Search for pair production of first or second generation leptoquarks in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC


Published in:
Physical Review D. Particles, Fields, Gravitation, and Cosmology

DOI:
10.1103/PhysRevD.83.112006

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Search for pair production of first or second generation leptoquarks in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC

G. Aad et al.*
(ATLAS Collaboration)
(Received 25 April 2011; published 15 June 2011)

This paper describes searches for the pair production of first or second generation scalar leptoquarks using 35 pb$^{-1}$ of proton-proton collision data recorded by the ATLAS detector at $\sqrt{s} = 7$ TeV. Leptoquarks are searched in events with two oppositely-charged muons or electrons and at least two jets, and in events with one muon or electron, missing transverse momentum and at least two jets. After event selection, the observed yields are consistent with the predicted backgrounds. Leptoquark production is excluded at the 95% CL for masses $M_{LQ} < 376$ (319) GeV and $M_{LQ} < 422$ (362) GeV for first and second generation scalar leptoquarks, respectively, when assuming the branching fraction of a leptoquark to a charged lepton is equal to 1.0 (0.5).

DOI: 10.1103/PhysRevD.83.112006 PACS numbers: 14.80.Sv, 13.85.+t

I. INTRODUCTION

The standard model is extremely successful at describing the elementary particles and their interactions, yet it gives no explanation for the striking symmetry between quarks ($q$) and leptons ($\ell$). This has, in part, been the motivation for many beyond-the-standard-model theories that posit the existence of leptoquarks (LQ), particles that carry both lepton and baryon quantum numbers, which couple to both quarks and leptons, and carry the triplet charge of quantum chromodynamics (QCD), color. Theories that predict leptoquarks include models that contain quark and lepton substructure [1], theories that seek grand unification [2], and models of extended technicolor [3]. Experimental bounds from searches for flavor-changing-neutral currents and lepton-family-number violation place restrictive limits on leptoquark decays to different generations of quarks and leptons [4,5]. Direct leptoquark searches at electron–proton colliders have sensitivity to first generation leptoquarks, but have a non-negligible dependence on the value of the LQ $- \ell - q$ coupling [6], whereas second and third generation leptoquarks are produced via lepton-flavor-violation mechanisms [7]. The relatively large cross sections [8] for scalar leptoquark production from proton-proton collisions lead to the expectation that early LHC data offer sensitivity to a mass range beyond that probed by other accelerators. Currently, the most stringent limits come from the Tevatron [9,10] and from the CMS experiment at the LHC [11,12].

Scalar leptoquarks can be produced in proton–proton collisions either in leptoquark–antileptoquark pairs (LQLQ) or singly. Single LQ production involves an unknown LQ $- \ell - q$ coupling, whereas the pair-production reaction occurs mostly via QCD processes which involve only the strong coupling constant. Therefore, results from LQLQ searches have negligible model dependence, and the only relevant parameter for scalar LQ production is the leptoquark mass $M_{LQ}$ [13]. Final states from LQ pair production have either two same-flavor oppositely-charged leptons and two jets (lljj); a lepton, a neutrino, and two jets (lujj); or two neutrinos and two jets. The leptons l here and throughout the paper are either electrons for first generation LQ or muons for the second generation [14]. The cross sections for the leptoquark-mediated processes $pp \rightarrow lljj$ and $lljj$ can be written as $\sigma_{LQ} \times \beta^2$ and $\sigma_{LQ} \times 2\beta(1 - \beta)$, respectively, where $\beta$ is the branching fraction for a single leptoquark to decay into a charged lepton and a quark.

II. ANALYSIS STRATEGY

This paper reports a search for scalar leptoquark pair production carried out using 35 pb$^{-1}$ of data recorded by the ATLAS detector during the 2010 LHC proton-proton running period. The analysis is performed separately in the lljj final state and in the lujj final state. These searches are combined, leading to final results presented as a function of $\beta$ and $M_{LQ}$ for the first and second generations.

Analyses for both final states begin by selecting event samples that have high acceptance for signal production. At this initial stage, these samples are dominated by the major backgrounds, $Z +$ jets and $t\bar{t}$ for the lljj case, and $W +$ jets and $t\bar{t}$ for the lujj case. The samples are then subdivided into orthogonal control and signal regions.

The control regions are used to validate the background modeling by the Monte Carlo (MC) simulation, and the signal region is used to search for evidence of LQ production. The signal region is defined using an a priori optimization procedure based on simulated background and signal events.
A major difference between LQ events and backgrounds is the presence of jet-lepton (also jet-neutrino in the case of $1\nu jj$) pairs coming from the decay of the parent LQ, giving a peak in the reconstructed jet-lepton mass spectrum for the signal. The approximate reconstruction of these masses provides the most important variables used to distinguish signal and background events. In addition, large LQ masses give rise to larger total measured transverse energy in LQ pair events than is seen for background events, giving another means to distinguish signal from background. Finally, reconstructed boson masses can be used to reject the dominant backgrounds from $V(V = W, Z) + \text{jets}$ production where the boson decays into leptons.

In the $lljj$ channel, an average reconstructed leptoquark mass $M_{\text{LQ}}$ is defined for each event by computing the average of the masses from the two lepton-jet combinations in the event. Both possible assignments of the two leptons and the two leading jets to LQ parents are considered. The chosen assignment is that which gives the smallest absolute difference between the two reconstructed masses. The probability to get the correct pairing with this method is of the order of 90%. The transverse energy in an event $S_T^f$ is defined as the scalar sum of the transverse energy (momentum) of the two electrons (muons) and of the two leading jets, $S_T^f = p_T^e + p_T^\mu + +$, where the transverse plane is defined as relative to the beam axis [15]. The invariant mass of the dilepton pair $M_{ll}$ provides rejection of the $Z + \text{jets}$ background.

