CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2016/03/23

Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 13 TeV

The CMS Collaboration

Abstract

A model-independent search for a narrow-width resonance decaying into two Higgs bosons, each having a mass of 125 GeV and decaying into a $b\bar{b}$ pair, is presented. The search is performed using proton-proton collision data corresponding to an integrated luminosity of 2.3 fb^{-1} at $\sqrt{s} = 13$ TeV recorded by the CMS detector at the LHC. No evidence for a signal is observed and upper limits at a 95% confidence level on the production cross section for such a resonance, in the mass range from 260 to 1200 GeV, are set.

1 Introduction

The discovery of a Higgs boson (H) with mass around 125 GeV and properties consistent with the standard model (SM) of particle physics at the Large Hadron Collider (LHC) [1, 2], opens new possibilities to search for higher mass resonances via their decays into Higgs bosons. Several well-motivated hypotheses of physics beyond the standard model posit narrow-width resonances decaying into pairs of Higgs bosons. One such resonance is the first Kaluza Klein (KK) excitation of the graviton in Randall-Sundrum (RS1) models of Warped Extra Dimensions [3]. This paper reports the results of a search for such a resonance in the 260-1200 GeV mass range, with both Higgs bosons decaying into bottom quarks. This search is performed using 2.3 fb^{-1} proton-proton collision data collected at $\sqrt{s}$ = 13 TeV by the Compact Muon Solenoid (CMS) detector at the LHC. The main challenge of this search is to discriminate the signature of four bottom quarks in the final state that hadronize into jets from the overwhelming multi-jet quantum chromodynamic (QCD) background. This is addressed by suitable event selection criteria that include b-jet identification techniques and by a model of the multi-jet background that is tested in control regions of data.

The results of this analysis at 13 TeV complement those for similar searches in the four bottom quarks final state conducted by the ATLAS [4] and CMS [5] experiments at 8 TeV.

2 CMS Detector

The central feature of the CMS apparatus is a superconducting solenoid with an internal diameter of 6 m that generates a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter. Muons are measured in gas-ionization detectors embedded in a steel flux return yoke outside the solenoid. The pixel tracker provides an impact parameter resolution for charged tracks of about 15 μm. This is essential for properly reconstructing secondary vertices used for the b-jet identification techniques used in this analysis. The first level of the CMS trigger system, consisting of custom hardware processors, uses information from the calorimeters to select the events for this analysis. The second level of the CMS trigger or the High Level Trigger (HLT), consisting of generic PC processor farms, further selects events using information from the calorimeters and trackers before sending them downstream for detailed processing and storage. Particles produced in the pp collisions are detected in the pseudorapidity range $|\eta| < 5$, where $\eta = -\ln\tan(\theta/2)$ and θ is the polar angle with respect to the direction of the proton beam. A more detailed description of the CMS detector can be found elsewhere [6].

3 Data and Simulated Samples

The final state of this analysis consists of four jets originating from the hadronization of b quarks, also called b-jets. To maximize the signal efficiency while reducing the overwhelming rate of QCD multi-jet events, two different trigger selections were combined, both exploiting the Combined Secondary Vertex (CSV) algorithm [7] to identify b-jets, that relies on quantities measured with the pixel detector. For the first trigger selection, four jets with $p_{\text{T}} > 30$ GeV are required. Of those four jets two are required to have $p_{\text{T}} > 90$ GeV and three jets are required to be tagged as b-jets. In the second trigger selection, for the high level trigger, four jets with $p_{\text{T}} > 45$ GeV are required and of those, three jets have to be identified as b-jets. The logical "OR" between these two selections is used to collect data corresponding to an integrated luminosity

of 2.3 fb^{-1} at $\sqrt{s} = 13$ TeV.

The production of a narrow-width resonance is simulated using the leading order (LO) MADGRAPH 5.1 [8] Monte Carlo generator. The spin-2 RS1 KK-Graviton is forced to decay to a pair of Higgs bosons with parameters reported in [9], where both Higgs bosons decay to $\text{b}\bar{\text{b}}$. The PYTHIA parameters for the underlying event are set to the Z2Star tune [10]. The response of the CMS detector is modeled using GEANT4 [11].

