Improving third-generation leptoquark searches with combined signals and boosted top quarks

by

Arvind Bhaskar, Tanumoy Mandal, Subhadip Mitra, Mohit Sharma

in

PHYSICAL REVIEW D

Report No: IIIT/TR/2021/-1

Centre for Computational Natural Sciences and Bioinformatics
International Institute of Information Technology
Hyderabad - 500 032, INDIA
October 2021
Improving third-generation leptoquark searches with combined signals and boosted top quarks

Arvind Bhaskar, Tanumoy Mandal, Subhadip Mitra, and Mohit Sharma

1Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500 032, India
2Indian Institute of Science Education and Research Thiruvananthapuram, Vithura, Kerala 695 551, India

(Received 23 June 2021; accepted 24 September 2021; published 27 October 2021)

Search strategies for the third-generation leptoquarks (LQs) are distinct from other LQ searches, especially when they decay to a top quark and a τ lepton. We investigate the cases of all TeV-scale scalar and vector LQs that decay to either a top-tau pair (charge-1/3 and charge-5/3 LQs) or a top-neutrino pair (charge-2/3 LQs). One can then use the boosted top (which can be tagged efficiently using jet-substructure techniques) and high-p_T τ leptons to search for these LQs. We consider two search channels with either one or two taus along with at least one hadronically decaying boosted top quark. We estimate the high luminosity LHC (HL-LHC) search prospects of these LQs by considering both symmetric and asymmetric pair and single production processes. Our selection criteria are optimized to retain events from both pair and single production processes. The combined signal has better prospects than the traditional searches. We include new three-body single production processes to enhance the single production contributions to the combined signal. We identify the interference effect that appears in the dominant single production channel of charge-1/3 scalar LQ (S_{1/3}). This interference is constructive if S_{1/3} is a weak-triplet and destructive for a singlet one. As a result, their LHC prospects differ appreciably.

DOI: 10.1103/PhysRevD.104.075037

I. INTRODUCTION

Recently leptoquarks (LQs) are receiving a lot of attention in the literature. Since they connect the quark and the lepton sectors of the Standard Model (SM), their phenomenology is interesting in general. However, the recent attention is mainly due to their ability to explain the long-standing anomalies observed in B-meson decays and the muon anomalous magnetic moment measurements (see, e.g., [1–34] for a few recent studies). Generally, the models explaining the anomalies contain LQs that couple with the third-generation quarks and leptons. This gives additional motivation to search for such LQs (commonly known as the third-generation LQs in the literature) at the Large Hadron Collider (LHC).

Various strategies have already been applied in the searches for third-generation LQs at the LHC [35–42].

In this paper, we consider a new search strategy for the third-generation scalar and vector LQs (sLQs and vLQs) that couple to a top quark and produce a boosted-top signature. We consider two signatures to estimate their discovery/exclusion prospects at the high-luminosity LHC (HL-LHC)—either one or two τ-leptons in association with at least one hadronically-decaying boosted top quark. This paper, thus, complements our previous prospect studies on LQs that dominantly decay to a top quark and a light charged lepton (either e or μ), producing a top-lepton resonance [11,19]. The τ-leptons are different from light leptons—they are unstable, and their decays always involve some missing energy making the τ-tagging efficiency lower than that for the light leptons. Therefore, analyzing the signatures involving τ-leptons needs additional care.

The two-τ signature is for charge-1/3 and charge-5/3 LQs, whereas the one-τ signature is for LQs with charge-2/3 and charge-1/3. We look for at least one hadronically-decaying boosted top-jet that can be tagged with the jet-substructure technique in both cases. Using this technique, one can tag a top-jet with good significance and enhance the LHC reach. In a previous paper [43], we argued that it is possible to improve the LHC discovery prospect of LQs (or any other new colored particles, for that matter) or exclusion limits if we systematically combine pair and single production channels. In this case, the two signatures capture combinations of pair and single production processes in

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP³.
the signal. Later, in Ref. [10], we highlighted the importance of considering asymmetric channels for LQ searches. Here, we consider both symmetric and asymmetric channels and optimize the selection criteria to retain a substantial fraction of pair and single production events. We note, recently, the CMS Collaboration has also used a similar asymmetric channel ($tb\tau\nu$ signature) and combined the pair and single production channels in their search for the third-generation LQs [40].

Generally, the simplest two-body single production processes are considered for the LQ searches in single production channels. However, there are three-body single production processes of LQs that are also important and have comparable or larger cross sections than the two-body single productions [43]. In this paper, we consider the three-body single productions and show that the combined signal can lead to better exclusion or discovery prospects.

Since the single-production cross sections depend on the LQ couplings, our results depend on LQ models. We have found an important point; for the same strength of the LQ couplings and decays, the HL-LHC discovery reach for the charge-1/3 component of the triplet scalar $S_3$ is higher than that of the singlet $S_1$, another charge-1/3 sLQ. This happens because of constructive interference among some subprocesses in the $pp \rightarrow S_3^{3/3}\tau_j l_j$ single production process. For $S_1$, this interference is destructive in nature. Obviously, a similar difference also appears in the exclusion limits of these two sLQs. Hence, one has to be mindful about the type of LQ and look beyond branching ratios while reading the exclusion limits on charge-1/3 sLQs. This interference effect is observed only in some channels and in case of charge-1/3 LQ models; other LQ species we have considered in this paper do not show such an effect.

Before we proceed further, we note that since this paper is a follow up to Refs. [11,19], we frequently refer to these papers and omit some common details while ensuring our current presentation is self-contained. The rest of the paper is a follow up to Refs. [11,19], we frequently refer to these papers [44,45] but the current symbols are taken from Ref. [46]). To avoid proton-decay constraints, we assume the diquark operators. For simplicity, we also assume both the Pontecorvo-Maki-Nakagawa-Sakata neutrino-mixing matrix and the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix to be unity. This is justified since the LHC is blind to the flavor of the neutrinos and the small off-diagonal terms in the CKM matrix would not play a significant role in our analysis.

