DRAFT
CMS Physics Analysis Summary

2016/06/07
Head Id: 345175
Archive Id: 345982M
Archive Date: 2016/06/01
Archive Tag: trunk

Search for electroweak production of a vectorlike quark decaying to a top quark and a Higgs boson using boosted topologies in all-hadronic final state

The CMS Collaboration

Abstract

A search is performed for the electroweak production of a vectorlike top quark partner T of charge $+2/3$ in association with a standard model top or bottom quark, using 2.3 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected by the CMS experiment. The search targets T quarks decaying to a top quark and a Higgs boson in the fully-hadronic final state. For a T quark with mass above 1 TeV, the daughter top quark and the Higgs boson are highly Lorentz-boosted and each can be reconstructed as a single hadronic jet. Jet substructure and b tagging techniques are used to identify the top and Higgs jets, and to suppress the standard model backgrounds, which is found to be consistent with observations. Upper limits at 95% CL are set on the product of the single T quark production cross section and the $\mathcal{B}(\mathrm{T} \to \mathrm{tH})$ that vary between $0.31 - 0.93$ pb for T quark masses in the range $1000 - 1800$ GeV.

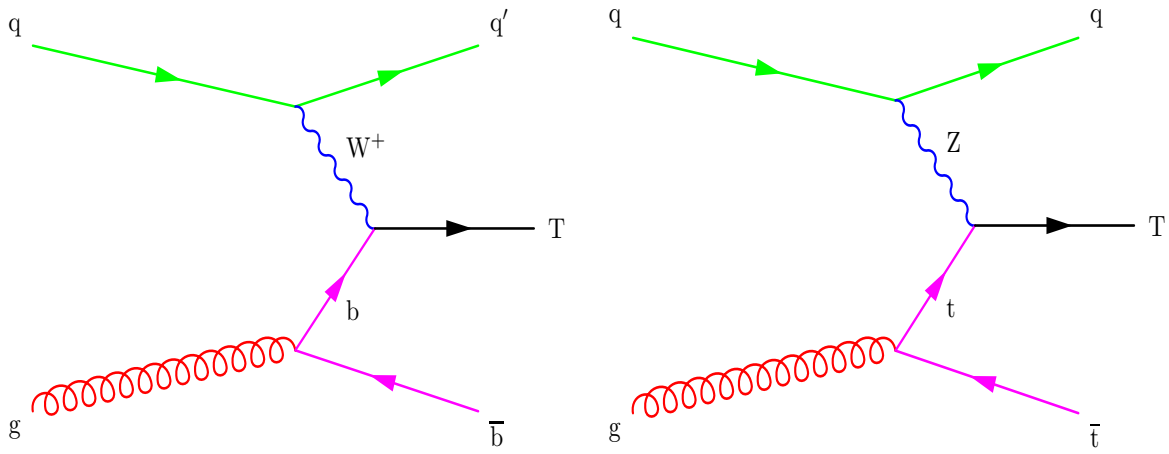

Figure 1: Example production diagrams. Charged current (left) and neutral current (right).

1 Introduction

We report on the search for a vectorlike quark (VLQ) of charge $+2/3$ (T). These particles appear in many extensions of the standard model (SM) and usually mix with the SM third generation top quark [1–6]. Such a particle could have a role in regularizing the Higgs boson mass and thus offers a potential solution to the hierarchy problem [7, 8]. These particles can be produced in pairs through the strong interaction, or singly, via electroweak interaction, in association with the SM t or b quarks. The latter production mode, shown in Fig. 1, is highly model dependent and probing such processes could shed light on the mixing of the VLQs with the SM 3rd generation quarks. In these models, the decays $\mathrm{T} \to \mathrm{bW}$, $\mathrm{T} \to \mathrm{tZ}$, and $\mathrm{T} \to \mathrm{tH}$ are the most relevant [9, 10].

The VLQs being non-chiral, their masses are not constrained by their Yukawa coupling to the Higgs field, hence these particles are not ruled out by the recent Higgs boson discovery [11–13]. The single production modes are more relevant for extending the search range beyond the TeV scale as the cross sections fall off slower than the pair-production process due to kinematic constraints. Many models predict a higher single production rate for high T quark masses than the pair-production. Searches for a T quark produced in pairs were conducted by the ATLAS and the CMS collaborations and have placed limits on its mass between 720 and 950 GeV, depending on its decay mode [14–19]. In this letter we report on the search for a singly produced T quark using 2.3 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV using the CMS experiment. The production processes considered are $\mathrm{pp} \to \mathrm{Tbq}$ or $\mathrm{pp} \to \mathrm{Ttq}$. We consider the decay mode $\mathrm{T} \to \mathrm{tH}$ with the top quark decaying fully hadronically to three partons $\mathrm{t} \to \mathrm{bq\overline{q}'}$ and the Higgs boson decaying to $\mathrm{H} \to \mathrm{b\overline{b}}$. The search for such a mode exemplifies the possibilities for exploring new physics in new ways with the recently discovered Higgs boson. This search also showcases the importance of the boosted techniques in a challenging all-hadronic final state, to reconstruct hadronically decaying SM particles using their highest branching ratio decay modes.

