CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2014/08/08

Search for narrow resonant pair production of the Higgs boson in the final state with four bottom quarks

The CMS Collaboration

Abstract

A model-independent search for narrow width resonant pair production of the Higgs boson where both 125 GeV bosons decay into bottom quarks is presented. The search is performed in proton-proton collision data corresponding to an integrated luminosity of 17.93 fb^{-1} at $\sqrt{s}$ = 8 TeV recorded by the CMS detector at the LHC. No evidence for a signal is observed. Upper limits on the production cross section for such a resonance, in the mass range of 270 GeV to 1100 GeV, are reported.

1 Introduction

Following the discovery of the Higgs boson (H) with mass around 125 GeV and properties consistent with the standard model of particle physics (SM) at the Large Hadron Collider (LHC) [1, 2], it has become experimentally feasible and important to search for any resonant pair production mechanism for it. 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 radion in Randall-Sundrum (RS1) models of Warped Extra Dimensions [3]. It is the quantum of the scalar field with a non-zero vacuum expectation value introduced in order to stabilize the radius of the extra dimension. This paper reports the results of a search for such a resonance between masses of 270 GeV and 1100 GeV where both Higgs bosons decay into bottom quarks. We perform this search in 17.93 ± 0.47 fb^{-1} of proton-proton collision data acquired at $\sqrt{s}$ = 8 TeV by the Compact Muon Solenoid (CMS) detector at the LHC. The main challenge of this search is to distinguish the signature of four bottom quarks in the final state that hadronize into jets from the copious multi-jet quantum chromodynamic (QCD) background in proton-proton collisions. This is addressed by suitable event selection criteria that include b-jet identification techniques and a model of the multi-jet background that is tested in control regions of data.

The results of this analysis may be compared with a similar search in the four bottom quark final state conducted by the ATLAS experiment [4].

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 [5].

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. In order to overcome the overwhelming rate of QCD multi-jet events, the trigger used for this analysis exploits a b-jet identification algorithm. The first level of the trigger requires two jets to exceed p_{T} thresholds of 56 GeV or 64 GeV, depending on luminosity conditions. At the HLT, four jets are required with $p_{\mathrm{T}} > 30$ GeV, two of which are required to have $p_{\mathrm{T}} > 80$ GeV. Furthermore, at the HLT, two b-jets are identified using the Combined Secondary Vertex (CSV) algorithm [6] that relies mainly on quantities measured with the pixel

detector. In 85% of simulated signal events that pass this trigger, the two jets that satisfy b-tagging requirements at the trigger also satisfy b-jet requirements at event reconstruction after the trigger selection. Using this trigger, the CMS experiment has collected data corresponding to an integrated luminosity of 17.93 fb^{-1} at $\sqrt{s}$ = 8 TeV.

The signal process used to optimize this analysis is the spin-0 RS1 radion forced to decay to a pair of Higgs bosons with parameters reported in [7], where both Higgs bosons decay to $\text{b}\bar{\text{b}}$. MADGRAPH 5.1 [8] is used to generate the signal, di-boson, W+jets, Z+jets, and $\text{t}\bar{\text{t}}$ samples. Single-top samples are produced with POWHEG [9–11] and the QCD multi-jet samples with PYTHIA [12]. Samples of the Higgs boson produced in association with a Z boson are produced using the POWHEG event generator interfaced with HERWIG++ [13] for parton showering and hadronization. The PYTHIA parameters for the underlying event are set to the Z2Star tune [14]. The response of the CMS detector is modeled using GEANT4 [15].

4 Event Reconstruction

Jets are reconstructed from particle-flow candidates using the anti-k_T clustering algorithm [16], with a distance parameter of 0.5, as implemented in the FASTJET package [17, 18]. Jet energy corrections, as a function of pseudorapidity and transverse momentum of the jet, are applied [19]. 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 [20].

Techniques for identifying b-jets, also known as b-tagging, is the fulcrum 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 and charged particles in their decay products form secondary vertices with tracks that are displaced from the primary interaction point with impact parameters of the same distance scale. Measurements of the properties of such secondary vertices, lifetimes of the B-hadron, and low-p_{T} lepton information when available for a jet are used by the Combined MultiVAriate algorithm (CMVA) [21] for b-tagging events reconstructed after the trigger selection. It determines secondary vertices using the Inclusive Vertex Finder algorithm [22]. The CMVA algorithm outputs a continuous discriminant between 0 and 1. A working point that yields a 74.6% efficiency for tagging true b-jets and a 3.4% efficiency for tagging light flavor jets is chosen by optimizing it against the sensitivity of this analysis.

Simulated events are weighted to match the number of primary vertices per event in data. Corrections to account for differences in efficiencies between data and simulation for the trigger are evaluated in a $\text{t}\bar{\text{t}}$-enriched control region of the data. The corrections are derived as functions of reconstructed jet p_{T} and the jet CMVA discriminants, and no significant correlation is observed between their dependence on these two quantities. This $\text{t}\bar{\text{t}}$-enriched control region is obtained using a control trigger which requires a muon with $p_{\text{T}} > 24$ GeV. A $\text{t}\bar{\text{t}}$-enriched control region in data and simulation is also used to estimate differences in b-tagging efficiencies between data and simulation. Event weights to correct for these differences in data and simulation due to the trigger efficiency and b-tagging are applied to simulated events.

