CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2016/08/04

Search for a narrow heavy resonance decaying to bottom quark-antiquark pairs at $\sqrt{s} = 13\,\text{TeV}$

The CMS Collaboration

Abstract

An inclusive search for a narrow resonance produced in proton-proton collisions and decaying to a bottom quark-antiquark pair is presented. A data sample collected at $\sqrt{s} = 13$ TeV and comprising 2.69 fb^{-1} of collision events recorded in 2015 with the CMS experiment at the CERN LHC has been analysed. The search focuses on the production of scalar resonances through gluon-gluon fusion, and of Randall-Sundrum gravitons, both with a negligible natural width relative to the experimental dijet mass resolution. A maximum likelihood fit to the dijet mass spectrum is performed to separate the signal from the continuum multijet background. No significant excess over the expectation from the background is observed. Limits on production cross sections times branching ratio are obtained for values of the resonance mass ranging from 550 to 1200 GeV.

1 Introduction

The observation of a Higgs scalar particle [1–3] with the LHC Run 1 proton-proton collision data [4, 5] motivates the search for additional high-mass resonances. Models with a second doublet of scalar fields are known as two-Higgs-doublet models (2HDM) [6] and feature the existence of a total of five scalar or pseudo-scalar resonances. The search for high-mass resonances is also motivated by models with extended gauge symmetries, from which heavy Z′ gauge bosons could arise [7], and by models with additional space-like dimensions, where spin-2 resonances would appear as graviton excitations [8]. A benchmark for possible spin-2 graviton productions is provided by the Randall-Sundrum (RS) models [9] where the mass separation between graviton excitations is large enough to allow for their independent observation [10]. While the additional scalar Higgs boson in 2HDM models is assumed to be mostly produced through gluon-gluon and $b\bar{b}$ fusion, the Z′ and the spin-2 graviton can be also generated through light quark-antiquark annihilation. In all cases, decays to a bottom quark-antiquark pair ($b\bar{b}$) can occur with high rate.

This document describes the search for the decay of a narrow boson resonance X into $b\bar{b}$ pairs for values of the resonance mass (m_X) in the range 550-1200 GeV. The present analysis extends previous results on dijet searches at $\sqrt{s} = 13$ TeV [11–13] by focusing on the range $m_X \lesssim 1$ TeV, which is experimentally challenging due to the high background rates. Dijet resonances in a similar mass range have been searched for in 8 TeV proton-proton collisions by using data collection techniques with reduced event content (scouting) [14], resulting in exclusion limits on the visible cross section for a 750 GeV resonance of about 0.8-2.0 pb depending on the relative fraction between quark and gluon decays. Owing to the reduced amount of information available in the event record, no attempt at tagging the jet flavour was made in Ref. [14]. A preliminary analysis of the $\sqrt{s} = 13$ TeV data with a similar scouting technique has been recently released by the CMS Collaboration [15], resulting in an upper limits of about 3 pb on the visible cross section for a 750 GeV resonance decaying to quarks. Differently from the previous studies, the analysis described in this document achieves the rate reduction required to store the full event content by deploying b tagging at the trigger level, thus primarily selecting $X \to b\bar{b}$ decays. A similar analysis of the $\sqrt{s} = 13$ TeV data has been recently presented by the ATLAS Collaboration [16]. This analysis utilises two benchmark spin-0 and spin-2 scenarios for modeling the signal features and for optimising the event selection. The search is based on proton-proton collision data collected at $\sqrt{s} = 13$ TeV in 2015 by the CMS experiment and corresponding to an integrated luminosity of $2.69 \pm 0.06\,\text{fb}^{-1}$ [17].

The general analysis strategy is to select events with a pair of highly energetic jets, which are further identified to originate from the hadronisation of b quarks. The online selection uses dedicated trigger paths, originally developed to search for signatures with up to four b jets, like those expected from heavy Higgs bosons produced via $b\bar{b}$ fusion and decaying to $b\bar{b}$ pairs [18]. The signal under consideration is expected to produce a peak in the dijet mass spectrum compared to the standard model (SM) backgrounds, which is dominated by multijet production and is expected to follow instead a smoothly falling function in the dijet mass. Besides a veto on isolated leptons, this analysis does not attempt at tagging extra particles in the event, thus preserving full efficiency also to gluon-gluon and quark-antiquark initiated signal events. Quantum interference with the non-resonant QCD production of $b\bar{b}$ pairs is neglected.

This document is organised as follows. Section 2 highlights the features of the CMS detector needed to perform this analysis. Section 3 details the production of simulated samples used to study the signal and the main backgrounds. The event reconstruction, trigger logic, and the offline selection are described in Sections 4, 5, and 6, respectively. The analysis strategy is

discussed in Section 7 and the associated systematic uncertainties in Section 8. The final results are presented in Section 9, and a summary is given in Section 10.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. A silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter are located within the axial field. Muons are measured in gas-ionisation detectors embedded in the steel flux-return yoke of the solenoid. Forward calorimetry ($3 < |\eta| < 5$) complements the coverage provided by the barrel ($|\eta| < 1.3$) and endcap ($1.3 < |\eta| < 3$) detectors. The first level (L1) of the CMS trigger system, composed of specialised processors, uses information from the calorimeters and muon detectors to select the most interesting events in a time interval of less than 4 μs. The high-level trigger (HLT) processor farm decreases the event rate from about 100 kHz to about 1 kHz, before data storage. A more detailed description of the CMS apparatus and the main kinematic variables used in the analysis can be found in Ref. [19].

3 Simulated samples

Samples of simulated signal and background events are employed to guide the analysis optimisation and to evaluate the signal acceptance. Several event generators are used to produce such samples. The NNPDF3.0 parton distribution functions (PDF) [20] are used for the Monte Carlo generation.

Samples of spin-0 and spin-2 signal events are generated at leading order (LO) with the PYTHIA 8 [21, 22] program, also for the parton shower (PS) and hadronisation processes and for the modelling of the underlying event, using the CUETP8M1 tune [23]. The signal samples are generated for values of m_X in the range 550-1200 GeV, and assume only $X \to b\bar{b}$ decays. For the spin-0 case, the signal samples are generated with a fixed width-over-mass ratio $\Gamma_X/m_X = 10^{-2}$, which is negligible compared to the dijet resolution of around 10% within the search region. For the spin-2 case, signal samples are generated with a RS effective coupling constant $\tilde{\kappa} = 0.1$ [9], corresponding to a width-over-mass ratio $\Gamma_X/m_X = 1.4 \cdot 10^{-2}$, again negligible compared to the experimental dijet mass resolution.

