

# Search for the production of vector-like quarks at CMS

---

**J. M. Hogan\*** on behalf of the CMS collaboration

*Brown University*

*E-mail:* [j.hogan@cern.ch](mailto:j.hogan@cern.ch)

Three recent searches are presented for electroweak production of a vector-like top quark partner  $T$  of charge  $2e/3$  in association with a standard model top or bottom quark, using  $2.3 \text{ fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13 \text{ TeV}$  collected by the CMS experiment at the CERN LHC. The searches, the first of their kind at this energy, are targeted to three decay modes of the  $T$  quark:  $tH$ , in a fully hadronic final state;  $tZ$ , in a dilepton final state; and  $bW$ , in a single lepton final state. All searches make use of algorithms to identify  $b$  quark jets and boosted  $W$ ,  $H$  boson and top quark decays. Across a range of  $T$  quark masses from  $700 - 1800 \text{ GeV}$ , these searches place 95% C.L. upper limits on the production cross section times branching fraction of  $0.37 - 0.93 \text{ pb}$  ( $T \rightarrow tH$ ),  $0.11 - 2.0 \text{ pb}$  ( $T \rightarrow bW$ ), and  $0.16 - 0.97 \text{ pb}$  ( $T \rightarrow tZ$ ).

*38th International Conference on High Energy Physics  
3-10 August 2016  
Chicago, USA*

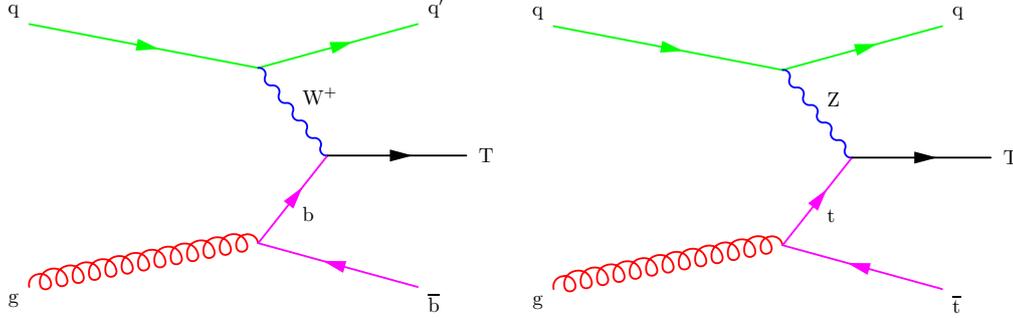
---

\*Speaker.

## 1. Introduction

The discovery of a Higgs boson with a low mass completed the standard model (SM), but suggests new physics is needed to avoid large corrections from top quark loops. Several theories that seek to preserve an electroweak Higgs mass predict the existence of TeV-scale particles called vector-like quarks (VLQ) (see [1] and references within).

VLQs can be produced in pairs via the strong interaction, or singly in association with top or bottom quarks, via the electroweak interaction. For high VLQ masses, single production (shown in Fig. 1) is expected to dominate. The VLQs are non-chiral particles with left- and right-handed (LH and RH) components that share a weak isospin multiplet. Therefore they do not receive mass through a Yukawa coupling. A fourth generation of chiral quarks is strongly disfavoured by precision measurements of the SM [2], but VLQs are not similarly constrained. An electroweak singlet T quark has three decay channels:  $bW$ ,  $tZ$  and  $tH$ , with  $\mathcal{B}(bW) = 2\mathcal{B}(tZ, tH)$ . A T doublet has decay modes  $tZ$  and  $tH$  with branching fractions of 0.5. Other searches for VLQs at CMS are listed in [3].



**Figure 1:** Example production diagrams for (left)  $pp \rightarrow Tbq$  and (right)  $pp \rightarrow Ttq$  production.

## 2. Searches for VLQ at CMS

A 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. [4]. The data used in these searches were recorded during 2015 in  $pp$  collisions at  $\sqrt{s} = 13$  TeV. The total integrated luminosity is  $2.3 \text{ fb}^{-1}$ . Signal samples were generated for VLQ masses between 700 and 1800 GeV, with LH and RH couplings to the SM. All masses are generated with a narrow width which is negligible compared to the experimental resolution for couplings below  $\approx 0.5$ . Cross sections are evaluated at next-to-leading order [5]. SM background processes are modeled using both Monte Carlo (MC) simulation and data-driven estimations, described below for each search.

