CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch 2016/08/05

Search for pair-production of first generation scalar leptoquarks in pp collisions at $\sqrt{s} = 13$ TeV with 2.6 fb^{-1}

The CMS Collaboration

Abstract

A search for pair production of first generation scalar leptoquarks is presented, considering a final state containing two electrons and at least two jets, using proton-proton collision data at $\sqrt{s} = 13$ TeV. The data were recorded with the CMS detector at the LHC during 2015, and correspond to an integrated luminosity of 2.6 fb^{-1}. First generation scalar leptoquarks with masses less than 1130 (920) GeV are excluded for $\beta = 1.0$ ($\beta = 0.5$), where β is the branching fraction of a leptoquark decaying to a charged lepton and a quark.

1 Introduction

The standard model of particle physics contains hints that the lepton and quark sectors are related by a fundamental symmetry. Many beyond the standard model theories [1–5], include such a symmetry, which gives rise to new bosons called leptoquarks (LQ). Leptoquarks couple to both leptons and quarks, carrying lepton and baryon numbers, and fractional electric charge. They can be either scalar or vector particles (carrying zero or one unit of spin) and are color triplets. Current experimental searches for rare processes, such as lepton number violation and flavor-changing neutral currents, suggest that leptoquarks come in three generations which do not mix [6, 7].

Leptoquark pair-production at the LHC proceeds mainly through gluon-gluon fusion (gg) and quark anti-quark annihilation ($q\bar{q}$), with gluon-gluon fusion being the dominant production process. Due to the gg dominance, along with the fact that only one $q\bar{q}$ Feynman diagram contains the LQ-quark-lepton vertex, scalar LQ pair-production has negligible dependence on Yukawa LQ-quark-lepton coupling, usually denoted as λ, and LQ searches are essentially independent of λ. The production cross section does depend on the leptoquark mass M_{LQ}.

Leptoquarks decay to a quark and a lepton, giving rise to lepton-jet resonances. A first generation leptoquark will decay to an electron and a quark with branching faction β, or an electron neutrino and a quark with branching fraction $1-\beta$; β is treated as a free parameter in the analysis. This gives rise to three distinct final states for pair-produced leptoquarks: two electrons and two jets; one electron, missing transverse energy, and jets; and no electrons, missing transverse energy, and jets. The final state branching fractions are β^2, $2\times\beta(1-\beta)$, and $(1-\beta)^2$, respectively, and the first and second final states are denoted as $eejj$ and $e\nu jj$. This analysis considers a $eejj$ final state only.

Previous experimental limits from CMS with 19.7 fb^{-1} of data at $\sqrt{s}=8$ TeV yielded lower limits on first generation leptoquark masses of 1010 (850) GeV for $\beta=1(0.5)$ [8]. The ATLAS experiment released results from $\sqrt{s}=13$ TeV collision data taken in 2015, using 3.2 fb^{-1} of data, and reported observed (expected) mass limits of 1100 (1160) GeV for $\beta=1$ [9]. Here, the corresponding search using data collected by the CMS detector is presented, using 2.6 fb^{-1} of data.

The results are presented in the following manner. Section 2 is a description of the CMS detector. Section 3 gives an overview of the data and Monte Carlo simulated events used in the analysis. Section 4 details the physics objects used, as well as the event selection. Section 5 describes the major backgrounds and how they are estimated. Section 6 discusses the signal significance optimization procedure. Section 7 details the systematic uncertainties. The results are presented in Section 8, and summarized in Section 9.

2 The CMS Detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the superconducting solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity [10] coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid.

The electromagnetic calorimeter consists of 75 848 lead tungstate crystals, which provide cov-

erage in pseudorapidity $|\eta| < 1.479$ in a barrel region (EB) and $1.479 < |\eta| < 3.0$ in two endcap regions (EE). A preshower detector consisting of two planes of silicon sensors interleaved with a total of $3X_0$ of lead is located in front of the EE.

The electron momentum is estimated by combining the energy measurement in the ECAL with the momentum measurement in the tracker. The momentum resolution for electrons with $p_T \approx 45$ GeV from $Z \to ee$ decays ranges from 1.7% for nonshowering electrons in the barrel region to 4.5% for showering electrons in the endcaps [11].

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [10].

3 Data and Monte Carlo Simulation Samples

In this search, the data used correspond to 2.6 fb^{-1} of integrated luminosity, and were collected by the CMS detector from proton-proton collisions at the LHC in 2015. Events passing a single electron high-level trigger (HLT) as well as single photon HLT paths were used, as is described below in Section 4.

Monte Carlo (MC) simulated events were used to study both leptoquark signal and most standard model backgrounds. The CMS detector simulation is implemented using GEANT4 [12]. For scalar leptoquark signals, MC event samples were generated for masses between 200 and 2000 GeV, at 50 GeV intervals. The generator used was PYTHIA [13, 14] version 8.1, with the NNPDF2.3LO parton distribution function (PDF) set [15]. Krämer et al. have calculated the leptoquark cross sections at next-to-leading order (NLO) [16, 17] using the CTEQ6.6M PDF set [18]. The signal cross sections are shown in Table 1.