In the $1\nu jj$ channel, LQ mass equivalents are also defined. The neutrino transverse momentum $p_T^\nu$ is inferred from the missing transverse momentum in the event $E_T^{\text{miss}}$ as described in Sec. V. As in the $lljj$ final state, two pairings of lepton and jet and $E_T^{\text{miss}}$ and jet are possible. However, because the component of the neutrino momentum along the beamline is undetermined, only one mass $M_{\text{LQ}}$ from the charged lepton and a jet can be reconstructed. The $E_T^{\text{miss}}$ and the remaining jet are used to compute a transverse mass, $M_{\text{LQ}}^T = \sqrt{2p_T^e E_T^{\text{miss}}(1 - \cos \phi)}$ in which $p_T^j$ is the transverse momentum of a jet, and $\phi$ is the angle between the $p_T^e$ and the $E_T^{\text{miss}}$ vectors in the transverse plane. In analogy with the $lljj$ final state, the chosen pairing is that which gives the smallest absolute difference between $M_{\text{LQ}}$ and $M_{\text{LQ}}^T$, also resulting in the correct assignment more than 90% of the time. The measured transverse energy is defined as the scalar sum of the lepton transverse momentum, the missing transverse momentum, and the momentum of the two leading jets in the event, $S_T^f = p_T^e + E_T^{\text{miss}} + +$. An additional transverse mass variable $M_T = \sqrt{2p_T^e E_T^{\text{miss}}(1 - \cos \phi)}$ provides rejection against the dominant $W + \text{jets}$ background. Here, $p_T^e$ is the measured lepton transverse momentum, and $\phi$ is the angle between the $p_T^e$ and $E_T^{\text{miss}}$ vectors.

### III. ATLAS DETECTOR

ATLAS [16,17] is a multipurpose detector with a forward-backward symmetric cylindrical geometry and nearly $4\pi$ coverage in solid angle. The three major subcomponents of ATLAS are the tracking detector, the calorimeter, and the muon spectrometer. Charged-particle tracks and vertices are reconstructed in silicon based tracking detectors that cover $|\eta| < 2.5$ and transition radiation detectors extending to $|\eta| < 2.0$. The inner tracking system is immersed in a homogeneous 2 T magnetic field provided by a solenoid. Electron, photon, and jet transverse energies are measured in the calorimeter. The ATLAS calorimeter system is segmented into a central barrel and endcaps collectively covering the pseudorapidity range $|\eta| < 4.9$, and is equipped with both electromagnetic and hadronic calorimeters. Surrounding the calorimeter, a muon spectrometer with air core toroids, a system of precision tracking chambers, and detectors with triggering capabilities provides precise muon identification and measurements. A three-level event-triggering system allows for the selection of inclusive muon events with muon transverse momentum ($p_T^\mu$) greater than 13 GeV and inclusive electron events with electron transverse energy ($E_T$) greater than 15 GeV.

### IV. SIMULATED SAMPLES

Event samples processed through a detailed detector simulation [18] based on GEANT4 [19] are used to determine most background and all signal yields. As the recorded data contain a non-negligible fraction of events with more than one interaction per beam crossing, simulated events passing the initial lepton selection criteria are reweighted in order to bring the vertex multiplicity distributions in data and simulation into agreement. All background samples use ATLAS tune MC09 [20]. The vector-boson processes, $W + \text{jets}$ and $Z + \text{jets}$, are generated using ALPGEN V2.13 [21] with CTET6L1 [22] for the parton distribution functions (PDFs), interfaced to HERWIG V6.510 [23] and JIMMY V4.31 [24]. Exclusive samples with zero to four additional partons ($np$ with $n = 0, 1, 2, 3, 4$) and an inclusive sample with five or more additional partons ($np$ with $n = 5$) are used. The cross sections are computed using the ALPGEN cross sections scaled so that the sum of the $np$ sample cross sections are equal to the next-to-next-to-leading-order inclusive cross sections times branching fraction to a single lepton species: $\sigma(W \rightarrow \ell\nu) = 10.46 \pm 0.42$ nb and $\sigma(Z/\gamma^* \rightarrow \ell\ell) = 1.070 \pm 0.054$ nb for $M_{\ell\ell} > 40$ GeV [25,26]. Additional samples for cross checks are generated using SHERPA 1.1.3 [27] with PDF set CTEQ6L1, and PYTHIA 6.421 [28] with PDF MRST 2007 LO [29]. Samples of $\ell\ell$ events are produced using MC@NLO V3.41 [30] and POWHEG 1.01 patch 4 [31] with PDF CTEQ6.6M [32] interfaced to HERWIG for parton showering. Cross sections are scaled to the next-to-next-to-leading-
prediction of $165^{+14}_{-16}$ pb [33,34]. Single top events are generated using MC@NLO with cross sections of 3.94 pb, 58.7 pb, and 13.1 pb for the $s$-, $t$- and $tW$-channels, respectively. Their uncertainties are $\pm 10\%$ [35,36]. Diboson events are generated using HERWIG. Next-to-leading order (NLO) cross sections are calculated with MCFM [37]: $44.9 \pm 2.2$ pb, $18.0 \pm 1.3$ pb, and $5.96 \pm 0.30$ pb for $WW$, $WZ (M_\ell \ell > 40$ GeV) and $ZZ (M_\ell \ell > 60$ GeV), respectively. Cross section uncertainties take into account scale and PDF uncertainties, as well as differences with other generators.

