PROCEEDINGS OF SCIENCE

Search for heavy BSM particles coupling to third generation quarks at CMS

Anna Benecke*
on behalf of the CMS Collaboration
Université catholique de Louvain
E-mail: anna.benecke@cern.ch

Recent results from searches for beyond standard model particles decaying to quarks of the third generation with data collected by the CMS experiment are presented. For all searches proton-proton collision data at a center-of-mass-energy of 13 TeV with an integrated luminosity of 137 fb^{-1} are analyzed. In addition, the state-of-the-art techniques in the identification of boosted objects and pileup mitigation which are used in the analyses are discussed. The searches can be grouped in three different categories: excited bottom quarks, heavy gauge boson and vector-like quarks.

**** The European Physical Society Conference on High Energy Physics (EPS-HEP2021), ****
**** 26-30 July 2021 ****
**** Online conference, jointly organized by Universität Hamburg and the research center DESY ****

*Speaker

1. Introduction

The standard model (SM) of particle physics is a very successful theory that is tested experimentally up to high precision. However, unanswered questions like the inclusion of gravity or the hierarchy problem remain. The hierarchy problem points out the unnatural amount of fine tuning needed to cancel the corrections to the bare Higgs boson mass. Many theories beyond the SM that aim to solve the hierarchy problem predict new heavy spin-1 resonances (e.g. heavy partners of the SM bosons) and a non-chiral fourth generation of quarks. The left- and right-handed component of such non-chiral hence vector-like quarks (VLQs) transform in the same way under the electroweak symmetry and no coupling to the Higgs boson is needed to generate the mass of these particles. Excited quarks are predicted by compositness models such as the Randall–Sundrum models [1, 2] and models with a heavy gluon [3].

The searches presented in this report can be grouped in three different signatures: excited bottom quarks (b*), heavy spin-1 resonances (W') and VLQ. All searches analyze proton-proton collision data collected by the CMS experiment [4] corresponding to an integrated luminosity of $137\,\mathrm{fb}^{-1}$. Due to the high resonance masses probed, the searches consider highly Lorentz-boosted particles. This means that the decay products are better reconstructed in one large-cone jet rather than individual small-cone jets. Novel techniques used to identify these boosted objects, such as HOTVR (*Heavy Object Tagging with Variable R*), DeepAK8 and imageTop as well as recent developments in the pileup mitigation techniques are briefly discussed.

2. Searches for excited bottom quarks

In this section two analyses searching for an excited bottom quark b^* decaying to a top quark and a Wboson (tW) are presented. The combination of the all-hadronic and lepton+jets (l+jets) final state is discussed in Ref. [5]. Both searches analyze the full data set collected between 2016 to 2018 corresponding to an integrated luminosity of $137\,\mathrm{fb}^{-1}$. The search in the all-hadronic final state has a dijet signature: one large-cone jet originating from the W boson and one from the t quark. In order to identify these two jets, standard substructure techniques like the SoftDrop [6] mass and N-subjettiness [7] are used.

The search in the l+jets final state reconstructs a large-cone jet originating from a t quark recoiling against a W boson reconstructed from a lepton (electron or muon) and missing transverse momentum. The large-cone jet is reconstructed with the *Heavy Object Tagging with Variable R* (HOTVR) algorithm [8]. Unlike other jet clustering algorithms, HOTVR adapts the jet radius based on the transverse momentum p_T of the jet. This makes the algorithm more sensitive in the low p_T region (due to a larger jet radius) and stable against soft and wide angular radiation at high p_T (due to a smaller jet radius). The identification efficiency for large-cone jets originating from a t quark as function of the p_T of the t quark can be seen in Fig. 1 (left). In addition, HOTVR does the grooming with the mass jump [9, 10] criterion during the jet clustering procedure.

During each bunch crossing at the Large Hadron Collider not only one proton-proton collision (the collision of interest) happens but several. Particles from these additional soft collisions are called pileup particles. These particles can add additional energy to the detector. Both analyses use the *Pileup Per Particle Identification* (PUPPI) algorithm [11] to mitigate the effect of pileup.

