A search for pair production of first-generation scalar leptoquarks is performed in the final state containing two electrons and two jets using proton-proton collision data at \( \sqrt{s} = 7 \) TeV. The data sample used corresponds to an integrated luminosity of 33 pb\(^{-1}\) collected with the CMS detector at the CERN LHC. The number of observed events is in good agreement with the predictions for the standard model background processes, and an upper limit is set on the leptoquark pair production cross section times \( \beta^2 \) as a function of the leptoquark mass, where \( \beta \) is the branching fraction of the leptoquark decay to an electron and a quark. A 95% confidence level lower limit is set on the mass of a first-generation scalar leptoquark at 384 GeV for \( \beta = 1 \), which is the most stringent direct limit to date.

DOI: 10.1103/PhysRevLett.106.201802

PACS numbers: 14.80.Sv, 12.60.+i, 13.85.Rm

Although the standard model (SM) of fundamental particles and their interactions is in excellent agreement with most collider data, there are compelling reasons to believe new physics should appear at high energy scales. Some well-motivated theories of physics beyond the SM, including grand unified theories [1], composite models [2], technicolor [3–5], and superstring-inspired \( E_6 \) models [6], postulate the existence of a symmetry, beyond that of the SM, relating quarks and leptons and implying the existence of new bosons, called leptoquarks (LQ). An LQ carries color, has fractional electric charge, can have spin 0 (scalar) or spin 1 (vector), and couples to a lepton and a quark with coupling strength \( \lambda \). An LQ would decay to a charged lepton and a quark, with an unknown branching fraction \( \beta \), or a neutrino and a quark, with branching fraction \( 1 - \beta \). A review of LQ phenomenology and searches can be found in [7,8]. Constraints from experiments sensitive to flavour-changing neutral currents, lepton-family-number violation, and other rare processes [9] favor LQs that couple to quarks and leptons within the same SM generation, for LQ masses accessible at current colliders. The first-generation scalar LQs studied in this Letter couple only to an electron or an electron neutrino and a light quark. Measurements at electron-proton colliders constrain the coupling \( \lambda \) to be comparable to or less than the electromagnetic coupling \( \lambda_{\text{EM}} = \sqrt{4\pi \alpha_{\text{EM}}} = 0.3 \), for a first-generation LQ mass, \( M_{\text{LQ}} \), less than 300 GeV [10,11]. Prior to this work, the D0 Collaboration set the most stringent limit for a broad range of the coupling \( \lambda \) on the mass of the first-generation scalar LQ, namely, \( M_{\text{LQ}} > 299 \) GeV for \( \beta = 1 \) [12].

This Letter presents the results of a search for pair production of first-generation scalar LQs using events containing two electrons and two jets from a data sample of \( pp \) collisions at \( \sqrt{s} = 7 \) TeV collected in 2010 with the Compact Muon Solenoid (CMS) detector at the LHC. The data sample corresponds to an integrated luminosity of 33.2 ± 3.7 pb\(^{-1}\). In \( pp \) collisions at this energy, LQs are predominantly produced in pairs via gluon-gluon fusion and quark-antiquark annihilation with a cross section that depends on the strong coupling constant \( \alpha_s \) but is nearly independent of \( \lambda \). This cross section depends on the spin and the mass of the LQ and, for scalar LQs, has been calculated including next-to-leading-order (NLO) quantum chromodynamics (QCD) corrections [13]. In this study we did not consider possible contributions from single LQ production, which has a cross section that is dependent on \( \lambda \).

