Results are presented from a search for the pair production of first- and second-generation scalar leptoquarks in proton-proton collisions at $\sqrt{s} = 7$ TeV. The data sample corresponds to an integrated luminosity of 5.0 fb$^{-1}$, collected by the CMS detector at the LHC. The search signatures involve either two charged leptons of the same flavor (electrons or muons) and at least two jets or a single charged lepton (electron or muon), missing transverse energy, and at least two jets. If the branching fraction of the leptoquark decay into a charged lepton and a quark is assumed to be $\beta = 1$, leptoquark pair production is excluded at the 95% confidence level for masses below 830 GeV and 840 GeV for the first and second generations, respectively. For $\beta = 0.5$, masses below 640 GeV and 650 GeV are excluded. These limits are the most stringent to date.

I. INTRODUCTION

The structure of the standard model (SM) of particle physics suggests a fundamental relationship between quarks and leptons. There are many models beyond the SM that predict the existence of leptoquarks (LQ), hypothetical particles that carry both baryon number and lepton number and couple to both quarks and leptons. Among these scenarios are grand unified theories [1,2], composite models [3], extended technicolor models [4–6], and superstring-inspired models [7]. Leptoquarks are color triplets with fractional electric charge and can be either scalar or vector particles. A leptoquark couples to a lepton and a quark with a coupling strength $\lambda$, and it decays to a charged lepton and a quark with an unknown branching fraction $\beta$ or to a neutrino and a quark with branching fraction $1 - \beta$. To satisfy constraints from bounds on flavor-changing neutral currents and from rare pion and kaon decays [3,8], it is assumed that leptoquarks couple to quarks and leptons of a single generation. Leptoquarks are classified as first, second, or third generation, depending on the generation of leptons to which they couple. The dominant mechanisms for the production of leptoquark pairs at the Large Hadron Collider (LHC) are gluon-gluon ($gg$) fusion and quark-antiquark ($q\bar{q}$) annihilation, shown in Fig. 1. The dominant processes depend only on the strong coupling constant and have been calculated at next-to-leading order (NLO) [9]. The cross section for production via the unknown Yukawa coupling $\lambda$ of a leptoquark to a lepton and a quark is typically smaller.

This paper reports on a search for pair production of scalar leptoquarks. Several experiments have searched for pair-produced scalar leptoquarks, but none has obtained evidence for them [10–15]. This search uses a data sample corresponding to an integrated luminosity of 5.0 fb$^{-1}$ recorded with the Compact Muon Solenoid (CMS) detector during the 2011 proton-proton run of the LHC at $\sqrt{s} = 7$ TeV. The analysis performed in this paper considers the decay of leptoquark pairs into two charged leptons of the same flavor (either electrons or muons) and two quarks; or into a charged lepton, a neutrino, and two quarks. As a result, two distinct classes of events are selected: one with two high-transverse-momentum ($p_T$) electrons or muons and at least two high-$p_T$ jets ($\ell\ell jj$) and the other with one high-$p_T$ electron or muon, large missing transverse energy ($E_T^{\text{miss}}$), and at least two high-$p_T$ jets ($\ell\nu jj$).

The CMS detector, described in detail elsewhere [16], uses a cylindrical coordinate system with the $z$ axis along the counterclockwise beam axis. The detector consists of an inner tracking system and electromagnetic (ECAL) and hadron (HCAL) calorimeters surrounded by a 3.8 T solenoid. The inner tracking system consists of a silicon pixel...
and strip tracker, providing the required granularity and precision for the reconstruction of vertices of charged particles in the range $0 \leq \phi \leq 2\pi$ in azimuth and $|\eta| < 2.5$, where the pseudorapidity $\eta$ is defined as $\eta = -\ln(\tan(\theta/2))$, and $\theta$ is the polar angle measured with respect to the $z$ axis. The crystal ECAL and the brass/scintillator sampling HCAL are used to measure with high resolution the energies of photons, electrons, and hadrons for $|\eta| < 3.0$. The three muon systems surrounding the solenoid cover a region $|\eta| < 2.4$ and are composed of drift tubes in the barrel region ($|\eta| < 1.2$), of cathode strip chambers in the endcaps (0.9 $|\eta| < 2.4$), and of resistive plate chambers in both the barrel region and the end caps ($|\eta| < 1.6$). Events are recorded based on a trigger decision using information from either the calorimeter or muon systems. The final trigger decision is based on the information from all subsystems, which is passed on to the high-level trigger (HLT), consisting of a farm of computers running a version of the reconstruction software optimized for fast processing.

The $\ell\ell jj$ and $\ell\nu jj$ analyses are performed separately and the results are combined as a function of the branching fraction $\beta$ and the leptoquark mass $M_{LQ}$ for first and second generations independently. The analysis in all four decay channels searches for leptoquarks in an excess of events characteristic of the decay of heavy objects. Various triggers are used to collect events depending on the decay channel and the data taking periods as described in Sec. II. An initial selection isolates events with high-$p_T$ final-state particles (two or more isolated leptons and two or more jets; or one isolated lepton, two or more jets, and large $E_T^{miss}$ indicative of the emission of a neutrino). Kinematic variables are then identified to further separate a possible leptoquark signal from the expected backgrounds, and optimized thresholds on the values of these variables are derived to maximize the sensitivity to the possible presence of a signal in each decay mode. The variables used in the optimization are the invariant mass of jet-lepton pairs ($M_{jj}$), the scalar sum ($S_T$) of the $p_T$ of each of the final-state objects, and either the invariant mass of the dilepton pair ($M_{\ell\ell}$) in the $\ell\ell jj$ channels or $E_T^{miss}$ in the $\ell\nu jj$ channels.

Major sources of SM background are $Z/\gamma^* + $ jets, $W + $ jets processes, and $t\bar{t}$. Smaller contributions arise from single-top production, diboson processes, and QCD multijet processes. The major backgrounds are determined either from control samples in data or from Monte Carlo (MC) simulated samples normalized to data in selected control regions.