2 Signal and background simulation

The single production cross section σ and the branching ratio $\mathcal{B}(\mathrm{T} \to \mathrm{tH})$ are highly model dependent. The production cross sections are evaluated using the Simplest Simplified Model framework of Ref. [20]. The charged and neutral current processes, leading to the $\mathrm{pp} \to \mathrm{Tbq}$

and pp $\rightarrow$ Ttq production modes respectively, depend quadratically on coupling parameters c^{bW} and c^{tZ}. The cross sections are evaluated at approximate next-to-leading order (NLO) with either the left (LH) or right (RH) coupling set to 1 and the other set to zero using the techniques of Ref. [21].

Signal events for the process pp $\rightarrow$ Tbq, T $\rightarrow$ tH are generated using the leading order Monte Carlo (MC) event generator MADGRAPH 5.1.3.30 [22] for various T quark masses. The MADGRAPH samples are generated with an additional parton. The NNPDF3.0 parton distribution function (PDF) is used [23]. The samples have both the top quark t and the Higgs boson H decaying inclusively, with the mass of H boson set to 125 GeV and that of the t quark set to 172.5 GeV.

The main SM background processes are $t\bar{t}$+jets and multijet production through the strong interaction. Small quantities of W+jets events may also be present in the final state. Single top quark tW events are found to have negligible contribution to the overall background composition. The $t\bar{t}$+jets, W+jets and tW background events are estimated using MC simulations. The multijet process being difficult to model using a simulation, is estimated from the data.

The $t\bar{t}$+jets events are simulated using POWHEG V2 [24–27] with inclusive t quark decays, and interfaced with PYTHIA8.212 [28]. The mass of the top quark is set to 172.5 GeV. A next-to-next-to-leading order (NNLO) cross section for the $t\bar{t}$+jets is used [29] which gives $831.76^{+46.45}_{-50.85}$ pb for this top quark mass. The W+jets sample includes only hadronic W decays and is simulated with Madgraph and PYTHIA8.212. The cross section of the W+jets is 95.14 ± 3.57 pb, calculated at NNLO using the FEWZ 3.1 [30] cross section calculator, with the appropriate filter efficiency for the requirement on the sum of the parton transverse momenta p_T. The sum of the p_T of the partons is required to be > 600 GeV in the simulated sample. The tW cross section is taken as 35.85 pb [31, 32]. The multijet samples are generated using MADGRAPH with up to four partons using the MLM matching scheme [33] with PYTHIA8.212 and are used to optimize the event selection and background estimation procedure.

Additional proton-proton interactions in the same and adjacent bunch crossings (pileup) are simulated by overlaying low p_T QCD interactions using the PYTHIA MC generator, and assuming that the total inelastic proton-proton cross section is 69 mb. The distribution of the number of pileup events in the simulated samples were reweighted to match that in the data due to the running conditions of the LHC.

3 The CMS detector and event reconstruction

The CMS detector, its coordinate system, and kinematic variables are detailed in Ref. [34]. The detector consists of a superconducting solenoid of 6 m internal diameter at its core, providing a magnetic field of 3.8 T. Within the field volume are housed a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass/scintillator hadron calorimeter (HCAL). The tracker extends from -2.5 to $+2.5$ in pseudorapidity η while the ECAL and HCAL extend up to $|\eta| < 3$. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors.

This analysis uses proton-proton data collected by the CMS experiment at a center-of-mass energy of $\sqrt{s} = 13$ TeV and 25 ns bunch spacing starting in the summer of 2015. A total integrated luminosity of 2.3 fb^{-1} is used based on data collected with the magnetic field on. Events are selected using a two-stage trigger system, requiring the presence of large hadronic jet activity in the detector. The level 1 trigger retains events with loose jet requirements for

further processing by the high-level trigger (HLT). The HLT reconstructs jets clustered using the particle flow (PF) algorithm [35] described below, and with $p_T > 40$ GeV and $|\eta| < 3$. The scalar sum of the jet p_T is required to be greater than 800 GeV for the event to be selected by the HLT for further processing.

Charged particle tracks are used to reconstruct interaction vertices with the one having the highest $\sum p_T^2$ of the tracks considered the primary vertex. The PF event reconstruction algorithm reconstructs and identifies stable particles in the detector using an optimized combination of all subdetector information. Jets are reconstructed using calibrated particle flow candidates using the infrared and collinear safe anti-k_T algorithm [36] implemented using the FastJet algorithm [37]. Charged particles not originating from the primary vertex are not clustered in the jets to remove particles from pileup. Further corrections based on the jet area technique [38–40] are applied on a per event basis to mitigate the effect of pileup contribution to the jet energy. The energies of the resulting PF jets are corrected to account for the non-linear detector response, and difference with respect to the simulation. The jet energy corrections derived from simulation and confirmed with in situ measurements are applied. From simulations, the jet momentum which is the vector sum of all calibrated particle momenta in the jet, is found to be within 5-10% of the true momentum over the whole range of detector acceptance [40]. Jet energy resolution is found typically to be from 15% at 10 GeV to 4% at 1 TeV. Two non-mutually exclusive jet collections are reconstructed using distance parameters in the $\eta - \phi$ plane of 0.4 (AK4 jets) and 0.8 (AK8 jets).Jets are required to pass a standard set of quality criteria to reject detector and electronics noise misidentified as jets [41].