The CMS experiment [14] uses a right-handed coordinate system, with the origin at the nominal interaction point, the \( x \) axis pointing to the center of the LHC ring, the \( y \) axis pointing up (perpendicular to the LHC plane), and the \( z \) axis along the anticlockwise-beam direction. The polar angle, \( \theta \), is measured from the positive \( z \) axis and the azimuthal angle, \( \phi \), is measured in the \( x-y \) plane. The pseudorapidity is given by \( \eta = -\ln(\tan(\theta/2)) \). The central feature of the CMS apparatus is a superconducting solenoid, of 6 m internal diameter, providing a field of 3.8 T. Within the field volume are the silicon pixel and strip tracker, the crystal electromagnetic calorimeter (ECAL), which includes a silicon sensor preshower detector in front of the ECAL endcaps, and the brass-scintillator hadron calorimeter (HCAL). Muons are measured in gas-ionization detectors embedded in the steel return yoke. In addition to the barrel and endcap detectors, CMS has extensive forward calorimetry. The ECAL has an ultimate
energy resolution of better than 0.5% for unconverted photons with transverse energies above 100 GeV. The energy resolution is 3% or better for the range of electron energies relevant for this analysis. The HCAL, when combined with the ECAL, measures jets with a resolution $\Delta E/E \approx 100\%/\sqrt{E_{\text{GeV}} \cdot \text{GeV}} \oplus 5\%$. The inner tracker measures charged particles within an impact parameter resolution of about 15 $\mu$m and a transverse momentum ($p_T$) resolution of about 1.5% for 100 GeV particles. The relative luminosity is measured using the forward calorimeters. Collision events were selected by a first level trigger made of a system of fast electronics and a higher level trigger that consists of a farm of commercial CPUs running a version of the offline reconstruction optimized for fast processing.

Events used in this analysis are collected with an efficiency greater than 99.9% by single and double electron triggers with various thresholds depending on the instantaneous luminosity. Offline, events are required to contain at least one primary vertex with $z$ position within 24 cm of the nominal center of the detector. Electron candidates are required to have an electromagnetic cluster in ECAL that is spatially matched to a reconstructed track in the central tracking system in both $\eta$ and $\phi$. Electron candidates are further required to have a shower shape consistent with that of an electromagnetic shower, have a ratio between the hadronic and electromagnetic energy of less than 5%, and be isolated within a cone $\sqrt{\Delta \eta^2 + \Delta \phi^2} < 0.3$ from other energy deposits in the calorimeter and from additional reconstructed tracks (beyond the matched track) in the central tracking system. More information about electron triggering and identification at CMS can be found elsewhere [15]. Jets, the experimental signature of the hadronization of partons, are reconstructed in this analysis from calorimetry information by the anti-$k_T$ algorithm [16] with the distance parameter set to 0.5. The energy response of the jets is adjusted by applying a correction determined from Monte Carlo (MC) simulated events and a residual correction derived from data by analyzing the $p_T$ balance in di-jet events [17].

The collision data were compared to samples of MC generated events, where the response of the detector was simulated using GEANT4 [18]. The selection procedure as well as the electron and jet reconstructions described for the data are also applied to the MC simulation samples. Signal samples for LQ masses from 200 to 500 GeV were generated with PYTHIA [19], version 6.422, tune D6T [20,21].

An initial sample containing at least two electrons and at least two jets is selected. The dominant SM processes that produce such events are $Z/\gamma^* +$ jets and $t\bar{t}$, which are simulated, respectively, using ALPGEN [22] and MADGRAPH [23,24] interfaced with PYTHIA for parton showering and hadronization. Other backgrounds include multijet production with two jets misidentified as electrons and $W +$ jets events with one jet misidentified as an electron. There is also a small contribution from di-boson and single top production. The two leading (in $p_T$) electrons and two leading jets are used in the analysis and, to reduce the backgrounds, required to have $p_T > 30$ GeV. Selected electrons and jets have pseudorapidities $|\eta| < 2.5$ and $|\eta| < 3.0$, respectively, and if any of the selected electrons are closer than $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.7$ to any of the selected jets, the event is rejected. In addition, the preliminary requirements $M_{ee} > 50$ GeV and $S_T > 250$ GeV are applied, where $M_{ee}$ is the di-electron invariant mass and $S_T$ is defined as the sum of the magnitudes of the $p_T$ of the two leading electrons and two leading jets. At this stage of the selection, referred to as preselection, there are sufficient data to compare with the MC predictions. Good data-MC agreement is observed in the shape of all kinematic distributions of the selected electrons and jets. Figure 1 shows the $M_{ee}$ and $S_T$ distributions. The $Z/\gamma^* +$ jets MC distributions have been normalized to the data at the $Z$ boson mass, as described later.