5 Analysis Strategy

We search for a narrow-width $\text{X} \rightarrow \text{H}(\text{b}\bar{\text{b}})\text{H}(\text{b}\bar{\text{b}})$ resonance between masses of 270 GeV and 1100 GeV. The momenta and angles of the decay products of such a resonance change substantially over this range and in order to obtain the lowest expected upper limits we use different event selection criteria in two main kinematic regions: the Low Mass Region (LMR) for reso-

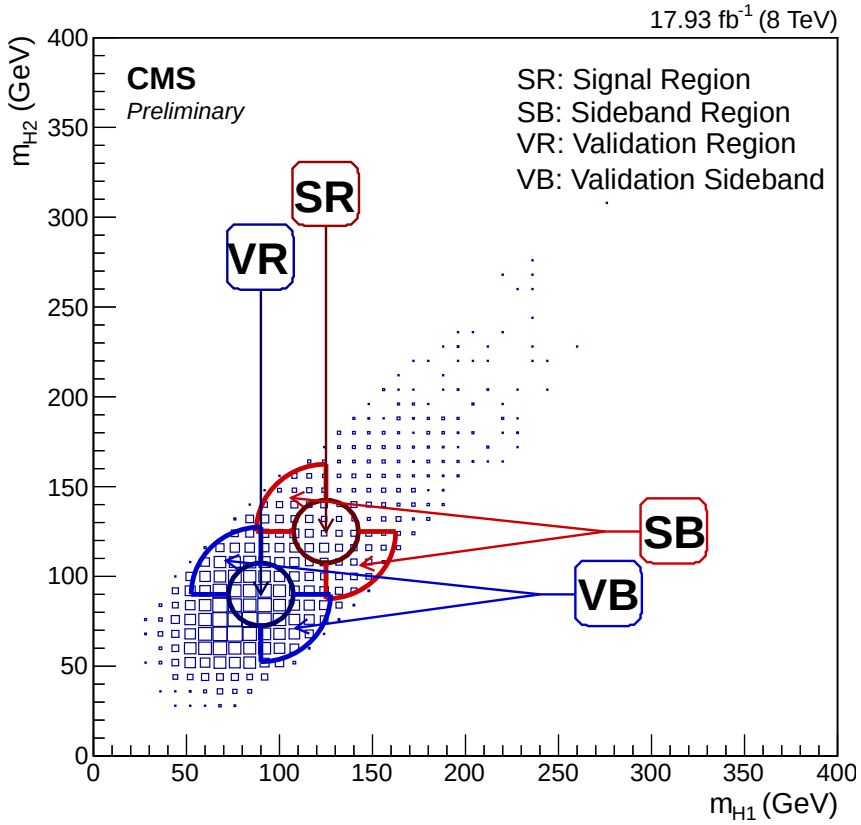

Figure 1: The distribution of events after b-tagging and kinematic selections as a function of the two reconstructed Higgs boson masses m_{H1} and m_{H2} for data.

nance mass hypothesis from 270 GeV to 440 GeV, and the High Mass Region (HMR) for masses from 450 GeV to 1100 GeV.

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 4 central ($|\eta| < 2.5$) b-tagged jets with $p_{\mathrm{T}} > 40$ GeV. Amongst these jets, two pairs are chosen according to appropriate criteria and considered to be the Higgs boson candidates H_1 and H_2. In the two dimensional space defined by the reconstructed masses of the two Higgs boson candidates, m_{H1} and m_{H2}, the Signal Region (SR) is defined as $\sqrt{\Delta m_{H1}^2 + \Delta m_{H2}^2} < 17.5$ GeV, where $\Delta m_{H_{1,2}} = m_{H_{1,2}} - 125$ GeV. In order to avoid observer-expectancy bias, this region is kept blinded during the optimization of this analysis. 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. We test the parametric form that will be used to fit the m_X distribution of multi-jet QCD in SR in the Sideband Region (SB) defined as $17.5 \text{ GeV} < \sqrt{\Delta m_{H1}^2 + \Delta m_{H2}^2} < 35$ GeV and $\Delta m_{H1} \cdot \Delta m_{H2} < 0$. The veracity of employing the parametric form used to fit QCD events in SB to fit QCD events in SR is demonstrated in a neighboring Validation Region (VR) and Validation Region Sideband (VB) similarly defined with $\Delta m_{H_{1,2}} = m_{H_{1,2}} - 90$ GeV. This method is further demonstrated in an additional Control Region where one of the jets is required to not be a b-jet by reversing the b-tag requirement. Data in this Control Region corresponding to SR is unblinded for this demonstration.