Of particular importance to VLQ searches is identifying the flavor content of jets as well as jets containing highly boosted objects from a heavy particle decay. To identify “AK4” jets as originating from a  $b$  quark the Combined Secondary Vertex (CSV) algorithm is used [6]. In larger radius “AK8” jets grooming techniques (pruning [7] with  $z_{\text{cut}} < 0.1$  and  $D_{\text{cut}} < 0.5$ , soft-drop [8] with  $z_{\text{cut}} < 0.1$  and  $\beta = 0$ ) are applied to remove soft and wide-angle contributions to the jet energy,

and reveal subjects from the hadronization of the hard partons. The mass of the groomed jet is thus closer to that of the parent particle. Subjects identified by the soft-drop algorithm can also be tagged as b jets using the CSV algorithm, which is highly effective for identifying decays such as  $H \rightarrow b\bar{b}$  or t quark decays.

To quantify concentrations of energy flow inside a jet the shape variables  $\tau_N$  of the N-subjettiness algorithm are used [9]. A jet compatible with two substructures (such as from a W or Higgs boson decay) would have values of the ratio  $\tau_2/\tau_1$  much less than unity while a jet from a boosted  $t \rightarrow bqq'$  decay, with three substructures, would have the value of  $\tau_3/\tau_2$  much less than one. In contrast, jets with no substructure would exhibit larger values for both  $\tau_2/\tau_1$  and  $\tau_3/\tau_2$ . Thus these variables provide good discrimination against multijet backgrounds.

## 2.1 Search for $T \rightarrow tH$

This search considers the decay mode  $T \rightarrow tH$  with  $t \rightarrow bqq'$  and  $H \rightarrow b\bar{b}$  [10]. Jet substructure and b-tagging techniques are employed to identify the highly Lorentz-boosted t quark and Higgs boson. The main SM background processes are  $t\bar{t} + \text{jets}$  and multijet production, with smaller contributions from W + jets events and single t quark events. The multijet contribution is estimated from data and other background processes are estimated using MC.

In this search  $H_T = \sum p_T(\text{jets})$  for all AK4 jets with  $p_T > 30$  GeV and  $|\eta| < 5$ , requiring  $H_T > 1100$  GeV. The AK8 jets with  $p_T > 300$  GeV and  $|\eta| < 2.4$  are considered H tagged if they have pruned mass within 105–135 GeV,  $\tau_2/\tau_1 < 0.6$ , and two b-tagged subjects. The t-tagged jets require  $p_T > 400$  GeV, soft-drop mass within 110–210 GeV,  $\tau_3/\tau_2 < 0.54$ , and one b-tagged subject. Boosted jets with both subjects failing the b tagging criteria but otherwise satisfying the H tagging criteria (“anti-H-tagged”) are used to define control regions.

It is observed that the background MC simulations do not model well the observed jet  $p_T$  and  $H_T$  distributions. The observed data/MC ratio of the  $H_T$  distribution is well described by a linear fit with a negative slope parameter, used to correct MC-based backgrounds. After preselection, the leading H-tagged and t-tagged jets in each event (which must be spatially separated) are paired to form the T quark candidate with invariant mass  $M(T)$ . The search is performed by looking for a localized excess in the T quark candidate mass distribution above the SM background.

The multijet background is estimated from data by using sideband distributions in four regions A, B, C, and D. Regions A and B require an anti-H-tagged jet rather than an H-tagged jet. Regions A and C invert the requirement of a t-tagged jet. Region D is the signal region. If the two variables that span regions ABCD are uncorrelated, the number of events  $N_{A,B,C,D}$  should follow the relation  $N_A/N_B = N_C/N_D$ . The ABCD method is also used to obtain the background  $M(T)$  distribution for the signal region using the anti-H and t-tagged jets in region B and multiplying the yield by the ratio  $N_C/N_A$ . After subtracting simulated backgrounds from data in each region,  $N_C/N_A$  is found to be  $13.6 \pm 0.2$ . The total estimated background from all sources is  $35.7 \pm 5.6$  events, and 30 events are observed. The  $H_T$  and  $M(T)$  distributions are shown in Fig. 2(left, middle).