Table 1: Scalar first-generation LQ signal cross sections calculated at NLO.

M_{LQ}	$\sigma(\mu = M_{LQ})$ [pb]	M_{LQ}	$\sigma(\mu = M_{LQ})$ [pb]
200	6.06E+01	850	1.82E-02
250	2.03E+01	900	1.23E-02
300	8.04E+00	950	8.45E-03
350	3.59E+00	1000	5.86E-03
400	1.74E+00	1050	4.11E-03
450	9.06E-01	1100	2.91E-03
500	4.96E-01	1150	2.08E-03
550	2.84E-01	1200	1.50E-03
600	1.69E-01	1250	1.09E-03
650	1.03E-01	1300	7.95E-04
700	6.48E-02	1350	5.85E-04
750	4.16E-02	1400	4.33E-04
800	2.73E-02	1450	3.21E-04

Standard model processes give rise to an *eejj* final state, and are thus considered as background in this analysis. The dominant processes are top quark pair-production ($t\bar{t}$) and associated W or Z vector boson production with jets. Other processes (including QCD multijets, single top quark, diboson, and γ+jets) contribute less significantly. MC samples are used to study them:

- $t\bar{t}$+jets: generated with MADGRAPH [19, 20]
- Z/γ+jets and W+jets: generated with MADGRAPH [19, 20] in bins of H_T
- single top quark: generated with POWHEG [21]

- diboson (WW, ZZ, or WZ): generated with PYTHIA
- γ+jets: generated with MADGRAPH in bins of H_T

where H_T is defined as the scalar p_T sum over all jets in the event. Pileup events are included in the MC simulation, and they are reweighted to reproduce the distributions as seen in the data.

4 Physics Objects and Event Selection

4.1 Object Reconstruction

Electrons used for this search are identified from tracks in the tracking system which are matched to energy deposits in the electromagnetic calorimeter (ECAL) [11]. Their associated ECAL deposit must have transverse energy (E_T) > 50 GeV, and pseudorapidity $|\eta| < 2.5$, excluding the ECAL barrel-endcap transition region, $1.4442 < |\eta| < 1.566$. For high-momentum electrons, a dedicated set of identification requirements is applied [11]. The requirements ensure that the ECAL deposit is consistent with that expected from an electron and that the track is well matched in η and ϕ (azimuthal angle) and has no more than one inner tracker hit missing. In addition, the hadronic calorimeter energy in a cone of $\Delta R < 0.15$ around the electron must be less than 5% of the electron's energy (after correcting for event activity unrelated to the electron). Finally, the tracker and calorimeter energies in a cone of $\Delta R < 0.3$ around the electron must be low enough so that the electron can be said to be isolated (again, after correcting for other event activity).

Muons are used as a veto in the analysis in order to suppress standard model backgrounds. They are reconstructed by matching inner tracker tracks to tracks in the muon system [22]. They must pass a p_T cut of 45 GeV and have $|\eta| < 2.1$. Identification criteria optimized for high-momentum muons are applied, including requirements on the muon track quality, the relative track isolation, and the number of hits in the muon system and inner tracker. Primary vertex matching is also imposed.

Jet reconstruction is accomplished by using the anti-k_T algorithm [23] with a distance parameter $R = 0.4$. The particles used in the jet clustering are the result of the particle flow technique [24], which utilizes an optimized combination of all CMS sub-detectors to reconstruct photons, electrons, muons, and charged and neutral hadrons in each event. Identification criteria are applied to jets based on their particle flow constituent energy fractions, and they must have at least two constituents [25]. The jet energy calibration is achieved by studying dijet and photon+jet events [26]. The jets must have reconstructed $p_T > 50$ GeV and $|\eta| < 2.4$. Jets are vetoed if they are closer than $\Delta R = 0.3$ to a lepton. Missing energy E_T^{miss} is the magnitude of the projection onto the plane orthogonal to the beam line of the negative vector sum of all the particle flow candidates in the event.

4.2 Event Selection

We use a single electron trigger to select candidate signal events in the data, where a passing electron must satisfy $p_T > 27$ GeV, and a series of loose identification and isolation requirements. In the Monte Carlo simulation, we do not use the simulated trigger but instead apply the efficiency as measured from the data. A series of event filters are applied to the Monte Carlo simulation and the data in order to reject backgrounds from, e.g., detector noise and beam halo events.

The events used in this search contain exactly two electrons and at least two jets, which must satisfy the criteria detailed above. The two leading jets are taken along with the electrons to

form two leptoquark candidates. In addition, invariant mass of the electrons M_{ee} must be greater than 50 GeV, and the S_{T}, defined as the scalar p_{T} sum of the two leading jets and two electrons, must be greater than 300 GeV. These requirements constitute a preselection, which we use to study background estimates and perform a signal significance optimization.