Signal events for LQ masses of 250 to 400 GeV with a 50 GeV binning are generated with PYTHIA and tune D6 [38] with cross sections and uncertainties determined from Ref. [8] using the CTEQ6.6 [39] PDF set. The unknown LQ $- \ell - q$ coupling value is set to $0.01 \times \sqrt{4\pi\alpha_{EM}}$. This corresponds to an LQ full width of less than 1 MeV, and a negligible decay length [40,41].

V. OBJECT IDENTIFICATION

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter. Electron identification [42] is performed using the transverse shower shape, the longitudinal leakage into the hadronic calorimeter, and the requirement that a good-quality track points to the cluster. The electron transverse energy is measured in the calorimeter, while its direction is obtained from the track. Further rejection against hadrons is achieved by using the energy deposit patterns in the first layer of the electromagnetic calorimeter. In order to suppress the background from photon conversions, a hit in the first layer of the pixel detector is required. Electrons used in this analysis are required to have $ET > 20$ GeV and $|\eta| < 2.47$, with the exclusion of the poorly instrumented region between the barrel and the end-cap calorimeters at $1.35 < |\eta| < 1.52$. A small fraction of events with electrons near problematic regions of the electromagnetic calorimeter readout is removed.

Muons selected for this analysis are required to have $p_T^\mu > 20$ GeV and $|\eta| < 2.4$. Muon tracks are reconstructed independently in the inner detector and in the muon spectrometer, with a minimum number of hits required in each. A good match is required between the tracks found in the inner tracker and the muon spectrometer. In order to reject the cosmic ray background, tracks from muon candidates must extrapolate back to the reconstructed event vertex, satisfying $|d_0| < 0.1$ mm and $|z_0| < 1$ cm, where $d_0$ is the minimum distance between the muon trajectory and the event primary vertex in the plane perpendicular to the beam direction, and $z_0$ is the corresponding distance parallel to the beam direction.

Finally, both electrons and muons are required to be isolated from other energy in the calorimeters by imposing $E^{cone}_T/E_T > 0.2$ and $E^{cone}_T/p_T < 0.25$, for electrons and muons, respectively. Here $E^{cone}_T$ is the transverse energy in the calorimeter in a cone of size $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.2$ centered on the lepton direction, excluding the lepton contribution.

The presence of neutrinos is inferred from the missing transverse momentum $E_T^{miss}$. The $E_T^{miss}$ is defined from the vector sum of the transverse energy in calorimeter cells included in topological clusters [16] and the transverse momentum of the muon, $E_T^{miss} = -(\Sigma E_T^{cells} + \vec{p}_T^\mu)$. The clusters are corrected to take into account the different response to hadrons compared to electrons or photons, as well as dead material and out-of-cluster energy losses [43].

Jets are reconstructed from calorimeter energy clusters using the anti-$k_t$ [44] algorithm with a radius parameter $R = 0.4$. After applying quality requirements based on shower shape and signal timing with respect to the beam crossing [45], jets selected for this analysis must satisfy $p_T > 20$ GeV and $|\eta| < 2.8$ and must be separated from leptons by $\Delta R > 0.5$. Calibrations of lepton and jet transverse momenta, which are mostly derived from control samples in data, are applied prior to making the kinematic selections.

VI. PRESELECTION

Initial selection criteria define event samples with high signal acceptance and yields which are dominated by the major backgrounds $V + jets$ and $t\bar{t}$. Events in both electron and muon analyses are required to have at least one reconstructed proton-proton interaction vertex with at least three charged-particle tracks associated to it. The dilepton channels additionally require exactly two electrons (muons) with $E_T(p_T^\ell) > 20$ GeV. The single lepton channels require exactly one electron (muon) with $E_T(p_T^\ell) > 20$ GeV, $E_T^{miss} > 25$ GeV, and $M_T > 40$ GeV. For the $e\nu jj$ channel only, a triangle cut $\Delta \phi_j^\text{jet},E_T^{miss} > 1.5 \times (1 - E_T^{miss}/45)$, where $\phi$ is in radians and $E_T^{miss}$ in GeV is applied, in order to reduce the multijet background contamination. All analyses require at least two jets satisfying $p_T > 20$ GeV. The two highest $p_T$ jets are used to calculate LQ masses.

The acceptance for both signal and background is estimated by applying all selection criteria to simulated events. Expected yields are obtained by scaling the acceptance by the predicted cross section and integrated luminosity.

After preselection, good agreement is observed between simulation and data in both the number of events and the shape of distributions of kinematic variables. The determination of the expected backgrounds is discussed in Secs. VII and VIII. The expected and observed yields for the four different channels are summarized in Table I.

VII. BACKGROUND DETERMINATION

Small differences between data and simulation for resolutions and trigger and reconstruction efficiencies are determined in control data samples and applied to the
The predicted and observed yields for the preselected sample for all channels. Both statistical and systematic uncertainties are included.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Predicted Yield</th>
<th>Observed Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>eejj</td>
<td>610 ± 240</td>
<td>626</td>
</tr>
<tr>
<td>evjj</td>
<td>6100 + 1000 − 1100</td>
<td>6088</td>
</tr>
<tr>
<td>μμjj</td>
<td>830 + 200 − 150</td>
<td>853</td>
</tr>
<tr>
<td>μνjj</td>
<td>9500 ± 2500</td>
<td>9248</td>
</tr>
</tbody>
</table>

simulated events. These corrections influence the obtained yields by less than 2%. The background arising from lepton misidentification is determined using data-driven methods in the four channels, as is the contribution from the Z+jets background to lljj final states. All other backgrounds are modeled with Monte Carlo simulations and tested with data in control regions defined to enhance their contributions, as described in Sec. VIII.

