CMS Physics Analysis Summary

Contact: cms-pag-conveners-b2g@cern.ch 2017/07/11

Search for a singly produced vector-like quark B decaying to a b quark and a Higgs boson in a fully hadronic final state using boosted topologies

The CMS Collaboration

Abstract

A search is presented for the single production of a heavy vector-like quark (B) decaying to a Higgs boson and a bottom quark, B $\rightarrow$ Hb, with the Higgs boson decaying to a pair of bottom quarks. The decay products of the Higgs boson are highly boosted, hence typically collimated. They are reconstructed as a single, massive jet, with heavy flavour content. The single production of vector-like quarks is characterised by the presence of a light flavour quark emitted in the forward region of the detector. The analysis is performed using a data sample collected in 2016 by the CMS experiment at the LHC in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 35.9 fb^{-1}. The observation is consistent with background expectation and upper limits are placed on the production cross section times the branching ratio of a vector-like quark B decaying to a Higgs boson and a bottom quark. Values of cross section times branching ratio above 0.07–1.28 pb are excluded at 95% confidence level for masses of 700–1800 GeV, assuming a resonance with negligible width with respect to experimental resolution. Similar sensitivity is observed for different assumptions on the intrinsic width of the vector-like quark B.

1 Introduction

With the discovery of the Higgs boson by the ATLAS [1] and CMS [2] experiments, the standard model (SM) of particle physics confirmed its ability to predict experimental observations with high precision. However, the SM does not address problems related to the nature of the electroweak symmetry breaking, such as the hierarchy problem. An extension of the SM can address this problem by introducing new particles that play an important role in the cancellation of loop corrections to the Higgs mass [3]. Supersymmetric theories propose bosonic partners of the top quark in order to address the hierarchy problem; others such as little Higgs [4–6], or composite Higgs models [6, 7] overcome the hierarchy problem by introducing heavy fermionic resonances called vector-like quarks (VLQs) [3, 8–11]. Unlike SM chiral fermions, these particles do not receive mass through a Yukawa coupling term, and are not excluded by present searches, while the existence of a fourth generation of SM quarks has been ruled out by electroweak precision measurements [12] and the discovery of the Higgs boson [11]. Previous searches for VLQs have been performed by the ATLAS [13–19] and CMS [20–27] experiments using proton-proton collision data recorded at centre-of-mass energies of 7, 8 and 13 TeV.

In this note, we present the first search at the LHC for electroweak production of a single VLQ B with charge $-1/3e$, decaying to a bottom quark and a Higgs boson. The search uses proton-proton collision data collected in 2016 by the CMS experiment at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 35.9 fb^{-1}. The fully hadronic final state with the Higgs boson decaying to a pair of b quarks is considered. Figure 1 illustrates the electroweak production of a B quark, where the new particle is produced in association with a bottom quark. The B quark could also be produced in association with a top quark, although with a smaller cross section.

The B decay channel considered in this analysis is B$\rightarrow$Hb, however, the B quark can also decay into Zb or Wt with branching ratios depending on the specific model. Under the hypothesis of a Bquark being a singlet of the SM, the branching fractions for these three decay modes are 0.25, 0.25 and 0.50, for the B decaying to Hb, Zb and Wt, respectively.

The cross section for single production of the B quark depends on its mass as well as on its width, and therefore on its coupling to SM particles. The benchmark model considered in this note assumes a coupling coefficient to the Z boson and the b quark, c(bZ), equal to 0.5. Under this assumption the intrinsic resolution of the signal is negligible with respect to the experimental resolution and the approximation of narrow width is valid. For the first time the case in which the B quark has a non negligible resonance width is also considered, with a value up to 30% of the resonance mass. Previous CMS searches for VLQ B relied on narrow width approximation only.

In this note, Section 2 describes the CMS detector and the reconstruction of physics objects; Section 3 gives an overview of the data and simulations used in this analysis. Section 4 and 5 provide a description of the analysis event selection and of the background estimation method. Systematic uncertainties are discussed in Section 6. The results of the search are presented in Section 7.

2 CMS detector and physics object reconstruction

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintilla-

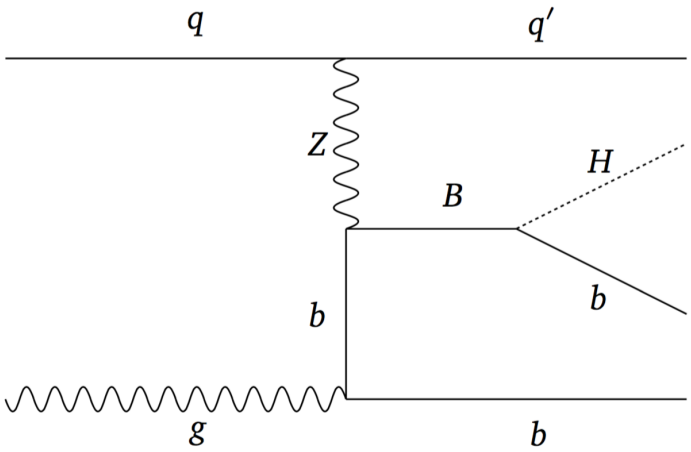

Figure 1: Leading order Feynman diagram for the production of a single vector-like quark B in association with a b quark, and its decay to an H boson and a b quark.

tor hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [28].

Events of interest are selected using a two-tiered trigger system [29]. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a time interval of less than 4 μs. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to less than 1 kHz before data storage.