4 Event Reconstruction

Jets are reconstructed from particle-flow candidates using the anti-k_T clustering algorithm [12], with a distance parameter of 0.4, as implemented in the FASTJET package [13, 14]. Jet energy corrections, as a function of pseudorapidity and transverse momentum of the jet, are applied [15]. Jet identification criteria are also applied to reject fake jets from detector noise and jets originating from primary vertices not associated with the hard interaction [16].

Techniques for identifying b-jets are the core of this analysis. Bottom quarks hadronize into b-hadrons that decay only through the weak interaction. Thus, they have lifetimes of the order of $c\tau = 450$ μm. Daughter particles so can have a sizable impact parameter with respect to the B hadron point of origin and result in reconstructed secondary vertices. Measurements of the properties of such secondary vertices, lifetimes of the b-hadron, and low-p_{T} lepton information, when available, are used by the Combined MultiVAriate algorithm (CMVA) [17] for b-tagging events. This algorithm determines secondary vertices using the Inclusive Vertex Finder algorithm [18]. The CMVA algorithm outputs a continuous discriminant between -1 and 1. The chosen operating point corresponds to a 72% efficiency for tagging jets originating from b hadrons and a mistagging rate of 1% for light-flavor jets.

Simulated events are weighted to match the number of primary vertices per event in data. These events are also weighted to account for differences in b-tagging efficiency between data and simulation. The trigger efficiencies for signal are evaluated by passing generated events through a trigger simulation.

5 Analysis Strategy

The search for a narrow-width $\text{X} \to \text{H}(\text{b}\bar{\text{b}})\text{H}(\text{b}\bar{\text{b}})$ resonance is performed between masses of 260 GeV and 1200 GeV. The momenta and angles of the decay products of such a resonance change substantially over this range and in order to maximize the sensitivity of this search, different event selection criteria are used for the two main kinematic regions: the low-mass region (LMR) for resonance mass hypothesis from 260 GeV to 400 GeV, and the medium-mass region (MMR) for masses from 400 GeV to 1200 GeV. The 400 GeV transition threshold between the LMR and the MMR has been chosen to maximize the expected sensitivity. Above 900 GeV the Higgs bosons have a momentum considerably higher than their mass. Thus, each Higgs to $\text{b}\bar{\text{b}}$ decay would be reconstructed more efficiently as one hadronic jet with a larger anti-k_T distance parameter (0.8). The mass range above 1200 GeV (high-mass region) is out of the scope and not covered in this search.

The QCD multi-jet background is modeled in data by studying parametric fits in sideband regions as illustrated in Fig. 1 and described below. Event selection begins with identifying events containing at least four b-tagged jets with $p_{\text{T}} > 30$ GeV and $|\eta| < 2.5$. Among these jets, two pairs are chosen according to the criteria defined in Sec. 6 and considered to be the Higgs boson candidates H_1 and H_2. In the two dimensional space defined by the reconstructed

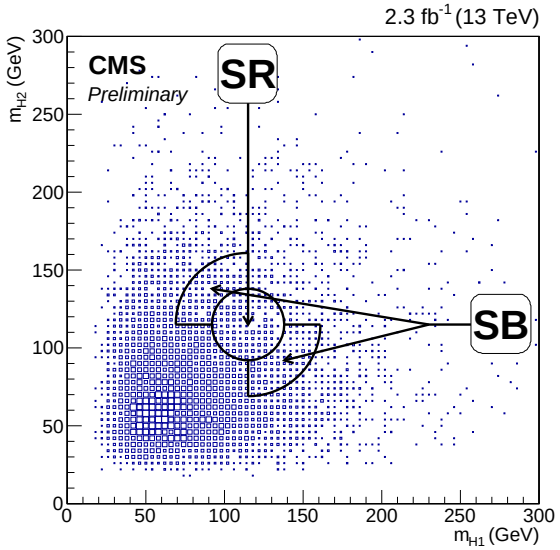

Figure 1: Illustration of SR and SB in the (m_{H1}, m_{H2}) plane used to motivate and validate the parametric model for the QCD multijet background. The quantities m_{H1} and m_{H2} are the two reconstructed Higgs boson masses after b-tagging and kinematic selections for data in medium-mass region.

