Study of the Calorimeter Performance for Future Colliders with the High Granularity Calorimeter in CMS and the Full Simulation based on FCC

Chih-Hsiang Yeh

Department of Physics and Center for High Energy and High Field Physics, National Central University, Chung-Li, Taiwan

Abstract

The discovery of Higgs boson at Large Hadron Collider(LHC) in 2012 opened a new door for particle physics. Since then, people have been hard-working on upgrading the LHC to the higher luminosity and center-of-mass (C.M.) energy, so to study the properties of Higgs boson and to discover new particles predicted by physics beyond the standard model. Some people even proposed the idea to build new colliders, such as Future Circular Collider(FCC), etc. near the working LHC space now. For my study, I focus on three topics which are related to the detector performance. First, because the radiation tolerance problem and increasing pileup at High-Luminosity LHC(HL-LHC), the Compact Muon Solenoid(CMS) experiment at LHC will replace the current endcap calorimeter with a silicon-pad High Granularity Calorimeter(HGCAL). I did some studies on the performance for them using data from test beam in CMS and the Hexaboard of its electronics in National Taiwan University(NTU). Second, I used the GEANT4 full simulation to simulate the FCC detector under the very high energy condition and study the jet performance. We expect to help the future facilities to design a high resolution detector, and could have a large potential to find the new physics with the new detector.

Contents

1 Introduction

1.1 Motivation and Expectation

After identifying the Higgs boson at LHC in 2012, people are eager to explore the new particle and the physics Beyond Standard Model(BSM) , such as dark matter candidates, Z' bosons, heavy Higgs, and so on. CMS and ATLAS collaboration are also upgrading the detectors to improve the functions to cope with the new challenging conditions in the future colliders. For the operation in the future, one of the most important issues of detector performance is: How can we boost the efficiency of the distinguishability of signal from the background?

In the upcoming HL-LHC era, the pileup will be raised significantly compared with the condition the LHC is operating at currently. Therefore, first of all, we need to identify the particle very well. Otherwise, we are not able to reject the unwanted contaminant particles. My first topic with HGCAL in CMS focuses on the particle identification with electrons and pions. Because we will keep searching on more new physics with the electron final states , but some pions will be misidentified as the electrons, our requirements are removing these mistagged "fake" electrons. It is expected that we can find out the optimized cuts for them and apply in the cut-based analysis, aiding the signal and background studies that are very sensitive to pions. Furthermore, I studied some cross-talk phenomena on the PCB boards which are attached with the sensor used in the HGCAL, and try to find out the correlation between the injection pulse strength and cross-talk. We expect to understand the noise created in the electronics and reduce them in the future.

Some have proposed colliders with the higher C.M. energy for the next generation, such as Future Circular Collider (FCC), Circular Electron-Positron Collider (CEPC), Super proton-proton Collider (SppC) etc. When those colliders operate, other crucial problems with higher C.M. energy will arise. Under these circumstances, the jets from the segmentation will be very boosted. In this case, we need to explore other ways to look into this boosted jets structure. Otherwise, we can't separate the signal from the background very well. My second topic with FCC detector focuses on the boosted conditions, using different jet substructures with various configurations of the FCC detector to see whether the smallest detector configuration will give the best separation power to distinguish signal from the background.

1.2 Summary for research topics

The research activities of Chih-Hsiang Yeh from July 1, 2018 to Febraury 28, 2019 includes the following:

- Research with electronics of HGCAL : Cross-talk studies with the no sensor based Hexaboard in NTU

- Research with HGCAL in CMS : Pion-rejection studies with a CMS HGCAL test-beam prototype EM calorimeter

- Research with simulation FCC detector : Studies of granularity of a hadronic calorimeter for tens-of-TeV jets at a 100 TeV pp collider

I will describe in more detail for all of the items above.

2 Cross-talk studies with the no sensor based Hexaboard in NTU

Cross-talk is the one basic phenomenon exist in the electronics, which is a signal transmitted on one circuit or channel of a transmission system creates an unwanted effect in another circuit or channel. Then, it will lead to giving out the wrong signal from the electronics. We need to quantify this value and know more about this kind of noise, otherwise, it will be a big problem when we use this electronics to reconstruct the events from the sensor.

I did this study with Prof.Stathes Paganis, Prof.Rong-Shayang Lu, master student Chia-Hong Chein from NTU, and Prof.Shin-Shan Eiko Yu from NCU. The contributions of Chih-Hsiang Yeh to this study includes the following:

- Study the correlation between the stability of pedestals(SCA0) and other SCA sampling.
- Study the correlation between the injection pulse strength and the cross-talk.

I will describe the detail as following.

2.1 Experiment apparatus and Event reconstruction

For the experiment, we used the electronics which are composed of three parts, including the module (version 2), RPI Hexa and Raspberry Pi 3(RPI). In the Figure1, it presents the electronics and data flow we applied. For the module, it also be called "Hexaboard", because of its shape is hexagonal-like. The main power of it is to collect analog signal and convert them into the digital format. It can be bonded to silicon sensor, such as the HGCAL, to collect real charge from particle go through the sensor and do physical analysis. It also can be given the electrical pulse manually to simulate the sensor-attached condition. For the RPI Hexa, it is the bridge between FPGAs and Raspberry Pi(RPI). When the signal come out from the four CHIPs of hexaboard, first of all, we will collect them to the "slave" FPGA on the HexaBoard. Then after finishing collecting the data, the "slave" FPGA will transfer the data to MAX10 FPGA("master" FPGA) on the RPI Hexa. For Raspberry Pi 3(RPI), acting like the small computer, can do many things communicate with the hexaboard. The main power for it is to receive the MAX10 data and can encode to read back the data, and finish reconstructing the events.

Every SKIROC2cms CHIP on the HexaBoard corresponds to 64 channels, while each channel has its own readout circuit(Pre-amplifier, shapers,etc.). In each event(run), it recorded 30 numbers to reconstruct the event for every channel. Totally 30 numbers are given by 13 ADC counts in both highgain(HG) and lowgain(LG) plus 2 TOTgain(Time Over Threshold) and 2 TOAs(Time Of Arrival). The 13 ADC counts of HG and LG come from 13 SCA(Switched Capacitor Array) units in the circuit, which sample the
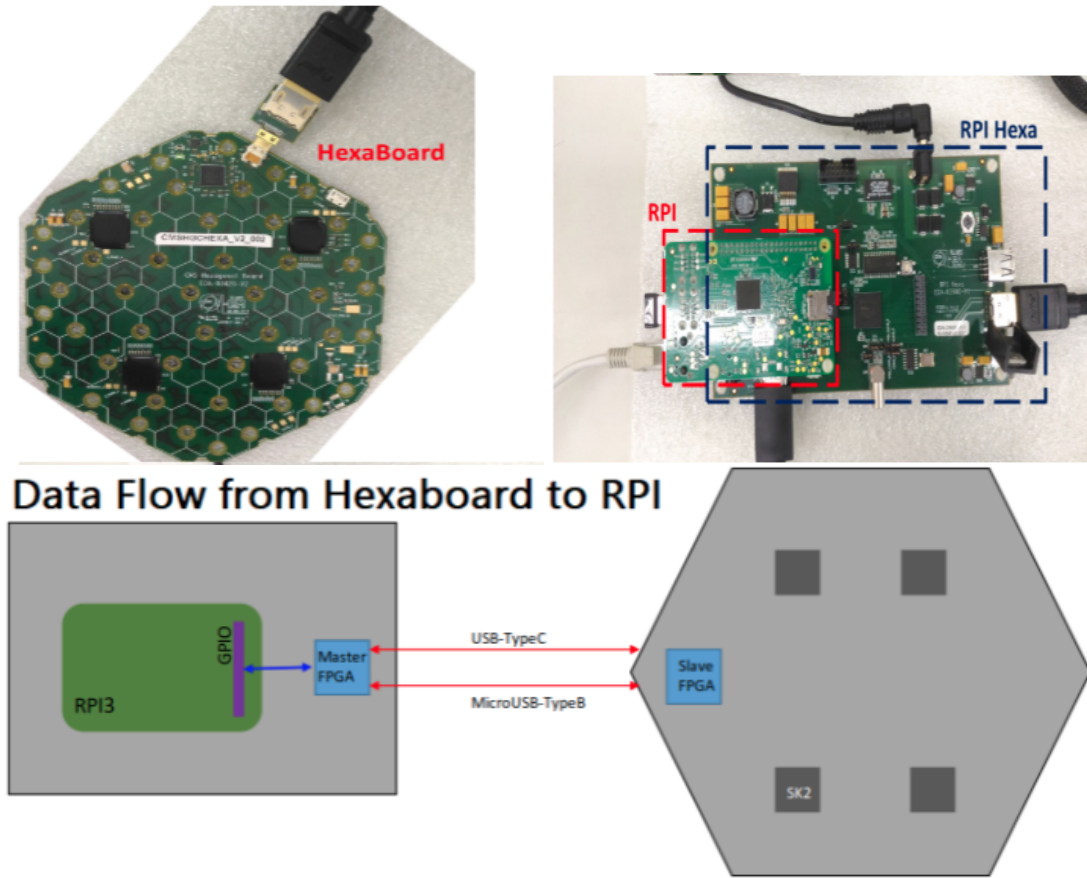
Figure 1: (Upper left) The Hexaboard which was applied in the study. On the board, there are four SKIROC2cms CHIPs (black hexagonal) used to receive the electrical pulse which is given by the cable, and give out the signal to the Hexa RPI. (Upper right) The Hexa RPI and RPI which is applied to transform the data and communicate with Hexaboard. (Down) The data flow pictures for introducing the data transfom from Hexaboard to Hexa RPI and RPI.

input signal every 25ns[1]. They can be used to define the hardware noise and differences between each capacitor. At the same time, we recorded the time stamp with the pulse trigger[2] for each event. Time stamp label the whole pulse in order, so we can use the time stamp to mark the pulse location, specially for the peak of pulse.

2.2   Events and Methods for analysis

For simplifying the study, we used the Hexaboard without the sensor on it. In this studies, we used two types of runs to do the researches through the whole process:

---

[1]There exist thirteen different capacitors with the number SCA0, SCA1....SCA12, and they will record one ADC value in order every 25ns for the each channel of CHIPs. When going to SCA12 and finishing recording, it will return to SCA0. The order of SCA numbers are not related to the order of the electrical pulse we give, but the time stamps are.

[2]In real case, it recorded the starting SCA and the end of SCA with roll position array when the pulse trigger is on and off, and rearrange the SCAs with this array as the time stamp, so it can be specified as the pulse information.

- Pedestal run: Run without injecting the electrical pulse, and record the non-pulse run to be our reference (intrinsic noise).
- Charged run: Run with injecting the electrical pulse to the certain channel, and record the with-pulse run to do the study.

Be attention with, because we simulated the sensor-attached condition, and sensor only apply the 32 channels on the Hexaboard, we used 32 channels of the Hexaboard to study. In the Figure2 is the channel map that we used in the study.
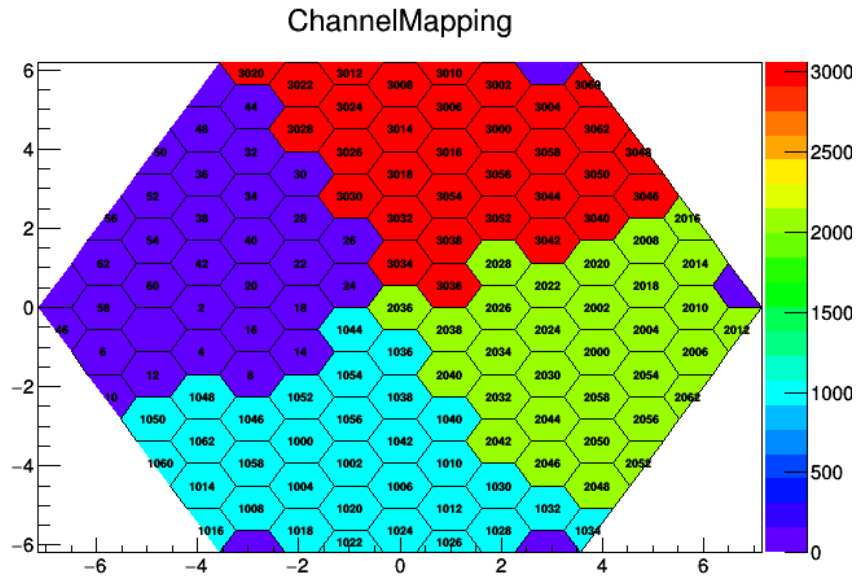

Figure 2: The channel map which was applied in the research, there are four regions with the different color represent the different SKIROC2cms CHIPs. A digit in thousands is the order of the CHIPs, and a digit in tens and ones represent the channels on each CHIPs.

The following three cases are shown in the fig.3. (1)Pedestal run were used to evaluate the pedestals and noise for each channel. The value of mean of ADC-counts will be defined as the pedestal value from each channel, and it is the hardware-dependent(SCA-dependent) value. (2)Charged run were used to record the responses with ADC-counts values from every channel with the electrical pulse, and simulated the real physics. (3)In the end, the "Pedestal-subtracted" value will be used in our study for subtracting the "hardware difference". The mean value of the ADC-counts for it is calculated after Charged run ADC-counts subtract the mean of ADC-counts from Pedestal run with same SCA number. Note, in the study, we always used SCAs of HG to do. For simplifying the case, we fixed the injected-channel to number 20 in CHIP 0 and saw all cases.
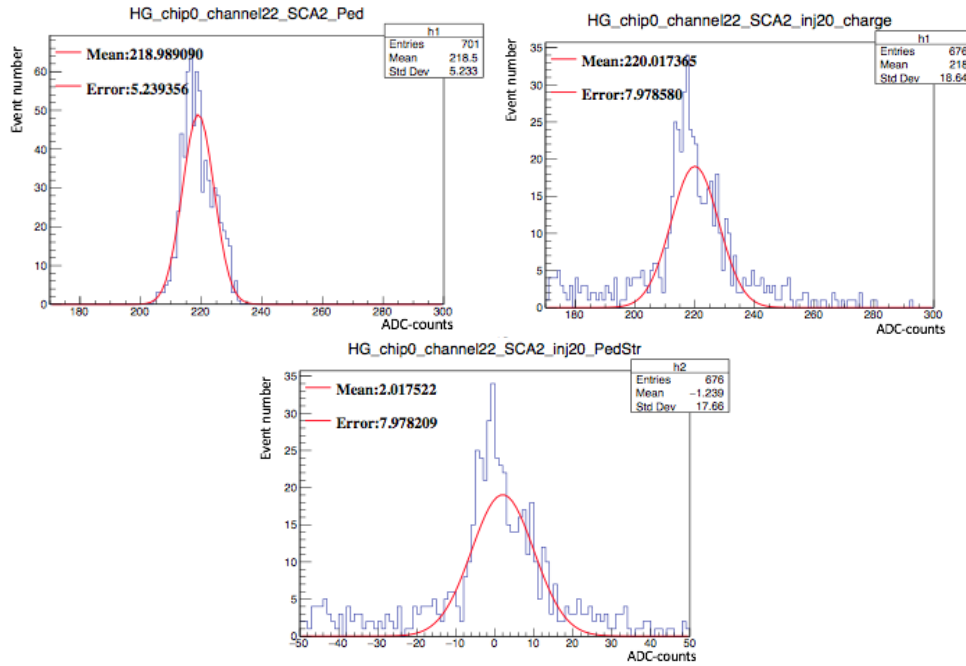
Figure 3: These figures perform the histograms that using in the study. All this cases are shown with fixing at SCA=2 and channel=22. (Upper left) The pedestal run histograms which is used to calculate the pedestal and noise. Fitting the line is to get the mean of ADC-counts. (Upper right) The charged run for example. (Down) The pedestal-subtracted case is shown. Also fit the line to get the mean and error of ADC-counts, and use in the study.

2.3   Results and Conclusion

First of all, because the first SCA (the 0th) is useful for making on-line pedestal subtraction (because there may be some low frequency noise), and usually our signal comes after SCA0, we need to explore the correlation between the stability of pedestals(SCA0) and other SCA sampling. In the Fig.4, we can see that the nearest channel of the injection channel (channel 20) with channel 18 and 22 as examples, the stability of them are very good. Because the fitting lines are pretty flat in both of them, and that means in the different SCA number other than SCA0, they are slightly different.

Second, we wanted to explore the correlation between injection pulse strength and the cross-talk. In real physics, because particles will have different energies individually in the collider, they will see the different pulse strength in the real case. At there, we wanted to see whether the different pulse strength will give out more or less the cross-talk. In the fig.5, they show that the mean and error values from Pedestal-subtracted value for different DAC of injection pulse. The results for them are quiet obviously showing that the channels near the injection channel, when the DAC is higher, the cross-talk will be more significant. Oppositely, when the channel is farther from the injection channel, it won't be influenced by the increased cross-talk when the DAQ go much higher.

For the conclusion, we can see from our study that the cross-talk exist in the Hexaboard, and we used the ADC-counts from the Pedestal-subtracted values to quantify
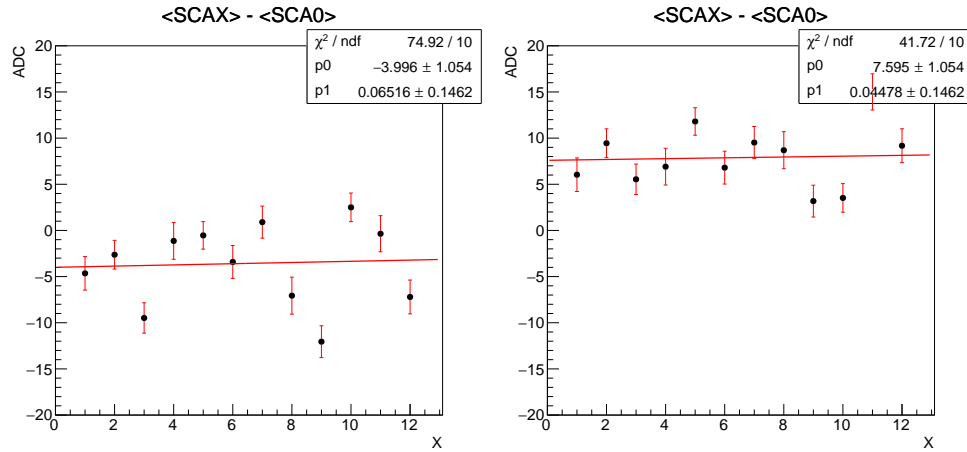
Figure 4: The figures show the difference mean of ADC-counts between the first SCA and other SCAs. (Left) The channel 18 response and (Right) The channel 22 response.
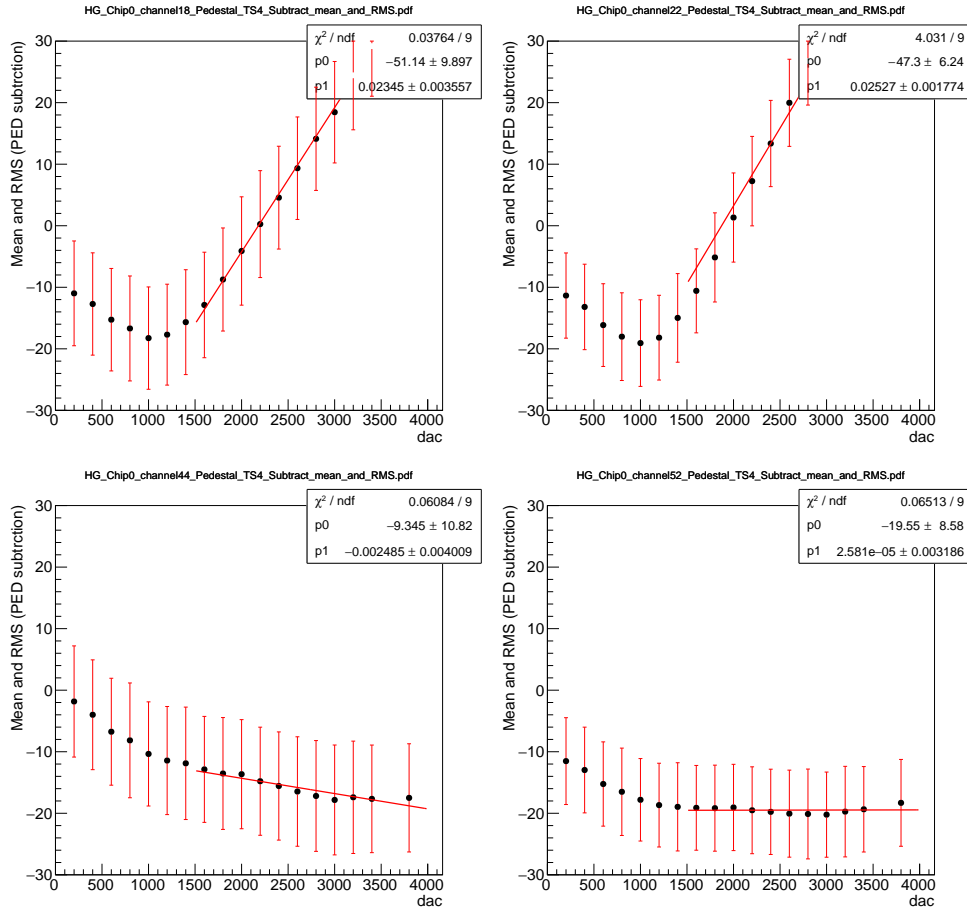
Figure 5: These figures present the correlation between the different DAQ input and Pedestal-subtracted mean and error. The upper left and right pictures perform the channel of 18 and 22, which are closed to the injection channel. The down left and right pictures show channel of 44 and 52, which are farther from the injection channel.

the cross-talk. We expected that this study can help to quantify the cross-talk and can decrease it in the future.

