Search for vectorlike light-flavor quark partners in proton-proton collisions at $\sqrt{s} = 8$ TeV

CMS Collaboration; Canelli, Florencia; Kilminster, Benjamin; Aarestad, Thea; Brzhechko, Danyyl; Caminada, Lea; De Cosa, Anna Paoloa; Del Burgo, Riccardo; Donato, Silvio; Galloni, Camilla; Hreus, Tomas; Leontsinis, Stefanos; Mikuni, Vinicius Massami; Neutelings, Izaak; Rauco, Giorgia; Robmann, Peter; Salerno, Daniel; Schweiger, Korbinian; Seitz, Claudia; Takahashi, Yuta; Wertz, Sebastien; Zucchetta, Alberto; et al

Abstract: A search is presented for heavy vectorlike quarks (VLQs) that couple only to light quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV at the LHC. The data were collected by the CMS experiment during 2012 and correspond to an integrated luminosity of 19.7 fb−1. Both single and pair production of VLQs are considered. The single-production search is performed for down-type VLQs (electric charge of magnitude $1/3$), while the pair-production search is sensitive to up-type (charge of magnitude $2/3$) and down-type VLQs. Final states with at least one muon or one electron are considered. No significant excess over standard model expectations is observed, and lower limits on the mass of VLQs are derived. The lower limits range from 400 to 1800 GeV, depending on the single-production cross section and the VLQ branching fractions $B(W), B(Z), B(H)$. When considering pair production alone, VLQs with masses below 845 GeV are excluded for $B(W)=1.0$, and below 685 GeV for $B(W)=0.5, B(Z)=B(H)=0.25$. The results are more stringent than those previously obtained for single and pair production of VLQs coupled to light quarks.

DOI: https://doi.org/10.1103/PhysRevD.97.072008

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-160238
Journal Article
Published Version

The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:
CMS Collaboration; Canelli, Florencia; Kilminster, Benjamin; Aarestad, Thea; Brzhechko, Danyyl; Caminada, Lea; De Cosa, Anna Paoloa; Del Burgo, Riccardo; Donato, Silvio; Galloni, Camilla; Hreus, Tomas; Leontsinis, Stefanos; Mikuni, Vinicius Massami; Neutelings, Izaak; Rauco, Giorgia; Robmann, Peter; Salerno, Daniel; Schweiger, Korbinian; Seitz, Claudia; Takahashi, Yuta; Wertz, Sebastien; Zucchetta, Alberto; et al (2018). Search for vectorlike light-flavor quark partners in proton-proton collisions at $\sqrt{s} = 8$ TeV. Physical Review. D, Particles, fields, gravitation and cosmology, D97:072008.
Search for vectorlike light-flavor quark partners in proton-proton collisions at $\sqrt{s} = 8$ TeV

A. M. Sirunyan et al.
(CMS Collaboration)

(Received 8 August 2017; published 11 April 2018)

A search is presented for heavy vectorlike quarks (VLQs) that couple only to light quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV at the LHC. The data were collected by the CMS experiment during 2012 and correspond to an integrated luminosity of 19.7 fb$^{-1}$. Both single and pair production of VLQs are considered. The single-production search is performed for down-type VLQs (electric charge of magnitude $1/3$), while the pair-production search is sensitive to up-type (charge of magnitude $2/3$) and down-type VLQs. Final states with at least one muon or one electron are considered. No significant excess over standard model expectations is observed, and lower limits on the mass of VLQs are derived. The lower limits range from 400 to 1800 GeV, depending on the single-production cross section and the VLQ branching fractions for the different decay modes depend on the multiplet in which the VLQ resides. In most models, the neutral-current branching fractions $B(Q \rightarrow Zq)$ and $B(Q \rightarrow Hq)$ are roughly the same size, and the charged-current branching fraction $B(Q \rightarrow Wq)$ can vary between 0 and 1. Other decay modes are assumed to be negligible, so the sum of the three branching fractions is one.

I. INTRODUCTION

Vectorlike quarks (VLQs) are hypothetical spin-1/2 fermions, whose left- and right-handed chiral components transform in the same way under the standard model (SM) symmetries, and hence have vector couplings to gauge bosons. Such VLQs appear in a number of models that extend the SM to address open questions in particle physics. These models include: beautiful mirrors [1], little-Higgs models [2–4], composite-Higgs models [5], theories invoking extra dimensions [6], grand unified theories [7], and models providing insights into the SM flavor structure [8].

Owing to the possible role of third-generation quarks in the solution of problems in electroweak symmetry breaking, the VLQs in many of the aforementioned models mix predominantly with third generation quarks. In addition, indirect experimental constraints on the quark couplings of the lighter generations from precision electroweak measurements are typically stronger than those on third-generation couplings [9]. However, the coupling corrections from several different VLQs may cancel, which can significantly relax constraints on the mixing of VLQs with the first and second generations. In this paper, we consider the pair production of heavy VLQs, denoted by $Q$, with electric charge of magnitude $1/3$ or $2/3$, that are partners of the first-generation SM quarks. We also consider the electroweak single production of vectorlike down-type quarks with electric charge of magnitude $1/3$, which we denote by $D$ in this context.

Figure 1 shows examples of Feynman diagrams for the leading-order electroweak single production and strong pair production of VLQs coupled to first-generation quarks. In order to describe the production processes, new couplings of the VLQs to light-flavor quarks via $W$, $Z$, and Higgs bosons are introduced, whereas no new coupling to gluons is considered. Assuming a short enough lifetime, the new quarks do not hadronize before decaying to $Wq$, $Zq$, or $Hq$, where $q$ indicates a SM quark. The branching fractions for the different decay modes depend on the multiplet in which the VLQ resides [10]. In most models, the neutral-current branching fractions $B(Q \rightarrow Zq)$ and $B(Q \rightarrow Hq)$ are roughly the same size, and the charged-current branching fraction $B(Q \rightarrow Wq)$ can vary between 0 and 1. Other decay modes are assumed to be negligible, so the sum of the three branching fractions is one.

*Full author list given at the end of the article.

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 SCOAP$^3$. DOI: 10.1103/PhysRevD.97.072008
of about 8 TeV using the CMS detector at the CERN LHC. The signal model parameter space.

Previous searches for single and pair production of such VLQs have been performed by the ATLAS experiment at √s = 7 and 8 TeV [12,13]. These searches exclude singly produced VLQs with masses below 900 (760 GeV), with Qq → Wqq (Qq → Zqq), and pair-produced VLQs with masses below 690 GeV, with B(Q → Wq) = 1, at 95% confidence level (C.L.).

For high VLQ masses above 1 TeV, the kinematically favored single-production mode may be the dominant production mode, since the pair-production cross section via the strong interaction drops rapidly as a function of the VLQ mass. Nevertheless, since the single-production cross section depends on unknown model-dependent parameters, the pair-production mode may be dominant for sub-TeV VLQ masses. Furthermore, the VLQs may decay to W, Z, and Higgs bosons with unknown decay branching fractions. These considerations motivate searches for VLQs over a wide mass range with search methods optimized for both singly and pair produced VLQs, decaying in a variety of modes.

In this paper we report results of a search for VLQs in proton-proton collisions at a center-of-mass energy of 8 TeV using the CMS detector at the CERN LHC. The data set analyzed corresponds to an integrated luminosity of about 19.7 fb⁻¹. Events with one or more isolated leptons are used for the search. The signal channels considered are listed in Table I. The processes Dq → Hqq and Q̅Q̅ → HqHq have not been considered because of the low efficiency for selecting isolated leptons in such decay modes. The search for singly produced VLQs is performed only for vector-like down-type quarks. The search for pair-produced VLQs is also applicable to up-type quarks, as their decay products are experimentally indistinguishable from those of down-type VLQs.

This is the first search for VLQs coupled to light-flavor quarks that simultaneously considers the single and pair production modes, in a scan over the branching fractions of the VLQs to W, Z, and Higgs bosons. Furthermore, for the first time in these topologies, kinematic fits using boosted jet substructure techniques in single-lepton events are applied to improve the VLQ mass reconstruction, and events with at least two leptons are analyzed to retain sensitivity to VLQs that have a high probability of decaying to a Z boson.

II. ANALYSIS STRATEGY

In this analysis, the search for singly produced vector-like D quarks involves the reconstruction of a VLQ resonance in final states with exactly one or two leptons and two or three jets. In the search for pair produced VLQs, in the final state with one lepton, missing transverse momentum, and four jets, a kinematic fit is performed to reconstruct the VLQ mass. Final states with two, three or four leptons and at least two jets are also considered, using reconstructed observables sensitive to the VLQ mass. The results of all channels, which are mutually exclusive, are combined in the calculation of the limits on the VLQ masses and the production cross sections.

The searches are performed without assuming that the hypothetical quark belongs to a particular SU(2) multiplet structure. Therefore the analysis is not optimized for a combined search for all quarks in a given multiplet. As such, the exclusion limits presented in this analysis are expected to be more conservative than those that would be obtained in a dedicated model-dependent search combining the signal from all quarks within a multiplet. On the other hand, the approach used here allows a more model-independent interpretation.

| Production          | Channel                  |
|---------------------|--------------------------|
| Single (electroweak)| Dq → Wqq                |
|                     | Dq → Zqq                |
| Pair (strong)       | Q̅Q̅ → WqWq              |
|                     | Q̅Q̅ → WqZq              |
|                     | Q̅Q̅ → WqHq              |
|                     | Q̅Q̅ → ZqZq              |
|                     | Q̅Q̅ → ZqHq              |

FIG. 1. Vectorlike quarks (denoted Q) can be produced in proton-proton collisions either singly through electroweak interactions (the channel mode (upper) is shown as an example), or in pairs via the strong interaction (lower). For single production we consider in the present work only vectorlike quarks with electric charge of magnitude 1/3 (denoted D).

B(Q → Zq) = B(Q → Hq) = 0.25, implied by the equivalence theorem [11], it is chosen as a benchmark point in the signal model parameter space.

Table I. Decay channels of vector-like quarks considered in the analysis.
A. Search for single production

We consider the electroweak charged-current and neutral-current modes of the single production of vector-like $D$ quarks. The interaction Lagrangian density for the vector-like $D$ quarks contains three unknown parameters, corresponding to the couplings to the three bosons, $\kappa_W$, $\kappa_Z$, and $\kappa_H$ [9,14]:

$$\mathcal{L}_{\text{interaction},D} = \frac{g_W}{\sqrt{2}} \kappa_W W^\mu R \bar{d}_R \gamma^\mu d_R + \frac{g_W}{2 \cos \theta_W} \kappa_Z Z_R \bar{d}_R \gamma^\mu d_R - \frac{m_Q}{v} \kappa_H H \bar{d}_R d_L + \text{h.c.}$$  \hspace{1cm} (1)

Here $v \approx 246$ GeV is the Higgs field vacuum expectation value, $m_Q$ is the VLQ mass, $\theta_W$ is the weak mixing angle and $g_W$ is the coupling strength of the weak interaction. In Eq. (1) the terms for just one chirality are given (the R and L field indices refer to right- and left-handed helicities, respectively), but there are equivalent terms for the other helicities.

The coupling parameters, $\kappa$, are model dependent, and originate from the mixing between SM quarks and VLQs. These couplings can be reparametrized as $\kappa = v \tilde{\kappa}/\sqrt{2} m_Q$, with the new parameter $\tilde{\kappa}$ being naturally of order unity in a weakly coupled theory [9].