**A. Scalar leptoquark models**

1. $S_1 = (\bar{3}.1.1/3)$

   The interaction Lagrangian of the sLQ $S_1$ can be written as follows:

   \[
   \mathcal{L} \supset y_{3ij}^{LL} \bar{Q}_L^{Ci} S_1 i\sigma^2 L_j^i + y_{3ij}^{RR} \bar{u}_R^{Ci} S_1 l_R^i + \text{h.c.},
   \]

   where $u_R$ and $l_R$ are a SM right-handed up-type quark and a charged lepton, respectively and $Q_L$ and $L_L$ are the SM left-handed quark and lepton doublets, respectively. The superscript $C$ denotes charge conjugation and $\sigma^2$ is the second Pauli matrix. The generation indices are denoted by $i, j = \{1, 2, 3\}$. The color indices are suppressed. Since the LHC cannot distinguish between the flavors of the neutrinos, we write them as $\nu$. The terms relevant for our analysis are given as

   \[
   \mathcal{L} \supset y_{3ij}^{LL} (\bar{\tau}_1^L \nu_L + \bar{\tau}_2^L \tau_L) S_1 + y_{3ij}^{RR} \bar{\tau}_1^L \tau_L S_1 + \text{h.c.}
   \]

2. $S_3 = (\bar{3}.3.1/3)$

   There is only one type of renormalizable operator one can write for $S_3$ that is invariant under the SM gauge group

   \[
   \mathcal{L} \supset y_{3iij}^{LL} \bar{Q}_L^{Ci} e^{ab} (\bar{\tau}_3^L S_3 )^{bc} L_L^i + \text{h.c.}
   \]

   Here, the SU(2) indices are denoted by $a, b, c = \{1, 2\}$. Expanding and keeping only the relevant interaction terms, we get,

   \[
   \mathcal{L} \supset -y_{3iij}^{LL} (\bar{\tau}_1^L \nu_L + \bar{\tau}_2^L \tau_L) S_3^{1/3} + \sqrt{2} (\bar{\tau}_1^L \tau_3 L_L^{S_3^{4/3}} - \bar{\tau}_2^L \nu_L S_3^{2/3}) + \text{h.c.}
   \]

3. $R_2 = (3.2.7/6)$

   Similarly, for $R_2$ we have the following terms,

   \[
   \mathcal{L} \supset -y_{2iij}^{RL} \bar{u}_R^{Ci} R_2^{ab} L_L^b + y_{2iij}^{RR} \bar{\tau}_R^{Ci} Q_R^{ia} + \text{h.c.},
   \]

   which, after expansion and considering the relevant terms, can be written as

   \[
   \mathcal{L} \supset -y_{2iij}^{RL} \bar{\tau}_R^{Ci} R_2^{5/3} + y_{2iij}^{RR} \bar{\tau}_R^{Ci} Q_R^{ia} + y_{2iij}^{RL} \bar{\tau}_R^{Ci} Q_R^{ia} + \text{h.c.}
   \]
B. Vector leptoquark models

1. $U_1 = (3.1.5/3)$

The electric charge of $U_1$ is $5/3$. Hence, it couples exclusively with the right-handed leptons:

$$L \supset \bar{x}^{RL}_{ij} \gamma^{ij}_R \nu^\mu U_{1,\mu} l_R^\mu + \text{H.c.}$$

For our purpose, we have

$$L \supset \bar{x}^{RR}_{i3} \gamma_R \epsilon_{i3} U_{1,\bar{\nu}} + \text{H.c.}$$

2. $U_1 = (3.1.2/3)$

The necessary interaction terms for the charge-2/3 $U_1$ can be written as

$$L \supset x_{1ii} \tilde{U}^i_R \nu^\mu U_{1,\nu} l_R^\mu + x_{2ii} \widetilde{\nu}^i_R \nu^\mu U_{1,\nu} l_R^\mu + \text{H.c.}$$

The $i, j = 3$ terms can be written explicitly as

$$L \supset x_{1ii} \{ \tilde{R} \nu \cdot U_{1,\nu} l_R^\mu + \tilde{L} \nu \cdot U_{1,\nu} l_R^\mu \}$$

3. $V_2 = (\bar{3}, 2.5/6)$

For $V_2$, the Lagrangian is as follows:

$$L \supset \bar{x}^{RL}_{ij} \gamma^{ij}_L \nu^\mu V_{aL}^{ab} l_R^b + \bar{x}^{RR}_{ij} \gamma^{ij}_L \nu^\mu V_{aL}^{ab} l_R^b + \text{H.c.}$$

Expanding the Lagrangian we get the terms relevant for our analysis as

$$L \supset -x_{23}^{RL} \nu \gamma^{2/3} L + \gamma^{2/3} \nu L \gamma + \text{H.c.}$$

4. $V_2 = (\bar{3}, 2, -1/6)$

For $V_2$, the Lagrangian is given as

$$L \supset \bar{x}^{RL}_{ij} \gamma^{ij}_L \nu^\mu V_{aL}^{ab} l_R^b + \text{H.c.}$$

The terms with the third-generation fermions can be written as

$$L \supset x_{23}^{RL} \gamma^{2/3} L + \gamma^{2/3} L \gamma + \text{H.c.}$$

5. $U_3 = (3.3.2/3)$

The necessary interaction terms for the triplet $U_3$ are given as

$$L \supset x_{3ii} \tilde{O}^i_R \nu^\mu \epsilon^{ijk} \tilde{L}^{j} + \text{H.c.}$$

where $\sigma^k$ denotes the Pauli matrices. The terms for the third-generation fermions can be written explicitly as

$$L \supset x_{333}^{RL} \{ -\tilde{b} \nu \cdot U \nu L + \tilde{l} \nu \cdot U \nu L \}$$

$$+ \text{H.c.}$$

C. Simplified models and benchmark scenarios

The models listed above can also be expressed in terms of some simple phenomenological Lagrangians,

$$L \supset \lambda \{ \sqrt{\eta^2} \tilde{L} \nu \cdot U \nu L + \text{H.c.} \}$$

$$+ \lambda \tilde{L} \nu \cdot U \nu L + \text{H.c.}$$

$$L \supset \lambda \{ \sqrt{\eta^2} \tilde{L} \nu \cdot U \nu L + \text{H.c.} \}$$

$$+ \lambda \tilde{L} \nu \cdot U \nu L + \text{H.c.}$$

In the simple expressions, a sLQ (vLQ) with charge $= \pm 1/3$ is denoted as $\phi_n (\phi_n)$. Here, $\eta_1$ and $\eta_2$ are the charged lepton chirality fractions [11,43]. The $\epsilon_R = \pm 1$ in Eq. (20) is a sign term to accommodate a possible relative sign between the left-handed and right-handed terms [see Eq. (11)]. We shall consider only real couplings in our analysis for simplicity.

