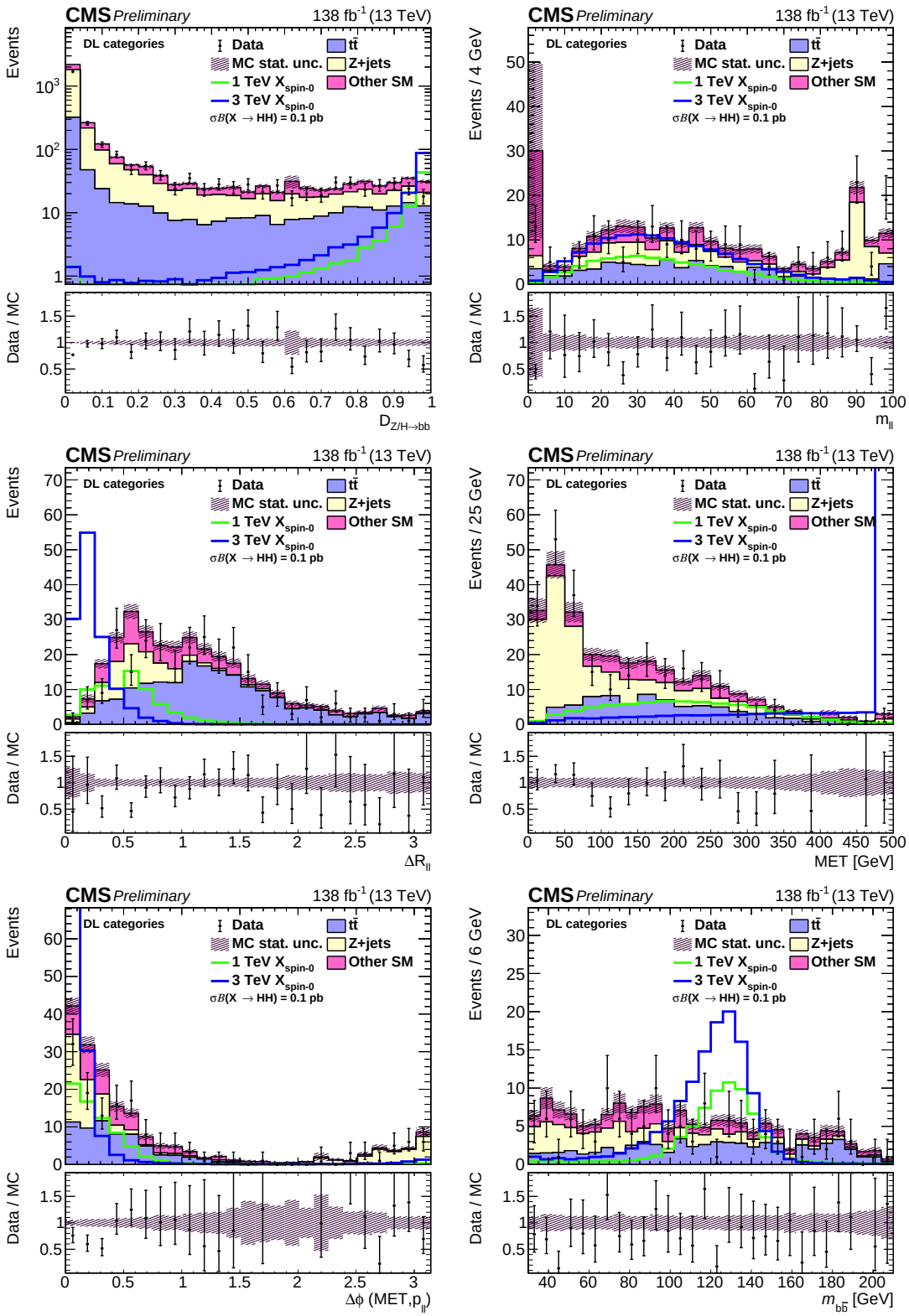

Figure 2: Dilepton channel variables: distributions of important variables are shown for data (points), simulated SM processes (filled histograms), and simulated signal (solid lines). The statistical uncertainty of the simulated sample is shown as the hatched band. Spin-0 signals for m_X of 1.0 and 3.0 TeV are displayed. For both signal models, $\sigma\mathcal{B}(X \to HH)$ is set to 0.1 pb. The bottom panes of each plot show the ratio of the data to the sum of all background processes.

discriminator $D_{\ell\nu q\bar{q}}$ that was first introduced in Eq. 1. The ratio τ_2/τ_1 measures how consistent the jet substructure is with a two-prong decay versus a single-prong decay, with lower values more strongly indicating a two-prong decay. The distributions of $D_{\ell\nu q\bar{q}}$ and τ_2/τ_1 are shown in Fig. 1.

All events in the SL but not in the DL search region are required to satisfy both $\tau_2/\tau_1 \leq 0.75$ and $D_{\ell\nu q\bar{q}} \leq 11.0$. We construct a low-purity category (labeled "LP") with events that satisfy either $0.45 < \tau_2/\tau_1 \leq 0.75$ or $2.5 < D_{\ell\nu q\bar{q}} \leq 11.0$ and a high-purity category (labeled "HP") with events that satisfy both $\tau_2/\tau_1 \leq 0.45$ and $D_{\ell\nu q\bar{q}} \leq 2.5$. In 2016 data, the lower working point for τ_2/τ_1 is 0.55 rather than 0.45.

As a result, events are divided into all combinations of categories for a total of 12 exclusive selections (eight SL and four DL). When describing a single selection, the category label is a combination of those listed above. For example, in the SL channel the tightest $b\bar{b}$ jet tagging category with a low-purity τ_2/τ_1 selection in the electron channel is: "e, bT, LP." The categories and their corresponding labels are summarized in Table 1 and 2.

Table 1: SL channel event categorization and corresponding category labels. All combinations of the two lepton flavor, two $b\bar{b}$ jet tagging, and two $H \to WW^*$ decay purity selections are used to form eight independent event categories. The lower τ_2/τ_1 working point is 0.55 in 2016 and 0.45 in 2017 and 2018.

Categorization type	Selection	Category label
Lepton flavor	Electron	e
	Muon	μ
$b\bar{b}$ jet tagging	$0.8 \leq D_{Z/H\to b\bar{b}} < 0.97$	bL
	$D_{Z/H\to b\bar{b}} \geq 0.97$	bT
$H \to WW^*$ purity	$0.45(0.55) < \tau_2/\tau_1 \leq 0.75$ or $2.5 < D_{\ell\nu q\bar{q}} \leq 11.0$	LP
	$\tau_2/\tau_1 \leq 0.45(0.55)$ and $D_{\ell\nu q\bar{q}} \leq 2.5$	HP

Table 2: DL channel event categorization and corresponding category labels. All combinations of the two lepton flavor and two $b\bar{b}$ jet tagging selections are used to form four independent event categories.

Categorization type	Selection	Category label
Lepton flavor	Two electrons or two muons	SF
	One electron and one muon	OF
$b\bar{b}$ jet tagging	$0.8 \leq D_{Z/H\to b\bar{b}} < 0.97$	bL
	$D_{Z/H\to b\bar{b}} \geq 0.97$	bT

The search is performed in these categories for $30 < m_{b\bar{b}} < 210$ GeV. Extending down to 30 GeV helps to capture the background in the fit, but events below 30 GeV would be relatively difficult to model since these are events for which the SD algorithm removes nearly all of the jet energy. The $m_{b\bar{b}}$ distribution is displayed in Fig. 1. Events with $700 < m_{HH} < 5050$ GeV are considered. The lower bound is chosen such that the m_{HH} distribution is monotonically decreasing for the full background. The upper bound is several hundred GeV above the highest mass event observed in data.

For spin-0 scenarios, the selection efficiency for SL channel events to pass the criteria of any event category is 9% at $m_X = 0.8$ TeV. The efficiency increases with m_X to 23% at $m_X = 1.5$ TeV because the Higgs boson decays become more collimated. Above 1.5 TeV the selection efficiency decreases to a minimum of 14% at $m_X = 4.5$ TeV for two main reasons: the b tagging

efficiency degrades at high-p_T for jets and the lepton isolation worsens for extremely collimated Higgs boson decays. For DL channel events, the combined selection efficiency to pass the criteria of any event category is 9% at $m_X = 0.8$ TeV, increases sharply with m_X to 30% at $m_X = 1.5$ TeV, and then increases more slowly to 36% at $m_X = 4.5$ TeV. The efficiency grows over the full range of m_X because the leptons, with no nearby jet, become easier to select at high p_T.

Tables 3 and 4 show the efficiencies for each individual selection requirement with the full selection otherwise applied in the denominator.

Table 3: SL channel: the efficiency of each selection criterion with the rest of the full selection applied. The efficiencies for the total expected SM background and signal at 1.0 TeV and 3.0 TeV are shown.

SL channel selection	Bkg eff.	1 TeV eff.	3 TeV eff.
b-jet veto	0.31	0.87	0.82
$D_{Z/H\to b\bar{b}} \geq 0.8$	0.07	0.81	0.84
$\tau_2/\tau_1 \leq 0.75$	0.69	0.91	0.92
$D_{\ell\nu q\bar{q}} \leq 11.0$	0.63	0.87	0.83
$p_T/m \geq 0.3$	0.87	0.97	0.86

.

Table 4: DL channel: the efficiency of each selection criterion with the rest of the full selection applied. The efficiencies for the total expected SM background and signal at 1.0 TeV and 3.0 TeV are shown.