5 Background Determination

The major standard model background processes which give rise to *eejj* final states are Z/γ^*+jets and $\mathrm{t\bar{t}}$+jets, with smaller contributions from W+jets, γ+jets, single top, and diboson events. QCD multijet events, in which a jet can be misidentified as an electron passing the identification criteria, are also an important background. The Z/γ^*+jets background is estimated using the shape from Monte Carlo simulation, normalizing to the data in the preselection region around the Z peak: $80 < M_{ee} < 100$ GeV. The ratio of data over MC events in this region is $R_Z = 1.00 \pm 0.03$ (stat). Examining a tightening of the S_{T} threshold alone (so that we remain away from the signal region), we observe a significant variation of R_Z. Therefore we apply corrections to the Z/γ^*+jets MC events as a function of S_{T}. The corrections range from close to unity at low S_{T} to up to about 0.3 at high S_{T}. The systematic uncertainty is taken as the full value of the correction itself. For preselection the corrections are derived and applied in S_{T} bins, while for final selections, corrections are derived using the corresponding one-sided S_{T} thresholds. The $\mathrm{t\bar{t}}$+jets background shape is estimated from the MC simulation, normalized using a sample of $e\mu jj$ events in the data at preselection with a cut on the S_{T} of 340 GeV (lowest LQ final selection threshold). From these $e\mu jj$ data events, we subtract the contributions from non-$\mathrm{t\bar{t}}$+jets background using MC, and apply rescaling to account for the difference in electron and muon efficiencies and acceptances, and for the difference in the single electron trigger efficiency between final states containing one or two electrons. This prediction can then be compared with *eejj* events in the $\mathrm{t\bar{t}}$+jets MC to yield a scale factor $R_{t\bar{t}} = 0.88 \pm 0.03$ (stat). We examine $R_{t\bar{t}}$ in the MC and the data-driven prediction while tightening the S_{T} threshold. We take the maximum relative difference of $R_{t\bar{t}}$ calculated using the data-driven prediction with respect to the MC prediction as the systematic uncertainty.

The QCD multijet background arises from misidentified jets which pass the electron identification criteria. We estimate this background by estimating the probability for a misidentified electron passing a loose identification to also pass the full identification (as described in Section 4.1). This probability is calculated using events firing single photon triggers, and is obtained in the ECAL barrel and two endcap regions, with several jet multiplicities. To obtain the QCD multijet background prediction at each final selection point, the probability is applied to events passing the nominal final selections, but where the loose electron identification criteria are used. The method is tested by applying the probability to a sample of events where both electrons have failed the full selection. This is compared to a sample of events where one electron has passed the full selection and the other has failed it. To assign a systematic uncertainty, we repeat the same comparison, but with the S_{T} threshold tightened. The difference in the predicted and observed number of events, compared between the two S_{T} regions, is taken as the systematic uncertainty on the QCD multijet predicted background per electron, which is about 20%.

We proceed to compare the data and background predictions at preselection. Figure 1 shows the p_{T} distributions of the two leading electrons and jets in the event. For final selections, we use three variables: M_{ee}, S_{T}, and $M_{min}(ej)$. The $M_{min}(ej)$ is calculated by first obtaining the electron-jet pairing that minimizes the leptoquark anti-leptoquark invariant mass difference, and then taking the smaller of the two resulting $M(ej)$. Figure 2 shows the spectra of these

three variables at preselection.

The background prediction is found to describe the data well at preselection.