3 Pion-rejection studies with a CMS HGCAL test-beam prototype EM calorimeter

Hypotheses for physics beyond the standard model (SM) predict the existence of heavy resonances that decay to any combination of two among the massive vector bosons (W or Z, collectively referred to as V) or to a V and the scalar SM Higgs boson (H). Among the considered models are those dealing with warped extra dimensions (WED) [1, 2] and composite-Higgs bosons [3–6]. The benchmark models considered are a heavy vector triplet (HVT) model [7] and the bulk scenario [8–10] (G_{bulk} graviton) in the Randall–Sundrum (RS) WED model [1, 2]. The HVT model generalizes a large number of models that predict spin-1 resonances, such as those in composite-Higgs theories, which can arise as a singlet, either W' or Z' [11–13], or as a V' triplet (where V' represents W' and Z' bosons) [7]. The HVT and G_{bulk} models are considered as benchmarks for diboson resonances with spin 1 ($W' \rightarrow$ WZ or WH, $Z' \rightarrow$ WW or ZH), and spin 2 ($G_{bulk} \rightarrow$ WW or ZZ), respectively, produced via quark-antiquark annihilation ($q\overline{q}^{(\prime)} \rightarrow W'$, $q\overline{q} \rightarrow Z'$) and gluon-gluon fusion (gg $\rightarrow G_{bulk}$).

This analysis presents a search for resonances with masses above 0.6 TeV decaying into a pair of vector bosons. The analysis is based on data collected in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS experiment at the CERN LHC during 2015, corresponding to an integrated luminosity of 2.3 and 2.7 fb^{-1} for the $\ell\nu q\overline{q}$, where $\ell = \mu$ or e, and $q\overline{q}q\overline{q}$ final state, respectively. The $\ell\nu$+jet search also includes the $W \rightarrow \tau\nu$ contribution from the decay $\tau \rightarrow \ell\nu\overline{\nu}$. The gain in sensitivity from τ leptons is limited by the small branching fractions involved. The key challenge of the analyses is the reconstruction of the highly energetic decay products. Since the resonances under study have masses of order 1 TeV, their decay products, i.e. the bosons, have on average transverse momenta (p_T) of several hundred GeV or more. As a consequence, the particles emerging from the boson decays are very collimated. In particular, the jet-decay products of the bosons cannot be resolved using the standard algorithms, but are instead reconstructed as a single jet object. Dedicated techniques, called jet "V tagging" techniques, are applied to exploit the substructure of such objects, to help resolve jet decays of massive bosons [14, 15]. The V tagging also helps suppress SM backgrounds, which mainly originate from the production of multijet, W+jets, and nonresonant VV events.

The final states considered are either $\ell\nu q\overline{q}$ or $q\overline{q}q\overline{q}$, where the hadronic decay products of the V decay are reconstructed in a single jet. They result in events with either a charged lepton, a neutrino, and a single reconstructed jet ($\ell\nu$+jet channel), or two reconstructed jets (dijet channel). As in the analyses of previous data [16, 17], the aim is to reconstruct all decay products of the new resonance to be able to search for a localized enhancement in the diboson invariant mass spectrum. While the analyses in general aim at large resonance masses, we conduct two exclusive searches in the $\ell\nu$+jet final state, separately optimized for the mass ranges 0.6–1.0 TeV ("low-mass") and > 1 TeV ("high-mass").

The AK8 jets are used to reconstruct the W jet and Z jet candidates from their decays to highly boosted quark jets. To discriminate against m ultijet backgrounds, we exploit

both the reconstructed jet mass, which is required to be close to the W or Z boson mass, and the two-prong jet substructure produced by the particle cascades of two high-p_T quarks that merge into one jet [15]. Jets that are i dentified as arising from the merged decay products of a V boson are hereafter referred to as "V jets".

As a first step in exploring potential substructure, the jet constituents are subjected to a jet grooming algorithm that improves the resolut ion in the jet mass and reduces the effect of pileup [18]. The goal of the algorithm is to recluster the jet constituents, while applying additional requirements that eliminate soft, large-angle QCD r adiation that increases the jet mass relative to the initial V boson mass. Different jet grooming algorithms have been explored at CMS, and their perf ormance on jets in multijet processes has been studied in detail [18]. In this analysis, we use the *jet pruning* [19, 20] algorithm for the main analysis and the *jet trimmin g* algorithm [21] at the trigger level as well as for cross checks. Jet pruning reclusters each AK8 jet starting from all its original constituents, through the implementation of the Cambridge-Aachen (CA) algo rithm [22, 23] to discard "soft" recombinations in each step of the iterative CA procedure. The pruned jet mass, m_{jet}, is computed from the sum of the four-momenta of the constituents that are not removed by the pruning; it is then scaled by the same factor as that used to correct the original jet p_T. The jet is considered as a V jet candidate if m_{jet} falls in the rang e $65 < m_{jet} < 105\,\text{GeV}$, which we define as the signal jet mass window. In the low-mass analysis, only W jet candidates are considered , thus the mass window applied is $65 < m_{jet} < 95\,\text{GeV}$.

Additional discrimination against jets from gluon and single-quark hadronization is obtained from the quantity called *N-subjettiness* [24]. The constituents of the jet before the pruning procedure are reclustered using the k_T algorithm [22,], until N joint objects (*subjets*) remain in the iterative combination procedure of the algorithm. The N-subjettiness , τ_N, is then defined as

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k}), \tag{1}$$

where the index k runs over the PF constituents of the jet, and the distances $\Delta R_{n,k}$ are calculated relative to the axis of the n-th subjet. The normalization factor d_0 is calculated as $d_0 = \sum_k p_{T,k} R_0$, setting R_0 to the distance parameter used in t he clustering of the original jet. The variable τ_N quantifies the compatibility of the jet clustering with the hypothesis that exactly N subjets are present, with sma ll values of τ_N indicating greater compatibility. The ratio between 2-subjettiness and 1-subjettiness, $\tau_{21} = \tau_2/\tau_1$, is found to be a powerful discriminant between jets originat ing from hadronic V decays and from gluon and single-quark hadronization. Jets from W or Z decays in signal events are characterized by lower values of τ_{21} relative to SM backgrounds. We reject V jet candidates with $\tau_{21} > 0.75$. The remaining events are further categorized according to their value of τ_{21} to enhance t he sensitivity of the analysis.

Since data/simulation discrepancies in the jet substructure variables m_{jet} and τ_{21} can bias the signal efficiency estimated from simulated samples, the modelling of signal efficiency is cross-checked in a signal-free sample with jets having characteristics that are similar to those expected for a genuine signal. A sample of high-p_T W bosons that decay to quarks, and are reconstructed as single AK8 jets, is studied in t$\bar{\text{t}}$ and single

top quark event s. Scale factors for the τ_{21} selection efficiency are extracted following Ref. [15]. In this method, the pruned jet mass distributions of events that pass and fail the τ_{21} selection are fitted simultaneously to separate t he W boson signal from the combinatorial components in the top quark enriched sample in both data and simulation. The scale factors are used to correct the total signal efficiency and the VV backgroun d normalization predicted by the simulation. The uncertainties in the scale factors quoted for the τ_{21} selection include two systematic uncertainties. One comes from the modelling o f the nearby jets and p_T spectrum in $t\bar{t}$ MC events, obtained by comparing LO and NLO $t\bar{t}$ simulation. The other is due to the choice of the models used to fit signal and background. The quadratic sum of these systematic uncertainties is found to be smaller than half of the st atistical uncertainty in the scale factor. An additional uncertainty is calculated to account for the extrapolation of the scale factor from $t\bar{t}$ events with an average jet p_T $ox200\,$GeV to higher momenta. This is estimated from the difference between PYTHIA and HERWIG++ [25] showering models wi th resulting factors of $4.5\%\ln(p_T/200\,\text{GeV})$ and $5.9\%\ln(p_T/200\,\text{GeV})$ for $\tau_{21} < 0.6$ and $\tau_{21} < 0.45$, respectively. For the $0.45 < \tau_{21} < 0.75$ selection, this uncertainty is increased by the ratio of the uncertainties in the scale factors, and treated as anti-correlated with the uncertainty for $\tau_{21} < 0.45$. The mean $\langle m_{\text{jet}} \rangle$ and resolution σ value of the Gaussian component of the fitted W jet mass are also extracted to obtain co rrections that are applied to the simulated pruned jet mass. The mass peak position is slightly shifted relative to the W boson mass because of the extra energy deposited in the jet cone from pileup, un derlying event, and initial-state radiation not completely removed in the jet pruning procedure. For events with top quarks, additional energ y contributions arise also from the possible presence of a b jet close to the W jet candidate. Because the kinematic properties of W jets and Z jets are very similar, the same corrections are also used when the V jet is assumed to arise from a Z boson.

The m_{VV} distribution observed in data is dominated by SM background processes where single quark or gluon jets are falsely identified as V jets. Depending on the final state, the dominant processes are multijets (dijet channel) and inclusive W boson production ($\ell\nu$+jet channel). Subdominant backgrounds include $t\bar{t}$, single top quark, and nonresonant diboson SM production.

In the $\ell\nu$+jet channel, the multijet background is predicted to be negligible from MC simulation, whereas it represents the major contributio n in the dijet analysis. For the latter, we assume that the SM background can be described by a smooth, parametrizable, monotonically decreasing distribution. The search is performed by separately fitting the background function to each search region and simultaneously adding resonant Breit–Wigner (BW) forms across all search regions to represent the signal. The background probability function is defined by empirical functional forms of 3 and 2 parameters, respectively:

$$\frac{\mathrm{d}N}{\mathrm{d}m_{\text{jj}}} = \frac{P_0(1 - m_{\text{jj}}/\sqrt{s})^{P_1}}{(m_{\text{jj}}/\sqrt{s})^{P_2}} \qquad \text{and} \qquad \frac{\mathrm{d}N}{\mathrm{d}m_{\text{jj}}} = \frac{P_0}{(m_{\text{jj}}/\sqrt{s})^{P_2}}, \qquad (2)$$

where m_{jj} is the dijet invariant mass (equivalent to the diboson candidate mass m_{VV} for this channel), $\sqrt{s}$ is the pp collision energy in the centre of mass, P_0 is a normalization parameter, and P_1 and P_2 parametrize the shape of the m_{VV} distribution. Starting from

the two-parameter functional form, a Fisher F-test is used to check at 10% confidence level (CL) if additional parameters are needed to model the background distribution. For the WW categories and the WZ HP category, the two-parameter form is found to describe the data spectrum sufficiently well, while for all other channels the three-parameter functional form is preferable. Alternative parametrizations and functions with up to five parameters are also studied as a cross-check.

The backgrounds from $t\bar{t}$ and single top quark production in the $\ell\nu$+jet channel are estimated from data-based correction factors in the normalization of the simulation. A top quark enriched control sample is selected by applying all the analysis requirements in $\ell\nu$+jet events except that the b jet veto is inverted by requiring, instead, at least one b-tagged AK4 jet in the event. From the comparison between data and simulation, normalization correction factors for $t\bar{t}$ and single top quark background processes are ev aluated in the pruned jet mass regions $65 < m_{\text{jet}} < 105$ GeV and $65 < m_{\text{jet}} < 95$ GeV, for the electron and muon channels, and for the low- and high-mass selections, separately. The scale factors include both the W boson signal and the combinatorial components mainly due to events where the extra b jet from the top quark decay is in the proximity of the W, and are used to correct the normalization of the ar and single top quark simulated background predictions in the signal regions.

The W+jets background in the $\ell\nu$+jet channel is estimated through the *α ratio method*. This method assumes that the correlation between m_{jet} and m_{VV} for the dominant W+jets background can be adequately modelled by simulation . A signal-depleted control region (sideband) is defined by requiring the mass of the V jet to lie below or above the nominal selection; the VV distribution observed in this region is then extrapolated to the nominal region through a transfer function estimated from simulation. O ther minor sources of background, such as $t\bar{t}$, single top quark, and SM diboson production, are estimated from simulation after applying c orrection factors based on control regions in data. The sideband region is defined around the jet mass window. The lower and upper sidebands correspond to t he m_{jet} ranges 40–65 and 135–150 GeV, respectively. The Higgs boson mass region, defined by the range 105--135 GeV, corresponds to the signal region of searches for diboson in final states with highly Lorentz-boosted Higgs bosons [26], and is therefore not used to estimate the background.

The overall normalization of the W+jets background in the signal region is determined from a fit to the m_{jet} distribution in the lower and upper sidebands of the data. The analytical form of the fitting function is chosen from simulation studies, as are the contributions from minor backgrounds.

The m_{VV} distribution observed in data and the SM background prediction are compared to check for the presence of a new resonance decaying to vector bosons. No bins with an excess with significance larger than three standard deviations are observed. We set upper limits on the production cross section of such resonances by combining the event categories of the dijet and $\ell\nu$+jet analyses. We follow the asymptotic approximation [27] of the CL_S criterion described in Refs. [28, 29]. The limits computed following this approach are found to agree with the results obtained using the modified frequentist prescription [28, 29]. For each channel and each signal hypothesis a likeli-

hood function is built from the reconstructed m_{VV} mass distribution observed in data, the background prediction, and the signal resonance shape. A maximum-likelihood fit to the data is then performed to obtain the best estimate of the signal cross section. Systematic uncertainties are profiled [30] as log-normal nuisance parameters in the statistical interpretation, and for each possible value of the fitted signal cross section they are all refitted to maximize the likelihood.

Exclusion limits are set in the context of the bulk graviton model and of the HVT Models A and B, under the assumption of a natural width negligible compared to the experimental resolution (narrow-width approximation). To maximize the sensitivity of the search we combine the results from all the analysis channels in each of the considered signal hypotheses. In the combination, the systematic uncertainties in jet momentum scale and resolution, V tagging efficiency scale factors, and integrated luminosity are assumed to be 100% correlated.

Figure 6 shows the resulting expected and observed exclusion limits at 95% CL on the signal cross section as a function of the resonance mass for all signal hypotheses. The limits are compared with the product of cross section and branching fraction ($\sigma \mathcal{B}$) to WW and ZZ for a bulk graviton with $k/\overline{M}_{\text{Pl}} = 0.5$, and with $\sigma \mathcal{B}$ for WZ and WW for spin-1 particles predicted by the HVT Mo dels A and B. In this context, we consider a scenario where we expect the W' and Z' bosons to be degenerate in mass (triplet hypothesis). In addition, we consider also a scenario where only a charged (W') or a neutral (Z') resonance is expected at a given mass (singlet hyp othesis). Combining the analyses leads to about 10–30% more stringent expected upper limits on the cross section compared to the most sensitive individual channel, depending on the resonance mass and the signal hypothesis. For G_{bulk}, Z' and triplet signals (W' signal) with masses <0.8 TeV (<0.75 TeV), the limits are obtained from the lo w-mass $\ell\nu$+jet channel, while for the higher masses they are obtained from the high-mass $\ell\nu$+jet and dijet channels. The dijet analysis provides more stringent expected upper limits on the cross sections than the $\ell\nu$+jet analysis for resonance masses above 1.7 TeV for Z' and >1.3 TeV for W', because of the larger branching fractions: $\mathcal{B}(\text{WW} \rightarrow \text{q}\overline{\text{q}}\text{q}\overline{\text{q}}) = 44\%$, $l\mathcal{B}(\text{WW} \rightarrow \ell\nu\text{q}\overline{\text{q}}) = 31\%$, $\mathcal{B}(\text{WZ} \rightarrow \text{q}\overline{\text{q}}\text{q}\overline{\text{q}}) = 46\%$, and $\mathcal{B}(\text{WZ} \rightarrow \ell\nu\text{q}\overline{\text{q}}) = 16\%$. In f act, the combination of high- and low-purity categories, together with the weak dependence of tagging efficiency on p_{T} [] improves the sensitivity for most potential signal models. In the narrow-width bulk graviton model, the combined sensitivity of the searches is not large enough to set mass limits, however, cross sections are excluded in the range 3–1200 fb. For HVT Model A (B), the combined data exclude singlet W' resonances with masses <2.0 (2.2) TeV and Z' resonances with masses below <1.6 (1.7) TeV. Under the triplet hypothesis, spin-1 resonances with masses <2.3 and <2.4 TeV are excluded for HVT Models A and B, respectively.