In the particular scenario where the VLQ couples only to the first-generation quarks, it can be shown [14] that the neutral-current coupling strength parameter, $\tilde{\kappa}_Z$, may be expressed approximately through the charged-current coupling strength parameter, $\tilde{\kappa}_W$, and the branching fractions of the decays of the VLQ to $W$ and $Z$ bosons, $\mathcal{B}_W = \mathcal{B}(Q \rightarrow Wq)$ and $\mathcal{B}_Z = \mathcal{B}(Q \rightarrow Zq)$, via:

$$\tilde{\kappa}_Z \approx \sqrt{\frac{2 \mathcal{B}_Z}{\mathcal{B}_W}} \tilde{\kappa}_W, \hspace{1cm} (2)$$

if $\mathcal{B}_W \neq 0$. It is therefore sufficient to determine limits on the cross section and mass as a function of the three free parameters, $\tilde{\kappa}_W, \mathcal{B}_W$ and $\mathcal{B}_Z$, producing cross section and mass limits that then depend only on these parameters. If $\mathcal{B}_W$ approaches 0, with $\tilde{\kappa}_W$ fixed to a nonzero value, Eq. (2) implies that $\tilde{\kappa}_Z$ diverges, and when $\mathcal{B}_W$ is exactly zero, Eq. (2) is no longer applicable. Results for an alternative single-production coupling parametrization that does not exhibit divergent behavior throughout the parameter scan are available in the Supplemental Material [15].

The expected signal topologies are listed in the upper two rows of Table I. It should be noted that singly produced VLQs are produced in association with a forward-going first-generation quark. As will be explained in Sec. VI A, we define two event categories corresponding to these two topologies, based on whether one or two isolated leptons are present in the final state. In these event categories we employ the reconstructed mass of the D quark decaying into a $W$ or $Z$ boson and a quark to search for a signal.

B. Search for pair production

In the search for strongly produced VLQ pairs, $Q\bar{Q}$, several event categories are defined that are optimized for the decay modes of pair produced VLQs listed in Table I. Signal events do not often contain $b$ jets, except in the cases where a Higgs boson is produced.

The single-lepton event categories are optimized for the following decay modes of VLQ pairs:

$$Q\bar{Q} \rightarrow WqWq \rightarrow \ell vq\ell q'q_h; \hspace{1cm} (3)$$

$$Q\bar{Q} \rightarrow WqZq \rightarrow \ell vq\ell qq_h; \hspace{1cm} (4)$$

$$Q\bar{Q} \rightarrow WqHq \rightarrow \ell vq bbq_h. \hspace{1cm} (5)$$

In these events the $W$ boson decays leptonically into a muon or an electron plus a neutrino and the other boson ($W$, $Z$, or $H$) decays into a pair of quarks. These events are classified as either $\mu + \pi$ or $e + \pi$ jets events. A light quark, $q_r$, is produced in association with the leptonically decaying $W$ boson, and $q_h$ is the equivalent for the hadronically decaying boson. We perform a constrained kinematic fit for each event to reconstruct the mass of the VLQ. The full kinematic distributions of the final state are reconstructed, and the mass of the Q quark, $m_{fit}$, is obtained, as detailed in Sec. VI B 1. In addition, the $S_T$ variable is defined as the scalar sum of the transverse momenta $p_T^\ell$ of the charged lepton, the transverse momenta $p_T^{jet}$ of the jets, and the $p_T^{miss}$ value:

$$S_T = \sum p_T^\ell + \sum p_T^{jet} + p_T^{miss}. \hspace{1cm} (6)$$

The variable $p_T^{miss}$, referred to as the missing transverse momentum, is defined as the magnitude of the missing transverse momentum vector, which is the projection on the plane perpendicular to the beams of the negative vector sum of the momenta of all reconstructed particles in the event. The $S_T$ variable, calculated after the fit, is used to define a phase space region where the signal-to-background ratio is enhanced.

In the dilepton event categories we employ two variables to search for a VLQ signal, as will be discussed in Sec. VI B 2. The first variable is the reconstructed mass of the $Q$ quark decaying into a $Z$ boson and a quark, the second one is the $S_T$ variable defined in Eq. (6).

In the multilepton event category, in this analysis defined as containing three or four leptons, the number of expected events is small. Here, rather than using a kinematic variable to identify a possible signal, events are counted after imposing the selection criteria.
III. THE CMS DETECTOR

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid.

The CMS detector is nearly hermetic, allowing momentum balance measurements to be made in the plane transverse to the beam direction. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [16].

IV. EVENT SAMPLES

The data used for this analysis were recorded during the 2012 data taking period, at a proton-proton center-of-mass energy of 8 TeV. The total integrated luminosity of the data sample is 19.7 fb\(^{-1}\) (19.6 fb\(^{-1}\) in the categories optimized for single VLQ production and those requiring at least two leptons optimized for pair production). The trigger used to select the muon data sample is based on the presence of at least one muon with a pseudorapidity satisfying \(|\eta| < 2.1\) and transverse momentum \(p_T > 40\) GeV (in the single-lepton pair-production category), or at least one isolated muon with \(p_T > 24\) GeV (in all other categories). For the electron data sample, events must pass a trigger requiring the presence of one isolated electron with \(p_T > 27\) GeV.

Simulated samples are used to estimate signal efficiencies and background contributions. The processes \(pp \rightarrow Dq\) and \(pp \rightarrow Q\) are simulated using the MadGraph 5.1.5.3 event generator [17] with CTEQ6L1 parton distribution functions (PDFs) [18], with a decay width of 1\% of the VLQ mass and without extra partons, and then passed to Pythia 6.424 [19] with the ZZ* tune [20,21] for hadronization. The following SM background processes are simulated: \(t\bar{t}\) production (including \(t\bar{t}\) production in association with a vector boson and one or more jets, denoted \(t\bar{t}Z + \text{jets}\) and \(t\bar{t}W + \text{jets}\)); single top quark production via the \(tW\), \(s\)-channel, and \(t\)-channel processes; single-boson and diboson production (\(W + \text{jets}\), \(Z + \text{jets}\), \(WW\), \(WZ\), and \(ZZ\)), triboson processes (\(WWW\), \(WWZ\), \(WZZ\), \(ZZZ\)), and multijet events.

Samples of the SM background processes, \(t\bar{t}\) + jets, and single top quark production via \(tW\), \(s\)-, and \(t\)-channels, are simulated using the Powheg 1.0 [22–24] event generator. The diboson processes (\(WW\), \(WZ\), and \(ZZ\)) and multijet events are generated using the Pythia event generator. The \(t\bar{t}Z + \text{jets}\), \(t\bar{t}W + \text{jets}\), \(W + \text{jets}\), \(Z + \text{jets}\) and triboson samples are simulated using the MadGraph event generator. The Pythia generator is used for parton shower development and hadronization, for all simulated background processes. The CTEQ6M PDFs are used for Powheg, while for the other generators the CTEQ6L1 PDFs are used.

The VLQ single-production cross sections are calculated at leading order (LO) with the MadGraph generator, and the pair-production cross sections, at next-to-next-to-LO (NNLO) [25]. The production cross sections for the background processes are taken from the corresponding cross section measurements made by the CMS experiment [26–29]: \(t\bar{t} + \text{jets}\), single top quark production in the \(tW\) mode, \(WW\), \(WZ\), and \(ZZ\); and are in agreement with theoretical calculations at next-to-LO (NLO) or NNLO accuracy. The cross section for multijet processes is calculated at leading order by Pythia. The cross sections of the remaining processes mentioned above are calculated either at NLO or at NNLO.

All simulated events are processed through the CMS detector simulation based on Geant4 [30]. To simulate the effect of additional proton-proton collisions within the same or adjacent bunch crossings (pileup), additional inelastic events are generated using Pythia and superimposed on the hard-scattering events. The Monte Carlo (MC) simulated events are weighted to reproduce the distribution of the number of pileup interactions observed in data, with an average of 21 reconstructed collisions per beam crossing.

V. EVENT RECONSTRUCTION

The event reconstruction uses the particle flow (PF) algorithm [31] which reconstructs and identifies each individual particle with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, muon, electron, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy. Muons are identified by tracks or hits in the muon system that are associated with the extrapolated trajectories of charged particles reconstructed in the inner tracker and have small energy deposits in the traversed calorimeter cells. Electrons are identified as charged-particle tracks that are associated with potentially several ECAL clusters that result from the showering of the primary particles and from secondary bremsstrahlung photons produced in the tracker material [32]. Charged hadrons are identified as charged-particle tracks associated with energy deposits in the HCAL, and identified as neither electrons nor muons. Finally, neutral hadrons are identified as HCAL energy clusters not linked to any charged hadron trajectory, or as ECAL and HCAL energy excesses with respect to the expected charged-hadron energy deposit.

The energy of each muon is obtained from the corresponding track momentum. The energy of each electron is determined from a combination of the track momentum at the interaction vertex, the corresponding ECAL cluster energy, and the energy sum of all bremsstrahlung photons...
attached to the track. The energy of each charged hadron is determined from a combination of the track momentum and the corresponding ECAL and HCAL energies, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of each neutral hadron is obtained from the corresponding corrected ECAL and HCAL energies.

Particles found using the PF algorithm are clustered into jets using the direction of each particle at the interaction vertex. Charged hadrons that are associated with pileup vertices are excluded, using a method referred to as charged-hadron subtraction. Particles that are identified as charged leptons, isolated according to criteria discussed later, are removed from the jet clustering procedure. In the analysis, two types of jets are used: jets reconstructed with the infrared- and collinear-safe anti-$k_T$ algorithm [33] operated with a distance parameter $R = 0.5$ (AK5 jets) and jets reconstructed with the Cambridge–Aachen algorithm [34] using a distance parameter $R = 0.8$ (CA8 jets), as implemented in FASTJET version 3.0.1 [35,36]. An event-by-event jet-area-based correction [37–39] is applied to remove, on a statistical basis, pileup contributions that have not already been removed by the charged-hadron subtraction procedure.

The momentum of each jet is determined from the vector sum of all particle momenta in the jet, and is found from simulation to be within 5% to 10% of the true momentum for all values of $p_T$ and over the whole detector acceptance. Jet energy corrections varying with $p_T$ and $\eta$ are applied to each jet to account for the combined response function of the calorimeters. They are derived from simulation, and are confirmed with in situ measurements of the energy balance of dijet and photon + jet events [40]. The jet energy resolution amounts typically to 15%–20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV.

As the mass of the heavy VLQ increases, the Lorentz boosts of the decay products also increase. The quark pairs from the hadronic decays of $W$, $Z$, or Higgs bosons become increasingly collimated and eventually the resulting hadronic showers cannot be resolved as separate jets. The CA8 jets are used to identify these merged hadronic boson decays and a jet pruning algorithm, which removes soft/wide-angle radiation [41–43] is then applied to resolve the merged subjects.

Charged leptons originating from decays of heavy VLQs are expected to be isolated from nearby jets. Therefore, a relative isolation ($I_{\text{rel}}$) criterion is used to suppress backgrounds from non-prompt leptons or hadrons misidentified as leptons inside jets. Relative isolation is calculated as the sum of the $p_T$ of the charged hadrons, neutral hadrons, and photons in a cone of $\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}$ around the lepton, with the lepton track itself removed from the sum, divided by the lepton $p_T$. Here $\Delta \phi$ and $\Delta \eta$ are the azimuthal angle and pseudorapidity differences with respect to the lepton direction. In the calculation of $I_{\text{rel}}$ using PF reconstruction, the isolation cone size is taken to be $\Delta R = 0.4$ for muons and $\Delta R = 0.3$ for electrons. In the calculation of $I_{\text{rel}}$, pileup corrections are applied.