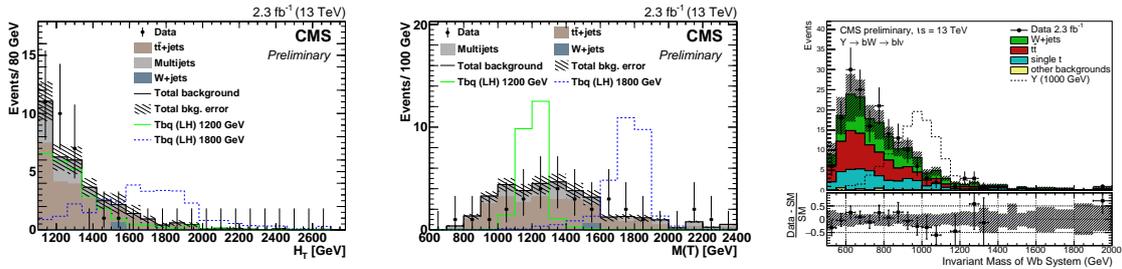
## 2.2 Search for $T, Y \rightarrow bW$

In this analysis, we search for single production of a Y (charge  $-4e/3$ ) or T quark that decays into a W boson and a b quark [11]. The associated light flavor quark produces a jet in the forward region of the detector, and it serves as a signature of single VLQ production.

The search requires events with only one lepton (electron or muon) with  $p_T > 40$  GeV and  $|\eta| < 2.1$ . Events must have at least two jets, one in the central ( $|\eta| < 2.4$ ) and one in the forward region ( $2.4 < |\eta| < 5.0$ ) of the detector, that are well separated from the lepton by  $\Delta R > 1.5$ . The central jet must have  $p_T > 200$  GeV and be b tagged. The forward jet must have  $p_T > 30$  GeV. In single VLQ production, the b quark and W boson tend to be in opposite directions, so the azimuthal angle between the central b jet and the lepton must satisfy  $\Delta\phi(\ell, b) > 2$ . Events must also have missing transverse energy  $E_T^{\text{miss}} > 50$  GeV due to the neutrino from the W boson decay, and transverse mass  $M_T(\ell, E_T^{\text{miss}}) < 130$  to suppress  $t\bar{t}$  dilepton events. Finally, events must have  $S_T > 500$  GeV, where  $S_T = p_T(\ell) + E_T^{\text{miss}} + \sum p_T(\text{jets})$ .

The mass of the VLQ candidate is reconstructed as an invariant mass  $M_{\text{inv}}$  of the lepton, the leading central jet, and the neutrino, where the x,y-components of the neutrino momentum are given by the missing transverse momentum, while the z-component is determined from a W mass constraint. New physics from  $Y/T \rightarrow bW$  would result in a peak in the  $M_{\text{inv}}$  distribution. The search is thus performed by fitting the observed  $M_{\text{inv}}$  distribution to a combination of SM background processes plus new physics signal.

The dominant backgrounds in this search are  $t\bar{t}$  and W + jets events. The modeling of these processes from simulation is validated by studying dedicated background-enriched samples. The top  $p_T$  spectrum in the  $t\bar{t}$  simulation is known to be mis-modeled and it is reweighted using an empirical function. We observe that in the W + jets-enriched region the number of events in the tails of the jet  $p_T$  distributions are overpredicted. We derive a correction to the W + jets simulation in a similar manner to the search for  $T \rightarrow tH$ . In the electron channel  $95.3^{+30.8}_{-27.9}$  events are expected with 78 events observed, and in the muon channel  $115.2^{+34.9}_{-33.3}$  events are expected with 95 events observed. The  $M_{\text{inv}}$  spectrum is fitted to a combination of SM background processes plus the VLQ signal. The post-fit  $M_{\text{inv}}$  distribution, with the shape and background normalizations corresponding to the maximum likelihood values, is presented in Fig. 2(right).