Figure 7 shows a scan of the coupling parameters and the corresponding observed 95% CL exclusion contours in the HVT model for the combined analyses. The parameters are defined as $g_V c_H$ and $g^2 c_F / g_V$, related to the coupling strengths of the new resonance to the Higgs boson and to fermions. The range of the scan is limited by the assumption that the new resonance is narrow. A contour is overlaid, representing the region where the theoretical width is larger than the experimental resolution of the searches, and hence where the narrow-resonance assumption is not satisfied. This con-
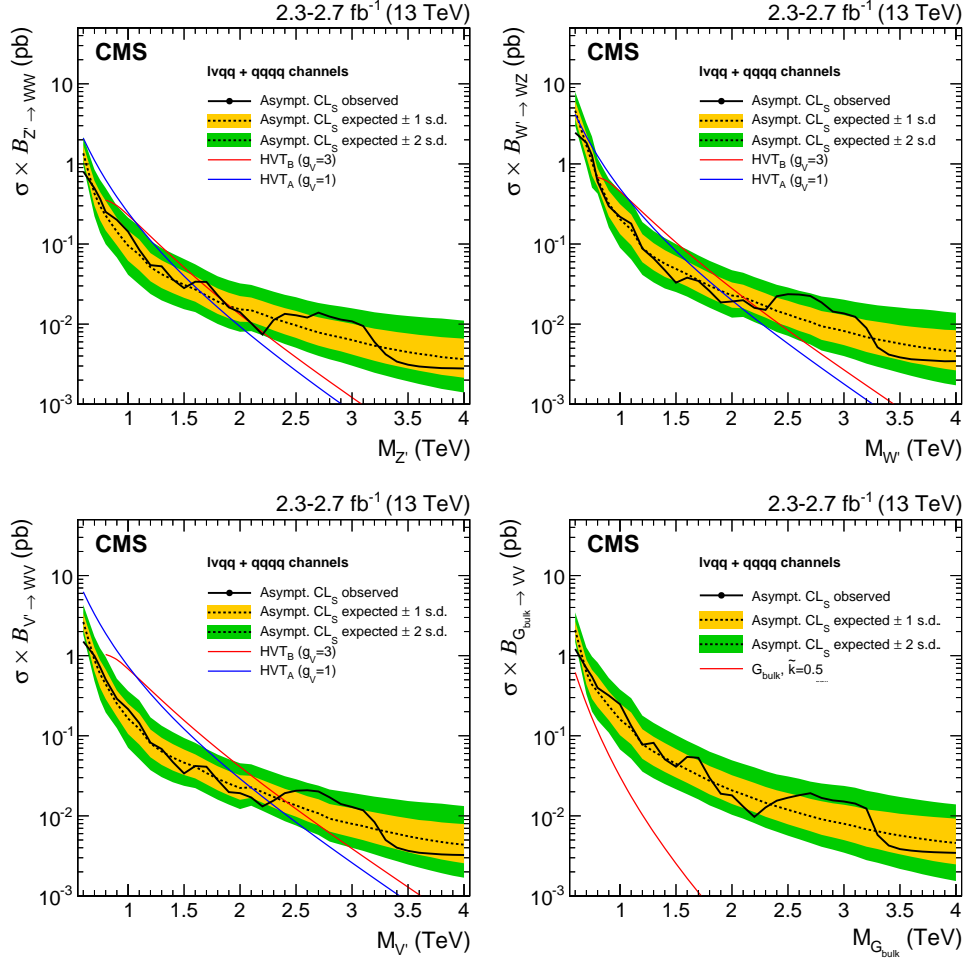
Figure 6: Observed (black solid) and expected (black dashed) 95% CL upper limits on the production of a narrow-width resonance decaying to a pair of vector bosons for different signal hypotheses. In the upper plots, limits are set in the context of a spin-1 neutral Z′ (left) a nd charged W′ (right) resonances, and compared with the prediction of the HVT Models A and B. In the lower left plot, limits are set in the same model under the tr iplet hypothesis (W′ and Z′). In the lower right plot, limits are set in the context of a bulk graviton with $k/\overline{M}_{\text{Pl}} = 0.5$ and compared with the prediction. For G_{bulk}, Z′ and triplet signals (W′ signal) with masses <0.8 TeV (<0.75 TeV), the limits are obtained from the lo w-mass $\ell\nu$+jet channel, while for the higher masses they are obtained from the high-mass $\ell\nu$+jet and dijet channels.

tour is defined by a predicted resonance width of 5%, corresponding to the narrowest resonance mass resolution of the searches.

4   Search for heavy resonances decaying into a vector boson and a Higgs boson in the (ll, lnu, nunu) bb final state at $\sqrt{s} = 13$ TeV

The contribution of the group of Prof. Yu to this analysis includes the following:
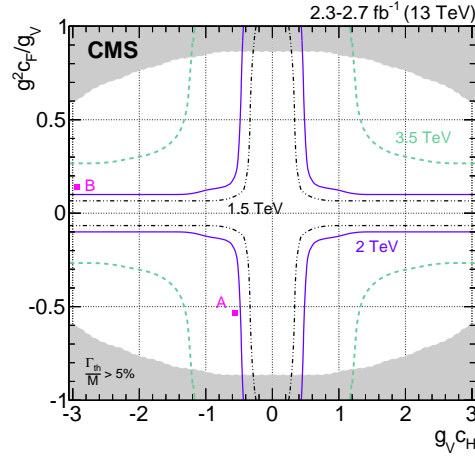
Figure 7: Exclusion regions in the plane of the HVT couplings ($g^2 c_F / g_V, g_V c_H$) for three resonance masses, 1.5, 2.0, and 3.5 TeV. Model points A and B of the benchmarks used in the analysis are also shown. The solid, dashed, and dashed-dotted lines represent the boundaries of the regions excluded by this search for different resonance masses (th e region outside these lines is excluded). The areas indicated by the solid shading correspond to regions where the resonance width is predicted to be more than 5% of the resonance mass and the narrow-resonance assumption is not satisfied.

- Optimization of the pruned Higg-mass selection
- Optimization of the lepton p_T, η, $\Delta\eta$, $\Delta\phi$ selection
- Generation of the signal simulation
- Generation of the background simulation to cover the high p_T range
- Estimation of the systematic uncertainty on the Higgs-tagging signal efficiency due to the use of different parton showering
- Estimation of the systematic uncertainty due to fit bias

Within the group, the major contributors are Prof. Shin-Shan Yu, Dr. Vieri Candelise, Dr. Raman Khurana, the PhD student Yu-Hsiang Chang, and the master students Ching-Wei Chen, Jun-Yi Wu, and Henry Yee-Hsian Tong. The 2015 analysis results were pubishled in Phys Lett. B [26]. I will briefly summarize the analysis below.

The discovery of a Higgs boson h at the CERN LHC [31–33] suggests that the standard model (SM) mechanism that connects electroweak (EW) symmetry breaking to the generation of particle masses is largely correct. However, the relatively light value of the Higgs boson mass $m_H = 125.09 \pm 0.21\,\text{(stat)} \pm 0.11\,\text{(syst)}\,\text{GeV}$ [30, 34–36] leaves the hierarchy problem unsolved [37], pointing to phenomena beyond the SM, which could be unveiled by searches at the LHC. Many theories that incorporate phenomena beyond the SM postulate the existence of new heavy resonances coupled to the SM bosons. Among them, weakly coupled spin-1 Z′ [38, 39] and W′ models [11] or strongly coupled Composite Higgs [3, 40, 41], and Little Higgs models [42–44] have been widely discussed.

A large number of models are generalized in the heavy vector triplet (HVT) framework [45], which introduces one neutral (Z′) and two electrically charged (W′) heavy

resonances. The HVT model is parametrized in terms of three parameters: the strength g_V of a new interaction; the coupling c_H between the heavy vector bosons, the Higgs boson, and longitudinally polarized SM vector bosons; and the coupling c_F between the HVT bosons and the SM fermions. In the HVT scenario, model B with parameters $g_V = 3$, $c_H = 0.976$, and $c_F = 1.024$ [45] is used as the benchmark. With these values, the couplings of the heavy resonances to fermions and to SM bosons are similar, yielding a sizable branching fraction for the heavy resonance decay into a SM vector boson W or Z (generically labeled as V) and a Higgs boson [45].

Bounds from previous searches [46–49] require the masses of these resonances to be above 1 TeV in the HVT framework. In this mass region, the two bosons produced in the resonance decay would have large Lorentz boosts in the laboratory frame. When decaying, each boson would generate a pair of collimated particles, a distinctive signature, which can be well identified in the CMS experiment. Because of the large predicted branching fraction, the decay of high-momentum Higgs bosons to $b\bar{b}$ final states is considered. The Higgs boson is reconstructed as one unresolved jet, tagged as containing at least one bottom quark. Backgrounds from single quark and gluon jets are reduced by a jet mass requirement. In order to discriminate against the large multijet background, the search is focused on the leptonic decays of the vector bosons (Z $\to \nu\nu$, W $\to \ell\nu$, and Z $\to \ell\ell$, with $\ell = $ e, μ).

The main SM background process is the production of vector bosons with additional hadronic jets (V+jets). The estimation of this background is based on events in signal-depleted jet mass sidebands, with a transfer function, derived from simulation, from the sidebands to the signal-enriched region. Top quark production also accounts for a sizable contribution to the background in 1ℓ final states, and is determined from simulation normalized to data in dedicated control regions. Diboson production processes, including pairs of vector bosons (VV) and the SM production of Higgs boson and vector boson (Vh), represent minor contributions to the overall background and are estimated from simulation. A signal would produce a localized excess above a smoothly falling background in the distribution of the kinematic variable m_{Vh}, whose definition and relationship to the resonance mass M_X depends on the final state. Results are interpreted in the context of HVT models in the benchmark scenario B [45].

The set of criteria used to identify the Higgs boson candidate is the same for each event category. The highest-p_T AK8 jet in the event is required to have $p_T > 200$ GeV and $|\eta| < 2.5$. The pruned jet mass m_{jet} must satisfy $105 < m_{jet} < 135$ GeV. The region $65 < m_{jet} < 105$ GeV is not used, to avoid overlaps with searches targeting resonant VV final states. In order to discriminate against the copious vector boson production in association with light-flavored jets, events are classified according to the number of subjets (1 or 2) passing the loose b tagging selection; those failing this requirement are discarded.

Events are divided into categories depending on the number (0, 1, or 2) and flavor (e or μ) of the reconstructed charged leptons, and the presence of either 1 or 2 b-tagged subjets in the AK8 jet. The two categories with no charged leptons are referred to collectively as the zero-lepton (0ℓ) channel. Similarly, the single-lepton (1ℓ) and double-lepton (2ℓ) channels each comprise four categories. In total, 10 exclusive categories are

defined.

In the 0ℓ channel, candidate signal events are expected to have a large $p_{\mathrm{T}}^{\mathrm{miss}}$ due to the boosted Z boson decaying into a pair of neutrinos, which escape undetected. Data are collected using triggers that require $p_{\mathrm{T}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$ greater than 90 GeV, without including muons in the $p_{\mathrm{T}}^{\mathrm{miss}}$ or $H_{\mathrm{T}}^{\mathrm{miss}}$ computation. A stringent selection is applied to the reconstructed $p_{\mathrm{T}}^{\mathrm{miss}}$, which is required to be greater than 200 GeV, to ensure that the trigger is fully efficient. The copious multijet production is greatly suppressed by imposing requirements on the minimum azimuthal angular separations between jets and the missing transverse momentum vector, $\Delta\phi(\mathrm{jet}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$. All the AK8 and AK4 jets in the event must satisfy $\Delta\phi(\mathrm{jet}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$. The Higgs boson jet candidate must fulfill the tighter requirement $\Delta\phi(\mathrm{jet}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}) > 2$ and additional criteria designed to remove events arising from detector noise. Events containing isolated leptons with $p_{\mathrm{T}} > 10$ GeV, hadronically-decaying τ leptons with $p_{\mathrm{T}} > 18$ GeV, and photons with $p_{\mathrm{T}} > 15$ GeV are removed in order to reduce the contribution of other SM processes. The $\mathrm{t\bar{t}}$ background contribution is reduced by removing events in which any AK4 jet, excluding the Higgs boson jet candidate, is b tagged using the loose operating point. Because of the lack of visible decay products from the Z boson, reconstruction of the resonance mass is not directly viable. Instead, the Higgs boson jet momentum and the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ are used to compute the transverse mass $m_{\mathrm{Vh}}^{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\mathrm{miss}} E_{\mathrm{T}}^{\mathrm{jet}} \left[1 - \cos\Delta\phi(\mathrm{jet}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})\right]}$. This variable is utilized as an estimator of $\mathrm{M_X}$ for the 0ℓ channel.

Events in the 1ℓ channel are collected requiring one lepton to be reconstructed online. The p_{T} threshold at trigger level is 105 GeV for electrons and 45 GeV for muons. Offline, events are accepted if there is exactly one reconstructed electron or muon with p_{T} larger than 135 GeV or 55 GeV, respectively, passing restrictive selection criteria. Events with additional leptons passing looser selections, or hadronically decaying τ leptons, are discarded. In the single-electron channel, multijet background is reduced by requiring $p_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV. Azimuthal angular separations $\Delta\phi(\ell, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}) < 2$ and $\Delta\phi(\mathrm{jet}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}) > 2$ are required to select a back-to-back topology. As for the 0ℓ selection, events with additional b-tagged AK4 jets are vetoed. The four-momentum of the W boson candidate is quantified using a kinematic reconstruction of the neutrino momentum. The components of the neutrino momentum in the transverse plane are assumed to be equal to $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$. By constraining the invariant mass of the charged lepton and neutrino to be equal to the W boson mass, a quadratic equation is derived for the longitudinal component of the neutrino momentum, p_z^{ν}. The reconstructed p_z^{ν} is chosen to be the real solution with the lower magnitude or, where both the solutions are complex, the real part with the lowest value. If the W boson has a transverse momentum greater than 200 GeV, it is used to construct the resonance candidate mass m_{Vh}, otherwise the event is discarded.

The 2ℓ channel accepts events collected with the same triggers as in the 1ℓ channel. An additional isolated electron or muon with $p_{\mathrm{T}} > 20$ GeV, with the same flavor as the leading one and opposite charge, is required to be reconstructed and identified. In order to increase the signal efficiency, a looser identification requirement is applied to both electrons, and one of the two muons is allowed to be identified only in the tracker.

If the isolation cones of the two muons overlap, the contribution of one is subtracted from the isolation calculation of the other in each case. The Z boson candidates are retained only if the dilepton invariant mass lies between 70 and 110 GeV. The transverse momentum of the Z boson candidate is required to be at least 200 GeV, otherwise the event is removed. Additionally, the separation in η and ϕ between the Z boson candidate and the Higgs boson jet is required to satisfy $|\Delta\eta(Z, \text{jet})| < 5$ and $\Delta\phi(Z, \text{jet}) > 2.5$. Since the $t\bar{t}$ contribution is small, no veto on additional b-tagged AK4 jets is applied. The resonance candidate mass m_{Vh} is defined as the invariant mass of the Z boson and the AK8 jet.

The signal efficiency for the combined 0ℓ, 1ℓ, and 2ℓ channels following these selections is 20–30% for the 2 b-tagged subjet categories for a resonance mass $M_X = 1$ TeV, decreasing to about 10% for $M_X = 4$ TeV. This reduction is due to the degradation of track reconstruction and b tagging performances at very large p_T, and to the smaller angle between the two b quarks, which tend to be reconstructed in one single subjet. The loss of efficiency is recovered by the 1 b-tagged subjet categories, which provide an additional 10% signal efficiency at $M_X = 1$ TeV, and 20% at $M_X = 4$ TeV.

The background estimation follows the similar method as the semileptonic channel of the VV analysis in Section **??**. The shape and normalization of the m_{jet} distribution for the main V+jets background is extracted from a fit of the sum of all contributing processes to the SB data, after fixing the shape and normalization of the subdominant backgrounds. The fits to the m_{jet} distributions are shown in Fig. 8. The normalization of the diboson processes is derived from simulation, while the top quark normalization is taken from the control regions with the exception of the dilepton channels. The procedure is repeated selecting an alternative function to model the m_{jet} distribution for the main background. The difference between the results obtained with the main and the alternative function is considered as a systematic uncertainty. A deficit of 2.4 standard deviations is observed in the 1μ, 2 b tag category.

The shape of the V+jets background distribution in the m_{Vh} variable is obtained via a transfer function determined from simulation as:

$$\alpha(m_{Vh}) = \frac{N_{SR}^{sim,V+jets}(m_{Vh})}{N_{SB}^{sim,V+jets}(m_{Vh})} \tag{3}$$

where $N_{SR}^{sim,V+jets}(m_{Vh})$, $N_{SB}^{sim,V+jets}(m_{Vh})$ are two-parameter probability density functions determined from the m_{Vh} spectra in the SR and the SB of the simulated V+jets sample, respectively. The ratio $\alpha(m_{Vh})$ accounts for the correlations and the small kinematic differences involved in the interpolation from the sidebands to the SR, and is largely independent of the shape uncertainties and the assumptions on the overall cross section. The shape of the main background is extracted from data in the m_{jet} sidebands, after multiplying the obtained distribution by the $\alpha(m_{Vh})$ ratio. The overall predicted background distribution in the SR, $N_{SR}^{pred}(m_{Vh})$, is given by the following relation:

$$N_{SR}^{pred}(m_{Vh}) = N_{SB}^{obs,V+jets}(m_{Vh})\,\alpha(m_{Vh}) + N_{SR}^{sim,t\bar{t}}(m_{Vh}) + N_{SR}^{sim,VV}(m_{Vh}) \tag{4}$$

where $N_{SB}^{obs,V+jets}(m_{Vh})$ is the probability distribution function obtained from a fit to
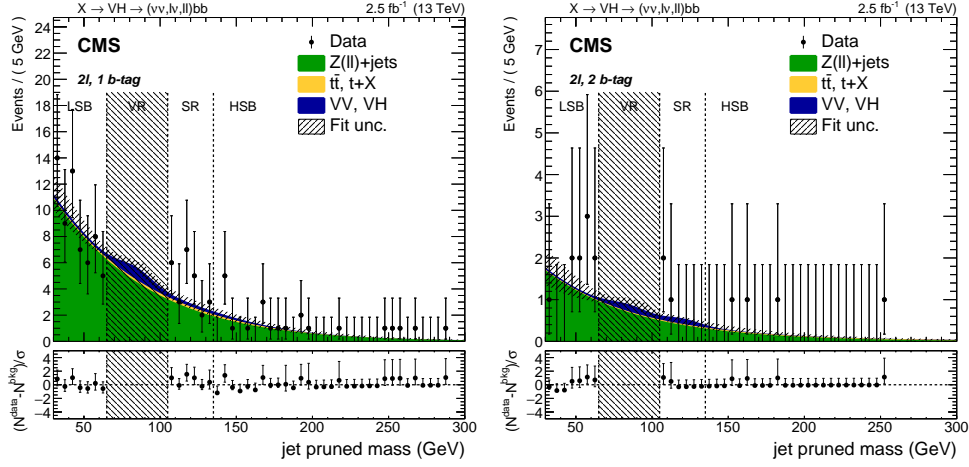
Figure 8: Pruned jet mass distribution of the leading AK8 jet in the 2ℓ category, and separately for the 1 (left) and 2 (right) b-tagged subjet selections. The shaded band represents the uncertainty from the fit to data in the pruned jet mass sidebands. The observed data are indicated by black markers. The dashed vertical lines separate the lower (LSB) and upper (HSB) sidebands, the W and Z bosons mass region (VR), and the signal region (SR). The bottom panels report the pulls in each bin, $(N - N^{\mathrm{bkg}})/\sigma$, where σ is the Poisson uncertainty in data. The error bars represent the normalized Poisson errors on the data.

data in the m_{jet} sidebands of the sum of the background components, and $N_{\mathrm{SR}}^{\mathrm{sim,t\bar{t}}}(m_{\mathrm{Vh}})$ and $N_{\mathrm{SR}}^{\mathrm{sim,VV}}(m_{\mathrm{Vh}})$ are the $t\bar{t}$ and diboson components, respectively, fixed to the shapes and normalizations derived from the simulated samples and control regions. The observed data in the SR are in agreement with the predicted background, as shown in Fig. 9.

The validity and robustness of this method is tested on data by splitting the lower m_{jet} sideband in two and predicting shape and normalization of the intermediate sideband from the lower and upper sidebands. The number of events and distributions found in data are compatible with the prediction within the systematic uncertainties.