We illustrate how one can map the standard LQ models to the simplified Lagrangians in Table I with some benchmark scenarios. The scenarios are named according to lepton chiralities following Refs. [11,19]. For example, in the left-coupling (LC) scenarios, the lepton that couples to the LQ along with a top quark is left-handed. Similarly, in the right-coupling (RC) scenarios, it is right-handed. In LCSS (LCOS) scenarios, the LQ couples to a left-handed $\tau$ or a $\nu$ via couplings of equal magnitudes with the same (opposite) sign. In the RLCS and RLCOS scenarios, the LQ couples with $\nu$ and $\tau_R$ with same-sign or opposite-sign couplings. Note, however, the couplings need not be of the same magnitude in these cases.
TABLE I. Summary of the benchmark scenarios showing the map between the known LQ models [Eqs. (6)–(15)] and the simple models [Eqs. (16)–(21)]. The branching ratio $\beta$ for a $\phi/\chi$ to decay to a top quark is fixed in all scenarios, except for $U_1$ in the vector LCSS scenario ($\beta \leq 50\%$) and $R_{1/2}^{23/3}$ ($V_{12}^{33/3}$) in the scalar (vector) RLCSS/OS scenarios where $0 \leq \beta < 100\%$ [for $\beta = 100\%$, the scalar (vector) RLCSS/OS scenarios become the same as the scalar LC (vector RC) scenario]. An asterisk marks exceptional scenarios. Here, $\lambda/\Lambda$ is a generic free-coupling parameter. We have chosen only this one coupling to control all of the nonzero new couplings in every benchmark for simplicity. This essentially means also choosing $\beta$ to be $50\%$ in the exceptional scenarios.

<table>
<thead>
<tr>
<th>Benchmark scenario</th>
<th>Possible charge(s) of LQ couplings equal to chirality fraction $\lambda/\Lambda$</th>
<th>Nonzero coupling Type of LQ</th>
<th>Charged lepton decay mode(s) $(s)$ ${\beta, 1 - \beta}$</th>
<th>Branching ratios $(%)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar LC</td>
<td>$2/3 \phi_2, \tilde{\phi}_2$</td>
<td>$\tilde{\lambda}_v = \lambda_v$</td>
<td>$\tau$</td>
<td>${100%, 0}$</td>
</tr>
<tr>
<td>LCS1</td>
<td>$1/3 \phi_1, \tilde{\phi}_1$</td>
<td>$\tilde{\lambda}<em>\ell = \lambda</em>\ell$</td>
<td>$\tau$</td>
<td>${50%, 50%}$</td>
</tr>
<tr>
<td>vector RC</td>
<td>$1/3 \phi_1, \tilde{\phi}_1$</td>
<td>$\tilde{\lambda}<em>\ell = \lambda</em>\ell$</td>
<td>$\tau$</td>
<td>${100%}$</td>
</tr>
<tr>
<td>RLCSS*</td>
<td>$2/3 \phi_2, \tilde{\phi}_2$</td>
<td>$\tilde{\lambda}<em>\ell = \lambda</em>\ell$</td>
<td>$\tau$</td>
<td>${100%, 0}$</td>
</tr>
<tr>
<td>RLCOS*</td>
<td>$2/3 \phi_2, \tilde{\phi}_2$</td>
<td>$\tilde{\lambda}<em>\ell = \lambda</em>\ell$</td>
<td>$\tau$</td>
<td>${100%, 0}$</td>
</tr>
<tr>
<td>Vector LC</td>
<td>$1/3 \chi_1, \tilde{\chi}_1$</td>
<td>$\tilde{\lambda}_v = \lambda_v$</td>
<td>$\nu$</td>
<td>${50%, 50%}$</td>
</tr>
<tr>
<td>vector RC</td>
<td>$1/3 \chi_1, \tilde{\chi}_1$</td>
<td>$\tilde{\lambda}_v = \lambda_v$</td>
<td>$\nu$</td>
<td>${50%, 50%}$</td>
</tr>
<tr>
<td>RLCSS*</td>
<td>$2/3 \chi_2, \tilde{\chi}_2$</td>
<td>$\tilde{\lambda}<em>\ell = \lambda</em>\ell$</td>
<td>$\nu$</td>
<td>${100%, 0}$</td>
</tr>
<tr>
<td>RLCOS*</td>
<td>$2/3 \chi_2, \tilde{\chi}_2$</td>
<td>$\tilde{\lambda}<em>\ell = \lambda</em>\ell$</td>
<td>$\nu$</td>
<td>${100%, 0}$</td>
</tr>
</tbody>
</table>

In the table, we also show the decays and branching ratios (BRs) of each LQ assuming it has no other decay modes. In the LC and RC scenarios, the LQ decays to a top quark with 100% BR. In contrast, in the LCS1/OS scenarios, it does so with about 50% BR [except for the vector LCSS scenario where a $U_1$ can decay to a top quark and a $\nu$ with less than 50% BR depending on the $x_{133}$ coupling, see Eq. (9)]. For the RLCSS/OS scenarios, we choose the top-quark decay mode BRs to be 50%. The discovery/exclusion potential for these LQs mostly depend on the BRs and masses. Hence, scenarios with equal BRs would have similar potentials, except for the scalar LCSS/OS scenarios where the interference between some single production diagrams becomes constructive or destructive depending on the relative sign of the couplings [11]. In the table, we still show these scenarios to demonstrate how the simple Lagrangians (suitable for a straightforward interpretation of experimental results) can cover all the possibilities.