masses of the two Higgs boson candidates, m_{H1} and m_{H2}, the Signal Region (SR) is defined as the circular region with $\chi < 1$ where χ is defined in Eq. 1, where 115 GeV is the average mean of m_{H1} and m_{H2} distributions for the different signal hypotheses, and σ_H has been optimized for the sensitivity and it is 17 GeV and 23 GeV for LMR and MMR respectively.

$$\chi^2 = \left(\frac{m_{H1} - 115\ \text{GeV}}{\sigma_H}\right)^2 + \left(\frac{m_{H2} - 115\ \text{GeV}}{\sigma_H}\right)^2 \tag{1}$$

The mass of the resonance, m_X, is computed as the invariant mass of H_1 and H_2 modified by corrections to the jet p_{T} discussed in the next paragraph. The parametric form that will be used to fit the m_X distribution of multi-jet QCD in the SR, is tested in the Sideband Region (SB), defined as $1 < \chi < 2$ and $(m_{H1} - 115) \cdot (m_{H2} - 115) < 0$. The veracity of employing the parametric form used to fit background events in SB to fit background events in SR is demonstrated in a control region where one of the jets is required to not be a b-jet by inverting the b-tag requirement.

In order to improve the resolution on the invariant mass of the resonance, the constraint on the invariant mass of the Higgs boson candidates to 125 GeV is exploited to correct the momenta of the reconstructed b-jets. Since jet direction is reconstructed with better resolution than jet p_{T}, this constraint mainly benefits the latter. This improves the invariant mass resolution of the reconstructed signal resonance by 20% – 40% depending on the mass hypothesis.

6 Event Selection

For the LMR, HH candidates are chosen from the four selected jets such that $|m_H - 115\,\text{GeV}| < 34\,\text{GeV}$ for each candidate Higgs boson. For the MMR, HH candidates are chosen from the four selected jets such that jets associated with a H candidate remain confined within a cone of

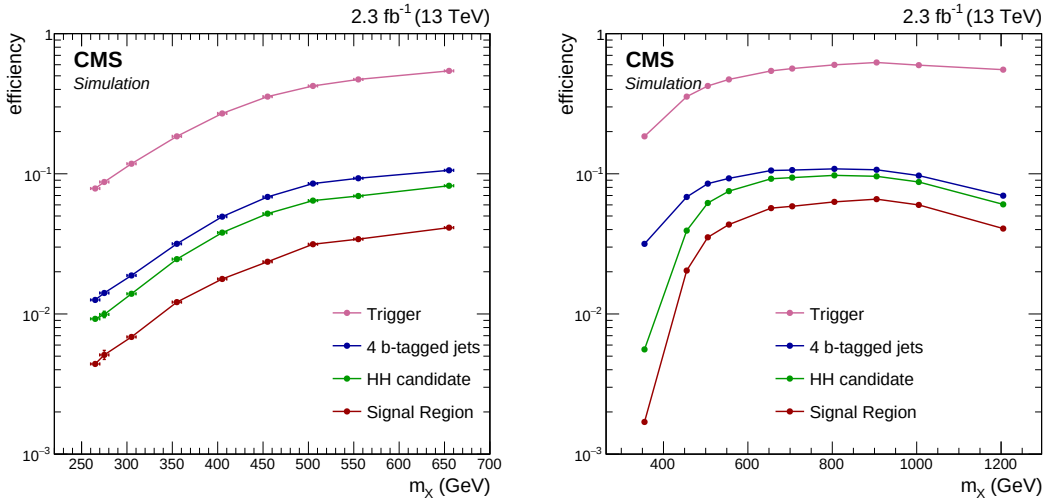

Figure 2: The selection efficiency for simulated $X \to H(b\bar{b})H(b\bar{b})$ events (X is a spin-2 RS1 KK-Graviton) at different stages of the event selection for each mass hypothesis, for the low-mass region (left) and the medium-mass region (right).