DL channel selection	Bkg eff.	1 TeV eff.	3 TeV eff.
b-jet veto	0.45	0.86	0.84
$D_{Z/H\to b\bar{b}} \geq 0.8$	0.05	0.81	0.83
MET ≥ 85 GeV	0.55	0.88	0.97
$6 \leq m_{\ell\ell} \leq 75$ GeV	0.62	0.95	0.94
$\Delta R_{\ell\ell} \leq 1.0$	0.51	0.93	0.998
$\lvert\Delta\phi(\vec{p}_T^{\,miss}, \vec{p}_{\ell\ell})\rvert \leq \pi/2$	0.83	0.98	0.97

Due to the properties of X boson production, the Higgs bosons in spin-2 signal events are produced at lower values of $|\eta|$ than those from spin-0 signal, resulting in a larger selection efficiencies in both channels. The relative increase in efficiency for spin-2 signal is larger at low mass (≈40%) than at high mass (≈15%).

5.5 Control region

Two control regions (CRs) are used to validate the SM background estimation and systematic uncertainties. These regions are signal-depleted by construction, and the events within them are not used to search for signal. The first, labeled "top CR," targets background events with top quarks, particularly $t\bar{t}$. Such events are selected by inverting the AK4 jet b-tag veto. To increase the statistical power of the sample, the p_T/m selection is removed for SL channel events, and the $\Delta R_{\ell\ell}$ selection is altered to $\Delta R_{\ell\ell} > 0.4$ for DL channel events. Events in this control region are then divided into the 12 categories previously described in Sec. 5.4. The $m_{b\bar{b}}$ and m_{HH} distributions in this control region are similar to the distributions in the signal region for backgrounds with top quarks. The top quark p_T spectrum in $t\bar{t}$ events has been shown to be mis-modelled in simulation [100, 101]. A small p_T-dependent correction on the order of a

few percent is measured in an expanded version of this control region and applied to the $\mathrm{t\bar{t}}$ simulation.

While the top CR is an adequate probe of processes that involve top quarks, it is not sensitive to background from Z/γ^*+jets, W+jets, or QCD multijet processes. Instead, a second control region, labeled "non-top CR," is used to study the modeling of these processes. The selection of events in this control region is the same as for the signal region, except that the $\mathrm{b\bar{b}}$ jet is required to be inconsistent with having $\mathrm{b\bar{b}}$ substructure, i.e., $0.01 \leq D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} \leq 0.04$. We exclude events with $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} < 0.01$ due to heavy mismodelling in that region. As a result, events in this control region are not categorized by $\mathrm{b\bar{b}}$ jet tagging, yielding half as many (six) categories here than in the top CR. Due to having fewer categories, the non-top CR cannot in principle test the modeling of the b-tagging of light-flavored jets. Instead, we rely on the top CR to verify that this modeling is well-behaved.

Ultimately, the final values of the normalization and shape and their corresponding uncertainties come from the two-dimensional fit to signal and background in the search region.

6 Background and signal modeling

The search is performed by simultaneously estimating the signal and background yields with a 2D maximum likelihood fit of the data in the 12 event categories. The data are binned in two-dimensions, $m_{\mathrm{b\bar{b}}}$ and m_{HH}, within the ranges $30 \leq m_{\mathrm{b\bar{b}}} \leq 210\,\mathrm{GeV}$ and $700 \leq m_{\mathrm{HH}} \leq 5050\,\mathrm{GeV}$. The $m_{\mathrm{b\bar{b}}}$ bin width is 6 GeV, and the m_{HH} bin width is variable: 25 GeV width at the low end of the mass range, 50 GeV width in the middle of the mass range, and 75 GeV at high mass. These bin widths are smaller than the mass resolutions but large enough to keep the number of bins computationally tractable. Signal and background mass distributions are modeled using two-dimensional templates that are smoother than the distributions directly from simulation. Independent templates are used for each event category. The templates are created using only simulation, and shape and normalization uncertainties that account for possible differences between data and simulation are included while executing the fit. This fitting method was previously presented in Ref. [51].

6.1 Background component classification

To perform the fit to data, we split the background into components and then generate 2D templates in the $m_{\mathrm{b\bar{b}}}$ and m_{HH} mass plane for each component independently. Each component is then allowed to vary in the fit to the total background in each search category.

Rather than split by SM process, we distinguish four components by generator-level information such that they each have distinct $m_{\mathrm{b\bar{b}}}$ distribution shapes. The background is divided by counting in simulation the number of generator-level quarks from the immediate decay of a top quark or vector boson within $\Delta R < 0.8$ of the $\mathrm{b\bar{b}}$ jet axis. The first component is the "m_{t} background", in which all three quarks from a single top decay fulfill this criterion. The second component is the "m_{W} background", identified as the events that do not fulfill the m_{t} background criterion but in which both quarks from either a Z or W boson fall within the jet cone. Both of these backgrounds contain resonant peaks in the $m_{\mathrm{b\bar{b}}}$ shape corresponding to either the top quark or W boson mass. The "lost t/W background" contains events in which at least one quark is contained within the $\mathrm{b\bar{b}}$ jet cone, but not the full set needed to satisfy one of the previous two requirements. Finally, the "q/g background" designates all other events. The first three categories are primarily composed of $\mathrm{t\bar{t}}$ events, while the q/g background is

composed mostly of W+jets and QCD multijet processes in the SL channel and Z/γ^*+jets in the DL channel. This is summarized in Table 5.

Table 5: The four background components with their kinematical properties and defining number of generator-level quarks within $\Delta R < 0.8$ of the $b\bar{b}$ jet axis.

Bkg. category	Dominant SM processes	Resonant in $m_{b\bar{b}}$	Num. of gen-level quarks
m_t	$t\bar{t}$	top mass	3 from top
m_W	$t\bar{t}$	W mass	2 from W
lost t/W	$t\bar{t}$	No	1 or 2
q/g	V+jets and multijet	No	0

6.2 Template construction strategy

For each of the four background components, a unique template in the $m_{b\bar{b}}$ and m_{HH} mass plane is produced for each of the 12 event categories. First, we produce a small set of inclusive templates that have more statistical power than the set of events in the search categories. These inclusive templates are made by combining events in multiple categories and by relaxing selections, provided that the inclusive shape remains consistent with the shape for the full selection. Then, for each of the 12 event categories, the inclusive templates are fit to the simulated mass distributions to produce a unique template. This fit is performed in a similar manner and with a similar parameterization of the template shape as is done for the fit to data. The background templates and associated systematic uncertainties are ultimately validated by fitting to data in dedicated control regions, a procedure described in Sec. 6.5.

In the SL channel, extra effort is undertaken to smooth the QCD multijet simulation before including it in the template-building procedure. This simulated process has large statistical uncertainties because large enough samples cannot be produced with the needed effective luminosity. However, the W+jets simulation has much more statistical power but also has similar $b\bar{b}$ jet reconstruction compared to the QCD simulation. Both processes contribute significantly to the q/g background, with large light-flavored jets that are misidentified as b jets, yielding a very similar falling shapes in the $m_{b\bar{b}}$ spectrum and similar $b\bar{b}$ jet tagging distributions. Instead of using the QCD simulation directly in the q/g background modeling, a combined distribution is created by measuring the ratio of QCD to W+jets event yields as a function of m_{HH} and then using these corrections to scale up the W+jets simulation. Corrections and distributions are obtained for each lepton flavor and $H \to WW^*$ purity category, since the $b\bar{b}$ jet tagging between W+jets and QCD is equivalent. This distribution is then used as input to the q/g background modeling to account for both processes.

6.3 Background modeling

The background templates are modeled using conditional probabilities of $m_{b\bar{b}}$ as a function of m_{HH} so that the templates include the correlation of these two variables, fully described in Ref. [58]. The full two-dimensional template is modelled as:

$$P_{bkg}(m_{b\bar{b}}, m_{HH}) = P_{b\bar{b}}(m_{b\bar{b}}|m_{HH}, \theta_1) P_{HH}(m_{HH}|\theta_2), \tag{5}$$

where $P_{b\bar{b}}$ is a two-dimensional conditional probability distribution, P_{HH} is a one-dimensional probability distribution, and θ_1 and θ_2 are sets of nuisance parameters used to account for background shape uncertainties. The sets θ_1 and θ_2 do not share any common nuisance parameters.

The P_{HH} templates are produced by smoothing one-dimensional m_{HH} histograms with kernel density estimation (KDE) [102–104]. To produce these templates, we use Gaussian kernels with adaptive bandwidths, which are parameters of the KDE that control the smoothing and are dependent on the local event density. We do this to apply less smoothing to regions of the distribution with many events and more smoothing to regions with few events. For $m_{\mathrm{HH}} \gtrsim 2\,\mathrm{TeV}$, where there are very few events even in simulation, m_{HH} tail is further smoothed by fitting with an exponential function.

The two-dimensional templates $P_{\mathrm{b\bar{b}}}$ are obtained with different methods for the resonant and non-resonant background components. For each of the m_{t} and m_{W} resonant backgrounds, we fit the $m_{\mathrm{b\bar{b}}}$ distributions with a double Crystal Ball function [105, 106] centered around m_{t} and m_{W}, respectively. This function has a Gaussian core, which is used to model the bulk of the $m_{\mathrm{b\bar{b}}}$ resonance, and power-law tails, which account for the effects of jet misreconstruction. The fits are performed for events binned in m_{HH} to capture the dependence of the $m_{\mathrm{b\bar{b}}}$ shape on m_{HH}. For the non-resonant lost t/W and q/g backgrounds, the $P_{\mathrm{b\bar{b}}}$ are estimated from two-dimensional histograms using two-dimensional KDE. Independent KDE parameters are used for each dimension and each background when building the $P_{\mathrm{b\bar{b}}}$. As done for the P_{HH} tail modeling, the high mass m_{HH} distribution tail here is exponentially smoothed. The normalizations from simulation are used as the initial values for the background normalizations in the fit to data.