Because the Z+jets yield in the signal region for the lljj final state arises in the tails of distributions, it is determined using a data-assisted method. The yield in the signal region \(N^\text{sig}_D\) is calculated as

\[
N^\text{sig}_D = \frac{N^Z_D}{N^Z_{MC}} N^\text{sig}_{MC},
\]

where \(N^Z_D\) and \(N^Z_{MC}\) are the observed numbers of events in data and MC, respectively, in a 20 GeV wide dilepton mass window around the nominal Z boson mass, and \(N^\text{sig}_{MC}\) is the expected Z+jets contribution to the signal region (defined in Table IV). To estimate a systematic uncertainty, the prediction of \(N^\text{sig}_D\) was derived with and without requirements on the number of jets in the event, as well as with the different Monte Carlo generators described in Sec. IV. The 3% background in the 20 GeV mass window from \(t\bar{t}\) and diboson events, estimated with Monte Carlo, is subtracted from the data, as is the <1% contamination from fake leptons. The data-assisted methods (with and without a cut on the number of jets and with various mass window definitions) and the purely Monte Carlo-based estimates agree within 10%. The largest difference is observed between results obtained with ALPGEN and SHERPA, and is used as a systematic uncertainty. The final estimate is obtained using ALPGEN for \(N^Z_{MC}\) and \(N^\text{sig}_{MC}\).

Jets misidentified as leptons and leptons from semileptonic decays of hadrons are referred to as fake leptons. In this paper, events with fake leptons are referred to as QCD background [46]. The QCD background in the lljj channel is estimated using a fitting method. Templates for both real and fake electrons (muons) are made for the \(E^\text{cone}_T/E_T\) \((E^\text{cone}_T/p_T)\) variable, before applying the isolation cut. Simulation is used to determine templates for real leptons. Templates for fake leptons are derived from the data by selecting events with exactly one lepton. Events containing W candidates are rejected by requiring \(E^\text{miss}_T < 10\) GeV. The small residual real lepton contamination is estimated from simulated events and is subtracted from the data. In the \(eejj\) channel, additional templates are made for the fraction of hits associated with transition radiation from electrons. For both channels, fits with these templates are performed on dilepton events with two or more jets giving the probability that an event contains at least one fake lepton. Fits are made inside the Z mass window, as well as in the signal region. Both fits indicate a small expected contribution from fake leptons. In the signal region, an independent upper limit on the number of fakes is set by extrapolating from the number of fakes determined in lower jet multiplicity bins and with lower dilepton mass. The QCD background contribution to the \(\mu\nujj\) analysis is determined using a scaled control sample method. In this method, a pair of variables that cleanly separate real muons from vector boson or leptoquark decay from fake muons is chosen to divide the full sample into four statistically independent regions. One of the four regions corresponds to the signal region, which also contains muons from vector-boson decays. The background in the signal region is then found by scaling the yield in one of the nonsignal regions by the ratio of yields in the remaining two regions. This analysis uses \(E^\text{miss}_T\) and muon impact parameter \(|d_0|\) to define the different regions. The three background regions are defined as \(|d_0| < 0.1\) mm and \(E^\text{miss}_T < 25\) GeV (A), \(|d_0| > 0.1\) mm and \(E^\text{miss}_T < 25\) GeV (B) and \(|d_0| > 0.1\) mm and \(E^\text{miss}_T > 25\) GeV (C). The signal region (D) is the region with \(|d_0| < 0.1\) mm and \(E^\text{miss}_T > 25\) GeV. The correlation between these two variables is less than 10% in simulated samples, giving stability to the method. The yield \(N_D\) in the signal region is given by \(N_D = N_Z N^\text{sig}_{MC} N_D\), where \(N\) refer to the observed numbers of events in the four statistically independent regions after subtraction of the small, residual contributions from events with real muons in the background regions, which are estimated using simulated events.

In the \(evjj\) channel, the normalization of the QCD background is determined from a fit to data of the sum of the \(M_T\) distributions for QCD events and all other backgrounds, primarily \(W +\) jets and \(t\bar{t}\). This sum is constrained to equal the total data yield, with the QCD fraction being the fit parameter. The template for the QCD background shape is determined using a QCD enriched sample constructed by taking the difference in shapes between two samples, a loose sample selected using only the electrons passing the trigger requirements, and the nominal sample, selected using the full electron identification requirements. Residual contamination from real electrons in the QCD enriched sample is estimated to be 7%. The latter is estimated with a loose-tight matrix method [33], and used to perform a shape-dependent subtraction. For both muon channels, the background from events containing cosmic ray muons is negligible.
VIII. CONTROL REGIONS

Control samples are used to validate the background determination based on MC. The two most important control samples for the $lljj$ analyses are (1) $Z + \text{jet}$: events for which the dilepton invariant mass lies within the $Z$ mass window $81 \leq M_l \leq 101$ GeV and (2) $t\bar{t}$: events that contain at least two jets and both an electron and a muon selected, as defined in Sec. V. By definition, the $t\bar{t}$ control region is common for both the $eejj$ and $\mu \mu jj$ channels. The most important control samples for the $lvjj$ analyses are (1) $W + 2$ jets: events with exactly two jets, a charged lepton, and $E_T^{\text{miss}}$ such that $M_{lT}$ is in the region of the $W$ Jacobian peak (40 GeV $\leq M_{lT} \leq 150$ GeV), (2) $W + 3$ jets: as in (1) but with at least 3 jets, and (3) a $t\bar{t}$ enriched sample which requires at least four jets with $>50$ GeV, $>40$ GeV, and $>30$ GeV. The expected signal contamination in the control regions is at most 1%. The predicted and observed yields in these control samples are shown in Tables II and III. Distributions of predicted and observed yields in these control samples are shown in Fig. 2 for the $lljj$ samples in the $Z + \text{jet}$ and $t\bar{t}$ control regions, and distributions for the reconstructed $M_{LQ}$ are shown in Fig. 2 for the $lvjj$ samples in the $W + 2$ jets and the $t\bar{t}$ control regions.