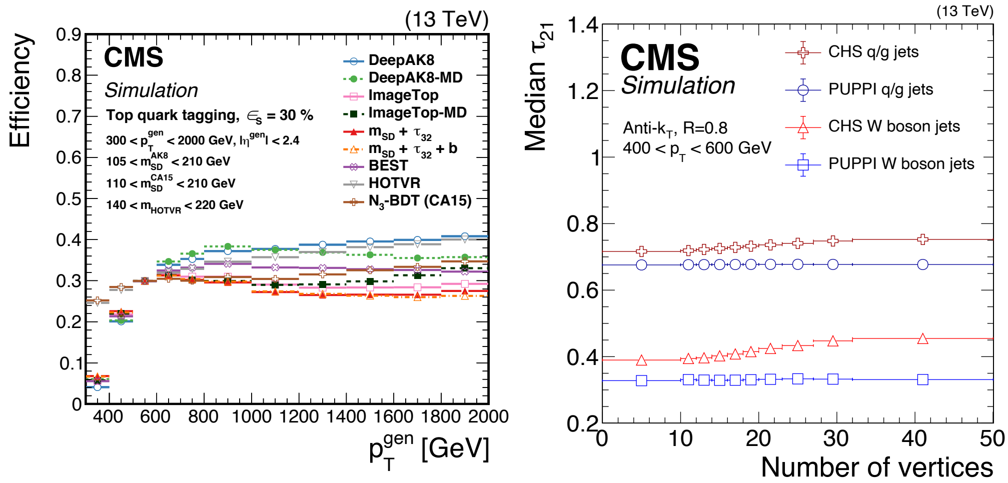

Figure 1: (Left) Identification efficiency of large-cone jets originating from a t quark as function of p_T for the different state-of-the-art algorithms used in the CMS Collaboration. Taken from Ref. [13]. (Right) Median τ_{21} as a function of the number of vertices for PUPPI (blue) and CHS (red). Taken from Ref. [12].

The performance of this algorithm was studied extensively in Ref. [12]. PUPPI shows a stable performance against pileup for key variables used to identify boosted objects. An example and the comparison to the *Charged Hadron Subtraction* (CHS) algorithm for the n-subjettiness ratio $\tau_2/\tau_1 = \tau_{21}$ used to identify large-cone jets originating from a W boson can be seen in Fig. 1 (right).

The sensitive variables for both analyses are the invariant mass of the reconstructed t-W-system, m_{tW}, and the mass of the reconstructed tquark, m_t. An example of m_{tW} for the l+jets analysis can be seen in Fig. 2 (left). The dominant SM backgrounds for the two analyses are inclusive t quark pair production (l+jets analysis) and QCD multijet (all-hadronic analysis). The QCD multijet background of the all-hadronic analysis is extrapolated from a control region (inverted t-tag requirement) with a two dimensional pass-fail ratio. The non-$t\bar{t}$ background of the l+jets analysis is estimated with the α-ratio method [14]. The inclusive t quark pair production background is taken from simulation for both analyses.

No deviation from the expected backgrounds is seen and upper limits on the production cross section are set for left-handed (LH), right-handed (RH) excited bottom quarks and the combination (LH+RH). The result for the combination (LH+RH) can be seen in Fig. 2 (right). With the combination of the l+jets and all-hadronic analysis the results are improved by almost a factor of two compared to previous results. Excited bottom quarks are excluded at 95% CL up to 2.95 (LH), 3.03 (RH) and 3.22 TeV (LH+RH).

3. Searches for heavy gauge boson

Two searches for a heavy gauge boson W' are presented in this section. The first search in Ref. [15] features the decay into SM particles, $W' \to tb$. The second analysis in Ref. [16] searches for a W' decaying to a VLQ (T or B) and a SM bottom (b) or t quark. The VLQ decays into a H/Z boson and a t/b quark leading to a final state with b quark, t quark and a H or Z boson. Both analyses consider the all-hadronic final state.