After final selection, the data are well described by the SM background predictions, and upper limits on the leptoquark pair-production cross section are set using a CL$_S$ modified frequentist approach [17,18]. Using Poisson statistics, 95% confidence level (CL) upper limits are obtained on the leptoquark pair-production cross section times branching fraction as a function of leptoquark mass ($M_{LQ}$). This is compared with the NLO predictions [9] to determine lower limits on $M_{LQ}$ for $\beta = 1$ and $\beta = 0.5$. The $\ell\ell jj$ and $\ell\nu jj$ channels are combined to further maximize the exclusion in $\beta$ and $M_{LQ}$, especially for the case $\beta \sim 0.5$, where combining the two channels increases the sensitivity of the search.

The data and the initial event selection are detailed in Sec. II of this paper, followed by a description of the signal modeling and background estimates in Secs. III and IV, respectively. Section V contains the final event selection, and Sec. VI describes the systematic uncertainties. The results of the search are presented in Sec. VII and summarized in Sec. VIII.

II. DATA SET AND OBJECT RECONSTRUCTION

For the first-generation $eejj$ analysis, events are required to pass a double-electron trigger or a double-photon trigger, with an electron or photon $p_T > 33$ GeV. For the first-generation $evjj$ analysis, events are required to pass either a single-electron trigger or a trigger based on the requirement of one electron with $p_T$ threshold between 17 and 30 GeV, missing transverse energy threshold between 15 and 20 GeV, and two jets with $p_T$ threshold between 25 and 30 GeV. The trigger thresholds vary according to the run period. For the second-generation leptoquark analyses, events are required to pass a single-muon trigger without isolation requirements and with a $p_T$ threshold of 40 GeV. For the $eejj$ channel, the trigger efficiency is greater than 99%. For the $evjj$ channel, the electron trigger efficiency is measured to be 95%. For the $\mu\mu jj$ and $\mu\nu jj$ channels, the single-muon trigger efficiency is measured to be 92% per muon.

Electron candidates [19] are required to have an electromagnetic cluster with $p_T > 40$ GeV and pseudorapidity $|\eta| < 2.5(2.2)$ for the $eejj$ ($evjj$) analysis, excluding the transition region between the barrel and the end cap detectors, 1.44 $|\eta| < 1.57$. The $evjj$ analysis requires lower electron $|\eta|$ to reduce the QCD multijet background, with negligible reduction of the signal acceptance. Electron candidates are required to have an electromagnetic cluster in the ECAL that is spatially matched to a reconstructed track in the central tracking system in both $\eta$ and the azimuthal angle $\phi$, and to have a shower shape consistent with that of an electromagnetic shower. Electron candidates are further required to be isolated from additional energy deposits in the calorimeter and from reconstructed tracks beyond the matched track in the central tracking system. In addition, to reject electrons coming from photon conversion in the tracker material, the track associated with the reconstructed electron is required to have hits in all inner tracker layers.

Muons are reconstructed as tracks in the muon system that are matched to the tracks reconstructed in the inner tracking system [20]. Muons are required to have

052013-2
**III. SIGNAL AND BACKGROUND MODELING**

The MC samples for the signal processes are generated for a range of leptoquark mass hypotheses between 250 and 900 GeV, with a renormalization and factorization scale $\mu = M_{LQ}$. The MC generation uses the PYTHIA generator [24] (version 6.422) and CTEQ6L1 parton distribution functions (PDF) [25]. The MC samples used to estimate the contribution from SM background processes are $t\bar{t}$ + jets events, generated with MADGRAPH [26]; single-top events ($s$, $t$, and $tW$ channels), generated with POWHEG [27]; $Z/\gamma^*$ + jets events and $W +$ jets events, generated with SHERPA [28]; $VV$ events, where $V$ represents either a $W$ or a $Z$ boson, generated with PYTHIA; QCD muon-enriched multijet events, generated with PYTHIA in bins of transverse momentum of the hard-scattering process from 15 GeV to the kinematic limit. The simulation of the CMS detector is based on GEANT4 [29] and includes multiple collisions in a single bunch crossing corresponding to the luminosity profile of the LHC during the data taking periods of interest.

**IV. BACKGROUND ESTIMATE**

The main processes that can mimic the signature of a leptoquark signal in the $\ell\ell jj$ channels are $Z/\gamma^*$ + jets, $t\bar{t}$, $V+$ jets, $W+$ jets, and QCD multijets. The $Z/\gamma^*$ + jets background is determined by comparing events from data and MC samples in two different regions: in the region of low ($L$) dilepton invariant mass around the $Z$ boson mass (70 $< M_{\ell\ell} < 100$ GeV for electrons and 80 $< M_{\ell\ell} < 100$ GeV for muons) and in the region of high ($H$) mass $M_{\ell\ell} > 100$ GeV. The low mass scaling factor $R_Z = N_L / N_L^{MC}$ is measured to be $1.27 \pm 0.02$ for the $eejj$ channel and $1.29 \pm 0.02$ for the $\mu\mu jj$ channel, where $N_L$ and $N_L^{MC}$ are the number of data and MC events, respectively, in the $Z$ mass window. The number of $Z/\gamma^*$ + jets events above 100 GeV is then estimated as

$$N_H = R_Z N_H^{MC},$$ (1)
where \(N_{MC}^{M''}\) is the number of MC events with \(M'' > 100\) GeV. The estimated number of \(Z/\gamma^* + \) jets events is obtained with the selection criteria optimized for different leptoquark mass hypotheses, and it is used in the limit setting procedure.

The number and kinematic distributions for the \(t\bar{t}\) events with two leptons of the same flavor are estimated from the number of data events that contain one electron and one muon. This type of background is expected to produce the \(ee\) final state or the \(\mu\mu\) final state with half the probability of the \(e\mu\) final state. In the data the number of \(ee\) or \(\mu\mu\) events is estimated to be

\[
N_{ee(\mu\mu)} = \frac{1}{2} \frac{e_\mu(e_\mu)}{e_\mu(e_\mu)}/N_{e\mu},
\]

where \(e_\mu\) and \(e_e\) are the muon and electron reconstruction and identification efficiencies and \(e^{\text{trig}}\) are the HLT efficiencies to select \(ee\), \(\mu\mu\), and \(e\mu\) events.