Event reconstruction is based on the CMS Particle Flow (PF) algorithm [30]. It reconstructs and identifies each individual particle with an optimized combination of information from the various elements of the CMS detector. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with the origin of the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energy. Jets are reconstructed from the clustering of PF candidates by means of the anti-k_{T} algorithm [31] with a distance parameter of 0.4 and 0.8, referred to as AK4 and AK8 jets, after passing the charged hadron subtraction (CHS) pile up mitigation algorithm [32]. The CHS algorithm discards charged hadrons not originating from the primary vertex. The reconstructed vertex with the largest value of summed physics-object transverse momentum squared, p_{T}^2, is taken to be the primary pp interaction vertex. The physics objects are the objects returned by a jet finding algorithm [31, 33] applied to all charged tracks associated with the vertex, plus the corresponding associated missing transverse momentum. The residual contamination removed is proportional to the event energy density and the jet area estimated using the FASTJET package [33]. Jet momentum is determined as the vectorial sum of all particle momenta in the jet. Energy scale calibrations (JES) determined

from simulation are then applied to correct the four momenta of the jets. Residual corrections, determined from collision data, are applied in data events only. The jet energy resolution (JER) for the simulated jets is degraded in order to reproduce the resolution observed in data. Jet candidates with cone radius of 0.4 in this analysis are required to have $p_T > 30$ GeV and an absolute pseudorapidity, $|\eta|$, lower than 4, and to satisfy a tight set of identification criteria designed to reject spurious detector and reconstruction effects [34]. Larger cone jet candidates are selected with $p_T > 300$ GeV and $|\eta| < 2.4$.

The combined secondary vertex (CSVv2) b tagging algorithm [35] is used to identify jets arising from the hadronization of a b quark. In this analysis PF jets are considered as b tagged if the CSVv2 discriminant for the jet passes the medium working point requirement, for which the misidentification probability averaged over the jet kinematics in $t\bar{t}$ events is around 1% for light-flavour jets with p_T above 30 GeV while the corresponding tagging efficiency for b jets is around 70%.

Higgs boson jet candidates are identified by looking at the heavy-flavour content of the candidate jet. A pruning algorithm [36], used to remove soft and wide-angle radiation so that the mass of the hard constituents can be measured more precisely, is applied to AK8 jets. The mass of the jet, after the pruning is performed, is required to be in the mass window of [105, 135] GeV. Both subjets reconstructed inside the AK8 jet cone with the soft-drop algorithm [37] are required to pass the medium working point of the CSVv2 b tagging algorithm.

Finally, an interesting feature of the signal under study is the presence of an additional jet that is typically produced in the forward direction. Forward jets are reconstructed as AK4 jets with the same selection criteria and corrections as defined above, but with $|\eta| > 2.4$.

3 Physics Modelling and Simulation

The production of B $\rightarrow$ Hb provides a high jet-multiplicity signature rich in heavy-flavour content and characterised by a boosted Higgs boson. The dominant background in this search is from multijet production; additional minor backgrounds arise from $t\bar{t}$ process, as well as from the production of W and Z bosons in association with jets.

Simulated events are used throughout the analysis to define the selection strategy and determine its expected sensitivity. The background from multijet events is estimated with data in control regions. However, simulation is used to cross-check the method and evaluate its validity. The contributions from other background processes, such as $t\bar{t}$, Z and W boson production in association with jets, are estimated with Monte Carlo (MC) simulation.

Multijet, as well as electroweak background processes, Z /γ^*+jets and W+jets, are simulated at the leading order (LO) using the generator MADGRAPH_aMC@NLO 2.2.2 [38] interfaced to PYTHIA 8 [39] for the parton-shower simulation. The $t\bar{t}$ background MC samples are produced with POWHEG v2 [40–43] and interfaced with PYTHIA 8. The mass of the top quark is set to 172.5 GeV and the cross section is calculated at next-to-next-to-leading order and next-to-next-to-leading-logarithmic order (NNLO+NNLL) in perturbative QCD using the Top++v2.0 program [44]. The cross sections of Z /γ^*+jets and W+jets processes are calculated at next-to-next-to-leading order (NNLO) using FEWZ [45].

The signal process B $\rightarrow$ Hb $\rightarrow$ $b\bar{b}$ b is simulated using the leading order, LO, MC event generator MADGRAPH 5.2 interfaced to PYTHIA 8 for the parton-shower simulation. Signal cross sections have been calculated at leading order with approximate next-to-leading order corrections [46]. Several benchmark models have been considered with B quark mass ranging from

700 up to 1800 GeV in steps of 100 GeV. For values of coupling coefficient c(bZ) smaller than 0.5, the theoretical width of the new heavy resonance is negligible compared to the experimental resolution. Signal samples for B quarks with larger width (10%, 20% and 30% of the mass hypothesis) were also generated in the same mass range.

Simulations using LO and NLO calculations respectively use the LO and NLO NNPDF3.0 [47] Parton Distribution Function (PDF) sets. All signal and background samples are processed using GEANT 4 [48] to provide a full simulation of the CMS detector. The samples are reweighted to account for the distribution of additional pileup interactions observed in 2016 collisions.

4 Event selection

This analysis searches for a Higgs boson and a bottom quark arising from the decay of a B quark, exploiting the hadronic decay of the Higgs boson into a pair of bottom quarks. Given the large mass of the new particle, its decay products are expected to be produced with large transverse momentum, p_T. Therefore, the two bottom quarks, originating from the Higgs boson decay, emerge very close to each other, giving rise to a single large-cone jet.

The data used for this analysis are collected through an online selection based on the hadronic activity (H_T), defined as the scalar sum of the transverse momenta of all AK4 jets with $p_T > 30$ GeV and $|\eta| < 3.0$. The H_T threshold used for this trigger is 900 GeV. Data containing at least one jet reconstructed by the HLT with $p_T > 450$ GeV, are selected as well in order to increase the trigger efficiency. A H_T requirement applied at analysis level ($H_T > 950$ GeV, where the H_T is now computed using AK4 jets with $p_T > 50$ GeV and $|\eta| < 2.4$) ensures trigger efficiency better than 87%.