To reduce the background from $Z \rightarrow ee$ production, a minimal value of $M_{ee}$ well above the mass of the $Z$ boson is required, and, to reduce all SM backgrounds, $S_T$ is required

![Figure 1](color online). Top: the $M_{ee}$ distribution for events that have passed the preselection requirements. Bottom: the $S_T$ distribution for events that have passed the preselection requirement, except the preselection requirement on $S_T$ itself ($S_T > 250$ GeV), and have $M_{ee} > 125$ GeV. The MC distributions for the signal ($\beta = 1$) and the contributing backgrounds are shown. The $Z/\gamma^* +$ jets MC has been normalized as described in the text. Other backgrounds include $W +$ jets, di-boson, and single top. All background histograms are cumulative.
to be large. While the LQ signal is expected to appear as a peak in the mass distribution of the electron-jet pairs, we find that the $S_T$ variable is more powerful with the present statistics as it is not affected by combinatorics. The minimal values required for $M_{ee}$ and $S_T$ were optimized by minimizing the expected upper limit on the leptoquark cross section in the absence of an observed signal using a Bayesian approach [8,25] that is well suited for counting experiments in the Poisson regime. The optimized lower value of $M_{ee}$ is found to be 125 GeV for all the LQ hypotheses under test, while the lower value of $S_T$ varies as indicated in Table I. Table I shows the number of surviving events for MC signal, MC background, and data samples after applying the full optimized selection. The reported product of signal selection efficiency and acceptance is estimated from MC simulated events. The product of the di-electron efficiency and acceptance, prior to any $M_{ee}$ and jet requirements, varies from 58.7% to 68.0% for LQ masses from 200 to 500 GeV.

The $Z/\gamma^* + \text{jets}$ background dominates the preselection sample. After the preselection, the ratio between data and MC events with $80 < M_{ee} < 100$ GeV (where the contamination from other SM processes is 3%) is $1.20 \pm 0.14$. This ratio is used to normalize the $Z/\gamma^* + \text{jets}$ MC events. The statistical uncertainty on this normalization factor is used as an uncertainty on the MC estimate of the $Z/\gamma^* + \text{jets}$ background after the full selection. In addition, a systematic uncertainty of 20% due to the modeling of the shape of this background is determined by comparing the number of $Z/\gamma^* + \text{jets}$ events surviving final $S_T$ cut selections in MADGRAPH samples with the renormalization or factorization scales and matching thresholds varied by a factor of 2. The $t\bar{t}$ background is estimated from MC calculations normalized to the CMS measurement of the $t\bar{t}$ cross section [26]. An uncertainty, 41%, is also taken from the same measurement. The small contribution from other background processes containing vector bosons is estimated by MC calculations. The multijet background is determined