### TABLE III. Optimized thresholds for different mass hypotheses of the \(\ell\ell jj\) signal.

<table>
<thead>
<tr>
<th>(M_{LQ}) [GeV]</th>
<th>250</th>
<th>350</th>
<th>400</th>
<th>450</th>
<th>500</th>
<th>550</th>
<th>600</th>
<th>650</th>
<th>750</th>
<th>850</th>
<th>900</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S'' &gt; ) [GeV]</td>
<td>450</td>
<td>570</td>
<td>650</td>
<td>700</td>
<td>800</td>
<td>850</td>
<td>900</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>(E^{\text{miss}}_{T} &gt; ) [GeV]</td>
<td>100</td>
<td>120</td>
<td>120</td>
<td>140</td>
<td>160</td>
<td>160</td>
<td>180</td>
<td>180</td>
<td>220</td>
<td>240</td>
<td></td>
</tr>
<tr>
<td>(M'' &gt; ) [GeV]</td>
<td>100</td>
<td>110</td>
<td>120</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>150</td>
<td></td>
</tr>
</tbody>
</table>

The number and kinematic distributions for the \(t\bar{t}\) events with two leptons of the same flavor are estimated from the number of data events that contain one electron and one muon. This type of background is expected to produce the \(ee\) final state or the \(\mu\mu\) final state with half the probability of the \(e\mu\) final state. In the data the number of \(ee\) or \(\mu\mu\) events is estimated to be

\[
N_{ee(\mu\mu)} = \frac{1}{2} \frac{e_\mu(e_\mu)}{e_\mu(e_\mu)}/N_{e\mu},
\]

where \(e_\mu\) and \(e_e\) are the muon and electron reconstruction and identification efficiencies and \(e^{\text{trig}}\) are the HLT efficiencies to select \(ee\), \(\mu\mu\), and \(e\mu\) events.

### FIG. 2 (color online). \(eejj\) channel: the distributions of \(S'^{\ell}_{T}\) (top) and of \(M_{ej}\) for each of the two electron-jet pairs (bottom) for events that pass the initial selection level. The data are indicated by the points, and the SM backgrounds are given as cumulative histograms. The expected contribution from a leptoquark signal with \(M_{LQ} = 400\) GeV is also shown.

### FIG. 3 (color online). \(evjj\) channel: the distributions of \(S'^{\ell}_{T}\) (top) and of \(M_{ej}\) (bottom) for events that pass the initial selection level. The data are indicated by the points, and the SM backgrounds are given as cumulative histograms. The expected contribution from a leptoquark signal with \(M_{LQ} = 400\) GeV is also shown.

---

**S. CHATRCHYAN et al.**

**PHYSICAL REVIEW D** **86**, 052013 (2012)
the above probability to a sample with two electron candidates, two or more jets, and low $E_{T\text{miss}}$. The expected contribution from a leptoquark signal with $M_{LQ} = 400$ GeV is also shown.

No QCD multijet MC events pass the $\mu\mu jj$ final selection. A cross-check made using a data control sample containing same-sign muons confirms that the QCD multijet background is negligible in this channel.

For the first-generation leptoquark analyses the multijet background contribution is estimated from a data control sample as follows. The probability that an electron candidate passing loose electron requirements additionally passes all electron requirements is measured as a function of $p_T$ and $\eta$ in a data sample with one and only one electron candidate, two or more jets, and low $E_{T\text{miss}}$. This sample is dominated by QCD multijet events and is similar in terms of jet activity to the $eejj$ and $evjj$ analysis samples. A correction for a small contamination of genuine electrons passing all electron requirements is derived from MC simulations. The QCD multijet background in the final $eejj$ ($evjj$) selection is predicted by applying twice (once) the above probability to a sample with two electron candidates (one electron candidate and large $E_{T\text{miss}}$), and two or more jets, which satisfy all the requirements of the signal selections. The resulting estimate is $\sim1\%$ ($\sim8\%$) of the total background for the selections corresponding to the region of leptoquark masses where the exclusion limits are placed.

Contributions to the $\ell\ell jj$ background from $VV + \text{jets}$ processes and single-top production are small, and they are estimated using MC simulation.

In the $\nu jj$ channel, the main backgrounds come from three sources: processes that lead to the production of a genuine $W$ boson such as $W + \text{jets}$, $t\bar{t}$, single-top production, diboson processes ($WW$, $WZ$); instrumental background, mostly caused by the misidentification of jets as leptons in multijet processes, thus creating misidentified electrons or muons and misreconstructed $E_{T\text{miss}}$ in the final state; and $Z$ boson production, such as $Z/\gamma^* + \text{jets}$ and $ZZ$ processes. The contribution from the principal backgrounds, $W + \text{jets}$ and $t\bar{t}$, is estimated with MC simulation normalized to data at the initial selection level in the region $50 < M_T < 110$ GeV, where $M_T$ is the transverse mass calculated from the lepton $p_T$ and $E_{T\text{miss}}$.