Background samples of QCD multijet events are simulated using the leading-order MADGRAPH 5 [24] generator interfaced with PYTHIA 8. Top-pair ($t\bar{t}$) and single-top quark productions are simulated using the POWHEG [25–29] program interfaced with PYTHIA 8. The production cross sections for the $t\bar{t}$ and single-top samples are rescaled to next-to-next-to-leading-order (NNLO) plus soft-gluon resummation at next-to-next-leading log (NNLL) cross sections calculations [30, 31].

To accurately simulate the LHC luminosity conditions during data taking, additional simulated pp interactions overlapping in the same or neighbouring bunch crossings of the main interaction, denoted as pileup, are added to the simulated events with a multiplicity distribution that matches the one in data. Small residual differences between the data pileup profile and the one contained in the simulated samples are accounted for by reweighing the simulated events.

4 Event reconstruction

The offline analysis uses reconstructed charged-particle tracks and candidates from the particle-flow (PF) algorithm [32–34]. In the PF event reconstruction all stable particles in the event, i.e. electrons, muons, photons, and charged and neutral hadrons, are reconstructed as PF candidates using information from all CMS subdetectors to obtain an optimal determination of their direction, energy, and type. The PF candidates are then used to reconstruct the jets and missing transverse energy. Among the vertices identified in the event, the one with highest $\sum p_{\mathrm{T}}^2$ and magnitude of $\sum p_{\mathrm{T}}$ from the associated physics objects is designated as primary vertex [35].

Jets are reconstructed by clustering PF candidates with the anti-k_{T} algorithm [36, 37] with a distance parameter of 0.4. Reconstructed jets require a small additional energy correction, mostly due to thresholds on reconstructed tracks and clusters in the PF algorithm and various reconstruction inefficiencies [38]. Jet identification criteria are also applied to reject misreconstructed jets resulting from detector noise, as well as jets heavily contaminated with pileup energy (clustering of energy deposits not associated with a parton from the primary pp interaction) [39]. The efficiency of the jet identification criteria is greater than 99%, with the rejection of 90% of background pileup jets with p_{T} of about 50 GeV.

The identification of jets that originate from the hadronisation of b quarks (b tagging) is performed by the CSVv2 algorithm [40, 41], which is also implemented at the HLT level, as described in Section 5. The CSVv2 algorithm combines the information from track impact parameters and secondary vertices identified within a given jet, providing a continuous discriminator output. When deployed at the trigger level, the algorithm makes use of regional tracking around the jets. Operative points of increasing signal purity are defined by requiring a minimum value of the b tagging output. The loose (CSVv2L), medium (CSVv2M), and tight (CSVv2T) operative points are used throughout the analysis. They correspond to a probability of about 85%, 70%, and 50% to identify b jets, with mistagging probabilities on light jets of about 10%, 1%, and 0.1%, respectively. Differences between the efficiencies and mistagging probabilities measured in data control samples and those predicted by the simulation are accounted for by reweighing the simulated events with scale factors that depend on the p_{T} and $|\eta|$ of the jets.

5 Triggers

The data events used for this analysis are collected by exploiting two dedicated trigger paths. Both paths use a common L1 selection (seed) of events with two jets, and a common HLT selection, differing only in the requirements on the two jet kinematics and on their b-tagging values.

The L1 seed requires the presence of at least two jets with $|\eta| < 2.6$ and with transverse momentum in excess of 100 GeV. The HLT event selection requires the presence of at least two calorimeter-based jets (CaloJets) with p_{T} in excess of 100 GeV and $|\eta|$ less than 2.5, two CaloJets with p_{T} in excess of 80 GeV and passing a minimum b-tagging selection, and two particle-flow jets (PFJets).

The first path applies a looser threshold on the online b-tagging output, and requires that both PFJets have p_{T} in excess of 160 GeV. The second path applies instead a tighter b-tagging selection but a looser p_{T} threshold of 100 GeV on both PFJets. In the latter case, the pseudorapidity distance between the two PFJets is further required to be smaller than 1.6. The efficiency of the loose (tight) online b-tagging selection for jets that pass the offline CSVv2T operative point

ranges between 90% (80%) and 95% (90%) for transverse momenta in the range 100-400 GeV. The analysis makes use of the inclusive combination of both trigger paths.

In order to measure the efficiency of the kinematic part of the triggers, two prescaled control paths are used, requiring one PFJet with p_T in excess of 60 and 80 GeV, respectively. Simulated trigger efficiencies are then corrected to match those measured in data by using scale factors that are parametrised as a function of the p_T and $|\eta|$ of the two jets. A control sample of events recorded by prescaled dijet triggers is then used to measure the efficiency of the online b-tagging part of the triggers. The efficiency measured on data is then compared to the one predicted by simulation, resulting in a scale factor of 0.89 with a 0.05 uncertainty coming from the limited size of the control sample.

6 Event selection

The offline event selection criteria are optimised on the simulated samples in order to maximise the analysis sensitivity to the presence of a narrow resonance decaying to $b\bar{b}$ pairs. The figure of merit for the optimisation is chosen to be the expected upper limit at 95% CL on cross section times branching ratio. More details on the limit setting procedure are provided in Section 9.

The final selection criteria define the signal region used for the search. An event is accepted in the signal region if it satisfies the following:

- it passes either one of the two trigger paths described in Section 5;
- it has at least two jets that pass the CSVv2M operative point, and at least one that passes the CSVv2T one;
- among the jets with $|\eta| < 2.4$ and p_T above 20 GeV, the two jets with the highest b-tagging output must in addition have p_T in excess of 100 GeV and a pseudorapidity distance ($\Delta\eta_{b\bar{b}}$) smaller than 1.6.

Events are vetoed if they contain one or two opposite-sign same-flavour isolated electrons or muons with p_T in excess of 20 GeV. No veto on the presence of isolated photons is applied.

About 1.3M events satisfy the selection criteria. Figure 1 (top-left) shows the acceptance times efficiency for the simulated $X \to b\bar{b}$ signal samples as a function of the resonance mass for the two spin assumptions. The drop in efficiency for increasing values of m_X is mostly caused by the degradation of the b tagging efficiency for high-p_T jets.