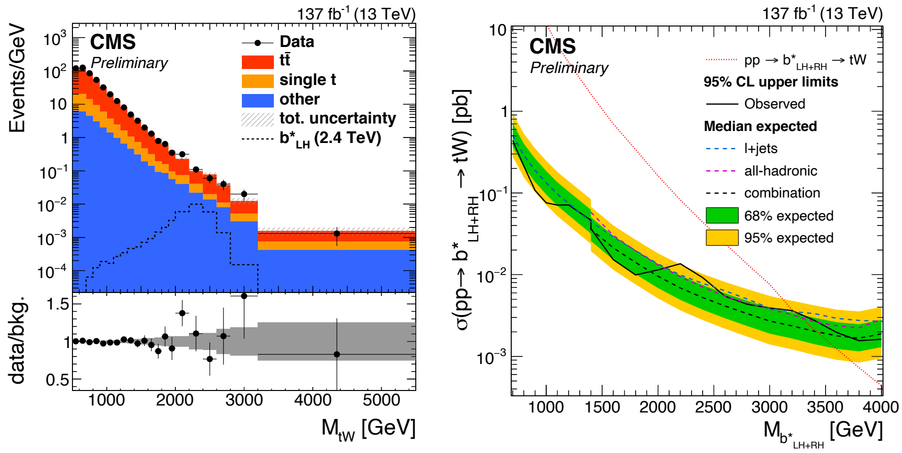

Figure 2: (Left) Invariant mass of the reconstructed t-W-system for the l+jets analysis. (Right) Upper limits on the production cross section of excited bottom quarks as function of the invariant mass of the reconstructed t-W-system, m_{tW}. Taken from Ref. [5].

The first search into SM particles, $\mathrm{W}' \to \mathrm{tb}$, has a dijet signature of the large-cone jet originating from the t quark recoiling against the small-cone jet originating from the b quark. The small-cone jet is identified with DeepJet [17], a b-tagger. The large-cone jet is identified with DeepAK8-MD [13], where MD means mass-decorrelated. DeepAK8 is a neural network with 2 input lists: one list with particle-level information of the jet constituents and one list with secondary vertex information. The challenge with each tagger is the correlation with $\mathrm{p_T}$ or mass resulting in a shaping of the SM QCD multijet background. This shaping can harm the sensitivity of a bump hunt search. Therefore, DeepAK8 has a mass-decorrelated version which is achieved by including a penalty function in the training. The shaping of the QCD multijet background without and with the mass decorrelation can be seen in Fig. 3 (left). The blue curve (DeepAK8, without mass decorrelation) shows a peak around the top mass, while the mass-decorrelated version (green curve, DeepAK8-MD) shows a smoothly falling behavior and a similar shape compared to the inclusive spectrum in dark gray (inclusive AK8).

The sensitive variable of this search is the invariant mass of the t-b-system, m_{tb}. The main background is QCD multijet production which is estimated with a pass-fail-ratio from an orthogonal region which failed the b-tagging requirement on the small-cone jet. Due to the absence of a deviation from the expect SM background, upper limits on the production cross section are set. The results can be seen in Fig. 3 (right). Left- and right-handed W' are excluded at 95% CL below 3.4 TeV.

The second analysis searching for a new heavy gauge boson W' considers the decay into a vector-like quark and a SM particle: $\mathrm{W}' \to \mathrm{Tb/Bt}$. The VLQ decays further into H/Z boson and a SM t/b quark resulting in a final signature of a H/Z boson, a b and a t quark. The small-cone jet originating from the b quark is identified with DeepJet [17]. The large-cone jet originating from the Z boson is identified with standard cut-based tagging requirements on the SoftDrop mass of the jet ($65 < m_{\mathrm{SD}}(Z) < 105$ GeV) and a requirement on τ_{21} with a signal efficiency of around 85%. The large-cone jet originating from the H boson is identified by the Double-B tagging algorithm [18].