6.4 Signal modeling

The signal templates are also modeled following Ref. [58] using conditional probabilities:

$$P_{\mathrm{signal}}(m_{\mathrm{b\bar{b}}}, m_{\mathrm{HH}}|m_{\mathrm{X}}) = P_{\mathrm{HH}}(m_{\mathrm{HH}}|m_{\mathrm{b\bar{b}}}, m_{\mathrm{X}}, \theta_1')P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{X}}, \theta_2'). \tag{6}$$

The sets θ_1' and θ_2' do not share any common nuisance parameters. However, θ_2' and θ_1 from Eq. 5 do share two nuisance parameters corresponding to the scale and resolution uncertainties of soft-drop jet resonances in the $m_{\mathrm{b\bar{b}}}$ dimension. This is discussed in more detail in Sec. 7.1.2.

The P_{signal} distributions are first obtained for discrete m_{X} values by fitting histograms of the signal mass distributions. Models continuous in m_{X} are then produced by interpolating the fit parameters. The one-dimensional $P_{\mathrm{b\bar{b}}}$ templates are created by fitting the $m_{\mathrm{b\bar{b}}}$ spectra with a double Crystal Ball function, and the mass resolution is slightly larger than 10%, largest at low mass. The modeling of events in the bL category also contains an exponential component to model the small fraction of signal events with no resonant peak in the distribution.

The two-dimensional P_{HH} templates are built to account for correlations between m_{HH} and $m_{\mathrm{b\bar{b}}}$. These m_{HH} distributions are also modeled with a double Crystal Ball function, but with an additional linear dependence on $m_{\mathrm{b\bar{b}}}$, parameterized by $\Delta_{\mathrm{b\bar{b}}} = (m_{\mathrm{b\bar{b}}} - \mu_{\mathrm{b\bar{b}}})/\sigma_{\mathrm{b\bar{b}}}$. Here, $\mu_{\mathrm{b\bar{b}}}$ and $\sigma_{\mathrm{b\bar{b}}}$ are the mean and width parameters, respectively, in the fit to the $m_{\mathrm{b\bar{b}}}$ spectra. To accomplish this, the mean parameter μ_{HH} along m_{HH} in the Crystal Ball function fit is then

$$\mu_{\mathrm{HH}} = \mu_0(1 + \mu_1 * \Delta_{\mathrm{b\bar{b}}}), \tag{7}$$

where μ_0 and μ_1 are fit parameters. Doing this, we can account for mismeasurements of the $\mathrm{b\bar{b}}$ jet that result in mismeasurements of m_{HH}. The resolution of the m_{HH} resonance, denoted σ_{HH}, is also dependent on $m_{\mathrm{b\bar{b}}}$ such that

$$\sigma_{\mathrm{HH}} = \begin{cases} \sigma_0(1 + \sigma_1 * \Delta_{\mathrm{b\bar{b}}}) & \Delta_{\mathrm{b\bar{b}}} < 0 \\ \sigma_0 & \Delta_{\mathrm{b\bar{b}}} \geq 0 \end{cases} \quad (8)$$

where σ_0 and σ_1 are fit parameters. An undermeasurement of $m_{\mathrm{b\bar{b}}}$ can result from the SD algorithm removing too much energy from the Higgs boson decay. In this case, the correlation worsens, and the m_{HH} resolution grows wider. For $|\Delta_{\mathrm{b\bar{b}}}| > 2.5$, we use the value at the boundary since the correlation does not hold for severe mismeasurements. The m_{HH} resolution is $\approx$5%.

The product of the acceptance and efficiency for X $\rightarrow$ HH events to fall into any of the individual search categories is taken from simulation. As done for the signal shape parameters, the efficiency is interpolated along m_{X}. Uncertainties in the relative acceptances and in the integrated luminosity of the sample are included in the 2D maximum likelihood fit that is used to obtain confidence intervals on the X $\rightarrow$ HH process. The signal modeling is tested using toy experiments in which we fit the templates to pseudo-data that contains a fixed amount of signal; no significant bias in the fitted signal yield is found.

6.5 Validation of background models with control region data

The background models are validated in the top CR and non-top CR data samples. For both control regions, background templates are constructed in the same way as for the search region, except using the control region selection. The background templates are then fit to the control region data with the same nuisance parameters that are used in the standard 2D maximum likelihood fit. In the non-top control region, the m_{t} background is negligible and not included in the modeling. The result of the simultaneous fit is shown in Fig. 3 for both control regions. To improve visualization, the displayed binning in this and subsequent histograms is coarser than that used in the maximum likelihood fit. The projections in both mass dimensions are shown for the combination of all event categories. The fit results model the data well, indicating that the shape uncertainties can account sufficiently for potential differences between data and simulation.

7 Systematic uncertainties

Systematic uncertainties that affect the normalization and shape of the signal and background are modelled with nuisance parameters in the 2D maximum likelihood fit to data. Nuisance parameters for shape uncertainties have Gaussian constraints, while normalization uncertainties have log-normal constraints. In certain cases a single nuisance parameter may affect both the normalization and the shape of a resonance, in which case the nuisance parameter constraint is Gaussian. Detailed methods of parameterizing the background and signal uncertainties are described in Sec. 7.1 and 7.2, respectively.

To implement non-resonant mass shape uncertainties, templates are first generated with modified event weights, which include multiplicative parameters proportional to $m_{\mathrm{b\bar{b}}}$, m_{HH}, $1/m_{\mathrm{b\bar{b}}}$, and $1/m_{\mathrm{HH}}$. For each parameter, this produces two alternative templates that represent an upward and downward shift from the nominal model. The 2D fit then interpolates between these two alternative templates to constrain the magnitudes of these parameters. Resonant mass shape uncertainties are implemented as uncertainties on the mean and width parameters of a double Crystal Ball function. In most cases, different nuisance parameters are used for the background shape uncertainties from those used for the signal uncertainties.

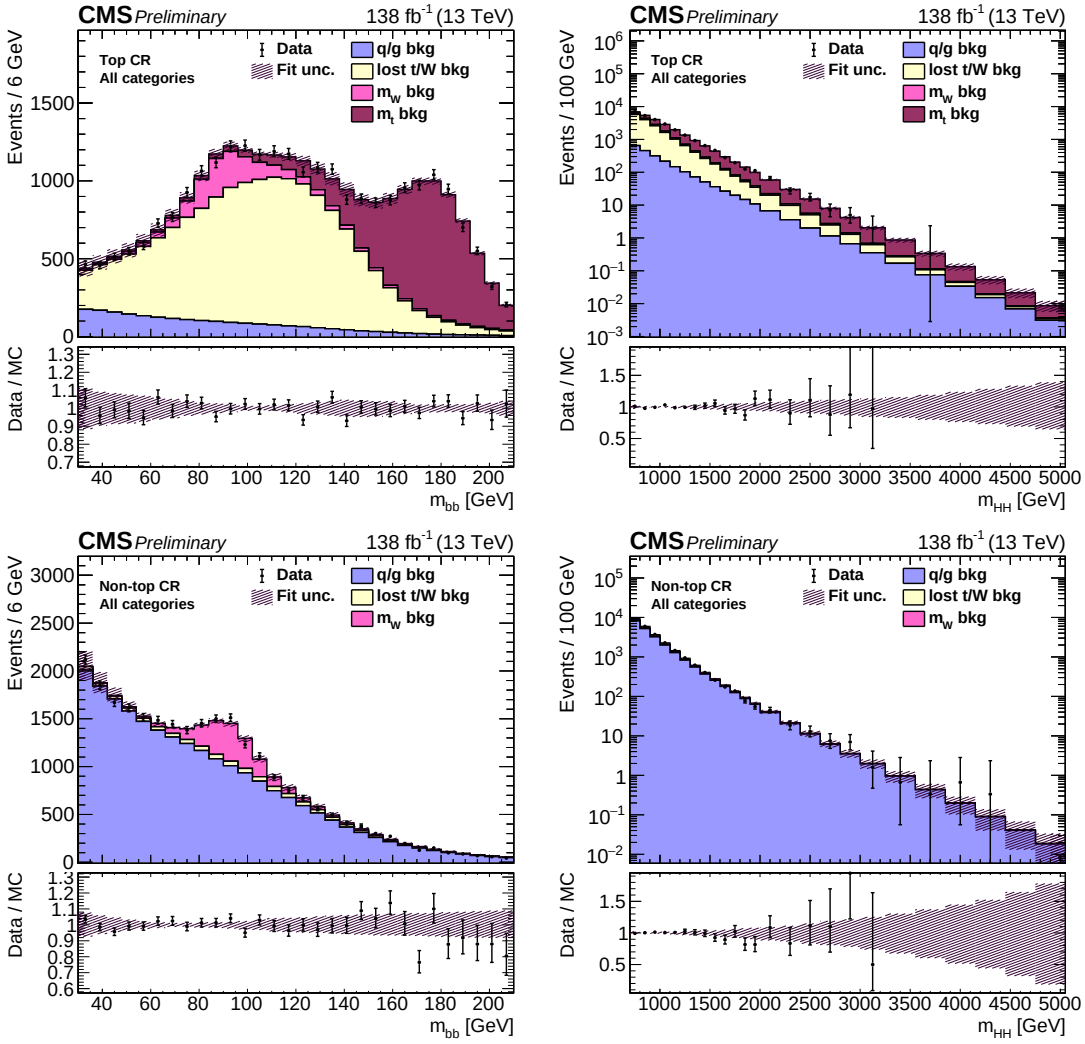

Figure 3: The post-fit model compared to data in the top CR (upper plots) and non-top CR (lower plots), projected into $m_{\mathrm{b\bar{b}}}$ (left) and m_{HH} (right). Events from all categories are combined. The fit result is the filled histogram, with the different colors indicating different background categories. The background shape uncertainty is shown as the hatched band. The bottom panes of each plot show ratio of the data to the fit result.