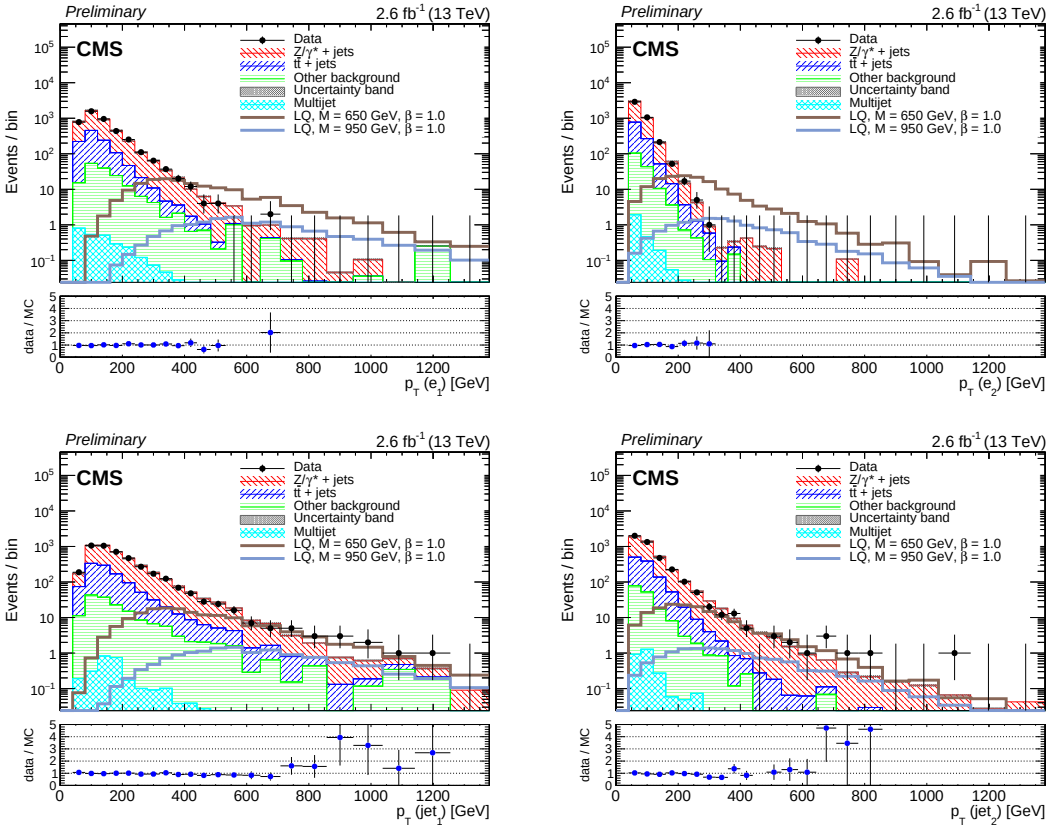

Figure 1: Distributions of p_{T} for the two leading electrons and two leading jets in the events passing preselection. Two leptoquark signal mass hypotheses are shown in addition to the data and background predictions. The category denoted as "Other background" includes single top, W+jets, diboson, and γ+jets events.

6 Signal Significance Optimization

Given the good data and background agreement observed at preselection, we proceed to optimize the signal significance. This is done using the signal MC samples and the background predictions derived as described above. The three final selection variables M_{ee}, S_{T}, and $M_{min}(ej)$, are scanned in the complete 3-D parameter space. We then choose the thresholds which maximize the significance criterion proposed by Punzi, using a significance of five standard deviations [27]. This is done for each leptoquark mass point considered. The resulting thresholds on the three selection variables are fit with a second-degree polynomial to smooth away any statistical fluctuations. The results from the fit are then used as the final thresholds in the analysis. These thresholds are shown in Table2.

After optimization, the signal efficiencies times acceptances range from about 30% at low LQ masses to about 50% at high LQ masses. Applying the final selection thresholds results in the event yields given in Table 3.