Charged leptons are categorized by the stringency of their selection criteria in two types, namely “tight” and “loose” leptons, as defined in Table II. In the analysis, events with at least one tight muon or electron are selected, while the loose lepton criteria are used to identify and exclude the presence of additional leptons in the event. Additional requirements for tight and loose leptons used in the single-lepton channel optimized for VLQ pair production are described in Sec. VI B 1.

To identify jets as originating from a $b$ quark (b-tagged jets), the combined secondary vertex (CSV) algorithm is used [44,45]. This algorithm combines variables that distinguish $b$ jets from non-$b$ jets, such as the track impact parameter significance and properties of the secondary vertex. The algorithm uses a likelihood ratio technique to compute a $b$ tagging discriminator. We use two operating points (with different thresholds applied to the $b$ tagging discriminator): medium and loose, which are designated as CSVM and CSVL, respectively [45]. The medium (loose) CSV discriminant operating point corresponds to a light-quark or gluon mistag rate of about 1% (10%) and a $b$ tagging efficiency of about 70% (84%). B-tagging is applied to AK5 jets and to subjets of CA8 jets.

Data-to-simulation $b$ tagging efficiency and mistag rate scale factors correct for the small differences between the efficiencies observed in data and in simulation. We use scale factors that depend on both jet $p_T$ and $\eta$ [45].

### VI. ANALYSIS

#### A. Search for single production

We use two collections of AK5 jets with $p_T > 30$ GeV. The first collection consists of all jets that satisfy $|\eta| < 2.4$; these jets are referred to as selected central jets. The second collection contains all jets that satisfy $2.4 < |\eta| < 5.0$; these jets are referred to as selected forward jets. In order to exploit the presence of first-generation quarks in the final state of VLQ processes, we require the presence of a number of selected central jets for which the $b$-tag CSV

| Table II. Initial selection requirements for tight and loose leptons. |
|-----------------|-----------------|
| Muons           | Loose           |
| $p_T > 20$ GeV  | $p_T > 10$ GeV  |
| $|\eta| < 2.1$    | $|\eta| < 2.5$   |
| $I_{\text{rel}} < 0.12$ | $I_{\text{rel}} < 0.2$ |
| Electrons       |                 |
| Tight           | Loose           |
| $p_T > 20$ GeV  | $p_T > 15$ GeV  |
| $|\eta| < 2.5$    | $|\eta| < 2.5$   |
| $I_{\text{rel}} < 0.1$ | $I_{\text{rel}} < 0.15$ |


TABLE III.

The event categories as optimized for the VLQ single production. The categories are based on the number of tight muons or electrons present in the event, along with additional criteria optimized for specific VLQ topologies. Events containing any additional loose leptons are excluded.

| Event category | Tight leptons ($\mu, e$) | Additional selection criteria |
|----------------|-------------------------|-------------------------------|
| $W^- qq$       | 1 with $p_T > 30$ GeV   | 1 or 2 selected central jets, all antitagged |
|                | Negative charge         | Leading $p_T > 200$ GeV       |
|                |                         | 1 selected forward jet        |
|                |                         | $p_T(W \rightarrow \ell\nu) > 150$ GeV |
|                |                         | $\Delta R(\ell, \nu) < 1.5$  |
|                |                         | $p^\text{miss}_T > 60$ GeV, $M_T > 40$ GeV |
| $Zqq$          | 2 opposite-sign same-flavor | 1 or 2 selected central jets, all antitagged |
|                | Leading $p_T > 30$ GeV  | Leading $p_T > 200$ GeV       |
|                | Subleading $p_T > 20$ GeV | 1 selected forward jet        |
|                |                         | $|m_{\ell\ell} - m_Z| < 7.5$ GeV |
|                |                         | $p_T(Z \rightarrow \ell\ell) > 150$ GeV |

discriminant lies below the CSVL threshold. These jets are referred to as “antitagged” jets. Events with one or two tight muons or electrons are selected. The leptons (jets) in each event are ordered by transverse momentum. The lepton (jet) with the largest $p_T$ is labeled as the leading lepton (jet) and the others are labelled as subleading leptons (jets). We define two event categories that are sensitive to the single production topologies presented in Table I, $W^- qq$ and $Zqq$. In order to enhance the signal sensitivity to the corresponding category, we require the lepton charge in the corresponding category, indicated as $W^- qq$, to be negative. For a $D$ mass of 1100 GeV, this choice approximately doubles the signal-to-background ratio. The production rate for $D$ quarks is higher than that for $\bar{D}$ quarks [9] because of the proton PDFs. The production of $W$ bosons in the SM is also charge asymmetric for the same reason, with more $W^+$ bosons produced than $W^-$ bosons. We therefore use only the $W^- qq$ category in this search, and do not consider the corresponding category with a positively charged lepton, $W^+ qq$, to search for a signal. The definition of the event categories used to search for single production of VLQs is summarized in Table III.

The leptonically decaying $W$ and $Z$ boson candidates are reconstructed and thresholds are imposed on their transverse momenta, $p_T(W)$ or $p_T(Z)$. A $W$ boson candidate is reconstructed as follows. The $z$ component of the neutrino momentum is obtained by imposing the $W$ boson mass constraint on the lepton-neutrino system, resulting in a quadratic equation in the neutrino $p_z$. If the solution is complex, the real part is taken as the $z$ component. If both solutions are real we take the one where the total reconstructed neutrino momentum has the largest difference in $\eta$ with respect to the leading central jet in the event. We require the separation between the lepton and the reconstructed neutrino to satisfy $\Delta R < 1.5$, because

TABLE IV.

Event yields in the muon and electron channels for the event categories optimized for the single production search. The $W^- qq$ event category is not used in the search, but is shown for comparison, in order to demonstrate the expected lepton charge asymmetry. For the separate background components the indicated uncertainties are statistical only, originating from the limited number of MC events, while for the total background yield the combined statistical and systematic uncertainty is given. The prediction for the signals is shown assuming branching fractions of $B_W = 0.5$ and $B_Z = B_H = 0.25$. The label “Other” designates the background originating from $\bar{t}tW, \bar{t}tZ$ and triboson processes.

| Channel                      | Muon | Electron | Muon | Electron | Muon | Electron |
|------------------------------|------|----------|------|----------|------|----------|
| $W^+ qq$                     |      |          |      |          |      |          |
| Estimated backgrounds        |      |          |      |          |      |          |
| $\bar{t} +$ jets            | 26 ± 2 | 23 ± 2 | 28 ± 3 | 24 ± 2 | <1 | <1 |
| $W +$ jets                   | 2069 ± 43 | 1906 ± 41 | 1191 ± 36 | 1082 ± 32 | <1 | <1 |
| $Z +$ jets                   | 17 ± 3 | 10 ± 3 | 22 ± 4 | 8.7 ± 1.9 | 541 ± 20 | 428 ± 18 |
| Single top quark             | 20 ± 3 | 20 ± 3 | 11 ± 2 | 12 ± 2 | <1 | <1 |
| VV                           | 28 ± 2 | 27 ± 2 | 31 ± 2 | 31 ± 2 | 9.9 ± 0.7 | 7.6 ± 0.6 |
| Multijet                     | 3.9 ± 0.9 | 8.5 ± 2.5 | 2.8 ± 0.8 | 5.7 ± 2.0 | <1 | <1 |
| Other                        | <1 | <1 | <1 | <1 | <1 | <1 |
| Total background             | 2170 ± 440 | 2000 ± 400 | 1290 ± 240 | 1160 ± 230 | 550 ± 110 | 436 ± 87 |
| Observed                     | 2082 | 1838 | 1112 | 1027 | 527 | 421 |
| Signal ($m_Q = 600$ GeV, $\hat{\kappa}_w = 0.1$) | 1.8 | 1.5 | 4.6 | 4.1 | 1.5 | 1.2 |
| Signal ($m_Q = 1100$ GeV, $\hat{\kappa}_w = 1$) | 8.9 | 6.7 | 44.4 | 43.6 | 12.1 | 11.4 |
these two particles, when produced in the decay of a boosted $W$ boson, are expected to be close to each other. A requirement on the transverse mass $M_T = \sqrt{2p_T^l p_T^{\text{miss}}(1 - \cos[\Delta \phi(l, p_T^{\text{miss}})])} > 40$ GeV is imposed to suppress the multijet background. A $Z$ boson candidate is reconstructed from two same-flavor opposite-sign dileptons, and requirements on the mass, $m_{ll}$, of the dilepton system are imposed, as described in Table III.

The event yields for the observed data as well as for the expected SM backgrounds are shown in Table IV for the muon channel and the electron channel. The respective uncertainties, and the highest bin contains the overflow.

In each of the event categories we reconstruct the mass of the VLQ candidate from the $W$ or $Z$ boson candidates and the leading central jet in the event. The reconstructed mass normalizations of the simulated $W$ and $Z$ boson production processes in association with either light-flavor jets or heavy-flavor jets are derived from data by fitting the CSVL b-tagged jet multiplicity distribution in control samples. A deficit of data events compared to simulation is observed in both the signal-depleted $W^+qq$ and the signal-enriched $W^-qq$ categories, motivating a dedicated background prediction in the $W^-qq$ category as described below.

FIG. 2. The reconstructed mass of the VLQ candidate in the $W^+qq$ event category (upper) and the $W^-qq$ event category (lower), in the muon channel (left) and the electron channel (right). The contributions of simulated events where the $W$ boson is produced in association with light-flavor (LF) jets and heavy-flavor (HF) jets are shown separately. The distribution for a heavy VLQ signal (indicated as $Dq$ representing a down-type VLQ produced in association with a SM quark) of mass 1100 GeV and $\bar{\kappa}_W = 1$ (for $B_W = 0.5$ and $B_Z = B_B = 0.25$) is scaled up by a factor of 15 for visibility. The enhanced $D$ quark signal contribution in the $W^-qq$ event category in comparison to the $W^+qq$ event category is clearly shown. The hatched bands represent the combined statistical and systematic uncertainties, and the highest bin contains the overflow.
can be used to efficiently discriminate between the SM background and the VLQ processes.

Figure 2 shows the reconstructed mass of the VLQ candidate for the $W^{-}qq$ category (upper) and the $W^{+}qq$ category (lower), comparing data to simulation. The distributions of the reconstructed VLQ candidate mass comparing data to the prediction derived from a control region in data are shown in Fig. 3 for the muon channel (left) and the electron channel (right), for the background prediction and the data. The solid bold (blue) line is the background distribution estimated from data, with a final shape correction that accounts for the difference between the $W +$ jets simulation in the control region and the $W^{-}qq$ signal region. The dashed (blue) line is the same, but without the shape correction. The dotted (grey) line represents the SM prediction from simulation. The lower panel shows the ratio of the data to the data-driven background distribution with shape corrections. For bins from 1000 GeV onwards, a wider bin width is chosen to reduce statistical uncertainties in the background estimation from the data control region. The horizontal error bars on the data points indicate the bin width.

FIG. 3. The reconstructed VLQ candidate mass in the $W^{-}qq$ category for the muon channel (left) and the electron channel (right), for the background prediction and the data. The solid bold (blue) line is the background distribution estimated from data, with a final shape correction that accounts for the difference observed in simulation, of negatively charged $W$ boson events to positively charged $W$ boson events. Finally, we apply a shape correction to account for the difference observed in the $W +$ jets simulation between the control region and the $W^{-}qq$ signal region, which originates from the different forward jet and lepton charge requirements.

The reconstructed mass of the VLQ candidate in the $Zqq$ category is shown for data and the simulated signal sample in Fig. 4, for the muon and electron channels. The SM background is completely dominated by the $Z +$ jets process.