4 Event selection

Jet grooming techniques [42] aim to identify Lorentz-boosted massive hadronically decaying particles like $H \to b\bar{b}$ or $t \to q\bar{q}'b$, and reject single partonic jets from QCD interactions. The grooming algorithms pruning [43] and soft drop [44] are employed to remove soft contributions to the jet energy due to the underlying event and pileup, and to reveal subjets formed from the hadronization of the hard partons arising from the massive particle decay. The mass of the groomed jet is thus closer to that of the parent massive particle, and the subjets can be associated with its decay products.

The optimal pruning parameters are found to be $z_{cut} = 0.1$ and $D_{cut} = 0.5$. The choices of soft drop parameters are set to $z_{cut} = 0$ and $\beta = 0$. Higgs boson tagged (H tagged) jets are required to have a pruned mass in the range $105 - 135$ GeV while top quark tagged (t tagged) jets are required to have a soft drop mass within $110 - 210$ GeV. The soft drop subjets are used for both the H and t tagging. This combination of soft drop and pruning is found to have the best performance against pileup and other background rejection [45, 46].

Besides the soft drop and pruning algorithms, the N-subjettiness algorithm [47] is also used, which computes the variables τ_N for a jet. The variables τ_N quantify lobes of energy flow inside a jet. A jet compatible with two substructures would have values of the ratio τ_2/τ_1 much less than unity, as in a boosted $H \to b\bar{b}$. Likewise, a jet from boosted $t \to bW$, with three substructures, would have the value of τ_3/τ_2 much less than one. In contrast, jets with no substructure would exhibit larger values for both τ_2/τ_1 and τ_3/τ_2. Thus these variables provide a good discriminating power against multijet backgrounds. The ratio $\tau_2/\tau_1 < 0.6$ and $\tau_3/\tau_2 < 0.54$ and are used for H and t jet tagging respectively.

Subjet b tagging is used to further suppress backgrounds. The combined secondary vertex algorithm (CSVv2) identifies jets and subjets containing B hadrons using a combination of track

and secondary vertex-related variables [48]. It produces a neural network based discriminator output in the range $0-1$ with higher values indicating larger b jet probability. For H tagging, two subjets are required to be b tagged with CSVv2 discriminator greater than 0.605, which has a background rate of 10% per subjet and a signal efficiency of $40-70\%$, depending on the subjet p_{T} [49]. An otherwise H tagged jet but with both subjets failing the b tagging criteria (called "anti-Higgs tagged") is used to define a control region for background estimation. For t tagging, one subjet is required to have the CSVv2 discriminator greater than 0.79, which gives an overall t tagging mistag rate of 0.1% [50].

The H_{T} for each event is defined as $\sum p_{\mathrm{T}}$ of all AK4 jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 5$. $H_{\mathrm{T}} > 1100$ GeV is required, for which the trigger efficiency in data is found to be 100% with negligible error. The AK8 jets are required to have $p_{\mathrm{T}} > 300$ GeV and $|\eta| < 2.4$. The preselection criteria requires events to have at least 4 AK4 and at least 1 AK8 jets in addition to the H_{T}. The main analysis selection is based on the reconstruction of the T candidate invariant mass $M(\mathrm{T})$ using H and t tagged jets.

A discrepancy between the data and MC is observed, after the preselection, for variables like the jet p_{T} and H_{T}. A correction to the MC-based backgrounds estimated from simulations is extracted based on a two-parameter linear fit to the observed data/MC ratio of the H_{T} distribution. Extensive cross checks are performed in different control regions on the validity of this correction factor. The H_{T} reweighting is applied to the backgrounds estimated using MC simulations. The impact of the reweighting on the background $M(\mathrm{T})$ distribution was found to be small. Simulated signal events are not reweighted since such an effect cannot be verified using the data for the signal events. Besides, it was found that the effect of the reweighting on the signal would be well within the precision achievable using the current data.

The H and t tagged jets are required to have p_{T} greater than 300 GeV and 400 GeV, respectively. The analysis proceeds by choosing the highest p_{T} H and t tagged jet in the event. A separation in the $\eta - \phi$ plane, $\Delta R(\mathrm{H},\mathrm{t}) > 2.0$ is applied so as not to tag the same jet as both a Higgs and a top jet. The leading H and t tagged jet in each event are paired to form the T candidate with the total invariant mass being assigned as the T mass $M(\mathrm{T})$. The search is performed by looking for a localized excess in the T candidate mass distribution above the SM background. The reconstructed T quark mass distribution in the pp $\rightarrow$ Tbq production channel for a few representative masses is shown in Fig. 2. Table 1 gives some representative signal efficiencies for the pp $\rightarrow$ Tbq and pp $\rightarrow$ Ttq processes with LH couplings for different T quark masses. The efficiencies for the RH couplings are comparable to the LH couplings of the corresponding models.

