Search for pair-produced scalar and vector leptoquarks decaying into third-generation quarks and first- or second-generation leptons in pp collisions with the ATLAS detector

The ATLAS Collaboration

DOI
10.1007/JHEP06(2023)188

Publication date
2023

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):
Search for pair-produced scalar and vector leptoquarks decaying into third-generation quarks and first- or second-generation leptons in $pp$ collisions with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

Abstract: A search for pair-produced scalar and vector leptoquarks decaying into quarks and leptons of different generations is presented. It uses the full LHC Run 2 (2015–2018) data set of 139 fb$^{-1}$ collected with the ATLAS detector in proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Scalar leptoquarks with charge $-(1/3)e$ as well as scalar and vector leptoquarks with charge $+(2/3)e$ are considered. All possible decays of the pair-produced leptoquarks into quarks of the third generation ($t, b$) and charged or neutral leptons of the first or second generation ($e, \mu, \nu$) with exactly one electron or muon in the final state are investigated. No significant deviations from the Standard Model expectation are observed. Upper limits on the production cross-section are provided for eight models as a function of the leptoquark mass and the branching ratio of the leptoquark into the charged or neutral lepton. In addition, lower limits on the leptoquark masses are derived for all models across a range of branching ratios. Two of these models have the goal of providing an explanation for the recent $B$-anomalies. In both models, a vector leptoquark decays into charged and neutral leptons of the second generation with a similar branching fraction. Lower limits of 1980 GeV and 1710 GeV are set on the leptoquark mass for these two models.

Keywords: Exotics, Hadron-Hadron Scattering

ArXiv ePrint: 2210.04517
1 Introduction

Leptoquarks (LQs) have already been discussed for a few decades, as they provide a connection between the quark and lepton sectors, which exhibit similar structures. They are predicted by many extensions of the Standard Model (SM), e.g. in unified theories [1–3] and technicolor [4–6] or composite models [7–9]. Recent hints of a potential violation of lepton flavour universality in various measurements of $B$-meson decays (‘$B$-anomalies’) [10–21] can also be attributed, if confirmed, to the exchange of leptoquarks [22–29]. In addition, some of these models [28, 29, 32] introducing LQs aim to simultaneously provide an explanation for the longstanding discrepancy between the measured and the predicted anomalous magnetic moment of the muon [33]. LQs are bosons carrying colour charge and fractional electrical charge. They possess non-zero baryon and lepton numbers and are assumed to decay into a quark–lepton pair. The branching ratio into a quark and a charged lepton is denoted by $\mathcal{B}$, and that into a quark and neutrino by $1 - \mathcal{B}$. Leptoquarks can be scalar or vector bosons and can be produced singly or in pairs in proton–proton collisions.

\footnote{After submission of the paper a new result from LHCb has been reported [30, 31] superseding refs. [17, 20, 21] and indicating compatibility with the SM.}
The assumption that LQs can only interact with leptons and quarks of the same generation [34], and are spin-0 particles, has been used in most of the searches for LQs. Recently, however, searches for LQs with couplings to quarks and leptons of different generations have aroused interest because these couplings are required in order to explain the \( B \)-anomalies mentioned above. The theoretical explanations usually require LQs with couplings to third-generation quarks and second-generation leptons. In particular, vector LQs with a charge of \( 2/3 \), in units of the elementary charge \( e \), have been identified recently as promising candidates [35, 36].

First results of searches for LQs with couplings to quarks and leptons of different generations were published recently for pair-produced LQs decaying into charged leptons, i.e. in dilepton final states. Results from the full LHC Run 2 data set (2015–2018) of proton–proton collisions at \( \sqrt{s} = 13 \text{ TeV} \), corresponding to an integrated luminosity of 139 fb\(^{-1}\), are available for pair-produced scalar LQs decaying into a top-quark and an electron or muon from the ATLAS [37] and the CMS Collaborations [38] and for the decay into a \( b \), \( c \) or light-quark and an electron or muon [39] from ATLAS. In addition, using a partial (35.9 fb\(^{-1}\)) Run 2 data set, CMS has published a result for pair-produced vector LQs decaying into a top-quark and a muon, obtained by scaling the results for scalar LQs to the larger production cross-sections expected for vector LQs, assuming no kinematic differences [40].

Given the good coverage of large branching ratios \( \mathcal{B} \) by the dileptonic measurements described above, the results presented here use a single-lepton (electron or muon) final state optimised for medium to small \( \mathcal{B} \). In this case, one of the LQs decays into a neutrino and the other decays into a charged lepton, or both decay into neutrinos and the charged lepton arises from a leptonically decaying top-quark (\( t_{\text{lep}} \)). The results are interpreted as searches for pair-produced LQs with charges of either \( \pm (2/3)e \) (up-type) or \( \pm (1/3)e \) (down-type). All possible decays of the pair-produced up-type and down-type LQs into a quark (\( t, b \)) of the third generation and a lepton (\( \ell, \nu \)) of the first or second generation are considered, as seen in figure 1. With flavour off-diagonal couplings allowed, the model used for up-type (LQ\(_{\text{up}}^{\text{mix}}\)) and down-type (LQ\(_{\text{down}}^{\text{mix}}\)) scalar LQs is an extension of that [41] used in previous ATLAS searches [42], where all possible decays of the pair-produced up-type and down-type scalar LQs into a quark (\( t, b \)) and a lepton (\( \tau, \nu \)) of the third generation were considered. The present search for up-type LQs is also optimised for a vector LQ (vLQ) model [43] designed to provide an explanation for the various \( B \)-anomalies.