The two jets in the event with the highest b-tagging output are selected to be the two b-jet candidates. According to the simulation, this sorting prescription yields a correct event interpretation in over 95% of signal events for the resonance mass range considered in this analysis. In order to recover harder final-state radiation (FSR) outside of the b-jet cone, the four-momenta of any other jet with p_T in excess of 15 GeV and adjacent to the b-jet axis within an η-ϕ distance of 0.8 are added to the b-jet four-momentum when evaluating the invariant mass of the b-jet pair ($m_{b\bar{b}}$). The FSR-recovery algorithm was optimised for the SM Higgs boson searches in $b\bar{b}$ final states [42] and turns out to be also effective for the case of interest. For values of the resonance mass in the range considered in this analysis, the FSR-recovery algorithm induces a shift of the peak position of the dijet mass distribution of about 1-3% while reducing the full width at half maximum (FWHM) by about 15-30%. Figure 1 (top-right) compares the distribution of the dijet mass with and without the FSR-recovery algorithm for a few selected signal mass points.

A set of strictly looser selection criteria (preselection) is introduced for validation purposes. The preselection criteria are similar to the ones used to define the signal region, differing only

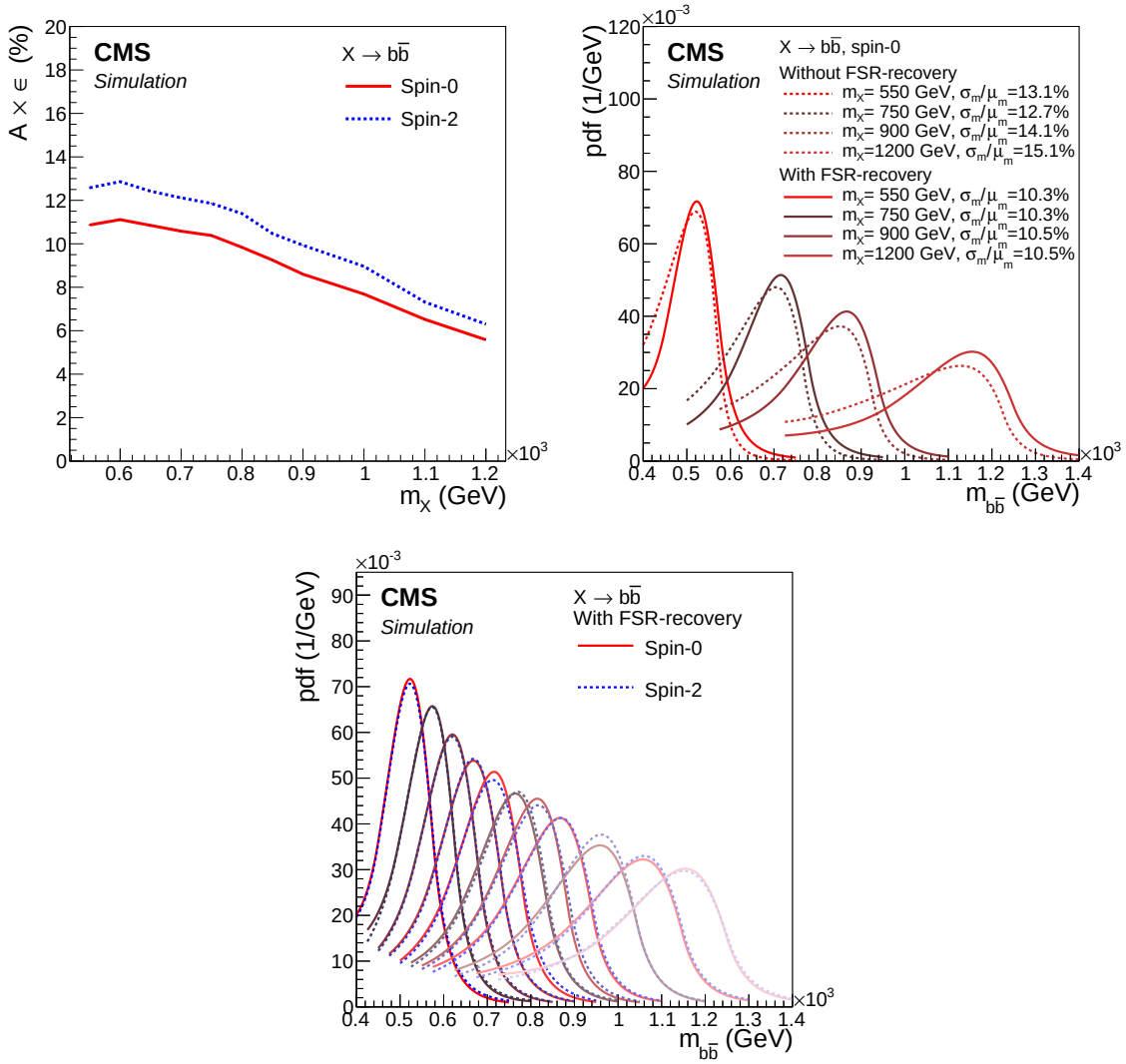

Figure 1: Top-left: acceptance (A) times reconstruction efficiency (ϵ) for the simulated $X \to b\bar{b}$ signal samples as a function of the resonance mass (m_X). For each value of m_X, the acceptance includes the requirement that the dijet mass is contained in the range utilised for the final signal extraction as described in Section 7. Top-right: comparison between the probability density function used to model the dijet mass ($m_{b\bar{b}}$) distribution for the signal in the case where $m_{b\bar{b}}$ is computed with (solid line) or without (dashed line) the FSR-recovery algorithm. The ratio σ_m/μ_m, where the σ_m is proportional to the full width at half maximum of the distribution and μ_m is the peak position, is also reported. Bottom: the same probability density functions for different values of the resonance mass and for the two spin hypotheses, after FSR recovery.

by the looser selection $\Delta\eta_{b\bar{b}} < 2.0$, and by the requirement that the b jet with the smallest b-tagging score has to pass at least the CSVv2L operative point. Figure 2 shows the distribution of a sample of kinematic variables for events that pass the preselection. The observed distributions are compared to the expectation from the multijet simulation scaled by a flat k-factor of 1.53. The latter is defined as the ratio between the data yield after subtracting the expected top contribution and the multijet background yield normalised to the LO cross section. The multijet simulation is not expected to provide a satisfactory modelling of the multijet kinematics for all relevant variables, in particular for $m_{b\bar{b}}$. Indeed, only a mild agreement between the observed distributions and the multijet simulation is observed for several jet-related variables, as shown in the top and middle panels of Fig. 2. It should be stressed that the lack of a proper modelling of the multijet background is not affecting the validity of the results since this analysis does not rely on the simulation to predict the relevant features of the background.