In the case of vLQs, the kinetic terms of the Lagrangian contains an additional parameter $\kappa$ [46]

$$\mathcal{L} \supset - \frac{1}{2} \chi^{\dagger}_{\mu \nu} \chi^{\mu \nu} + M_{\chi} \chi^{\dagger}_{\mu \nu} - i g_\kappa \chi^{\dagger}_{\mu} T^a \chi_{\nu} G^{a \mu \nu},$$

(22)

where $\chi^{\dagger}_{\mu \nu}$ stands for the field-strength tensor of $\chi$. We found that the cross section of the pair and single production processes depends on $\kappa$. Here, we consider two benchmark values, $\kappa = 0$ and $\kappa = 1$.

III. LHC PHENOMENOLOGY AND SEARCH STRATEGY

We summarize the computational packages used in our analysis. The Lagrangian terms in Eqs. (16)–(21) have been encoded in FeynRules [47] to generate UFO [48] model files. We use MADGRAPH5 [49] to generate the signal and background processes at the leading order (LO). Only the pair production cross sections of sLQs are corrected with a next-to-leading order (NLO) QCD $K$-factor of 1.3 [50,51]. The higher-order corrections to the SM background processes are known in the literature and their cross sections are corrected using the appropriate $K$-factors. We use NNPDF2.3LO [52] parton distribution functions with dynamical renormalization and factorization scales set equal to the mass of the LQs. The parton level events are passed through PYTHIA6 [53] for showering and hadronization. We then use DELPHES3 [54] for simulating the detector effects with the default CMS card (we only modify the $b$-tagging efficiency to set it at a flat 85%). We reconstruct fatjets from DELPHES tower objects using the anti-$k_t$ [55] clustering algorithm (with $R = 0.8$) in FastJet [56].
A. Production at the LHC

Both the sLQs and vLQs can be produced resonantly at the LHC via the pair production and the single production processes. The pair production processes are almost independent of $\lambda$; they mostly depend on the universal QCD coupling. In the case of vLQs, the free parameter $\kappa$ from Eq. (22), can affect both pair and single production cross sections. The pair production channels can have the following final states,

$$pp \rightarrow \begin{cases} \phi_1/\chi_1 \tau \rightarrow (\tau \tau)/(\tau \nu)/(\nu \nu) \\ \phi_1/\chi_1 b \nu \rightarrow (\tau \nu)/(\nu \nu)/(\nu b) \\ \phi_1/\chi_1 \tau j \rightarrow (\tau \nu j)/(\nu j) \\ \phi_1/\chi_1 j \nu \rightarrow (\nu j)/(j j) \end{cases}.$$  

(23)

In our previous papers [11,19], we considered the symmetric pair production channels where both the LQs decay to the same final state and considered single productions yielding similar final states. In this paper, we also explore the asymmetric pair production channels with $(\tau \tau)/(\tau \nu)/(\nu \nu)/(\nu b)/b/(b b)$ final states while ignoring final states without a top quark or a $\tau$. We include single production channels giving similar final states (we refer to these as the asymmetric single production channels). Though it is generally assumed that the symmetric modes are more promising for discovery [57], due to combinatorics, the asymmetric modes can also provide comparable (or, in some cases, even better) reach. From Table I, we see that the charge-1/3 or charge-5/3 LQs can have the $(\tau \tau)/(\tau \nu)/(\nu \nu)/(\nu b)/b/(b b)$ final states, whereas and $t b \nu$ can come from charge-1/3 or charge-2/3 LQs.

A three-body single production, where a LQ is produced in association with a $\tau$ and either a jet or a top quark, can produce similar final states as the pair production. This allows us to combine pair and single productions in the same analysis and enhance the discovery reach [11,19,43,58,59]. We show the possible symmetric and asymmetric single production modes below.

$$pp \rightarrow \begin{cases} \phi_1/\chi_1 \tau \rightarrow (\tau \nu)/(\nu \nu) \\ \phi_1/\chi_1 b \nu \rightarrow (\tau \nu)/(\nu \nu)/(\nu b) \\ \phi_1/\chi_1 \tau j \rightarrow (\tau \nu j)/(\nu j) \\ \phi_1/\chi_1 j \nu \rightarrow (\nu j)/(j j) \end{cases}.$$  

(24)

$$pp \rightarrow \begin{cases} \phi_2/\chi_2 \nu \rightarrow (\nu \nu)/(\nu \nu)/(\nu \nu) \\ \phi_2/\chi_2 \nu \rightarrow (\nu \nu)/(\nu \nu)/(\nu \nu) \\ \phi_2/\chi_2 \nu \rightarrow (\nu j)/(j j) \end{cases}.$$  

(25)

$$pp \rightarrow \begin{cases} \phi_5/\chi_5 \tau \rightarrow (\tau \tau)/(\tau \tau)/(\tau \tau) \\ \phi_5/\chi_5 \tau j \rightarrow (\tau \nu j)/(\nu j) \end{cases}.$$  

(26)

Of these, we ignore the modes without at least one top quark and/or a $\tau$. In principle, there are two $bg$ initiated two-body single production processes also where a LQ is produced in association with a $\tau$ or a $\nu$. We ignore them to avoid double counting with the $gg$ initiated three-body processes where a gluon splits into a pair of $b$ quarks. This is justified since the $b$ quark has to be hard in order to produce a TeV-scale LQ.