TABLE II. The predicted and observed yields in the control samples for the electron final states. Top refers to both single top and $t\bar{t}$ events. Both statistical and systematic uncertainties are included.

<table>
<thead>
<tr>
<th>Event Source</th>
<th>$eejj$ Control Region</th>
<th>$eejj$ Control Region</th>
<th>$evjj$ Control Region</th>
<th>$evjj$ Control Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>$Z + \geq 2$ jets</td>
<td>$t\bar{t}$</td>
<td>$W + 2$ jets</td>
<td>$W + \geq 3$ jet</td>
</tr>
<tr>
<td>$V + \text{jets}$</td>
<td>150 $\pm$ 23</td>
<td>0.3 $\pm$ 0.1</td>
<td>2100 $\pm$ 700</td>
<td>580 $\pm$ 190</td>
</tr>
<tr>
<td>Top</td>
<td>20 $\pm$ 0.3</td>
<td>24 $\pm$ 4</td>
<td>21 $\pm$ 4</td>
<td>44 $\pm$ 9</td>
</tr>
<tr>
<td>Diboson</td>
<td>2.0 $\pm$ 0.3</td>
<td>0.8 $\pm$ 0.1</td>
<td>17 $\pm$ 4</td>
<td>8.3 $\pm$ 1.9</td>
</tr>
<tr>
<td>QCD</td>
<td>4.0 $^{+14.0}_{-4.0}$</td>
<td>0.0 $^{+0.1}_{-0.0}$</td>
<td>64 $\pm$ 14</td>
<td>68 $\pm$ 15</td>
</tr>
<tr>
<td>Total Bkg</td>
<td>158 $\pm$ 25</td>
<td>25 $\pm$ 4</td>
<td>2200 $\pm$ 700</td>
<td>700 $\pm$ 200</td>
</tr>
<tr>
<td>Data</td>
<td>140</td>
<td>22</td>
<td>2344</td>
<td>722</td>
</tr>
</tbody>
</table>

TABLE III. The predicted and observed yields in the control samples for the muon final states. Top refers to both single top and $t\bar{t}$ events. Both statistical and systematic uncertainties are included.

<table>
<thead>
<tr>
<th>Event Source</th>
<th>$\mu \mu jj$ Control Region</th>
<th>$\mu \mu jj$ Control Region</th>
<th>$\mu \nu jj$ Control Region</th>
<th>$\mu \nu jj$ Control Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>$Z + \geq 2$ jets</td>
<td>$t\bar{t}$</td>
<td>$W + 2$ jets</td>
<td>$W + \geq 3$ jet</td>
</tr>
<tr>
<td>$V + \text{jets}$</td>
<td>190 $\pm$ 24</td>
<td>0.3 $\pm$ 0.1</td>
<td>3300 $\pm$ 1100</td>
<td>900 $\pm$ 300</td>
</tr>
<tr>
<td>Top</td>
<td>2.7 $\pm$ 0.5</td>
<td>24 $\pm$ 4</td>
<td>14 $\pm$ 3</td>
<td>53 $\pm$ 1</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.2 $\pm$ 0.1</td>
<td>28 $\pm$ 6</td>
<td>14 $\pm$ 3</td>
<td>3.0 $\pm$ 0.7</td>
</tr>
<tr>
<td>QCD</td>
<td>6.0 $^{+11.0}_{-6.0}$</td>
<td>0.0 $^{+0.1}_{-0.0}$</td>
<td>300 $\pm$ 100</td>
<td>130 $\pm$ 50</td>
</tr>
<tr>
<td>Total Bkg</td>
<td>200 $\pm$ 25</td>
<td>25 $\pm$ 4</td>
<td>3600 $\pm$ 1100</td>
<td>1100 $\pm$ 330</td>
</tr>
<tr>
<td>Data</td>
<td>216</td>
<td>22</td>
<td>3588</td>
<td>1120</td>
</tr>
</tbody>
</table>

IX. SELECTION OPTIMIZATION

The signal regions in the four final states are defined using a random grid search optimization method [47]. The optimization is performed using simulated signal and background samples (determined as described in Sec. VII and VIII) after preselection, with the final cuts determined by minimizing the Poisson probability that the predicted background fluctuates to at least the predicted signal plus background. A 10% overall scale uncertainty on the yields, roughly matching the uncertainty on the luminosity, is included in the optimization as a systematic uncertainty. Optimization for the single lepton (dilepton) channels nominally is done using signal samples with LQ masses of 300 (400) GeV. Cuts derived using different LQ masses typically differ from the nominal values within statistical uncertainties. A variety of input variables are considered. The requirements that give the best separation between signal and background are shown in Table IV. The yields obtained after final selection for the various channels are shown in Table V, along with the predicted background and signal.
X. SYSTEMATIC UNCERTAINITIES

Systematic uncertainties are derived for a variety of sources, including data-simulation differences in trigger and reconstruction efficiencies and in the energy and momentum resolutions for leptons, jets, and $E_{\text{T}}^{\text{miss}}$, instantaneous and integrated luminosity, modeling of the underlying event, variations in the method used to determine the QCD background, and uncertainty on the LQ pair-production cross section. A summary of the systematic uncertainties for the four final states is shown in Table VI.