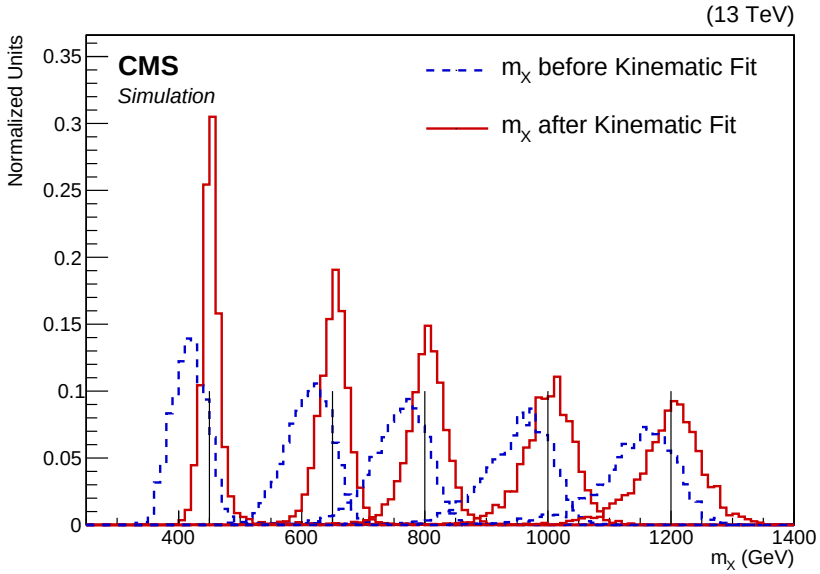

Figure 3: The m_X distribution of signal simulated events (spin-2 RS1 KK-Graviton) after the event selection criteria for each of mass hypothesis, with and without the correction by the kinematic constraint to m_H.

$\Delta R < 1.5$. In case of multiple HH candidates in an event, the combination that minimizes the χ^2 defined in Eq. 1 is chosen. The invariant masses of the the two Higgs boson candidates in each event, are reported in the two-dimensional plane of Fig. 1. Requiring events to fall within the SR defined in Fig. 1 completes the signal selection criteria. The cumulative selection efficiency of these criteria for the graviton signal benchmark is reported in Fig. 2. The reconstructed invariant mass distributions for the signal with different mass hypotheses are shown in Fig. 3.

7 Signal and Background Modeling

A parametric signal model is built for each mass hypothesis by fitting the m_X distribution in the signal Monte Carlo sample. A sum of two Gaussians, requiring five parameters, is used in the LMR to account for tails in the distribution from incorrect combinations of jets. In the MMR, an "ExpGaussExp" function with four parameters [5], is used to model the signal.

By comparing the numbers of data events and simulated Monte Carlo events of top quark pair production ($t\bar{t}$) in SR, $t\bar{t}$ is estimated to contribute approximately 10% and 15% of the selected events in the LMR and the MMR, respectively. The Z+jets, ZZ and ZH processes are found, through Monte Carlo studies, to contribute less than 1% of the background and are therefore neglected in this analysis. Since the $t\bar{t}$ contribution to the total background is relatively small and its shape is found not to affect the final background shape, it is not treated as a separate component in the data driven estimate.

The shape of the m_X distribution of the multi-jet events is estimated from data events in the sideband regions. The GaussExp function [5] is used to fit the m_X distributions in SB, and the SR of the invert b-tag control regions. Fits to the SB distribution in LMR are shown in Fig. 4. Examples of the test of this background model in the control region with one invertd b-tag, in SB and SR for MMR are shown in Fig 5. In both cases the GaussExp parametric form fits well the m_X distribution and therefore it is used to model the multi-jet background in SR.