<table>
<thead>
<tr>
<th>$M_{LQ}$</th>
<th>Signal samples (MC)</th>
<th>Standard model background samples (MC)</th>
<th>Events in data</th>
<th>Obs./Exp. 95% C.L. u.l. on $\sigma$ [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{LQ}$ (ST Cut) [GeV]</td>
<td>Selected Events</td>
<td>Acceptance $\times$ Efficiency</td>
<td>$t\bar{t}$ + jets</td>
<td>$Z/\gamma^* + \text{jets}$</td>
</tr>
<tr>
<td>200 ($S_T &gt; 340$)</td>
<td>117.5 $\pm$ 0.8</td>
<td>0.297 $\pm$ 0.002</td>
<td>2.6 $\pm$ 0.1</td>
<td>2.0 $\pm$ 0.2</td>
</tr>
<tr>
<td>250 ($S_T &gt; 400$)</td>
<td>43.8 $\pm$ 0.2</td>
<td>0.380 $\pm$ 0.002</td>
<td>1.3 $\pm$ 0.1</td>
<td>1.3 $\pm$ 0.1</td>
</tr>
<tr>
<td>280 ($S_T &gt; 450$)</td>
<td>24.4 $\pm$ 0.1</td>
<td>0.403 $\pm$ 0.002</td>
<td>0.69 $\pm$ 0.05</td>
<td>0.87 $\pm$ 0.07</td>
</tr>
<tr>
<td>300 ($S_T &gt; 470$)</td>
<td>17.3 $\pm$ 0.09</td>
<td>0.430 $\pm$ 0.002</td>
<td>0.52 $\pm$ 0.05</td>
<td>0.75 $\pm$ 0.07</td>
</tr>
<tr>
<td>320 ($S_T &gt; 490$)</td>
<td>12.3 $\pm$ 0.06</td>
<td>0.451 $\pm$ 0.002</td>
<td>0.43 $\pm$ 0.04</td>
<td>0.65 $\pm$ 0.07</td>
</tr>
<tr>
<td>340 ($S_T &gt; 510$)</td>
<td>8.88 $\pm$ 0.04</td>
<td>0.469 $\pm$ 0.002</td>
<td>0.32 $\pm$ 0.04</td>
<td>0.56 $\pm$ 0.06</td>
</tr>
<tr>
<td>370 ($S_T &gt; 540$)</td>
<td>5.55 $\pm$ 0.02</td>
<td>0.496 $\pm$ 0.002</td>
<td>0.26 $\pm$ 0.03</td>
<td>0.47 $\pm$ 0.06</td>
</tr>
<tr>
<td>400 ($S_T &gt; 560$)</td>
<td>3.55 $\pm$ 0.02</td>
<td>0.522 $\pm$ 0.002</td>
<td>0.20 $\pm$ 0.03</td>
<td>0.41 $\pm$ 0.05</td>
</tr>
<tr>
<td>450 ($S_T &gt; 620$)</td>
<td>1.70 $\pm$ 0.01</td>
<td>0.539 $\pm$ 0.002</td>
<td>0.12 $\pm$ 0.02</td>
<td>0.28 $\pm$ 0.05</td>
</tr>
<tr>
<td>500 ($S_T &gt; 660$)</td>
<td>0.868 $\pm$ 0.003</td>
<td>0.565 $\pm$ 0.002</td>
<td>0.08 $\pm$ 0.02</td>
<td>0.23 $\pm$ 0.05</td>
</tr>
</tbody>
</table>

TABLE II. Summary of the systematic uncertainties affecting the number of signal and background events for the hypothesis of LQ with mass of 300 GeV.

<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>Magnitude [%]</th>
<th>Effect on $N_{signal}$ [%]</th>
<th>Effect on $N_{all}$ [kg] [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data-driven uncertainty</td>
<td>⋯</td>
<td>⋯</td>
<td>22</td>
</tr>
<tr>
<td>$Z/\gamma^* + \text{jets}$ Background Shape</td>
<td>20</td>
<td>⋯</td>
<td>11</td>
</tr>
<tr>
<td>Jet Energy Scale</td>
<td>5</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Elec. Energy Scale Barrel/Endcap</td>
<td>1/3</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Electron Pair Reco/ID/Isol</td>
<td>10</td>
<td>10</td>
<td>⋯</td>
</tr>
<tr>
<td>MC Statistics</td>
<td>⋯</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Integrated Luminosity</td>
<td>11</td>
<td>11</td>
<td>⋯</td>
</tr>
<tr>
<td>Total</td>
<td>⋯</td>
<td>15</td>
<td>28</td>
</tr>
</tbody>
</table>
from data. The probability that an isolated electromagnetic cluster is reconstructed as an electron is measured in a background sample requiring a single cluster, a jet multiplicity similar to the analysis final state, and small missing transverse energy. This probability and a data sample with two or more of these clusters and two or more jets were used to determine the multijet contribution to the final selection sample. The resulting systematic uncertainty is determined to be 20%. This background accounts for <1% of the total background for all LQ masses hypotheses with a decreasing trend for increasing LQ mass hypothesis, and is not considered any further.