A control region where the simulation is instead expected to provide a satisfactory description of the data is defined by requiring the preselection criteria with inverted lepton veto, resulting in a sample of events that is more than 95% pure in $t\bar{t}$ and single-top events (top-quark sideband). The top-quark sideband can then be used to validate the emulation of the triggers and of the object reconstruction in the simulated signal samples. The observed event yield in this region amounts to 6626 events, compatible with an expectation of (6610 ± 100) from $t\bar{t}$ and single-top backgrounds. A good modelling of the data is observed for all relevant variables. As an example, the expected and observed distribution of the pseudo-rapidity difference between the b jets and of their invariant mass are reported in the bottom panels of Fig. 2.

7 Signal and background modeling

The $m_{b\bar{b}}$ distribution for the signal is modeled with an empirical probability density function (pdf) obtained from the convolution of a Gaussian with an exponential [43]. This model depends on five parameters, two of which are taken to be the peak position (μ_m) and the standard deviation of the Gaussian core ($\sigma_m = \mathrm{FWHM}/2\sqrt{2\ln(2)}$), while the other three are related to the modelling of the low- and high-mass tails. This pdf is found to provide a good modelling of both the core and the tails of the $m_{b\bar{b}}$ spectrum for all simulated signal samples. Figure 1 (bottom) shows the signal dijet mass pdf for different values of m_χ and for both spin assumptions. The peak position of the signal pdf is systematically shifted to lower values compared to the generated mass because of gluon radiation not caught by the FSR-recovery algorithm and because of semileptonic B hadron decays with undetected neutrinos. Discrepancies between the pdfs for signals generated at the same value of m_χ, but different spin numbers, can in principle arise from differences in the b-jet kinematics.

The $m_{b\bar{b}}$ distribution for background events that satisfy the final selection criteria features a global maximum around 250 GeV, with an inflection located about 100 GeV ahead and no resonant structure expected beyond that point. Therefore, the background functional form is modeled with a parametric smooth function, whose coefficients are treated as nuisance parameters and will be eventually determined by the fit. The usage of simulated samples to predict the background pdf is undesirable by both the lack of a suitable amount of events, i.e. larger than the observed data set, and by the observation that the LO multijet simulation does not provide an adequate modelling of the data, as shown in Section 6. In the absence of a theoretical prediction for the background pdf, different families of pdf's are studied, and the comparison between them drives the assessment of a systematic uncertainty due to the choice of a particular *ansatz* model. Similar to other high-mass searches [11, 44, 45], the following parametrisation:

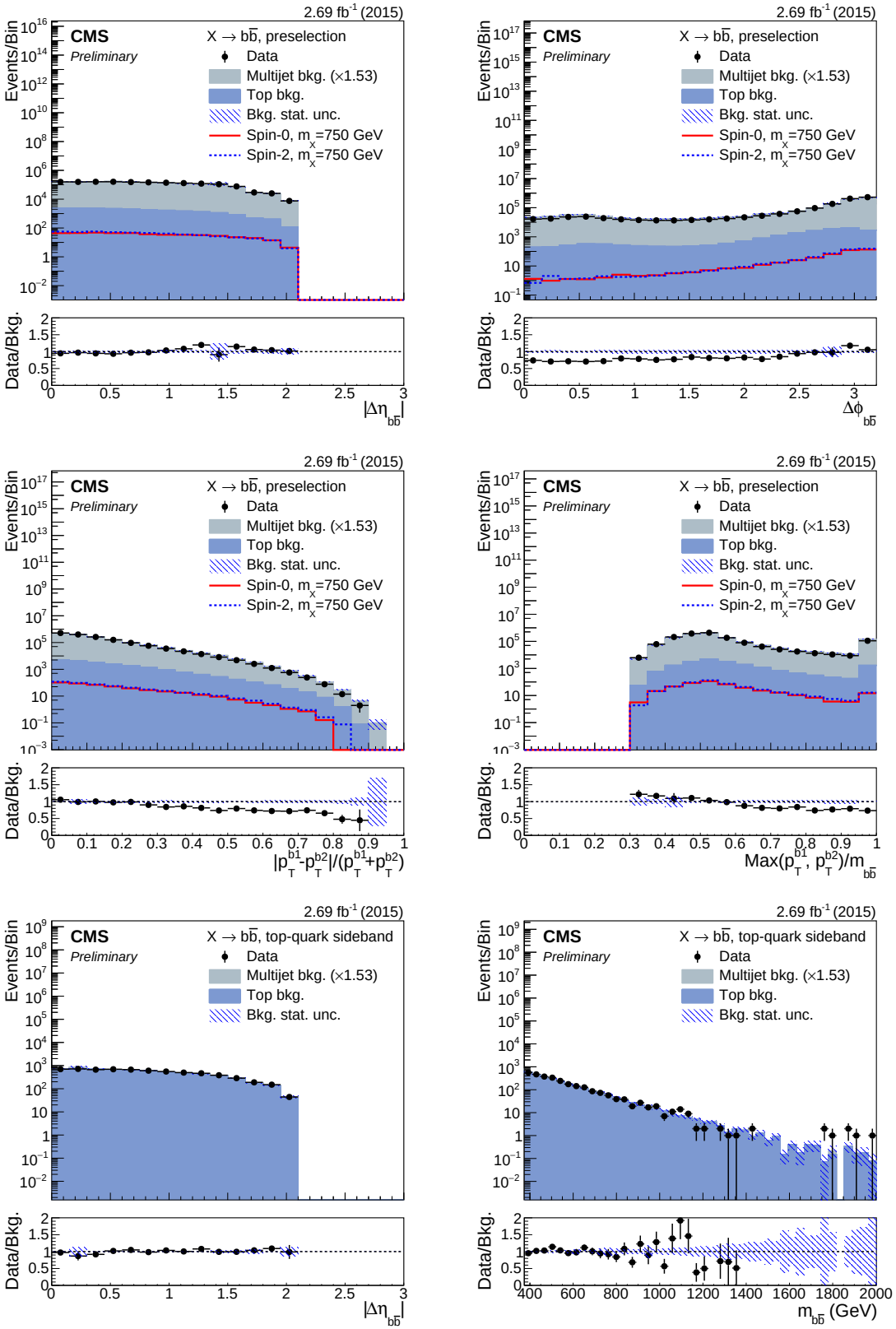

Figure 2: The distributions after preselection of (top-left) the pseudorapidity difference between the b jets, (top-right) their distance in azimuth angle, (centre-left) their p_{T} balance, and (centre-right) leading b-jet p_{T} in units of the dijet mass. For each plot, the expected contribution from a spin-0 (solid-red line) and spin-1 (dashed-blue line) signal with $m_{\mathrm{X}} = 750\,\mathrm{GeV}$, and normalised to a cross section of 1 pb, have been overlayed for comparison. Distributions of (bottom-left) the pseudo-rapidity difference between the b jets, and (bottom-right) the dijet mass in the top-quark sideband.