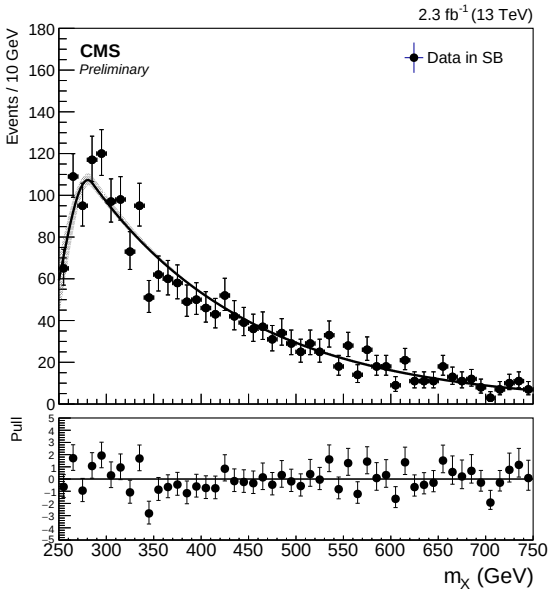

Figure 4: The m_X distributions in data in the Signal Region Sideband (SB) of the low-mass region. The distributions are fitted to the GaussExp function and the shaded regions correspond to 1 σ variations of the parametrized form.

8 Systematic Uncertainties

The signal yield for a given production cross section is affected by a 2.7% systematic uncertainty in the measurement of integrated luminosity at CMS [19]. Sources of systematic uncertainties that affect the signal efficiencies are listed in Table 1. The jet energy scale [15] is varied within one standard deviation as a function of jet p_T and η, and the efficiency of the selection criteria recomputed. It is found to affect signal efficiencies up to 4.2%. The effect of the uncertainty in

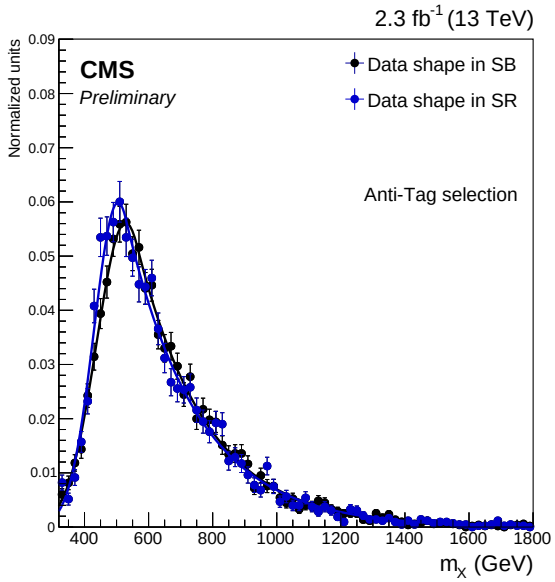

Figure 5: The m_X distributions in the control region of data where one of the four jets is required to not be a b-jet. The fits in the SR and SB regions of the medium-mass region are presented.

Table 1: Impact of systematic uncertainties on the signal efficiencies in the low-mass region (LMR) and the medium-mass region (MMR).

Source of systematic uncertainty	Impact in LMR (%) Signal	Impact in MMR (%) Signal
Jet energy scale	1.2 – 4.2	0.5 – 2.8
Jet energy resolution	0.1 – 1.4	0.5 – 1.1
b-tagging scale factor	9.6 – 10.4	9.3-10.6
Trigger efficiency	10.2 – 20.9	5.2 – 10.3

the jet energy resolution is evaluated by smearing the jet energies according to the measured uncertainty. This is found to affect signal efficiencies between 0.1% and 1.4%. The data-driven estimate for the trigger efficiency, has systematic uncertainties that are found to impact signal efficiencies between 5% and 20%. Uncertainties stemming from the b-tagging scale factor of the CMVA algorithm [17] is evaluated to be about 10%. The impact of these systematic uncertainties on the parametric models of the signal are also considered.

An alternative background model, based on polynomial functions has been used to compute biases in the reconstructed signal strength associated with the choice of the background parametric model. These biases are found to be negligible within the uncertainty on the measured signal strength for all mass hypotheses. Therefore, this is not included in the computation of expected upper limits.