All background and signal uncertainties are listed in Tab. 6 and 7, respectively, with their initial sizes. A single uncertainty type can be applied to multiple event categories with independent nuisance parameters per category. The background model contains 104 total nuisance parameters, while the signal model contains 27, with two parameters shared between signal and background. This and the descriptions of each uncertainty and their correlations are also described in the rest of this section.

7.1 Background uncertainties

Background uncertainty parameters are chosen considering possible discrepancies between data and simulation, such as in the relative background composition or in the jet energy scale. Studies of the two control regions are used to verify that the chosen uncertainties cover such differences. The fit result does not depend strongly on the sizes of the pre-fit uncertainties because they serve as loose constraints on the fit. We verify this by inflating all pre-fit background uncertainties by a factor of two and observing that the final result does not change. Therefore,

Table 6: Background systematic uncertainties included in the maximum likelihood fit. The N_p column indicates the number of nuisance parameters used to model the uncertainty. In the last two columns, σ_I refers to the initial estimate of the uncertainty, and σ_C refers to the constrained uncertainty obtained post-fit. For the q/g, $t\bar{t}$, and lost t/W shape uncertainties, "scale" uncertainties are those implemented with alternative templates with multiplicative parameters proportional to mass m, and "inverse scale" are for proportional to $1/m$.

Uncertainty type	Processes	N_p	σ_I	σ_C/σ_I
SD jet $m_{b\bar{b}}$ scale	m_W, m_t, signal	2	0.54%, 2.0% (m_t)	98%, 19% (m_t)
SD jet $m_{b\bar{b}}$ resolution	m_W, m_t, signal	2	8.6%, 17.2% (m_t)	95%, 25% (m_t)
q/g normalization	q/g	12	50% (1ℓ), 100% (2ℓ)	37-78%
q/g m_{HH} scale	q/g	10	$\pm 0.5 * m_{HH}$ / TeV	78-99%
q/g m_{HH} inverse scale	q/g	10	$\pm 1.4 *$ TeV$/m_{HH}$	64-99%
q/g $m_{b\bar{b}}$ scale	q/g	4	$\pm 0.00375 * m_{b\bar{b}}$ / GeV	81-99%
q/g $m_{b\bar{b}}$ inverse scale	q/g	4	$\pm 15 *$ GeV$/m_{b\bar{b}}$	77-99%
Lost t/W $m_{b\bar{b}}$ scale	lost t/W	4	$\pm 0.003 * m_{b\bar{b}}$ / GeV	71-99%
Lost t/W $m_{b\bar{b}}$ inverse scale	lost t/W	4	$\pm 18 *$ GeV$/m_{b\bar{b}}$	88-99%
$t\bar{t}$ normalization	lost t/W, m_W, m_t	12	35% (1ℓ), 70% (2ℓ)	19-68%
$t\bar{t}$ relative norm.	lost t/W, m_W, m_t	8	35% (1ℓ), 70% (2ℓ)	9-96%
$t\bar{t}$ m_{HH} scale	lost t/W, m_W, m_t	12	$\pm 0.25 * m_{HH}$ / TeV	84-99%
$t\bar{t}$ m_{HH} relative scale	lost t/W, m_W, m_t	8	$\pm 0.25 * m_{HH}$ / TeV	74-99%
$t\bar{t}$ m_{HH} inverse scale	lost t/W, m_W, m_t	12	$\pm 0.7 *$ TeV$/m_{HH}$	61-99%

the pre-fit uncertainties are sufficiently large to account for discrepancies between data and simulation in the control regions. More complex background models, such as those with more nuisance parameters or higher-order shape distortions, were previously studied in Ref. [58] but were not found to be necessary.

In the following subsections, we detail the parameterization of the different uncertainties for the background.

7.1.1 Background normalization uncertainties

The m_W, m_t, and lost t/W backgrounds all primarily arise from $t\bar{t}$ production; consequently, some uncertainties are applied by treating these three backgrounds together, referred to collectively as the $t\bar{t}$ background in Tab. 6. We account for differences between data and simulation in the $t\bar{t}$ normalization by including independent nuisance parameters for each category that allow the normalizations of these backgrounds to vary in a correlated manner ("$t\bar{t}$ normalization"). However, the three $t\bar{t}$-dominated background components exhibit differences in the b tagging efficiency and the $b\bar{b}$ jet p_T spectrum, so we include additional nuisances ("$t\bar{t}$ relative normalization") that allow the relative normalizations of each of these to vary within the absolute normalization, which itself also varies. Separate nuisance parameters are used to control the q/g background normalization, as this is the only background component to arise primarily from non-$t\bar{t}$ processes.

7.1.2 Background shape uncertainties

The shape uncertainties for the backgrounds are modelled differently depending on whether or not the shape is resonant in the $m_{b\bar{b}}$ dimension. All backgrounds are non-resonant in the m_{HH}

Table 7: Signal systematic uncertainties included in the maximum likelihood fit. The N_p column indicates the number of nuisance parameters used to model the uncertainty. In the "Uncertainty values" column, some uncertainties are noted to affect both the yield (Y) and m_{HH} shape (S for scale, R for resolution) of the signal. All other uncertainties, except the SD jet mass uncertainties, are uncertainties on the signal yield.

Uncertainty type	N_p	Uncertainty values
SD jet $m_{b\bar{b}}$ scale	1	0.54%
SD jet $m_{b\bar{b}}$ resolution	1	8.6%
Luminosity	1	1.6%
PDF+scale	1	spin-0: 2.0%, spin-2: 2.5%
Trigger	6	1ℓ: 2.0%, 2ℓ: 3.0%
Pileup	1	1ℓ: 1.0%, 2ℓ: 0.6%
Electron reconstruction	1	1ℓ: 0.5%, 2ℓ: $< 0.8\%$
Electron identification	2	1ℓ: 4.2%, 2ℓ: $< 2.6\%$
Muon identification	2	1ℓ: 2.3%, 2ℓ: $< 2.3\%$
Electron isolation	1	1ℓ: 6%, 2ℓ: 3% for each electron
Muon isolation	1	1ℓ: 6%, 2ℓ: 2% for each muon
Jet energy scale	1	Y: 2%, $S(m_{HH})$: 0.8%, $R(m_{HH})$: 3%
Jet energy resolution	1	Y: 0.5%. $S(m_{HH})$: 0.3%, $R(m_{HH})$: 4%
Unclustered energy	1	Y: 0.5%, $S(m_{HH})$: 0.1%, $R(m_{HH})$: 1.5%
Other detector effects	2	Y: 0.6%, $R(m_{HH})$: 1.0%
AK4 b tag efficiency	1	$< 4.0\%$
AK4 b tag fake rate	1	$< 2.5\%$
$b\bar{b}$ jet b tagging	1	bL: 8.5%, bT: 11.5%
$q\bar{q}$ jet τ_2/τ_1 efficiency	1	LP: 26% HP: 6.7%

dimension, and mismodelling of the background p_T spectrum can manifest as an incorrect m_{HH} scale. To account for this, the m_{HH} shape uncertainties are implemented with alternative background templates built with parameters proportional to m_{HH} ("scale") and $1/m_{HH}$ ("inverse scale"), as described in the beginning of Sec. 7. For the q/g background, a pair of these nuisance parameters is included for each category in the SL channel and each b tagging category in the DL channel. For the top backgrounds, we include a pair of these nuisances for each search category. Furthermore, to allow the top backgrounds to be anti-correlated, we include nuisances for the relative m_{HH} scale (alternative templates built with factors $\sim m_{HH}$) for each b tagging category, separately for the SL channel and the DL channel.

The q/g and lost t/W backgrounds are non-resonant in $m_{b\bar{b}}$, and so alternative templates are also used to encode the $m_{b\bar{b}}$ systematics with factors proportional to $m_{b\bar{b}}$ or $1/m_{b\bar{b}}$. The uncertainties account for mismodelling in the simulated jet energy scale and resolution. For both of these non-resonant backgrounds, the $m_{b\bar{b}}$ shape does not depend on the b tagging, and so there is a pair of nuisances for each background and each b tagging category, separately for the SL channel and the DL channel.

For the m_W and m_t backgrounds in the $m_{b\bar{b}}$ dimension, where resonances are constructed using large soft-drop jets, the jet mass uncertainties are dependent on the jet substructure. Because of this, the jet mass uncertainties for the signal and the m_W background, respectively from the two-prong decays $H \to b\bar{b}$ and $W \to q\bar{q}$, are correlated. This is the only such instance where signal and background are correlated, sharing nuisance parameters. Uncertainties that

have been measured in data for W boson decays into merged jets in $t\bar{t}$ events are found to cover discrepancies between our simulation and data for the m_W background but not for the m_t background. We do not expect these uncertainties to cover discrepancies in the m_t background because the SD algorithm behaves differently for the three-prong top jets ($t \to bq\bar{q}$) in this background. Thus, these uncertainties are larger than for two-prong jets and are not correlated with the m_W jet mass shape uncertainties, as seen in Tab. 6.