The shape of the reconstructed signal mass distribution is extracted from the simulated signal samples. The signal shape is parametrized separately for each channel with a Gaussian peak and a power law to model the lower tails. The resolution of the reconstructed m_{Vh} is given by the width of the Gaussian core for the 1ℓ and 2ℓ channels and by the RMS of the $m_{\mathrm{Vh}}^{\mathrm{T}}$ distribution in the 0ℓ channel, and is found to be 10–16%, 8–5%, 5–3% of M_X in the 0ℓ, 1ℓ, and 2ℓ channels, respectively, when going from low to high resonance masses.

Results are obtained from a combined signal and background fit to the unbinned m_{Vh} distribution, based on a profile likelihood. Systematic uncertainties are treated as nuisance parameters and are profiled in the statistical interpretation. The background-only hypothesis is tested against the $X \to Vh$ signal in the ten categories. The asymptotic modified frequentist method is used to determine limits at 95% confidence level (CL) on the contribution from signal. Limits are derived on the product of the cross section for a heavy vector boson X and the branching fractions for the decays $X \to Vh$

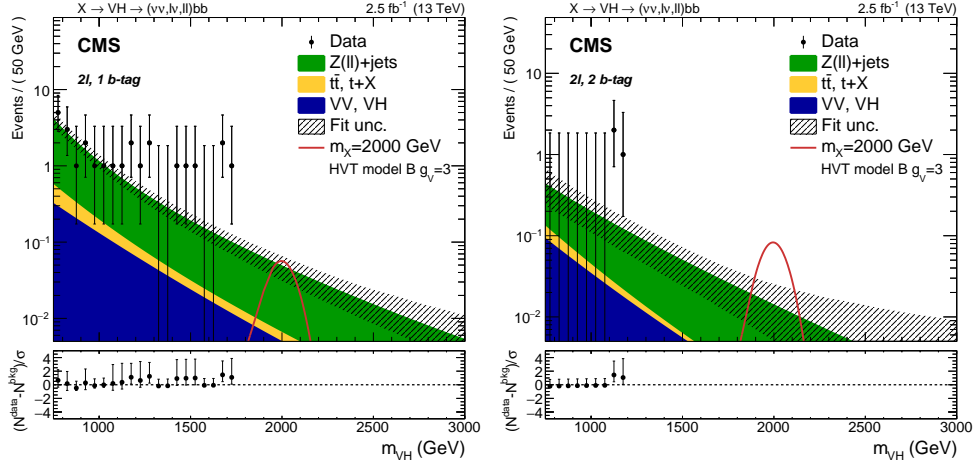Figure 9: Resonance candidate mass m_{Vh} distributions in the 2ℓ category, and separately for the 1 (left) and 2 (right) b-tagged subjet selections. The expected background events are shown with the filled area, and the shaded band represents the total background uncertainty. The observed data are indicated by black markers, and the potential contribution of a resonance with $M_X = 2000$ GeV produced in the context of the HVT model B with $g_V = 3$ is shown with a solid red line. The bottom panels report the pulls in each bin, $(N^{\mathrm{data}} - N^{\mathrm{bkg}})/\sigma$, where σ is the Poisson uncertainty in data. The error bars represent the normalized Poisson errors on the data.

and h $\rightarrow$ b$\overline{\mathrm{b}}$, denoted $\sigma(X)\,\mathcal{B}(X \rightarrow \mathrm{Vh})\,\mathcal{B}(\mathrm{h} \rightarrow \mathrm{b}\overline{\mathrm{b}})$. No specific assumption is made on $\mathcal{B}(\mathrm{h} \rightarrow \mathrm{b}\overline{\mathrm{b}})$, since this decay channel has not yet been measured.

The 0ℓ and 2ℓ categories are combined to provide upper limits for the case where X is a heavy spin-1 vector singlet Z$'$, in the narrow-width approximation. Similarly the 1ℓ categories are combined to provide limits for the case where X is a heavy W$'$. All the 0ℓ, 1ℓ, and 2ℓ channels are combined to put stringent exclusion limits on the HVT model, scenario B, assuming the Z$'$ and W$'$ cross sections as predicted by the model. There are normalization increases caused by event migration between the leptonic channels, which are estimated to be 5–10% in the 0ℓ channel, due to mis-assigned W$'$ events, and less than 1% in the 1ℓ channel, due to mis-assigned Z$'$ events. Figure 10 presents the exclusion limits as a function of the heavy triplet mass. A resonance with $M_X \lesssim 2.0$ TeV is excluded at 95% CL in the HVT model B.

5   Search for dark matter in association with a Higgs boson decaying into a pair of bottom quarks at $\sqrt{s} = 13$ TeV with the CMS detector

This analysis is driven by the group of Prof. Shin-Shan Yu and the contribution includes the following:

- Established the benchmark signal model for the whole mono-h group (including other Higgs decay channels)
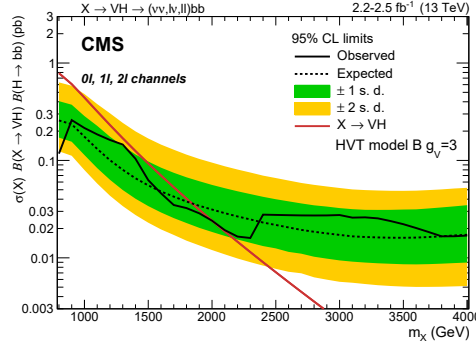- Generated the signal simulation for the whole mono-h group

Figure 10: Observed and expected 95% CL upper limit with the ± 1 and ± 2 standard deviation uncertainty bands on $\sigma(X)\,\mathcal{B}(X \to \mathrm{Vh})\,\mathcal{B}(\mathrm{h} \to \mathrm{b\bar{b}})$ in the HVT model B benchmark scenario with $g_V = 3$ as a function of the resonance mass, for the combination of all the considered channels.

- Study the kinematic distributions and estimated the cross sections of models that predict mono-h signature, such as Z'-2HDM, baryonic Z' model, SMM, 2HDM+a, and the scalar models
- Performed the full 2015 analysis and served as the analysis and paper contacts
- Proved that the bb channel is the most sensitive channel
- Performed data-driven background estimation

The manpower includes Prof. Shin-Shan Yu, Dr. Raman Khurana, former postdoc Yun-Ju Lu, former research assistant Fang-Ying Tsai, master student Ching-Wei Chen, and undergraduate student Shu-Xiao Liu. We worked together with a CERN scientist Michele de Gruttola on this analysis. Our 2015 results are combined with the diphoton channel and published in JHEP [50]. Currently our 2016 analysis has been pre-approved. Again, I will briefly summarize our analysis results below, with focus on the bb channel.

Dark matter (DM) is one of the most compelling pieces of evidence for physics beyond the standard model (SM) [51–53]. Cosmological observations demonstrate that around 85% of the matter in the Universe is comprised of DM (see e.g. [54]). These observations make it likely that DM is primarily composed of weakly interacting massive particles (WIMPs). If non-gravitational interactions exist between DM and SM particles, DM could be produced by colliding SM particles at high energy. In many models, the pair production of DM particles in hadron collisions proceeds through a spin-0 or spin-1 bosonic mediator, with the DM particles leaving the detector without a measurable signature. However, if they are produced in association with a detectable SM particle, which could be emitted as initial state radiation (ISR) off the interacting constituents of the colliding protons, or as part of new effective vertex couplings between DM and SM particles, a large missing transverse momentum (p_T) imbalance could be observed in the event.

Unlike in the case of W bosons, Z bosons, hadronic jets, or photons, the ISR of the SM Higgs boson [55–57] is highly suppressed due to the Yukawa-like nature of its coupling strength. However, the associated production of a Higgs boson and DM particles can

occur in a scheme where the Higgs boson is part of the interaction producing the DM particles, thereby directly probing the structure of the effective DM-SM coupling.

The search presented is interpreted in terms of one benchmark simplified model: a Type-2 Two Higgs Doublet Model with an additional vector Z′ boson (Z′-2HDM) [58–60], where the Z′ boson is produced resonantly and decays into a Higgs boson plus an intermediate heavy pseudoscalar particle A, which decays into two DM particles, $\chi\bar{\chi}$.

In the Z′-2HDM, the Type-2 2HDM is used to formulate the Higgs sector and the gauge symmetry of the SM is extended by a $U(1)_{Z'}$ group with a corresponding new massive Z′ gauge boson. After electroweak symmetry breaking, the Higgs doublets attain vacuum expectation values v_u and v_d, resulting in five physical Higgs bosons: a light neutral CP-even scalar h, assumed to be the observed 125 GeV Higgs boson; a heavy neutral CP-even scalar H; a neutral CP-odd scalar A; and two charged scalars H$^\pm$. Only the masses m_A and $m_{Z'}$ affect the kinematic properties of the physics objects studied in this search.

The mass of the DM particles m_χ is assumed to be less than or equal to 100 GeV. The ratio of the vacuum expectation values of the two Higgs fields coupling to the up-type and down-type quarks, $\tan\beta$, and the coupling of the A particle to DM particles, g_χ, are both fixed at unity. The signal cross sections are calculated for the fixed value of Z′ gauge coupling $g_{Z'} = 0.8$. For the masses of the other Higgs bosons, $m_H = m_{H^\pm} = m_A$ is assumed.

The quantity $\vec{p}_T^{\text{miss}}$, calculated as the negative vectorial sum of the transverse momentum (p_T) of all objects identified in an event, represents the total momentum carried by the DM particles. The magnitude of this vector is referred to as p_T^{miss}. For a given value of $m_{Z'}$, the p_T of the A decreases as m_A increases. Therefore, the p_T^{miss} spectrum softens with increasing m_A.

In the h $\rightarrow$ b$\bar{\text{b}}$ channel, the fact that the p_T of the h should increase with $m_{Z'}$ and decrease with m_A is exploited. The minimum separation in the pseudorapidity and azimuth (η, ϕ) plane between the decay products of h scales as m_h/p_T^h, where p_T^h is the transverse momentum of the h boson. The allowed mass ranges of $m_{Z'}$ and m_A imply a very wide range of values for p_T^h and consequently a wide range in the separation of the decay products. Analysis in this channel is therefore divided into two regimes: (i) a resolved regime where the h decays to two distinct reconstructed b jets, and (ii) a Lorentz-boosted regime where the h is reconstructed as a single fat jet. For each mass point, the analysis with best sensitivity for the expected limit is used as the final result. The signal extraction is performed through a simultaneous fit to the signal- and background-enriched control regions.

The trigger used in the selection of signal-like events requires $p_T^{\text{miss}} > 90$ GeV and $H_T^{\text{miss}} > 90$ GeV, where H_T^{miss} is defined as the magnitude of the vectorial sum of the p_T of all jets in the event with $p_T > 20$ GeV. An additional trigger with a $p_T^{\text{miss}} > 170$ GeV requirement is used to achieve higher efficiency. In this way, events with either high p_T^{miss} or high H_T^{miss} will pass the trigger. For events passing the selection criteria that have $p_T^{\text{miss}} > 170(200)$ GeV for the resolved (boosted) analysis, the trigger efficiency is found to be greater than 98%. The p_T^{miss} threshold for the analysis of the resolved

regime is set slightly lower to enhance the signal efficiency in this region of phase space, where the p_T^{miss} distribution is softer.

Event filters are used to remove spurious high p_T^{miss} events caused by instrumental noise in the calorimeters, or beam halo muons. It has been verified that the efficiency of these filters for accepting signal events is very close to 100%. The main part of the event selection consists of Higgs boson tagging. This selection is different for the resolved and boosted analyses. In the resolved regime, events are required to have two AK4 jets with $p_T > 30\,\text{GeV}$ and $|\eta| < 2.4$. These two jets are used to reconstruct the Higgs boson candidate, which is required to have $p_T > 150\,\text{GeV}$. Each of the two AK4 jets in the resolved regime is required to pass the b tagging selection, whereas in the boosted regime, the two subjets inside an AK8 jet must both pass the b tagging selection. In the boosted regime, the decay products from the Higgs boson are merged. Therefore, an AK8 jet with p_T greater than $200\,\text{GeV}$ is used to reconstruct the Higgs boson. If more than one Higgs boson candidate is reconstructed, the ambiguity is resolved by selecting the candidate with the highest p_T. Backgrounds due to hadronic jets are further reduced by constraining the reconstructed Higgs boson candidate mass, m_{bb}, to be between 100 and $150\,\text{GeV}$. For the resolved regime, the Higgs boson candidate mass is reconstructed using two b-tagged AK4 jets. For the boosted regime, the corrected pruned mass of the AK8 jet with two b-tagged subjets is used as the Higgs boson candidate mass.

Multijet events can act as a source of background when the energy of one of the jets is mismeasured. Therefore, the absolute difference between the azimuthal angles of the vector $\vec{p}_T^{\,\text{miss}}$ and any other AK4 jet with $p_T > 30\,\text{GeV}$ is required to be greater than 0.4 radians. Multijet background is further reduced in the resolved analysis by requiring the azimuthal angle difference between the $\vec{p}_T^{\,\text{miss}}$ and $\vec{p}_{T,\text{trk}}^{\,\text{miss}}$ to be less than 0.7 radians.

Events are rejected if they have any isolated electron (muon) with $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5\,(2.4)$ or any candidates with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.3$ [61–63]. In addition, the events must not have any additional loose AK4 b-tagged jet or more than one additional AK4 jet with $p_T > 30\,\text{GeV}$ and $|\eta| < 4.5$. These vetoes considerably reduce the background from semileptonic top decay modes and leptonic decays of W+jets.

The product of the detector acceptance and selection efficiency varies from 1 to 29%, depending on the values of $m_{Z'}$ and m_A. The average p_T^{miss} increases with $m_{Z'}$ and decreases with m_A.

Several CRs are used to correct the background normalizations with dedicated scale factors. For both resolved and boosted regimes, the selection criteria of these CRs are kept as close as possible to those of the SR, except for the inversion of the additional object vetoes (leptons, jets) and the Higgs boson mass window. This makes the CRs orthogonal to the SR.

For the resolved regime, three CRs are specified: $Z(\to \nu\bar{\nu})$+jets, top quark, and W+jets. The b tagging selection in all the CRs is the same as in the SR in order to minimize the b tagging systematic uncertainties when extrapolating the background scale factors measured in the CRs to the SR. The $Z(\to \nu\bar{\nu})$+jets CR is defined with the same selection as the SR, except for the inversion of the reconstructed Higgs boson mass requirement.

The W+jets and top quark CRs are defined by removing the mass selection and requiring exactly one isolated electron (muon) with $p_T > 10$ GeV and $|\eta| < 2.5$ (2.4). Events with one additional AK4 jet are placed in the top quark CR, whereas events with no additional AK4 jets enter the W+jets CR.

For the boosted regime, the $Z(\to \nu\bar{\nu})$+jets CR is defined by inverting the mass requirement for the AK8 jet. Owing to the low event count and very similar topology between the W+jets and top quark backgrounds it is difficult to construct two separate CRs for W+jets and top quark backgrounds. Hence, the single-lepton CR, a combination of mainly W+jets and top quark events, is defined using the same selection as that for the signal, but requiring exactly one isolated electron (muon) with $p_T > 10$ GeV and $|\eta| < 2.5$ (2.4) and removing the mass requirement.

Since no excess of events has been observed over the SM background expectation in the signal region, the results of this search are interpreted in terms of an upper limit on the production of DM candidates in association with a Higgs boson in the process $Z' \to Ah \to \chi\bar{\chi}h$. The upper limits are computed at 95% confidence level (CL) using a modified frequentist method (CL_s) [64–66] computed with an asymptotic approximation [67]. A profile likelihood ratio is used as the test statistic in which systematic uncertainties are modeled as nuisance parameters. These limits are obtained as a function of $m_{Z'}$ and m_A for both Higgs boson decay channels and for the combination of the two. The two decay channels are combined using the branching ratios predicted by the SM. In the combination of the two analyses, all signal and p_T^{miss}-related systematic uncertainties as well as the systematic uncertainty in the integrated luminosity are assumed to be fully correlated.

Figure 11 (left) shows the 95% CL expected and observed limits on the dark matter production cross section $\sigma(Z' \to Ah \to \chi\bar{\chi}h)$, for $h \to b\bar{b}$ and $h \to \gamma\gamma$ for $m_A = 300$ GeV. These results, obtained with $m_\chi = 100$ GeV, can be considered valid for any dark matter particle mass below 100 GeV since the branching fraction for decays of A to DM particles, $\mathcal{B}(A \to \succ\bar{\prec})$, decreases as m_χ increases. As shown in Figure 11, for the phase space parameters considered for this model (g_χ and $\tan\beta$ equal to unity), results of the combined analysis are mainly driven by the $h \to b\bar{b}$ channel. The combination with the $h \to \gamma\gamma$ channel provides a 2-4% improvement in terms of constraints on the model for the low Z' mass values. Future iterations of this search will explore additional phase spac e regions of the Z'-2HDM model, i.e. larger values of $\tan\beta$, where the $h \to \gamma\gamma$ channel becomes more sensitive than $h \to b\bar{b}$ [58].

Figures 11 (right) and 12 show the 95% CL expected and observed upper limits on the signal strength $\sigma_{95\%CL}(Z' \to Ah \to \chi\bar{\chi}h)/\sigma_{theory}(Z' \to Ah \to \chi\bar{\chi}h)$.

Figures 11 (right) and 12 show the 95% CL expected and observed upper limits on the signal strength $\sigma_{95\%CL}(Z' \to Ah \to \chi\bar{\chi}h)/\sigma_{theory}(Z' \to Ah \to \chi\bar{\chi}h)$. For $m_A = 300$ GeV, the Z' mass range from 600 to 1780 GeV is expected to be excluded with a 95% CL when the signal model cross section is calculated using $g_{Z'} = 0.8$, while the observed data, for $m_A = 300$ GeV, exclude the Z' mass range from 600 to 1860 GeV. When the signal model cross section is calculated using the constrained $g_{Z'}$, the expected exclusion range is 830 to 1890 GeV, and the observed exclusion range is 770 to 2040 GeV. Figure 12 shows the

expected and observed upper limits on the signal strength for the $h \to b\bar{b}$ and $h \to \gamma\gamma$ decay channels. Figure 13 shows the upper limits on the signal strength combining the results from both the $h \to b\bar{b}$ and $h \to \gamma\gamma$ decay channels.
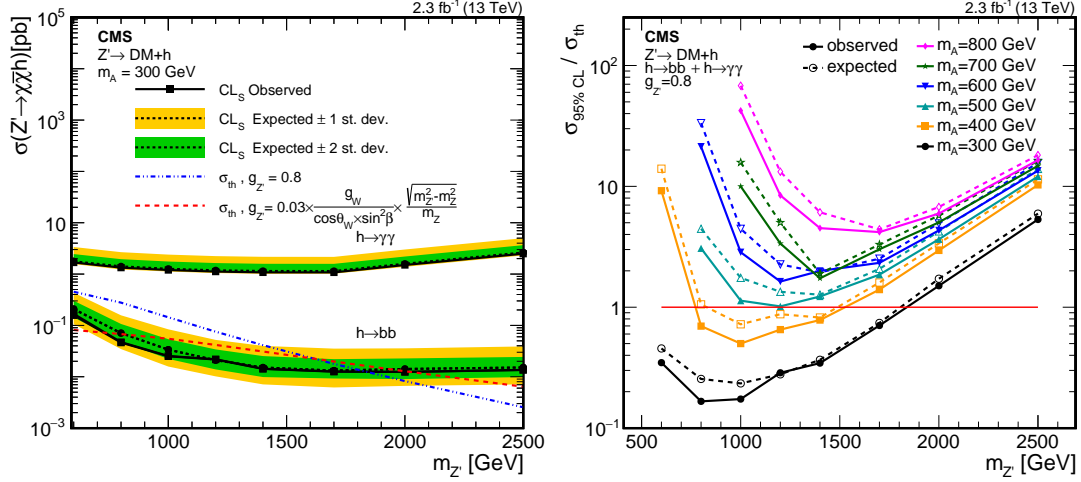

Figure 11: Left: The expected and observed 95% CL limits on dark matter production cross sections for $h \to b\bar{b}$ and $h \to \gamma\gamma$ for $m_A = 300\,\mathrm{GeV}$. The exclusion region is shown for two $g_{Z'}$ values. The dark green and light yellow bands show the 68% and 95% uncertainties on the expected limit. Right: The expected and observed 95% CL limits on the signal strength for $m_A = 300\text{--}800\,\mathrm{GeV}$ are shown. Other parameters for this model are fixed to $m_\chi = 100\,\mathrm{GeV}$ and $\tan\beta = g_\chi = 1$. The theoretical cross section (σ_{th}) used for the right-hand plot is calculated using $g_{Z'} = 0.8$.