The analysis strategy is based on a final-state signature with one lepton, high missing transverse momentum and at least four jets due to a hadronically decaying top-quark (\( t_{\text{had}} \)) and a \( b \)-quark. Dedicated neural networks (NNs), trained in a common training region, are used for the separation of signal and background. This is done separately for scalar and vector LQ pair-production because they exhibit different kinematic behaviour for small values of \( \mathcal{B} \), i.e. when the charged lepton arises mostly from the top-quark decay. For each of the models and for various branching ratios, a signal region (SR) based on the NN output is defined. Control regions (CRs) are defined so as to be enriched in the various background processes. They are orthogonal to the SR, and orthogonal to each other. The statistical interpretation is based on a simultaneous fit to the CRs and the SR, in which the
normalisations for top-quark pair (tt), W + jets, and single top-quark production as well as a possible signal contribution are determined, while taking into account the experimental and theoretical systematic uncertainties. The results are presented as limits on the leptoquark mass as a function of the branching ratio.

2 ATLAS detector

The ATLAS experiment [44] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle.\textsuperscript{2} It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner tracking detector (ID) covers the pseudorapidity range |η| < 2.5. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A steel/scintillator-tile hadron calorimeter covers the central pseudorapidity range (|η| < 1.7). The endcap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to |η| = 4.9. The muon spectrometer (MS) surrounds the calorimeters and is based on three large superconducting air-core toroidal magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector. The muon spectrometer includes a system of precision tracking chambers and fast detectors for triggering. A two-level trigger

\textsuperscript{2}ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, φ) are used in the transverse plane, φ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as η = −ln tan(θ/2). Angular distance is measured in units of \(\Delta R = \sqrt{(\Delta η)^2 + (\Delta φ)^2}\).
system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to accept events at a rate below 100 kHz. This is followed by a software-based trigger that reduces the accepted event rate to 1 kHz on average depending on the data-taking conditions. An extensive software suite [45] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

3 Data and simulated event samples

In this search, data from proton–proton collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 139 fb$^{-1}$, collected in the years 2015 to 2018 with the ATLAS detector, are analysed. Data are required to have been collected during stable beam conditions and with all detector subsystems operational [46]. The average number of simultaneous $pp$ interactions per bunch crossing, referred to as pile-up, is approximately 34, averaged over the whole data set.

Monte Carlo (MC) simulated event samples are used to model the signal and background processes. In all samples except those produced with SHERPA 2.2.1 or SHERPA 2.2.2 [47], decays of heavy-flavour hadrons were modelled with EvtGen 1.2.0 or EvtGen 1.6.0 [48], depending on the process. Pile-up was modelled by overlaying minimum-bias events generated with PYTHIA 8.186 [49] and the A3 [50] set of tuned parameters (referred to as the ‘tune’) onto the simulated hard-scatter events. A reweighting procedure was applied in order to match the pile-up profile of the recorded data. The ATLAS simulation infrastructure [51] was used to simulate the detector and its response. Nominal SM background samples were produced with a detailed GEANT4 [52] detector simulation, whereas a faster calorimeter simulation [51] was applied for the signal samples and systematic variations of the backgrounds. The same offline reconstruction methods used for data were applied to the simulated samples. Corrections were applied to the simulated events in order to match the selection efficiencies and energy and mass scales and resolutions of reconstructed simulated particles to those measured in data control samples.

Simulated events with pair-produced scalar LQs were generated at next-to-leading order (NLO) in quantum chromodynamics (QCD) with MADGRAPH5_AMC@NLO 2.6.0 [53] and the NNPDF3.0nlo parton distribution function (PDF) [54] set with $\alpha_s = 0.118$. An extension of the LQ model of ref. [41] was used, allowing flavour off-diagonal couplings. The model is based on previous fixed-order NLO QCD calculations [55, 56]. To retain information about spin correlations, the decays of LQs as well as top-quarks were handled with MadSpin [57]. MADGRAPH5_AMC@NLO was interfaced with PYTHIA 8.230 [58] to model the parton shower (PS), hadronisation, and underlying event (UE). Here and in the following, PYTHIA was used with the A14 tune [59] and the NNPDF2.3LO [60] set of PDFs. The coupling strength $\lambda$ was set to 0.3, leading to a signal width of approximately 0.2%. The model parameter $\beta \in [0, 1]$ modifies the coupling of LQs to leptons, such that the coupling to charged leptons is given by $\sqrt{\beta} \lambda$ and to neutrinos by $\sqrt{1-\beta} \lambda$. It differs from the branching fraction $B$ into charged leptons because of phase-space corrections arising mainly from the large top-quark
The parameter $\beta$ was set to 0.5 in the simulation. Different values for $B$ were achieved by reweighting the MC events according to their decay, as described in ref. [42]. The LQ pair-production cross-sections were obtained from the calculation of direct top-squark pair production, as this process has the same production modes, computed at approximate next-to-next-to-leading order (NNLO) in QCD with resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [61–64]. The cross-sections do not include lepton $t$-channel contributions, which are neglected in ref. [41] and may lead to corrections at the percent level [65]. For this analysis, signal samples were produced separately for electrons and muons and for both up- and down-type LQs with a mass spacing of 100 GeV from 300 GeV to 800 GeV and from 1600 GeV to 2500 GeV, and with a finer spacing of 50 GeV between 800 GeV and 1600 GeV to improve the resolution around the expected mass exclusion limit.