$$\frac{x^{p_0+p_1\ln(x)}}{\int_{x_L}^{x_H}\xi^{p_0+p_1\ln(\xi)}d\xi}, \tag{1}$$

where $x = m_{b\bar{b}}/\sqrt{s}$, is used as the nominal background pdf. In formula (1), p_0 and p_1 are real-valued coefficients and $x_{L(H)}$ are the lower (upper) ranges of the fit range such that the pdf is properly normalised. It has been checked that this pdf provides a good modelling of the simulated mass distributions, albeit with large statistical uncertainty. Extensions of formula (1) are obtained by including a multiplicative polynomial modulation, i.e. $x^{p_0} \to (\sum_{k\neq 1}^{n} p_k x^k)$, or by expanding the exponent into a series of $\ln(x)$, i.e. $x^{p_1\ln(x)} \to x^{\sum_{k=1}^{n} p_k \ln^k(x)}$. In both cases, the number of coefficients is determined by comparing the maximum likelihood values obtained for a fit with n and $n+1$ coefficients, where the asymptotic property of the log-likelihood ratio [46] are used to assess whether the inclusion of the additional coefficient is statistically significant. A p-value of 5% is conventionally chosen for the convergence of the test. It is found that a linear polynomial modulation is required only for the lowest fit range, while no need for a cubic or higher-order terms in $\ln(x)$ is found.

Three different family of pdf's are studied as alternative background models: the power of a polynomial, the exponential of a polynomial, and the product of a polynomial times an exponential. In all cases, the maximum number of coefficients is determined by the same procedure used for the nominal pdf. For each family, a systematic bias on the extraction of a signal with mass m_X is determined using toy experiments. Prior to the toy generation, each alternative pdf is fitted to the observed data. Toy experiments are then generated by sampling the alternative pdf thus obtained with Monte Carlo techniques. For each toy experiment, an extended maximum-likelihood fit under the signal-plus-background hypothesis is performed by using the nominal pdf as background model, resulting in a estimate N_S^{fit} of signal events with statistical uncertainty σ_S^{fit}. The mean value of the pull distribution obtained from the toy experiments, where the pull is defined as the ratio N_S^{fit}/σ_S^{fit}, is considered as a systematic bias associated with the alternative model under consideration. It has been verified that an unbiased estimation of the signal normalisation is instead obtained when using the nominal pdf both for generating and fitting toy experiments that include a signal contribution of variable strength.

Biases in excess of one are observed when a search range much larger than the width of the signal distribution is fitted. For this reason, the search ranges are adjusted as a function of the resonance mass under test. The lower and upper edges of the ranges are parametrised as

$$x_{L(H)}(m_X) = \left[a_{L(H)}m_X + b_{L(H)}\right]/\sqrt{s}, \tag{2}$$

where $a_{L(H)} = 0.50\ (1.0)$ and $b_{L(H)} = 125\ (200)$ GeV are chosen such that the resulting range corresponds to roughly the same quantile of the signal pdf at each value of m_X. Although still sizable, the biases obtained with this choice of the fit range are found to be $\lesssim 100\%$ for all families of alternative pdf's and across all values of m_X. In the likelihood function used to extract the measurement, the potential bias is accounted for by including an additional contribution on the background with identical shape as the signal at a given m_X. The size of this contribution is constrained, using a Gaussian constraint, to be within 100% of the statistical uncertainty of the fitted signal yield at that value of m_X.

8 Systematic uncertainties

The sources of systematic uncertainty related to both the background and the signal modelling are summarised in Table 1. The leading systematic uncertainty comes from the background description. In particular, the uncertainty associated with the choice of the $m_{\mathrm{b\bar{b}}}$ pdf for the background has the largest impact on the analysis sensitivity.

An uncertainty of 2.7% is assigned to the total integrated luminosity measurement [17]. The estimation of the pileup distribution in data depends on the assumed value of the inelastic pp cross section. The effect of varying this cross section by $\pm 5\%$ is propagated to the signal yields, resulting in a negligible impact.

The trigger efficiency in the simulation is corrected by using data-to-simulation scale factors as described in Section 5. Since the decay of the high-mass resonances considered in this analysis gives rise to highly energetic jets, which are typically on the plateau of the trigger turn-on curve, the impact of the jet energy trigger correction is found to be sub-dominant. The efficiency of passing the online b tagging filters is instead quite uniform versus the jet energy and is treated as a constant normalisation uncertainty for all signal samples.

The experimental uncertainties on the jet energy scale (JES) and jet energy resolution (JER) affect the signal acceptance and the shape of the signal pdf. Their impact is assessed by varying the JES and JER in the simulated samples within their measured uncertainties [38] and recomputing the acceptance and mass pdf accordingly. These variations affect the acceptance by 1-2%, while the μ_{m} parameter is shifted by 1% and σ_{m} by 2-5%. An additional uncertainty on the signal modelling is assigned by summing in quadrature the JEC (JER) variations to the statistical uncertainty on μ_{m} (σ_{m}) obtained from the fit of the simulated distributions.

Additional uncertainties are assigned to the efficiency of tagging the b jets by changing the b-tagging scale factors within their measured uncertainty. The corresponding impact on the signal acceptance is estimated to be about 6-7% depending on m_{X}. The lepton veto selection has a sub-percent efficiency on the signal samples, and no systematic uncertainty is therefore assigned with it.

Theoretical uncertainties affect the prediction of the signal acceptance by the simulation. The impact of the PDFs and strong coupling constant (α_{S}) uncertainties are estimated for the particular kinematic phase space of the search comparing the results to those obtained when using the Monte Carlo replicas of the NNPDF set; the total PDF and α_{S} uncertainty amounts to about 6%.

9 Results

The number of potential signal events in the selected data sample is extracted by performing a binned maximum likelihood fit to the $m_{\mathrm{b\bar{b}}}$ distribution. In the fit, the background coefficients and the total background normalisation are all treated as free parameters, whereas the signal model coefficients are constrained by Gaussian priors added to the log-likelihood function. Uncertainties that affect the signal normalisation are included in a similar way, but using log-normal priors. A uniform binning with a width of 0.1 GeV, which is negligible compared to the dijet mass resolution, is used for all mass windows. Figures 3 and 4 report the results of the fit in the various windows. No significant excess over the background-only expectation is observed. Upper limits at the 95% confidence level (CL) on the cross section times branching ratio ($\sigma \times \mathrm{BR}(\mathrm{X} \to \mathrm{b\bar{b}})$) are set using the asymptotic $\mathrm{CL_s}$ method [47–49]. The test-statistic chosen is a profile likelihood ratio in which systematic uncertainties are modelled as nuisance

Table 1: Summary of the systematic uncertainties affecting the signal and background modelling. The last column indicates whether a given source of systematic uncertainty affects the shape of the $m_{b\bar{b}}$ distribution.