The lepton trigger and reconstruction efficiency systematic uncertainties are derived by varying the selection of the Z event sample used to measure in situ efficiencies, and by varying the treatment of the (small) background in this sample. In addition, a 1% uncertainty is included on the muon isolation requirement that accounts for the difference of the efficiency of the isolation requirement in data and $t\bar{t}$ and LQ simulated samples. Finally, a 3% uncertainty is added to account for differences in the muon $|d_{0}|$ distributions between data and simulation. Lepton momentum scale and resolution uncertainties are obtained by comparing the peak and width of $Z \rightarrow \ell \ell$ events in data and Monte Carlo.

The jet energy scale and resolution are varied by their uncertainties [48,49] for all simulated events, and their impact estimated independently. In addition, a 5% uncertainty is added in quadrature to the jet energy scale uncertainties to account for differences in response for quark and gluon jets. These variations in scale and resolution are also propagated to the $E_{\text{T}}^{\text{miss}}$. The systematic uncertainty from instantaneous luminosity effects is evaluated by comparing the results from simulated samples with and without additional minimum bias events (pileup) added. This is an overestimate of the systematic, but is still small (2%–6%, depending on the sample) compared to other sources.

Systematic uncertainties on the QCD background are determined by comparing results from alternate normalizations to those described in Sec. VII. The uncertainty on the $lljj$ final states is determined by varying the estimate between the nominal prediction and the upper level from the extrapolation method described in Sec. VII. The uncertainty in the $evjj$ final state is determined by comparing the default method, based on fits to the $M_{T}$ distribution, to alternate fits using the $E_{T}^{\text{miss}}$ or the electron $E_{T}$ distributions. The largest fractional difference (22%) between the nominal fit and fits using the alternate variables is taken as the systematic uncertainty. The uncertainty in the background to the $\mu \nu jj$ final state is determined by comparing the default method to one that uses the muon isolation instead of the $d_{0}$ variable. The difference in yield between the two estimates is used to determine the systematic uncertainty, giving 27%.

The systematic uncertainties for the production models of $W +$ jets and $Z +$ jets events in the single lepton

FIG. 1 (color online). $S_{\ell}^{T}$ distributions for the $eejj$ (top) and the $\mu \mu jj$ final states (middle) in the Z control region. The bottom figure shows the $S_{\ell}^{T}$ distribution for the dilepton $t\bar{t}$ control region. The data are indicated by the points and the standard model backgrounds are shown with cumulative distributions. The QCD background is estimated from data, while the other background contributions are obtained from simulated samples. The top background includes both $t\bar{t}$ and single top events. The expected contribution from a potential LQ signal is negligible and is not included in the figure. The MC uncertainty band shows the combined statistical and systematic uncertainties on the prediction in each bin.
analysis are determined by varying parameters at the generator level as described in [26], including variations of the renormalization and factorization scales and minimum parton \( p_T \). The largest absolute deviation (40%) is used as a systematic uncertainty. In the dilepton analysis, the systematic uncertainty is determined by varying the dilepton mass window and replacing the ALPGEN event generator with SHERPA. The difference between ALPGEN and SHERPA dominates, giving an uncertainty in the \( eejj(\mu\mu jj) \) channel of 34 (45)% on the \( Z + \) jets prediction in the signal region.

Systematic uncertainties are evaluated for the \( t\bar{t} \) production model by comparing the predictions obtained with MC@NLO and POWHEG. The result is a 35% uncertainty for both \( lljj \) and \( lvjj \) channels.

An integrated luminosity uncertainty of 11% [50] is applied to all backgrounds determined from simulated events and to the signal. Additional signal uncertainties are obtained from varying the renormalization scale parameter to 0.5\( M_{LQ} \) and to 2\( M_{LQ} \) (15%), from different PDF choices (13%-17%) for LQ masses in the range 300–400 GeV, and from initial- and final-state-radiation effects (2%). The systematic uncertainties on the theory cross sections used to normalize backgrounds are given in Sec. IV.

TABLE IV. The selection requirements used to define the signal region, as obtained from the optimization procedure. \( p_T^{i1} > 30 \) GeV implies that the \( p_T \) of both the two leading leptons and the two leading jets in the dilepton samples should exceed 30 GeV.

\[
\begin{array}{ccc}
\text{\textbf{eejj and } } & \text{eejj} & \text{\textbf{\textmu\mu jj}} \\
M_{ll} > 120 \text{ GeV} & M_T > 200 \text{ GeV} & M_T > 160 \text{ GeV} \\
M_{1LQ} > 150 \text{ GeV} & M_{1LQ} > 180 \text{ GeV} & M_{1LQ} > 150 \text{ GeV} \\
p_T^{i1} > 30 \text{ GeV} & M_{LQ} > 180 \text{ GeV} & M_{LQ} > 150 \text{ GeV} \\
S_T^l > 450 \text{ GeV} & S_T^l > 410 \text{ GeV} & S_T^l > 400 \text{ GeV}
\end{array}
\]

XI. RESULTS

The expected and observed numbers of events are shown in Table V. Data and standard model expectations are in good agreement. 95% CL upper limits on LQ pair-production cross sections are determined using a modified
TABLE V. The predicted and observed yields in the signal region for all channels. The $lljj$ ($l\nu jj$) channel signal yields are computed assuming $\beta = 1.0(0.5)$. Both statistical and systematic uncertainties are included. Top refers to both single top and $tt$ events.