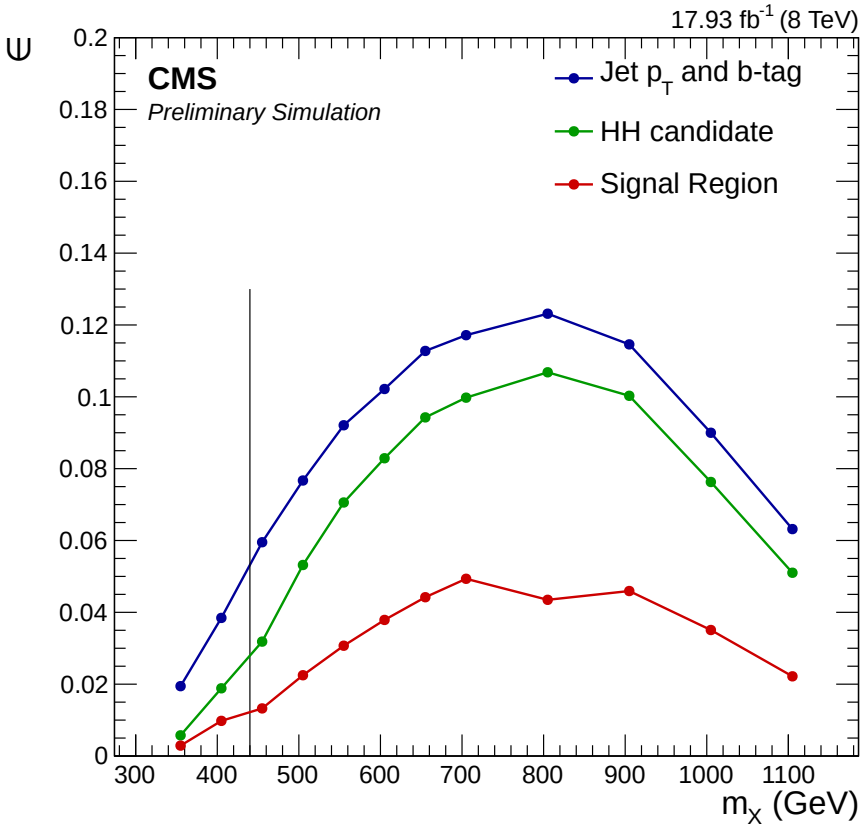

Figure 2: The selection efficiency (ϵ) for $X \to H(b\bar{b})H(b\bar{b})$ events at different stages of the event selection for each mass hypothesis. The vertical line separates event selection criteria for the Low Mass Region on the left from the High Mass Region on the right.

In order to improve the resolution on the invariant mass of the resonance, a constraint on the invariant mass of the Higgs boson candidates to 125 GeV is used to correct the momenta of the reconstructed b-jets. Since jet direction is reconstructed with better resolution than jet p_{T}, this constraint mainly affects the latter. This improves the invariant mass resolution of the reconstructed signal resonance by 20% – 40% depending on the mass hypothesis. Higgs boson candidates in events within VR and VB regions are instead constrained to a mass of 90 GeV.

6 Event Selection

For the LMR, HH candidates are paired from the 4 selected jets such that $|m_H - 125\,\mathrm{GeV}| < 35\,\mathrm{GeV}$ for each candidate Higgs boson. Selected HH candidates are required to have at least two jets with $p_T > 90$ GeV. Thus, the minimum invariant mass of the resonance that this analysis can probe is 260 GeV. The lowest mass hypothesis simulated for this analysis is 270 GeV and the signal efficiency for it is 0.013%. This is extremely low and does not offer a stringent upper limit, so we do not simulate lower mass hypotheses.

For the HMR, HH candidates are paired from the 4 selected jets such that jets associated with a H candidate remain confined within a cone of $\Delta R < 1.5$. For mass hypotheses of 740 GeV to 1100 GeV, an additional requirement on the Higgs boson candidate $p_{\mathrm{T}} > 300$ GeV is placed.

In case of multiple HH candidates in an event, the combination that minimizes $|m_{H1} - m_{H2}|$ is

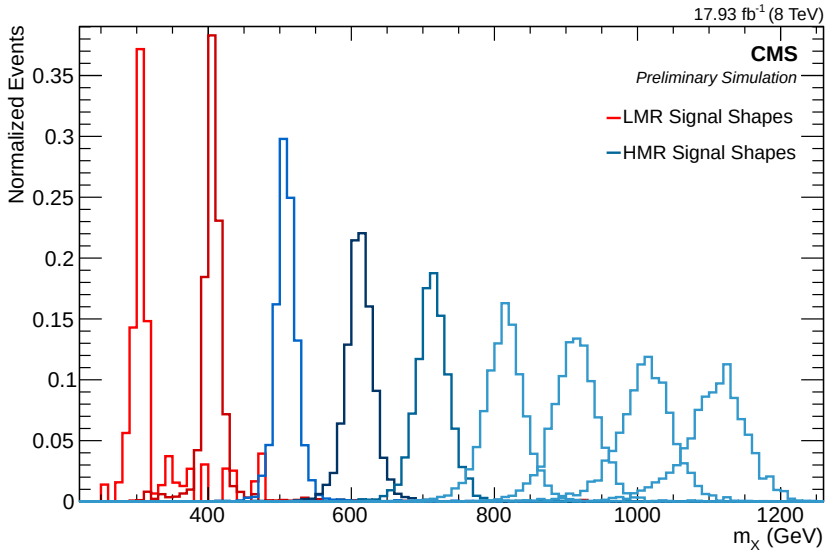

Figure 3: The m_X distribution of signal events after the event selection criteria for each of mass hypothesis. Momenta of b-jets have been corrected by the kinematic constraint to m_H.

chosen. Having identified the two Higgs boson candidates in each event, we plot their masses on a two dimensional histogram as shown in 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 radion signal samples is shown in Fig. 2. The reconstructed invariant mass distributions for the signal with different mass hypotheses are shown in Fig. 3.

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 5 parameters, is used in the LMR to account for tails in the distribution from incorrect combinations of jets. In the HMR, an "ExpGaussExp" function with 4 parameters as described in the Appendix A, is used to fit the signal. Examples of these fits are shown in Fig. 4.