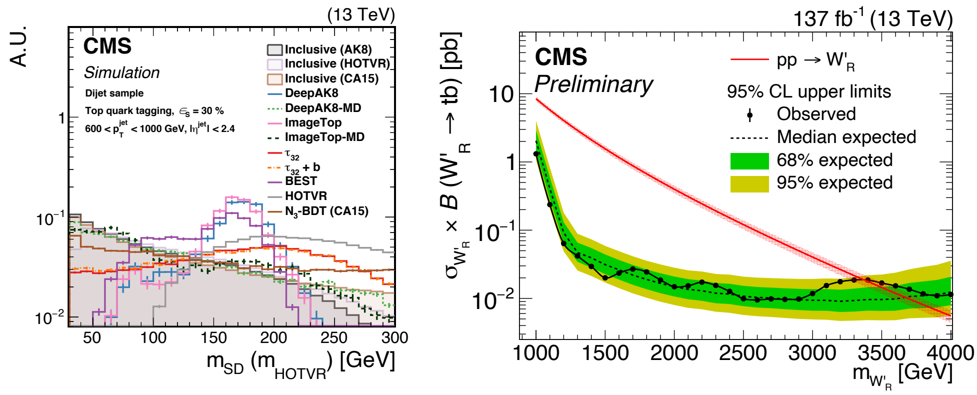

Figure 3: (Left) Mass distribution of the large-cone jet at a t-tagging efficiency of 30% for QCD multijet simulation. Taken from Ref. [13]. (Right) Upper limits on the production cross section of a heavy right-handed gauge boson, W′, as a function of the invariant mass distribution of the t-b-system. Taken from Ref. [15].

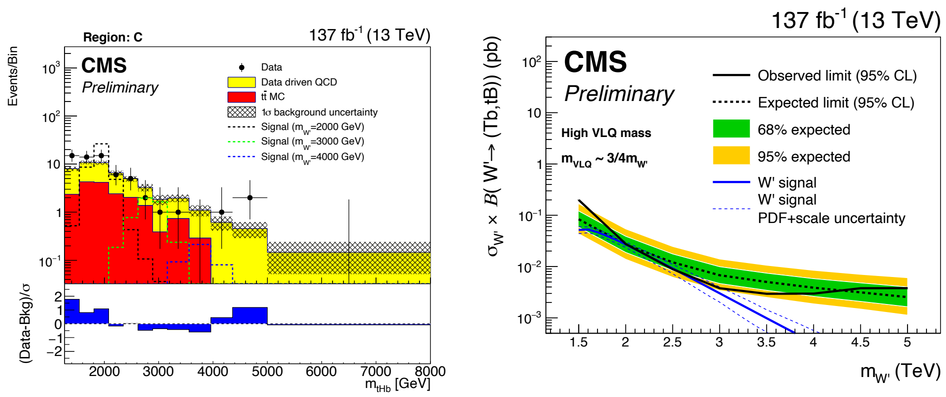

Figure 4: (Left) The invariant mass of the three jet system of m_{tHb}. (Right) Upper limit on the production cross section as a function of the heavy new gauge boson mass, $m_{\mathrm{W'}}$, for $m_{\mathrm{VLQ}} \sim 2/3 m_{\mathrm{W'}}$. Taken from Ref. [16].

The large-cone jet originating from the t quark is identified with the novel t-tagging technique imageTop-MD [13]. The imageTop-MD algorithm pixelizes the energy deposit of a jet. The image is then flipped horizontally and vertically such that the energy maximum is in the lower-left corner. In order to decorrelate the outcome from the p_T of the jet an equally shaped p_T spectrum for QCD multijet production (background) and t jets (signal) is used. In addition, a DeepJet b-tag requirement on the subjets is added to enhance the sensitivity. The sensitive variable of this search is the invariant mass of the three jet system, m_{tZb} or m_{tHb}. An example of this variable can be seen in Fig. 4 (left).

The main background of this analysis is QCD multijet production, which is taken from data using a transfer function derived in a control region. The transfer function transforms the QCD multijet background from a region with a loose t-tag requirement to the signal region with a tight ttag requirement.