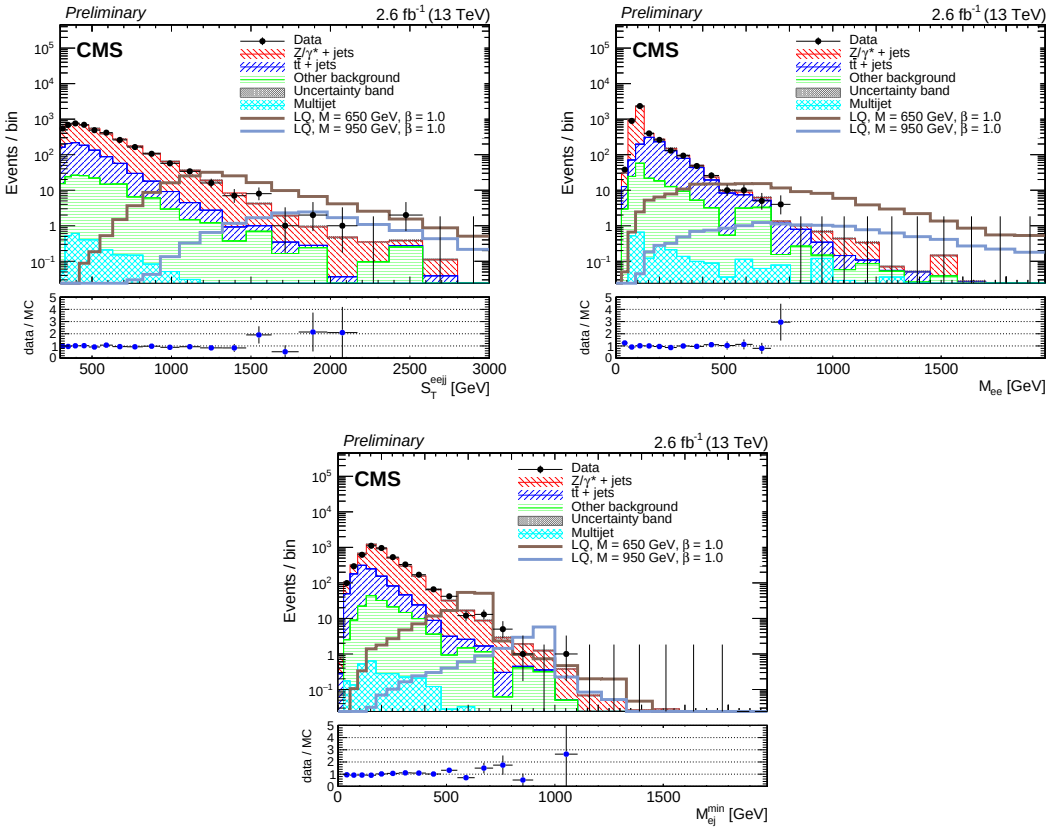

Figure 2: Distributions of the final selection variables S_{T}, M_{ee}, and $M_{min}(ej)$, in the events passing preselection. Two leptoquark signal mass hypotheses are shown in addition to the data and background predictions. The category denoted as "Other background" includes single top, W+jets, diboson, and γ+jets events.

Table 2: Optimized final selection thresholds for the $eejj$ analysis.

M_{LQ} [GeV]	S_{T}[GeV] >	m_{ee} [GeV] >	m_{ej}^{min} [GeV] >
200	340	130	160
250	405	140	205
300	470	155	245
350	535	165	285
400	595	175	325
450	660	185	360
500	720	195	400
550	780	205	435
600	840	210	470
650	900	220	500
700	960	230	535
750	1015	235	565
800	1075	245	595
850	1130	250	625
900	1190	255	650
950	1245	265	675
1000	1300	270	705
1050	1355	275	725
1100	1410	280	750
1150	1460	285	775
1200	1515	285	795
1250	1565	290	815
1300	1615	295	830
1350	1670	300	850
1400	1720	300	865
1450	1770	300	880
1500	1815	305	895

Table 3: Event yields after the optimized *eejj* selections for signal, observed data, and background predictions. All uncertainties are statistical except for the total background, where statistical and systematic uncertainties are reported separately.