7 Background Modeling

By comparing the numbers of data events and simulated Monte Carlo events of top quark pair production ($t\bar{t}$) in SR, we estimate that $t\bar{t}$ contributes approximately 22% and 27% of the selected events between m_X of 270 GeV and 1100 GeV in LMR and HMR, respectively, as shown in Fig. 5. 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 background is significant, we decide to treat it as a separate component that is studied through simulation. It is modeled in the m_X distribution using the parametric function "GaussExp" as described in Appendix B and illustrated in Fig. 6.

The shape of the m_X distribution of QCD multi-jet events, the main component of the background, is estimated by subtracting simulated $t\bar{t}$ events, as shown in Fig. 5, from data events in the sideband regions. The GaussExp function with all parameters floating is demonstrated to fit well the m_X distributions in SB, VR, VB, and the corresponding areas in the reverse b-tag Control Regions. Such fits to the VR and VB distributions of m_X in LMR, and the SB distribution in HMR are shown in Fig. 7. Examples of the test of this background model in the Control Region with one reversed b-tag, in VR and VB for LMR, and SR and SB for HMR are shown in

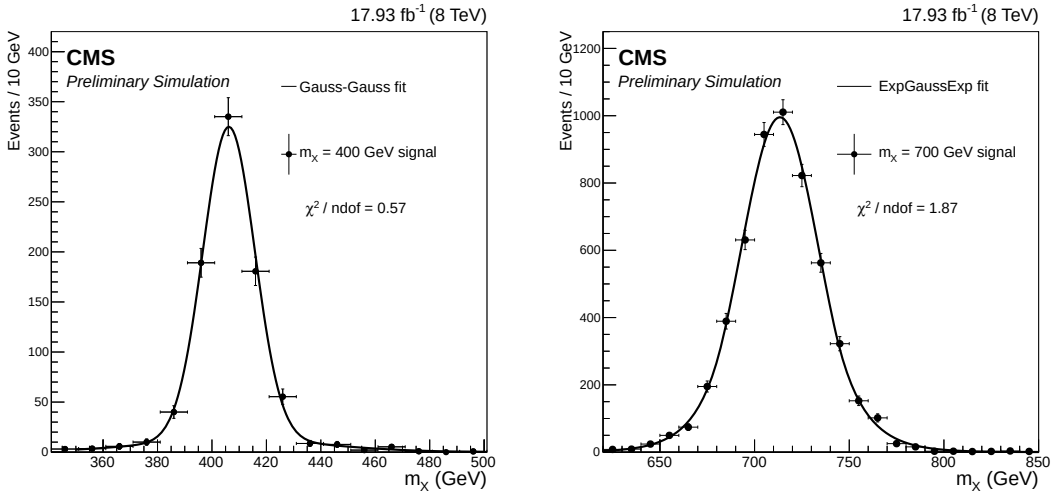

Figure 4: The fit to the m_X distribution of simulated signal events after the event selection criteria for $m_X = 400$ GeV (left) and 700 GeV (right). Momenta of b-jets have been corrected by the kinematic constraint to m_H. The distribution on the left is fitted to the sum of two Gaussian functions, while that on the right is fitted to the ExpGaussExp function described in the text.

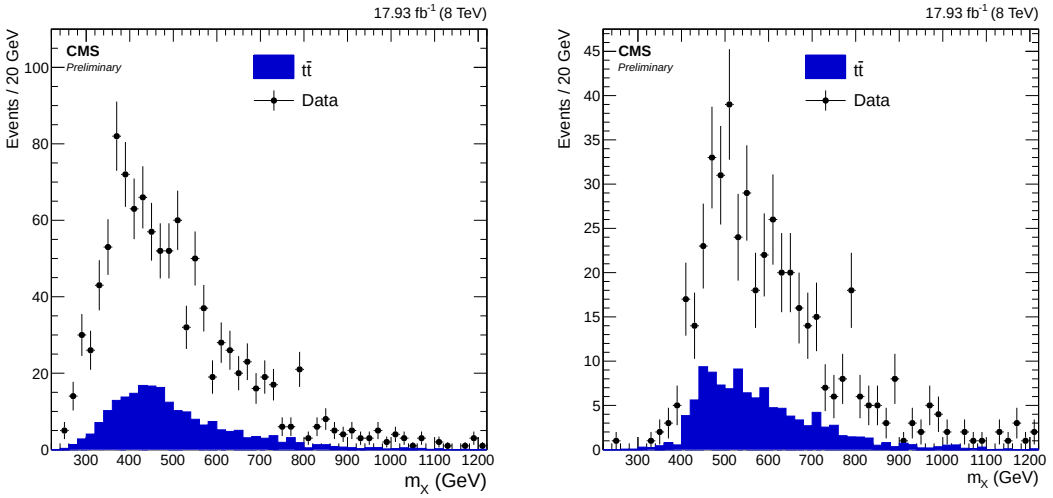

Figure 5: The $t\bar{t}$ composition (filled colors) of the estimated background (markers) in the SR region for the Low Mass Region (left), and the High Mass Region (right).

Fig 8. Thus, we conclude that the GaussExp parametric form with free parameters can be used to fit the QCD multi-jet distribution of m_X in SR.