B. Search for pair production

1. Single-lepton channel

In the single-lepton event categories optimized for the search for pair produced VLQs, each of the selected events must contain exactly one charged lepton (muon or electron) and at least four jets. The jet multiplicity requirement ensures that there is no overlap with the single-lepton $W^{-}qq$ category selection outlined in Sec. VI A, which selects events with at most three jets. The jet collection may consist of AK5 jets or also of the subjets of a $V$-tagged CA8 jet, where $V$ indicates a $W$, $Z$, or Higgs boson.

For heavy VLQs the quark pair from the hadronic decay of the $V$ boson may become so collimated that the overlapping hadronic showers cannot be resolved as separate jets. This means it is not possible to perform a kinematic fit to the final state and therefore the signal reconstruction efficiency drops. The CA8 jets with $p_T > 200$ GeV are used to identify the merged hadronic $V$ boson decays by applying a jet pruning algorithm, which resolves the merged jets into subjets. The efficiency drop caused by the jet merging at high VLQ masses can be recovered by using the subjets in the kinematic fit.

A pruned CA8 wide-jet mass is equal to the invariant mass of the subjets. A CA8 jet is considered to be: $W$-tagged if the pruned jet mass satisfies $60 < M_{jet} < 100$ GeV, $Z$-tagged if it satisfies $65 < M_{jet} < 115$ GeV, or $H$-tagged
if $100 < M_{\text{jet}} < 140$ GeV. If two subjets cannot be resolved, no $V$ tagging is done. The three different $V$ tagging selections overlap, such that the same event can be selected in different categories. As explained at the end of this section, the overlap is removed in the final distributions and each event is only counted once.

If the $V$-tagged jet overlaps with any AK5 jets, the AK5 jet is replaced with the two subjets of the matched CA8 jet.

Jets are considered as overlapping if $\Delta R < 0.04$, where $\Delta R$ is constructed using the directions of the CA8 and AK5 jets. The $b$ tagging of subjets is used in case of H-tagged CA8 jets.

Muon (electron) candidates in selected events contain tight muons (electrons) with $p_T > 45(30)$ GeV. Events in the $\mu$+jets (e+jets) channel must satisfy $p_T^{\text{miss}} > 20(30)$ GeV. Events having a loose muon or electron in addition to a tight lepton are vetoed. For this selection, loose leptons are defined as in Table II, except that loose electrons have relative isolation $I_{\text{rel}} < 0.2$ and $p_T > 20$ GeV. The jet collection described previously is used in a kinematic fit after the following additional selection requirements. Selected AK5 jets must have $p_T > 30$ GeV, while CA8 jets should have $p_T > 200$ GeV. All jets should satisfy $|\eta| < 2.4$. We require the presence of at least four jets, and the highest four $p_T$-ordered jets in the collection must satisfy $p_T > 120, 90, 50,$ and $30$ GeV, respectively.

We perform constrained kinematic fits of the selected events to the hypotheses described by Eqs. (3), (4) and (5). The kinematic reconstruction of events is performed using the HitFit package [46], which was developed by the D0 experiment at Fermilab [47] for the measurement of the top quark mass in the lepton + jets channel.

The fit is performed by minimizing a $\chi^2$ quantity constructed from the differences between the measured value of each momentum component for each reconstructed object and the fitted value of the same quantity divided by the corresponding uncertainties. The four-momenta of the final-state particles are subject to the following constraints:

$$m(\ell \nu) = m_W,$$  \hspace{1cm} (7)

$$m(q\bar{q}) = m_W,$$ or $$m(q\bar{q}) = m_Z,$$ or $$m(b\bar{b}) = m_H,$$ \hspace{1cm} (8)

$$m(\ell \nu q_\ell) = m_{\text{hadr}} = m_{\text{fit}},$$ \hspace{1cm} (9)

where $m_W$ denotes the $W$ boson mass, $m_Z$ the $Z$ boson mass, and $m_H$ the Higgs boson mass, with the values taken from the PDG [48]. The $m_{\text{hadr}}$ variable represents the mass of the three quarks on the hadronic side of the decay $[m(q\bar{q} q_h), m(q\bar{q} q_h), \text{or } m(b\bar{b} q_h)]$. The kinematic fit is performed for each $V$ hypothesis in parallel.

The $z$ component of the neutrino momentum is estimated from one of the two constraints given above that contain the neutrino momentum, with a two-fold quadratic ambiguity. The solutions found for the $z$ component of the neutrino momentum are used as starting values for the fit. If there are two real solutions, they are both taken in turn, doubling the number of fitted combinations. In the case of complex solutions, the real part is taken as a starting value. Using one constraint for calculation of $z$ component of the neutrino momentum leaves only two constraints for the kinematic fit. Only the combinations for which the $\chi^2$
probability of the fit exceeds 0.1% are accepted. If the jet collection contains more than four jets, then the five highest \( p_T \) jets are considered, and all possible combinations of four jets are checked.

In order to distinguish between jets originating from quarks and from gluons, we use the quark-gluon likelihood discrimination tagger (QGT) [49]. To reduce the combinatorial background in the assignment of jets to final-state quarks, \( V \) tagging, QGT tagging, and \( b \) tagging information is used. If a \( V \) tag is present, only combinations where the subjets of the \( V \)-jet match decay products of the corresponding boson are considered. The QGT tag requirements are then applied to those jets which are assigned to the \( \{ q_r, q_h \} \) quark pair. To suppress jets that may have originated from gluons we require the QGT discriminant values to satisfy the requirements QGT\(_{q_r} > 0.4 \) or QGT\(_{q_h} > 0.4 \). This excludes combinations in which both light quark jets have discriminant values favoring gluons.

A b-tagged jet veto is applied to the jets that have been assigned to the quark pair \( \{ q_r, q_h \} \). Since the \( V \)-tagged events have a better signal-to-background ratio, we apply softer b-tag selection requirements for this event category, as described in Table V. A more stringent requirement is applied on events without a \( V \) tag.

Additional \( b \) tagging requirements are applied to the jets associated with a Higgs boson decay. For H-tagged events, at least one jet from the Higgs boson decay must have a CSVL \( b \) tag, and for non-H-tagged events, at least one jet must have a CSVM \( b \) tag.

After applying the kinematic fit we impose an additional threshold on \( S_T \): \( S_T > 1000 \) GeV, where \( S_T \) is calculated from jets selected during the kinematic fit, using post-fit transverse momentum values. The \( S_T \) requirement strongly suppresses the remaining background.

Table VI presents the event yields obtained after applying the selections described above. There is good agreement between data and the expected SM background. The number of expected signal events is also presented.

The result of the kinematic fit is one mass distribution per reconstruction hypothesis and lepton channel, as shown in Fig. 5. The mass distributions are presented for the \( \mu + \) jets channel in the plots on the left, and for the \( e + \) jets channel in the plots on the right. In the case of \( e + \) jets events, the contribution from multijets is estimated from control samples in data. Events are selected that pass the electron trigger, but contain objects that satisfy inverted electron identification requirements. The normalisation of the multijet contribution is determined from a maximum likelihood fit of the observed \( p_T^{\text{miss}} \) distribution. The shapes in this fit are predicted by the MC simulation, where electroweak backgrounds are constrained to their expected cross sections and float within uncertainties, while the multijet normalization is allowed to float freely.

The uppermost row of distributions in Fig. 5 are those associated with the \( W_qW_q \) reconstruction, while the middle row corresponds to the \( W_qZ_q \) reconstruction, and the lowest row, to the \( W_qH_q \) reconstruction. For both

### Table V. Combinations of pairs of jets that have not been identified as \( V \)-jet matches, which can be accepted for matching to the quark pair \( \{ q_r, q_h \} \). In the left column, the group with the lowest available b-tag content is chosen, and within that group, the combination with the lowest \( \chi^2 \) is selected. In the right column, only the antitagged category is accepted.

| Events with V-tag | Events without V-tag |
|-------------------|----------------------|
| 0 CSVL \( b \) tags | 0 CSVL \( b \) tags |
| 1 CSVL \( b \) tag only; no CSVM \( b \) tags |
| 2 CSVL \( b \) tags; no CSVM \( b \) tags |

### Table VI. Numbers of expected background events from simulation and of data events in the \( W_qW_q \), \( W_qZ_q \), and \( W_qH_q \) channels after applying the single-lepton event selection in the search for pair produced VLQs. For the separate background components the indicated uncertainties are statistical only, originating from the limited number of MC events, while for the total background yield the combined statistical and systematic uncertainty is given.

| Channel | \( W_qW_q \) | \( W_qZ_q \) | \( W_qH_q \) |
|---------|-------------|-------------|-------------|
| Background process | \( \mu + \) jets | \( e + \) jets | \( \mu + \) jets | \( e + \) jets | \( \mu + \) jets | \( e + \) jets |
| \( t\bar{t} + \) jets | 257 ± 5 | 269 ± 5 | 295 ± 6 | 304 ± 7 | 224 ± 6 | 241 ± 6 |
| \( W+E_{T} \geq 3 \) jets | 396 ± 13 | 462 ± 14 | 426 ± 12 | 497 ± 14 | 42 ± 4 | 42 ± 4 |
| Single top quark | 13 ± 2 | 25 ± 3 | 13 ± 2 | 30 ± 4 | 11 ± 2 | 17 ± 3 |
| \( Z/\gamma^{*} + \geq 3 \) jets | 27 ± 2 | 27 ± 2 | 30 ± 2 | 30 ± 2 | 2.8 ± 0.5 | 2.9 ± 0.5 |
| \( WW, WZ, ZZ \) | 10 ± 1 | <1 | 10 ± 1 | <1 | 1.7 ± 0.6 | <1 |
| Multijet | <1 | 59 ± 4 | <1 | 59 ± 4 | <1 | 11 ± 2 |
| Total observed | 703 ± 80 | 840 ± 100 | 773 ± 86 | 920 ± 110 | 282 ± 37 | 314 ± 41 |
| \( m_Q = 600 \) GeV | 112 | 117 | 63 | 64 | 36 | 35 |
| \( m_Q = 800 \) GeV | 20 | 20 | 11 | 11 | 6.5 | 5.7 |
| \( m_Q = 1000 \) GeV | 3.3 | 3.3 | 1.8 | 2.0 | 1.1 | 0.8 |
FIG. 5. Reconstructed mass distributions for $WqWq$ (uppermost), $WqZq$ (middle), and $WqHq$ (lowest) reconstruction from a kinematic fit. Plots on the left are for the $\mu +$ jets channel and on the right, for the $e +$ jets channel. The distribution for pair-produced VLQs of mass 700 GeV for $B_W = 1.0$ (uppermost), $B_W = B_Z = 0.5$ (middle) and $B_W = B_H = 0.5$ (lowest) are scaled up for visibility by a factor of 5, 10 and 5, respectively. The hatched bands represent the combined statistical and systematic uncertainties.
the $WqZq$ and $WqHq$ reconstruction, the expected pair-produced VLQ signals are shown for $B(Q \to Wq) = 0.5$ and $B(Q \to Zq) = 0.5$ or $B(Q \to Hq) = 0.5$, respectively. These distributions show good agreement between data and the expected SM background.

Following the strategy described in Ref. [50] we then further tighten the $S_T$ requirement to $S_T > 1240$ GeV. This improves the signal-to-background ratio. At the same time we combine the $\mu^+ + \text{jets}$ and $e^+ + \text{jets}$ events, and use the resulting $m_\text{fit}$ distributions for the cross section limit calculations. Figure 6 shows these $m_\text{fit}$ distributions for the $WqWq$ (uppermost), $WqZq$ (middle), and $WqHq$ (lowest) reconstruction.