9 Results

The m_X distribution in data within the SR and results of the fit with the parametric background model are shown in Fig. 6. Parameters controlling the shapes and yields of the signal are allowed to float within ranges determined by systematic uncertainties. The parameters and normalization of the multi-jet background shape are left free to float. Fits of the data to the

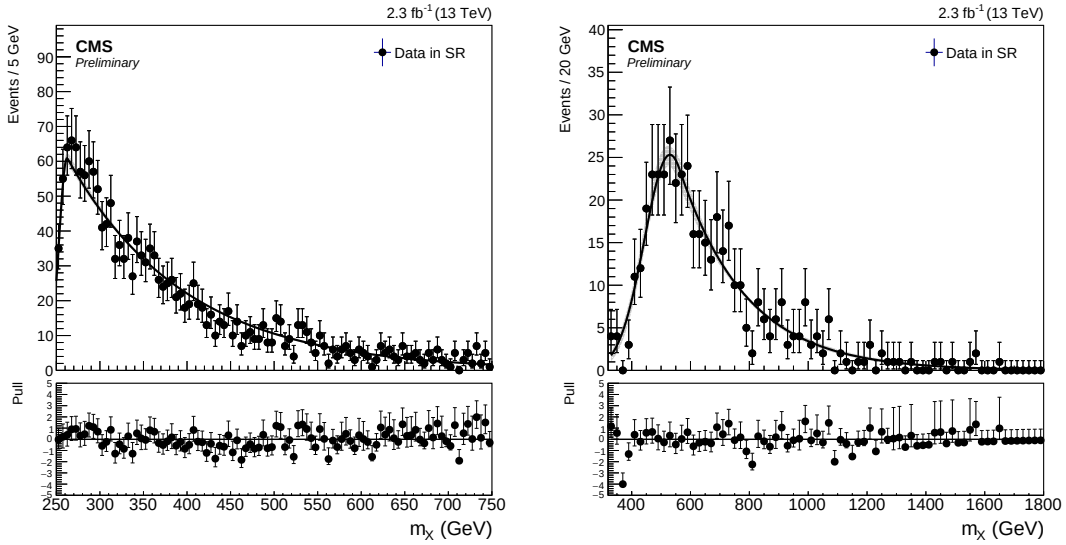

Figure 6: The m_X distribution in the Signal Region (SR) of data in the LMR (left) and the MMR (right). A fit to the background-only hypothesis, which consists of the QCD multi-jet shape is shown. The shaded region corresponds to a ± 1 σ variation of this parametrized form. The number of degrees of freedom (n) corresponds to the number of fit parameters (4) subtracted from the number of bins in the histogram.

background-only hypothesis are also shown in Fig. 6 and the shapes of the fit are found to adequately interpolate the data in both the LMR and MMR.

The observed and expected upper limits on the cross section for $pp \to X \to H(b\bar{b})H(b\bar{b})$ at a 95% confidence are computed using the modified frequentist CL_S method [20, 21]. These limits are shown in Fig. 7 and in more details in Tab. 2, where the green and yellow bands respectively represent the 1σ and 2σ confidence intervals around the expected limits. The observed upper limits are found to remain within 2σ band of the expected upper limits. The LO theoretical cross section for the gluon fusion production of a RS1 KK-Graviton decaying to a pair of Higgs bosons [22] each decaying to a $b\bar{b}$ pair with a branching fraction of 57% [9], is overlaid to the experimental limits on the cross section. The RS1 warped extra dimension scenario for this graviton has the product of the curvature, k, and half the circumference of the extra dimension, L, set to 35, and k set to 10% of the Planck mass (M_{Pl}).