5 Background estimation

The main backgrounds are the $\mathrm{t\bar{t}}$+jets and multijets, with traces of W+jets, the tW background being completely eliminated after the full event selection. These backgrounds are estimated using simulations. The multijet background is estimated from the data by using sideband distributions in four regions A, B, C, and D. Region D is the signal region, as defined in the previous section. Region B contains at least one anti-Higgs tagged jets and zero H tagged jet, along with a t tagged jet. Regions A and C contain zero t tagged jet and at least one anti-Higgs tagged and H tagged jets, respectively.

The two variables that span the region $ABCD$, the H/ anti-Higgs tagging and the t tagging criteria, are uncorrelated. Hence the number of events $N_{A,B,C,D}$ for the corresponding regions would follow the relation $N_A/N_B = N_C/N_D$. Thus, the number of background events in the

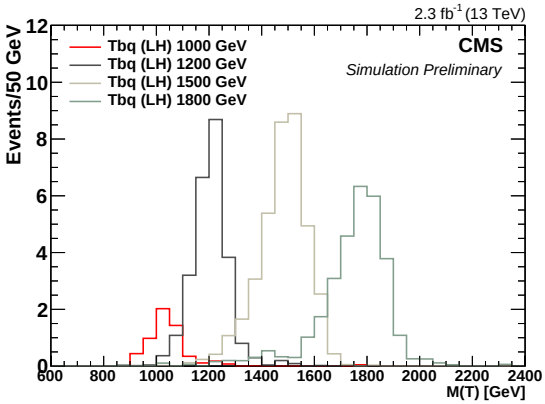

Figure 2: The $M(\mathrm{T})$ distributions for the pp $\rightarrow$ Tbq production mode with LH coupling for the T quark masses 1000, 1200, 1500, and 1800 GeV after all the selection criteria. The signal $\sigma\,\mathcal{B}(\mathrm{T} \rightarrow \mathrm{tH})$ is set to 1 pb.

Table 1: The signal efficiencies for successive event selections in percent for the pp $\rightarrow$ Tbq (LH) and pp $\rightarrow$ Ttq (LH) models. The statistical size of the simulated signal samples are much larger than the total integrated luminosity and hence the statistical errors on the efficiencies are negligible.

Mass (GeV)	pp $\rightarrow$ Tbq, T $\rightarrow$ tH (LH)			pp $\rightarrow$ Ttq, T $\rightarrow$ tH (LH)		
	Preselection	H tag	t tag	Preselection	H tag	t tag
1000	15.7	3.14	0.25	51.5	9.13	0.75
1100	24.9	5.74	0.76	61.1	11.4	1.07
1200	36.4	8.67	1.12	69.1	12.8	1.30
1300	44.6	10.6	1.37	74.8	13.8	1.43
1400	52.0	11.8	1.63	79.4	14.9	1.58
1500	57.1	12.6	1.66	82.1	15.1	1.61
1700	63.3	12.9	1.53	86.3	15.8	1.71
1800	65.6	13.2	1.35	87.9	15.5	1.52

signal region D would be given by counting the number of events in the three control regions: $N_D = N_B \times N_C / N_A$. The validity of the non-correlation was tested using MC simulations.

More than just predicting the background event yields, the $ABCD$ method is used to obtain the background $M(\mathrm{T})$ distribution for the signal region. The anti-Higgs and t tagged jets are paired to reconstruct the $M(\mathrm{T})$ shape in the control region B. This, when multiplied by the ratio N_C / N_A gives the background $M(\mathrm{T})$ shape in the signal region. A full closure test of the procedure is performed using simulations. The compatibility of the shapes in the regions B and D are verified using simulated QCD multijet samples. Moreover, the shapes of the distribution in the region B between the simulation and data are found to be consistent, and thus, a closure using simulated samples is deemed to hold also for the data.

Since only the multijets background is estimated using the $ABCD$ method, the simulated $\mathrm{t\bar{t}}$+jets, W+jets, and tW background are subtracted from the data in each of the A, B, and C regions to obtain only the predicted multijet background in data for that region. Then this multijet event distribution from the B region is scaled by the ratio of the events in region C to region A to obtain the predicted background distribution in the signal region. The number of events in the control regions are given in Table 2. The ratio N_C / N_A is found to be 13.6 ± 0.2.

Table 2: Numbers of events for the control region A, B, and C in the data, and the non-multijets background. The errors are statistical only.