The region $50 < M_T < 110$ GeV is used to determine both the $W + \text{jets}$ and the $t\bar{t}$ normalization factors using two mutually exclusive selections (less than four jets or at least four jets with $p_T > 40$ GeV and $|\eta| < 2.4$) that separately enhance the samples with $W + \text{jets}$ and with $t\bar{t}$ events. The results of these two selections are used to form a system of equations:
\[N_1 = R_{\text{id}}N_{1,\text{id}} + R_{\text{w}}N_{1,\text{w}} + N_{1,\text{QCD}} + N_{1,\text{O}};\]
\[N_2 = R_{\text{id}}N_{2,\text{id}} + R_{\text{w}}N_{2,\text{w}} + N_{2,\text{QCD}} + N_{2,\text{O}},\]

where \(N_i, N_{i,\text{w}}, N_{i,\text{id}}, N_{i,\text{O}},\) and \(N_{i,\text{QCD}}\) are the number of events in data, \(W + \ell\ell\) MC simulation, \(t\bar{t}\) MC simulation, other MC-simulated backgrounds, and QCD multijet events obtained from data, passing selection \(i\). In the \(\mu\nu jj\) analysis, the solution of the system yields the following normalization factors: \(R_{\text{id}} = 0.82 \pm 0.04\) (stat) \(\pm 0.02\) (syst), where the systematic errors reflect the uncertainties on the QCD multijet background estimate. In the \(\mu\nu jj\) analysis, the normalization factors are \(R_{\text{id}} = 0.84 \pm 0.03\) (stat) and \(R_{\text{w}} = 1.24 \pm 0.02\) (stat).

The contribution from QCD multijet processes after all selection criteria are applied to the \(\mu\nu jj\) sample is estimated to be negligible. No QCD MC events survive the full selection criteria optimized for any leptoquark mass hypothesis.

### TABLE IV. Individual background (BG) sources, expected signal, data, and total background event yields after the initial (first row) and final selections for the \(eejj\) analysis. Other BG includes single top, \(W + \ell\ell\), \(\gamma + \ell\ell\), and \(VV + \ell\ell\). Only statistical uncertainties are reported.

<table>
<thead>
<tr>
<th>(M_{\ell Q})</th>
<th>(Z + \text{ jets})</th>
<th>(t\bar{t})</th>
<th>QCD</th>
<th>Other BG</th>
<th>LQ signal</th>
<th>Data</th>
<th>Total BG</th>
</tr>
</thead>
<tbody>
<tr>
<td>⋯</td>
<td>6234 ± 24</td>
<td>768 ± 19</td>
<td>49.59 ± 0.43</td>
<td>147.6 ± 2.3</td>
<td>⋯</td>
<td>7201</td>
<td>7199 ± 31</td>
</tr>
<tr>
<td>400</td>
<td>35.7 ± 1.8</td>
<td>19.1 ± 3.1</td>
<td>0.877 ± 0.022</td>
<td>3.12 ± 0.56</td>
<td>487.4 ± 2.2</td>
<td>55</td>
<td>58.8 ± 3.6</td>
</tr>
<tr>
<td>500</td>
<td>6.55 ± 0.70</td>
<td>2.45 ± 1.10</td>
<td>0.192 ± 0.012</td>
<td>1.03 ± 0.42</td>
<td>109.30 ± 0.46</td>
<td>14</td>
<td>10.2 ± 1.4</td>
</tr>
<tr>
<td>550</td>
<td>4.65 ± 0.58</td>
<td>0.98 ± 0.69</td>
<td>0.139 ± 0.012</td>
<td>0.84 ± 0.42</td>
<td>57.35 ± 0.23</td>
<td>11</td>
<td>6.60 ± 0.99</td>
</tr>
<tr>
<td>600</td>
<td>3.04 ± 0.46</td>
<td>0.49 ± 0.49</td>
<td>0.088 ± 0.011</td>
<td>0.72 ± 0.41</td>
<td>30.95 ± 0.14</td>
<td>8</td>
<td>4.34 ± 0.79</td>
</tr>
<tr>
<td>650</td>
<td>2.14 ± 0.38</td>
<td>0.49 ± 0.49</td>
<td>0.073 ± 0.011</td>
<td>0.48 ± 0.40</td>
<td>16.998 ± 0.065</td>
<td>6</td>
<td>3.18 ± 0.74</td>
</tr>
<tr>
<td>750</td>
<td>1.04 ± 0.26</td>
<td>0.000 ± 0.56 / 0.020</td>
<td>0.0092 ± 0.00022</td>
<td>0.41 ± 0.40</td>
<td>5.526 ± 0.023</td>
<td>0</td>
<td>1.45 ± 0.47</td>
</tr>
<tr>
<td>850</td>
<td>0.81 ± 0.23</td>
<td>0.000 ± 0.56 / 0.000101</td>
<td>0.00101 ± 0.00022</td>
<td>0.40 ± 0.40</td>
<td>1.9679 ± 0.0078</td>
<td>0</td>
<td>1.21 ± 0.46</td>
</tr>
</tbody>
</table>

### TABLE V. Individual background (BG) sources, expected signal, data, and total background event yields after the initial (first row) and final selections for the \(eejj\) analysis. Other BG includes single top, \(Z + \text{ jets}\), \(\gamma + \ell\ell\), and \(VV + \ell\ell\). Only statistical uncertainties are reported.