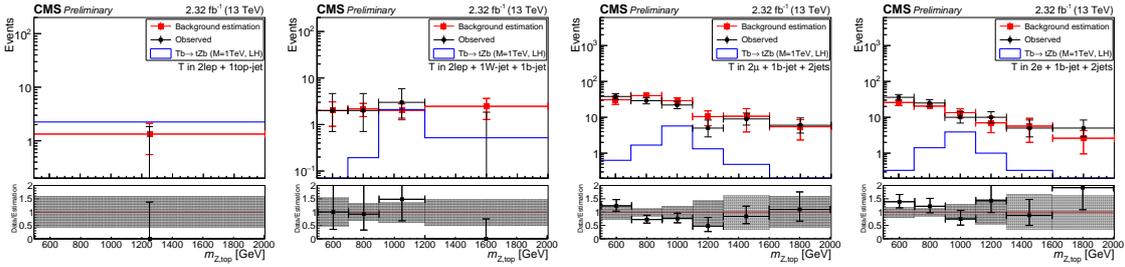
**Figure 2:** The (left)  $H_T$  and (middle)  $M(T)$  distributions after the  $T \rightarrow tH$  selection. The T quark signal distributions for two masses are shown with  $\sigma \mathcal{B}(T \rightarrow tH) = 1$  pb. (Right) The invariant mass  $M_{\text{inv}}$  distribution for  $T \rightarrow bW$  after the fit. The data overflow event has an electron and  $M(T) = 2.2$  TeV.

### 2.3 Search for $T \rightarrow tZ$

This analysis searches for single production of  $T \rightarrow tZ$ , with  $Z \rightarrow \ell^+ \ell^-$  and  $t \rightarrow bqq'$  [12]. Decays of B quarks and T quarks produced via a  $Z'$  decay are also presented in [12].

In this search AK8 jets are labeled as W tagged if they have  $p_T > 200$  GeV, pruned mass of 65 – 105 GeV, and  $\tau_2/\tau_1 < 0.6$ . Jets identified as t tagged have  $p_T > 400$  GeV, softdrop mass of 110 – 210 GeV and  $\tau_3/\tau_2 < 0.69$ . Each event must have a pair of opposite sign electrons or muons with  $M(\ell^+\ell^-)$  within 70 – 110 GeV. Events are categorized as follows: events with a fully merged t-tagged jet, events with a b jet and a W-tagged jet, muon events with resolved t quark decays, and electron events with resolved t quark decays. The t candidate is reconstructed by summing the Lorentz vectors of the jet(s) containing its decay products. Candidates must have  $p_T > 150$  GeV. The T candidate is reconstructed by summing the Lorentz vectors of the Z boson and of the t quark. Other selection criteria include: at least one b tagged jet,  $\Delta R(\ell^+, \ell^-) < 0.9 - 1.1$  (depending on the category), and leading lepton  $p_T > 115$  GeV.

In order to rely as little as possible on the simulation, a data-driven prediction, called the  $\alpha$ -ratio method, is performed. The number of events in the signal region is extrapolated from a control region with zero b-tagged jets in each event. The background in the signal region is evaluated with the following formula:  $N_{bkg}(M_{T/B}) = N_{cr}(M_{T/B}) \cdot \alpha(M_{T/B})$ , where  $N_{cr}(M_{T/B})$  is the number of events in data in the control region, in bins of  $M_{T/B}$ , and  $\alpha(M_{T/B})$  is a binned ratio between the shapes of the  $M_{T/B}$  distributions in MC in the signal and control regions. Fig. 3 compares the prediction from the data-driven method and data.



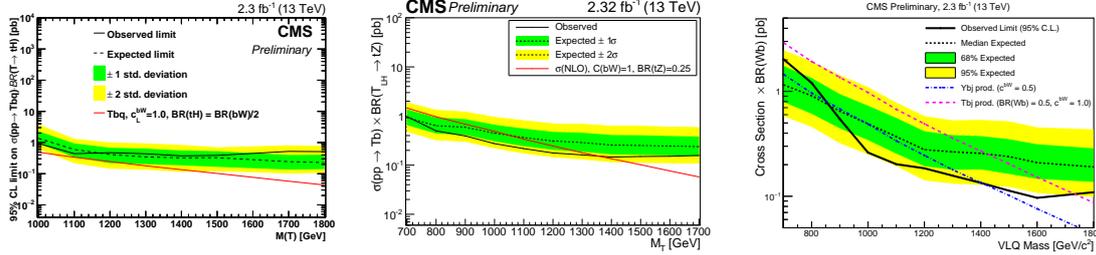
**Figure 3:** Comparison of the alpha method prediction and data for (left) fully merged events, (middle-left) partially merged events, and resolved events with (middle-right) muons or (right) electrons.