	A	B	C
Data	94126 ± 307	207 ± 14	7251 ± 85
$\mathrm{t\bar{t}}$+jets	812 ± 6	53.9 ± 1.5	366 ± 4
W+jets	1258 ± 25	5.96 ± 1.76	109 ± 8
tW	27 ± 2	0.52 ± 0.20	12.7 ± 1.1
Data - $\mathrm{t\bar{t}}$+jets - W+jets - tW	92029 ± 308	147 ± 15	6764 ± 86

The total estimated background from all sources is given in Table 3, along with the number of observed events in the data. Finally, the H_{T} and $M(\mathrm{T})$ distributions in the data, estimated backgrounds, and the signal are shown in Fig. 3. The overall agreement between the observed number of events and the predicted background is good within the estimated uncertainties (discussed in Section 6).

Table 3: Estimated background and observed events in the signal region after all selection criteria. The combined statistical and systematic uncertainty is shown. The systematic uncertainty on the data-driven multijets background is anticorrelated with that on the MC-driven $\mathrm{t\bar{t}}$+jets background. Hence the uncertainty on the total background is less than what one would obtain if the uncertainties on the individual backgrounds are added in quadrature.

Process	Events
Estimated multijets (using data)	10.8 ± 5.5
Estimated $\mathrm{t\bar{t}}$+jets (using MC)	24.3 ± 8.1
Estimated W+jets (using MC)	0.6 ± 0.6
Estimated total background	35.7 ± 5.6
Observed events	30

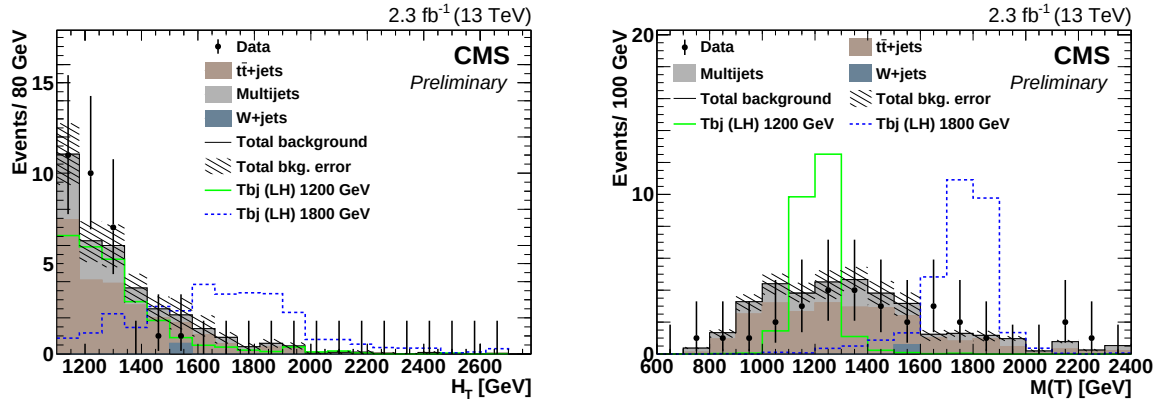

Figure 3: The H_{T} (left) and $M(\mathrm{T})$ (right) distributions after full event selection. The black markers with error bars are the data. The various background components are shown as filled histograms, and are estimated using simulations ($\mathrm{t\bar{t}}$+jets and W+jets) and the data (non-$\mathrm{t\bar{t}}$+jets and non-W+jets multijets component). The T quark signal distribution for two T quark masses are also shown. The signal $\sigma\,\mathcal{B}(\mathrm{T} \to \mathrm{tH})$ is set to 1 pb.

6 Systematic uncertainties

The systematic uncertainties can be classified into those that affect the total rate of the predicted background and the signal events, and those that change the shape of the $M(\mathrm{T})$ distributions. Among the former are the luminosity uncertainty of 2.7% [51]; the pileup reweighting uncertainty of 5% on the total inelastic proton-proton collision cross section; the cross section uncertainties on the MC background predictions; and the parton distribution functions at $1-3\%$, using the PDF4LHC recommended procedure [52]. The scale factor uncertainty due to the N-subjettiness selection for H tagging is 12.5%.

The jet energy and mass correction and resolution uncertainties affect the shapes of the $M(\mathrm{T})$ for both the simulated signal and background processes. The jet energy scale and resolution uncertainties are a few percent while jet mass correction uncertainty is 10%. The H_{T}-reweighting has an uncertainty of $1-3\%$.

The subjet b tagging and the t tagging scale factor uncertainties also affect the $M(\mathrm{T})$ shape. The t tagging scale factor uncertainty is the largest at about $15-30\%$ over the entire p_{T} range. The subjet b tagging scale factor systematic uncertainties are $2-5\%$ for subjets from b quarks, twice that for those from c quarks, and about 10% for light quark subjets.

The uncertainty on the predicted multijet background arises from the statistical uncertainty of the data sample used in the control regions as well as the above uncertainties which are propagated while subtracting the $\mathrm{t\bar{t}}$+jets, W+jets, and the single top contributions. Thus, the systematic uncertainties on the simulated backgrounds is anticorrelated with the data-driven multijets background estimation.