In Fig. 1, we show the parton-level cross sections of different production processes of scalar (left column) and vector (right column) LQs as functions of their masses. We compute the cross sections with $\lambda = 1$ for the sLQs, and $\Lambda = 1$ and $\kappa = 1$ for the vLQs. From these figures, we get an idea about the relative magnitudes of the cross sections and the relative importance of processes (a priori to the cuts) in the mass range of our interest. In Fig. 1(a), two types of LC scenarios, namely the LCOS and LCSS scenarios are separately shown for the $pp \rightarrow \phi_1 \tau j$ single production process. Both scenarios have the same BRs of $\phi_1$, i.e., 50% each in $\tau \tau$ and $b \nu$ modes. But the cross section for the LCSS scenario is bigger than that in the LCOS scenario. This is because, the $\lambda_{\mu} = \lambda_{\nu}$ relation in the LCSS scenario results in a constructive interference between some of the single production diagrams which becomes destructive in the LCOS scenario where $\lambda_{\nu} = -\lambda_{\mu}$. Such difference is not observed for other LQ species.

Single production cross sections vary as the square of the new coupling $\lambda$ or $\Lambda$. In the figures all processes have been computed for $\lambda = \Lambda = 1$. With order one $\lambda$ (or $\Lambda$), it is possible for some single production modes to have bigger cross sections than the pair production in the mass range of our interest. For example, we see that in the LC scenario, $pp \rightarrow \chi_1 \tau j$ overtakes the pair production processes at 2.6 TeV for $\kappa = 1$ (Fig. 1(b)) but the cross section for the similar $\chi_1 \tau j$ process remains lower. Similar behavior can be seen for the $pp \rightarrow \phi_1 \tau j$ process also. However, in some scenarios, these processes have much smaller cross section. For example, in the RC scenario, the $pp \rightarrow \phi_1 \tau j \chi \tau j \chi \tau j$ processes have much smaller cross section than the rests in the entire mass range considered. Similarly, $pp \rightarrow \phi_3 \tau j$ have small cross section in the LC scenario. This is mainly because, in these scenarios, the LQs are produced from a right-handed top quark generated in the charge current interaction via a chirality flip. In the figure, we just show the plots for the vLQs with $\kappa = 1$. If $\kappa = 0$, the pair production cross sections become smaller and hence the crossover points appear at lower masses.

B. Signal topologies and the background

From Eqs. (23)–(26) we see that if we disregard the final states without a single top quark or a $\tau$, the remaining ones can be identified as of two types—one with one top quark and two $\tau$’s and the other with one top quark and only one $\tau$. As before, we will consider only hadronically-decaying top quarks in our analysis to exploit its boosted nature. A $\tau$ can decay either hadronically or leptonically with branching
FIG. 1. The parton-level cross sections of different production channels of $\phi_i$ [(a), (c), (e)] and $\chi_i$ [(b), (d), (f)] at the 14 TeV LHC as functions of $M_{\phi_i}/\phi_i$. The single production cross sections are computed for a benchmark coupling $\lambda, \Lambda = 1$ (see Table I). For the VLQs, we have shown the scenarios with QCD coupling $\kappa = 1$. Here, we consider $\tau j$ and $t j$ in the single production processes and the $j$ includes all the light jets including $b$-jets. Their cross sections are generated with a cut on the transverse momentum of the jet, $p_T > 20$ GeV.
fractions of about 65% and 35%, respectively. This gives three possible final states in terms of $\tau$ decays—$t \tau_h$, when both of them decay hadronically with a probability about 42%, $t \tau_h \tau$, when one of them decays hadronically and the other leptonically with about 46% probability, and $\tau \tau_h \tau$, when both of them decay leptonically with a probability of only 12%. In our analysis, we consider both hadronic and semileptonic $\tau$s and study the following signatures:

(A) at least one hadronically-decaying top quark along with either two hadronically-decaying $\tau$ leptons ($t_1 \tau_h \tau_h$) or a hadronically-decaying $\tau$ and a leptonically-decaying one ($t_1 \tau_h \tau$),

(B) at least one hadronically-decaying top quark with only one $\tau$ decaying hadronically and some missing energy ($t_2 \tau_h + \text{MET}$).

The first signature would capture the symmetric final states, while the second one, the asymmetric ones. The other symmetric final state with a pair of top quarks leading to the $t_2 t_2 + \text{MET}$ signature ($ttu\bar{u}$) has been studied in Ref. [60]. As already discussed, the above signal topologies can arise from both pair and single production processes. For instance, the signal $t_1 \tau_h \tau_h$ can come from $pp \rightarrow \chi_1, \phi_1 \chi'_{1,5}$ (or $pp \rightarrow \phi_1, \phi_1 \phi_1$) and $pp \rightarrow \chi_1, \phi_1 \phi_1$ (or $pp \rightarrow \phi_1, \phi_1, \phi_1$) processes. If not carefully done, this might lead to double counting while generating the signal processes. To circumvent this issue, we ensure that the LQ and its antiparticle are never on shell simultaneously while generating single production events [43].

The SM background of our signatures is very large and requires carefully designed kinematic cuts to make the signal observable. The SM background for both $(t_1 \tau_h \tau_h + t_2 \tau_h \tau)$ and $(t_1 \tau_h \tau_h + \text{MET})$ signals would contain at least one toplike fatjet. In the SM, it can arise directly from a top quark decaying hadronically or indirectly when a bunch of QCD jets combine to mimic the features of a toplike fatjet. In addition, the SM background should also contain $\tau$-tagged high-$p_T$ jets (from hadronic $\tau$ decays) and high-$p_T$ light leptons (since there is a leptonic $\tau$ decay in the signal). Although small, a QCD jet can sometime appear as a $\tau_h$ due to mistagging. Since some background processes have huge cross sections, even small mistagging rates might lead to a substantial number of background events at the end. The relevant background processes for the two signatures are listed in Table II where the available highest-order values of their total cross sections are shown. We generate all the background processes at LO and scale the cross sections with the appropriate $K$ factors. Notice that some of the background processes have very high cross sections. For these processes, we apply generation-level cuts to save computation time and improve the statistics. We briefly discuss the background processes below.