<table>
<thead>
<tr>
<th>Source</th>
<th>$eejj$</th>
<th>$e\nu jj$</th>
<th>$\mu\mu jj$</th>
<th>$\mu\nu jj$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V + \text{jets}$</td>
<td>$0.50 \pm 0.28$</td>
<td>$0.65 \pm 0.38$</td>
<td>$0.28 \pm 0.22$</td>
<td>$2.6 \pm 1.4$</td>
</tr>
<tr>
<td>Top</td>
<td>$0.51 \pm 0.23$</td>
<td>$0.67 \pm 0.39$</td>
<td>$0.52 \pm 0.23$</td>
<td>$1.6 \pm 0.9$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$0.03 \pm 0.01$</td>
<td>$0.10 \pm 0.03$</td>
<td>$0.04 \pm 0.01$</td>
<td>$0.10 \pm 0.03$</td>
</tr>
<tr>
<td>QCD</td>
<td>$0.02^{+0.03}_{-0.02}$</td>
<td>$0.06 \pm 0.01$</td>
<td>$0.00^{+0.01}_{-0.00}$</td>
<td>$0.0 \pm 0.0$</td>
</tr>
<tr>
<td>Total Bkg</td>
<td>$1.1 \pm 0.4$</td>
<td>$1.4 \pm 0.5$</td>
<td>$0.8 \pm 0.3$</td>
<td>$4.4 \pm 1.9$</td>
</tr>
<tr>
<td>Data</td>
<td>$2$</td>
<td>$2$</td>
<td>$0$</td>
<td>$4$</td>
</tr>
<tr>
<td>LQ(250 GeV)</td>
<td>$38 \pm 8$</td>
<td>$9.6 \pm 2.1$</td>
<td>$45 \pm 10$</td>
<td>$13 \pm 3$</td>
</tr>
<tr>
<td>LQ(300 GeV)</td>
<td>$17 \pm 4$</td>
<td>$5.1 \pm 1.1$</td>
<td>$21 \pm 5$</td>
<td>$6.4 \pm 1.4$</td>
</tr>
<tr>
<td>LQ(350 GeV)</td>
<td>$7.7 \pm 1.7$</td>
<td>$2.6 \pm 0.6$</td>
<td>$9.4 \pm 2.1$</td>
<td>$3.0 \pm 0.7$</td>
</tr>
<tr>
<td>LQ(400 GeV)</td>
<td>$3.5 \pm 0.8$</td>
<td>$\ldots$</td>
<td>$4.4 \pm 1.0$</td>
<td>$\ldots$</td>
</tr>
</tbody>
</table>

TABLE VI. Systematic uncertainties for the $lljj$ and $l\nu jj$ final states. The lepton trigger, identification, and momentum (energy) scale and resolution uncertainties are small and grouped together. Single top and $tt$ events are also grouped together. Uncertainties marked with * (**) are only for the electron (muon) channels. QCD modeling systematics are not shown. All numbers are in percentages.

<table>
<thead>
<tr>
<th>Channel</th>
<th>$lljj$</th>
<th>$l\nu jj$</th>
<th>$lljj$</th>
<th>$l\nu jj$</th>
<th>$lljj$</th>
<th>$l\nu jj$</th>
<th>$lljj$</th>
<th>$l\nu jj$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production Cross Section</td>
<td>$\ldots$</td>
<td>$4$</td>
<td>$13$</td>
<td>$13$</td>
<td>$5$</td>
<td>$5$</td>
<td>$18$</td>
<td>$18$</td>
</tr>
<tr>
<td>Modeling</td>
<td>$34^{*}$, $45^{**}$</td>
<td>$40$</td>
<td>$35$</td>
<td>$35$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>Electron Energy Scale &amp; Resolution*</td>
<td>$+13$, $-0.2$</td>
<td>$5$</td>
<td>$10$</td>
<td>$2$</td>
<td>$7$</td>
<td>$1$</td>
<td>$8$</td>
<td>$1$</td>
</tr>
<tr>
<td>Muon Momentum Scale &amp; Resolution**</td>
<td>$20$</td>
<td>$5$</td>
<td>$2$</td>
<td>$8$</td>
<td>$1$</td>
<td>$6.7$</td>
<td>$1$</td>
<td></td>
</tr>
<tr>
<td>Jet Energy Scale</td>
<td>$6$</td>
<td>$+22$, $-13$</td>
<td>$+9$, $-18$</td>
<td>$32$</td>
<td>$+16$, $-6$</td>
<td>$+17$, $-24$</td>
<td>$2$</td>
<td>$3$</td>
</tr>
<tr>
<td>Jet Energy Resolution</td>
<td>$16$</td>
<td>$10$</td>
<td>$0.3$</td>
<td>$26$</td>
<td>$4$</td>
<td>$14$</td>
<td>$0.3$</td>
<td>$3$</td>
</tr>
<tr>
<td>Luminosity</td>
<td>$0.3$</td>
<td>$11$</td>
<td>$11$</td>
<td>$11$</td>
<td>$11$</td>
<td>$11$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pile up</td>
<td>$&lt;0.1$</td>
<td>$5$</td>
<td>$&lt;0.1$</td>
<td>$4$</td>
<td>$&lt;0.1$</td>
<td>$6$</td>
<td>$&lt;0.1$</td>
<td>$2$</td>
</tr>
<tr>
<td>Total Systematics</td>
<td>$39^{*}$</td>
<td>$+49$, $-45$</td>
<td>$47^{*}$</td>
<td>$57$</td>
<td>$(+22$, $-16)$</td>
<td>$+26$, $-31$</td>
<td>$22$</td>
<td>$22$</td>
</tr>
<tr>
<td></td>
<td>$52^{**}$</td>
<td>$+(+49$, $-44)^{**}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIG. 3 (color online). $S_T^L$ distribution for the $eejj$ (left) and the $\mu\mu jj$ final states (right) after all selections. The data are indicated by the points and the standard model backgrounds are shown with cumulative distributions. The expected LQ signals for various masses are also shown.
SEARCH FOR PAIR PRODUCTION OF FIRST OR…

Monte Carlo studies show that a better sensitivity is achieved when using kinematic shapes rather than just the total number of events. For the eejj and μμjj final states, the observed and predicted $S_T$ distributions are used.

frequentist approach [51,52]. Systematic uncertainties are incorporated in the limit calculation as nuisance parameters and are integrated out. The limits are translated into bounds in the $\beta$ versus LQ mass plane.