Events with three or more AK4 jets with $p_T > 30$ GeV and $|\eta| < 4.0$, among which at least one b tagged jet with $|\eta| < 2.4$, are selected. In addition, a veto is applied to events with one or more leptons to ensure that the selection criteria do not overlap with those used in possible future searches in leptonic final states. This selection will be referred to as preselection throughout the note. Selected events are further required to have at least one Higgs-tagged AK8 jet. To fulfill the Higgs-tagging requirement, a large radius jet must have a pruned mass between 105 and 135 GeV and two subjets identified as arising from b quark hadronization. The Higgs-tagging efficiency is $\sim$ 10–20%, depending on the B mass hypothesis. Figure 2 shows the b tagged subjet multiplicity expected for simulated background and signal processes compared to data.

The B quark is reconstructed from the Higgs-jet candidate along with the non-overlapping leading-p_T b tagged jet. The b quark from B decay is produced with high energy and selecting the b jet with the highest p_T significantly reduces the combinatorial background. Furthermore, in order to avoid overlaps with the decay products of the Higgs boson, a condition is applied on the spatial distance between the two objects, requiring ΔR(b,H) to be greater than 1.2, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$.

To further reduce the multijet background and the contamination arising from gluon-like jets, H_T is required to be larger than 950 GeV for low B quark mass values, $800 < m_B < 1500$ GeV. For high B quark masses, $1500 \leq m_B \leq 1800$ GeV, we only select events with $H_T > 1250$ GeV.

The signal-background discrimination is enhanced by exploiting the distinctive presence of a forward jet in signal events. Events are separated into categories based on the forward jet multiplicity. A high-purity category is obtained by requiring at least one jet in the forward region. A second category that contains a large fraction of events from both signal and background

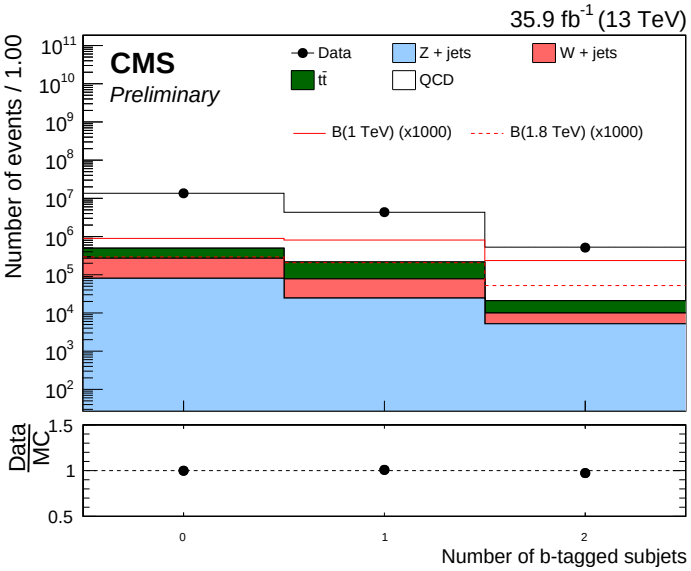

Figure 2: B tagged subjet multiplicity of large radius jets after the preselection and the low hadronic activity requirement are applied. The lower panel shows the data to background ratio. Signal distributions are multiplied by a factor of 1000. Background simulated events are normalised to data. The uncertainties on the background prediction and on data are statistical only.

processes is defined requiring exactly zero forward jets. The forward jet multiplicity expected for background and signal events at preselection, compared to data is presented in Fig. 3. The disagreement observed between data and simulation does not affect the measurement as the background contribution in the signal region is estimated from data. Moreover, the effect on signal events would be negligible as the measurement is dominated by events containing zero or one forward jets, for which the simulated and observed yields after preselection are consistent.

5 Signal extraction

A potential signal would manifest as a localised excess over background expectation in the reconstructed B quark mass, m_{B}, spectrum. To improve the sensitivity of the search to a B quark signal, this analysis exploits the additional information provided by the shape of the observed m_{B} distribution, as the signal is expected to peak around the mass of the B quark. A template fit is, hence, performed to the m_{B} distribution.

Multijet events costitute the dominant source of background for this search. An additional contribution arises from $\mathrm{t\bar{t}}$ events ($\sim$ 5–7%). To reduce the dependence on the multijet process modeling in simulation, the contribution from this background is derived from data. The ABCD method is used to estimate the expected yield of multijet background in the signal region, while the shape of its distribution in m_{B} is taken from a background enriched control region. A minor contribution ($\sim$ 1%) to background events arises from other SM processes, Z+jets and W+jets. Both $\mathrm{t\bar{t}}$ events and these minor backgrounds are estimated from simulation. The normalisation of multijet events in the signal region is estimated using four regions, out of which three are enriched in background events. These regions, in addition to the signal enriched one, are identified in a two-dimensional phase space defined by two variables that are

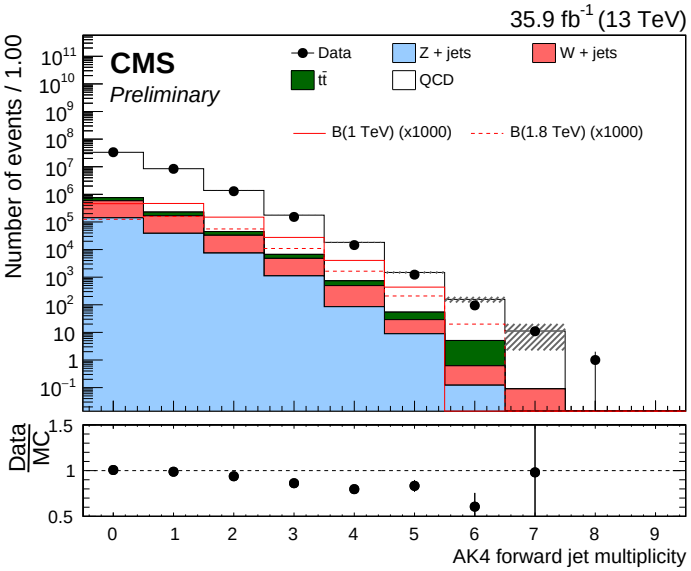

Figure 3: Forward jet multiplicity distribution before the event categorization is performed. Signal distributions are multiplied by a factor of 1000. Background simulated events are normalised to data. The lower panel shows the simulated background to data ratio. The uncertainties in the background prediction and on data are statistical only.