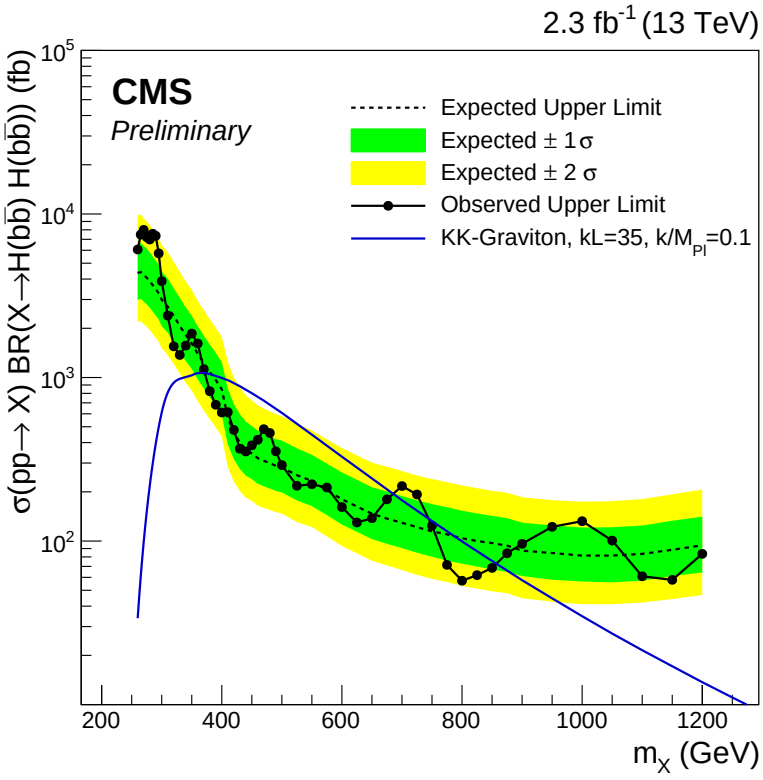

Figure 7: The observed and expected upper limits on the cross section for a spin-2 resonance $X \rightarrow H(b\bar{b})H(b\bar{b})$ at a 95% confidence level using data corresponding to an integrated luminosity of 2.3 fb^{-1} at $\sqrt{s}$ =13 TeV using the asymptotic CL_S method. Theoretical cross sections for the RS1 KK-Graviton, with k/M_{Pl} =0.1, kL =35, decaying to four b-jets via Higgs bosons are overlaid.

10 Conclusion

A model-independent search for a narrow-width resonance is presented. No evidence for a signal is observed in the explored mass range between 260 GeV and 1200 GeV. According to these results the RS1 KK-Graviton with $kL = 35$, $k/M_{Pl} = 0.1$ and mass above 350 GeV and below 725 GeV, and in the 775–850 GeV mass range is excluded at a 95% confidence level. A similar search, exploiting 17.9 fb^{-1} collected at 8 TeV, has comparable sensitivity [5] and excluded the same model in the 380–800 GeV mass range.

Table 2: The observed and expected upper limit of $\sigma(pp \to X \to H(b\bar{b})H(b\bar{b}))$ at a 95% confidence level using 2.3 fb^{-1} of data for the low and medium-mass regimes (LMR and MMR).

m_X (GeV)	Observed Upper Limit (fb)	Expected Upper Limit (fb)	-1σ (fb)	+1σ (fb)
LMR				
260	6065.3	4359.4	1382.5	2206.8
265	7480.2	4421.9	1390.5	2238.4
270	7996.4	4234.4	1326.3	2126.7
275	7217.7	4078.1	1277.3	2064.5
280	6982.3	3890.6	1218.6	1938.5
285	7572.3	3671.9	1135.9	1829.5
290	7351.6	3484.4	1091.4	1708.3
295	5743.9	3265.6	1002.2	1627.1
300	3884.3	2992.2	929.9	1455.1
310	2392.9	2710.9	832.0	1318.4
320	1551.4	2382.8	740.5	1130.3
330	1373.9	2070.3	630.4	998.5
340	1568.0	1820.3	561.2	863.5
350	1857.2	1617.2	488.5	767.1
360	1614.0	1394.5	421.2	661.5
370	1121.2	1230.5	373.2	573.8
380	824.9	1082.0	326.8	513.3
390	681.0	964.8	291.4	450.0
400	611.0	847.7	256.1	395.3
MMR				
400	759.6	993.0	248.4	401.4
410	614.1	568.4	178.8	287.7
420	478.2	462.9	146.1	226.9
430	367.1	400.4	124.4	196.3
440	352.8	364.3	113.2	172.7
450	384.8	342.8	103.6	162.6
460	416.5	321.3	97.1	152.4
470	483.3	309.6	93.5	144.3
480	458.0	299.8	90.4	137.4
490	353.7	288.1	85.6	132.0
500	291.5	282.2	83.9	129.4
525	217.8	252.9	75.8	116.0
550	222.5	233.4	69.8	103.3
575	212.6	205.1	61.0	92.4
600	161.4	180.2	54.0	82.6
625	129.9	162.6	48.7	74.5
650	137.9	147.0	44.2	66.2
675	179.9	137.2	41.1	64.0
700	216.9	129.4	38.8	59.3
725	192.7	121.6	36.7	55.7
750	124.2	114.7	34.5	53.6
775	71.5	109.9	34.0	51.2
800	57.2	104.0	31.3	49.3
825	61.9	100.1	30.2	47.5
850	68.5	97.2	30.3	45.3
875	84.2	93.8	29.0	44.4
900	96.2	87.4	26.3	42.2
950	122.2	84.5	26.5	40.0
1000	132.3	81.5	24.9	39.4
1050	100.7	81.5	25.6	40.0
1100	60.9	83.5	26.2	40.9
1150	57.9	88.4	27.3	44.7
1200	83.5	94.2	30.0	47.0