Monte Carlo studies show that a better sensitivity is achieved when using kinematic shapes rather than just the total number of events. For the eejj and μμjj final states, the observed and predicted $S_T$ distributions are used.

FIG. 4 (color online). $M_{LQ}$ distribution for the $e\nu j j$ (left) and the $\mu \nu j j$ final state (right) after all selections. The data are indicated by the points and the standard model backgrounds are shown with cumulative distributions. The expected LQ signals for various masses are also shown.

Monte Carlo studies show that a better sensitivity is achieved when using kinematic shapes rather than just the total number of events. For the eejj and μμjj final states, the observed and predicted $S_T$ distributions are used.

FIG. 5 (color online). 95% CL upper bounds on the LQ pair production cross sections as a function of LQ mass for the combination of the eejj and the eνjj channels (top), and for the $\mu \mu jj$ and the $\mu \nu jj$ channels (bottom). The figures on the left (right) show the limit for $\beta = 1.0 \ (0.5)$. The expected limit is indicated by the dashed line. The light yellow (dark green) solid band contains 68% (95%) of possible outcomes from pseudoexperiments in which the yield is Poisson fluctuated around the background-only expectation. Systematic uncertainties are included. The observed limit is indicated by the solid line. The theory prediction is indicated by the dotted line, which includes systematic uncertainties due to the choice of the PDF and to the renormalization and factorization scales ($\Delta \sigma_{\mu}$).
in the limit-setting procedure, and for the $e\nu jj$ and $\mu\nu jj$ final states the observed and predicted $M_{LQ}$ distributions are used. The $S_T$ distributions are shown in Fig. 3 for the $lljj$ channels. Figure 4 shows the $M_{LQ}$ distributions for the single lepton final states.

The 95% CL upper bounds on the cross section for first and second generation LQ pair production as a function of mass are shown in Fig. 5 for $\beta = 1.0$ and $\beta = 0.5$. The combined limits are also shown in the $\beta$ vs $M_{LQ}$ plane in Fig. 6 for both generations. The expected and observed combined limits including all systematic uncertainties are shown in Table VII for both $\beta = 1.0$ and $\beta = 0.5$. The systematic uncertainties, of the order of 50%, only change the limits by 5 to 10 GeV, depending on the value of $\beta$.

### XII. CONCLUSIONS

This paper reports the results of searches for pair production of first or second generation scalar leptoquarks using a data sample corresponding to an integrated luminosity of 35 pb$^{-1}$. The data in the high signal-to-background signal region are in good agreement with the standard model expectations. 95% CL upper bounds on the production cross section are determined. These are translated into lower bounds for first (second) generation leptoquark masses of $M_{LQ} > 376$ (422) GeV and $M_{LQ} > 319$ (362) GeV for $\beta = 1.0$ and $\beta = 0.5$, respectively. These are the most stringent bounds to date from direct searches for leptoquarks in much of the phase space.

### ACKNOWLEDGMENTS

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC, and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST, and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR, and VSC CR, Czech Republic; DNRF, DNSRC, and Lundbeck Foundation, Denmark; ARTEMIS, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNAS, Georgia; BMBF, DFG, HGF, MPG, and AvH Foundation, Germany; GSRT, Greece; ISF, MINERVA, GIF, DIP, and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNiSW, Poland; GRICES and FCT, Portugal; MERSYS (MECTS), Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; AARRS and MVZT, Slovenia; DST/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF, and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society, and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

### TABLE VII. 95% CL expected and observed limits on the first and second generation LQs under different assumptions on $\beta$.

<table>
<thead>
<tr>
<th>Type ($\beta$)</th>
<th>Expected limit (GeV)</th>
<th>Observed limit (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st generation (1.0)</td>
<td>387</td>
<td>376</td>
</tr>
<tr>
<td>1st generation (0.5)</td>
<td>348</td>
<td>319</td>
</tr>
<tr>
<td>2nd generation (1.0)</td>
<td>393</td>
<td>422</td>
</tr>
<tr>
<td>2nd generation (0.5)</td>
<td>353</td>
<td>362</td>
</tr>
</tbody>
</table>
America. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular, from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK), and BNL (USA) and in the Tier-2 facilities worldwide.

[13] The LQ – ℓ – q coupling determines the LQ lifetime and width. For LQ masses in the range considered here, 200 GeV ≤ M_{LQ} ≤ 400 GeV, couplings greater than q_{e} × 10^{-6}, with q_{e} the electron charge, correspond to decay lengths less than roughly 1 mm. In addition, to be insensitive to the coupling, the width cannot be larger than the experimental resolution of a few GeV. This sets the approximate sensitivity to the unknown coupling strength.
[14] By convention, we place the electron and its neutrino, and the up and down quarks in the first generation.
[15] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z-axis coinciding with the axis of the beam pipe. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, φ) are used in the transverse plane, φ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as η = −ln(tan(θ/2)).

PHYSICAL REVIEW D 83, 112006 (2011)
SEARCH FOR PAIR PRODUCTION OF FIRST OR...
SEARCH FOR PAIR PRODUCTION OF FIRST OR... PHYSICAL REVIEW D 83, 112006 (2011)