7.2 Signal uncertainties

Noted in Tab. 7, uncertainties are applied to the normalization of the signal to account for mismeasurements in the total integrated luminosity, the pileup profile, the trigger efficiency, the lepton selection efficiencies, and other detector effects. Signal acceptance uncertainties from the choices of parton distribution function (PDF), factorization scale, and renormalization scale are also applied. The scale uncertainties are obtained following Refs. [107, 108], and the PDF uncertainty is evaluated using the NNPDF 3.1 PDF set [79].

The signal acceptance and the m_{HH} resonance scale and resolution all have uncertainties due to the jet energy scale and resolution, the unclustered energy resolution, and other detector effects. The same $m_{b\bar{b}}$ resonance scale and resolution uncertainties that are applied for the m_W background are applied to the signal because they are both soft-drop jets with two-prong substructure.

The $q\bar{q}$ jet τ_2/τ_1 selection efficiency is measured in a $t\bar{t}$ data sample enriched with hadronically decaying W bosons. The uncertainties in this measurement are included as normalization uncertainties on the $H \to WW^*$ decay purity categories, and the LP and HP uncertainties are anticorrelated. Normalization uncertainties are also applied to account for the efficiency and fake rate of AK4 jet b tagging used to identify and reject jets from $t\bar{t}$. The uncertainty on the $b\bar{b}$ jet tagging efficiency is included as a single nuisance parameter that varies the signal normalization and is dependent on both the b tagging category and m_X. These $b\bar{b}$ jet tagging uncertainties are the dominant uncertainties on the signal normalization, followed by the uncertainties on the τ_2/τ_1 efficiencies.

8 Results

The data are interpreted by performing a maximum likelihood fit in the two-dimensional ($m_{b\bar{b}}$, m_{HH}) mass plane for one model containing only background processes and one containing both background and signal processes. We find that the background-only model fits the data well. We interpret the results as upper limits at 95% confidence level (CL) on $\sigma\mathcal{B}(X \to HH)$.

The quality of the fit is quantified with a likelihood ratio goodness-of-fit test using the saturated model [109]. The probability distribution function of the test statistic is obtained with toy experiments, and the observed value is within the central 68% interval of expected results. The best-fit values of the nuisance parameters are consistent with the initial $\pm 1\sigma$ range of uncertainty.

The fit result and the data are projected in $m_{b\bar{b}}$ for each event category in Fig. 4. The $m_{b\bar{b}}$ shape is modeled well, with each background component contributing over some subspace of the mass range. Particularly, the resonant peaks corresponding to the W boson and top quark are correctly modeled by the fit. Similarly, the m_{HH} projections of the fit are shown in Fig. 5. There is good agreement for the full m_{HH} mass range in these figures as well.

Upper limits are shown at 95% CL in Fig. 7 for $0.8 < m_X < 4.5\,\text{TeV}$ and both the spin-0 and

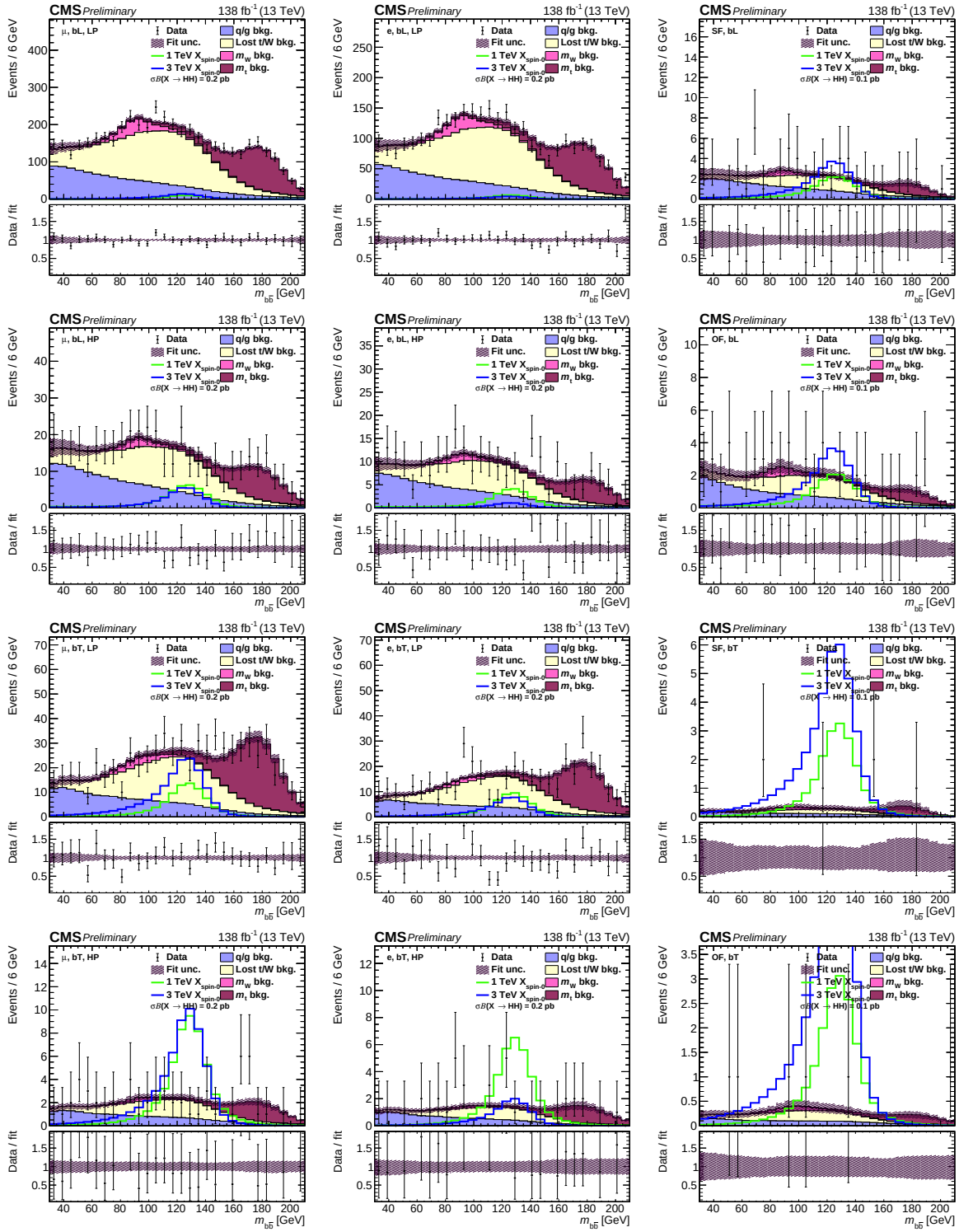

Figure 4: The fit result compared to data projected into $m_{\mathrm{b\bar{b}}}$ for both the single- and dilepton channels. The label for each search category is in the upper left of each plot. The fit result is the filled histogram, with the different colors indicating different background categories. The background shape uncertainty from the fit is shown as the hatched band. Example spin-0 signal distributions for $m_{\mathrm{X}} = 1.0$ and 3.0 TeV are shown as solid lines, with $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH})$ set to 0.2 and 0.1 pb for the SL and DL channels, respectively. The bottom panes show the ratio of the data to the fit result.