References

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Physics Letters B* **716** (2012) doi:http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[2] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* **716** (2012) doi:http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[3] C. Csáki, J. Hubisz, and S. J. Lee, "Radion phenomenology in realistic warped space Model", *Phys. Rev. D* **76** (Dec, 2007) 125015, doi:10.1103/PhysRevD.76.125015.

[4] ATLAS Collaboration, "Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from pp collisions at $\sqrt{s} = 8$ TeVwith the ATLAS detector", *Eur. Phys. J.* **C75** (2015) 412, doi:10.1140/epjc/s10052-015-3628-x, arXiv:1506.00285.

[5] CMS Collaboration, "Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 8 TeV", *Phys. Lett.* **B749** (2015) 560–582, doi:10.1016/j.physletb.2015.08.047, arXiv:1503.04114.

[6] CMS Collaboration, "The CMS experiment at the CERN LHC", *Journal of Instrumentation* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[7] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *Journal of Instrumentation* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[8] J. Alwall et al., "MadGraph 5: going beyond", *Journal of High Energy Physics* **06** (2011) 128, doi:10.1007/JHEP06(2011)128.

[9] M. Gouzevitch et al., "Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production", *Journal of High Energy Physics* **07** (2013) 148, doi:10.1007/JHEP07(2013)148.

[10] CMS Collaboration, "Measurement of the underlying event activity at the LHC with $\sqrt{s}$ = 7 TeV and comparison with $\sqrt{s}$ = 0.9 TeV", *Journal of High Energy Physics* **09** (2011) 109, doi:10.1007/JHEP09(2011)109.

[11] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[12] M. Cacciari, G. P. Salam and G. Soyez, "The anti-k_t jet clustering algorithm", *Journal of High Energy Physics* **04** (2008) 063.

[13] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[14] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett.* **B641** (2006) 57–61, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[15] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *Journal of Instrumentation* **6** (2011).

[16] CMS Collaboration, "Pileup Jet Identification", *PAS JME-13-005* (2013).

[17] CMS Collaboration, "Identification of b-quark Jets at the CMS Experiment in the LHC Run2 Startup", *CMS Physics Analysis Summary* **CMS-PAS-15-001** (2016).

[18] CMS Collaboration, "Measurement of $B\bar{B}$ angular correlations based on secondary vertex reconstruction at $\sqrt{s} = 7$ TeV", *Journal of High Energy Physics* **03** (2011) 136, doi:10.1007/JHEP03(2011)136.

[19] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", *CMS Physics Analysis Summary* **CMS-PAS-LUM-13-001** (2013).

[20] A. L. Read, "Presentation of search results: the CLs technique", *Journal of Physics G: Nuclear and Particle Physics* **28** (2002).

[21] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, "Procedure for the LHC Higgs boson search combination in Summer 2011", technical report, CERN, Geneva, Aug, 2011.

[22] V. Barger and M. Ishida, "Randall-Sundrum Reality at the LHC", *Phys. Lett. B* **709** (2012) doi:10.1016/j.physletb.2012.01.073, arXiv:1110.6452.