8 Systematic Uncertainties

The signal yield for a given production cross section and the $t\bar{t}$ yield are both affected by a 2.6% systematic uncertainty in the measurement of integrated luminosity at CMS [23]. Sources of systematic uncertainties that affect the signal and $t\bar{t}$ efficiencies are listed in Table 1. The jet energy scale 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 0.2%

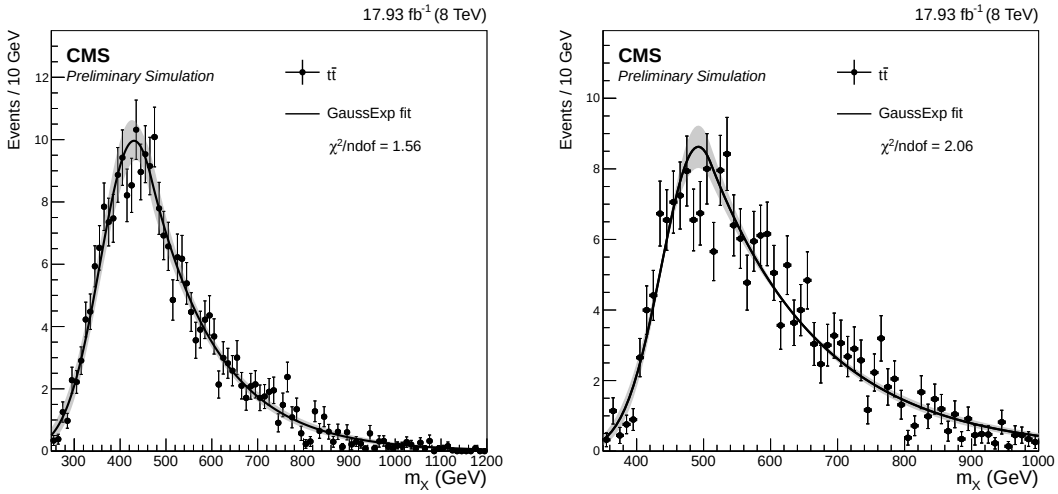

Figure 6: The m_X of simulated $t\bar{t}$ events after the event selection criteria for the Low Mass Region (left), and the High Mass Region (right). The distribution is fitted to the GaussExp function. The shaded region corresponds to a 1 σ variation of this parametrized form.

Table 1: Impact of systematic uncertainties on the signal and $t\bar{t}$ efficiencies in the Low Mass Region (LMR) and the High Mass Region (HMR).

Source of	Impact in LMR (%)		Impact in HMR (%)	
systematic uncertainty	Signal	$t\bar{t}$	Signal	$t\bar{t}$
Jet energy scale	0.1 – 0.2	0.8	0.0 – 0.2	0.1
Jet energy resolution	2.4 – 7.0	2.7	5.5 – 6.7	2.1
b-tagging scale factor	12.7	12.7	12.7	12.7
Trigger scale factor	6.0 – 18.8	9.1	6.1 – 8.0	7.2

and $t\bar{t}$ efficiencies up to 0.8%. The effect of uncertainty in the jet energy resolution is evaluated by smearing the jet energies according to the measured uncertainty. This is found to affect signal efficiencies between 2% and 7%, and $t\bar{t}$ efficiencies between 2% and 3%. Data-to-simulation scale factors for the trigger efficiency, as described in Section 4, have systematic uncertainties that are found to impact signal efficiencies between 6% and 18%, and $t\bar{t}$ efficiencies between 7% and 9%. Uncertainties stemming from the b-tagging scale factor of the CMVA algorithm is evaluated to be 12.7%. The impact of these systematic uncertainties on the parametric models of the signal and $t\bar{t}$ are also considered.

Additionally, a systematic uncertainty of 15% is placed on the $t\bar{t}$ production cross section, that propagates directly to its yield.

The bias in reconstructed signal strength from the possibility that the true distribution of the QCD multi-jet background is not the GaussExp function is estimated to be 2% – 26% of the signal cross section excluded at 95% confidence for LMR and 15% – 32% of the same for HMR. This is applied as a systematic uncertainty to the background model.

9 Results

We unblind the Signal Region (SR) to observe the distribution of m_X in data as shown in Fig. 9. This is fit to parametric models of the signal, $t\bar{t}$-background and QCD multi-jet background. Pa-