Simulated events with pair-produced up-type vector LQs were generated at leading order (LO) in QCD with MadGraph5_AMC@NLO 2.8.1 and the NNPDF3.0NLO PDF set with $\alpha_s = 0.118$. Decays of the vLQs and top-quarks were handled with MadSpin, while the PS and hadronisation were simulated with Pythia 8.244. The $U_1$ vLQ model in ref. [43] is used for the muon channel, directly seeking an explanation for the various $B$-anomalies. Couplings to electrons are assumed to vanish in this model because of existing tight bounds from low-energy observables, mainly in the lepton-flavour-violating sector. Nevertheless, an extension of the model [66] was used in this analysis to also probe the electron channel. The samples were produced with a coupling strength of $g_U = 3.0$, leading to a signal width of approximately 11%. The large value of $g_U$ is motivated by a suppression of the production cross-section for additional mediators in an ultraviolet-complete model, which might otherwise be in tension with existing LHC limits. The model accommodates both left- and right-handed couplings to fermions, where in the case of only left-handed couplings the coupling strengths to charged leptons and neutrinos are equal ($\beta = 0.5$). Although only left-handed couplings were used in producing the MC samples, the analysis uses the same reweighting of MC events as in the scalar LQ case to also probe different values of $B$. The model allows use of either the minimal (vLQ$_{\text{min}}^\text{YM}$) or the Yang–Mills (vLQ$_{\text{mix}}^\text{YM}$) coupling scenario, where in the latter the vLQ is a heavy gauge boson resulting in enhanced cross-sections. Kinematic differences between the minimal and the Yang–Mills coupling scenario were found to be negligible, except for masses below about 500 GeV. Samples were produced separately for the two coupling scenarios and for both muons and electrons, with a mass spacing of 100 GeV from 300 GeV to 1400 GeV and from 2300 GeV to 2500 GeV, and with a finer spacing of 50 GeV between 1400 GeV and 2300 GeV. No higher-order cross-section computations are available for this model. Therefore, the cross-sections computed at leading order by MadGraph5_AMC@NLO are used in the analysis.

Dominant background processes in the search include $t\bar{t}$, $W$+jets and single top-quark production, the last being mainly associated production of a top-quark and $W$ boson ($tW$). In addition, $t\bar{t}+V$ ($V = W, Z$), diboson, $t\bar{t}+H$, and $Z$+jets processes are also considered in the analysis. Contributions from multi-jet background with a jet misidentified as a lepton are negligible in the phase space of interest.
The production of $t\bar{t}$ events was modelled using the Powheg Box [67–70] v2 generator at NLO with the NNPDF3.0nlo set of PDFs and the $h_{\text{damp}}$ parameter\(^3\) set to 1.5 $m_t$ [71]. The cross-section was corrected to the theory prediction at NNLO including resummation of NNLL soft-gluon terms calculated using Top++ 2.0 [72]. The events were interfaced to Pythia 8.230 to model the PS, hadronisation, and UE.

The associated production of a top-quark and a $W$ boson was modelled using the Powheg Box v2 generator at NLO in QCD using the five-flavour scheme and the NNPDF3.0nlo set of PDFs. The diagram removal scheme [73] was used to avoid overlap with $t\bar{t}$ production because of interference. Single-top $t$-channel ($s$-channel) production was modelled using the Powheg Box v2 generator at NLO in QCD using the four-flavour (five-flavour) scheme and the corresponding NNPDF3.0nlo PDF set. The events were interfaced with Pythia 8.230 in all cases, except for $tW$ events with large missing transverse momenta, which were interfaced with Pythia 8.235.

The production of $V+$jets was simulated with the Sherpa 2.2.1 generator using NLO-accurate matrix elements for up to two jets, and LO-accurate matrix elements for up to four jets, calculated with the Comix [74] and OpenLoops [75, 76] libraries. They were matched with the Sherpa PS [77] using the MEPS@NLO prescription [78–81] using the set of tuned parameters developed by the Sherpa authors. The NNPDF3.0nnlo set of PDFs was used and the samples were normalised to a NNLO prediction [82].

Samples of diboson final states ($VV$) were simulated with the Sherpa 2.2.1 or 2.2.2 generator, depending on the process, including off-shell effects and Higgs boson contributions, where appropriate. Fully leptonic final states and semileptonic final states, where one boson decays leptonically and the other hadronically, were generated using matrix elements at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. The matrix element calculations were matched and merged with the Sherpa PS. The NNPDF3.0nnlo set of PDFs was used, along with the internal Sherpa tune.

The production of $t\bar{t}+W$ and $t\bar{t}+Z$ events was modelled using the MadGraph5\_aMC@NLO 2.3.3 generator at NLO with the NNPDF3.0nlo PDF set. The events were interfaced to Pythia 8.210.

The production of $t\bar{t}+H$ events was modelled using the Powheg Box v2 generator at NLO with the NNPDF3.0nlo PDF set, interfaced to Pythia 8.230.

An overview of the matrix element (ME) generator, PDF, shower generator and UE tune for signal and background samples is given in table 1.

4 Event reconstruction

Events studied in this analysis are required to have at least one reconstructed $pp$ interaction vertex with at least two associated tracks with transverse momentum $p_T > 0.5$ GeV. The primary vertex is selected as the one with the largest sum of squared transverse momenta.

---

\(^3\)The $h_{\text{damp}}$ parameter is a resummation damping factor and one of the parameters that controls the matching of Powheg matrix elements to the PS and thus effectively regulates the high-$p_T$ radiation against which the $t\bar{t}$ system recoils.
Table 1. List of ME generator and the order of the strong coupling constant in the perturbative calculation, PDF, shower generator and tune for the different signal and background processes.