M_{LQ}	LQ Signal	Z+Jets	$t\bar{t}$	QCD	Other	Total BG (stat.) (syst.)	Data
preselection	–	3207 ± 19	952.4 ± 2.8	2.70 ± 0.04	160.6 ± 4.3	$4323 \pm 19 \pm 240$	4252
LQ200	27772 ± 270	109.8 ± 3.3	223.1 ± 1.3	0.92 ± 0.02	40.9 ± 2.5	$374.8 \pm 4.3 \pm 130$	380
LQ250	12120 ± 99	60.1 ± 2.3	114.99 ± 0.94	0.68 ± 0.02	$24.5^{+1.8}_{-1.7}$	$200.3^{+3.1}_{-3.0} \pm 70$	221
LQ300	5690 ± 42	30.3 ± 1.4	57.32 ± 0.65	0.47 ± 0.01	$15.9^{+1.4}_{-1.3}$	$104.0^{+2.1}_{-2.0} \pm 37$	119
LQ350	2776 ± 19	18.5 ± 1.0	29.46 ± 0.46	0.24 ± 0.01	$9.5^{+1.1}_{-1.0}$	$57.7^{+1.6}_{-1.5} \pm 22$	63
LQ400	1480.7 ± 9.5	10.83 ± 0.64	15.15 ± 0.33	0.18 ± 0.01	$5.89^{+0.91}_{-0.76}$	$32.1^{+1.2}_{-1.0} \pm 13$	36
LQ450	800.3 ± 5.1	7.44 ± 0.54	8.18 ± 0.24	0.14 ± 0.01	$2.53^{+0.71}_{-0.53}$	$18.29^{+0.92}_{-0.79} \pm 9.7$	23
LQ500	461.6 ± 2.8	4.45 ± 0.28	4.56 ± 0.18	0.12 ± 0.01	$1.32^{+0.60}_{-0.37}$	$10.44^{+0.68}_{-0.50} \pm 6.1$	14
LQ550	268.7 ± 1.6	3.13 ± 0.21	2.86 ± 0.14	0.05 ± 0.00	$1.08^{+0.54}_{-0.32}$	$7.12^{+0.59}_{-0.41} \pm 4.6$	8
LQ600	166.51 ± 0.94	2.05 ± 0.11	1.57 ± 0.10	0.03 ± 0.00	$0.84^{+0.47}_{-0.28}$	$4.50^{+0.49}_{-0.32} \pm 3.2$	5
LQ650	107.72 ± 0.59	1.60 ± 0.10	1.03 ± 0.08	0.03 ± 0.00	$0.65^{+0.46}_{-0.19}$	$3.31^{+0.48}_{-0.23} \pm 2.5$	1
LQ700	69.54 ± 0.37	1.16 ± 0.09	0.66 ± 0.07	0.02 ± 0.00	$0.54^{+0.45}_{-0.18}$	$2.38^{+0.47}_{-0.21} \pm 1.9$	0
LQ750	45.07 ± 0.24	0.80 ± 0.07	0.47 ± 0.06	0.01 ± 0.00	$0.51^{+0.43}_{-0.13}$	$1.78^{+0.44}_{-0.16} \pm 1.4$	0
LQ800	30.69 ± 0.16	0.56 ± 0.05	0.29 ± 0.05	0.01 ± 0.00	$0.30^{+0.42}_{-0.11}$	$1.17^{+0.43}_{-0.13} \pm 1.1$	0
LQ850	20.69 ± 0.10	0.38 ± 0.04	0.24 ± 0.05	0.01 ± 0.00	$0.29^{+0.43}_{-0.13}$	$0.93^{+0.44}_{-0.14} \pm 0.84$	0
LQ900	14.40 ± 0.07	0.26 ± 0.03	0.18 ± 0.04	0.00 ± 0.00	$0.29^{+0.43}_{-0.13}$	$0.73^{+0.44}_{-0.13} \pm 0.64$	0
LQ950	10.10 ± 0.05	0.15 ± 0.02	0.15 ± 0.04	0.00 ± 0.00	$0.29^{+0.43}_{-0.12}$	$0.59^{+0.43}_{-0.13} \pm 0.49$	0
LQ1000	7.06 ± 0.03	0.13 ± 0.02	0.12 ± 0.03	0.00 ± 0.00	$0.03^{+0.42}_{-0.03}$	$0.28^{+0.42}_{-0.10} \pm 0.41$	0
LQ1050	5.09 ± 0.02	0.12 ± 0.02	0.06 ± 0.03	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.18^{+0.41}_{-0.08} \pm 0.36$	0
LQ1100	3.62 ± 0.02	0.11 ± 0.02	$0.04^{+0.03}_{-0.02}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.16^{+0.41}_{-0.08} \pm 0.29$	0
LQ1150	2.66 ± 0.01	0.09 ± 0.02	$0.04^{+0.03}_{-0.02}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.14^{+0.41}_{-0.08} \pm 0.26$	0
LQ1200	1.91 ± 0.01	0.07 ± 0.01	$0.04^{+0.03}_{-0.02}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.11^{+0.41}_{-0.07} \pm 0.23$	0
LQ1250	1.37 ± 0.01	0.03 ± 0.01	$0.02^{+0.02}_{-0.01}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.06^{+0.41}_{-0.06} \pm 0.11$	0
LQ1300	1.04 ± 0.00	0.04 ± 0.01	$0.02^{+0.02}_{-0.01}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.07^{+0.41}_{-0.06} \pm 0.11$	0
LQ1350	0.78 ± 0.00	0.03 ± 0.01	$0.02^{+0.02}_{-0.01}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.05^{+0.41}_{-0.05} \pm 0.06$	0
LQ1400	0.57 ± 0.00	0.02 ± 0.00	$0.00^{+0.02}_{-0.00}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.02^{+0.41}_{-0.02} \pm 0.04$	0
LQ1450	0.43 ± 0.00	0.02 ± 0.00	$0.00^{+0.02}_{-0.00}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.02^{+0.41}_{-0.02} \pm 0.04$	0
LQ1500	0.32 ± 0.00	0.01 ± 0.00	$0.00^{+0.02}_{-0.00}$	0.00 ± 0.00	$0.00^{+0.41}_{-0.00}$	$0.01^{+0.41}_{-0.01} \pm 0.02$	0

7 Systematic Uncertainties

The following systematic uncertainties are taken into account:

- An uncertainty of 2.7% on the determination of the integrated luminosity [28].
- A jet energy scale uncertainty, dependent on jet p_{T} and η of 2-15%.
- A jet resolution uncertainty [26], dependent on η and calculated by scaling the generated and reconstructed jet p_{T} difference typically between 11 and 23%.
- Z/γ^*+jets, $\mathrm{t\bar{t}}$+jets background shape uncertainty determined using MC weights representing factorization and renormalization scales varied by a factor of 0.5 and 2.
- Z/γ^*+jets, $\mathrm{t\bar{t}}$+jets normalization uncertainty: At each final selection, the Z+jets normalization is corrected based on the ratio of data to MC at the corresponding S_{T} threshold, as discussed above. The systematic uncertainty assigned is the distance of each correction from 1. For the $t\bar{t}$ normalization, as noted above, a comparison between a data-driven method and the MC, again at each final selection's S_{T} threshold, shows a maximum relative difference of about 20% from the nominal scale factor, which is taken as an S_{T} independent systematic uncertainty.
- A 40% uncertainty on the QCD multijet background prediction, as described in Section 5.
- An uncertainty on the signal acceptance, and background acceptance and cross section due to PDF uncertainty, which we evaluate using the PDF4LHC prescription [29].
- Modeling of pileup interactions in the MC is taken to have a $\pm$5% uncertainty.
- Electron trigger, reconstruction, and identification uncertainties: The trigger efficiency uncertainty is taken to be 6%, the ID efficiency uncertainty as 4% in the barrel and 6% in the endcap, and the reconstruction efficiency uncertainty as 5% [30].
- The electron energy scale uncertainty is taken to be 2% and the resolution uncertainty as 10% at high p_{T} [11].