6 Search for heavy resonances decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV

This is also one of the major analyses performed by Prof. Yu's group. The master student Ching-Wei Chen was also chosen to give the pre-approval talk on behalf of the whole dihiggs group, which was unusual for a master student. The manpower includes Prof. Shin-Shan Yu, Dr. Vieri Candelise, Dr. Raman Khurana, the master students Ching-Wei Chen and Gregorio Iii Tabbu De Leon. Our contribution includes the following:

- Production of signal simulation
- Study of kinematic distributions for the wide resonances and compare the RS and bulk Gravitions and Radions
- Production of the data and MC ntuples and provided them to the whole analysis group
- Optimization of the b-tagging selection/categories and higgs-mass window
- Performed the bump-hunt analysis
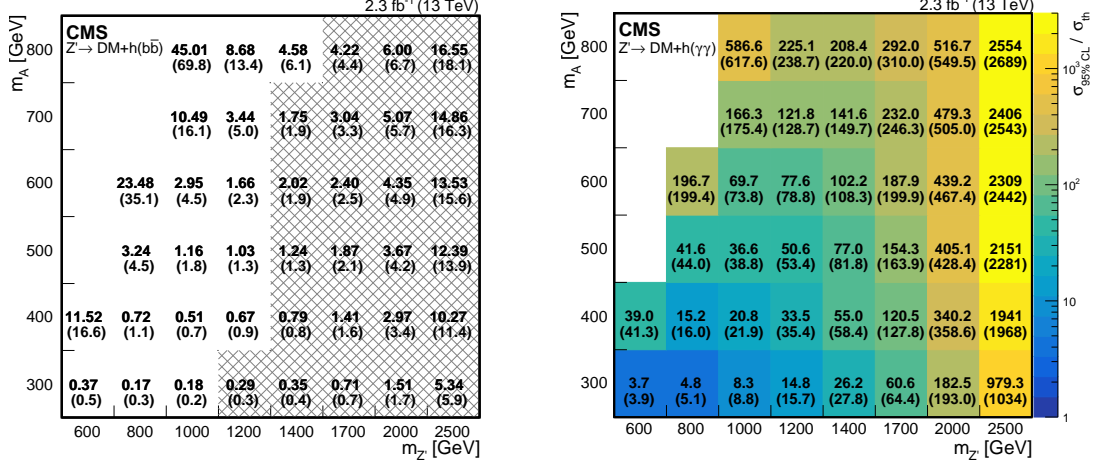
Figure 12: The observed (expected) 95% CL limits on the signal strength (as in Fig. 11 right), separately for the h $\rightarrow$ b$\bar{\text{b}}$ (left) and h $\rightarrow$ $\gamma\gamma$ (right) decay channels, and for m_A = 300–800 GeV and $m_{Z'}$ = 600–2500 GeV. Other parameters for this model are fixed to $m_{\chi=100\,\text{GeV}}$ and $\tan\beta = g_\chi = 1$. The theoretical cross sections are calculated using $g_{Z'}$ = 0.8. For h $\rightarrow$ b$\bar{\text{b}}$, the results for the resolved analysis are shown over a white background, whereas the boosted analysis results are shown over a hatched background.
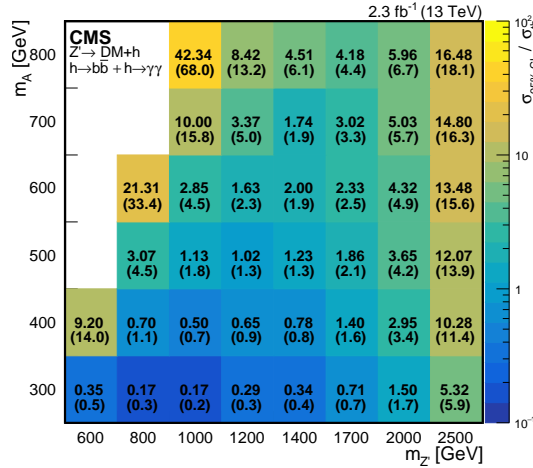

Figure 13: The observed (expected) 95% CL limits on the signal strength (as in Fig. 11 right) for the combination of G and h $\rightarrow$ b$\bar{\text{b}}$ decay channels, and for m_A = 300–800 GeV and $m_{Z'}$ = 600–2500 GeV. Other parameters for this model are fixed to m_χ = 100 GeV and $\tan\beta = g_\chi = 1$. The theoretical cross sections times branching fractions are calculated using $g_{Z'}$ = 0.8.

- Estimated the systematic uncertainties due to Higgs-tagging, PDF, scale, jet energy scale, b-tagging, etc.

The 2015 analysis was first published as a CMS Physics Analysis Summary [68] and then the 2016 results were submitted to Phys. Lett. B [69].

In the standard model (SM), the pair production of Higgs bosons (H) [70–72] in proton-proton collisions at 13 TeV is a rare process [73]. However, the existence of massive resonances decaying to Higgs boson pairs (HH) in many new physics models may enhance this rate to a level observable at the CERN LHC using the current data. For instance, models with warped extra dimensions (WED) [1] contain new particles such as the spin-0 radion [74–76] and the spin-2 first Kaluza–Klein (KK) excitation of the graviton [8, 77, 78], which have sizeable branching fractions to HH.

The WED models have an extra spatial dimension compactified between two branes, with the region between (called the bulk) warped via an exponential metric κl, κ being the warp factor and l the coordinate of the extra spatial dimension [79]. The reduced Planck scale ($\overline{M_{Pl}} \equiv M_{Pl}/8\pi$, M_{Pl} being the Planck scale) is considered a fundamental scale. The benchmarks of the model are $\kappa/\overline{M_{Pl}}$ and the ultraviolet cutoff of the theory $\Lambda_R \equiv \sqrt{6}e^{-\kappa l}\overline{M_{Pl}}$ [74]. In proton-proton (pp) collisions at the LHC, the graviton and the radion are produced primarily through gluon-gluon fusion and are predicted to decay to HH with a branching fraction of 10 and 23%, respectively [80].

We search for a narrow massive resonance decaying to an HH pair, in the $b\bar{b}b\bar{b}$ final state (with a branching fraction $\approx$33% [81]), performed using a data set corresponding to 35.9 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV. The search significantly improves upon the CMS analysis performed using the LHC data collected at $\sqrt{s} = 8$ TeV [82], and extends the searched mass range to 750–3000 GeV. This search is conducted for both the radion and the graviton, whereas the earlier search only considered the former.

In this search, the X $\rightarrow$ HH decay would result in highly Lorentz-boosted and collimated decay products of h $\rightarrow$ b$\bar{b}$, which are referred to as H jets. These are reconstructed using jet substructure and jet flavour-tagging techniques [83–85]. The background consists mostly of SM multijet events, with different characteristics in the search regions below and above the resonance mass $m_X = 1200$ GeV. The background is estimated using several control regions defined in the phase space of the masses and flavour-tagging discriminators of the two H jets, and the HH dijet invariant mass, allowing the background to be predicted over the entire range of m_X explored. The signal would appear as a peak in the HH dijet invariant mass spectrum above a smooth background distribution.

The H $\rightarrow$ b$\bar{b}$ system is reconstructed as a single high-p_T AK8 jet, where the decay products have merged within the jet, and the two highest p_T jets in the event are assumed to be the Higgs boson candidates. The jet is groomed [86] to remove soft and wide-angle radiation using the soft-drop algorithm [87, 88], with the soft radiation fraction parameter set to $z = 0.1$ and the angular exponent parameter to $\beta = 0$. The groomed jet is used to compute the soft-drop jet mass, which peaks at the Higgs boson mass for signal events and reduces the mass of background quark- and gluon-initiated jets. Dedicated mass corrections [89], derived from simulation and data in a region enriched

with $t\bar{t}$ events with merged $W \to q\bar{q}$ decays, are applied to the jet mass in order to remove residual dependence on the jet p_T, and to match the jet mass scale and resolution observed in data.
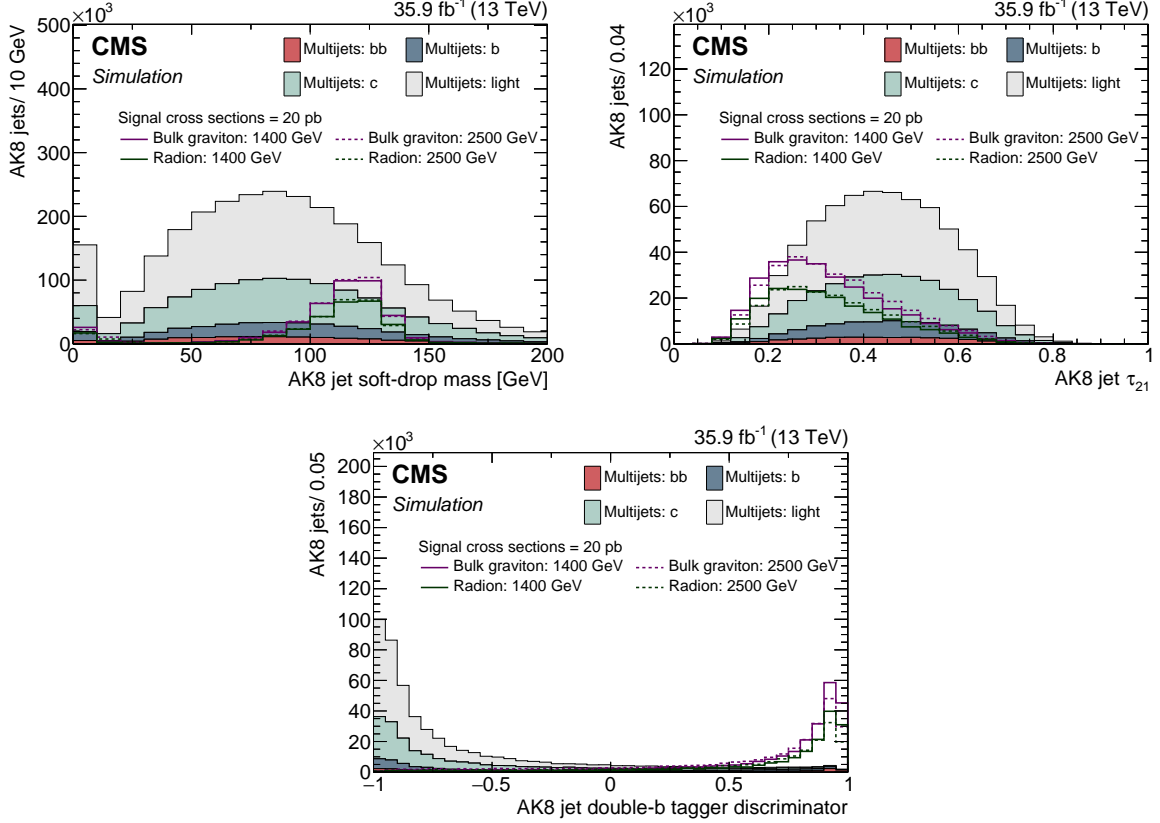

Figure 14: The soft-drop mass (upper left), the N-subjettiness τ_{21} (upper right), and the double-b tagger discriminator (lower) distributions of the selected AK8 jets. The multijet background components for the different jet flavours are shown, along with simulated bulk graviton and radion signals of masses 1400 and 2500 GeV. The number of signal and background events correspond to an integrated luminosity of 35.9 fb^{-1}. A signal cross section of 20 pb is assumed for all the mass hypotheses. The events are required to have passed the trigger selection, lepton rejection, the AK8 jet kinematic selections $p_T > 300$ GeV and $|\eta| < 2.4$, and $|\Delta\eta(j_1, j_2)| < 1.3$. The reduced dijet invariant mass $m_{jj}^{\rm red}$ is required to be greater than 750 GeV. The N-subjettiness requirement of $\tau_{21} < 0.55$ is applied to the upper left and lower figures. The soft-drop masses of the two jets are between 105–135 GeV for the upper right and lower figures.

The soft-drop masses of j_1 and j_2 are required to be within the range 105–135 GeV, resulting in an efficiency of about 60% for each H jet depending on the jet p_T. The "N-subjettiness" algorithm [90] is used to determine the consistency of the jet with two subjets from the two-pronged $h \to b\bar{b}$ decay, by computing the inclusive jet shape variables τ_N. The ratio $\tau_{21} \equiv \tau_2/\tau_1$ with a value much less than one indicates a jet with two subjets. The selection $\tau_{21} < 0.55$ is used, having a jet p_T-dependent efficiency of 50–70%.

For background events, j_1 and j_2 are often well separated in η, especially at high m_{jj}.
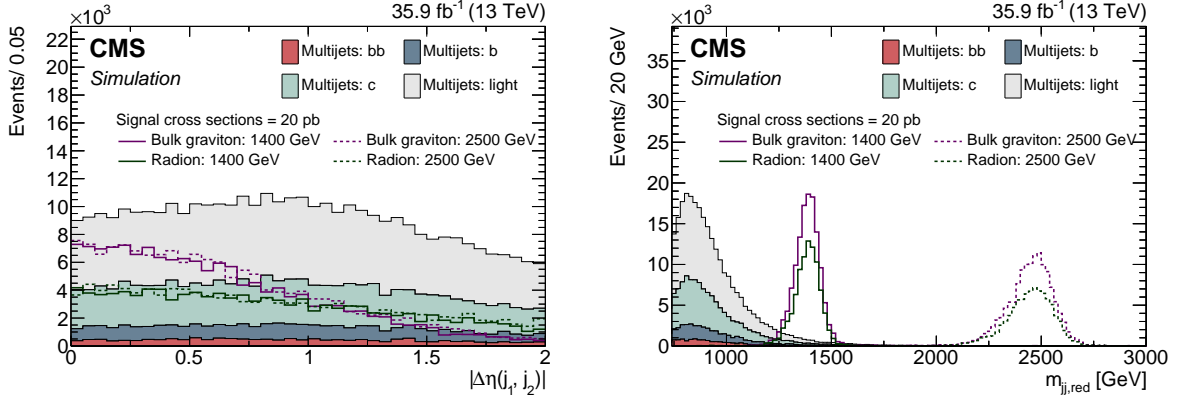
Figure 15: The jet separation $|\Delta\eta(j_1, j_2)|$ (left) and the reduced dijet invariant mass m_{jj}^{red} (right) distributions. The multijet background components for the different jet flavours are shown, along with simulated bulk graviton and radion signals of masses 1400 and 2500 GeV. The numbers of signal and background events correspond to an integrated luminosity of $35.9\,\text{fb}^{-1}$. The signal cross section is assumed to be 20 pb for all the mass hypotheses. The events are required to have passed the online selection, lepton rejection, the AK8 jet kinematic selections $p_T > 300\,\text{GeV}$, $|\eta| < 2.4$. The soft-drop masses of the two jets are between 105 and 135 GeV, and the N-subjettiness requirement of $\tau_{21} < 0.55$ is applied. The $|\Delta\eta(j_1, j_2)|$ distributions (left) require $m_{jj}^{red} < 750\,\text{GeV}$; the m_{jj}^{red} distributions (right) require $|\Delta\eta(j_1, j_2)| < 1.3$.

In contrast, signal events that contain a heavy resonance decaying to two energetic H jets are characterized by a small separation of the two jets. Events are thus required to have $|\Delta\eta(j_1, j_2)| < 1.3$.

The main method to suppress the multijet background is b tagging: since a true $H \to b\bar{b}$ jet contains two b hadrons, the H jet candidates are identified using the dedicated "double-b tagger" algorithm [91]. The double-b tagger exploits the presence of two hadronized b quarks inside the H jet, and uses variables related to b hadron lifetime and mass to distinguish between H jets and the background from multijet production; it also exploits the fact that the b hadron flight directions are strongly correlated with the axes used to calculate the N-subjettiness observables. The double-b tagger is a multivariate discriminator with output between -1 and 1, with a higher value indicating a greater probability for the jet to contain a $b\bar{b}$ pair. The double-b tagger discriminator thresholds of 0.3 and 0.8 correspond to H jet tagging efficiencies of 80 and 30% and are referred to as "loose" (L) and "tight" (T) requirements, respectively. Events must have the two leading p_T AK8 jets satisfying the loose double-b tagger requirement.

The main variable used in the search for a HH resonance is the "reduced dijet invariant mass" $m_{jj}^{red} \equiv m_{jj} - (m_{j_1} - m_H) - (m_{j_2} - m_H)$, where m_{j_1} and m_{j_2} are the soft-drop masses of the leading and subleading H-tagged jets in the event, and $m_H = 125\,\text{GeV}$ [30, 92] is the Higgs boson mass. The quantity m_{jj}^{red} is used rather than m_{jj} since by subtracting the soft-drop masses of the two H-tagged jets and adding back the exact Higgs boson mass m_H, fluctuations in m_{j_1} and m_{j_2} due to the jet mass resolution are corrected, leading to 8–10% improvement in the dijet mass resolution.

The soft-drop mass, τ_{21}, and double-b tagger discriminator distributions of the two leading p_T jets are shown in Fig. 14 for simulated events after passing the online selection, lepton rejection, kinematic selection, and the requirement $m_{jj}^{\mathrm{red}} > 750\,\mathrm{GeV}$. The N-subjettiness requirement of $\tau_{21} < 0.55$ is also applied for the soft-drop mass and the double-b tagger distributions. Since one of the triggers imposes a trimmed jet mass requirement, this affects the shape of the offline soft-drop jet mass, resulting in a steep rise above $\sim$20 GeV. The distributions of the $|\Delta\eta(j_1, j_2)|$ and the m_{jj}^{red} variables are shown in Fig. 15.

The double-b tagger discriminator of the two leading AK8 jets must exceed the loose threshold. In addition, if both discriminator values also exceed the tight threshold, events are classified in the "TT" category. Otherwise, they are classified in the "LL" category, which contains events with both j_1 and j_2 failing the tight threshold as well as events with either j_1 or j_2 passing the tight threshold while the other passes the loose threshold only.
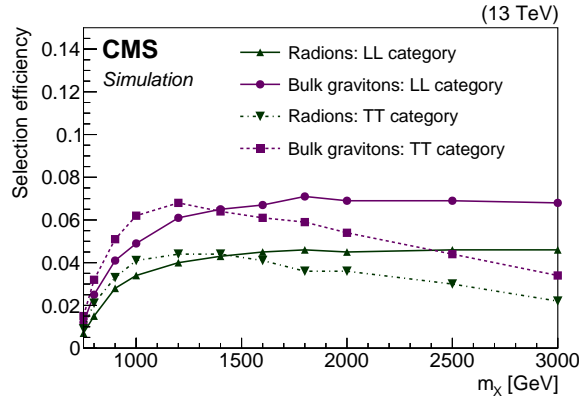

Figure 16: The signal selection efficiencies for the bulk graviton and radion models for different mass hypotheses of the resonances, shown for the LL and the TT signal event categories. Owing to the large sample sizes of the simulated events, the statistical uncertainties are negligible.