assumed to be uncorrelated: the b-tagged subjet multiplicity of the Higgs jet and its mass, m_{H}. The four regions are defined as follows:

- Region A: two b-tagged subjets and $105 < m_{\mathrm{H}} < 135$ GeV;
- Region B: two b-tagged subjets and $75 < m_{\mathrm{H}} < 105$ GeV or $m_{\mathrm{H}} > 135$ GeV;
- Region C: one b-tagged subjet and $105 < m_{\mathrm{H}} < 135$ GeV;
- Region D: one b-tagged subjet and $75 < m_{\mathrm{H}} < 105$ GeV or $m_{\mathrm{H}} > 135$ GeV.

Region A is the signal region, defined by the selection criteria described in the previous section. The multijet background event yield in the signal region is derived from regions B, C and D which are background-enriched. Assuming that the two variables, b tagged subjet multiplicity and Higgs-jet mass, are uncorrelated, the background event yields in the four regions follow the relation:

$$N_{\mathrm{A}}/N_{\mathrm{C}} = N_{\mathrm{B}}/N_{\mathrm{D}}, \tag{1}$$

where N_A, N_B, N_C, N_D are the yields in regions A, B, C and D respectively. Thus, the number of background events in the signal region A is given by:

$$N_{\mathrm{A}} = N_{\mathrm{C}} \times N_{\mathrm{B}}/N_{\mathrm{D}} \tag{2}$$

after subtracting the $\mathrm{t\bar{t}}$ contribution predicted by simulation from the observed yields. The contribution from Z +jets and W+jets is not subtracted as it is negligible.

The shape of the multijet background in the signal region is estimated from the event distribution in region C, since the reconstructed m_{B} spectrum is not expected to be correlated with the

b-jet multiplicity. The compatibility of the shapes in regions A and C is verified using simulated multijet events and cross-checked in data.

The reconstructed m_{B} distribution for the multijet background in region A, as estimated from the ABCD method, is found to be consistent with the prediction from simulation. Additionally, the method is validated using a signal-depleted region of high mass sidebands. Here, two regions A′ and C′ are defined, similarly to A and C, in the mass region $135 < m_{\mathrm{H}} < 165$ GeV. Two control regions B′ and D′, are defined requiring $75 < m_{\mathrm{H}} < 105$ GeV or $m_{\mathrm{H}} > 165$ GeV and with 2 or 1 b tagged subjets, respectively. The background distribution estimated in region A′, using the method described above, agrees well with the observation. The difference between the estimated and the observed yields in region A′ is used as a normalisation systematic uncertainty on the multijet estimate. In addition, the statistical uncertainty in region C due to the limited amount of data is considered as bin-by-bin shape variation on the final background estimate in the signal region.

6 Systematic uncertainties

The uncertainties on the event yield and the m_{B} distribution of signal and background processes are taken into account in the statistical inference procedure, and treated as nuisance parameters. The systematic effect of each source of uncertainty is evaluated by propagating the uncertainty on the input parameters to the reconstructed B quark mass distribution.

The uncertainty on the data-based background estimate for multijet process due to the limited amount of data and $\mathrm{t\bar{t}}$ simulation in the three control regions is propagated to the main observable, m_{B} in the signal region, by varying up and down ($\pm 1\sigma$) the observed event yields in regions B and D and propagating this shift to the expected background yield in the signal region. As the expected shape is estimated from region C, the statistical uncertainty of the observed m_{B} distribution in this region is also considered in the signal extraction. In addition to the normalization, this uncertainty affects the shape of the background m_{B} distribution in the signal region. Each bin of the observed m_{B} spectrum in region C is allowed to fluctuate independently according to its statistical uncertainty.

An additional systematic uncertainty on the estimated background yield is derived from the difference between the observation and the prediction found in the test performed on data and described in Section 5 ($\sim$ 5–10%).

The systematic uncertainties due to the limited size of the simulated signal samples as well as of background taken from simulation are included by allowing each bin of the m_{B} distribution to independently fluctuate according to Poisson statistics.

Additional systematic uncertainties on simulated signal and background distributions originate from the corrections applied to match simulated distributions to data, such as an uncertainty of 2.5% on the measured integrated luminosity, which affects the event yield only.

The corrections to account for the b-tagging efficiency difference between data and simulation are varied up and down by their uncertainties (4% to 15%) on both AK4 jets and subjets. The reconstructed jet four-momenta of both AK4 and AK8 are varied within the uncertainty on jet energy scale and resolution, and propagated to all observables, including m_{B}. In addition, the pruned mass scale and resolution of the Higgs-tagged jet is varied within uncertainties, affecting the m_{B} spectrum by $\sim$ 0.5–6.5%.

All simulated events are reweighted to match the distribution of pileup interactions in data.

The corresponding uncertainty is taken into account by varying the minimal bias cross section of 69 mb, used to calculate the data pileup distribution, by $\pm 4.6\%$. In this analysis, scale factors are applied to account for differences between the trigger efficiency measured in data and in simulated events. The uncertainty on these scale factors is applied as a function of and propagated to the m_{B} distribution affecting its shape.

An additional uncertainty is applied to account for discrepancies in the modelling of forward jet multiplicity. The magnitude of this effect is 0.5% for the category with zero forward jets and 2.0% for the one with at least one jet in the forward region.