To determine the uncertainties in the cases besides Z/γ^*+jets, $\mathrm{t\bar{t}}$+jets, and QCD normalization, the uncertainty source in question is shifted up and down by one standard deviation, and the signal and background yields are recalculated at each set of selection thresholds. The uncertainty is then taken as the maximum difference between the up/down shifted yields with respect to the yields without any shift. In the case of uncertainties on Z/γ^*+jets and $\mathrm{t\bar{t}}$+jets background, which are normalized to the data at preselection, before evaluating this difference, the yields with the up/down systematic shifts are also normalized to the data at preselection.

8 Results

We do not observe evidence of LQ signal in the data. We proceed to set upper limits on the scalar leptoquark production cross section using a full frequentist CLs method [31], using the yields in Table 3. The systematic uncertainties are modeled as nuisance parameters with a log-normal distribution, while statistical uncertainties are parameterized as log-normal functions, or, in the case of small yields, gamma functions with widths determined from the number of MC or data control region events.

The upper limits at 95% confidence level on $\sigma \times \beta^2$ along with the predicted scalar pair-production cross section at NLO are shown in Figure 3. The theoretical prediction at NLO is evaluated including the PDF uncertainty as well as variations in the renormalization and factorization scales, giving rise to a band around the nominal value. The intersection of the observed limits

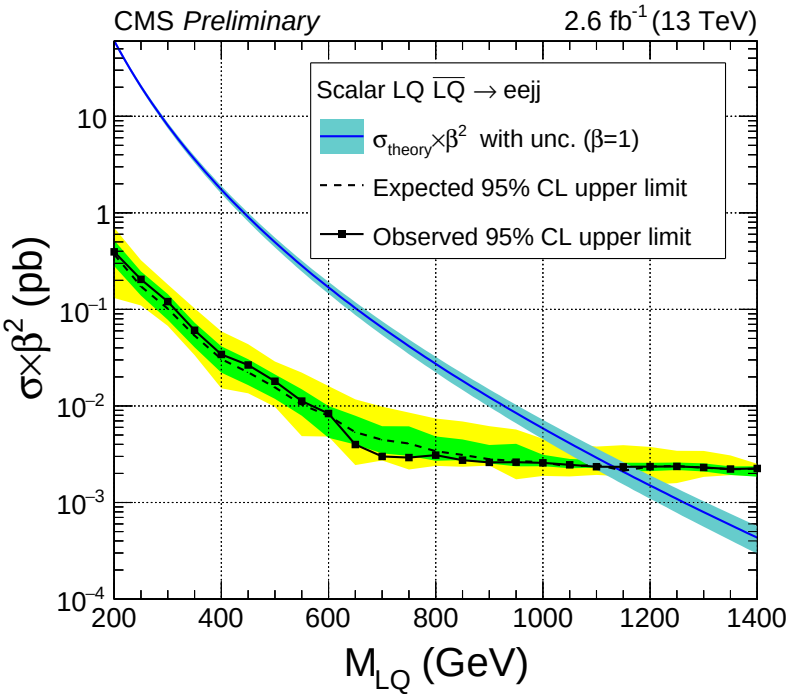

Figure 3: Observed and expected upper limits for scalar leptoquark pair-production cross section times β^2 at the 95% confidence level. The median (dashed line), 1σ (green), and 2σ (yellow) expected limits are shown. The theoretical prediction is varied by the adjusting the renormalization and factorization scales and combining with the PDF uncertainty, to producing the shaded band.

and the theoretical curve yields the lower limit on the leptoquark mass. The observed limit excludes first generation scalar leptoquarks with masses less than 1130 GeV, while the median expected limit is 1140 GeV. This assumes that $\beta = 1$. We also interpret the results in the case where $\beta = 0.5$. In this case, first generation scalar leptoquarks with masses less than 920 GeV are excluded, with a median expected limit of 910 GeV.

9 Conclusions

A search for pair production of first generation scalar leptoquarks has been presented, based on pp collision data collected in 2015 at $\sqrt{s} = 13$ TeV and corresponding to an integrated luminosity of 2.6 fb^{-1}. The search considers the leptoquark decay into a final state containing two electrons and two jets. After background prediction is complete, the selection thresholds are optimized using a scan of the full parameter space for each leptoquark mass point. At each set of optimized selection thresholds, the number of observed events agrees with the background expectation based on standard model processes. Limits at 95% confidence level are set assuming $\beta = 1$, excluding first generation scalar leptoquarks with masses less than 1130 GeV, compared to 1140 GeV for the median expected limit. Mass limits are also obtained for $\beta = 0.5$, in which case first generation scalar leptoquarks with masses less than 920 GeV are excluded, with a median expected limit of 910 GeV. This is the tightest CMS constraint on the first generation scalar leptoquark mass to date.