The systematic uncertainties affecting the number of expected signal and background events are summarized in Table II. The jet and electron energy scale uncertainties are given in the second column of Table II. The reconstruction, identification, and isolation efficiency for electrons is determined from MC simulated events and a systematic uncertainty is assessed using $Z \rightarrow \ell\ell$ events from collision data [27]. The statistical uncertainty on the number of MC events surviving the full event selection is reported in Table I for the signal and background. The uncertainty of the integrated luminosity of the data sample is dominated by the uncertainty derived from data. Figure 2 (top) shows the 95% confidence level (C.L.) upper limit on the LQ pair production cross section times $\beta^2$ as a function of the LQ mass. The systematic uncertainties reported in Table II are included in the calculation. The shaded region is excluded by the current D0 limit for $\beta = 1$. The $\sigma_{\text{theory}}$ curve and its band represent, respectively, the theoretical LQ pair production cross section and the uncertainties due to the choice of PDF renormalization-factorization scales [13].

The number of observed events in the collision data sample that pass the selection criteria optimized for each LQ mass considered is consistent with the prediction from SM processes, as reported in Table I. An upper limit on the LQ cross section in the absence of signal is therefore set using a Bayesian approach [25] that uses a Poisson likelihood, a flat prior for the signal cross section, and log-normal priors for the parameters used to model the systematic uncertainties. Systematic uncertainties for the signal are dominated by the integrated luminosity and the electron selection efficiencies, while the systematic uncertainties for the background are dominated by the uncertainty derived from data. Figure 2 (top) shows the 95% confidence level (C.L.) upper limit on the LQ pair production cross section times $\beta^2$ as a function of the leptoquark mass for 33.2 pb$^{-1}$ of integrated luminosity. The systematic uncertainties reported in Table II are included in the calculation. The upper limits are compared to an NLO prediction of the LQ pair production cross section [13] to set a 95% C.L. exclusion on LQ masses smaller than 384 GeV (expected 391 GeV), assuming $\beta = 1$. A theoretical uncertainty on the signal production cross sections due to the choice of renormalization/factorization scales (14%–15% for all LQ masses considered) has been

![FIG. 2 (color online). Top: the expected and observed upper limit at 95% C.L. on the LQ pair production cross section times $\beta^2$ as a function of the LQ mass. The systematic uncertainties reported in Table II are included in the calculation. The shaded region is excluded by the current D0 limit for $\beta = 1$. The $\sigma_{\text{theory}}$ curve and its band represent, respectively, the theoretical LQ pair production cross section and the uncertainties due to the choice of PDF renormalization-factorization scales [13]. Bottom: minimum $\beta$ for a 95% C.L. exclusion of the LQ hypothesis as a function of LQ mass. The observed (expected) exclusion curve is obtained using the observed (expected) upper limit and the central value of the theoretical LQ pair production cross section. The band around the observed exclusion curve is obtained by considering the observed upper limit while taking into account the uncertainties on the theoretical cross section. The shaded region is excluded by the current D0 limits, which combines results from searches in the two electron, electron-neutrino, and two neutrino channels.](201802-4)
calculated by varying the scales between half and twice the LQ mass, while a 90% C.L. PDF uncertainty (from 8 to 22% for LQ masses from 200 to 500 GeV) has been obtained from the CTEQ6.6 error PDF set following the standard prescription detailed in Ref. [29]. If the observed cross section upper limit is compared with the lower boundary of the cross section uncertainty band, the lower limit on the LQ mass for $\beta = 1$ becomes 370 GeV (expected 375 GeV). Figure 2 (bottom) shows the minimum $\beta$ for a 95% C.L. exclusion of the LQ hypothesis as a function of LQ mass.