We find that the $WqWq$ reconstruction gives a better expected mass limit than the $WqZq$ reconstruction even for high values of $B(Q \to Zq)$. The events selected and reconstructed for the $WqWq$ and $WqZq$ hypotheses are highly correlated, with an 82% overlap between the two. Furthermore, since the $WqWq$ reconstruction is more sensitive, we do not consider the $WqZq$ reconstruction further, and use only the $WqWq$ reconstruction for all branching fraction combinations of the VLQ decaying to a $W$ boson or a $Z$ boson. The $WqHq$ reconstruction improves the expected limits for large decay branching fractions of the VLQ into a Higgs boson. The events selected for the $WqHq$ reconstruction have a relatively small correlation.
FIG. 7. Mass distribution for the $WqHq$ reconstruction from a kinematic fit, for combined $\mu +$ jets and $e +$ jets channels and for events with $S_T > 1240$ GeV. Events appearing also in the $WqWq$ sample have been removed. The distribution for pair-produced VLQs of mass 800 GeV for $B_W = B_H = 0.5$ is scaled up by a factor of 15 for visibility. The hatched band represents the combined statistical and systematic uncertainties. The horizontal error bars on the data points only indicate the bin width.

with those selected for the $WqWq$ channel events, with only a 25% event overlap. We therefore use $WqHq$ reconstructed events and combine them with $WqWq$ events. Events in the $WqHq$ selection that also appear in the $WqWq$ selection are removed, so that there is no double counting. Figure 7 shows the reconstructed mass for $WqHq$ events where events overlapping with the $WqWq$ reconstruction have been removed. Table VII shows the number of selected events after applying the stricter $S_T$ requirement for both the $WqWq$ reconstruction and the $WqHq$ reconstruction, excluding those events that appear in both samples.

The distributions in Figs. 6 (upper left) and 7 of the reconstructed mass are used in the rest of the analysis. The binning in these distributions has been chosen such that the statistical uncertainty on the background expectation in each bin is less than 20%.

2. Dilepton and multilepton channels

The event categories with at least two leptons optimized for the search for pair produced VLQs make use of the collections of central jets and antitagged jets defined in Sec. VI A, in addition to b-tagged jets, which are required to have a $b$ tagging discriminant above the CSVM threshold.

We categorize the events according to the number of tight leptons along with selection criteria applied to the jets and the missing transverse momentum. Each of the event categories is designed to be particularly sensitive to one or more of the pair production topologies presented in Table I. This is reflected in the names used as identifiers for the categories: dileptonic $WqWq$, $ZqHq$, dileptonic $VqZq$, and multileptonic $VqZq$, where $V$ indicates a $W$ or $Z$ boson. For the decay channel $QQ \rightarrow WqHq$, no dedicated category has been defined, to avoid an overlap of selected events with the single-lepton categories described in the previous section.

The definition of each event category optimized for pair production is summarized in Table VIII. In all event categories except dileptonic $WqWq$, a leptonically decaying Z boson candidate is reconstructed, from two same-flavor opposite-sign dileptons, imposing a requirement on the dilepton mass $m_{\ell\ell}$, as described in Table VIII. Thresholds are imposed on the transverse momentum $p_T(Z)$ of the $Z$ boson candidate.

The event yields for the observed data as well as for the expected SM backgrounds are shown in Table IX for the muon channel and the electron channel. In the case of $\mu$-e dilepton events (for the dileptonic $WqWq$ event category only), the event is assigned to the muon channel or the electron channel depending on which trigger the event has passed online, with the priority given to the muon trigger. If the event has passed the muon trigger, the selected muon has $p_T > 30$ GeV and the electron has $p_T > 20$ GeV, then this event will be assigned to the muon channel, even if the event also passed the electron trigger. If the event has passed the electron trigger as well as the muon trigger, the selected electron has $p_T > 30$ GeV and the muon has $p_T$ in the range of 20–30 GeV, then the event will be assigned to

| Channel | $\mu +$ jets | $e +$ jets | $\mu +$ jets | $e +$ jets |
|---------|--------------|------------|--------------|------------|
| Background process | Events | Events | Events | Events |
| $t\bar{t}$ | 61 $\pm$ 3 | 65 $\pm$ 3 | 34 $\pm$ 3 | 46 $\pm$ 3 |
| $W + \geq 3$ jets | 103 $\pm$ 7 | 129 $\pm$ 8 | 8 $\pm$ 2 | 11 $\pm$ 3 |
| Single top quark | 2 $\pm$ 1 | 9 $\pm$ 2 | 2 $\pm$ 1 | 3 $\pm$ 1 |
| $Zf\bar{f} + \geq 3$ jets | 7 $\pm$ 1 | 6 $\pm$ 1 | $< 1$ | 1.0 $\pm$ 0.4 |
| $WW$, $WZ$, $ZZ$ | 3 $\pm$ 1 | $< 1$ | $< 1$ | $< 1$ |
| Multijets | $< 1$ | 15 $\pm$ 2 | $< 1$ | 3 $\pm$ 1 |
| Total background | 176 $\pm$ 21 | 224 $\pm$ 26 | 44 $\pm$ 7 | 64 $\pm$ 10 |
| Observed | 199 | 233 | 51 | 61 |
| Signal ($m_Q=600$ GeV) | 53 | 54 | 5.7 | 5.7 |
| Signal ($m_Q=800$ GeV) | 15 | 16 | 1.5 | 1.7 |
| Signal ($m_Q=1000$ GeV) | 2.9 | 3.1 | 0.3 | 0.2 |
the electron channel. In the final case where the event only passes the electron trigger, the selected electron has $p_T > 30$ GeV and the muon has $p_T > 20$ GeV, the event will be assigned to the electron channel.

In each of the mutually exclusive event categories an observable is constructed that efficiently discriminates SM background events from VLQ processes. In several of the event categories we reconstruct the mass of the VLQ.

### TABLE VIII. Event categories as optimized for the VLQ pair production, with at least two leptons. The categories are based on the number of tight muons or electrons present in the event, along with additional criteria optimized for specific VLQ topologies. Events containing any additional loose leptons are excluded.

| Event category | Tight leptons ($\mu, e$) | Additional selection criteria |
|----------------|--------------------------|-------------------------------|
| Dileptonic $WqWq$ | 2 opposite-sign | $\geq 2$ selected central jets, all antitagged |
|               | Leading $p_T > 30$ GeV | Leading $p_T > 200$ GeV |
|               | Subleading $p_T > 20$ GeV | Subleading $p_T > 100$ GeV |
|               | $|m_{\ell\ell} - m_Z| > 7.5$ GeV (same flavor) | $p_T^{\text{miss}} > 60$ GeV |
| $ZqHq$ | 2 opposite-sign same-flavor | $\geq 3$ selected central jets, $\geq 2$ antitagged |
|               | Leading $p_T > 30$ GeV | Leading $p_T > 200$ GeV |
|               | Subleading $p_T > 20$ GeV | Subleading $p_T > 100$ GeV |
|               | $|m_{\ell\ell} - m_Z| < 7.5$ GeV | $p_T(Z \rightarrow \ell\ell) > 150$ GeV |
|               | $p_T(Z \rightarrow \ell\ell) > 150$ GeV | $\geq 4$ selected central jets, $\geq 2$ antitagged |
| Dileptonic $VqZq$ | 2 opposite-sign same-flavor | Leading $p_T > 200$ GeV |
|               | Leading $p_T > 30$ GeV | Subleading $p_T > 100$ GeV |
|               | Veto events with b-tagged jets | $|m_{\ell\ell} - m_Z| < 7.5$ GeV |
|               | $p_T(Z \rightarrow \ell\ell) > 150$ GeV | $p_T(Z \rightarrow \ell\ell) > 150$ GeV |
|               | $\geq 2$ selected central jets, all antitagged |
| Multileptonic $VqZq$ | 3 or 4 | Leading $p_T > 200$ GeV |
|               | Leading $p_T > 30$ GeV | Subleading $p_T > 100$ GeV |
|               | Others $p_t > 20$ GeV | $|m_{\ell\ell} - m_Z| < 7.5$ GeV |
|               | $p_T(Z \rightarrow \ell\ell) > 150$ GeV | $p_T(Z \rightarrow \ell\ell) > 150$ GeV |
|               | $p_T^{\text{miss}} > 60$ GeV (3 leptons) | $\Delta R(\ell, \ell) > 0.05$ (other flavor) |

### TABLE IX. Event yields in the muon and electron channels for the event categories with at least two leptons, optimized for the pair production search. For the separate background components the indicated uncertainties are statistical only, originating from the limited number of MC events, while for the total background yield the combined statistical and systematic uncertainty is given. The prediction for the signals is shown assuming branching fractions of $B_W = 0.5$ and $B_Z = B_H = 0.25$. The label “Other” designates the background originating from $t\bar{t}W$, $t\bar{t}Z$ and triboson processes.

| Channel | Muon | Electron | Muon | Electron | Muon | Electron |
|---------|------|---------|------|----------|------|---------|
| $t\bar{t} +$ jets | 62 $\pm$ 4 | 22 $\pm$ 2 | 2.1 $\pm$ 0.7 | 1.2 $\pm$ 0.4 | $< 1$ | $< 1$ |
| $W +$ jets | $< 1$ | $< 1$ | $< 1$ | $< 1$ | $< 1$ | $< 1$ |
| $Z +$ jets | 79 $\pm$ 6 | 55 $\pm$ 5 | 53 $\pm$ 3 | 41 $\pm$ 2 | 238 $\pm$ 5 | 202 $\pm$ 4 |
| Single top quark | 4.6 $\pm$ 1.5 | 1.7 $\pm$ 0.8 | $< 1$ | $< 1$ | $< 1$ | $< 1$ |
| $VV$ | 8.5 $\pm$ 1.0 | 3.5 $\pm$ 0.6 | 1.0 $\pm$ 0.2 | $< 1$ | 3.7 $\pm$ 0.4 | 3.6 $\pm$ 0.4 |
| Multijet | 14 $\pm$ 2 | 9.2 $\pm$ 2.6 | $< 1$ | $< 1$ | $< 1$ | $< 1$ |
| Other | 1.8 $\pm$ 0.2 | $< 1$ | 1.3 $\pm$ 0.2 | $< 1$ | $< 1$ | $< 1$ |
| Total background | 170 $\pm$ 21 | 92 $\pm$ 17 | 58 $\pm$ 14 | 43 $\pm$ 10 | 243 $\pm$ 45 | 207 $\pm$ 37 |
| Observed | 174 | 95 | 54 | 48 | 249 | 201 |
| Signal ($m_Q = 600$ GeV, $\tilde{R}_W = 0.1$) | 11.7 | 4.2 | 3.9 | 3.4 | 9.1 | 7.4 |
| Signal ($m_Q = 1100$ GeV, $\tilde{R}_W = 1$) | 0.6 | 0.2 | 0.3 | 0.2 | 1.4 | 1.2 |
candidate. In other categories, where the mass of the VLQ candidate is poorly reconstructed, or where the event yield is too low, we use a simpler observable such as the $S_T$ variable defined in Eq. (6) or the event count.

The VLQ candidate mass is reconstructed in the $ZqHq$ and the dileptonic $VqZq$ event categories from two leptons forming a $Z$ boson candidate and a jet that potentially corresponds to the light quark from the VLQ decay. For the latter, we choose the highest $p_T$ antitagged jet with the largest $\Delta R$ separation from the $Z$ boson candidate. The resulting mass distributions are shown in Figs. 8 and 9, for the $ZqHq$ and dileptonic $VqZq$ categories, respectively. The background consists mainly of $Z +$ jets events with a large contribution from those in which the $Z$ boson is associated with heavy-flavor jets, because of the required presence of at least one b-tagged jet.