<table>
<thead>
<tr>
<th>Process</th>
<th>ME generator</th>
<th>ME order</th>
<th>PDF set</th>
<th>PS and hadronisation</th>
<th>UE tune</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar LQ</td>
<td>MadGraph5_aMC@NLO 2.6.0</td>
<td>NLO</td>
<td>NNPDF3.0nlo</td>
<td>Pythia 8.230</td>
<td>A14</td>
</tr>
<tr>
<td>Vector LQ</td>
<td>MadGraph5_aMC@NLO 2.8.1</td>
<td>LO</td>
<td>NNPDF3.0nlo</td>
<td>Pythia 8.244</td>
<td>A14</td>
</tr>
<tr>
<td>$t\bar{t}$/single top</td>
<td>Powheg Box v2</td>
<td>NLO</td>
<td>NNPDF3.0nlo</td>
<td>Pythia 8.230/8.235</td>
<td>A14</td>
</tr>
<tr>
<td>$V+$jets</td>
<td>Sherpa 2.2.1</td>
<td>MEPS@NLO</td>
<td>NNPDF3.0nnlo</td>
<td>Sherpa internal</td>
<td></td>
</tr>
<tr>
<td>Diboson</td>
<td>Sherpa 2.2.1/2.2.2</td>
<td>MEPS@NLO</td>
<td>NNPDF3.0nnlo</td>
<td>Sherpa internal</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}+V$</td>
<td>MadGraph5_aMC@NLO 2.3.3</td>
<td>NLO</td>
<td>NNPDF3.0nlo</td>
<td>Pythia 8.210</td>
<td>A14</td>
</tr>
<tr>
<td>$t\bar{t}+H$</td>
<td>Powheg Box v2</td>
<td>NLO</td>
<td>NNPDF3.0nlo</td>
<td>Pythia 8.230</td>
<td>A14</td>
</tr>
</tbody>
</table>

of tracks associated with the interaction vertex. In the analysis, a set of reconstructed objects is used, consisting of electrons, muons, and jets, as well as the missing transverse momentum. When identifying charged leptons, a staggered approach is used, where so-called baseline leptons fulfill less stringent requirements than signal leptons. Events are required to have exactly one signal lepton without any additional baseline leptons.

Electron candidates are reconstructed from energy deposits in the EM calorimeter matched to charged-particle tracks in the inner detector (ID). Requirements of $|d_0|/\sigma_{d_0} < 5$ on the transverse impact parameter $d_0$ (with uncertainty $\sigma_{d_0}$) and $|\Delta z_0 \sin \theta| < 0.5$ mm on the longitudinal track impact parameter $z_0$ ensure matching between track and vertex. Furthermore, electron candidates are required to lie within a pseudorapidity range of $|\eta| < 2.47$, excluding the EM calorimeter barrel–endcap transition region 1.37 < $|\eta| < 1.52$. Baseline electrons must have $p_T > 10$ GeV and fulfill loose identification criteria, using a likelihood-based discriminant that combines information about tracks in the ID and energy deposits in the calorimeter system [83]. In addition, baseline electrons are required to have a hit in the innermost layer of the pixel detector. Isolation requirements in both the calorimeter and the ID are imposed [83]. Electron candidates are rejected if the scalar sum of transverse momenta of tracks within a cone of size $\Delta R = \min(10 \text{ GeV}/p_T, 0.2)$, excluding the electron itself, is larger than 15% of the electron $p_T$. Similarly, an electron is removed if, after subtracting contributions from pile-up and the electron itself, the transverse energy deposited in the calorimeter within a cone of size $\Delta R = 0.2$ exceeds 20% of the transverse energy of the electron. To suppress backgrounds due to hadrons misidentified as electrons, signal electrons must in addition pass the ‘tight’ identification working point, and have $p_T > 30$ GeV.

Muon candidates are reconstructed from charged-particle tracks in the ID and the MS and from energy deposits in the calorimeters. For the reconstruction of muon candidates, tracks in the ID combined with tracks in the MS are used in the range $|\eta| < 2.5$. In addition, muons in the range $|\eta| < 0.1$ are reconstructed from ID tracks matched to an energy deposit in the calorimeter compatible with a minimally ionising particle. These muon candidates are called calorimeter-tagged (CT). Track-to-vertex matching is ensured by requiring $|d_0|/\sigma_{d_0} < 3$ for the transverse impact parameter and $|\Delta z_0 \sin \theta| < 0.5$ mm for the longitudinal track impact parameter. Baseline muons are required to have $p_T > 10$ GeV and to have compatible individual measurements in the ID and the MS. Signal muons are
required to have \( p_T > 30 \text{ GeV} \) and CT muons are not accepted \cite{84}. Additionally, signal muons are rejected if the scalar sum of transverse momenta of tracks within a cone of size \( \Delta R = \min(10 \text{ GeV}/p_T, 0.3) \) around the muon exceeds 6\% of its transverse momentum. In accordance with other searches for pair-produced LQs within ATLAS \cite{37, 39}, signal muons above a \( p_T \) threshold of 800 GeV must fulfil stricter requirements on the number of hits in the MS to ensure good momentum resolution.

Small-radius (small-\( R \)) jet candidates are built from particle-flow objects \cite{85, 86}, using the anti-\( k_t \) algorithm \cite{87, 88} with a radius parameter of \( R = 0.4 \). The particle-flow algorithm combines information about ID tracks and energy deposits in the calorimeters to form the input for jet reconstruction. Jets with \( p_T < 25 \text{ GeV} \) or \( |\eta| > 2.5 \) are rejected. To reduce contributions from pile-up, jet candidates with \( |\eta| < 2.4 \) and \( p_T < 60 \text{ GeV} \) are required to satisfy the ‘tight’ jet vertex tagger criterion \cite{89}. Small-\( R \) jets are categorised as \( b \)-tagged if they satisfy a requirement on the output of a multivariate algorithm, operating at a tagging efficiency of 77\% as determined with simulated \( t\bar{t} \) events \cite{90, 91}. The \( R = 0.4 \) jets are then reclustered with the anti-\( k_t \) algorithm with \( R = 1.0 \) to obtain large-\( R \) jets. Additionally, \( R = 0.4 \) jets are reclustered iteratively with the recursive method described in ref. \cite{92} to reconstruct hadronically decaying top-quark candidates (\( t_\text{had} \)). For this, the small-\( R \) jets are reclustered with an initial radius parameter of \( R = 3.0 \), which is iteratively reduced to \( R(p_T) = 2m_{\text{top}}/p_T \) to match the jet radius to the top candidate’s transverse momentum. Top candidates losing large fractions of their \( p_T \) in the shrinking process are discarded. Finally, only the leading-\( p_T \) candidate with a mass larger than 150 GeV is kept.