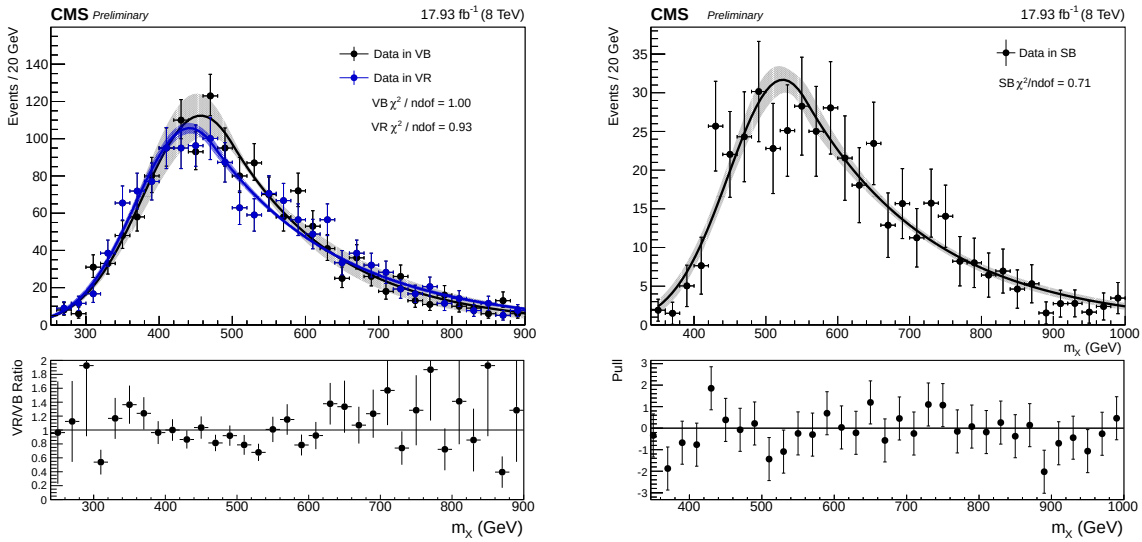

Figure 7: The m_X distributions of the QCD multi-jet component of the background in VR and VB of the Low Mass Region (left), and in the Signal Region Sideband (SB) of High Mass Region (right). The distributions are fitted to the GaussExp function and the shaded regions correspond to 1 σ variations of the parametrized form.

rameters controlling the shapes and yields of the signal and $t\bar{t}$-background models are allowed to float within ranges determined by systematic uncertainties. The parameters and normalization of the QCD multi-jet shape are allowed to float freely. Fits of the data to the background-only hypothesis are also shown in Fig. 9. We conclude that they adequately fit the unblinded data.

The observed and expected upper limits on the cross section for $pp \to X \to H(b\bar{b})H(b\bar{b})$ at 95% confidence are computed using the modified frequentist CL_S method [24, 25]. These limits are shown in Fig. 10 where the green and yellow bands represent the 1σ and 2σ confidence intervals around the expected limits. The observed upper limits are found to remain within 2σ of the expected upper limits. The theoretical cross section for the gluon fusion production of a radion (R) decaying to a pair of Higgs bosons that in turn decay to $b\bar{b}$ with a branching fraction of 57%, as calculated by MADGRAPH 5.1 is overlaid on the figure. In this calculation, the correction factor to account for next-to-leading order amplitudes of the heavy-quark loop is identical to that used for Higgs boson production through gluon fusion [26]. The warped extra dimension scenario for this radion has the product of the curvature and length of the extra dimension set to $kL = 35$, and the radion decay constant $\Lambda_R = 1$ TeV. The branching fraction $Br(R \to HH)$ is set to 25% for $m_R > 300$ GeV as an approximation to the theoretical value [3]. Such a radion is excluded at 95% confidence between masses of 300 GeV and 1000 GeV.

This analysis has been repeated with the spin-2 RS1 KK-graviton as the signal. An increase in signal efficiency of 20% – 30% is observed that results in a lower observed limit for this model. ATLAS limits [4] in the mass range of 500 GeV to 1500 GeV may be compared to such interpretations.

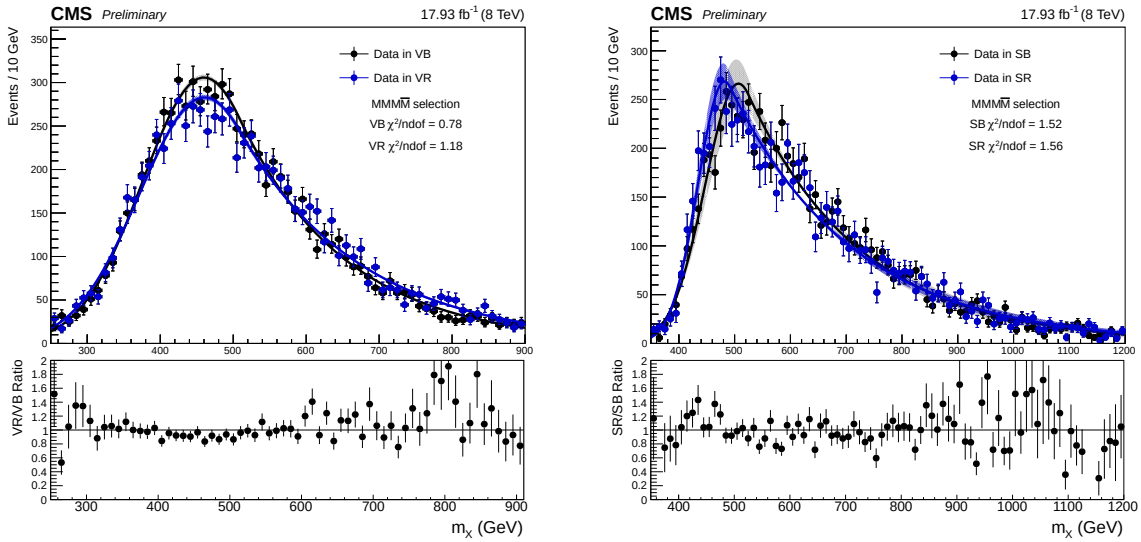

Figure 8: The m_X distributions of the QCD multi-jet component in the Control Region where one of the four jets is required to not be a b-jet. Presented are fits in the VR and VB regions of the Low Mass Region (left), and in the SR and SB regions of the High Mass Region (right). All distributions are fitted to the GaussExp function and the shaded regions correspond to 1 σ variations of the parametrized form.

10 Conclusion

A search for narrow width resonances between masses of 270 GeV and 1100 GeV decaying to pairs of Higgs bosons has been performed in the final state with four bottom quarks using 17.93 fb^{-1} of proton-proton collisions collected at $\sqrt{s} = 8$ TeV. No statistically significant signal is observed. Observed and expected upper limits at 95% confidence on the production cross section for such a resonance, in the mass range of 270 GeV to 1100 GeV, have been set.