In the dileptonic $WqWq$ event category we use the $S_T$ variable to discriminate between SM and VLQ processes as
in data to estimate the contribution from misidentified leptons passing the tight-lepton selection criteria. This contribution is very small.

**VII. COMBINATION**

We do not observe a significant excess of events over the background prediction, and combine the results of the single and pair production searches by calculating upper limits on the signal production cross sections and lower limits on the mass of the VLQs. The selection criteria defining the event categories optimized for single VLQ production and those optimized for pair VLQ production, are orthogonal. The discriminating observables for the different event categories and the methods by which they are reconstructed are summarized in Table XI. The distributions (templates) used in the limit calculation contain those of the observables in the single-lepton and dilepton event categories in the muon and the electron channel, shown in Figs. 3, 4, 8, 9, and 10, where the binning of the distributions is chosen in such a way that there are at least 10 expected background events per bin. In the single-lepton event category optimized for VLQ pair production, the distributions in Figs. 6 (upper left) and 7 of the reconstructed mass are used. In the event categories that require three and four leptons, we use the event counts of Table X.

The limit calculation is performed using a Bayesian interpretation [48]. Systematic uncertainties are taken into account as nuisance parameters. For uncertainties affecting the shapes of the variables used in the search, alternative templates are produced by varying each source of uncertainty within ±1 standard deviation, and associating the varied templates with Gaussian prior constraints of the corresponding nuisance parameters. Uncertainties affecting only the normalization are included, using log-normal prior constraints. A flat prior probability density function on the total signal yield is assumed. The likelihood function is marginalized with respect to the nuisance parameters representing the systematic uncertainties that arise from shape and global normalization variations. The shapes of the background and signal templates vary with the appropriate nuisance parameters. Statistical uncertainties

---

**Table X.** The total number of estimated background events compared to the number of observed events, in the multileptonic $VqZq$ event category, with either 3 or 4 leptons. The numbers of expected signal events for two different signal hypotheses are shown. The indicated uncertainties are statistical only, originating from the limited number of MC events.

| Source                | Observed | Total background | Misidentified lepton background | Irreducible background | Signal ($m_Q = 600 \text{ GeV}$, $B_W = 0.5$, $B_Z = 0.25$) | Signal ($m_Q = 400 \text{ GeV}$, $B_Z = 1.0$) |
|-----------------------|----------|-----------------|---------------------------------|------------------------|-------------------------------------------------|------------------------------------------|
| Data                  | 2        | 2               | 0.06 ± 0.06                     | 0.4 ± 0.1              | 2.1                                             | 4.9                                      |
| 0.5 ± 0.1             |          |                 |                                 |                        |                                                 |                                          |
| Observed              | 2        | 600             | 0.06                            | 0.4 ± 0.1              | 2.1                                             | 4.9                                      |
| Total background      | 4        | 600             | 0.06                            | 0.4 ± 0.1              | 2.1                                             | 4.9                                      |
| Misidentified lepton  | 2        | 600             | 0.06                            | 0.4 ± 0.1              | 2.1                                             | 4.9                                      |
| Irreducible background| 2        | 600             | 0.06                            | 0.4 ± 0.1              | 2.1                                             | 4.9                                      |

---

In the multileptonic $VqZq$ event category (three or four leptons), the number of events is too low to obtain a meaningful distribution. Instead, the event count is used as observable. The numbers of events observed and expected are summarized in Table X. The main SM background originates from irreducible diboson and triboson processes with three prompt charged leptons. We use control samples

---

**Figure 10.** The $S_T$ variable in the dileptonic $WqWq$ event category, in the muon channel (upper) and in the electron channel (lower). The contributions of simulated events where the $Z$ boson is produced in association with light-flavor (LF) jets and heavy-flavor (HF) jets are shown separately. The distribution for a heavy VLQ signal of mass 600 GeV and $\kappa_W = 0.1$ (for $B_W = 0.5$ and $B_Z = 0.25$) is scaled up by a factor of 10 for visibility. The hatched bands represent the combined statistical and systematic uncertainties.
TABLE XI. Discriminating variables used in the different event categories. The overlap of events in the \( WqWq \) and \( WqHq \) categories is removed, as explained in Sec. VI B 1.

| Event category | Discriminating variable | Reconstructed using | Shown in |
|----------------|-------------------------|---------------------|----------|
| \( W^{+}qq \)  | VLQ mass                | Lepton, neutrino, leading central jet | Figs. 2 and 3 |
| \( Zqq \)      | VLQ mass                | Two opposite-sign leptons, leading central jet | Fig. 4 |
| \( WqWq \)     | VLQ mass                | Kinematic fit, see Sec. VI B 1 | Figs. 5 and 6 |
| \( WqHq \)     | VLQ mass                | Kinematic fit, see Sec. VI B 1 | Figs. 5–7 |
| \( ZqHq \)     | VLQ mass                | Two opposite-sign leptons, high-\( p_T \) antitagged, jet with the largest \( \Delta R \) separation from the Z boson candidate | Fig. 8 |
| Dileptonic \( VqZq \) | VLQ mass                | Two opposite-sign leptons, high-\( p_T \) antitagged, jet with the largest \( \Delta R \) separation from the Z boson candidate | Fig. 9 |
| Dileptonic \( WqWq \) | \( S_T \)                  | See Sec. II A          | Fig. 10 |
| Multileptonic \( VqZq \) | Event count             | See Sec. VI B 2        | Table X |

A. Systematic uncertainties

The uncertainties in the \( \bar{t}t \) total cross section, electroweak and multijet background yields, integrated luminosity, lepton efficiencies, the choice of PDFs, and constant data-to-simulation scale factors affect only the normalization. Uncertainties that affect the shape and normalization of the distributions include those in the jet energy scale, jet energy resolution, \( p_T^{\text{miss}} \) resolution, \( b \) tagging efficiency, QGT tagging efficiency, number of additional \( pp \) interactions per bunch crossing, and the factorization and renormalization scales assumed in the simulation. Some of the uncertainties listed above have a negligible impact on the distributions and are neglected in the limit calculation.

The main backgrounds are \( \bar{t}t, W+jets, and Z+jets \) production. A 15% uncertainty in the cross section for \( \bar{t}t \) production is taken from the CMS measurement [52]. In the single-production event categories as well as the pair-production categories with multiple leptons, we use values for the normalization uncertainty in the \( W+jets \) and \( Z+jets \) background contributions, which are obtained from estimates based on data. The values are 20% for the light-flavor component, and 30% for the heavy-flavor component. These uncertainties are estimated to cover the changes in the normalizations induced by modifying the kinematic requirements that define the control samples. The uncertainties corresponding to the normalization of the smaller single top quark, diboson, \( \bar{t}tZ+jets, \bar{t}tW+jets, \) and triboson backgrounds in these categories are taken from the corresponding experimental measurements or the theoretical calculations. In the single-lepton pair-production categories, in which a kinematic fit is performed, the normalization of the non-\( \bar{t}t \) background processes has been assigned an uncertainty of 50%, reflecting the large uncertainty in the heavy-flavor component of the \( W+jets \) process and in other background processes, in the high-\( S_T \) signal region.

The integrated luminosity has an uncertainty of 2.6% [53]. Trigger efficiencies, lepton identification efficiencies, and data-to-simulation scale factors are obtained from data using the decays of Z bosons to lepton pairs. The uncertainties associated with all of these lepton related sources are included in the selection efficiency uncertainty, and together they amount to a total uncertainty of 3%.

The PDF uncertainties are estimated by varying up and down by one standard deviation the CTEQ6 PDF set parameters. Only the changes in acceptance caused by these uncertainties, not the change in total cross section, are propagated. For each simulated event, the weight corresponding to each varied PDF parameter is calculated, and an envelope for the distributions of the observables is created by taking the maximum and minimum of the variations bin by bin. This results in a normalization uncertainty of 1.4% for the signal and 8% for the background, with a negligible impact on the shape of the distributions.

The uncertainty in the jet energy scale is evaluated by scaling the jet energy in the simulation by the \( \eta \) and \( p_T \) dependent uncertainties, ranging from 0.5% to 2.3% [40]. The \( \eta \) dependent scale factors that smear the jet energy resolution are varied within their uncertainty, changing the scale factors between 2.4% and 3.8% depending on the absolute value of \( \eta \). Both AK5 and CA8 jet collections are subject to these variations. The systematic variations on the jet energy scale and resolution are applied before the splitting of the CA8 jets in subjets. The variations for subjets are done proportionally to the variations of their parent CA8 jet.

The changes in jet momentum resulting from the AK5 jet energy scale variations are propagated to the \( p_T^{\text{miss}} \). The effect of the unclustered energy uncertainty on \( p_T^{\text{miss}} \) is evaluated by varying the unclustered energy by \( \pm 10\% \), and is found to be negligible.

The systematic uncertainty in the \( b \) tagging efficiency is estimated by varying the data-to-simulation scale factors, for both medium and loose working points, within their uncertainty, separately for heavy-flavor (\( b \) and \( c \)) jets and
light-flavor jets. The relative precision on the heavy-flavor scale factors is 2–4% for $p_T$ below 120 GeV and about 5%–9% at the highest jet transverse momenta considered [45]. The scale factors for light-flavor jets are measured with a precision of about 5%–13%. For the evaluation of the systematic uncertainty originating from the QGT, the QGT discriminant values of the jets in the simulation are smeared, depending on the flavor, $p_T$, and $\eta$ of the jet [49]. The observed variations in the number of selected events, and in particular the variations in the signal inefficiencies, are very small and neglected.

To evaluate the uncertainty related to the modeling of multiple interactions in the same bunch crossing, the average number of interactions in the simulation is varied by ±5% relative to the nominal value. The impact of these variations on the distributions is found to be negligible.

The uncertainty in the factorization and renormalization scales assumed in the simulated $t\bar{t}$ sample is estimated by varying the scales simultaneously by a factor of two, and a factor of one half, relative to the nominal value.

Several systematic uncertainties affect the backgrounds estimated from control regions in data. The uncertainty in the estimated misidentified lepton background is considered in the multileptonic $VqZq$ category. In the $W^-q\bar{q}$ category, an uncertainty is assigned in the shape correction applied to the reconstructed mass distribution, which accounts for the different selection requirements between the control region and the signal region. The templates modeling this uncertainty are constructed by applying a shape correction twice the size of the nominal correction, and not applying a shape correction.

The systematic uncertainties in the integrated luminosity, the lepton efficiency scale factors, the jet energy scale, the jet energy resolution, and the $b$ tag efficiency and mistag rate scale factor uncertainties are considered as fully correlated across all channels. The uncertainties in the normalization of the different background processes are considered as uncorrelated among the event categories that make use of a kinematic fit and those that do not, because of the different signal selection procedures. The expected and observed mass limits change by less than 5 GeV when treating the $t\bar{t} + $ jets normalization uncertainty as completely correlated across all categories.

**VIII. RESULTS**

The 95% C.L. limit on the product of the production cross section and the branching fraction as a function of the VLQ mass, considering only single production of down-type VLQs and the corresponding optimized categories, is shown in Fig. 11. The upper (lower) plot shows the scenario where a nonzero $\bar{\kappa}_W$ ($\bar{k}_Z$) is considered while setting $\bar{k}_W = 0$ ($\bar{\kappa}_W = 0$) and including only the $W^-q\bar{q}$ ($Zq\bar{q}$) event category in the limit setting procedure. The LO theoretical predictions for the cross section are superimposed. The expected and observed mass limits change by less than 5 GeV when treating the $t\bar{t} + $ jets normalization uncertainty as completely correlated across all categories.