Source	Size	Shape
	Signal	
Luminosity	2.7%	No
Pileup reweighing	$\lesssim 1\%$	No
Jet kinematic trigger	1%	No
Online b-tagging	5%	No
Jet energy scale	1-2%	Yes
Jet energy resolution	10-15%	Yes
Offline b-tagging	6-7%	No
PDF + α_S	6%	No
	Background	
Choice of pdf	100% σ_S^{fit}	Yes
Background parameters	-	Yes

parameters [50]. The nuisance parameter featuring the largest impact on the sensitivity is the one related to the choice of the $m_{b\bar{b}}$ pdf for the background. For example, its inclusion into the likelihood function makes the expected limit at $m_X = 750\,\text{GeV}$ about 35% looser relative to the case where no bias uncertainty is considered.

Depending on the resonance mass and on its spin, upper limits in the range 2-11 pb are obtained. The results are summarised in Figure 5. The expected LO cross section times branching ratio for graviton production in the RS model [9] with $\tilde{\kappa} = 0.1$ is overlayed for illustration. Graviton masses in the range 550-680 GeV would be excluded at the 95% CL for this scenario.

10 Summary

A search for a narrow resonance with mass ranging from 550 to 1200 GeV, and decaying to bottom quark-antiquark pairs using $2.69\,\text{fb}^{-1}$ of collision events collected at $\sqrt{s} = 13\,\text{TeV}$ in 2015 by the CMS experiment has been presented. Events are recorded by dedicated trigger paths that require the coincidence of a pair of high-p_T and b-tagged jets. The offline selection is further optimised to enrich the data sample in events compatible with a narrow resonance decaying into b jets. The signal yield is extracted by performing a fit to the dijet mass distribution, where an analytical parametrisation of the continuum background is assumed. No significant excess is observed over the background-only expectation. Results are presented in the form of 95% CL limits on $\sigma \times \text{BR}(X \to b\bar{b})$ for two benchmark scenarios: a spin-0 resonance produced in gluon-gluon fusion, and a spin-2 Randal-Sundrum graviton, both with a negligible width-over-mass ratio. Upper limits ranging from 2 to 11 pb are obtained depending on the resonance mass and on its spin. This analysis extends previous results on dijet searches at $\sqrt{s} = 13\,\text{TeV}$.

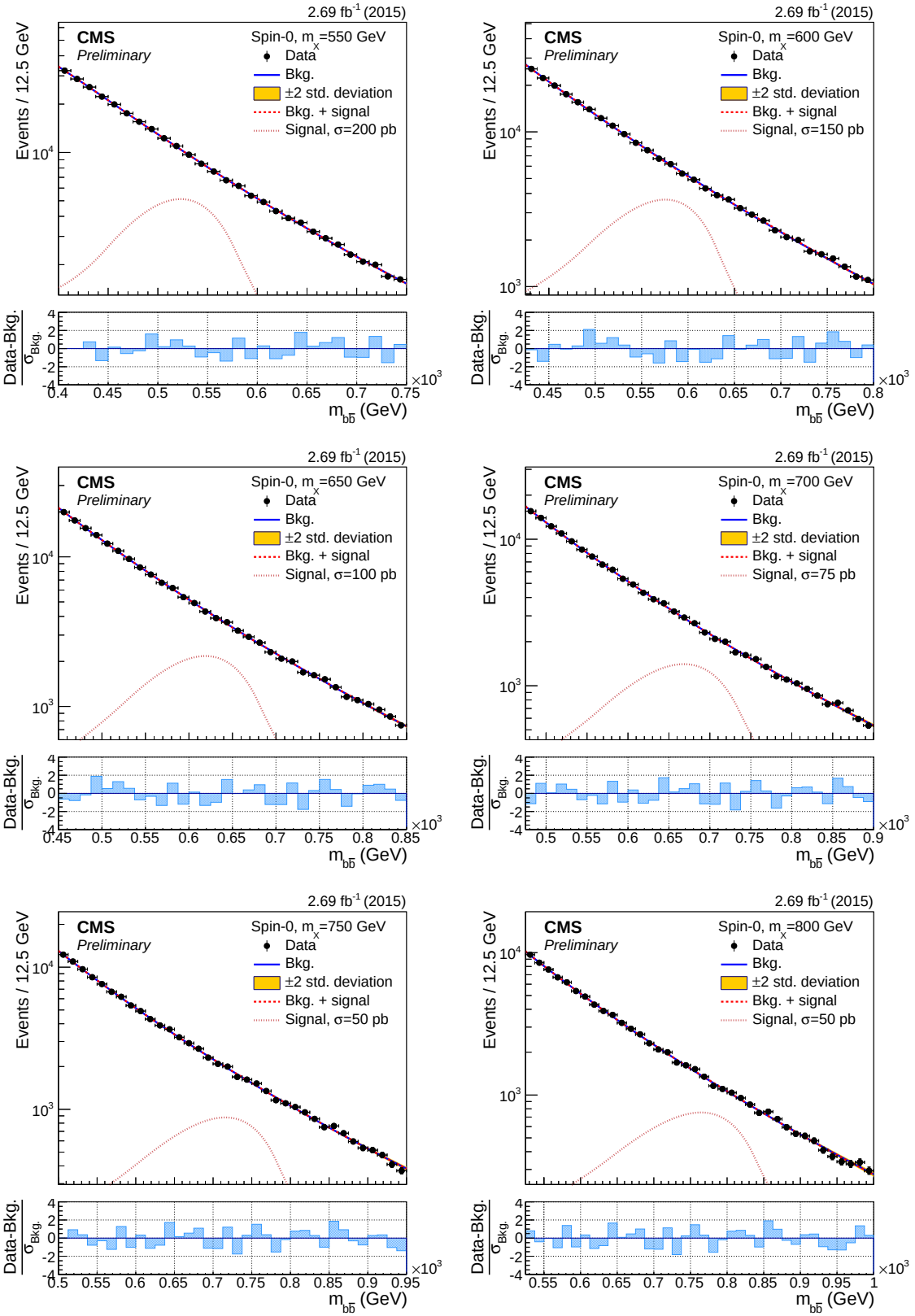

Figure 3: The dijet mass ($m_{b\bar{b}}$) distributions in the ranges used to search for a resonance with mass $m_X = 550, 600, 650, 700, 750, 800, 850\,\text{GeV}$. The results of the background-only fit are shown as a solid-blue line, while the signal plus background fit is shown as a dashed-red line. For illustrative purposes, the expected mass distribution for a spin-0 signal normalised to a large and arbitrary cross section is overlayed as a dotted light-red line. The bottom panels show the difference between the event counting and the fitted background yield in each bin, normalised to the total background uncertainty ($\sigma_{\text{Bkg.}}$) in that bin.