The missing transverse momentum (with magnitude \( E_{\text{miss}}^T \)) in an event is defined as the negative vectorial sum of the transverse momenta of all calibrated objects \cite{93}. It also includes an additional track-based soft term taking into account energy depositions not associated with any calibrated object.

An overlap removal procedure is applied to avoid ambiguities when reconstructing the objects described above, using the baseline lepton definitions. Electron–muon overlap is handled by removing any calorimeter-tagged muons sharing a track in the ID with an electron, and then removing any electrons sharing an ID track with a remaining muon. Subsequently, overlap between jets and leptons is removed by rejecting any jets within \( \Delta R = 0.2 \) of an electron and afterwards rejecting any electrons within \( \Delta R = 0.4 \) of a jet. Similarly, jets are discarded if they have fewer than three associated tracks and are within \( \Delta R = 0.2 \) of a muon candidate. Otherwise, the muon is rejected if it lies within \( \Delta R = \min(0.4, 0.04 + 10 \text{ GeV}/p_T(\mu)) \) of a jet.

5 Event selection and categorisation

Any event considered in this analysis must pass an \( E_{\text{miss}}^T \) trigger \cite{94}, as single-lepton triggers are found to be less efficient in particular in the muon channel. The \( E_{\text{miss}}^T \) trigger thresholds varied between 70 GeV and 110 GeV across the different data-taking periods. A requirement of \( E_{\text{miss}}^T > 250 \text{ GeV} \) is imposed in the offline event selection to ensure full efficiency. Events are required to contain exactly one signal lepton. Additionally, a veto is applied on further baseline leptons. Since the final states of interest contain one
Figure 2. Distributions of (a) \( m_{\text{eff}} \) and (b) \( p_T(\ell) \) in the top reweighting region after applying the top reweighting. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of the observed and expected numbers of background events are shown in the bottom panels. The last bin contains the overflow.

hadronically decaying top-quark and one additional jet, only events with at least four small-\( R \) jets are selected. Only one of those jets needs to be \( b \)-tagged to preserve high efficiency for the signal. To suppress contributions from fake \( \slashed{E}_T \) caused by mismeasurements of jets or leptons, events with \( m_T(\ell, \slashed{E}_T) = \sqrt{2p_T(\ell)\slashed{E}_T \cos \Delta \phi(p_T(\ell), \slashed{E}_T)} < 30 \text{ GeV} \) or \( \Delta \phi(\slashed{E}_T, j_{1,2}) < 0.4 \), where \( j_{1,2} \) indicates the highest-\( p_T \) and the second-highest-\( p_T \) small-\( R \) jet, respectively, are rejected.

The \( t\bar{t} \) background is known not to be modelled accurately at high transverse momenta [95, 96]. Reweighting factors are derived in bins of the jet multiplicity as a function of \( m_{\text{eff}} \), which is defined as the scalar sum of the transverse momenta of all reconstructed objects and the \( \slashed{E}_T \) and corresponds approximately to the mass of the LQ pair. This procedure is referred to as ‘top reweighting’ in the following. The reweighting factors are determined for the sum of the \( t\bar{t} \) and single-top backgrounds and are parameterised with a linear function, separately in each of the four jet multiplicity bins (4, 5, 6, \( \geq 7 \)). For this, a dedicated top reweighting region is defined with \( am_{T2} < 200 \text{ GeV} \), where the asymmetric transverse mass, \( am_{T2} \), is a variant of \( m_{T2} \) [97] and allows the reconstruction of dileptonic \( t\bar{t} \) events in which only one lepton is reconstructed [98, 99]. The reweighting factor is then applied to single-top and \( t\bar{t} \) events in each of the training and control regions defined in the following. Figure 2 shows the \( m_{\text{eff}} \) and lepton-\( p_T \) distributions in the top reweighting region after applying the reweighting procedure, in addition to the total background expectation before reweighting. The modelling of kinematic variables is improved by the correction. A potential signal contribution is negligible in this region as it amounts to, e.g., 0.5% in the last four bins in figure 2(a) for a scalar LQ with a mass of 1 TeV and \( B = 0.5 \).
Figure 3. Distributions of $m_{\text{eff}}$ in (a) the $W$+jets CR and (b) the single-top CR after applying the top reweighting, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of the observed and expected numbers of background events are shown in the bottom panels. The last bin contains the overflow.

Control regions (CRs) enriched in the various backgrounds and with negligible signal contamination are defined. They are orthogonal to the top reweighting region by requiring $a m_{T2} > 200$ GeV and are orthogonal to each other. The $W$+jets CR uses events in a window around the Jacobian peak, i.e. $50 \text{ GeV} \leq m_T(\ell, E_T^{\text{miss}}) < 120$ GeV, with a $b$-jet multiplicity of $n_b = 1$ and no hadronically decaying top-quark candidate. To increase the purity of selected $W$+jets events, only events with a positively charged lepton are considered, because the cross-section for $W^+ \text{ production is larger than that for } W^- \text{ production in } pp \text{ collisions.}$ For the single-top CR a requirement of $m_T(\ell, E_T^{\text{miss}}) < 120$ GeV is imposed. In order to reduce contributions from $W$+jets production, events must have exactly two $b$-tagged jets, with an angular separation $\Delta R(b_1, b_2) > 1.2$. Events containing a large-$R$ jet are vetoed. The purity of $W$+jets and single-top events in their respective CRs is $58\%$ and $38\%$. Distributions of $m_{\text{eff}}$ in these CRs are shown in figure 3.