<table>
<thead>
<tr>
<th>(M_{\ell Q})</th>
<th>(W + \text{ jets})</th>
<th>(t\bar{t})</th>
<th>QCD</th>
<th>Other</th>
<th>LQ signal</th>
<th>Data</th>
<th>Total BG</th>
</tr>
</thead>
<tbody>
<tr>
<td>⋯</td>
<td>20108 ± 99</td>
<td>9301 ± 42</td>
<td>3267 ± 26</td>
<td>1913 ± 53</td>
<td>⋯</td>
<td>34135</td>
<td>34590 ± 120</td>
</tr>
<tr>
<td>400</td>
<td>28.7 ± 3.6</td>
<td>17.5 ± 1.8</td>
<td>6.20 ± 0.46</td>
<td>6.01 ± 0.77</td>
<td>126.01 ± 0.82</td>
<td>43</td>
<td>58.4 ± 4.1</td>
</tr>
<tr>
<td>500</td>
<td>13.3 ± 2.4</td>
<td>6.3 ± 1.1</td>
<td>1.72 ± 0.22</td>
<td>2.80 ± 0.37</td>
<td>34.70 ± 0.23</td>
<td>18</td>
<td>24.2 ± 2.6</td>
</tr>
<tr>
<td>550</td>
<td>2.98 ± 0.95</td>
<td>3.38 ± 0.82</td>
<td>0.65 ± 0.10</td>
<td>1.46 ± 0.26</td>
<td>16.25 ± 0.10</td>
<td>10</td>
<td>8.5 ± 1.3</td>
</tr>
<tr>
<td>600</td>
<td>2.45 ± 0.87</td>
<td>2.33 ± 0.67</td>
<td>0.57 ± 0.10</td>
<td>1.29 ± 0.25</td>
<td>9.442 ± 0.056</td>
<td>6</td>
<td>6.6 ± 1.1</td>
</tr>
<tr>
<td>650</td>
<td>2.03 ± 0.83</td>
<td>1.01 ± 0.41</td>
<td>0.335 ± 0.079</td>
<td>0.76 ± 0.20</td>
<td>5.202 ± 0.032</td>
<td>4</td>
<td>4.14 ± 0.95</td>
</tr>
<tr>
<td>750</td>
<td>1.45 ± 0.65</td>
<td>0.62 ± 0.31</td>
<td>0.287 ± 0.080</td>
<td>0.65 ± 0.18</td>
<td>1.851 ± 0.010</td>
<td>4</td>
<td>3.01 ± 0.75</td>
</tr>
<tr>
<td>850</td>
<td>1.22 ± 0.61</td>
<td>0.62 ± 0.31</td>
<td>0.251 ± 0.078</td>
<td>0.61 ± 0.19</td>
<td>0.6973 ± 0.0037</td>
<td>4</td>
<td>2.70 ± 0.71</td>
</tr>
</tbody>
</table>

### TABLE VI. Individual background (BG) sources, expected signal, data, and total background event yields after the initial (first row) and final selections for the \(\mu\mu jj\) analysis. Other BG includes single top, \(W + \text{ jets}\), and \(VV + \ell\ell\). Only statistical uncertainties are reported.

<table>
<thead>
<tr>
<th>(M_{\ell Q})</th>
<th>(Z + \text{ jets})</th>
<th>(t\bar{t})</th>
<th>Other BG</th>
<th>LQ signal</th>
<th>Data</th>
<th>Total BG</th>
</tr>
</thead>
<tbody>
<tr>
<td>⋯</td>
<td>8644 ± 47</td>
<td>1218 ± 27</td>
<td>201.7 ± 2.8</td>
<td>⋯</td>
<td>9897</td>
<td>10063 ± 54</td>
</tr>
<tr>
<td>400</td>
<td>46.9 ± 2.4</td>
<td>30.1 ± 4.2</td>
<td>3.58±0.70 / -0.40</td>
<td>629.3 ± 4.0</td>
<td>68</td>
<td>80.6±0.49</td>
</tr>
<tr>
<td>500</td>
<td>10.4 ± 1.1</td>
<td>4.1 ± 1.6</td>
<td>0.89±0.62 / -0.21</td>
<td>136.73 ± 0.86</td>
<td>14</td>
<td>15.4±0.19</td>
</tr>
<tr>
<td>550</td>
<td>7.29 ± 0.94</td>
<td>2.4 ± 1.2</td>
<td>0.53±0.60 / -0.17</td>
<td>70.49 ± 0.40</td>
<td>9</td>
<td>10.2±0.53</td>
</tr>
<tr>
<td>600</td>
<td>5.03 ± 0.75</td>
<td>0.59 ± 0.59</td>
<td>0.47±0.60 / -0.16</td>
<td>37.39 ± 0.22</td>
<td>6</td>
<td>6.1±1.1</td>
</tr>
<tr>
<td>650</td>
<td>3.82 ± 0.65</td>
<td>0.59 ± 0.59</td>
<td>0.24±0.59 / -0.12</td>
<td>20.56 ± 0.13</td>
<td>5</td>
<td>4.7±1.1</td>
</tr>
<tr>
<td>750</td>
<td>2.03 ± 0.47</td>
<td>0.00±0.67 / -0.00</td>
<td>0.09±0.59 / -0.08</td>
<td>6.529 ± 0.038</td>
<td>1</td>
<td>2.1±1.0</td>
</tr>
<tr>
<td>850</td>
<td>1.56 ± 0.42</td>
<td>0.00±0.67 / -0.00</td>
<td>0.08±0.59 / -0.08</td>
<td>2.327 ± 0.014</td>
<td>0</td>
<td>1.6±0.4</td>
</tr>
</tbody>
</table>

052013-6
TABLE VIII. Systematic uncertainties and their effects on signal (S) and background (B) in all channels for the $M_{LQ} = 600$ GeV final selection. All uncertainties are symmetric.

<table>
<thead>
<tr>
<th>Systematic uncertainties</th>
<th>$eejj$</th>
<th>$\mu\mu jj$</th>
<th>$evjj$</th>
<th>$\mu\nu jj$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet energy scale</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Background modeling</td>
<td>⋮</td>
<td>⋮</td>
<td>11</td>
<td>⋮</td>
</tr>
<tr>
<td>Electron energy scale</td>
<td>1(3)</td>
<td>1</td>
<td>6</td>
<td>⋮</td>
</tr>
<tr>
<td>Muon momentum scale</td>
<td>1</td>
<td>⋮</td>
<td>⋮</td>
<td>⋮</td>
</tr>
<tr>
<td>Muon Reco/ID/Iso</td>
<td>1</td>
<td>⋮</td>
<td>⋮</td>
<td>⋮</td>
</tr>
<tr>
<td>Jet resolution (5–14)</td>
<td>0.5</td>
<td>⋮</td>
<td>⋮</td>
<td>⋮</td>
</tr>
<tr>
<td>Electron resolution</td>
<td>1(3)</td>
<td>0.5</td>
<td>1</td>
<td>⋮</td>
</tr>
<tr>
<td>Muon resolution</td>
<td>4</td>
<td>⋮</td>
<td>⋮</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td>Pileup</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>⋮</td>
</tr>
<tr>
<td>Integrated luminosity</td>
<td>2.2</td>
<td>2.2</td>
<td>⋮</td>
<td>⋮</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>13</td>
<td>3</td>
<td>11</td>
</tr>
</tbody>
</table>

However, as multijet processes are difficult to accurately model by MC simulation, several cross-checks are made with data control samples to ensure that the QCD multijet background in the $\mu\nu jj$ analysis is negligible. The method used to determine the QCD multijet background in the $evjj$ analysis is similar to the one used for the $eejj$ channel.