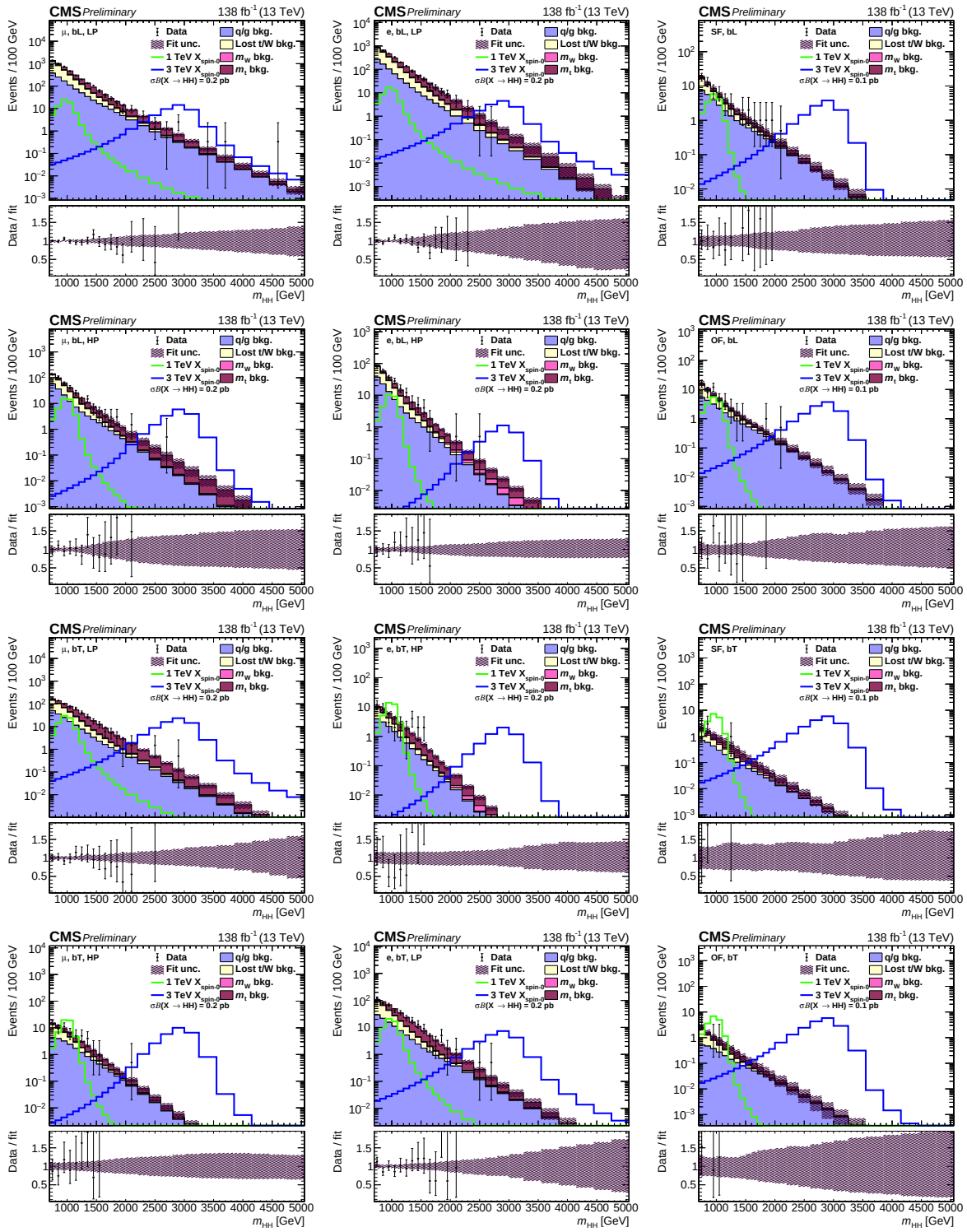

Figure 5: The fit result compared to data projected into m_{HH} for both the single- and dilepton channels. The label for each search category is in the upper left of each plot. The fit result is the filled histogram, with the different colors indicating different background categories. The background shape uncertainty from the fit is shown as the hatched band. Example spin-0 signal distributions for $m_{\mathrm{X}} = 1.0$ and 3.0 TeV are shown as solid lines, with $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH})$ set to 0.2 and 0.1 pb for the SL and DL channels, respectively. The bottom panes of each plot show the ratio of the data to the fit result.

spin-2 boson scenarios. The limits are evaluated using the asymptotic approximation [110] of the CL_s method [111, 112]. The observed exclusion limit is consistent with the expected limit. A spin-0 signal at $m_X = 0.8\,\text{TeV}$ is excluded for $\sigma\mathcal{B} > 24.5\,\text{fb}$, and the exclusion limit strengthens over the full mass range to $\sigma\mathcal{B} > 0.78\,\text{fb}$ at $m_X = 4.5\,\text{TeV}$. Spin-2 signals have larger acceptance, and so the exclusion limit on these signals is stronger: at $m_X = 0.8\,\text{TeV}$, we exclude $\sigma\mathcal{B} > 16.7\,\text{fb}$, and at $m_X = 4.5\,\text{TeV}$ we exclude $\sigma\mathcal{B} > 0.67\,\text{fb}$.

Table 8 shows the event yields for each search category that are observed in data and expected from a background-only fit, along with the associated uncertainty on the total background yield in each category. Figure 6 shows the expected exclusion limit at 95% CL for each search category alone. In general, the tight (bT) $b\bar{b}$ jet tagging categories are the most important over the full range of m_X, since these contain the most signal and the least amount of background. The DL categories are generally more sensitive than most SL categories as well since the background yields are much lower in the DL channel. At high m_X, the electron categories in the SL channel are the least sensitive because the electron reconstruction efficiency is severely degraded.

Table 8: Event yields broken down by search category. For each category, shown are the event yields observed in data, expected after a fit to the B-only model, and the corresponding relative uncertainty.

Search category	Observed	Expected (post-fit)	Post-fit uncertainty
μ bL LP	4542	4540.9	1.5%
μ bL HP	417	416.1	4.8%
μ bT LP	657	658.5	4.2%
μ bT HP	56	57.3	10.0%
e bL LP	2945	2945.4	1.9%
e bL HP	248	247.7	5.7%
e bT LP	423	423.9	4.2%
e bT HP	37	37.7	14.6%
SF bL	59	59.6	14.2%
OF bL	50	50.8	13.5%
SF bT	6	7.9	31.6%
OF bT	6	8.1	25.8%

Relative to the $X \to b\bar{b}\ell\nu q\bar{q}$ search in Ref. [58], this analysis ranges from 6 times more sensitive at low m_X to 14 times more sensitive at high m_X. The improvements in sensitivity arise primarily from three developments. First, an improvement in the expected upper limits by a factor of ≈3.5 is achieved due to the larger integrated luminosity alone. This level of improvement is expected because the number of background events is much smaller than the number of signal events under a typical signal peak, even at low m_X. Second, due to improved techniques in the SL channel alone, we achieve similar sensitivity at $m_X = 0.8\,\text{TeV}$ and up to a ≈2 times improvement at $m_X = 4.5\,\text{TeV}$. Finally, the addition of the DL channel provides significant improvement in sensitivity. At low m_X, the DL channel is ≈70% more sensitive than the SL channel, largely due to the background level that is over an order of magnitude smaller. At high m_X, where there is no background in any channel, the DL channel is similarly as sensitive as the SL channel. This occurs because the dilepton signal efficiency is largest at high mass, and in the SL channel, despite the larger branching fraction, the lepton efficiency (particularly for electrons) degrades at high mass due to the nearby boosted jet.

Predicted radion and bulk graviton cross sections [113] are also shown in Fig. 7 in the context of Randall-Sundrum models that allow the SM fields to propagate though the extra dimension.

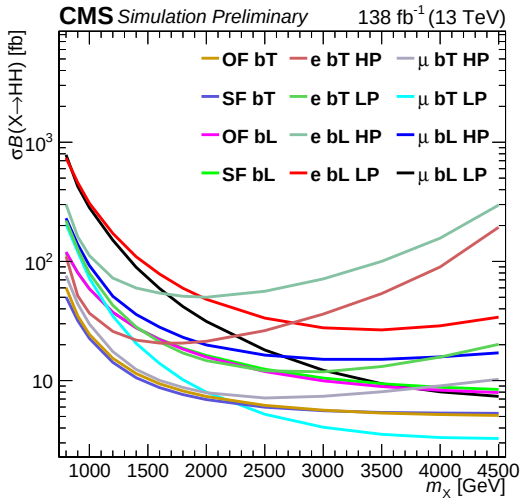

Figure 6: Expected upper limits at 95% confidence level (CL) for each of the 12 search categories individually.

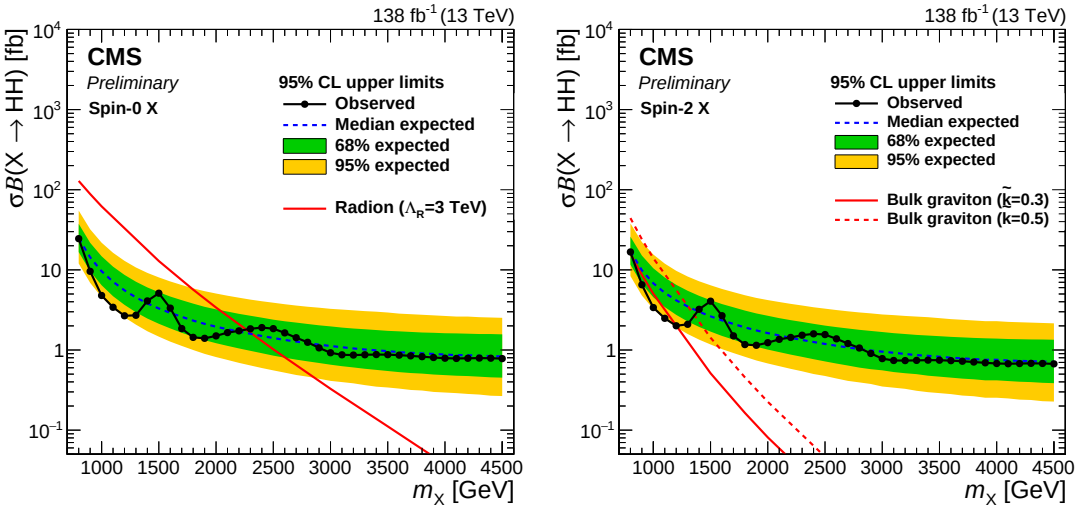

Figure 7: Observed and expected 95% CL upper limits on the product of the cross section and branching fraction to HH for a generic spin-0 (left) and spin-2 (right) boson X, as a function of mass. Example radion and bulk graviton predictions are also shown. The HH branching fraction is assumed to be 25 and 10%, respectively.

Typical model parameters are chosen as proposed in Ref. [114]. For radions, a branching fraction of 25% to HH and an ultraviolet cutoff $\Lambda_R = 3$ TeV are assumed. A 10% branching fraction is assumed for bulk gravitons, which occurs in scenarios that include significant coupling between the bulk graviton and top quarks. Bulk graviton production cross sections depend on the dimensionless quantity $\tilde{k} = \sqrt{8\pi}k/M_{\mathrm{Pl}}$, where k is the curvature of the extra dimension and M_{Pl} is the Planck mass. For this interpretation, we choose $\tilde{k} = 0.3$ and 0.5. For these particular signal parameters the radion and bulk graviton decay widths are larger than the 1 MeV width chosen for signal sample generation but much smaller than the detector resolution.