No deviation from the expected SM background is observed. Therefore, upper limits on the

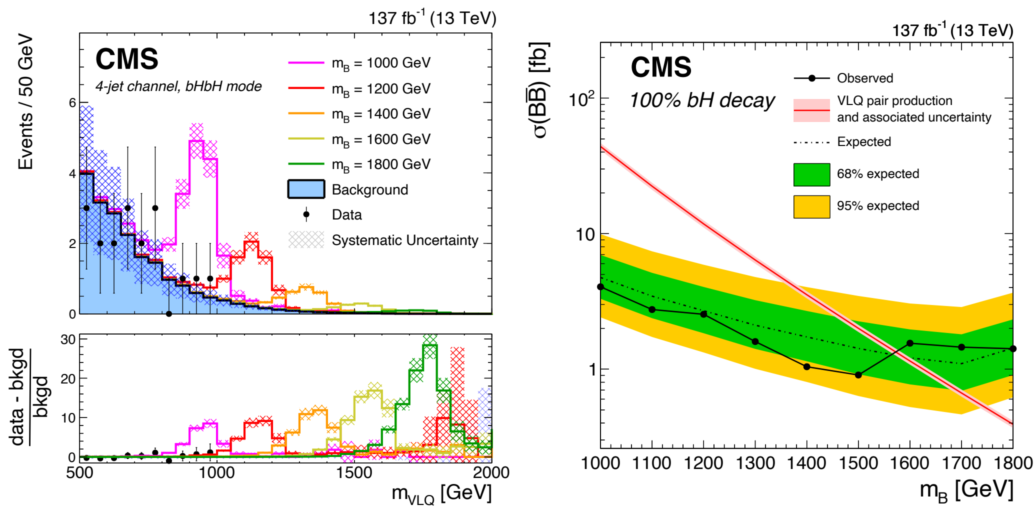

Figure 5: (Left) The average mass of the two reconstructed VLQ B in the bHbH final state in the merged topology. (Right) Upper limit on the production cross section in the bHbH final state as a function of the mass of the VLQ B. Taken from Ref. [19].

production cross section are set. The signal samples are grouped in three different mass bins of the VLQ: $m_{\mathrm{VLQ}} \sim 1/2 m_{\mathrm{W}'}$, $m_{\mathrm{VLQ}} \sim 2/3 m_{\mathrm{W}'}$ and $m_{\mathrm{VLQ}} \sim 3/4 m_{\mathrm{W}'}$. The upper limit on the production cross section for $m_{\mathrm{VLQ}} \sim 2/3 m_{\mathrm{W}'}$ can be seen in Fig. 4 (right). While a W$'$ with a mass below 3.2 TeV can be excluded at 95% CL if $m_{\mathrm{VLQ}} \sim 2/3 m_{\mathrm{W}'}$, the sensitivity to exclude a W$'$ if $m_{\mathrm{VLQ}} \sim 1/2 m_{\mathrm{W}'}$ or $m_{\mathrm{VLQ}} \sim 3/4 m_{\mathrm{W}'}$ is almost reached.

4. Searches for vector-like quarks

In this section searches for pair produced VLQs and singly-produced VLQs are presented. Both analyses cover the whole range from merged over partially merged to resolved topologies.

The search for pair produced VLQ Bs [19] considers the all-hadronic final state. The VLQ B can decay either in bZ or bH resulting in three different combinations: bZbZ, bZbH, bHbH. For all three combinations the boson is considered to decay into a pair of b quarks leading to a six b quark final state. The bosons can be resolved (four small-cone jets), partially merged (one of the bosons merged, the other resolved) and merged (two large-cone jets). Large-cone jets originating from the H or Z boson are identified by the DoubleB algorithm [18]. Small-cone jets originating from a b quark are identified with DeepJet [17]. The sensitive variable is the average mass of the two reconstructed VLQ B masses, m_{VLQ}. An example can be seen in Fig. 5 (left). No deviation from the SM has been found, therefore, upper cross section limits on the production cross section are set (Fig. 5, right) and VLQ B are excluded at 95% CL below 1570 (bHbH), 1390 (bZbZ) and 1450 GeV (bZbH).