The uncertainty on the choice of the factorization and renormalizarion scales, μ_F and μ_R, is taken into account by halving and doubling their values and taking the combination of μ_F and μ_R resulting in the maximum variation. Only the effects of acceptance is considered for signal events, leading to a variation up to the 1.3% depending on the mass hypothesis. Larger effects (15–25%) are observed on the simulated backgrounds considering the systematic effect on both normalization and acceptance. Furthermore, the uncertainty on the choice of PDFs is estimated by reweighting the samples with `NNPDF3.0` [47] replicas and propagated to the m_{B} distribution.

7 Results

A binned maximum likelihood fit is performed on the m_{B} distribution, Fig. 4, where the dominant multijet background is estimated from data, as discussed in Sec. 5. The observation is consistent with the background-only hypothesis in all the categories and upper limits are set on the production cross section times the branching fraction of a B quark produced in association with a bottom quark, and decaying to Hb, as a function of the B quark mass. The 95% confidence level exclusion limits are calculated using a modified frequentist approach (CLs) and a profile-likelihood ratio as test-statistic in an asymptotic approximation [49–51]. Signal and background templates are given by the distributions in Fig. 4. The two forward jet multiplicity based categories are combined to increase the sensitivity of the analysis. Their combination brings an improvement up to 20% with respect to just requiring at least one jet lying in the forward pseudorapidity region of the detector.

Systematic uncertainties described in Sec. 6 are treated as nuisance parameters affecting the rate or the shape of the expected m_{B} distribution of the concerned process. Uncertainties affecting normalisation only are modelled with a log-normal prior. For those uncertainties affecting shapes, the m_{B} distribution is allowed to vary using template morphing techniques [52].

Post-fit m_{B} distributions are presented in Fig. 5, while the expected yields are listed in Table 1 for the backgrounds and for two signal hypotheses together with observed yields, for the two signal mass categories.

The observed and expected combined upper limits from the two categories are shown in Fig. 6, for the singlet B quark production mode under the assumption of narrow width approximation. A coupling coefficient c(bZ)=0.5 and $\mathcal{B}(\mathrm{B} \to \mathrm{Hb}) = 25\%$ are considered. For this model we exclude values of cross section times branching fraction larger than 0.07–1.28 pb at 95% confidence level, for masses in the range 700–1800 GeV. Upper limits are compared with theoretical prediction calculated at NLO [46]. Figure 7 shows the observed and expected upper limits on the B quark production times branching fraction for a B quark with intrinsic width fixed to 10%, 20% and 30% of the resonant mass. Similar sensitivity is observed with respect to the case of negligible width. Observed exclusion limits vary between 0.08 and 1.97, 0.11 and

Table 1: Observed and expected event yields in the signal region in both low mass and high mass ranges. The multijet background expectation yields are derived from data, while the yields for the other sources of background are derived from simulation.

category	source	low mass yields	high mass yields
zero forward jets category	$t\bar{t}$	394 ± 46	117 ± 18
	W + jets	29 ± 13	10.5 ± 4.3
	Z + jets	43 ± 15	23 ± 23
	Multijet	5416 ± 60	1612 ± 24
	Backgrounds	5882 ± 42	1762 ± 26
	Data	5886 ± 77	1753 ± 42
	$m_B =$ 1000 GeV	29.3 ± 1.1	7.07 ± 0.53
	$m_B =$ 1800 GeV	4.88 ± 0.15	4.12 ± 0.14
at least one forward jet category	$t\bar{t}$	163 ± 20	58 ± 17
	W + jets	11.5 ± 4.2	4.3 ± 1.4
	Z + jets	2 ± 10	0.0 ± 0.0
	Multijet	1938 ± 23	549 ± 10
	Backgrounds	2115 ± 21	612 ± 15
	Data	2107 ± 46	608 ± 25
	$m_B =$ 1000 GeV	46.4 ± 1.4	8.51 ± 0.58
	$m_B =$ 1800 GeV	9.29 ± 0.21	7.78 ± 0.19

1.32, 0.10 and 1.22, respectively, for the 10%, 20% and 30% width/mass scenarios.

8 Summary

A search for electroweak production of a B quark with charge $-1/3e$, decaying to a b quark and a Higgs boson is presented. The analysis uses 35.9 fb^{-1} of proton-proton collision data collected in 2016 by the CMS experiment at $\sqrt{s} = 13$ TeV.

No significant deviations from SM expectations are observed and upper limits are set on the production cross section of a B quark. Expected (observed) limits vary between 0.07 (0.07) pb and 1.20 (1.28) pb at 95% confidence level, for the range of resonance mass considered. The search has been performed under the hypothesis of a singlet B quark decaying to Hb with branching ratio equal to 25% and narrow width resonance. The hypothesis of non-negligible width has also been studied, by considering production of a B quark with an intrinsic width of 10%, 20% and 30% of the resonance mass.