For the training of the NNs, a training region is defined so as to be orthogonal to the top reweighting region and both control regions by requiring $a m_{T2} > 200$ GeV and $m_T(\ell, E_T^{\text{miss}}) \geq 120$ GeV. An overview of the different selection requirements for the training and control regions is given in table 2. The product of signal acceptance and efficiency in the training region is similar for signal hypotheses with couplings to electrons or muons. For $B = 0.5$, it amounts to around 17% for up-type scalar LQs and to around 14% for down-type scalar LQs at $m_{LQ} = 1.4$ TeV. For vector LQs, it reaches 20% at the same mass for both the Yang–Mills coupling and the minimal coupling scenario.
Preselection

\[ E_T^{miss} \text{ triggers} \]

- exactly one signal lepton
- veto on additional baseline leptons
  \[ E_T^{miss} \geq 250 \text{ GeV} \]
- \( \geq 4 \) small-\( R \) jets
- \( m_T(\ell, E_T^{miss}) > 30 \text{ GeV} \)
- \( \Delta\phi(E_T^{miss}, j_{1,2}) > 0.4 \)

<table>
<thead>
<tr>
<th>Top reweighting region</th>
<th>( W^+\text{jets CR} )</th>
<th>Single-top CR</th>
<th>Training region</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_b \geq 1 )</td>
<td>( n_b = 1 )</td>
<td>( n_b = 2 )</td>
<td>( n_b \geq 1 )</td>
</tr>
<tr>
<td>( m_T(\ell, E_T^{miss}) \geq 120 \text{ GeV} )</td>
<td>( 50 \text{ GeV} \leq m_T(\ell, E_T^{miss}) &lt; 120 \text{ GeV} )</td>
<td>( m_T(\ell, E_T^{miss}) &lt; 120 \text{ GeV} )</td>
<td>( m_T(\ell, E_T^{miss}) \geq 120 \text{ GeV} )</td>
</tr>
<tr>
<td>( am_T2 &lt; 200 \text{ GeV} )</td>
<td>( am_T2 &gt; 200 \text{ GeV} )</td>
<td>( am_T2 &gt; 200 \text{ GeV} )</td>
<td>( am_T2 &gt; 200 \text{ GeV} )</td>
</tr>
<tr>
<td>--</td>
<td>( t_{\text{had}} ) candidate veto</td>
<td>large-( R ) jet veto</td>
<td>--</td>
</tr>
<tr>
<td>--</td>
<td>lepton charge = +1e</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>--</td>
<td>--</td>
<td>( \Delta R(b_1,b_2) &gt; 1.2 )</td>
<td>--</td>
</tr>
<tr>
<td>--</td>
<td>--</td>
<td>( NN_{out} &lt; 0.5 / \geq 0.5 )</td>
<td>--</td>
</tr>
</tbody>
</table>

**Table 2.** Overview of event selections applied in the different regions of the analysis.

### 6 Neural network training

Simulated signal and background events in the training region are used to train several NNs for the various signal hypotheses. The NNs are implemented using the NeuroBayes package [100, 101] which combines a three-layer feed-forward NN with a complex and robust preprocessing of the input variables prior to their presentation to the NN. The purpose of the preprocessing is to facilitate optimal network training by ordering the input variables according to their ability to discriminate between signal and background, taking correlations into account, and removing all but the most powerful ones.