The second analysis is a search for a singly-produced VLQ T decaying into tZ with the $\mathrm{Z} \to \nu\nu$ [20]. The t quark can be merged (one large-cone jet), partially merged (one large-cone jet and a small-cone jet) or resolved (three small-cone jets). Due to the two neutrinos the sensitive variable is the transverse mass, $M_T = \sqrt{2 \mathrm{p}_\mathrm{T}^\mathrm{t} \mathrm{p}_\mathrm{T}^\mathrm{miss} \left(1 - \cos \Delta\phi_{\mathrm{t},\vec{p}_\mathrm{T}^{\,\mathrm{miss}}}\right)}$. A dedicated control region is

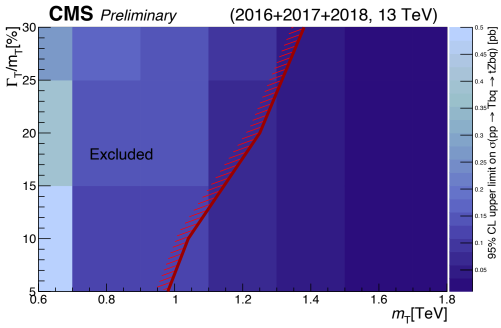

Figure 6: Upper limit on the production cross section for a singly-produced VLQ T $\rightarrow$ tZ$(\nu\nu)$. Taken from Ref. [20].

defined for each of the major backgrounds: inclusive t quark pair production, inclusive W boson production and inclusive Z boson production. The most significant excess between data and the expected SM backgrounds is around 2.5 standard deviations and is found for a signal with an expected m_T = 1.4 TeV and a narrow width. Upper limits on the production cross section are set and the limits for different widths can be seen in Fig. 6.

References

[1] 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.

[2] L. Randall and R. Sundrum, "An Alternative to compactification", *Phys. Rev. Lett.* **83** (1999) 4690–4693, doi:10.1103/PhysRevLett.83.4690, arXiv:hep-th/9906064.

[3] C. Bini, R. Contino, and N. Vignaroli, "Heavy-light decay topologies as a new strategy to discover a heavy gluon", *JHEP* **01** (2012) 157, doi:10.1007/JHEP01(2012)157, arXiv:1110.6058.

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

[5] CMS Collaboration, "Search for a heavy resonance decaying into a top quark and a W boson in the lepton+jets final state at 13 TeV", CMS-PAS-B2G-20-010, CERN, 2021.

[6] 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.

[7] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness", *JHEP* **03** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.

[8] T. Lapsien, R. Kogler, and J. Haller, "A new tagger for hadronically decaying heavy particles at the LHC", *Eur. Phys. J. C* **76** (2016), no. 11, 600, doi:10.1140/epjc/s10052-016-4443-8, arXiv:1606.04961.

PoS(EPS-HEP2021)675

[9] M. Stoll, "Vetoed jet clustering: The mass-jump algorithm", *JHEP* **04** (2015) 111, doi:10.1007/JHEP04(2015)111, arXiv:1410.4637.

[10] K. Hamaguchi, S. P. Liew, and M. Stoll, "How to decontaminate overlapping fat jets", *Phys. Rev. D* **92** (2015), no. 1, 015012, doi:10.1103/PhysRevD.92.015012, arXiv:1505.02930.

[11] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup Per Particle Identification", *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[12] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020), no. 09, P09018, doi:10.1088/1748-0221/15/09/P09018, arXiv:2003.00503.

[13] CMS Collaboration, "Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques", *JINST* **15** (2020), no. 06, P06005, doi:10.1088/1748-0221/15/06/P06005, arXiv:2004.08262.

[14] CMS Collaboration, "Search for massive resonances decaying into WW, WZ or ZZ bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2017) 162, doi:10.1007/JHEP03(2017)162, arXiv:1612.09159.

[15] CMS Collaboration, "Search for a W′ boson decaying to a top and a bottom quark at $\sqrt{s} = 13$ TeV", CMS-PAS-B2G-20-005, CERN, 2020.

[16] CMS Collaboration, "Search for W′ decaying to a vector-like quark and a top or bottom quark in the all-jets final state", CMS-PAS-B2G-20-002, CERN, 2021.

[17] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018), no. 05, P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[18] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018), no. 05, P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[19] CMS Collaboration, "A search for bottom-type, vector-like quark pair production in a fully hadronic final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **102** (2020) 112004, doi:10.1103/PhysRevD.102.112004, arXiv:2008.09835.

[20] CMS Collaboration, "Search for single production of a vector-like quark T decaying to a top quark and a Z boson with the Z boson decaying to neutrinos", CMS-PAS-B2G-19-004, CERN, 2021.