References

[1] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

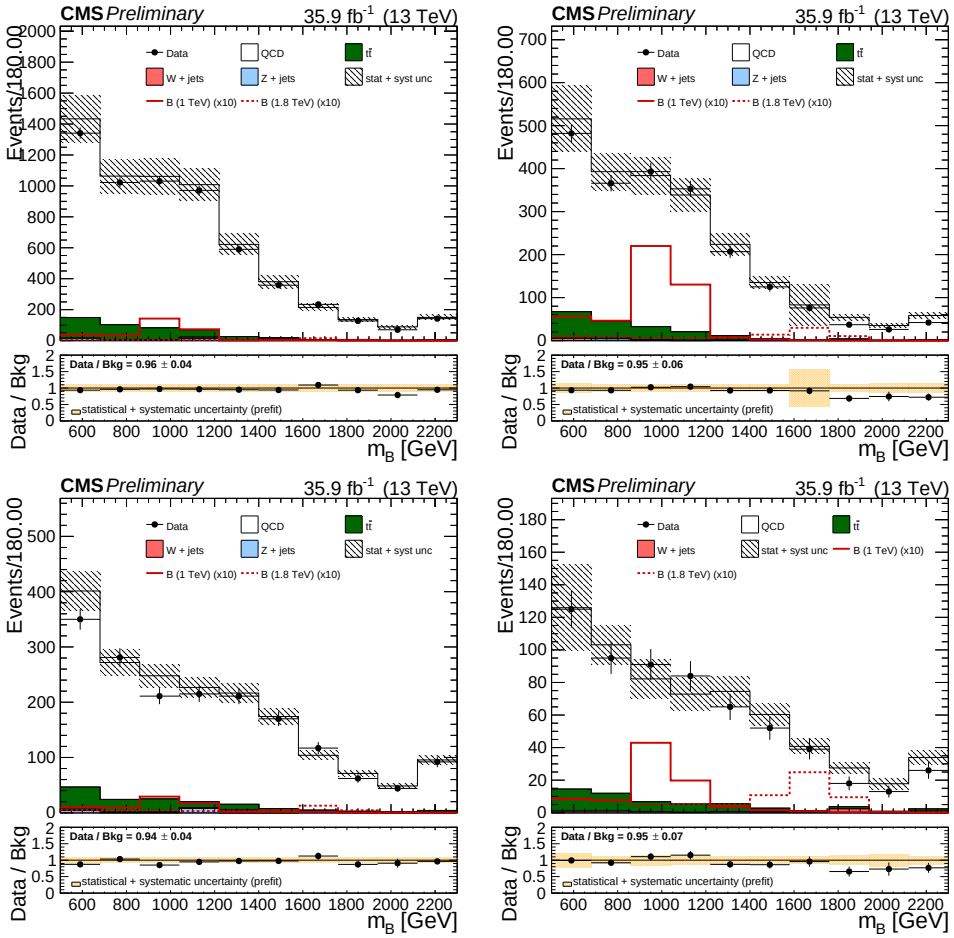

Figure 4: Pre-fit distributions of the reconstructed B quark mass in the categories with zero forward jets (left) and at least one forward jet (right) for the low (top) and high (bottom) mass analyses. The shaded error band in the ratio plot shows both statistical and systematics uncertainties.

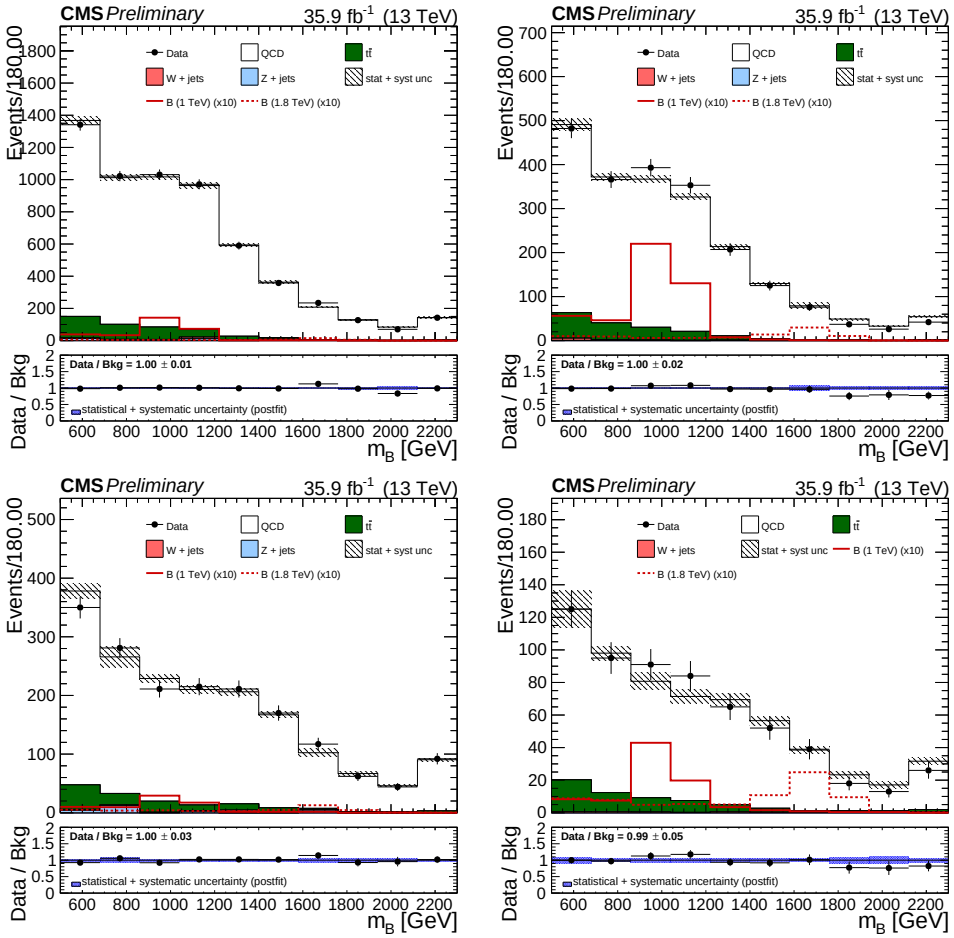

Figure 5: Post-fit distributions of the reconstructed B quark mass in the categories with zero forward jets (left) and at least one forward jet (right) for the low (top) and high (bottom) mass analyses. The shaded error band in the ratio plot shows both statistical and systematics uncertainties.