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), no. 1, 30 – 61, 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), no. 1, 1 – 29, 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] The ATLAS Collaboration Collaboration, "A search for resonant Higgs-pair production in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 8$ TeV", Technical Report ATLAS-CONF-2014-005, CERN, Geneva, Mar, 2014.

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

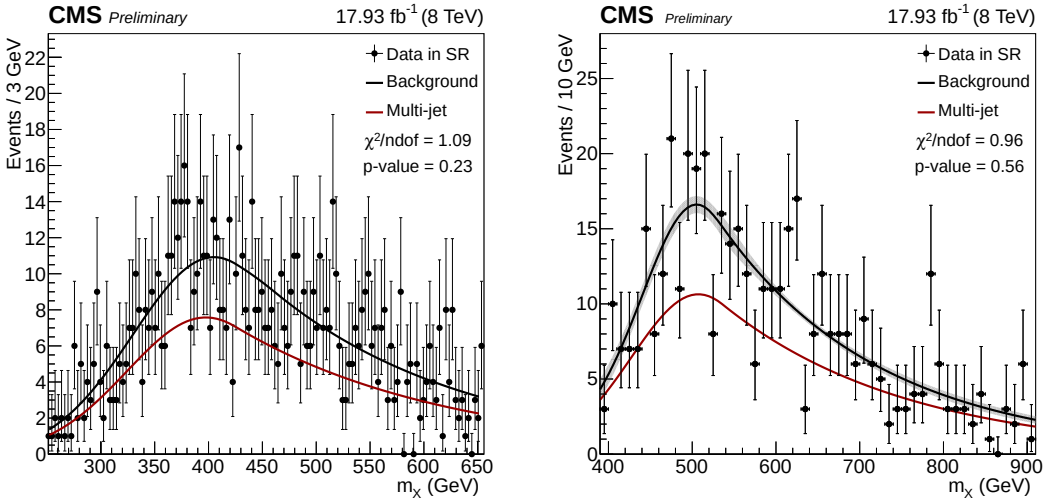

Figure 9: The m_X distribution in the Signal Region (SR) of data in the LMR (left) and the HMR (right) on unblinding. A fit to the background-only hypothesis, which consists of the $t\bar{t}$ shape stacked on top of the QCD multi-jet shape, is shown. The shaded region corresponds to a 1 σ variation of this parametrized form. The number of degrees of freedom (ndof) corresponds to the number of fit parameters (4) subtracted from the number of bins in the histogram. We conclude that the background-only hypothesis adequately fits the unblinded data.

[6] 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.

[7] M. Gouzevitch et al., "Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production", *Journal of High Energy Physics* **2013** (2013), no. 7, doi:10.1007/JHEP07(2013)148.

[8] J. Alwall et al., "MadGraph 5: going beyond", *Journal of High Energy Physics* **2011** (2011), no. 6, doi:10.1007/JHEP06(2011)128.

[9] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *Journal of High Energy Physics* **2007** (2007), no. 11, 070.

[10] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *Journal of High Energy Physics* **2004** (2004), no. 11, 040.

[11] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *Journal of High Energy Physics* **2010** (2010), no. 6, doi:10.1007/JHEP06(2010)043.

[12] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *Journal of High Energy Physics* **2006** (2006), no. 05, 026.

[13] S. Gieseke et al., "Herwig++ 2.0 Release Note", (2006). arXiv:hep-ph/0609306.

[14] 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* **2011** (2011), no. 9, doi:10.1007/JHEP09(2011)109.