The backgrounds are estimated separately for each category, and the combination of the likelihoods for the TT and LL categories gives the best signal sensitivity over a wide range of resonance masses, according to studies performed using simulated signal and multijet samples. The TT category has a good background rejection for m_X up to 2000 GeV. At higher resonance masses, where the background is small, the LL category provides better signal sensitivity. The full event selection efficiencies for bulk gravitons and radions of different assumed masses are shown in Fig. 16. The radion has a smaller efficiency than the bulk graviton because its $|\Delta\eta(j_1, j_2)|$ distribution is considerably wider than that of a bulk graviton of the same mass, as shown in Fig. 15 (left).

The method to determine the background modelling depends on whether the resonance mass m_X is above or below 1200 GeV. The background estimation relies on a set of control regions to predict the total background shape and normalization in the signal regions. The entire range of the m_{jj}^{red} distribution above 750 GeV is used for the prediction.
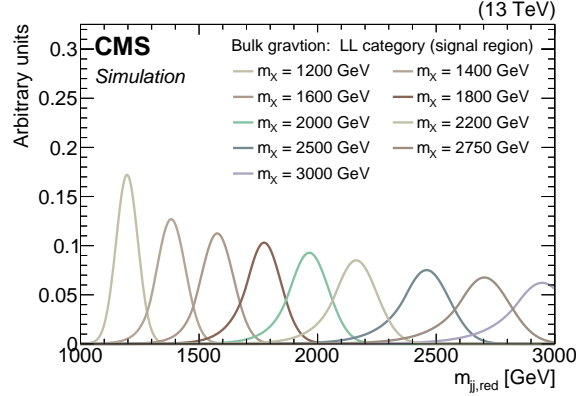
Figure 17: The bulk graviton signal m_{jj}^{red} distribution for the LL category, modelled using the sum of Gaussian and Crystal Ball functions. This modelling is performed for signals in the range $1100 < m_{jj}^{red} < 3000\,\text{GeV}$, where the background distribution falls smoothly. No events are observed above this value of m_{jj}^{red}.

For signals with $m_X \geq 1200\,\text{GeV}$, the underlying background distribution falls monotonically with m_{jj}^{red}. In this region, the background shape is modelled by a smooth function over which a localized signal is searched for. The parameters of the function and its total normalization are constrained by a simultaneous fit of the signal and background models to the data in the control and the signal regions. The m_{jj}^{red} distributions for the signal are modelled using the sum of a Gaussian and a Crystal Ball function [93], as shown in Fig. 17 for one signal category. The modelling is consistent across the different signal categories.
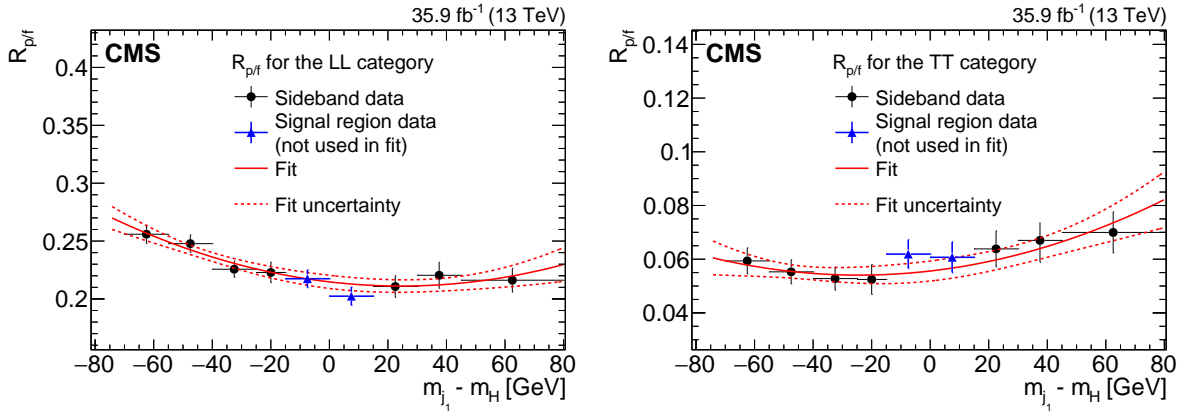

Figure 18: The pass-fail ratio $R_{p/f}$ of the leading p_T jet for the LL (left) and TT (right) signal region categories as a function of the difference between the soft-drop mass of the leading jet and the Higgs boson mass, m_{j_1} - m_H. The measured ratio in different bins of $m_{j_1} - m_H$ is used in the fit (red solid line), except in the region around $m_{j_1} - m_H = 0$, which corresponds to the signal region (blue triangular markers). The fitted function is interpolated to obtain $R_{p/f}$ in the signal region. The horizontal bars on the data points indicate the bin widths.

The signal and control regions are defined by two variables related to the leading p_T jet j_1: (i) its soft-drop mass m_{j_1} and (ii) the value of the discriminator of the double-
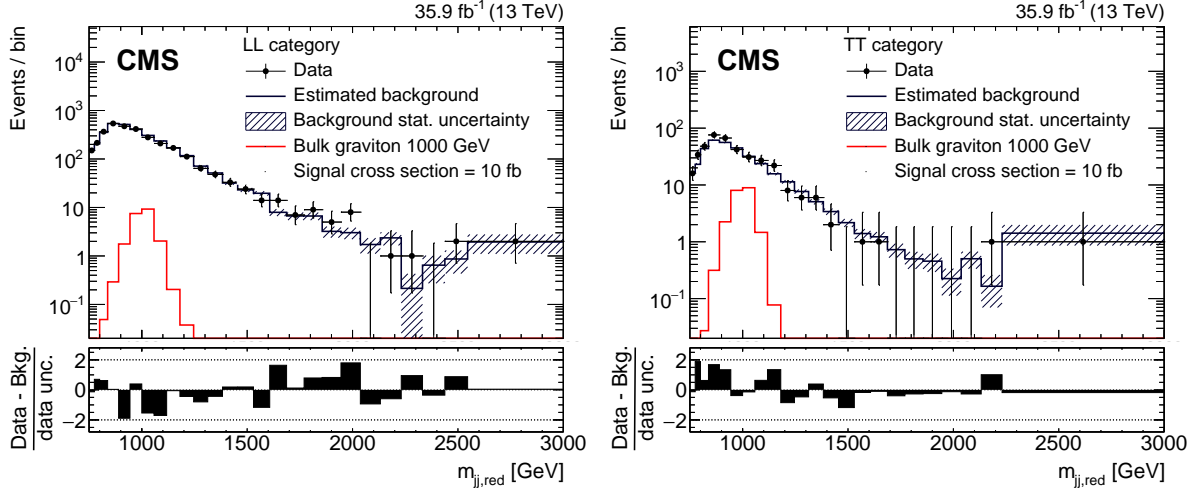
Figure 19: The reduced mass distributions m_{jj}^{red} for the LL (left) and TT (right) signal region categories. The points with bars show the data, the histogram with shaded band shows the estimated background and associated uncertainty. The m_{jj}^{red} spectrum for the background is obtained by weighting the m_{jj}^{red} spectrum in the antitag region by the ratio $R_{p/f}$ of Fig. 18. The signal predictions for a bulk graviton of mass 1000 GeV, are overlaid for comparison, assuming a production cross section of 10 fb. The last bins of the distributions contain all events with $m_{jj}^{\text{red}} > 3000$ GeV. The differences between the data and the predicted background, divided by the data statistical uncertainty (data unc.) as given by the Garwood interval [94], are shown in the lower panels.
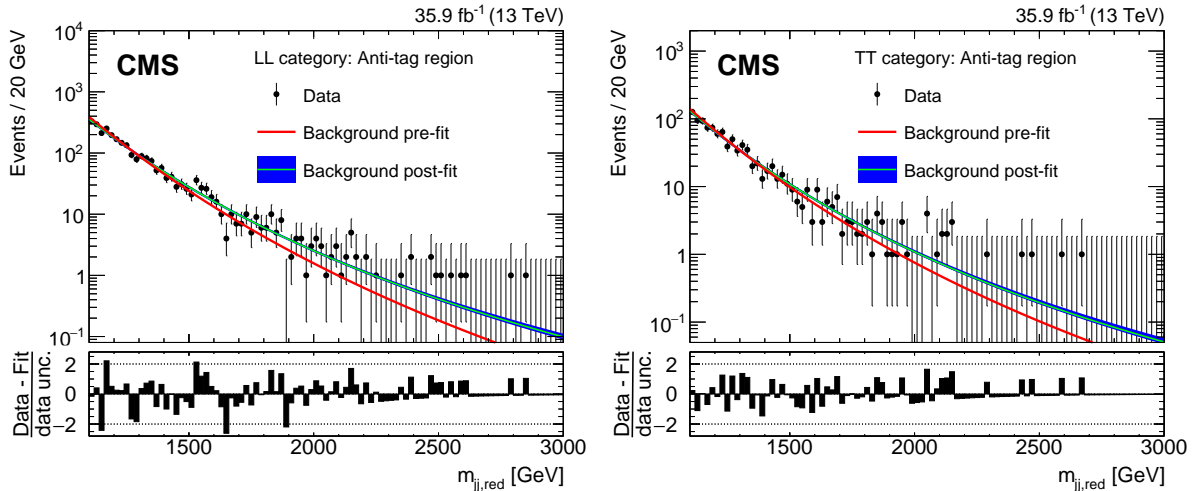

Figure 20: The reduced mass m_{jj}^{red} distributions in the antitag region for the LL (left) and TT (right) categories. The black markers are the data while the curves show the pre-fit and post-fit background shapes. The differences between the data and the predicted background, divided by the statistical uncertainty in the data (data unc.) as given by the Garwood interval [94], are shown in the lower panels.

Table 1: Definition of the signal, the antitag, and the sideband regions used for the background estimation. The regions are defined in terms of the soft-drop masses of the leading p_T H jet (j_1) and the subleading p_T H jet (j_2), and their double-b tagger discriminator values.

Event category	Jet	Soft-drop mass (GeV)	Double-b tagger discriminator
Signal (LL)	j_1		0.3–0.8
	j_2	105–135	
Signal (TT)	j_1		>0.8
	j_2		
Antitag (LL)	j_1		<0.3
	j_2	105–135	0.3–0.8
Antitag (TT)	j_1		<0.3
	j_2		>0.8
Sideband	j_1	<105 or >135	0.3–0.8
(LL, passing)	j_2	105–135	
Sideband	j_1	<105 or >135	>0.8
(TT, passing)	j_2	105–135	
Sideband	j_1	<105 or >135	<0.3
(LL, failing)	j_2	105–135	0.3–0.8
Sideband	j_1	<105 or >135	<0.3
(TT, failing)	j_2	105–135	>0.8

b tagger. The background is estimated in bins of the m_{jj}^{red} distribution. Considering these two variables, several regions are defined.

The *pre-tag* region includes events fulfilling the selection requirements described above apart from those on m_{j_1} and on the j_1 double-b tagger discriminator. The *signal* region is the subset of pre-tag events where m_{j_1} is inside the H jet mass window of 105–135 GeV, and with the j_1 double-b tagger discriminator greater than 0.3 or 0.8, for the LL and TT regions, respectively. The *antitag* regions require the j_1 double-b tagger discriminator to be less than 0.3, with the requirement on j_2 being the same as that for the corresponding LL or TT signal regions. The m_{j_1} *sideband* region consists of events in the pre-tag region, where m_{j_1} lies outside the H jet mass window. Based on whether j_1 passes or fails the double-b tagger discriminator threshold, the sideband region is divided into either "passing" or "failing", respectively. The antitag regions are dominated by the multijet background, and have identical kinematic distributions to the multijet background events in the signal region, according to studies using simulations. The definitions of the signal, the antitag, and the sideband regions are given in Table 1.

In the absence of a correlation between m_{j_1} and the double-b tagger discriminator values, one could measure in the m_{j_1} sideband the ratio of the number of events passing and failing the double-b tagger selection, $R_{\mathrm{p/f}} \equiv N_{\mathrm{pass}} / N_{\mathrm{fail}}$, i.e. the "pass-fail ratio". The yield in the antitag region (in each m_{jj}^{red} bin) could then be scaled by $R_{\mathrm{p/f}}$ to obtain an estimate of the background normalization in the signal region. However, there is a small correlation between the double-b tagger discriminator and m_{j_1}, which is taken
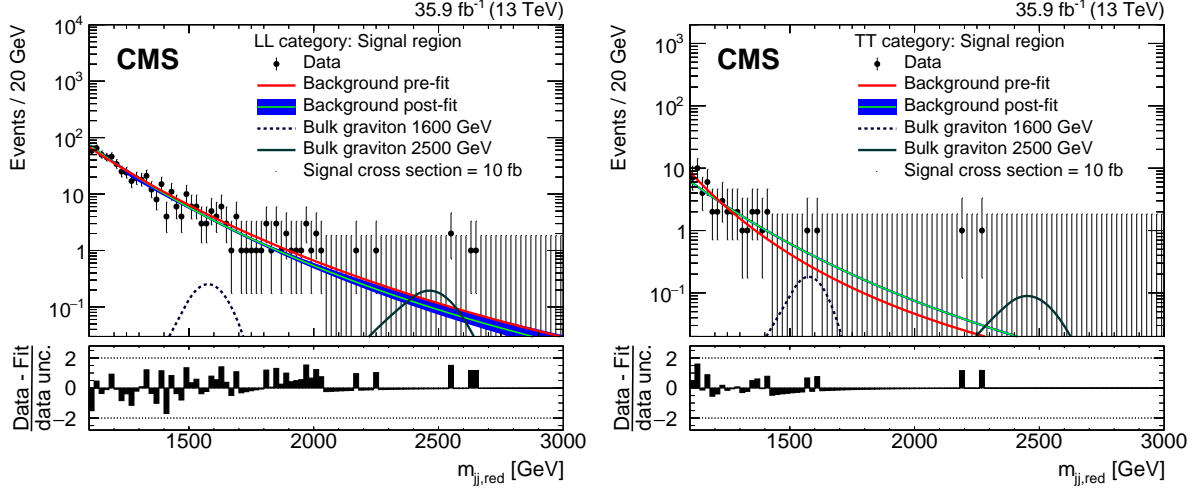
Figure 21: The reduced mass m_{jj}^{red} distributions in the signal region for the LL (left) and the TT (right) categories. The black markers are the data while the curves show the pre-fit and post-fit background shapes. The contribution of bulk gravitons of masses 1600 and 2500 GeV in the signal region are shown assuming a production cross section of 10 fb. The differences between the data and the predicted background, divided by the statistical uncertainty in the data (data unc.) as given by the Garwood interval [94], are shown in the lower panels.

into account by measuring the pass-fail ratio $R_{\text{p/f}}$ as a function of m_{j_1}.

The $R_{\text{p/f}}$ for the LL signal region is measured using ratio of the number of events in the "LL, passing" and "LL, failing" sideband regions, as defined in Table 1. Likewise, the $R_{\text{p/f}}$ for the TT signal region uses the ratio of the number of events in the "TT, passing" to the "TT, failing" sideband regions. The variation of $R_{\text{p/f}}$ as a function of m_{j_1} in each m_{j_1} sideband is fitted with a quadratic function. The fit to the pass-fail ratio is interpolated to the region where m_{j_1} lies within the H jet mass window of 105–135 GeV. An alternative fit using a third order polynomial was found to give the same interpolated value of $R_{\text{p/f}}$ in the Higgs jet mass window. Every event in the antitag region is scaled by the pass-fail ratio evaluated for the m_{j_1} of that event, to obtain the background prediction in the signal region.

Figure 18 shows the quadratic fit in the m_{j_1} sidebands of the pass-fail ratio $R_{\text{p/f}}$ as a function of m_{j_1}, as obtained in the data. The background prediction using this method, along with the number of observed events in the signal region is shown in Fig. 19.

For resonance masses of 1200 GeV and above, the background estimation is improved by simultaneously fitting a parametric model for the background and signal to the data in the signal and the antitag regions. In the fit, the ratio $R_{\text{p/f}}$ obtained from the sidebands is used to constrain the relative number of background events in the two regions. The signal normalization is unconstrained in the fit, while the uncertainties in the parameters of the functions used to model the background and $R_{\text{p/f}}$ are treated as nuisance parameters. The ratio $R_{\text{p/f}}$ is parametrized as a function of m_{jj}^{red}, to account for any possible residual $R_{\text{p/f}}$ dependency on the jet p_{T} at large m_{jj}^{red}. For the

background modelling, a choice among an exponential function $Ne^{-am_{jj}^{red}}$, a "levelled exponential" function $Ne^{-a\,m_{jj}^{red}/(1+a\,b\,m_{jj}^{red})}$, and a "quadratic levelled exponential function" $N\exp[(-a\,m_{jj}^{red}/\{1+a\,b\,m_{jj}^{red}\})-(-c\,m_{jj}^{red2}/\{1+b\,c\,m_{jj}^{red2}\})]$ was made, using a Fisher F-test [95]. At a confidence level of 95%, the levelled exponential function was found to be optimal. Since the background shapes in the antitag and the signal regions were found to be similar according to simulations, the background modelling was tested using the antitag region in the data.

The simultaneous fits to the signal and antitag regions are shown in Figs. 20 and 21, respectively, using the background model only. These are labelled as "post-fit" curves with the signal region background yields constrained to be $R_{p/f}$ times the background yields from the antitag regions. Also shown are the "pre-fit" curves which are obtained by fitting only the antitag regions to the background-only model. The pre-fit curves in the signal region are same as the antitag region pre-fit, scaled by $R_{p/f}$.

Among the four fitted regions, corresponding to the antitag and the signal regions in the LL and TT categories, the events with the highest value of m_{jj}^{red} occur in the antitag region of the LL category, at around $m_{jj}^{red} = 2850$ GeV. As the parametric background model is only reliable within the range of observed events, the likelihood is only evaluated up to $m_{jj}^{red} = 3000$ GeV. This results in a truncation of the signal distribution for resonances having m_X of 2800 GeV and above, with signal efficiency losses increasing to 30% for $m_X = 3000$ GeV, as shown in Fig. 17.