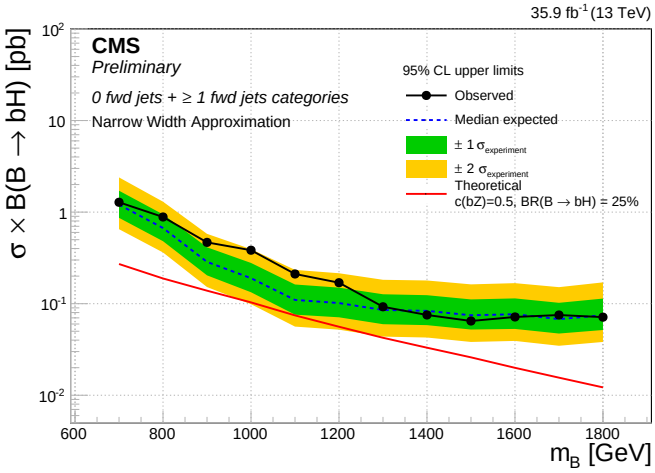

Figure 6: Observed and expected 95% confidence level upper limit on the B quark production cross section times branching fraction as a function of the signal mass hypothesis, under the assumption of narrow width resonance, for the combination of zero and at least one forward jet categories, assuming that the coupling coefficient c(bZ)=0.5 and $\mathcal{B}$(B→Hb)=0.25. The red solid curve corresponds to the theoretical cross section.

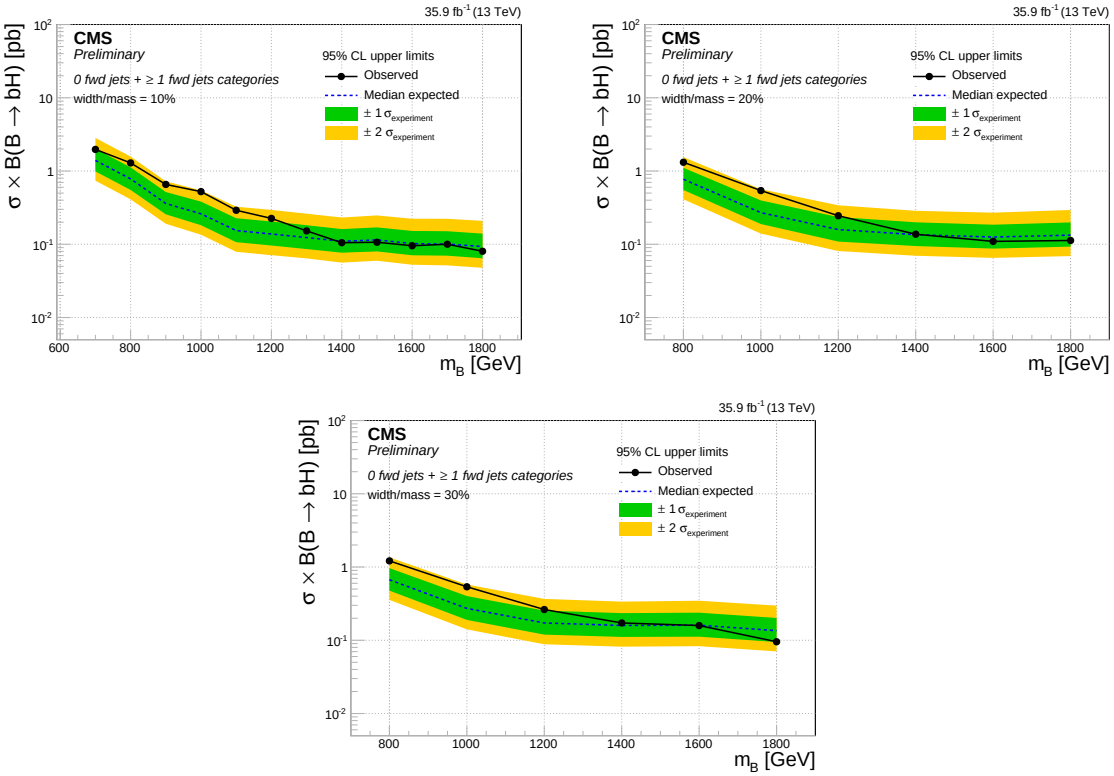

Figure 7: Observed and expected 95% confidence level upper limit on the B quark production cross section times branching fraction as a function of the signal mass hypothesis, where a width of 10% (left), 20% (center) and 30% (right) of the resonance mass is considered for the B quark. The results are shown for the combination of zero and at least one forward jet categories.

[3] J. A. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Prez-Victoria, "Handbook of vectorlike quarks: Mixing and single production", *Phys. Rev.* **D88** (2013), no. 9, 094010, doi:10.1103/PhysRevD.88.094010, arXiv:1306.0572.

[4] O. Matsedonskyi, G. Panico, and A. Wulzer, "Light Top Partners for a Light Composite Higgs", *JHEP* **01** (2013) 164, doi:10.1007/JHEP01(2013)164, arXiv:1204.6333.

[5] M. Perelstein, M. E. Peskin, and A. Pierce, "Top quarks and electroweak symmetry breaking in little Higgs models", *Phys. Rev.* **D69** (2004) 075002, doi:10.1103/PhysRevD.69.075002, arXiv:hep-ph/0310039.

[6] R. Contino, L. Da Rold, and A. Pomarol, "Light custodians in natural composite Higgs models", *Phys. Rev.* **D75** (2007) 055014, doi:10.1103/PhysRevD.75.055014, arXiv:hep-ph/0612048.

[7] R. Contino, T. Kramer, M. Son, and R. Sundrum, "Warped/composite phenomenology simplified", *JHEP* **05** (2007) 074, doi:10.1088/1126-6708/2007/05/074, arXiv:hep-ph/0612180.

[8] J. A. Aguilar-Saavedra, "Identifying top partners at LHC", *JHEP* **11** (2009) 030, doi:10.1088/1126-6708/2009/11/030, arXiv:0907.3155.

[9] A. De Simone, O. Matsedonskyi, R. Rattazzi, and A. Wulzer, "A First Top Partner Hunter's Guide", *JHEP* **04** (2013) 004, doi:10.1007/JHEP04(2013)004, arXiv:1211.5663.

[10] M. Buchkremer, G. Cacciapaglia, A. Deandrea, and L. Panizzi, "Model Independent Framework for Searches of Top Partners", *Nucl. Phys.* **B876** (2013) 376–417, doi:10.1016/j.nuclphysb.2013.08.010, arXiv:1305.4172.