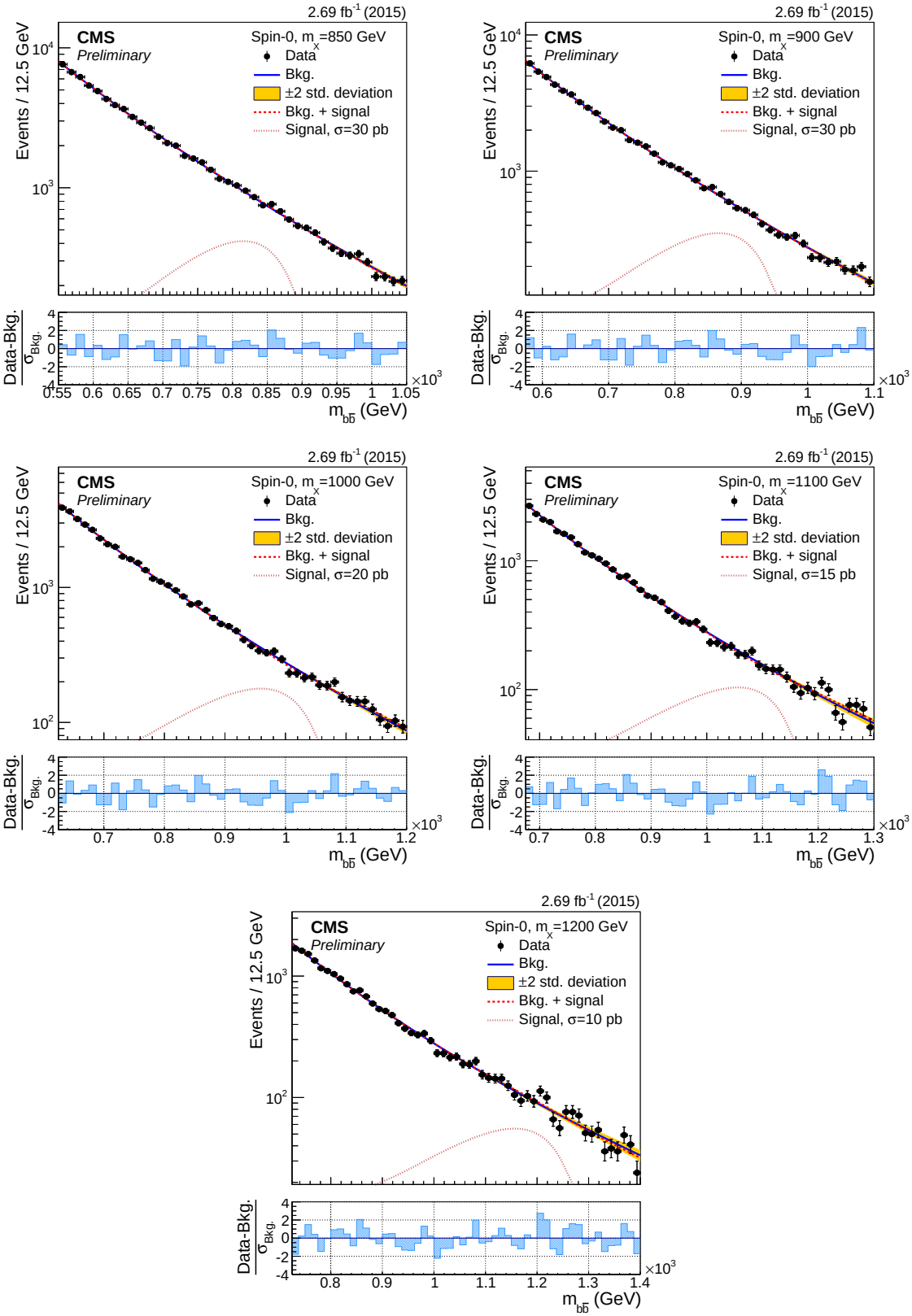

Figure 4: The dijet mass ($m_{b\bar{b}}$) distributions in the ranges used to search for a resonance with mass $m_X = 850, 900, 1000, 1100, 1200\,\text{GeV}$. The results of the background-only fit are shown as a solid-blue line, while the signal-plus-background fit is shown as a dashed-red line. For illustrative purposes, the expected mass distribution for a spin-0 signal normalised to a large and arbitrary cross section is overlayed as a dotted light-red line. The bottom panels show the difference between the event counting and the fitted background yield in each bin, normalised to the total background uncertainty ($\sigma_{\text{Bkg.}}$) in that bin.