7 Results

Given that no excess of events is seen over the estimated background, we proceed to set limits on the production cross section of the T quark produced in association with a t or a b quark through electroweak interactions. The shape of the $M(\mathrm{T})$ candidate distributions for the background and the signal are fit to the data to obtain an upper limit on the production cross section

of a single T quark decaying to T → tH . The systematic uncertainties are treated as nuisance parameters with log-normal priors. A binned likelihood fit is used, with a Bayesian approach to evaluate the best fit values of the nuisance parameters [53], to estimate the 95% confidence level (CL) upper limit on the signal strength. The expected and observed limits are shown in Fig. 4 for different T quark masses, and with left- and right-handed couplings to the 3rd generation SM quarks. A comparison is made with a theoretical model of Ref. [20, 21] assuming $\mathcal{B}(T \to bW) = 0.5$ and $\mathcal{B}(T \to tH) = 0.25$ in the case of pp → Tbq, and assuming only neutral current processes with $\mathcal{B}(T \to tZ) = \mathcal{B}(T \to tH) = 0.5$ in the case of pp → Ttq. The limits are listed in Table 4.

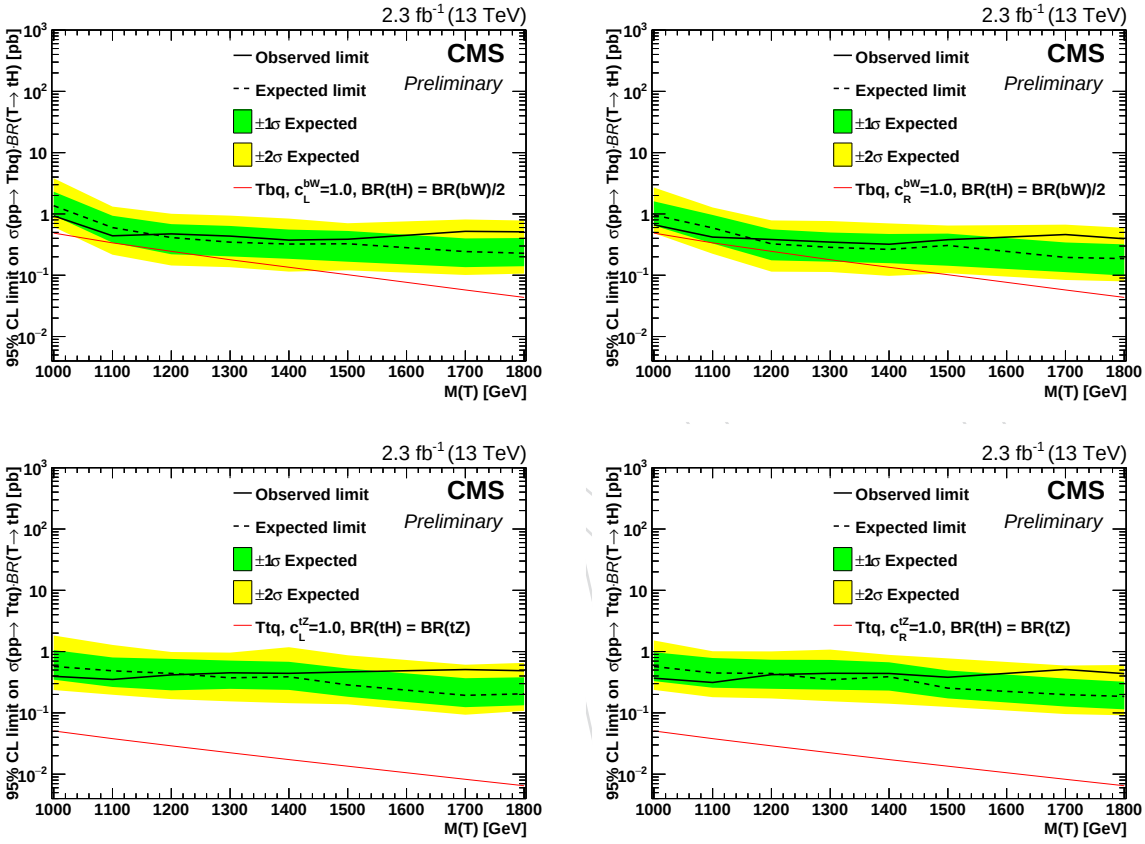

Figure 4: The expected and observed limits on the $\sigma\,\mathcal{B}(T \to tH)$ of the production processes pp → Tbq (top) and pp → Ttq (bottom), for different assumed values of the T quark mass, and with left handed (left) and right-handed (right) coupling to the SM 3rd generation.

8 Conclusion

A search for a vectorlike top quark partner T in the single production mode is performed using pp collision events at $\sqrt{s} = 13$ TeV collected by the CMS experiment in 2015. The T quarks are assumed to couple only to the SM 3rd generation quarks. The decay channel exploited is T → tH , with hadronic top quark decay, and H → $b\bar{b}$. Boosted H and t tagging are used to identify the top quark and Higgs boson decays in the final state, and the invariant mass of the two gives the T quark candidate mass. The background is mostly due to SM $t\bar{t}$+jets, with some amount of multijet and W+jets processes. No excess of data above the background is observed and 95% CL upper limits on $\sigma\,\mathcal{B}(T \to tH)$ are set varying between $0.31 - 0.93$ pb for T quark masses

Table 4: The observed and expected 95% CL upper limits on the cross sections (pb) of the T quark for the models pp $\rightarrow$ Tbq and pp $\rightarrow$ Ttq, and for different mass hypotheses.