V. EVENT SELECTION OPTIMIZATION

After the initial selection, the sensitivity of the search is optimized by maximizing the Gaussian signal significance $S/\sqrt{S+B}$ in all channels. Optimized thresholds on the following variables are applied for each leptoquark mass hypothesis in the $\ell\ell jj$ channels: $M_{\ell\ell}$, $S_T^{\ell\ell}$, and $M_{\ell j}^{\min}$. The invariant mass of the dilepton pair, $M_{\ell\ell}$, is used to remove the major contribution from the $Z/\gamma^* + jets$ background. The variable $M_{\ell j}^{\min}$ is defined as the smaller lepton-jet invariant mass for the assignment of jets and leptons to leptoquarks, which minimizes the LQ-LQ invariant mass difference.

Thresholds on the following variables are optimized for each leptoquark mass hypothesis in the $\ell\nu jj$ channels: $E_T^{miss}$, $S_T^p$, and $M_{\ell j}$. A minimum threshold on $E_T^{miss}$ is used, primarily to reduce the dominant $W + jets$ background. The variable $M_{\ell j}$ is defined as the invariant mass of the lepton-jet combination, which minimizes the LQ-LQ transverse mass difference. In addition, a lower threshold is applied on the transverse mass of the lepton (electron or muon) and $E_T^{miss}$ in the event, $M_T > 120$ GeV.

The resulting optimized thresholds are summarized in Tables II and III for the $\ell\ell jj$ and the $\ell\nu jj$ channels, respectively.

After the initial selection criteria are applied, the yields in data are found to be consistent with SM predictions. Distributions of variables used in the final selection for the $eejj$, $evjj$, $\mu\mu jj$, and $\mu\nu jj$ analyses are shown in Figs. 2–5.

The number of events selected in data and estimated backgrounds are then compared at different stages of selection. This information is shown in Tables IV, V, VI, and VII for the initial selection and for the final selection for each channel separately.

Data and background predictions after final selection are also shown in Figs. 6–9, which compare $S_T$ and the best combination for the lepton-jet invariant mass for a signal

![Graph 1](#)

![Graph 2](#)

FIG. 6 (color online). $eejj$ channel: the distributions of $S_T^p$ (top) and of $M_{\ell j}$ for each of the two electron-jet pairs (bottom) for events that pass the final selection criteria optimized for a signal leptoquark mass of 600 GeV. The data are indicated by the points, and the SM backgrounds are given as cumulative histograms. The expected contribution from a leptoquark signal with $M_{LQ} = 600$ GeV is also shown.
leptoquark mass of 600 GeV in the four decay channels considered.

VI. SYSTEMATIC UNCERTAINTIES

The uncertainty on the integrated luminosity is taken as 2.2% [30].

The statistical uncertainties on the values of $R_Z$ ($R_W$) after initial selection requirements are used as an estimate of the uncertainty on the normalization of the $Z/\gamma^* + \text{jets}$ ($W + \text{jets}$) backgrounds. The uncertainty on the shape of the $Z/\gamma^* + \text{jets}$ and $W + \text{jets}$ distributions is calculated to be 15% (10%) and 20% (11%) for first (second) generation, respectively, by comparing the predictions of MADGRAPH samples produced with factorization or renormalization scales and matrix element-parton shower matching threshold varied up and down by a factor of 2.
The uncertainty on the estimate of the $t\bar{t}$ background in the $eejj$ and $\mu\mu jj$ channels is derived from the statistical uncertainty of the $e\mu jj$ data sample and the ratio of electron and muon reconstruction uncertainties, which is calculated to be 2% and 3%, respectively. In addition, an uncertainty of 7% is assigned based on the estimated contamination from sources other than $t\bar{t}$ in the data sample containing one electron and one muon. A 5.5% (5%) uncertainty on the normalization of the estimated $t\bar{t}$ background in the $evjj$ ($\mu\nu jj$) channel is given by the statistical uncertainty on the value of $R_{jj}$ after the initial selection requirements. A 10% (10%) uncertainty on the shape of the $t\bar{t}$ background distribution in the $evjj$ ($\mu\nu jj$) channel is estimated by comparing the predictions of MADGRAPH samples produced with factorization or renormalization scales and matrix element-parton shower matching thresholds varied up and down by a factor of 2.

A systematic uncertainty of 50% (25%) on the QCD multijet background estimate for the $eejj$ ($evjj$) channel is estimated from the difference between the number of observed data and the background prediction in a QCD multijet-enriched data sample with lower jet multiplicity.

PDF uncertainties on the theoretical cross section of leptoquark production and on the final selection acceptance have been calculated using the PDF4LHC [31] prescriptions, with PDF and $\alpha_s$ variations of the MSTW2008 [32], CTEQ6.6 [33], and NNPDF2.0 [34] PDF sets taken into account. Uncertainties on the cross section vary from 10% to 30% for leptoquarks in the mass range of 200–900 GeV, while the effect of the PDF uncertainties on signal acceptance varies from 1% to 3%. The PDF uncertainties are not considered for background sources with uncertainties determined from data. An uncertainty on the modeling of pileup interactions in the MC simulation is determined by varying the mean of the distribution of pileup interactions by 8%.