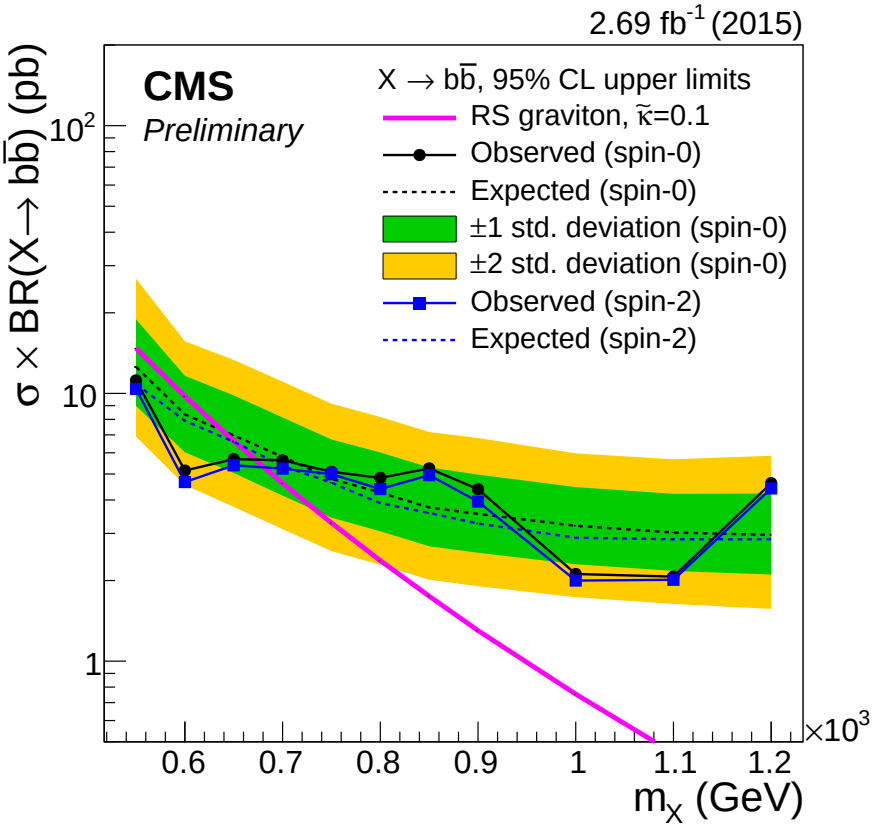

Figure 5: The expected and observed upper limits at 95% CL on $\sigma \times \mathrm{BR}(\mathrm{X} \to \mathrm{b\bar{b}})$ as a function of the resonance mass for the spin-0 (black) and spin-2 (blue) cases. The expected cross section times branching ratio for graviton production in the RS model [9] with $\tilde{\kappa} = 0.1$ is overlayed for illustration. The one- and two-standard deviation confidence interval bands for the background-only expected limit in the spin-0 case are shown as solid coloured bands. The corresponding limits on the visible cross section $\sigma \times \mathrm{BR}(\mathrm{X} \to \mathrm{b\bar{b}}) \times \mathrm{A} \times \epsilon$ can be obtained by using the acceptance-times-efficiency values reported in Fig. 1 (top-left).

References

[1] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321.

[2] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508.

[3] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particless", *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585.

[4] ATLAS and CMS Collaboration, "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.

[5] ATLAS, CMS Collaboration, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", doi:10.1103/PhysRevLett.114.191803, arXiv:1606.02266.

[6] N. Craig, J. Galloway, and S. Thomas, "Searching for Signs of the Second Higgs Doublet", arXiv:1305.2424.

[7] G. Langacke, "The Physics of Heavy Z' Gauge Bosons", *Rev. Mod. Phys* **81** (2009) 1199, doi:10.1103/RevModPhys.81.1199, arXiv:0801.1345.

[8] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "The Hierarchy problem and new dimensions at a millimeter", *Phys. Lett.* **B429** (1998) 263–272, doi:10.1016/S0370-2693(98)00466-3, arXiv:hep-ph/9803315.

[9] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension", *Phys. Rev. Lett.* **83** (1999) 3370–3373, doi:10.1103/PhysRevLett.83.3370, arXiv:hep-ph/9905221.

[10] 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.

[11] CMS Collaboration, "Search for narrow resonances decaying to dijets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phy. Rev. Lett.* **116** (2016) 071801, doi:10.1103/PhysRevLett.116.071801, arXiv:1512.01224.

[12] ATLAS Collaboration, "Search for resonances in the mass distribution of jet pairs with one or two jets identified as b-jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Physics Letters B* **759** (2016) 229–246, doi:10.1016/j.physletb.2016.05.064, arXiv:1603.08791.

[13] ATLAS Collaboration, ""Search for New Phenomena in Dijet Mass and Angular Distributions from pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector"", *Physics Letters B* **754** (2016) 302–322, doi:10.1016/j.physletb.2016.01.032, arXiv:1512.01530.

[14] CMS Collaboration, "Search for narrow resonances in dijet final states at $\sqrt{s} = 8$ TeV with the novel CMS technique of data scouting", *Phy. Rev. Lett.* **117** (2016) 031802, doi:10.1103/PhysRevLett.117.031802, arXiv:1604.08907.

[15] CMS Collaboration Collaboration, "Search for high-mass resonances in dijet final state with 2016 data", CMS Physics Analysis Summary CMS-PAS-EXO-16-032, 2016.

[16] ATLAS Collaboration Collaboration, "Search for resonances below 1.2 TeV from the mass distribution of b-jet pairs in proton-proton collisions at $\sqrt{s}$=13 TeV with the ATLAS detector", ATLAS Conference Note ATLAS-CONF-2016-031, 2016.

[17] CMS Collaboration, "CMS Luminosity Measurement for the 2015 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-15-001, 2016.

[18] CMS Collaboration, "Search for neutral MSSM Higgs bosons decaying into a pair of bottom quarks", *JHEP* **11** (2015) 071, doi:10.1007/JHEP11(2015)071, arXiv:1506.08329.

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

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

[21] T. Sjöstrand, S. Mrenna, and P. Skands, "A brief introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[22] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[23] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2015) 115, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

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

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

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

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

[28] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1007/JHEP02(2010)011,10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP02,011(2010)].

[29] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J.* **C71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[30] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through O(α_S^4)", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[31] N. Kidonakis, "Differential and total cross sections for top pair and single top production", (2012). arXiv:1205.3453.

[32] CMS Collaboration, "Particle–flow event reconstruction in CMS and performance for jets, taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[33] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[34] CMS Collaboration, "Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.

[35] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[36] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

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

[38] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[39] CMS Collaboration, "Pileup jet identification", CMS Physics Analysis Summary CMS-PAS-JME-13-005, 2013.

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

[41] CMS Collaboration Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, 2016.

[42] CMS Collaboration Collaboration, "Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$ with proton-proton collisions at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-16-003, 2016.

[43] A. Bukin, "Fitting function for asymmetric peaks", (2007). arXiv:0711.4449.

[44] CMS Collaboration, "Search for resonant production of high-mass photon pairs in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV", arXiv:1606.04093.

[45] ATLAS Collaboration, "Search for resonances in diphoton events at $\sqrt{s}$=13 TeV with the ATLAS detector", arXiv:1606.03833.

[46] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses", *Ann. Math. Statist.* **9** (1938) 60.

[47] T. Junk, "Confidence level computation for combining searches with small statistics", *Nulc. Instum. Meth.* **A434** (1999) 435–443, doi:doi:10.1016/S0168-9002(99)00498-2, arXiv:9902006.

[48] A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys.* **G28** (2002) 2693–2704, doi:doi:10.1088/0954-3899/28/10/313.

[49] 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, 10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727.

[50] ATLAS and CMS Collaboration, "Procedure for the LHC Higgs boson search combination in Summer 2011", CMS Physics Analysis Summary CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.