References

[1] B. Schrempp and F. Schrempp, "Light Leptoquarks", *Phys. Lett. B* **153** (1985) 101, doi:10.1016/0370-2693(85)91450-9.

[2] S. Dimopoulos and L. Susskind, "Mass Without Scalars", *Nucl. Phys. B* **155** (1979) 237, doi:10.1016/0550-3213(81)90304-7.

[3] H. Georgi and S. Glashow, "Unity of All Elementary-Particle Forces", *Phys. Rev. Lett.* **32** (1974) 438, doi:10.1103/PhysRevLett.32.438.

[4] J. C. Pati and A. Salam, "Lepton Number as the Fourth Color", *Phys. Rev. D* **10** (1974) 275, doi:10.1103/PhysRevD.10.275.

[5] J. L. Hewett and T. G. Rizzo, "Low-energy phenomenology of superstring-inspired E_6 models", *Phys. Lett.* **183** (1989) 193, doi:10.1016/0370-1573(89)90071-9.

[6] W. Buchmüller and D. Wyler, "Constraints on SU(5)-type Leptoquarks", *Phys. Lett. B* **177** (1986) 377, doi:10.1016/0370-2693(86)90771-9.

[7] O. Shanker, "πl2, Kl3, and K^0-$\bar{K}^0$ Constraints on Leptoquarks and Supersymmetric Particles", *Nucl. Phys. B* **204** (1982) 375, doi:10.1016/0550-3213(82)90196-1.

[8] CMS Collaboration, "Search for pair production of first and second generation leptoquarks in proton-proton collisions at sqrt(s) = 8 TeV", *Phys. Rev. D* **93** (2016), no. 3, 032004, doi:10.1103/PhysRevD.93.032004, arXiv:1509.03744.

[9] ATLAS Collaboration, "Search for scalar leptoquarks in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS experiment", arXiv:1605.06035. Submitted for publication in New Journal of Physics.

[10] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[11] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[12] GEANT4 Collaboration, "GEANT4: A simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[13] T. Sjöstrand et al., "High Energy Physics Event Generation with PYTHIA 6.1", *Comput. Phys. Commun.* **135** (2001) 238, doi:10.1016/S0010-4655(00)00236-8, arXiv:hep-ph/0010017.

[14] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[15] R. D. Ball et al., "Parton distributions with LHC data", *Nucl. Phys. B* **867** (2013) 244–289, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.

[16] M. Kramer, T. Plehn, M. Spira, and P. M. Zerwas, "Pair production of scalar leptoquarks at the Tevatron", *Phys. Rev. Lett.* **79** (1997) 341–344, doi:10.1103/PhysRevLett.79.341, arXiv:hep-ph/9704322.

[17] M. Krämer et al., "Pair production of scalar leptoquarks at the CERN LHC", *Phys. Rev. D* **71** (2005).

[18] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", *Phys. Rev. D* **78** (2008) 013004, doi:10.1103/PhysRevD.78.013004, arXiv:0802.0007.

[19] F. Maltoni and T. Stelzer, "Madevent: Automatic Event Generation with MadGraph", *JHEP* **02** (2003) 027, doi:10.1088/1126-6708/2003/02/027, arXiv:hep-ph/0208156.

[20] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **09** (2007) 028, doi:10.1088/1126-6708/2007/09/028, arXiv:0706.2334.

[21] S. Alioli et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[22] CMS Collaboration, "Muon Reconstruction and Identification Improvements for Run-2 and First Results with 2015 Run Data", CMS Detector Performance Summary CMS-DP-2015-015, 2015.

[23] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[24] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.

[25] CMS Collaboration, "Jet Performance in pp Collisions at 7 TeV", Technical Report CMS-PAS-JME-10-003, CERN, Geneva, 2010.

[26] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[27] G. Punzi, "Sensitivity of searches for new signals and its optimization", *eConf* **C030908** (2003) MODT002, arXiv:physics/0308063. [,79(2003)].

[28] CMS Collaboration, "CMS Luminosity Measurement for the 2015 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-15-001, 2015.

[29] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[30] CMS Collaboration, "Search for a Narrow Resonance Produced in 13 TeV pp Collisions Decaying to Electron Pair or Muon Pair Final States", Technical Report CMS-PAS-EXO-15-005, CERN, Geneva, 2015.

[31] A. L. Read, "Modified frequentist analysis of search results (the CL_s method)", CERN-OPEN 2000-205, 2000. 1st Workshop on Confidence Limits, CERN, Jan. 2000.