Energy and momentum scale uncertainties are estimated by assigning a 4% uncertainty on the jet energy scale, a 1% (3%) uncertainty on the electron energy scale for the barrel (end cap) region of ECAL, and a 1% uncertainty on the muon momentum scale. The effect of electron energy, muon momentum, and jet energy resolution on expected signal and backgrounds is assessed by smearing the electron energy by 1% and 3% in the barrel and end caps, respectively, by smearing the muon momentum by 4%, and by varying the jet energy resolution by an $\eta$-dependent value in the range 5%–14%. In the $e\nu jj$ analyses, the uncertainty on the energy and momentum scales and resolutions is propagated to the measurement of $E_{T}^{miss}$. The effect of these uncertainties is calculated for the (minor) background sources for which no data rescaling is applied. For the background sources for which data rescaling is applied, residual uncertainties are calculated (i.e. relative to the initial selection used to derive the rescaling factor).

Recent measurements of the muon reconstruction, identification, trigger, and isolation efficiencies using $Z \rightarrow \mu\mu$ events show very good agreement between data and MC events [35]. An $\sim$1% discrepancy is observed in the data-to-MC comparison of the muon trigger efficiency. This discrepancy is taken as a systematic uncertainty per muon, assigned to both signal and estimated background. The electron trigger and reconstruction and identification uncertainties contribute 3% (4%) to the uncertainty in both signal and estimated background for the $eejj$ ($evjj$) channel.

The systematic uncertainties and their effects on signal and background are summarized in Table VIII for all channels, corresponding to the final selection optimized for $M_{LQ} = 600$ GeV.

### VII. Results

The number of observed events in data passing the full selection criteria is consistent with the SM background prediction in all decay channels. An upper limit on the leptoquark pair-production cross section is therefore set using the CL$_S$ modified frequentist approach [17,18]. A log-normal probability function is used to integrate over the systematic uncertainties. Uncertainties of a statistical nature are described with $\Gamma$ distributions with widths determined by the number of events simulated in MC samples or observed in data control regions.

The 95% CL upper limits on $\sigma \times \beta^2$ or $\sigma \times 2\beta(1 - \beta)$ as a function of leptoquark mass are shown together with the NLO predictions for the scalar leptoquark...
Theoretical cross sections are represented for different values of the renormalization and factorization scale, $\mu^2$, varied between half and twice the leptoquark mass (blue shaded region). The PDF uncertainties are taken into account in the theoretical cross section values.

By comparing the observed upper limit with the theoretical cross section values, first-generation scalar leptoquarks with masses less than 830 (640) GeV are excluded with the assumption that $\beta = 1(0.5)$. Similarly,
second-generation scalar leptoquarks with masses less than 840 (620) GeV are excluded for $\beta = 1(0.5)$. This is to be compared with median expected limits of 790 (640) GeV for first-generation scalar leptoquarks and 800 (610) GeV for second-generation scalar leptoquarks.

The observed and expected limits on the branching fraction $\beta$ as a function of leptoquark mass can be further improved using the combination of the $e\ell jj$ and $\ell v jj$ channels, as shown in Fig. 12. These combinations lead to the exclusion of first- and second-generation scalar leptoquarks with masses less than 640 and 650 GeV for $\beta = 0.5$, compared with median expected limits of 680 and 670 GeV.

**VIII. SUMMARY**

In summary, a search for pair production of first- and second-generation scalar leptoquarks has been performed in decay channels with either two charged leptons of the same flavor (electrons or muons) and at least two jets, or a single charged lepton (electron or muon), missing transverse energy, and at least two jets, using 7 TeV proton-proton collisions data corresponding to an integrated luminosity of 5 fb$^{-1}$. The selection criteria have been optimized for each leptoquark signal mass hypothesis. The number of observed candidates for each hypothesis agree with the estimated number of background events. The CL$_S$ modified frequentist approach has been used to set limits on the leptoquark cross section times the branching fraction for the decay of a leptoquark pair. At 95% confidence level, the pair production of first- and second-generation leptoquarks is excluded with masses below 830 (640) GeV and 840 (650) GeV for $\beta = 1(0.5)$. These are the most stringent limits to date.

**ACKNOWLEDGMENTS**

We extend our thanks to Michael Kramer for providing the tools for calculation of the leptoquark theoretical cross section and PDF uncertainty. 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 the following: BMWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES (Croatia); RPF (Cyprus); MoER, SF0690030s09, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); SNSF (Switzerland); NSC (Taipei); TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA). Individuals have received support from the Marie-Curie programme and the European Research Council (European Union); the Leventis Foundation; the A.P. Sloan Foundation; the Alexander von Humboldt Foundation; the Austrian Science Fund (FWF); 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 Council of Science and Industrial Research, India; the Compaginia di San Paolo (Torino); and the HOMING PLUS programme of Foundation for Polish Science, cofinanced from European Union, Regional Development Fund.

(CMS Collaboration)
SEARCH FOR PAIR PRODUCTION OF FIRST- AND 

PHYSICAL REVIEW D 86, 052013 (2012)
SEARCH FOR PAIR PRODUCTION OF FIRST- AND ... PHYSICAL REVIEW D 86, 052013 (2012)