9 Summary

A search has been presented for new bosons decaying to a pair of Higgs bosons (H) where one decays into a bottom quark pair ($b\bar{b}$) and the other decays via one of three different modes into final states with leptons. The large Lorentz boost of the Higgs bosons produces a distinct experimental signature with one large-radius jet with substructure consistent with the decay $H \to b\bar{b}$. For the Higgs boson that does not decay to $b\bar{b}$, considered are the single-lepton decay $H \to WW^* \to \ell\nu q\bar{q}$ and the dilepton decays $H \to WW^* \to \ell\nu\ell\nu$ and $H \to \tau\tau \to \ell\nu\nu\ell\nu\nu$. In the single-lepton channel, the experimental signature also contains a second large-radius jet with a nearby lepton, which is consistent with the decay of $H \to WW^*$. In the dilepton channel, the experimental signature contains two leptons and significant missing transverse momentum. This search uses a sample of proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ collected by the CMS detector at the LHC. The primary Standard Model backgrounds — production of top quark pairs and Z/γ^*+jets in the dilepton channel only — are suppressed by reconstructing the HH decay chain and applying selections to discriminate signal from background. The signal and background yields are estimated by a two-dimensional template fit in the plane of the $b\bar{b}$ jet mass and the HH resonance mass. The templates are validated in a variety of data control regions and are shown to model the data well. The data are consistent with the expected Standard Model background. Upper limits are set on the product of the cross section and branching fraction for new bosons decaying to HH. The observed limit at 95% confidence level for a spin-0 boson ranges from 24.5 fb at 0.8 TeV to 0.78 fb at 4.5 TeV, while the limit for a spin-2 boson is 16.7 fb at 0.8 TeV and 0.67 fb at 4.5 TeV. This search produces the most stringent exclusion limits to date for $X \to HH$ production modes with leptons in the final state. The current sensitivity to $X \to HH$ production is stronger than or comparable to those from searches in other channels for HH resonances with masses above 800 GeV.

References

[1] ATLAS Collaboration, "Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 01, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[4] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", *Phys. Rev. Lett.* **13** (1964) 321, doi:10.1103/PhysRevLett.13.321.

[5] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 508, doi:10.1103/PhysRevLett.13.508.

[6] G. C. Branco et al., "Theory and phenomenology of two-Higgs-doublet models", *Phys. Rept.* **516** (2012) 1–102, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.

[7] P. Ramond, "Dual theory for free fermions", *Phys. Rev. D* **3** (1971) 2415, doi:10.1103/PhysRevD.3.2415.

[8] Y. A. Golfand and E. P. Likhtman, "Extension of the algebra of Poincaré group generators and violation of P invariance", *JETP Lett.* **13** (1971) 323.

[9] A. Neveu and J. H. Schwarz, "Factorizable dual model of pions", *Nucl. Phys. B* **31** (1971) 86, doi:10.1016/0550-3213(71)90448-2.

[10] D. V. Volkov and V. P. Akulov, "Possible universal neutrino interaction", *JETP Lett.* **16** (1972) 438.

[11] J. Wess and B. Zumino, "A Lagrangian model invariant under supergauge transformations", *Phys. Lett. B* **49** (1974) 52, doi:10.1016/0370-2693(74)90578-4.

[12] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nucl. Phys. B* **70** (1974) 39, doi:10.1016/0550-3213(74)90355-1.

[13] P. Fayet, "Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino", *Nucl. Phys. B* **90** (1975) 104, doi:10.1016/0550-3213(75)90636-7.

[14] H. P. Nilles, "Supersymmetry, supergravity and particle physics", *Phys. Rep.* **110** (1984) 1, doi:10.1016/0370-1573(84)90008-5.

[15] L. Randall and R. Sundrum, "A large mass hierarchy from a small extra dimension", *Phys. Rev. Lett.* **83** (1999) 3370, doi:10.1103/PhysRevLett.83.3370, arXiv:hep-ph/9905221.

[16] W. D. Goldberger and M. B. Wise, "Modulus stabilization with bulk fields", *Phys. Rev. Lett.* **83** (1999) 4922, doi:10.1103/PhysRevLett.83.4922, arXiv:hep-ph/9907447.

[17] O. DeWolfe, D. Z. Freedman, S. S. Gubser, and A. Karch, "Modeling the fifth dimension with scalars and gravity", *Phys. Rev. D* **62** (2000) 046008, doi:10.1103/PhysRevD.62.046008, arXiv:hep-th/9909134.

[18] C. Csaki, M. Graesser, L. Randall, and J. Terning, "Cosmology of brane models with radion stabilization", *Phys. Rev. D* **62** (2000) 045015, doi:10.1103/PhysRevD.62.045015, arXiv:hep-ph/9911406.

[19] C. Csaki, M. L. Graesser, and G. D. Kribs, "Radion dynamics and electroweak physics", *Phys. Rev. D* **63** (2001) 065002, doi:10.1103/PhysRevD.63.065002, arXiv:hep-th/0008151.

[20] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, "Phenomenology of the Randall-Sundrum gauge hierarchy model", *Phys. Rev. Lett.* **84** (2000) 2080, doi:10.1103/PhysRevLett.84.2080, arXiv:hep-ph/9909255.

[21] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, "Warped gravitons at the LHC and beyond", *Phys. Rev. D* **76** (2007) 036006, doi:10.1103/PhysRevD.76.036006, arXiv:hep-ph/0701186.

[22] A. L. Fitzpatrick, J. Kaplan, L. Randall, and L.-T. Wang, "Searching for the Kaluza-Klein graviton in bulk RS models", *JHEP* **09** (2007) 013, doi:10.1088/1126-6708/2007/09/013, arXiv:hep-ph/0701150.

[23] ATLAS Collaboration, "Searches for Higgs boson pair production in the $hh \to bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector", *Phys. Rev. D* **92** (2015) 092004, doi:10.1103/PhysRevD.92.092004, arXiv:1509.04670.

[24] ATLAS Collaboration, "Search for a new resonance decaying to a W or Z boson and a Higgs boson in the $\ell\ell/\ell\nu/\nu\nu + b\bar{b}$ final states with the ATLAS detector", *Eur. Phys. J. C* **75** (2015) 263, doi:10.1140/epjc/s10052-015-3474-x, arXiv:1503.08089.

[25] ATLAS Collaboration, "Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Eur. Phys. J. C* **75** (2015) 412, doi:10.1140/epjc/s10052-015-3628-x, arXiv:1506.00285.

[26] ATLAS Collaboration, "Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state using pp collision data at $\sqrt{s} = 8$ TeV from the ATLAS detector", *Phys. Rev. Lett.* **114** (2015) 081802, doi:10.1103/PhysRevLett.114.081802, arXiv:1406.5053.

[27] ATLAS Collaboration, "Search for WZ resonances in the fully leptonic channel using pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Lett. B* **737** (2014) 223, doi:10.1016/j.physletb.2014.08.039, arXiv:1406.4456.

[28] ATLAS Collaboration, "Search for heavy resonances decaying to a W or Z boson and a Higgs boson in the $q\bar{q}^{(\prime)}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett. B* **774** (2017) 494, doi:10.1016/j.physletb.2017.09.066, arXiv:1707.06958.

[29] ATLAS Collaboration, "Search for WW/WZ resonance production in $\ell\nu qq$ final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **03** (2018) 042, doi:10.1007/JHEP03(2018)042, arXiv:1710.07235.

[30] ATLAS Collaboration, "Search for resonant WZ production in the fully leptonic final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett. B* **787** (2018) 68, doi:10.1016/j.physletb.2018.10.021, arXiv:1806.01532.

[31] ATLAS Collaboration, "Search for heavy resonances decaying into WW in the $e\nu\mu\nu$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Eur. Phys. J. C* **78** (2018) 24, doi:10.1140/epjc/s10052-017-5491-4, arXiv:1710.01123.

[32] ATLAS Collaboration, "Searches for heavy ZZ and ZW resonances in the $\ell\ell qq$ and $\nu\nu qq$ final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **03** (2018) 009, doi:10.1007/JHEP03(2018)009, arXiv:1708.09638.

[33] ATLAS Collaboration, "Search for heavy ZZ resonances in the $\ell^+\ell^-\ell^+\ell^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Eur. Phys. J. C* **78** (2018) 293, doi:10.1140/epjc/s10052-018-5686-3, arXiv:1712.06386.

[34] ATLAS Collaboration, "Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector", *JHEP* **03** (2018) 174, doi:10.1007/JHEP03(2018)174, arXiv:1712.06518.

[35] ATLAS Collaboration, "Search for diboson resonances with boson-tagged jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett. B* **777** (2018) 91–113, doi:10.1016/j.physletb.2017.12.011, arXiv:1708.04445.

[36] ATLAS Collaboration, "Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector", arXiv:1811.04671.

[37] ATLAS Collaboration, "Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **01** (2019) 030, doi:10.1007/JHEP01(2019)030, arXiv:1804.06174.

[38] ATLAS Collaboration, "Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. Lett.* **121** (2018) 191801, doi:10.1103/PhysRevLett.121.191801, arXiv:1808.00336. [Erratum: doi:10.1103/PhysRevLett.122.089901].