Mass (GeV)	pp $\rightarrow$ Tbq (LH)		pp $\rightarrow$ Tbq (RH)		pp $\rightarrow$ Ttq (LH)		pp $\rightarrow$ Ttq (RH)	
	Limits in pb							
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
1000	0.93	1.36	0.66	0.96	0.40	0.57	0.37	0.57
1100	0.44	0.60	0.42	0.59	0.35	0.48	0.31	0.45
1200	0.47	0.41	0.38	0.32	0.42	0.44	0.42	0.44
1300	0.44	0.35	0.35	0.28	0.45	0.37	0.44	0.35
1400	0.37	0.32	0.32	0.26	0.44	0.39	0.44	0.39
1500	0.39	0.33	0.38	0.31	0.47	0.28	0.38	0.25
1700	0.52	0.24	0.46	0.20	0.51	0.19	0.51	0.20
1800	0.51	0.23	0.39	0.19	0.49	0.20	0.44	0.18

ranging from 1000 GeV to 1800 GeV in the pp $\rightarrow$ Tbq and pp $\rightarrow$ Ttq production channels with left and right handed couplings to the SM third generation quarks.

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centaurs and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Science, Research and Economy and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research Promotion Foundation, Cyprus; the Ministry of Education and Research, Estonian Research Council via IUT23-4 and IUT23-6 and European Regional Development Fund, Estonia; the Academy of Finland, Finnish Ministry of Education and Culture, and Helsinki Institute of Physics; the Institut National de Physique Nucléaire et de Physique des Particules / CNRS, and Commissariat à l'Énergie Atomique et aux Énergies Alternatives / CEA, France; the Bundesministerium für Bildung und Forschung, Deutsche Forschungsgemeinschaft, and Helmholtz-Gemeinschaft Deutscher Forschungszentren, Germany; the General Secretariat for Research and Technology, Greece; the National Scientific Research Foundation, and National Innovation Office, Hungary; the Department of Atomic Energy and the Department of Science and Technology, India; the Institute for Studies in Theoretical Physics and Mathematics, Iran; the Science Foundation, Ireland; the Istituto Nazionale di Fisica Nucleare, Italy; the Ministry of Science, ICT and Future Planning, and National Research Foundation (NRF), Republic of Korea; the Lithuanian Academy of Sciences; the Ministry of Education, and University of Malaya (Malaysia); the Mexican Funding Agencies (CINVESTAV, CONACYT, SEP, and UASLP-FAI); the Ministry of Business, Innovation and Employment, New Zealand; the Pakistan Atomic Energy Commission; the Ministry of Science and Higher Education and the

National Science Centre, Poland; the Fundação para a Ciência e a Tecnologia, Portugal; JINR, Dubna; the Ministry of Education and Science of the Russian Federation, the Federal Agency of Atomic Energy of the Russian Federation, Russian Academy of Sciences, and the Russian Foundation for Basic Research; the Ministry of Education, Science and Technological Development of Serbia; the Secretaría de Estado de Investigación, Desarrollo e Innovación and Programa Consolider-Ingenio 2010, Spain; the Swiss Funding Agencies (ETH Board, ETH Zurich, PSI, SNF, UniZH, Canton Zurich, and SER); the Ministry of Science and Technology, Taipei; the Thailand Center of Excellence in Physics, the Institute for the Promotion of Teaching Science and Technology of Thailand, Special Task Force for Activating Research and the National Science and Technology Development Agency of Thailand; the Scientific and Technical Research Council of Turkey, and Turkish Atomic Energy Authority; the National Academy of Sciences of Ukraine, and State Fund for Fundamental Researches, Ukraine; the Science and Technology Facilities Council, UK; the US Department of Energy, and the US National Science Foundation.

Individuals have received support from the Marie-Curie program and the European Research Council and EPLANET (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS program of Foundation for Polish Science, cofinanced from European Union, Regional Development Fund; the Compagnia di San Paolo (Torino); the Consorzio per la Fisica (Trieste); MIUR project 20108T4XTM (Italy); the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; and the National Priorities Research Program by Qatar National Research Fund.

References

[1] M. Schmaltz and D. Tucker-Smith, "Little Higgs review", *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229, doi:10.1146/annurev.nucl.55.090704.151502, arXiv:hep-ph/0502182.

[2] 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.

[3] I. Antoniadis, K. Benakli, and M. Quiros, "Finite Higgs mass without supersymmetry", *New J. Phys.* **3** (2001) 20, doi:10.1088/1367-2630/3/1/320, arXiv:hep-th/0108005.