Closure tests of the background estimation methods were performed using simulated multijet samples with signals of various cross sections. The tests indicated a good consistency between the expected and the assumed signal strengths.
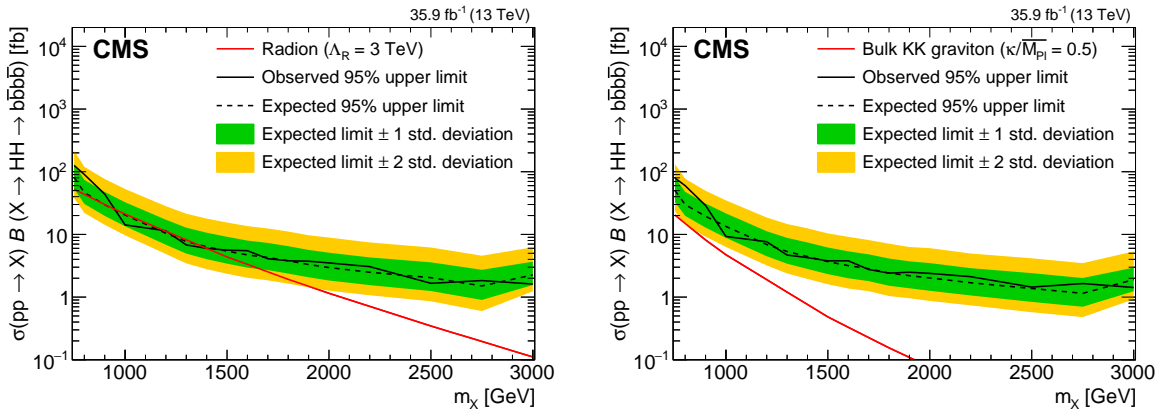

Figure 22: The limits for the spin-0 radion (left) and the spin-2 bulk graviton (right) models. The result for $m_X < 1200$ GeV uses the background predicted using the control regions, while for $m_X \geq 1200$ GeV the background is derived from a combined signal and background fit to the data in the control and the signal regions. The predicted theoretical cross sections for a narrow radion or a bulk graviton are also shown.

As shown in Figs. 19 and 21, for the signal regions, the observed m_{jj}^{red} distribution is consistent with the estimated background. The results are interpreted in terms of upper limits on the product of the production cross sections and the branching frac-

tions $\sigma(\text{pp} \to X)\mathcal{B}(X \to HH \to b\bar{b}b\bar{b})$ for radion and bulk graviton of various mass hypotheses. The asymptotic approximation of the modified frequentist approach for confidence levels, taking the profile likelihood as a test statistic [65, 66, 96], is used. The limits are shown in Fig. 23 for a narrow width radion or a bulk graviton. These are compared with the theoretical values of the product of the cross sections and branching fractions for the benchmarks $\kappa/\overline{M_{Pl}} = 0.5$ and $\Lambda_R = 3\,\text{TeV}$, where the narrow width approximation for signal is valid [80]. The expected limits on the bulk graviton are more stringent than those on the radion because of the higher efficiency of the $|\Delta\eta(j_1, j_2)|$ separation requirement for the former signal.

The upper limits on the production of the cross sections and branching fraction lies in the range 126–1.4 fb for a narrow resonance X of mass $750 < m_X < 3000\,\text{GeV}$. Assuming $\Lambda_R = 3\,\text{TeV}$, a bulk radion with a mass between 970 and 1400 GeV is excluded at 95% confidence level, except in a small region close to 1200 GeV, where the observed limit is 11.4 pb, the theoretical prediction being 11.2 pb.
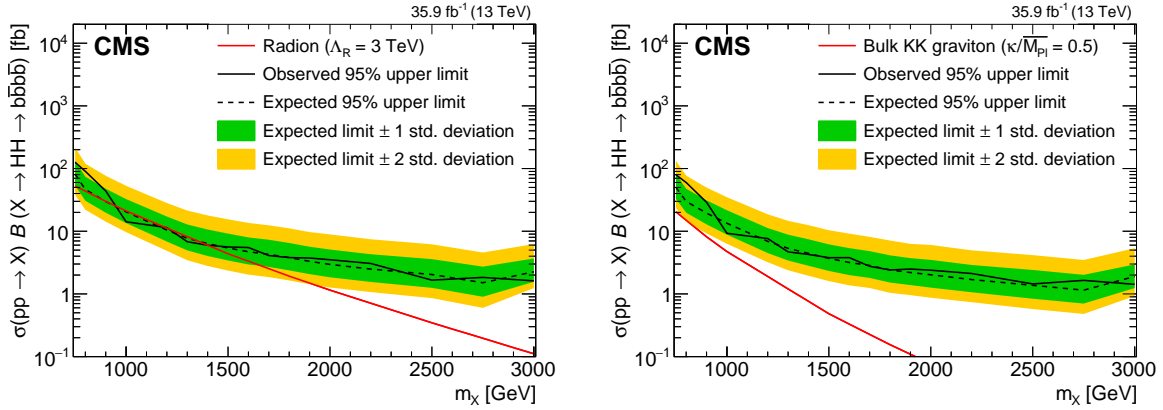

Figure 23: The limits for the spin-0 radion (left) and the spin-2 bulk graviton (right) models. The result for $m_X < 1200\,\text{GeV}$ uses the background predicted using the control regions, while for $m_X \geq 1200\,\text{GeV}$ the background is derived from a combined signal and background fit to the data in the control and the signal regions. The predicted theoretical cross sections for a narrow radion or a bulk graviton are also shown.

7 Work as the CMS contact of the 2HDM+a model in the LHC Dark Matter Working Group

Due to the previous work of mono-h models, Prof. Shin-Shan Yu was chosen as one of the CMS contacts of the new model 2HDM+a in the LHC Dark Matter Working Group. The model we focused on was documented in Ref. [97]. Prof. Yu first performed studies of p_T^{miss} distributions by scanning two model parameters, $\tan\beta$ and $\sin\theta$ and found a significant dependence on $\sin\theta$. See Ref [98] for the scan of $\tan\beta$ when fixing $M_H = 750$ GeV, and $M_A = 500$ GeV, $\sin\theta = 0.1$ and $\sin\theta = \frac{\sqrt{2}}{2}$, respectively; the mass of pseudoscalar a is also varied. See Ref [99] also for the scan of $\sin\theta$ when fixing $\tan\beta = 1$, $M_H = 750$ GeV, and $M_A = 500$ GeV; the mass of pseudoscalar a is also varied. Figures 24–26 show the corresponding widths of Higgs bosons and the cross sections.

Starting from 2017, a new mono-Higgs model 2HDM+a was proposed by the authors in Ref. [100]. ATLAS and CMS have come to a set of mass grids and I am comparing the cross sections with the ATLAS group (University of Heidelberg) for 124 mass points. In our trial of comparison, unlike the Z'-2HDM or Baryonic Z' model, more than one-third of the mass points have more than 1% difference. In addition, 3 mass points have more than 3% difference: (i) $m_A = m_{H,H^\pm} = 1500$ GeV, $m_a = 250$ GeV, (ii) $m_A = m_{H,H^\pm} = 200$ GeV, $m_a = 100$ GeV, and (iii) $m_A = m_{H,H^\pm} = 200$ GeV, $m_a = 150$ GeV. I have tried several different methods of event generation and always obtained consistent results: 10K events in one job, generating gridpack first and then running on gridpacks with 10×1K events, generating gridpack first and then running on gridpacks with 10K events in one job, or using the ATLAS gridpack to generate events with 10 jobs and 1K events each. After some investigation, it turns out that the ATLAS group was running on gridpack with 40 jobs $\times250$ events. Once they generate more events in one job, the cross sections become consistent with my numbers. Our final cross-checks are documented in Ref. [101]. We also discussed about the PDF set to use for final production of full simulation; the consensus was to use 5-flavor PDFs. However, it was not clear if we should use NNPDF 3.0 or NNPDF 3.1. ATLAS MC group does not enforce a use of NNPDF 3.1 however the CMS generator group strongly discourages a production with NNPDF 3.0. The advantage of NNPDF 3.1 is that it uses all LHC data in the PDF fits, including the differential $t\bar{t}p_T$ distributions. In addition, the PDF weights saved in the simulated samples still allow us to evaluate the cross sections with NNPDF 3.0.

In order to prepare for the 2017 MC production, I also learned how to include restrict cards in the model files so that the b quark mass could be set to zero by calling a modified model name "MODEL-no_b_mass".

8 Analysis of the HGC test beam data

Prof. Shin-Shan Yu and Dr. Vieri Candelise participated in the HGC test beam at the CERN H2 area in August and November 2016 and helped to analyze the data to find out the point where the readout electronics is not linear any more. Prof. Yu wrote the code to compute this turning point for both the one-layer and eight-layer HGC runs and now this code is in the official repository and is being used by the Fermlab test beam as well.

The Skiroc2 features dual-gain amplification: "high gain" (HG) for silicon signals up to about 100 MIPs with a good S/N ratio for single MIPs (so is used for the inter-cell calibration); "low gain" (LG) for silicon signals up to about a thousand MIPs but with a poor S/N ratio for single MIPs. In practice, the amplified signals from both HG and LG were digitized and read out. The analysis then used the most appropriate gain dependent on the signal size.

The idea is to analyze data and apply pedestal subtraction first. Once the pedestal is substracted, plot the high-gain ADC counts vs low-gain ADC counts in the two-dimensional histograms, see Fig. 27. One could see two different linear trends in the region with small low-gain and small high-gain ADC counts. We fit these two regions

to two different first-order polynomials. Then, we find out the interception points of these two lines. Prof. Yu had also cross-checked if the slopes and the saturation points are stable as a function of beam spot location, time, and beam energy. Our results will be published with other test beam analysis in a NIM or JINST paper.

In order to evaluate the crossover point (the threshold above which it is advantageous to use the LG instead of HG) we measured the sensor cell signals from the two gains in ADC counts after pedestal and common-mode subtraction on an event-by-event basis. We used electrons at various energies to explore the full dynamic range: E_e= [20, 70, 100, 150, 200, 250] GeV and for the two Skiroc2 chips separately per module. The output of the study was a set of two-dimensional histograms showing the expected linear trend between the two gains and the crossover point around the position [HG ; LG] = [1800 ; 200] ADC counts, above which the saturation of the HG is seen. In order to have a precise measurement of the relative slope, defined as the ratio between the HG and LG at the crossover point, we performed linear fits in two different LG ranges: [0-150] and [250-400] ADC counts. The HG saturation point was extracted by interpolating the two lines and finding the intersection point. The error on the intersection point was then evaluated by propagating the fit uncertainties accordingly.

An example of the HG-LG correlation plot for $E_e = 250$ GeV electrons is shown in Fig. 27. The LG and the HG cut off, extracted from the fit of the HG-LG correlation plot for each layer are presented in Table 2 for the first Skiroc2 as an example. Consistent results were found for the second Skiroc2 chip. For each layer and Skiroc2 we also calculated the gain ratio, defined as the ratio between the HG cut off and the LG cut off (HG/LG), which was used to normalize the signal response to the same gain. Results on the gain ratio for the first Skiroc2 are reported in Table 2. As can be seen from the table, no significant trends nor discrepancie s are found for the gain ratio results through the 8 layers. Also, no dependencies of the gain ratio as a function of the beam energies, runs nor Skiroc2 chips have been observed.

Layer	LG cut off	HG cut off	Gain Ratio (HG/LG)
1	186.2 ± 0.2	1897.1 ± 1.1	10.0 ± 0.4
2	187.1 ± 1.7	1872.5 ± 16.3	10.0 ± 0.4
3	204.9 ± 0.9	1990.3 ± 7.4	9.8 ± 1.6
4	196.9 ± 1.0	1915.0 ± 8.4	9.7 ± 1.7
5	196.7 ± 2.0	1910.6 ± 18.4	9.7 ± 2.2
6	193.0 ± 3.2	1893.4 ± 30.8	9.8 ± 2.2
7	199.1 ± 1.4	1957.9 ± 13.0	9.8 ± 1.6
8	226.3 ± 0.8	2085.0 ± 5.4	9.2 ± 0.2

Table 2: High Gain (HG) and Low Gain (LG) ADC counts values and the fit results on the gain ratio defined as the High-Low Gain ratio, evaluated for each of the 8 layers for the first Skiroc2 chip in the CERN setups.
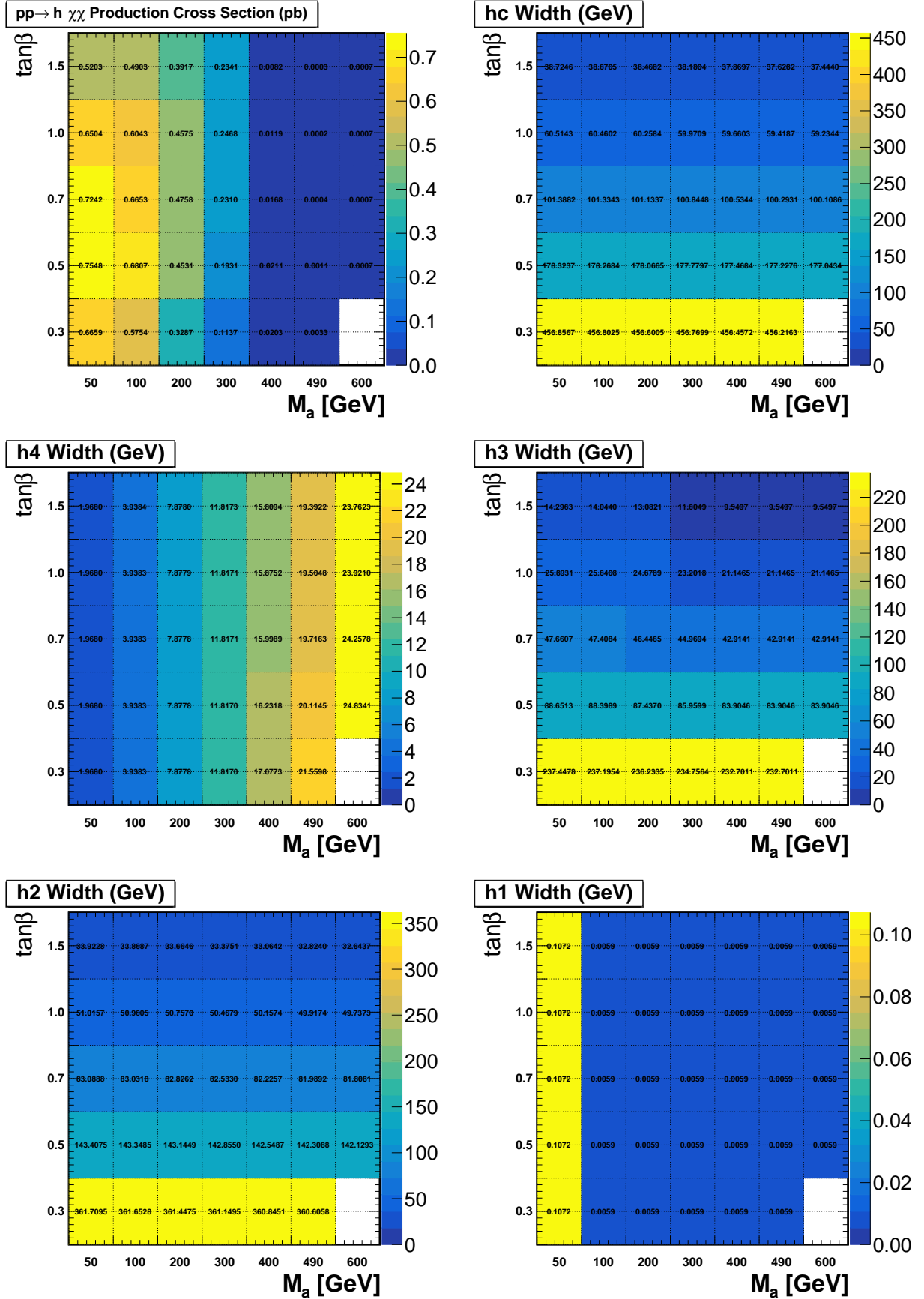
Figure 24: The cross sections and the widths of all Higgs bosons when scanning the $\tan\beta$ and M_a and fixing $\sin\theta = 0.1$.
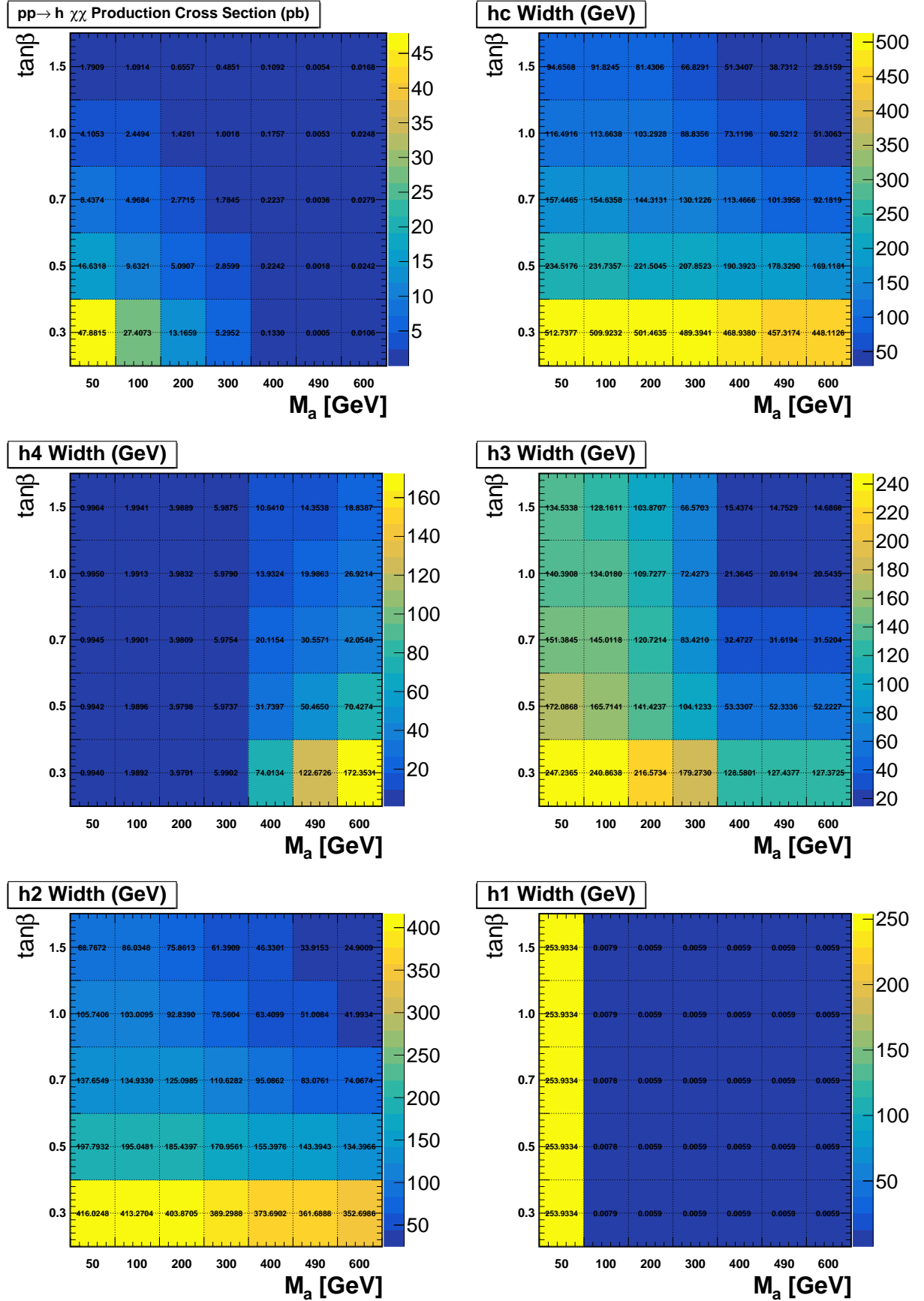
Figure 25: The cross sections and the widths of all Higgs bosons when scanning the $\tan\beta$ and M_a and fixing $\sin\theta = \frac{\sqrt{2}}{2}$.
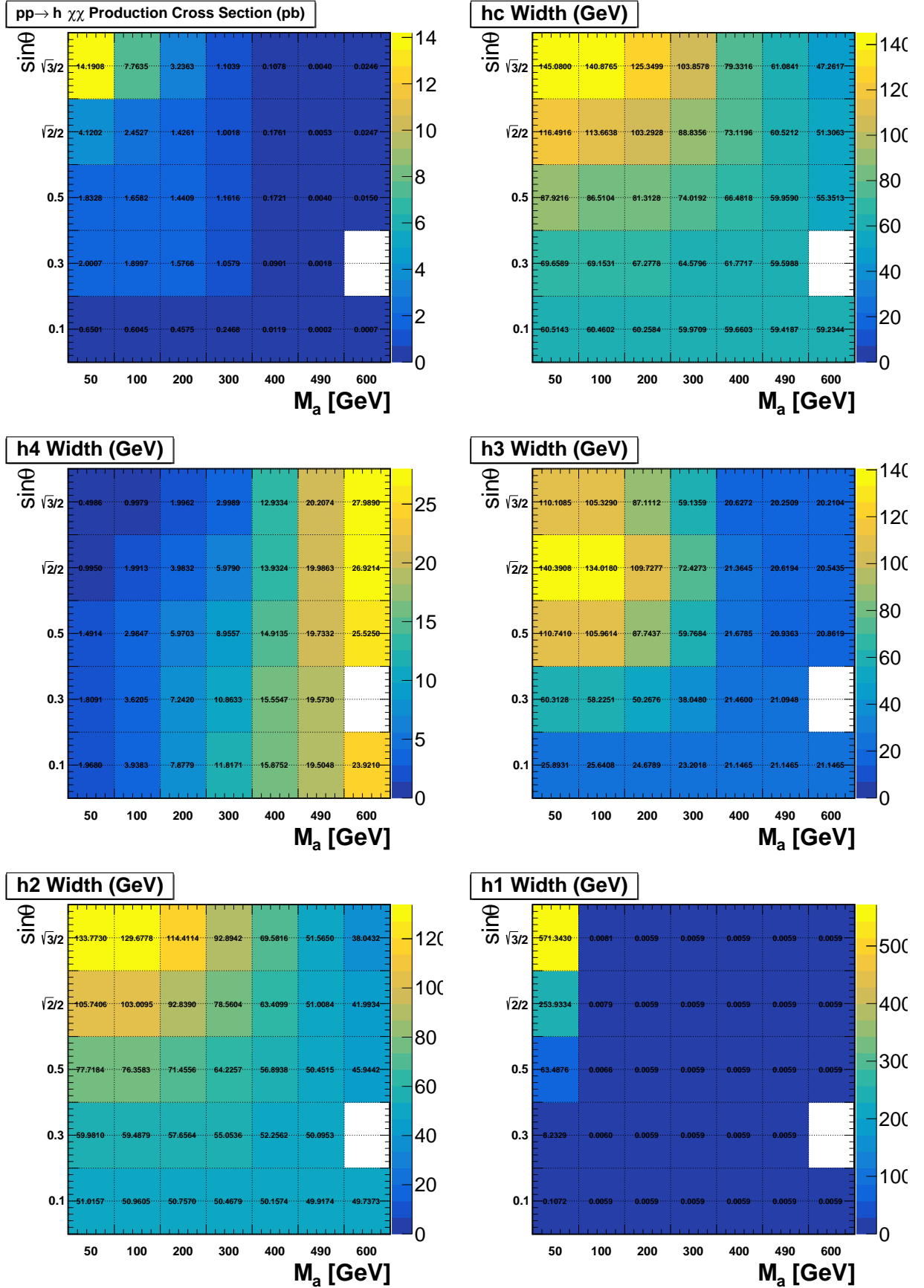
Figure 26: The cross sections and the widths of all Higgs bosons when scanning the sin θ and M_a and fixing tan β = 1.