### 3. Results

Given that no excess of events is seen over the estimated background in any analysis channel, we proceed to set 95% CL upper limits on the production cross section of the T quark produced through an electroweak interaction. In all analyses the shapes of the  $M(T)$  distributions for background and signal are fit to the data. The systematic uncertainties are treated as nuisance parameters. Depending on the search systematic uncertainties include uncertainties on data-driven background estimates, MC theory and heavy flavor contribution, lepton and tagging efficiencies, luminosity, inelastic pp cross section, jet energy and resolution corrections, forward jet efficiency, and modeling corrections. Limits are computed using either Bayesian or Asymptotic  $CL_s$  methods, and comparisons are made to the simplified model framework for a singlet or doublet T quark.

Figure 4 shows the upper limit from each search on  $\sigma(pp \rightarrow Tbq) \mathcal{B}(T \rightarrow tH, tZ, bW)$  compared to the benchmark scenario of a left-handed electroweak singlet T quark with unit coupling to bW. Across the relevant mass ranges, single T quark production is excluded for  $\sigma \mathcal{B} = 0.37 -$

0.93 pb ( $T \rightarrow tH$ ), 0.11 – 2.0 pb ( $T \rightarrow bW$ ), and 0.16 – 0.97 pb ( $T \rightarrow tZ$ ). These searches represent the advanced use of jet substructure techniques in heavy particle searches, and are the first searches of their kind in pp collisions at  $\sqrt{s} = 13$  TeV.



**Figure 4:** The expected and observed limits on the  $pp \rightarrow Tbq$  production process for a left-handed singlet  $T$  quark from (left) the  $T \rightarrow tH$  search, (middle) the  $T \rightarrow tZ$  search, and (right) the  $T/Y \rightarrow bW$  search.

## References

- [1] J. A. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer and M. Perez-Victoria, “Handbook of vectorlike quarks: Mixing and single production,” *Phys. Rev. D* **88** 094010 (2013) [hep-ph/13060572].
- [2] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, U. Nierste and M. Wiebusch, “Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations,” *Phys. Rev. Lett.* **109** 241802 (2012) [hep-ph/12091101].
- [3] CMS Collaboration, “CMS Beyond-two-generations Public Physics Results”, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsB2G>
- [4] S. Chatrchyan *et al.* [CMS Collaboration], “The CMS experiment at the CERN LHC,” *JINST* **3** S08004 (2008).
- [5] O. Matsedonskyi, G. Panico and A. Wulzer, “On the Interpretation of Top Partners Searches,” *JHEP* **1412** 097 (2014) [hep-ph/14090100].
- [6] S. Chatrchyan *et al.* [CMS Collaboration], “Identification of b-quark jets with the CMS experiment,” *JINST* **8** P04013 (2013) [hep-ex/12114462].
- [7] S. D. Ellis, C. K. Vermilion and J. R. Walsh, “Techniques for improved heavy particle searches with jet substructure,” *Phys. Rev. D* **80** 051501 (2009) [hep-ph/09035081].
- [8] A. J. Larkoski, S. Marzani, G. Soyez and J. Thaler, “Soft Drop,” *JHEP* **1405** 146 (2014) [hep-ph/14022657].
- [9] J. Thaler and K. Van Tilburg, “Maximizing Boosted Top Identification by Minimizing N-subjettiness,” *JHEP* **1202** 093 (2012) [hep-ph/11082701].
- [10] CMS Collaboration, “Search for a vectorlike top partner produced through electroweak interaction and decaying to a top quark and a Higgs boson using boosted topologies in the all-hadronic final state,” CMS-PAS-B2G-16-005.
- [11] CMS Collaboration, “Search for singly-produced vector-like quarks decaying into a W boson and a bottom quark using the single lepton final state,” CMS-PAS-B2G-16-006.
- [12] CMS Collaboration, “Search for single production of vector-like quarks decaying into final states with a Z boson and a top or a bottom quark,” CMS-PAS-B2G-16-001.