[11] O. Eberhardt et al., "Joint analysis of Higgs decays and electroweak precision observables in the Standard Model with a sequential fourth generation", *Phys. Rev.* **D86** (2012) 013011, doi:10.1103/PhysRevD.86.013011, arXiv:1204.3872.

[12] G. D. Kribs, T. Plehn, M. Spannowsky, and T. M. P. Tait, "Four generations and Higgs physics", *Phys. Rev. D* **76** (2007) 075016, doi:10.1103/PhysRevD.76.075016, arXiv:0706.3718.

[13] ATLAS Collaboration, "Search for vector-like B quarks in events with one isolated lepton, missing transverse momentum and jets at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev.* **D91** (2015), no. 11, 112011, doi:10.1103/PhysRevD.91.112011, arXiv:1503.05425.

[14] ATLAS Collaboration, "Search for pair production of a new heavy quark that decays into a W boson and a light quark in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev.* **D92** (2015), no. 11, 112007, doi:10.1103/PhysRevD.92.112007, arXiv:1509.04261.

[15] ATLAS Collaboration, "Search for pair and single production of new heavy quarks that decay to a Z boson and a third-generation quark in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *JHEP* **11** (2014) 104, doi:10.1007/JHEP11(2014)104, arXiv:1409.5500.

[16] ATLAS Collaboration, "Search for single production of a vector-like quark via a heavy gluon in the 4b final state with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett.* **B758** (2016) 249–268, doi:10.1016/j.physletb.2016.04.061, arXiv:1602.06034.

[17] ATLAS Collaboration, "Search for single production of a vector-like quark via a heavy gluon in the 4b final state with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett.* **B758** (2016) 249–268, doi:10.1016/j.physletb.2016.04.061, arXiv:1602.06034.

[18] ATLAS Collaboration, "Search for single production of vector-like quarks decaying into Wb in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Eur. Phys. J.* **C76** (2016), no. 8, 442, doi:10.1140/epjc/s10052-016-4281-8, arXiv:1602.05606.

[19] ATLAS Collaboration, "Search for the production of single vector-like and excited quarks in the Wt final state in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *JHEP* **02** (2016) 110, doi:10.1007/JHEP02(2016)110, arXiv:1510.02664.

[20] CMS Collaboration, "Search for vector-like charge 2/3 T quarks in proton-proton collisions at sqrt(s) = 8 TeV", *Phys. Rev.* **D93** (2016), no. 1, 012003, doi:10.1103/PhysRevD.93.012003, arXiv:1509.04177.

[21] CMS Collaboration, "Search for pair-produced vectorlike B quarks in proton-proton collisions at $\sqrt{s}$=8TeV", *Phys. Rev.* **D93** (2016), no. 11, 112009, doi:10.1103/PhysRevD.93.112009, arXiv:1507.07129.

[22] CMS Collaboration, "Search for top-quark partners with charge 5/3 in the same-sign dilepton final state", *Phys. Rev. Lett.* **112** (2014), no. 17, 171801, doi:10.1103/PhysRevLett.112.171801, arXiv:1312.2391.

[23] CMS Collaboration, "Search for single production of vector-like quarks decaying to a Z boson and a top or a bottom quark in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **05** (2017) 029, doi:10.1007/JHEP05(2017)029, arXiv:1701.07409.

[24] CMS Collaboration, "Search for single production of a heavy vector-like T quark decaying to a Higgs boson and a top quark with a lepton and jets in the final state", *Phys. Lett.* **B771** (2017) 80–105, doi:10.1016/j.physletb.2017.05.019, arXiv:1612.00999.

[25] CMS Collaboration, "Search for single production of vector-like quarks decaying to a Z boson and a top or a bottom quark in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **05** (2017) 029, doi:10.1007/JHEP05(2017)029, arXiv:1701.07409.

[26] CMS Collaboration, "Search for electroweak production of a vector-like quark decaying to a top quark and a Higgs boson using boosted topologies in fully hadronic final states", *JHEP* **04** (2017) 136, doi:10.1007/JHEP04(2017)136, arXiv:1612.05336.

[27] CMS Collaboration, "Search for single production of vector-like quarks decaying into a b quark and a W boson in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:1701.08328.

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

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

[30] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", (2017). arXiv:1706.04965. Submitted to *JINST*.

[31] 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.

[32] CMS Collaboration, "Pileup Removal Algorithms", CMS Physics Analysis Summary CMS-PAS-JME-14-001, CERN, Geneva, 2014.

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

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

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

[36] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure", *Phys. Rev.* **D80** (2009) 051501, doi:10.1103/PhysRevD.80.051501, arXiv:0903.5081.

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

[38] 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* **1407** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[39] T. Sjstrand 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.

[40] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **0411** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[41] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[42] 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.

[43] 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.

[44] M. Czakon and A. Mitov, "Top++: A program for the calculation of the top-pair cross-section at hadron colliders", *Computer Physics Communications* **185** (2014), no. 11, 2930 – 2938, doi:10.1016/j.cpc.2014.06.021.

[45] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in FEWZ", *Phys. Rev.* **D86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.

[46] O. Matsedonskyi, G. Panico, and A. Wulzer, "On the Interpretation of Top Partners Searches", *JHEP* **12** (2014) 097, doi:10.1007/JHEP12(2014)097, arXiv:1409.0100.

[47] NNPDF Collaboration, "Parton distributions from high-precision collider data", arXiv:1706.00428.

[48] S. Agostinelli et al., "Geant4a simulation toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303, doi:10.1016/S0168-9002(03)01368-8.

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

[50] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435–443, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[51] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727. [Erratum: Eur. Phys. J.C73,2501(2013)].

[52] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, pp. 115–120. 2011. arXiv:1103.0354. doi:10.5170/CERN-2011-006.115.