NeuroBayes uses Bayesian regularisation techniques for the training process to improve the generalisation performance and to avoid overtraining. In general, the network infrastructure consists of one input node for each input variable plus one bias node, an arbitrary, user-defined number of hidden nodes, and one output node which gives a continuous NN output score \( (NN_{out}) \) in the interval \((0, +1)\), where large values indicate signal-like events and small values background-like events. For the NNs of this analysis, 15 nodes are used in the hidden layer and the ratio of signal to background events in the training is 1:1. The different background processes are weighted according to their expected number of events. Only \( tt \), \( W^+\text{jets} \), single-top-quark, and \( tt+V \) events are used as background processes in the training. As a check for potential overtraining, only 80% of the simulated events serve as input to the training, while the remaining 20% are used as a test sample. No signs of overtraining are observed. After the training step, samples of simulated signal and background events, as well as the observed events, are processed by the NNs in order to get an \( NN_{out} \) value for each event. For each NN, the training region is divided into a low-\( NN_{out} \) control region with \( NN_{out} < 0.5 \), enriched mainly in \( tt \) events, and the signal region above 0.5.
Table 3. Input variables for the NN training, approximately sorted in descending ability to discriminate between signal and background. The order is not absolute as there is some dependence on the signal model and $B$. Some variables might not be defined in every event.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{T}(\ell, E_{T}^{\text{miss}})$</td>
<td>transverse mass of lepton and $E_{T}^{\text{miss}}$</td>
</tr>
<tr>
<td>$m_{\text{eff}}$</td>
<td>scalar sum of the transverse momenta of leptons, jets, and $E_{T}^{\text{miss}}$</td>
</tr>
<tr>
<td>Lepton flavour</td>
<td>flavour of the signal lepton</td>
</tr>
<tr>
<td>$p_{T}(\ell)$</td>
<td>transverse momentum of the lepton</td>
</tr>
<tr>
<td>$m_{\text{inv}}(b_{1}, \ell)$</td>
<td>invariant mass of the leading-$p_{T}$ b-jet and the lepton</td>
</tr>
<tr>
<td>$n_{\text{large}}$</td>
<td>reclustered large-$R$ jet multiplicity</td>
</tr>
<tr>
<td>$a_{\text{NT2}}$</td>
<td>asymmetric transverse mass</td>
</tr>
<tr>
<td>$E_{T}^{\text{miss}}$ significance</td>
<td>measure for assessing the compatibility of the observed $E_{T}^{\text{miss}}$ with zero, taking resolutions into account</td>
</tr>
<tr>
<td>$m_{T}(b_{1}, E_{T}^{\text{miss}})$</td>
<td>transverse mass of leading-$p_{T}$ b-jet and $E_{T}^{\text{miss}}$</td>
</tr>
<tr>
<td>$p_{T}(t_{\text{had}})$</td>
<td>transverse momentum of $t_{\text{had}}$</td>
</tr>
<tr>
<td>$\Delta \phi(E_{T}^{\text{miss}}, b_{2})$</td>
<td>azimuthal angle separation between $E_{T}^{\text{miss}}$ and subleading-$p_{T}$ b-jet</td>
</tr>
<tr>
<td>$m_{\text{inv}}(b_{2}, \ell)$</td>
<td>invariant mass of subleading-$p_{T}$ b-jet and lepton</td>
</tr>
<tr>
<td>$\Delta \phi(E_{T}^{\text{miss}}, b_{1})$</td>
<td>azimuthal angle separation between $E_{T}^{\text{miss}}$ and leading-$p_{T}$ b-jet</td>
</tr>
<tr>
<td>$\Delta \phi(t_{\text{had}}, \ell)$</td>
<td>azimuthal angle separation between $t_{\text{had}}$ and lepton</td>
</tr>
<tr>
<td>$p_{T}(b_{1})$</td>
<td>transverse momentum of leading-$p_{T}$ b-jet</td>
</tr>
</tbody>
</table>

The input variables are chosen because of their ability to discriminate between signal and background. In total, 15 input variables are provided for the training, including the lepton flavour in order to distinguish between electrons and muons. This is because a final state with one lepton flavour has some sensitivity to a LQ model with the other flavour if the lepton stems from a top-quark decay, i.e. mainly in the low $B$ region. Table 3 lists the input variables in order of decreasing ability to discriminate between signal and background. The order is not absolute, as there is some dependence on the signal model and $B$, e.g. the lepton flavour cannot discriminate at all in the region of low $B$, but is important otherwise. The modelling of the input variables in the training region is good in general, as can be seen for the most important ones in figure 4. Among the least well modelled of all input variables is $m_{\text{inv}}(b_{1}, \ell)$ shown in figure 4(c). This is due to the interference between single-top and $tt$ production, which is difficult to describe in MC simulations [102].

Figure 4 also displays three up-type signal hypotheses at a mass of 1.3 TeV. It can be seen that the signal shape depends not only on the value of $B$ but also on the spin of the LQ. The signal shape differences due to spin correlations are sizeable for low values of $B$ at small values of lepton $p_{T}$, where the lepton usually originates from a top-quark decay. Here, the lepton-$p_{T}$ distribution in the vector LQ model is found to be similar in shape to the background in contrast to the scalar case as can be seen figure 4(b). Therefore, separate NNs are trained for scalar and vector LQs as well as for various values of $B$. A total of four NNs per lepton flavour at $B = 0.0, 0.25, 0.5$, and 0.9 are used for up-type
Figure 4. Distributions of (a) $m_{\text{eff}}$, (b) $p_T(\ell)$, (c) $m_{\text{inv}}(b_1, \ell)$, and (d) $m_{T}(\ell, E_{\text{miss}})$ in the training region after applying the top reweighting, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. Signal distributions normalised to the total background expectation are overlaid for up-type scalar LQs with $B = 0.0$ and vector LQs with $B = 0.0$ and 0.5, each with $m_{LQ} = 1.3$ TeV. The ratios of the observed and expected numbers of background events are shown in the bottom panels. The last bin contains the overflow.

scalar and vector LQs. Since kinematic differences between the Yang–Mills coupling and the minimal coupling are negligible except for very low masses, the NNs for vector LQs are only trained on samples with the former coupling and then applied to the latter as well. In the case of down-type scalar LQs, only one NN is trained for LQs decaying into muons and another is trained for LQs decaying into electrons. For both trainings, a branching ratio of 0.5 is assumed, since only events with one LQ decaying into a charged lepton and the other decaying into a neutrino contribute significantly to the phase space under...
consideration. For scalar LQs, signal samples produced with $m_{LQ} = 500\, \text{GeV}, 900\, \text{GeV}, \text{ and } 1300\, \text{GeV}$ are combined in each training, as separate trainings for each signal mass did not yield a significant improvement for masses above about 600 GeV. For vector LQs, additional signal samples with $m_{LQ} = 1700\, \text{GeV}$ are used to take advantage of their higher production cross-sections and therefore higher expected mass limits when compared to scalar LQs.

7 Systematic uncertainties

The largest systematic uncertainties considered in the analysis are related to the modelling of the background processes. For $t\bar{t}$ production, the uncertainty due to the method used to match the matrix element to the parton shower is assessed by comparing a sample produced by MadGraph5_aMC@NLO with the nominal sample from Powheg Box, using the same parton shower. Conversely, an estimate of the uncertainties related to the underlying event, parton shower, and hadronisation is obtained by comparing a sample showered by Herwig 7 [103] with the nominal sample showered by Pythia, using the same ME generator for both. The effects of uncertainties in the renormalisation and factorisation scales are estimated by independently varying the scales by a factor of two. The impact of initial-state radiation (ISR) is estimated by varying $\alpha_s$ in the A14 tune. Similarly, the uncertainty related to final-state radiation (FSR) is assessed by varying the renormalisation scale for final-state parton-shower emissions by a factor of two. Additionally, an uncertainty related to the choice of value for the Powheg Box-specific $h_{\text{damp}}$ parameter is evaluated by using a varied value of $h_{\text{damp}} = 3.0 \, m_t$. PDF uncertainties are obtained from the PDF4LHC15 PDF set [104].