In conclusion, a search for pair production of first-generation scalar leptoquarks has been presented. The number of collision events, passing a selection optimized for exclusion of the LQ hypothesis, is in good agreement with the predictions for the SM background processes. A Bayesian approach that includes the treatment of the systematic uncertainties as nuisance parameters has been used to set an upper limit on the LQ cross section. By comparing this upper limit to a theoretical calculation of the LQ pair production cross section, the existence of first-generation scalar LQ with masses below 384 GeV has been excluded at 95% C.L., with a corresponding cross section limit of 0.265 pb. The lower limits on the LQ mass set for values of $\beta$ larger than about 0.4 are the most restrictive direct limits to date.

We wish to thank Michael Krämer for providing the NLO LQ pair production cross sections at $\sqrt{s} = 7 \text{ TeV}$. We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC machine. We thank the technical and administrative staff at CERN and other CMS institutes, and acknowledge support from: FMSR (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COCIB, (Colombia); MSES (Croatia); RPF (Cyprus); Academy of Sciences and NICPB (Estonia); Academy of Finland, ME, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NKTH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); PAEC (Pakistan); SCSR (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine, Uzbekistan); MST and MAE (Russia); MSTD (Serbia); MICINN and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); TUBITAK and TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA).

aDeceased.
bAlso at CERN, European Organization for Nuclear Research, Geneva, Switzerland.
cAlso at Universidade Federal do ABC, Santo Andre, Brazil.
dAlso at Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France.
eAlso at Suez Canal University, Suez, Egypt.
fAlso at Fayoum University, El-Fayoum, Egypt.
gAlso at Soltan Institute for Nuclear Studies, Warsaw, Poland.
hAlso at Massachusetts Institute of Technology, Cambridge, MA, USA.
iAlso at Université de Haute-Alsace, Mulhouse, France.
jAlso at Brandenburg University of Technology, Cottbus, Germany.
kAlso at Moscow State University, Moscow, Russia.
lAlso at Institute of Nuclear Research ATOMKI, Debrecen, Hungary.
mAlso at Eötvös Loránd University, Budapest, Hungary.
nAlso at Tata Institute of Fundamental Research - HECR, Mumbai, India.
oAlso at University of Visva-Bharati, Santiniketan, India.
pAlso at Facoltà Ingegneria Università di Roma “La Sapienza”, Roma, Italy.
qAlso at Università della Basilicata, Potenza, Italy.
rAlso at California Institute of Technology, Pasadena, USA.
sAlso at Faculty of Physics of University of Belgrade, Belgrade, Serbia.
tAlso at University of California, Los Angeles, Los Angeles, USA.
uAlso at University of Florida, Gainesville, USA.
vAlso at Université de Genève, Geneva, Switzerland.
wAlso at Scuola Normale e Sezione dell’ INFN, Pisa, Italy.
xAlso at INFN Sezione di Roma, Università di Roma “La Sapienza”, Roma, Italy.
yAlso at University of Athens, Athens, Greece.
zAlso at The University of Kansas, Lawrence, USA.
aaAlso at Institute for Theoretical and Experimental Physics, Moscow, Russia.
bbAlso at Paul Scherrer Institut, Villigen, Switzerland.
cAlso at University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia.
ddAlso at Gaziosmanpasa University, Tokat, Turkey.
eeAlso at Adiyaman University, Adiyaman, Turkey.
ffAlso at Mersin University, Mersin, Turkey.
ggAlso at Izmir Institute of Technology, Izmir, Turkey.
hhAlso at Kafkas University, Kars, Turkey.
iiAlso at Suleyman Demirel University, Isparta, Turkey.
jjAlso at Ege University, Izmir, Turkey.
kkAlso at Rutherford Appleton Laboratory, Didcot, United Kingdom.
llAlso at School of Physics and Astronomy, University of Southampton, Southampton, United Kingdom.
mmAlso at INFN Sezione di Perugia, Università di Perugia, Perugia, Italy.
nAlso at KFKI Research Institute for Particle and Nuclear Physics, Budapest, Hungary.
ooAlso at Institute for Nuclear Research, Moscow, Russia.
ppAlso at Horia Hulubei National Institute of Physics and Nuclear Engineering (IFIN-HH), Bucharest, Romania.