For both signal topologies, the dominant background is from the inclusive $t\bar{t} + \text{jets}$ process. All three $t\bar{t}$ decay modes, i.e., hadronic, semileptonic and leptonic contribute to the background—the semileptonic mode contributes dominantly followed by the leptonic and the hadronic modes. For the one-$\tau$ signature, the second dominant background is $W + \text{jets}$ whereas for the two-$\tau$ category, it is $tW$. We have included other relevant but subdominant background processes like $V + \text{jets}$, $V + W + \text{jets}$, $t + \text{jets}$, and $ttV$, etc., in our analysis.

### C. Analysis

For both signal topologies, we use the anti-$k_T$ clustering algorithm and make use of two different levels of jet information—we use two types of jets with different values of the jet radius parameter $R$. We call the jets with $R = 0.4$ “AK4-jets” and the $R = 0.8$ ones as “AK8-fatjets”. We use the Delphes tower objects to construct the AK8-fatjets and identify the hadronic top from them with the following criteria:

- Mass of the AK8-fatjet, $M_{fj} > 120$ GeV and its transverse momentum $p_T > 200$ GeV.
- The subtjetness ratios $\tau_{32} < 0.81$ and $\tau_{21} > 0.35$ where $\tau_{ij} \equiv \tau_i/\tau_j$.

The subtjetness criteria is motivated by the fact that $\tau_N$ assumes smaller values as the fatjet closely resembles a collection of $N$ subjets. The AK4-jets are used to identify the $\tau_{h}$- and $b$-tagged jets.

We list the selection cuts used for the two signal categories sequentially (cuts-flow) in Table III. We use cuts on the invariant mass and the transverse mass $M_T$ of the leptons. The transverse mass is defined as

\begin{align}
M_T^2(A, B) &= 2p_T^A p_T^B \{1 - \cos \Delta \phi(p_T^A, p_T^B)\}, \\
M_T^2(A, B, \mathcal{E}_T) &= M_T^2(A, B) + M_T^2(A, \mathcal{E}_T) \\
&\quad + M_T^2(B, \mathcal{E}_T).
\end{align}

To make sure that there is no overlap between a $\tau_h$ and the hadronic top, we demand a radial separation between the

<table>
<thead>
<tr>
<th>Background processes</th>
<th>$\sigma$ (pb)</th>
<th>QCD order</th>
</tr>
</thead>
<tbody>
<tr>
<td>V+ jets [61,62]</td>
<td>6.33 x 10^4</td>
<td>NNLO</td>
</tr>
<tr>
<td>W+ jets</td>
<td>1.95 x 10^5</td>
<td>NLO</td>
</tr>
<tr>
<td>$VV+\text{jets}$ [63]</td>
<td>124.31</td>
<td>NLO</td>
</tr>
<tr>
<td>$WZ+\text{jets}$</td>
<td>51.82</td>
<td>NLO</td>
</tr>
<tr>
<td>$ZZ+\text{jets}$</td>
<td>17.72</td>
<td>NLO</td>
</tr>
<tr>
<td>Single $t$ [64]</td>
<td>83.10</td>
<td>N^2LO</td>
</tr>
<tr>
<td>$tb$</td>
<td>248.00</td>
<td>N^2LO</td>
</tr>
<tr>
<td>$tj$</td>
<td>12.35</td>
<td>N^2LO</td>
</tr>
<tr>
<td>$t\bar{t}$ [65]</td>
<td>988.57</td>
<td>N^3LO</td>
</tr>
<tr>
<td>$ttV$ [66]</td>
<td>1.05</td>
<td>NLO + NNLL</td>
</tr>
<tr>
<td>$ttW$</td>
<td>0.65</td>
<td>NLO + NNLL</td>
</tr>
</tbody>
</table>
TABLE III. The sets of cuts applied for the two signatures. The cuts are motivated by the CMS analysis with 137 fb⁻¹ [40].

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>2τ (t₅t₅t₅ + t₅t₅τₑ)</td>
<td>Leptons and jets basic selection</td>
</tr>
<tr>
<td>1τ (t₅t₅ + MET)</td>
<td>Number of leptons/jets, Mass/energy cuts</td>
</tr>
</tbody>
</table>

where the number of signal and background events surviving the final selection cuts (as listed in Table III) are denoted by Nₛ and Nᵦ, respectively. We show the expected significance as functions of LQ masses in Fig. 2. Figs. 2(a)–2(d) are for the $t₅t₅t₅ + t₅t₅τₑ$ channel and Figs. 2(e)–2(h) are for the $t₅t₅ + MET$ channel. These plots are obtained for the 14 TeV LHC with 3 ab⁻¹ of integrated luminosity. We also display the masses corresponding to $\mathcal{Z} = 2, 3, \text{and } 5$ in Tables IV, V, and VI. As discussed earlier, the choice of $\kappa$ affects the pair and some single productions of vLQs and hence can significantly change the reach. In the figures we only present the results for $\kappa = 1$ but in Tables V and VI, we show the mass limits for another choice of, namely, $\kappa = 0$. We have used $\lambda, \Lambda = 1$ to estimate the significance for the combined signal (i.e., the pair and single production events together).

The figures are largely self-explanatory, and so, we only discuss some important points here. In Fig. 2(a), we show the expected significance $\mathcal{Z}$ in the $t₅t₅t₅ + t₅t₅τₑ$ channel as functions of $M_{LQ}$ (i.e., the mass of the charge-1/3 sLQ) in different scenarios. Notice that one can probe higher $M_{LQ}$ values in the combined LCSS scenario than the combined LCOS scenario even though the decays $\phi₁ \rightarrow ττ$ and $\phi₁ \rightarrow bν$ share 50% BR each in both scenarios (see Table I). The difference appears from the (constructive/destructive) interference in single production diagrams in these two scenarios that we discussed earlier. The destructive interference in the LCOS scenario takes the combined curve closer to the ‘pair (BR = 1)’ curve. In the RC scenario, $\phi₁$ has 100% BR in the $φ₁ \rightarrow ττ$ decay mode. However, the single production cross section is small in this scenario (see Fig. 1(a)). As a result, the combined reach is only marginally improved than the pair production only case with BR = 1.