![Graph showing the 95% C.L. exclusion limits on the product of the production cross section and the branching fraction](image)

**FIG. 11.** The 95% C.L. exclusion limits on the product of the production cross section and the branching fraction, considering only single production of down-type VLQs, and assuming a neutral current coupling of zero (upper) or a charged current coupling of zero (lower). The median expected and observed exclusion limits are indicated with a dashed and a solid line, respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68% and 95%, respectively, of the distribution of limits expected under the background-only hypothesis. The corresponding LO theory predictions are superimposed. The predictions are represented by a solid black line centered within a blue band, which shows the uncertainty of the calculation. The uncertainties are determined based on the choice of PDF set along with the renormalization and factorization scales.

Comparing the effect of either doubling or halving the central value of the scale. The PDF uncertainty is determined using the 44 eigenvectors of the CTEQ66 PDF set [54]. A mass of 1755 (1620) GeV is observed (expected) to be excluded at the 95% C.L. for $\bar{k}_W = 1.0$ and $B_z = 1$, and a mass of 1160 (1170) GeV is observed (expected) to be excluded at the 95% C.L. for $\bar{k}_Z = 1.0$ and $B_z = 1$. The results of the branching fraction scans for the charged-current VLQ single-production coupling parameter $\bar{\kappa}_W = 1.0$, $\bar{k}_W = 0.7$, $\bar{k}_W = 0.4$, $\bar{k}_W = 0.1$ are shown in Figs. 12–15. For values of $\bar{\kappa}_W = 1.0$ and 0.7, single
production is by far the dominant signal production mode, while the relative importance of the pair-production mode increases in much of the parameter space for $\tilde{\kappa}_W = 0.4$, and even more so for $\tilde{\kappa}_W = 0.1$. The black shaded region below $B_W \approx 0.1$ in each branching fraction triangle indicates the region where care should be taken with the interpretation of the results. In this region, $B_W$ approaches 0, and as explained in Sec. II A, the neutral-current single-production strength parameter $\tilde{\kappa}_Z$ diverges and the limits cannot be calculated. Results for an alternative single-production coupling parametrization that does not exhibit divergent behavior throughout the scan are available in tabulated form in the Supplemental Material [15]. The results from a branching fraction scan based on the pair-production data alone are shown in Fig. 16. The lower limits on the mass, together with the uncertainties in the median expected limits, are presented in Tables XII–XVI.

FIG. 12. The median expected (upper) and observed (lower) combined lower mass limits represented in a triangular form, where each point of the triangle corresponds to a given set of branching fractions for the decay of a VLQ into a boson and a first-generation quark. The limit contours are determined assuming that $\tilde{\kappa}_W = 1.0$, which means that the signal is dominated by electroweak single production. The black shaded band near $B_W = 0$ shows a region where the results cannot be reliably interpreted because $\tilde{\kappa}_Z$ diverges, as explained in the text.

FIG. 13. The median expected (upper) and observed (lower) combined lower mass limits represented in a triangular form, where each point of the triangle corresponds to a given set of branching fractions for a VLQ decaying into a boson and a first-generation quark. The limit contours are determined assuming $\tilde{\kappa}_W = 0.7$, which means that the signal will be dominated by electroweak single production for most of the parameter space represented by the triangles. The black shaded band near $B_W = 0$ represents a region where results cannot be reliably interpreted because $\tilde{\kappa}_Z$ diverges, as explained in the text.
The existence of a heavy vector-like $D$ quark with a mass below 1595 GeV is excluded at 95% C.L. when using the following choice of model parameters: $\tilde{\kappa}_W = 1.0$, $B_W = 0.5$, and $B_Z = 0.25$. This limit may be compared with the expected value of 1460 GeV. In the case where the VLQ couples only to the $W$ boson, the observed (expected) limit at 95% C.L. is 1745 (1620) GeV.

The sensitivity to pair production of VLQs for the event categories in which a kinematic fit is performed becomes more important for lower $\tilde{\kappa}_W$. In the extreme case where only pair production is considered (as shown in Fig. 16), the added sensitivity of the combined analysis when compared to the categories that use a kinematic fit or not is illustrated using some example parameter choices, as shown in Table XVII. When the branching fraction for the
decay to a $W$ boson becomes large, the event categories using the kinematic fit to the VLQ signal mass become more important. For lower $B_W$ and relatively large $B_Z$ and $B_H$, the dilepton $ZqHq$ event category drives the sensitivity. Figure 17 shows the 95% C.L. limit on the production cross section as a function of the VLQ mass, for the scenario where only pair production of the VLQs is considered, and for two different parameter choices. In
Fig. 17 (upper) the result is shown for $B_W = 0.5$ and $B_Z = 0.25$. For this set of parameters, we exclude VLQs with masses below 685 GeV at 95% C.L., compared to an expected exclusion limit of 720 GeV. In Fig. 17 (lower), the exclusion limits are shown for $B_W = 1$. In this case we exclude VLQs with masses below 845 GeV at 95% C.L., compared to an expected lower limit of 825 GeV.

In this search we use signal mass distributions simulated using the narrow-width approximation, where the decay width is about 1% of the mass of the VLQ and is significantly less than the experimental resolution. We have verified that this approximation does not affect the results. At smaller mass values ($\sim 700$ GeV) and for a parameter space with an exclusion limit close to this mass, the theoretically calculated width reaches up to about 4%, which is still well below the experimental resolution (about 9% in the $Wqg$ category, for example). For the highest mass values probed in the analysis ($\sim 1800$ GeV), the width approaches the experimental resolution in part of the parameter space. This does not change the results, as the width of the signal mass distributions remains smaller than the bin size at these high masses.

The results tighten the constraints on the masses, cross sections and decay branching fractions of VLQs coupled to light-flavor quarks. In the scenario where the VLQ couples to first-generation quarks only via the $W$ boson, the results can be compared to those obtained previously. The presented exclusion limits in this paper are more stringent than those obtained by the ATLAS experiment, when considering single production of VLQs alone at $\sqrt{s} = 7$ TeV [12] and pair production of VLQs alone at 8 TeV [13].

### IX. SUMMARY

A search has been performed for vectorlike quarks coupled to light quarks, produced in either single-production or pair-production processes, in proton-proton collisions at $\sqrt{s} = 8$ TeV at the LHC. In the single-production mode the search has been performed for down-type quarks (electric charge of magnitude 1/3), while in the pair-production
mode the search is sensitive to decays of vectorlike quarks into up, down and strange quarks. Approaches with and without kinematic fits have been used to perform this search. No significant excess over standard model expectations has been observed. Lower limits on the mass of the vectorlike quarks have been determined by combining the results from both the single- and pair-production searches. Limits have also been extracted using the data from the pair-production search alone. For all processes considered, the lower mass limits range from 400 to 1800 GeV, depending on the vectorlike quark branching fractions for decays to $W$, $Z$, and Higgs bosons and the assumed value of the electroweak single-production strength. When considering pair production alone, vector-like quarks with masses below 845 GeV (825 GeV expected) are excluded for $B_W = 1$ and $B_W = 0.25$ (upper), and $B_W = 1$ (lower) as a function of the hypothetical VLQ mass, and for the scenario where only strong pair production of the VLQs is considered. The median expected and observed exclusion limits are indicated with a dashed and a solid line, respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68% and 95%, respectively, of the distribution of limits expected under the background-only hypothesis. The cross section from a full NNLO calculation [25], including uncertainties in the PDF description and the renormalization and factorization scales, is shown by the magenta band.

TABLE XVII. Comparison of several expected 95% C.L. lower mass limits for signal pair production only, illustrating the added sensitivity in the combination of the event categories that use and do not use a kinematic fit.

| Signal benchmark | Dilepton and multilepton channels | Single-lepton channel using kinematic fit | Combination |
|------------------|-----------------------------------|------------------------------------------|-------------|
| $B_W = 1.0$, $B_Z = 0.0$ | 725 GeV 810 GeV 825 GeV | 725 GeV 810 GeV 825 GeV |
| $B_W = 0.5$, $B_Z = 0.2$ | 585 GeV 680 GeV 720 GeV | 585 GeV 680 GeV 720 GeV |
| $B_W = 0.1$, $B_Z = 0.2$ | 600 GeV 405 GeV 630 GeV | 600 GeV 405 GeV 630 GeV |
| $B_W = 0.1$, $B_Z = 0.1$ | 420 GeV <400 GeV 525 GeV | 420 GeV <400 GeV 525 GeV |
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 centers 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 Secretariat for Higher Education, Science, Technology and Innovation, Ecuador; 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, Council via IUT23-4 and IUT23-6 and European Regional Ministry of Education and Research, Estonian Research Science, Technology and Innovation, Ecuador; the Ministry of Education and Technology, Croatian Science Foundation; the Czech Republic, Czech Science Foundation; the Ministry of Education, Science and Technological Development of Serbia; the Secretaría de Estado de Investigación, Desarrollo e Innovación, Programa Consolider-Ingenio 2010, Plan de Ciencia, Tecnología e Innovación 2013–2017 del Principado de Asturias and Fondo Europeo de Desarrollo Regional, Spain; the Swiss Funding Agencies (ETH Board, ETH Zurich, PSI, SNF, Unizh, Canton Zurich, and SER); the Ministry of Science and Technology, Taiwan; 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 Turkmish 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 Horizon 2020 Grant, Contract No. 675440 (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 Scientific and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), Contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; European Social Fund”; the National Priorities Research Program by Qatar National Research Fund; the Programa Clarín-COFUND del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); and the Welch Foundation, Contract No. C-1845.
[1] D. Choudhury, T. M. P. Tait, and C. E. M. Wagner, Beautiful mirrors and precision electroweak data, Phys. Rev. D 65, 053002 (2002).

[2] N. Arkani-Hamed, A. G. Cohen, E. Katz, and A. E. Nelson, The littlest Higgs, J. High Energy Phys. 07 (2002) 034.

[3] M. Schmaltz, Physics beyond the standard model (theory): Introducing the little Higgs, Nucl. Phys. B, Proc. Suppl. 117, 40 (2003).

[4] M. Schmaltz and D. Tucker-Smith, Little Higgs review, Annu. Rev. Nucl. Part. Sci. 55, 229 (2005).

[5] D. Marzocca, M. Serone, and J. Shu, General composite Higgs models, J. High Energy Phys. 08 (2012) 013.

[6] H.-C. Cheng, B. A. Dobrescu, and C. T. Hill, Electroweak symmetry breaking and extra dimensions, Nucl. Phys. B589, 249 (2000).

[7] J. Kang, P. Langacker, and B. D. Nelson, Theory and phenomenology of exotic isosinglet quarks and squarks, Phys. Rev. D 77, 035003 (2008).

[8] D. Guadagnoli, R. N. Mohapatra, and I. Sung, Gauged flavor group with left-right symmetry, J. High Energy Phys. 04 (2011) 093.

[9] A. Atre, G. Azuelos, M. Carena, T. Han, E. Ozcan, J. Santiago, and G. Unel, Model-independent searches for new quarks at the LHC, J. High Energy Phys. 01 (2011) 080.

[10] J. A. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Pérez-Victoria, A handbook of vector-like quarks: mixing and single production, Phys. Rev. D 88, 094010 (2013).

[11] B. W. Lee, C. Quigg, and H. B. Thacker, Weak interactions and single production, Phys. Rev. D 22, 1519 (1977).