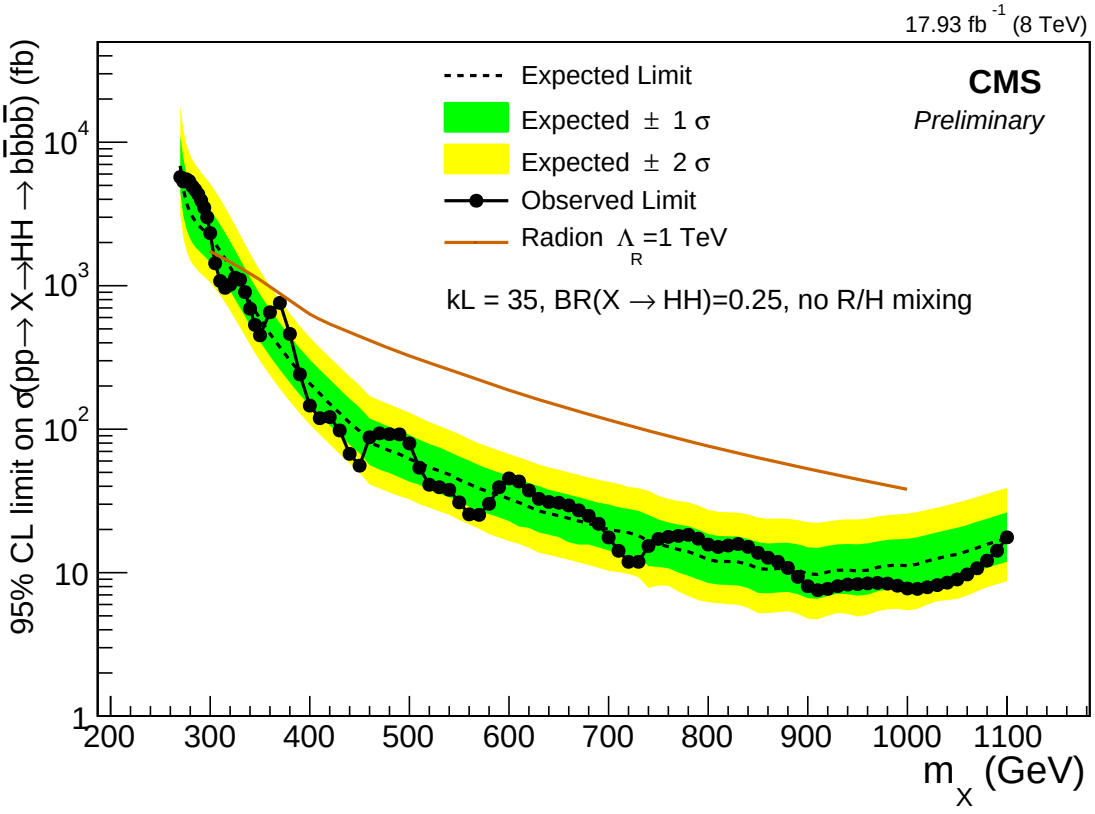

Figure 10: The observed and expected upper limits on the cross section for pp $\rightarrow$ X $\rightarrow$ H(b$\bar{\text{b}}$)H(b$\bar{\text{b}}$) at 95% confidence level using data corresponding to an integrated luminosity of 17.93 fb^{-1} at $\sqrt{s}$ =8 TeV using the asymptotic CL$_s$ method. Theoretical cross sections for the RS1-radion, with Λ_R =1 TeV, kL =35, and no radion-Higgs mixing, decaying to four b-jets via Higgs bosons are overlaid.

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

[16] M. Cacciari, G. P. Salam and G. Soyez, "The anti-k_t jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008), no. 04, 063.

[17] 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.

[18] 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.

[19] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *Journal of Instrumentation* **6** (2011), no. 11, P11002.

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

[21] CMS Collaboration Collaboration, "Algorithms for b Jet identification in CMS", Technical Report CMS-PAS-BTV-09-001, CERN, 2009. Geneva, Jul, 2009.

[22] CMS Collaboration, "Measurement of $B\bar{B}$ angular correlations based on secondary vertex reconstruction at $\sqrt{s} = 7$ TeV", *Journal of High Energy Physics* **2011** (2011), no. 3, doi:10.1007/JHEP03(2011)136.

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

[24] A. L. Read, "Presentation of search results: the CLs technique", *Journal of Physics G: Nuclear and Particle Physics* **28** (2002), no. 10, 2693.

[25] 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 CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011.

[26] S. Catani, D. de Florian, M. Grazzini, and P. Nason, "Soft-gluon resummation for Higgs boson production at hadron colliders", *Journal of High Energy Physics* **2003** (2003), no. 07, 028.

A The ExpGaussExp Function

The "ExpGaussExp" function has been developed to model the signal in the High Mass Region. This function has a Gaussian core differentiably continued on both sides to exponential tails. It is expressed in Eq. 1 and has four parameters as listed below.

- $\bar{x}$: The mean of the Gaussian core,
- σ: The standard deviation of the Gaussian core,
- k_L: The decay-coefficient of the lower exponential tail. This is also the number of standard deviations, on the low side, beyond which the Gaussian inflects into the exponential.
- k_H: The decay-coefficient of the higher exponential tail. This is also the number of standard deviations, on the high side, beyond which the Gaussian inflects into the exponential.

$$\begin{aligned} f(x;\bar{x},\sigma,k_L,k_H) &= \exp\left(\frac{k_H^2}{2} - \frac{k_H(x-\bar{x})}{\sigma}\right), \quad \text{for} \quad \frac{x-\bar{x}}{\sigma} > k_H \\ &= \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), \quad \text{for} \quad k_L \leq \frac{x-\bar{x}}{\sigma} \leq k_H \\ &= \exp\left(\frac{k_L^2}{2} + \frac{k_L(x-\bar{x})}{\sigma}\right), \quad \text{for} \quad \frac{x-\bar{x}}{\sigma} < k_L \end{aligned} \tag{1}$$

B The GaussExp Function

This function has been developed to fit the background m_X distributions. It has a Gaussian core differentiably continued on the high side to an exponential tail. It is motivated by the fact that the high side tails, well beyond the influence of trigger thresholds, and representative of multi-jet physics are observed to be exponentials on semi-log plots. The GaussExp function is expressed in Eq. 2 and has three parameters as listed below.

- $\bar{x}$: The mean of the Gaussian core,
- σ: The standard deviation of the Gaussian core,
- k: The decay-coefficient of the exponential tail. This is also the number of standard deviations beyond which the Gaussian inflects into the exponential on the high side.

$$\begin{aligned} f(m_X;\bar{x},\sigma,k) &= \exp\left(-\frac{1}{2}(\frac{x-\bar{x}}{\sigma})^2\right), \quad \text{for} \quad \frac{x-\bar{x}}{\sigma} \leq k \\ &= \exp\left(\frac{k^2}{2} - k\frac{x-\bar{x}}{\sigma}\right), \quad \text{for} \quad \frac{x-\bar{x}}{\sigma} > k \end{aligned} \tag{2}$$