For the charge-1/3 vLQ $χ₁$, there are no LCOS and LCSS scenarios. In Fig. 2(b), the LC50 and RC50 curves represent the cases where the BR of $χ₁ \rightarrow ττ$ mode is 50%. Such a scenario is possible if there are other decay modes of $χ₁$ that play no role in our analysis beyond modifying the BR. Hence, we show these plots to estimate how the significance would vary with the BR. The signal significance for $φ₅ (χ₅)$ is similar to that of $φ₁ (χ₁)$ with 100% BR to $ττ$ mode.

In the CMS analysis [40], the $tbτν + τν$ channel is considered for the charge-2/3 vLQ $Q₂$ or the charge-1/3 sLQ ($φ₁$). However, the same final state can also arise from a charge-2/3 sLQ ($φ₂$) or a charge-1/3 vLQ ($χ₂$) as well. We obtain the HL-LHC reach in the $t₅t₅ + MET$ channel for both charge-1/3 and 2/3 sLQs and vLQs ($φ₁₂$ and $χ₁₂$).
FIG. 2. Expected significance $Z$ for observing the $\phi_i/\chi_i$ signals over the SM backgrounds. They are plotted as functions of their masses for 3 ab$^{-1}$ of integrated luminosity at the 14 TeV HL-LHC for different coupling scenarios. The 'comb' implies the combined pair and single production processes. We have shown the pair production significance with BR = 50% and BR = 100%. We have considered $\lambda, \Lambda = 1$ when computing the signals.
In this channel, the interference effect for $\phi_1$ is not visible, and therefore, we club the LCOS and LCSS scenarios as LC50 in Fig. 2(e). The LC50 significance curve lies somewhere between the LCSS and LCOS curves in the $2\tau$ channel. For $\chi_1$, the RLCOS and RLCSS (clubbed as RLC50) lead to similar significance as BR = 50\% cases in the $2\tau$ channel. The reach for $\phi_2$ in the RLC50 scenario and for $\chi_2$ the LC50 scenario are shown in Figs. 2(g) and 2(h), respectively.

We can parametrize the coupling dependence of combined signal cross section (in any channel) as follows:

$$\sigma_{\text{combined}} \approx \sigma_{\text{pair}}(M_\phi) + \lambda^2 \sigma_{\text{single}}(\lambda = 1, M_\phi).$$  (30)

The above equation also holds for vLQs with a fixed $\kappa$. Using those relations, we can obtain the HL-LHC 5\sigma discovery reach and 2\sigma exclusion limits in the coupling-mass planes. We present these plots in Figs. 3 and 4 for the
FIG. 3. The $5\sigma$ ($2\sigma$) discovery (exclusion) reaches in the mass-coupling plane. These plots describe the lowest values of couplings needed to observe LQ signals with $5\sigma$ and $2\sigma$ significance as functions of masses with $3 \text{ ab}^{-1}$ of integrated luminosity. The pair-production-only regions for 50% and 100% BRs in the $\chi/\phi \rightarrow t\tau$ decay mode are shown in green. The pair production processes are insensitive to couplings.
FIG. 4. Same as Fig. 3 for the $t_b \tau_h +$ MET channel. The pair-production-only regions for 50% BRs in the $\chi/ϕ \to t\tau$ and $b\nu$ decay mode are shown in blue.
1τ and 2τ signatures. These plots show the lowest values of LQ-\( q - \ell \) couplings needed to observe the LQ signatures as functions of LQ masses with 5σ confidence level for discovery. For the exclusion plots, all points above the curves can be excluded with 95% confidence level at the HL-LHC. These plots are significant from the perspective of the \( B \)-meson anomalies. For example, the \( \lambda \) in the LCOS curves in Figs. 3(a) and 3(b) represent the \( x_{13}^{13} \) coupling of \( S_1 \) or the \( \Lambda \) in the LC50 curves in Figs. 4(g) and 4(h) is the \( x_{33}^{33} \) coupling of \( U_1 \). Hence, these plots show how far the LHC can probe the couplings required to explain the anomalies.

In our analysis so far, we have ignored the possible systematic errors. In practice, experiments have to account for them. However, as we argue in Appendix, our results can be taken as a reasonable conservative estimate of the projected experimental limits. Because of more precise (e.g., data-driven) estimation of the background processes and the use of profile likelihood ratios to estimate the CLs instead of approximate formulas of signal significance, we expect the actual experimental limits to be better than our estimations. The use of advanced analysis techniques (like machine learning) can improve the results further.

V. SUMMARY AND CONCLUSIONS

In this paper, we have considered all scalar and vector LQs that can decay to \( \tau \tau \) (applicable for charge-1/3 and charge-5/3 LQs) or \( \tau \nu \) pair (applicable for charge-2/3 LQs). Our choice of the decay modes involving a top quark is motivated from the fact that a top quark from a heavy LQ decay would be boosted. The features of the boosted top can then be used to search for the LQs. In our previous papers [11,19], we obtained the HL-LHC prospects for LQs that can decay to a boosted top quark and a light lepton (i.e., \( e \) or \( \mu \)). Here, we treat the case of the \( \tau \)-lepton separately since the hadronic and leptonic decays of \( \tau \)-leptons complicate the analysis. Therefore, the search strategies adopted here are completely different than those applied for the light leptons. Moreover, in this paper, we extend our study by including an asymmetric channel for charge-1/3 and charge-2/3 LQs.

We considered two different signatures: (a) at least one hadronically decaying boosted top quark and two high-\( p_T \) tau leptons. This signature is denoted as \( t_h \tau_h + t_h \tau_h \), according to the \( \tau \) decay modes involved, and (b) at least one hadronically decaying boosted top quark and one hadronically decaying high-\( p_T \) tau lepton with some missing energy \((t_h \tau_h + \text{MET})\). In both signatures, we used both standard jets (AK4-jets) and fatjets (AK8-fatjets) in the event selection criteria. For example, presence of \( \tau_h(s) \) and a \( b \)-jet in the signals are ensured by tagging the AK4-jets, and the AK8-fatjets are used to form the hadronic top quark. If an AK4-jet is tagged as \( \tau_h \), we make sure that it does not come from the AK8-fatjet identified as \( t_h \). But the \( b \)-tagged jet may or may not be a part of \( t_h \). The presence of \( b \)-jet helps us to tame the \( V+\) jets backgrounds.