[12] ATLAS Collaboration, Search for heavy vector-like quarks coupling to light quarks in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, Phys. Lett. B 712, 22 (2012).

[13] ATLAS Collaboration, Search for pair production of a new heavy quark that decays into a $W$ boson and a light quark in $pp$ collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, Phys. Rev. D 92, 112007 (2015).

[14] M. Buchkremer, G. Cacciapaglia, A. Deandrea, and L. Panizzi, Model-independent framework for searches of top partners, Nucl. Phys. B876, 376 (2013).

[15] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.97.072008 for results using an alternative parametrization of charged and neutral coupling strengths.

[16] CMS Collaboration, The CMS experiment at the CERN LHC, J. Instrum. 3, S08004 (2008).

[17] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, MadGraph 5: Going beyond, J. High Energy Phys. 06 (2011) 128.

[18] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, New generation of parton distributions with uncertainties from global QCD analysis, J. High Energy Phys. 07 (2002) 012.

[19] T. Sjöstrand, S. Mrenna, and P. Skands, PYTHIA 6.4 physics and manual, J. High Energy Phys. 05 (2006) 026.

[20] CMS Collaboration, Study of the underlying event at forward rapidity in $pp$ collisions at $\sqrt{s} = 0.9, 2.76$, and 7 TeV, J. High Energy Phys. 04 (2013) 072.

[21] CMS Collaboration, Event generator tunes obtained from underlying event and multiparton scattering measurements, Eur. Phys. J. C 76, 155 (2016).

[22] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, J. High Energy Phys. 11 (2004) 040.

[23] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with parton shower simulations: the POWHEG method, J. High Energy Phys. 11 (2007) 070.

[24] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: The POWHEG BOX, J. High Energy Phys. 06 (2010) 043.

[25] M. Czakon, P. Fiedler, and A. Mitov, Total Top-Quark Pair-Production Cross Section at Hadron Colliders through $O(a_s^2)$, Phys. Rev. Lett. 110, 252004 (2013).

[26] CMS Collaboration, Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 02 (2014) 024.

[27] CMS Collaboration, Observation of the Associated Production of a Single Top Quark and a $W$ Boson in pp Collisions at $\sqrt{s} = 8$ TeV, Phys. Rev. Lett. 112, 231802 (2014).

[28] CMS Collaboration, Measurement of the $W^+W^-$ and ZZ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV, Phys. Lett. B 721, 190 (2013).

[29] CMS Collaboration, Measurement of $W$ and $ZZ$ production in pp collisions at $\sqrt{s} = 8$ TeV in final states with b-tagged jets, Eur. Phys. J. C 74, 2973 (2014).

[30] S. Agostinelli et al. (GEANT4), Geant4—a simulation toolkit, Nucl. Instrum. Methods Phys. Res., Sect. A 506, 250 (2003).

[31] CMS Collaboration, Particle-flow reconstruction and global event description with the cms detector, J. Instrum. 12, P10003 (2017).

[32] CMS Collaboration, Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV, J. Instrum. 10, P06005 (2015).

[33] M. Cacciari, G. P. Salam, and G. Soyez, The anti-$k_t$ jet clustering algorithm, J. High Energy Phys. 04 (2008) 063.

[34] Yu. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, Better jet clustering algorithms, J. High Energy Phys. 08 (1997) 001.

[35] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, Eur. Phys. J. C 72, 1896 (2012).

[36] M. Cacciari and G. P. Salam, Dispelling the $N^3$ myth for the $k_t$ jet-finder, Phys. Lett. B 641, 57 (2006).

[37] CMS Collaboration, Determination of jet energy calibration and transverse momentum resolution in CMS, J. Instrum. 6, P11002 (2011).

[38] M. Cacciari, G. P. Salam, and G. Soyez, The catchment area of jets, J. High Energy Phys. 04 (2008) 005.

[39] M. Cacciari and G. P. Salam, Pileup subtraction using jet clustering algorithms, J. High Energy Phys. 08 (1997) 001.

[40] S. Ellis, C. K. Vermilion, and J. R. Walsh, Techniques for improved heavy particle searches with jet substructure, Phys. Rev. D 80, 051501 (2009).
A. M. Sirunyan et al.

PHYS. REV. D 97, 072008 (2018)
SEARCH FOR VECTORLIKE LIGHT-FLAVOR QUARK ...

PHYS. REV. D 97, 072008 (2018)

131 Catholic University of America, Washington DC, USA
132 The University of Alabama, Tuscaloosa, Alabama 35487, USA
133 Boston University, Boston, Massachusetts 02215, USA
134 Brown University, Providence, Rhode Island 02912, USA
135 University of California, Davis, California 95616, USA
136 University of California, Los Angeles, California 90095, USA
137 University of California, Riverside, California 92521, USA
138 University of California, San Diego, California 92093, La Jolla, USA
139 University of California, Santa Barbara, Department of Physics, Santa Barbara, California 93106, USA
140 California Institute of Technology, Pasadena, California 91125, USA
141 Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
142 University of Colorado Boulder, Boulder, Colorado 80309, USA
143 Cornell University, Ithaca, New York 14853, USA
144 Fairfield University, Fairfield, Connecticut 06430, USA
145 Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA
146 University of Florida, Gainesville, Florida 32611, USA
147 Florida International University, Miami, Florida 33199, USA
148 Florida State University, Tallahassee, Florida 32306, USA
149 Florida Institute of Technology, Melbourne, Florida 32901, USA
150 University of Illinois at Chicago (UIC), Chicago, Illinois 60607, USA
151 The University of Iowa, Iowa City, Iowa 52242, USA
152 Johns Hopkins University, Baltimore, Maryland 21218, USA
153 The University of Kansas, Lawrence, Kansas 66045, USA
154 Kansas State University, Manhattan, Kansas 66506, USA
155 Lawrence Livermore National Laboratory, Livermore, California 94551, USA
156 University of Maryland, College Park, Maryland 20742, USA
157 Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
158 University of Minnesota, Minneapolis, Minnesota 55455, USA
159 University of Mississippi, Oxford, Mississippi 38677, USA
160 University of Nebraska-Lincoln, Lincoln, Nebraska 68588, USA
161 State University of New York at Buffalo, Buffalo, New York 14260, USA
162 Northeastern University, Boston, Massachusetts 02115, USA
163 Northwestern University, Evanston, Illinois 60208, USA
164 University of Notre Dame, Notre Dame, Indiana 46556, USA
165 The Ohio State University, Columbus, Ohio 43210, USA
166 Princeton University, Princeton, New Jersey 08542, USA
167 University of Puerto Rico, Mayaguez, Puerto Rico 00681, USA
168 Purdue University, West Lafayette, Indiana 47907, USA
169 Purdue University Northwest, Hammond, Indiana 46323, USA
170 Rice University, Houston, Texas 77251, USA
171 University of Rochester, Rochester, New York 14627, USA
172 Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, USA
173 University of Tennessee, Knoxville, Tennessee 37996, USA
174 Texas A&M University, College Station, Texas 77843, USA
175 Texas Tech University, Lubbock, Texas 79409, USA
176 Vanderbilt University, Nashville, Tennessee 37235, USA
177 University of Virginia, Charlottesville, Virginia 22904, USA
178 Wayne State University, Detroit, Michigan 48202, USA
179 University of Wisconsin, Madison, Madison, Wisconsin 53706, USA

*Deceased.
1 Also at Vienna University of Technology, Vienna, Austria.
2 Also at State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China.
3 Also at Universidade Estadual de Campinas, Campinas, Brazil.
4 Also at Universidade Federal de Pelotas, Pelotas, Brazil.
5 Also at Université Libre de Bruxelles, Bruxelles, Belgium.
6 Also at Universidad de Antioquia, Medellin, Colombia.
7 Also at Joint Institute for Nuclear Research, Dubna, Russia.
8 Also at Suez University, Suez, Egypt.
9 Also at British University in Egypt, Cairo, Egypt.
Also at Fayoum University, El-Fayoum, Egypt.
Also at Helwan University, Cairo, Egypt.
Also at Université de Haute Alsace, Mulhouse, France.
Also at Skoltech Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia.
Also at CERN, European Organization for Nuclear Research, Geneva, Switzerland.
Also at RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany.
Also at University of Hamburg, Hamburg, Germany.
Also at Brandenburg University of Technology, Cottbus, Germany.
Also at Institute of Nuclear Research ATOMKI, Debrecen, Hungary.
Also at MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary.
Also at Institute of Physics, University of Debrecen, Debrecen, Hungary.
Also at IIT Bhubaneswar, Bhubaneswar, India.
Also at University of Visva-Bharati, Santiniketan, India.
Also at Institute of Physics, Bhubaneswar, India.
Also at University of Ruhuna, Matara, Sri Lanka.
Also at Isfahan University of Technology, Isfahan, Iran.
Also at Yazd University, Yazd, Iran.
Also at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran.
Also at Università degli Studi di Siena, Siena, Italy.
Also at Purdue University, West Lafayette, USA.
Also at International Islamic University of Malaysia, Kuala Lumpur, Malaysia.
Also at Malaysian Nuclear Agency, MOSTI, Kajang, Malaysia.
Also at Consejo Nacional de Ciencia y Tecnología, Mexico city, Mexico.
Also at Warsaw University of Technology, Institute of Electronic Systems, Warsaw, Poland.
Also at Institute for Nuclear Research, Moscow, Russia.
Also at St. Petersburg State Polytechnical University, St. Petersburg, Russia.
Also at University of Florida, Gainesville, USA.
Also at National Research Nuclear University ‘Moscow Engineering Physics Institute’ (MEPhI), Moscow, Russia.
Also at California Institute of Technology, Pasadena, USA.
Also at Budker Institute of Nuclear Physics, Novosibirsk, Russia.
Also at Faculty of Physics, University of Belgrade, Belgrade, Serbia.
Also at INFN Sezione di Roma, Sapienza Università di Roma, Rome, Italy.
Also at University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia.
Also at Scuola Normale e Sezione dell’INFN, Pisa, Italy.
Also at National and Kapodistrian University of Athens, Athens, Greece.
Also at Riga Technical University, Riga, Latvia.
Also at Institute for Theoretical and Experimental Physics, Moscow, Russia.
Also at Albert Einstein Center for Fundamental Physics, Bern, Switzerland.
Also at Gaziosmanpasa University, Tokat, Turkey.
Also at Istanbul Aydin University, Istanbul, Turkey.
Also at Mersin University, Mersin, Turkey.
Also at Cag University, Mersin, Turkey.
Also at Piri Reis University, Istanbul, Turkey.
Also at Adiyaman University, Adiyaman, Turkey.
Also at Ozyegin University, Istanbul, Turkey.
Also at Izmir Institute of Technology, Izmir, Turkey.
Also at Marmara University, Istanbul, Turkey.
Also at Kafkas University, Kars, Turkey.
Also at Istanbul Bilgi University, Istanbul, Turkey.
Also at Yildiz Technical University, Istanbul, Turkey.
Also at Hacettepe University, Ankara, Turkey.
Also at Rutherford Appleton Laboratory, Didcot, United Kingdom.
Also at School of Physics and Astronomy, University of Southampton, Southampton, United Kingdom.
Also at Instituto de Astrofísica de Canarias, La Laguna, Spain.
Also at Utah Valley University, Orem, USA.
Also at Beykent University.
Also at Erzincan University, Erzincan, Turkey.
Also at Mimar Sinan University, Istanbul, Istanbul, Turkey.
Also at Texas A&M University at Qatar, Doha, Qatar.
Also at Kyungpook National University, Daegu, Korea.