[4] Y. Hosotani, S. Noda, and K. Takenaga, "Dynamical gauge-Higgs unification in the electroweak theory", *Phys. Lett.* **B607** (2005) 276, doi:10.1016/j.physletb.2004.12.029, arXiv:hep-ph/0410193.

[5] K. Agashe, R. Contino, and A. Pomarol, "The Minimal composite Higgs model", *Nucl. Phys.* **B719** (2005) 165, doi:10.1016/j.nuclphysb.2005.04.035, arXiv:hep-ph/0412089.

[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] P. H. Frampton, P. Q. Hung, and M. Sher, "Quarks and leptons beyond the third generation", *Phys. Rept.* **330** (2000) 263, doi:10.1016/S0370-1573(99)00095-2, arXiv:hep-ph/9903387.

[8] 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.

[9] J. A. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Prez-Victoria, "Handbook of vectorlike quarks: Mixing and single production", *Phys. Rev.* **D88** (2013) 094010, doi:10.1103/PhysRevD.88.094010, arXiv:1306.0572.

[10] Y. Okada and L. Panizzi, "LHC signatures of vector-like quarks", *Adv. High Energy Phys.* **2013** (2013) 364936, doi:10.1155/2013/364936, arXiv:1207.5607.

[11] ATLAS, 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.

[12] 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.* **B716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[13] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2012) 301, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[14] CMS Collaboration, "Search for vector-like charge 2/3 T quarks in proton-proton collisions at sqrt(s) = 8 TeV", *Phys. Rev.* **D93** (2016) 012003, doi:10.1103/PhysRevD.93.012003, arXiv:1509.04177.

[15] ATLAS Collaboration, "Search for production of vector-like top quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", ATLAS Conference Note ATLAS-CONF-2016-013, CERN, 2016.

[16] 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", (2016). arXiv:1602.05606. Submitted to EPJC.

[17] CMS Collaboration, "Search for Single Production of a Vector Like T Quark Decaying to a Higgs Boson and a Leptonically Decaying Top Quark", CMS Physics Analysis Summary CMS-PAS-B2G-15-008, CERN, 2016.

[18] CMS Collaboration, "Search for pair production of vector-like T quarks in the lepton plus jets final state", CMS Physics Analysis Summary CMS-PAS-B2G-16-002, CERN, 2016.

[19] ATLAS Collaboration, "Search for production of vector-like quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *JHEP* **08** (2015) 105, doi:10.1007/JHEP08(2015)105, arXiv:1505.04306.

[20] 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.

[21] J. M. Campbell, R. K. Ellis, and F. Tramontano, "Single top production and decay at next-to-leading order", *Phys. Rev.* **D70** (2004) 094012, doi:10.1103/PhysRevD.70.094012, arXiv:hep-ph/0408158.

[22] 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.

[23] R. D. Ball et al., "Parton distributions for the LHC Run II", *JHEP* **2015** (04) 40, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[24] 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.

[25] 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.

[26] 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.

[27] S. Frixione, G. Ridolfi, and P. Nason, "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *Journal of High Energy Physics* **2007** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[28] T. Sjostrand, S. Mrenna, and P. Z. 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.

[29] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[30] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Comput. Phys. Commun.* **182** (2011) 2388, doi:10.1016/j.cpc.2011.06.008, arXiv:1011.3540.

[31] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-", *Phys. Rev.* **D82** (2010) 054018, doi:10.1103/PhysRevD.82.054018, arXiv:1005.4451.

[32] N. Kidonakis, "Top Quark Production", in *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)*, p. 139. 2014. arXiv:1311.0283. doi:10.3204/DESY-PROC-2013-03/Kidonakis.

[33] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

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

[35] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CERN, 2009.

[36] 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.

[37] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett.* **B641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[38] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[39] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.* **B659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[40] 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.

[41] CMS Collaboration, "Jet Performance in pp Collisions at $\sqrt{s}$=7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, CERN, 2010.

[42] G. P. Salam, "Towards Jetography", *Eur. Phys. J.* **C67** (2010) 637, doi:10.1140/epjc/s10052-010-1314-6, arXiv:0906.1833.

[43] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches", *Phys. Rev.* **D81** (2010) 094023, doi:10.1103/PhysRevD.81.094023, arXiv:0912.0033.

[44] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", *JHEP* **1405** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[45] CMS Collaboration, "Identification techniques for highly boosted W bosons that decay into hadrons", *JHEP* **12** (2014) 017, doi:10.1007/JHEP12(2014)017, arXiv:1410.4227.

[46] CMS Collaboration, "Boosted Top Jet Tagging at CMS", CMS Physics Analysis Summary CMS-PAS-JME-13-007, CERN, 2013.

[47] 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.

[48] 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.

[49] 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, 2016.

[50] CMS Collaboration, "Top Tagging with New Approaches", CMS Physics Analysis Summary CMS-PAS-JME-15-002, CERN, 2016.

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

[52] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys.* **G43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[53] The ATLAS and the CMS Collaborations, and the LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, 2011.