[39] CMS Collaboration, "Combination of searches for heavy resonances decaying to WW, WZ, ZZ, WH, and ZH boson pairs in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV", *Phys. Lett. B* **774** (2017) 533, doi:10.1016/j.physletb.2017.09.083, arXiv:1705.09171.

[40] CMS Collaboration, "Search for a heavy resonance decaying into a Z boson and a vector boson in the $\nu\bar{\nu}q\bar{q}$ final state", *JHEP* **07** (2018) 075, doi:10.1007/JHEP07(2018)075, arXiv:1803.03838.

[41] CMS Collaboration, "Search for heavy resonances decaying into two Higgs bosons or into a Higgs boson and a W or Z boson in proton-proton collisions at 13 TeV", arXiv:1808.01365.

[42] CMS Collaboration, "Search for a heavy resonance decaying into a Z boson and a Z or W boson in $2\ell 2q$ final states at $\sqrt{s} = 13$ TeV", *JHEP* **09** (2018) 101, doi:10.1007/JHEP09(2018)101, arXiv:1803.10093.

[43] CMS Collaboration, "Search for production of Higgs boson pairs in the four b quark final state using large-area jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:1808.01473.

[44] CMS Collaboration, "Searches for a heavy scalar boson H decaying to a pair of 125 GeV Higgs bosons hh or for a heavy pseudoscalar boson A decaying to Zh, in the final states with $h \to \tau\tau$", *Phys. Lett. B* **755** (2016) 217, doi:10.1016/j.physletb.2016.01.056, arXiv:1510.01181.

[45] CMS Collaboration, "Search for massive resonances decaying into WW, WZ, ZZ, qW, and qZ with dijet final states at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **97** (2018) 072006, doi:10.1103/PhysRevD.97.072006, arXiv:1708.05379.

[46] CMS Collaboration, "Search for heavy resonances that decay into a vector boson and a Higgs boson in hadronic final states at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **77** (2017) 636, doi:10.1140/epjc/s10052-017-5192-z, arXiv:1707.01303.

[47] CMS Collaboration, "Search for heavy resonances decaying into a vector boson and a Higgs boson in final states with charged leptons, neutrinos and b quarks at $\sqrt{s} = 13$ TeV", arXiv:1807.02826.

[48] CMS Collaboration, "Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 8 TeV", *Phys. Lett. B* **749** (2015) 560, doi:10.1016/j.physletb.2015.08.047, arXiv:1503.04114.

[49] CMS Collaboration, "Search for heavy resonances decaying into a vector boson and a Higgs boson in final states with charged leptons, neutrinos, and b quarks", *Phys. Lett. B* **768** (2017) 137, doi:10.1016/j.physletb.2017.02.040, arXiv:1610.08066.

[50] CMS Collaboration, "Search for a massive resonance decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **781** (2018) 244–269, doi:10.1016/j.physletb.2018.03.084, arXiv:1710.04960.

[51] CMS Collaboration, "Search for a heavy resonance decaying to a pair of vector bosons in the lepton plus merged jet final state at $\sqrt{s} = 13$ TeV", *JHEP* **05** (2018) 088, doi:10.1007/JHEP05(2018)088, arXiv:1802.09407.

[52] CMS Collaboration, "Search for massive resonances decaying into WW, WZ or ZZ bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2017) 162, doi:10.1007/JHEP03(2017)162, arXiv:1612.09159.

[53] CMS Collaboration, "Search for ZZ resonances in the $2\ell 2\nu$ final state in proton-proton collisions at 13 TeV", *JHEP* **03** (2018) 003, doi:10.1007/JHEP03(2018)003, arXiv:1711.04370.

[54] CMS Collaboration, "Search for new resonances decaying via WZ to leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **740** (2015) 83, doi:10.1016/j.physletb.2014.11.026, arXiv:1407.3476.

[55] CMS Collaboration, "Search for massive WH resonances decaying into the $\ell\nu b\bar{b}$ final state at $\sqrt{s} = 8$ TeV", *Eur. Phys. J. C* **76** (2016) 237, doi:10.1140/epjc/s10052-016-4067-z, arXiv:1601.06431.

[56] CMS Collaboration, "Search for a massive resonance decaying into a Higgs boson and a W or Z boson in hadronic final states in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JHEP* **02** (2016) 145, doi:10.1007/JHEP02(2016)145, arXiv:1506.01443.

[57] CMS Collaboration, "Search for narrow high-mass resonances in proton-proton collisions at $\sqrt{s} = 8$ TeV decaying to a Z and a Higgs boson", *Phys. Lett. B* **748** (2015) 255, doi:10.1016/j.physletb.2015.07.011, arXiv:1502.04994.

[58] CMS Collaboration, "Search for resonances decaying to a pair of higgs bosons in the $b\bar{b}q\bar{q}\ell\nu$ final state in proton-proton collisions at $\sqrt{s}$ = 13 tev", *Journal of High Energy Physics* **10** (2019) 125, doi:10.1007/JHEP10(2019)125, arXiv:1904.04193.

[59] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

[60] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[61] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR.

[62] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[63] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[64] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in FEWZ", *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.

[65] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 61, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[66] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[67] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[68] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[69] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[70] T. Melia, P. Nason, R. Röntsch, and G. Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG BOX", *JHEP* **11** (2011) 078, doi:10.1007/JHEP11(2011)078, arXiv:1107.5051.

[71] P. Nason and G. Zanderighi, "W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2", *Eur. Phys. J. C* **74** (2014) 2702, doi:10.1140/epjc/s10052-013-2702-5, arXiv:1311.1365.

[72] R. Frederix, E. Re, and P. Torrielli, "Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO", *JHEP* **09** (2012) 130, doi:10.1007/JHEP09(2012)130, arXiv:1207.5391.

[73] H. B. Hartanto, B. Jager, L. Reina, and D. Wackeroth, "Higgs boson production in association with top quarks in the POWHEG BOX", *Phys. Rev. D* **91** (2015) 094003, doi:10.1103/PhysRevD.91.094003, arXiv:1501.04498.

[74] M. Czakon and A. Mitov, "Top++: A program for the calculation of the top-pair cross-section at hadron colliders", *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[75] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[76] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[77] CMS Collaboration, "Extraction and validation of a new set of cms pythia8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

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

[79] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[80] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[81] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[82] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[83] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[84] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[85] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2.

[86] https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentificationRun2.

[87] https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2.

[88] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/p09018, arXiv:2003.00503.

[89] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[90] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms", *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.

[91] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Proceedings of the Workshop on Monte Carlo Generators for HERA Physics, Hamburg, Germany*, p. 270. 1998. arXiv:hep-ph/9907280.

[92] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.

[93] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC", *Phys. Rev. Lett.* **100** (2008) 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.

[94] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[95] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[96] E. Bols et al., "Jet Flavour Classification Using DeepJet", *JINST* **15** (2020), no. 12, P12012, doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.

[97] CMS Collaboration, "Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector", CMS Detector Performance Note CMS-DP-2018-058, 2018.

[98] CMS Collaboration, "Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques", *JINST* **15** (2020) P06005, doi:10.1088/1748-0221/15/06/P06005.

[99] J. Thaler and K. Van Tilburg, "Identifying boosted objects with N-subjettiness", *JHEP* **03** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.

[100] CMS Collaboration, "Measurement of normalized differential $\mathrm{t\bar{t}}$ cross sections in the dilepton channel from pp collisions at $\sqrt{s} = 13$ TeV", *JHEP* **04** (2018) 060, doi:10.1007/JHEP04(2018)060, arXiv:1708.07638.

[101] CMS Collaboration, "Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV", *Phys. Rev. D* **95** (2017) 092001, doi:10.1103/PhysRevD.95.092001, arXiv:1610.04191.

[102] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function", *Ann. Math. Stat.* **27** (1956) 832. http://www.jstor.org/stable/2237390.

[103] B. Silverman, "Density estimation for statistics and data analysis". Chapman and Hall, 1986. ISBN 0412246201.

[104] D. Scott, "Multivariate density estimation: theory, practice, and visualization". John Wiley and Sons, 1992. ISBN 0471547700.

[105] M. J. Oreglia, "A study of the reactions $\psi' \to \gamma\gamma\psi$". PhD thesis, Stanford University, 1980. SLAC Report SLAC-R-236.

[106] J. Gaiser, "Charmonium spectroscopy from radiative decays of the J/ψ and ψ'". PhD thesis, Stanford University, 1982. SLAC Report SLAC-R-255.

[107] M. Cacciari et al., "The $\mathrm{t\bar{t}}$ cross-section at 1.8 TeV and 1.96 TeV: A study of the systematics due to parton densities and scale dependence", *JHEP* **04** (2004) 068, doi:10.1088/1126-6708/2004/04/068, arXiv:hep-ph/0303085.

[108] S. Catani, D. de Florian, M. Grazzini, and P. Nason, "Soft gluon resummation for Higgs boson production at hadron colliders", *JHEP* **07** (2003) 028, doi:10.1088/1126-6708/2003/07/028, arXiv:hep-ph/0306211.

[109] S. Baker and R. D. Cousins, "Clarification of the use of chi square and likelihood functions in fits to histograms", *Nucl. Instrum. Meth.* **221** (1984) 437, doi:10.1016/0167-5087(84)90016-4.

[110] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[111] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[112] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[113] A. Oliveira, "Gravity particles from warped extra dimensions, predictions for LHC", arXiv:1404.0102.

[114] M. Gouzevitch et al., "Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production", *JHEP* **07** (2013) 148, doi:10.1007/JHEP07(2013)148, arXiv:1303.6636.