97 CERN, European Organization for Nuclear Research, Geneva, Switzerland
98 Paul Scherrer Institut, Villigen, Switzerland
99 Institute for Particle Physics, ETH Zurich, Zurich, Switzerland
100 Universitàt Zürich, Zurich, Switzerland
101 National Central University, Chung-Li, Taiwan
102 National Taiwan University (NTU), Taipei, Taiwan
103 Chulalongkorn University, Bangkok, Thailand
104 Cukurova University, Adana, Turkey
105 Middle East Technical University, Physics Department, Ankara, Turkey
106 Bogazici University, Istanbul, Turkey
107 Istanbul Technical University, Istanbul, Turkey
108 National Scientific Center, Kharkov Institute of Physics and Technology, Kharkov, Ukraine
109 University of Bristol, Bristol, United Kingdom
110 Rutherford Appleton Laboratory, Didcot, United Kingdom
111 Imperial College, London, United Kingdom
112 Brunel University, Uxbridge, United Kingdom
113 Baylor University, Waco, Texas, USA
114 The University of Alabama, Tuscaloosa, Alabama, USA
115 Boston University, Boston, Massachusetts, USA
116 Brown University, Providence, Rhode Island, USA
117 University of California, Davis, Davis, California, USA
118 University of California, Los Angeles, Los Angeles, California, USA
119 University of California, Riverside, Riverside, California, USA
120 University of California, San Diego, La Jolla, California, USA
121 University of California, Santa Barbara, Santa Barbara, California, USA
122 California Institute of Technology, Pasadena, California, USA
123 Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
124 University of Colorado at Boulder, Boulder, Colorado, USA
125 Cornell University, Ithaca, New York, USA
126 Fairfield University, Fairfield, Connecticut, USA
127 Fermi National Accelerator Laboratory, Batavia, Illinois, USA
128 University of Florida, Gainesville, Florida, USA
129 Florida International University, Miami, Florida, USA
130 Florida State University, Tallahassee, Florida, USA
131 Florida Institute of Technology, Melbourne, Florida, USA
132 University of Illinois at Chicago (UIC), Chicago, Illinois, USA
133 The University of Iowa, Iowa City, Iowa, USA
134 Johns Hopkins University, Baltimore, Maryland, USA
135 The University of Kansas, Lawrence, Kansas, USA
136 Kansas State University, Manhattan, Kansas, USA
137 Lawrence Livermore National Laboratory, Livermore, California, USA
138 University of Maryland, College Park, Maryland, USA
139 Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
140 University of Minnesota, Minneapolis, Minnesota, USA
141 University of Mississippi, University, Mississippi, USA
142 University of Nebraska-Lincoln, Lincoln, Nebraska, USA
143 State University of New York at Buffalo, Buffalo, New York, USA
144 Northeastern University, Boston, Massachusetts, USA
145 Northwestern University, Evanston, Illinois, USA
146 University of Notre Dame, Notre Dame, Indiana, USA
147 The Ohio State University, Columbus, Ohio, USA
148 Princeton University, Princeton, New Jersey, USA
149 University of Puerto Rico, Mayaguez, Puerto Rico, USA
150 Purdue University, West Lafayette, Indiana, USA
151 Purdue University Calumet, Hammond, Indiana, USA
152 Rice University, Houston, Texas, USA
153 University of Rochester, Rochester, New York, USA
154 The Rockefeller University, New York, USA
155 Rutgers, the State University of New Jersey, Piscataway, New Jersey, USA
156 University of Tennessee, Knoxville, Tennessee, USA
157 Texas A&M University, College Station, Texas, USA

052013-21
aDeceased.
bAlso at Vienna University of Technology, Vienna, Austria.
cAlso at National Institute of Chemical Physics and Biophysics, Tallinn, Estonia.
dAlso at Universidade Federal do ABC, Santo Andre, Brazil.
eAlso at California Institute of Technology, Pasadena, CA, USA.
fAlso at CERN, European Organization for Nuclear Research, Geneva, Switzerland.
gAlso at Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France.
hAlso at Suez Canal University, Suez, Egypt.
iAlso at Zewail City of Science and Technology, Zewail, Egypt.
jAlso at Cairo University, Cairo, Egypt.
kAlso at Fayoum University, El-Fayoum, Egypt.
lAlso at British University, Cairo, Egypt.
mNow at Ain Shams University, Cairo, Egypt.
nAlso at Soltan Institute for Nuclear Studies, Warsaw, Poland.
oAlso at Université de Haute-Alsace, Mulhouse, France.
pNow at Joint Institute for Nuclear Research, Dubna, Russia.
qAlso at Moscow State University, Moscow, Russia.
rAlso at Brandenburg University of Technology, Cottbus, Germany.
sAlso at Institute of Nuclear Research ATOMKI, Debrecen, Hungary.
tAlso at Eötvös Loránd University, Budapest, Hungary.
uAlso at Tata Institute of Fundamental Research—HECR, Mumbai, India.
vAlso at University of Visva-Bharati, Santiniketan, India.
wAlso at Sharif University of Technology, Tehran, Iran.
xAlso at Isfahan University of Technology, Isfahan, Iran.
yAlso at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Teheran, Iran.
zAlso at Facoltà Ingegneria Università di Roma, Roma, Italy.
aaAlso at Università della Basilicata, Potenza, Italy.
bbAlso at Università degli Studi Guglielmo Marconi, Roma, Italy.
ccAlso at Università degli Studi di Siena, Siena, Italy.
ddAlso at University of Bucharest, Faculty of Physics, Bucuresti-Magurele, Romania.
eeAlso at Faculty of Physics of University of Belgrade, Belgrade, Serbia.
ffAlso at University of Florida, Gainesville, FL, USA.
ffAlso at University of California, Los Angeles, Los Angeles, CA, USA.
hhAlso at Scuola Normale e Sezione dell’ INFN, Pisa, Italy.
iiAlso at INFN Sezione di Roma, Università di Roma “La Sapienza,” Roma, Italy.
jjAlso at University of Athens, Athens, Greece.
kkAlso at Rutherford Appleton Laboratory, Didcot, United Kingdom.
llAlso at The University of Kansas, Lawrence, KS, USA.
mmAlso at Paul Scherrer Institut, Villigen, Switzerland.
nnAlso at Institute for Theoretical and Experimental Physics, Moscow, Russia.
ooAlso at Gaziosmanpasa University, Tokat, Turkey.
ppAlso at Adiyaman University, Adiyaman, Turkey.
qqAlso at Izmir Institute of Technology, Izmir, Turkey.
rroAlso at The University of Iowa, Iowa City, IA, USA.
ssAlso at Mersin University, Mersin, Turkey.
ttAlso at Ozyegin University, Istanbul, Turkey.
uuAlso at Kafkas University, Kars, Turkey.
vvAlso at Suleyman Demirel University, Isparta, Turkey.
wwAlso at Ege University, Izmir, Turkey.
xxAlso at School of Physics and Astronomy, University of Southampton, Southampton, United Kingdom.