An important point in our analysis is that the two signatures are inclusive in nature. Hence, they can keep both pair and single production events with substantial fractions, i.e., our selection criteria are not optimized for pair or single productions separately, rather they are optimized for their combination. Such a combined-signal strategy can significantly improve the exclusion limits or discovery reach for the LQs. Because of the messy nature of final states due to the presence of \( \tau \), the reach for the third-generation LQs is lower than was obtained in Refs. [11,19] with the light leptons. However, the overall HL-LHC reach goes beyond 1.5 TeV for sLQs and 2 TeV for vLQs. We have also mimicked the CMS selection criteria for the boosted or “fully-merged” top quark category from Ref. [40] to compare their analysis with ours. With our analysis, the future LHC prospects for the combined signal stand slightly better than those obtained with the CMS cuts. Therefore, our proposed analysis can act as a good search strategy for the specific subset of LQs we considered.

The single production processes that enter in the combined signal are three-body single productions whereas the commonly considered ones are two-body single productions. We found an interesting interference between single production diagrams for the charge-1/3 sLQ in the \( pp \rightarrow \phi_1 \tau j \) channel. We introduced two scenarios (LCSS and LCOS) that can be realized in the singlet \((S_1)\) and triplet \((S_3)\) sLQ models to demonstrate this effect. In the LCSS scenario, due to constructive interference, the discovery reach is appreciably higher than that in the LCOS scenario where the interference is destructive in nature. This interference effect is not observed in other LQs.

ACKNOWLEDGMENTS

A. B. and S. M. acknowledge financial support from the Science and Engineering Research Board (SERB), India under Grant No. ECR/2017/000517. T. M. is supported by the SERB, India under research Grant No. CRG/2018/004889 and an intramural grant from Indian Institute of Science Education and Research Thiruvananthapuram.

APPENDIX: EFFECT OF SYSTEMATIC ERRORS

Here, we show the effect of systematic uncertainty on our results. For illustration, we consider two benchmark choices of 5% and 10% systematics on our background estimations. In the presence of total systematic error \( \sigma_B \), Eq. (29) generalizes to

\[
Z = \sqrt{2} \left( N_S + N_B \right) \ln \left[ \frac{N_S + N_B}{N_B + (N_S + N_B) \sigma_B^2} \right] - \frac{N_B}{\sigma_B^2} \ln \left[ 1 + \frac{\sigma_B^2 N_S}{N_B(N_B + \sigma_B^2)} \right]^{1/2}. \tag{A1} \]
whose approximated form is perhaps more familiar,

\[ Z \approx \frac{N_s}{\sqrt{N_B + \sigma_B^2}}. \]  

(A2)

A naive use of the above formulas would adversely affect our estimation of the mass limits, especially in the cases of sLQs where the signal cross sections are smaller than the vLQs [67]. However, so far, we have not utilized the signal and background distributions beyond the total number of events. Actually, it is possible to obtain comparable mass limits even in the presence of 5%–10% systematic errors from binned data.

For an illustration, we show the limits estimated from a binned-data analysis for a few benchmark scenarios of \( \phi_1 \) in Table VII. The numbers, obtained for both 5% and 10% systematic errors, are to be compared with those in Table IV. The bins are shown in Fig. 5. For the vLQs, we find similar to slightly improved numbers (∼5%–10%) than those in Tables V and VI.

### TABLE VII. The mass limits for some sample scenarios, with 5% and 10% systematic uncertainties.

| Comb \( \bar{Z} \) | Limit on \( M_{\phi_1} \) (TeV) |
|------------------|------------------------|------------------------|------------------------|
|                  | LCSS-Combined | Pair (BR = 0.5) | RC-Combined | Pair (BR = 1) | LCSS-Combined | Pair (BR = 0.5) |
| 5                | 1.06          | 0.94            | 0.84        | 0.75          | 1.42          | 1.37          | 1.40          | 1.36          | 1.19          | 1.11          | 1.15          | 1.05          |
| 3                | 1.34          | 1.26            | 1.12        | 1.04          | 1.56          | 1.51          | 1.54          | 1.49          | 1.37          | 1.30          | 1.32          | 1.24          |
| 2                | 1.53          | 1.47            | 1.27        | 1.22          | 1.67          | 1.65          | 1.66          | 1.63          | 1.47          | 1.43          | 1.44          | 1.38          |

FIG. 5. The bins show the number of signal and background events in the \( t_h \tau_h \tau_h + t_h \tau_h \tau_f \) channel [(a) and (b)] and the \( t_h \tau_h + \text{MET} \) channel [(c) and (d)]. The events are obtained applying all the cuts in Table III, except the ones on the variables used for binning, i.e., \( N(\tau_h) \) and \( M_T \) in the \( t_h \tau_h \tau_h + t_h \tau_h \tau_f \) channel and \( N(\tau_h) \) and MET in the \( t_h \tau_h + \text{MET} \) channel.
To estimate the signal significance from binned data, we have employed the Liptak-Stouffer (weighted) $Z$-score method where the metascore or the combined significance is given as

$$ Z = \frac{\sum_{i=1}^{N} w_i Z_i}{\sqrt{\sum_{i=1}^{N} w_i^2}}. \quad (A3) $$

Here, $Z_i$ denotes the signal significance in the $i$th bin ($i \in \{1, 2, 3, \ldots, N\}$) computed from Eq. (A1) and $w_i$ is the corresponding weight, which is taken to be equal to the variance in that bin. We have set the $w_i^2$ equal to the square of the total errors, i.e., $w_i^2 = (\text{statistical error})^2 + (\text{systematic error})^2 = N^2 + (\sigma_b^2)^2$.

---