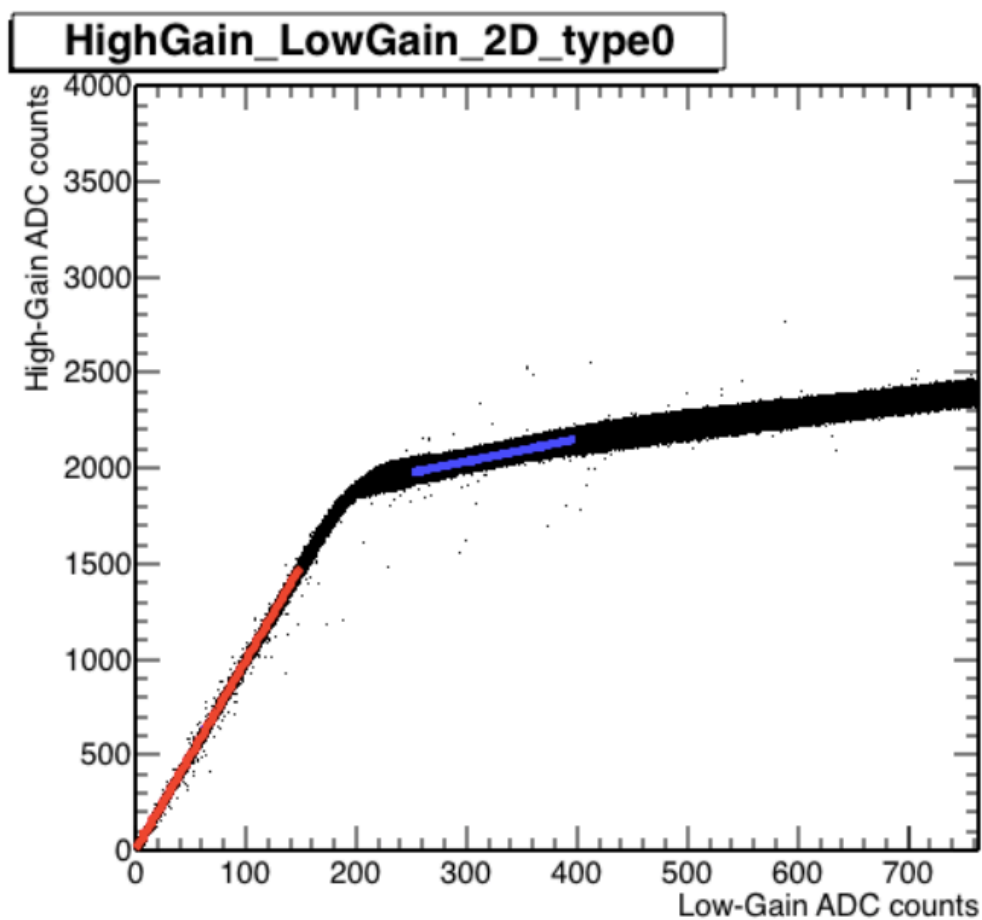Figure 27: The High-Low Gain correlation for an incoming electron beam of energy $E_e = 250\,\text{GeV}$.

References

[1] 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`.

[2] L. Randall and R. Sundrum, "An Alternative to Compactification", *Phys. Rev. Lett.* **83** (1999) 4690, `doi:10.1103/PhysRevLett.83.4690`, `arXiv:hep-th/9906064`.

[3] B. Bellazzini, C. Csaki, and J. Serra, "Composite Higgses", *Eur. Phys. J. C* **74** (2014) 2766, `doi:10.1140/epjc/s10052-014-2766-x`, `arXiv:1401.2457`.

[4] R. Contino, D. Marzocca, D. Pappadopulo, and R. Rattazzi, "On the effect of resonances in composite Higgs phenomenology", *JHEP* **10** (2011) 081, `doi:10.1007/JHEP10(2011)081`, `arXiv:1109.1570`.

[5] D. Marzocca, M. Serone, and J. Shu, "General composite Higgs models", *JHEP* **08** (2012) 013, `doi:10.1007/JHEP08(2012)013`, `arXiv:1205.0770`.

[6] D. Greco and D. Liu, "Hunting composite vector resonances at the LHC: naturalness facing data", *JHEP* **12** (2014) 126, `doi:10.1007/JHEP12(2014)126`, `arXiv:1410.2883`.

[7] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, "Heavy Vector Triplets: Bridging Theory and Data", *JHEP* **09** (2014) 060, `doi:10.1007/JHEP09(2014)060`, `arXiv:1402.4431`.

[8] 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`.

[9] 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`.

[10] O. Antipin, D. Atwood, and A. Soni, "Search for RS gravitons via $W_L W_L$ decays", *Phys. Lett. B* **666** (2008) 155, `doi:10.1016/j.physletb.2008.07.009`, `arXiv:0711.3175`.

[11] C. Grojean, E. Salvioni, and R. Torre, "A weakly constrained W′ at the early LHC", *JHEP* **07** (2011) 002, `doi:10.1007/JHEP07(2011)002`, `arXiv:1103.2761`.

[12] P. Langacker, "The physics of heavy Z′ gauge bosons", *Rev. Mod. Phys.* **81** (2009) 1199, `doi:10.1103/RevModPhys.81.1199`, `arXiv:0801.1345`.

[13] E. Salvioni, G. Villadoro, and F. Zwirner, "Minimal Z′ models: present bounds and early LHC reach", *JHEP* **11** (2009) 068, `doi:10.1088/1126-6708/2009/11/068`, `arXiv:0909.1320`.

[14] CMS Collaboration, "V Tagging Observables and Correlations", CMS Physics Analysis Summary CMS-PAS-JME-14-002, CERN, 2014.

[15] CMS Collaboration, "Identification techniques for highly boosted W bosons that decay into hadrons", *JHEP* **12** (2014) 017, `doi:10.1007/JHEP12(2014)017`, `arXiv:1410.4227`.

[16] CMS Collaboration, "Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at $\sqrt{s} = 8$ TeV", *JHEP* **08** (2014) 174, `doi:10.1007/JHEP08(2014)174`, `arXiv:1405.3447`.

[17] CMS Collaboration, "Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* **08** (2014) 173, `doi:10.1007/JHEP08(2014)173`, `arXiv:1405.1994`.

[18] CMS Collaboration, "Studies of jet mass in dijet and W/Z + jet events", *JHEP* **05** (2013) 090, `doi:10.1007/JHEP05(2013)090`, `arXiv:1303.4811`.

[19] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure", *Phys. Rev. D* **80** (2009) 051501, `doi:10.1103/PhysRevD.80.051501`, `arXiv:0903.5081`.

[20] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches", *Phys. Rev. D* **81** (2010) 094023, `doi:10.1103/PhysRevD.81.094023`, `arXiv:0912.0033`.

[21] D. Krohn, J. Thaler, and L.-T. Wang, "Jet trimming", *JHEP* **02** (2010) 084, `doi:10.1007/JHEP02(2010)084`, `arXiv:0912.1342`.

[22] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions", *Nucl. Phys. B* **406** (1993) 187–224, `doi:10.1016/0550-3213(93)90166-M`.

[23] 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`.

[24] 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`.

[25] M. Bähr et al., "Herwig++ physics and manual", *Eur. Phys. J. C* **58** (2008) 639, `doi:10.1140/epjc/s10052-008-0798-9`, `arXiv:0803.0883`.

[26] 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`.

[27] 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`].

[28] 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`.

[29] 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`.

[30] ATLAS and CMS Collaborations, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", *Phys. Rev. Lett.* **114** (2015) 191803, `doi:10.1103/PhysRevLett.114.191803`, `arXiv:1503.07589`.

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

[32] 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`.

[33] 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`.

[34] ATLAS Collaboration, "Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector", *Phys. Rev. D* **90** (2014) 052004, `doi:10.1103/PhysRevD.90.052004`, `arXiv:1406.3827`.

[35] CMS Collaboration Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", `arXiv:1412.8662`.

[36] CMS Collaboration, "Evidence for the direct decay of the 125 GeV Higgs boson to fermions", *Nat. Phys.* **10** (2014) 557, `doi:10.1038/nphys3005`, `arXiv:1401.6527`.

[37] R. Barbieri and G. F. Giudice, "Upper bounds on supersymmetric particle masses", *Nuc. Phys. B* **306** (August, 1988) 63, `doi:10.1016/0550-3213(88)90171-X`.

[38] V. D. Barger, W.-Y. Keung, and E. Ma, "A gauge model with light *W* and *Z* bosons", *Phys. Rev. D* **22** (1980) 727, `doi:10.1103/PhysRevD.22.727`.

[39] E. Salvioni, G. Villadoro, and F. Zwirner, "Minimal Z' models: present bounds and early LHC reach", *JHEP* **09** (2009) 068, `doi:10.1088/1126-6708/2009/11/068`, `arXiv:0909.1320`.

[40] R. Contino, D. Pappadopulo, D. Marzocca, and R. Rattazzi, "On the effect of resonances in composite Higgs phenomenology", *JHEP* **10** (2011) 081, `doi:10.1007/JHEP10(2011)081`.

[41] D. Marzocca, M. Serone, and J. Shu, "General composite Higgs models", *JHEP* **08** (2012) 13, `doi:10.1007/JHEP08(2012)013`.

[42] T. Han, H. E. Logan, B. McElrath, and L.-T. Wang, "Phenomenology of the little Higgs model", *Phys. Rev. D* **67** (2003) 095004, `doi:10.1103/PhysRevD.67.095004`, `arXiv:hep-ph/0301040`.

[43] M. Schmaltz and D. Tucker-Smith, "Little Higgs Theories", *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229, `doi:10.1146/annurev.nucl.55.090704.151502`.

[44] M. Perelstein, "Little Higgs models and their phenomenology", *Prog. Part. Nucl. Phys.* **58** (2007) 247, `doi:10.1016/j.ppnp.2006.04.001`, `arXiv:hep-ph/0512128`.

[45] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, "Heavy vector triplets: bridging theory and data", *JHEP* **09** (2014) 60, `doi:10.1007/JHEP09(2014)060`, `arXiv:1402.4431`.

[46] CMS Collaboration, "Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in $\ell^+\ell^- b\bar{b}$ final states", *Phys. Lett. B* **748** (2015) 221, `doi:10.1016/j.physletb.2015.07.010`, `arXiv:1504.04710`.

[47] 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`.

[48] 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`.

[49] 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) 1, `doi:10.1140/epjc/s10052-016-4067-z`, `arXiv:1601.06431`.

[50] CMS Collaboration, "Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s} = 13$ TeV", `arXiv:1703.05236`.

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

[52] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection", *Annual Review of Astronomy and Astrophysics* **48** (2010) 495, `doi:10.1146/annurev-astro-082708-101659`, `arXiv:https://doi.org/10.1146/annurev-astro-082708-101659`.

[53] T. A. Porter, R. P. Johnson, and P. W. Graham, "Dark Matter Searches with Astroparticle Data", *Annual Review of Astronomy and Astrophysics* **49** (2011) 155, `doi:10.1146/annurev-astro-081710-102528`, `arXiv:https://doi.org/10.1146/annurev-astro-081710-102528`.

[54] Planck Collaboration, "Planck 2015 results - XIII. Cosmological parameters", *AA* **594** (2016) A13, `doi:10.1051/0004-6361/201525830`.

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

[56] 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`.

[57] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **06** (2013) 81, `doi:10.1007/JHEP06(2013)081`.

[58] A. Berlin, T. Lin, and L.-T. Wang, "Mono-Higgs Detection of Dark Matter at the LHC", *JHEP* **06** (2014) 078, `doi:10.1007/JHEP06(2014)078`, `arXiv:1402.7074`.

[59] T. D. Lee, "A Theory of Spontaneous T Violation", *Phys. Rev.* **D 8** (1973) 1226, `doi:10.1103/PhysRevD.8.1226`.

[60] G. Branco et al., "Theory and phenomenology of two-Higgs-doublet models", *Phys. Rept.* **516** (2012) 01, `doi:10.1016/j.physrep.2012.02.002`, `arXiv:1106.0034`.

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

[62] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s}$ = 7 TeV at the LHC", *JINST* **8** (2013) P11002, `doi:10.1088/1748-0221/8/11/P11002`, `arXiv:1306.6905`.

[63] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS", *JINST* **11** (2016) P01019, `doi:10.1088/1748-0221/11/01/P01019`, `arXiv:1510.07488`.

[64] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group", Technical Report CERN-2013-004, Geneva, 2013. `doi:10.5170/CERN-2013-004`, `arXiv:1307.1347`.

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

[66] 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`.

[67] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.* **C71** (2011) 1554, `doi:10.1140/epjc/s10052-011-1554-0`.

[68] CMS Collaboration, "Search for heavy resonances decaying to a pair of Higgs bosons in four b quark final state in proton-proton collisions at sqrt(s)=13 TeV", technical report.

[69] 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", `arXiv:1710.04960`.

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

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

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

[73] D. de Florian and J. Mazzitelli, "Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD", *Phys. Rev. Lett.* **111** (2013) 201801, `doi:10.1103/PhysRevLett.111.201801`, `arXiv:1309.6594`.

[74] 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`.

[75] 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`.

[76] 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`.

[77] 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`.

[78] 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`.

[79] G. F. Giudice, R. Rattazzi, and J. D. Wells, "Graviscalars from higher dimensional metrics and curvature Higgs mixing", *Nucl. Phys. B* **595** (2001) 250, `doi:10.1016/S0550-3213(00)00686-6`, `arXiv:hep-ph/0002178`.

[80] A. Oliveira, "Gravity particles from Warped Extra Dimensions, a review. Part I - KK Graviton", (2014). `arXiv:1404.0102`.

[81] D. de Florian et al., "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", CERN Report CERN-2017-002-M, CERN, 2016. `doi:10.23731/CYRM-2017-002`, `arXiv:1610.07922`.

[82] CMS Collaboration, "Search for heavy resonances decaying to two Higgs bosons in final states containing four b quarks", *Eur. Phys. J. C* **76** (2016) 371, `doi:10.1140/epjc/s10052-016-4206-6`, `arXiv:1602.08762`.

[83] 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`.

[84] B. Cooper, N. Konstantinidis, L. Lambourne, and D. Wardrope, "Boosted $hh \to b\bar{b}b\bar{b}$: a new topology in searches for TeV-scale resonances at the LHC", *Phys. Rev. D* **88** (2013) 114005, `doi:10.1103/PhysRevD.88.114005`, `arXiv:1307.0407`.

[85] 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`.

[86] G. P. Salam, "Towards jetography", *Eur. Phys. J. C* **67** (2010) 637, `doi:10.1140/epjc/s10052-010-1314-6`, `arXiv:0906.1833`.

[87] 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`.

[88] 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`.

[89] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, CERN, 2017.

[90] J. Thaler and K. Van Tilburg, "Maximizing boosted top identification by minimizing N-subjettiness", *JHEP* **02** (2012) 093, `doi:10.1007/JHEP02(2012)093`, `arXiv:1108.2701`.

[91] CMS Collaboration, "Identification of double-b quark jets in Boosted Topologies", CMS Physics Analysis Summary CMS-PAS-BTV-15-002, CERN, 2016.

[92] CMS Collaboration, "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at sqrt(s) = 13 TeV", (2017). `arXiv:1706.09936`. Submitted to JHEP.

[93] M. Oreglia, "A Study of the Reactions $\psi' \to \gamma\gamma\psi$". PhD thesis, SLAC, 1980.

[94] F. Garwood, "(i) Fiducial Limits for the Poisson Distribution", *Biometrika* **28** (1936) 437, `doi:10.1093/biomet/28.3-4.437`.

[95] R. G. Lomax and D. L. Hahs-Vaughn, "Statistical concepts: a second course". Taylor and Francis, Hoboken, NJ, 2012.

[96] 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`].

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

[98] Shin-Shan Yu, "Scan of $\tan\beta$ in the 2HDM+a model", github, 2017.

[99] Shin-Shan Yu, "Scan of $\sin\theta$ in the 2HDM+a model", github, 2017.

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

[101] Shin-Shan Yu, "Cross-check of ATLAS and CMS 2HDM+a cross sections", gitlab, 2017.