For single top-quark production, uncertainties in ME-to-PS matching, the choice of parton shower, renormalisation and factorisation scales, ISR, FSR, and PDF are evaluated using the same procedures as for $t\bar{t}$ production. Large uncertainties arise due to interference effects between $t\bar{t}$ and $tW$ production. They are estimated by comparing the nominal sample based on the diagram removal scheme with a sample using the diagram subtraction scheme [73, 105].

In the top reweighting procedure, the parameters of the linear fit are varied within their $2\sigma$ uncertainty to account for potential non-linearities in addition to the statistical uncertainty, treating the normalisation and the shape component in each of the four jet multiplicity bins independently. Compared to the $1\sigma$ variation, this choice has a negligible impact on the results.

For $W+\text{jets}$, $Z+\text{jets}$, diboson, $t\bar{t}+V$, and $t\bar{t}+H$ processes, renormalisation and factorisation scale variations are considered, following the same procedures as for $t\bar{t}$. In addition, an uncertainty of 50% is assigned to the heavy-flavour component of the $W+\text{jets}$ background to cover differences in flavour composition between control and signal regions seen in MC studies.

Theoretical systematic uncertainties also include cross-section uncertainties for those background processes for which the normalisation is not determined in the fit. For diboson and $Z+\text{jets}$ production, this uncertainty is taken to be 6% [106] and 5% [107], respectively.
For $t\bar{t}+H$ production, it amounts to 11% [108] and for $t\bar{t}+Z$ production to 15% [108]. For $t\bar{t}+W$ production, the cross-section uncertainty is taken to be 50% to account for potential differences between predicted and measured values as reported in ref. [109].

Systematic uncertainties in the signal prediction arise from acceptance effects due to renormalisation and factorisation scale, ISR/FSR, and PDF and $\alpha_s$ variations. They were found to not exceed 5% in total across the whole mass range, and this is therefore taken as a conservative estimate.

Additionally, detector-related uncertainties are considered, the dominant ones being the small-$R$ jet energy scale and resolution uncertainties [86], Furthermore, systematic uncertainties related to the jet mass scale and resolution, the lepton identification, isolation, and reconstruction efficiencies as well as the lepton energy scale and resolution [83, 84], the $b$-tagging efficiencies [90], and the $E_T^{\text{miss}}$ reconstruction [110] are taken into account. Minor contributions to the total systematic uncertainty also come from the uncertainty of 1.7% in the integrated luminosity [111], obtained using the LUCID-2 detector [112] for the primary luminosity measurements, and from an uncertainty related to pile-up reweighting.

All sources of systematic uncertainty affect the total event yield, and all, except the ones where the size of the uncertainty is explicitly stated above, also affect the shape of the distributions used in the fit.

8 Statistical interpretation

The binned distributions of the NN output are used to test for the presence of a signal. Simultaneous binned profile-likelihood fits are performed for hypothesis testing, following a modified frequentist method implemented in RooStats [113] and using the $NN_{\text{out}}$ distribution in the signal region and the overall number of events in the low-$NN_{\text{out}}$, $W+$jets, and single-top control regions. Systematic uncertainties affecting signal and background expectations are accounted for by including them in the fit in the form of nuisance parameters. For uncertainties in the modelling of background processes for which the normalisation is determined in the likelihood fit, only shape effects and acceptance differences between the CRs and the SR are considered, in order to avoid double-counting of normalisation uncertainties. This procedure is also used for pre-fit uncertainty bands in order to have an equivalent treatment of systematic uncertainties. A smoothing algorithm is applied to certain systematic variations in the signal region in order to reduce statistical fluctuations between bins. To simplify the fitting procedure, for each region and each process a nuisance parameter is only considered if the overall effect on the normalisation of the process is larger than 1%. In the case of the signal region, which has multiple bins, nuisance parameters are also considered if their effect in any bin within the signal region is above 1%.

The binned likelihood function $L(\mu, \theta)$ is constructed as the product of Poisson probability terms over all bins considered in the analysis. It depends on the signal strength parameter $\mu$, a multiplicative factor applied to the theoretical signal production cross-section, and $\theta$, a set of nuisance parameters, implemented in the likelihood function as Gaussian priors for shape effects and as log-normal priors for normalisation effects. The expected number of events in a bin depends on $\mu$ and $\theta$. The nuisance parameters $\theta$ ad-
just the expectations for signal and background according to the corresponding systematic uncertainties, and their fitted values correspond to the amounts that best fit the data.

The test statistic \( q_\mu \) is defined as the profile likelihood ratio

\[
q_\mu = -2 \ln \frac{L(\hat{\mu}, \hat{\theta})}{L(\hat{\mu}, \hat{\theta})},
\]

where \( \hat{\mu} \) and \( \hat{\theta} \) are the values of the parameters that maximise the likelihood function (with the constraints \( 0 \leq \hat{\mu} \leq \mu \)), and \( \hat{\theta} \) are the values of the nuisance parameters (NPs) that maximise the likelihood function for a given value of \( \mu \). This test statistic is used to determine whether the observed data are compatible with the background-only hypothesis, i.e. with \( \mu = 0 \). Furthermore, by using the CL_s method [114], upper limits on the signal production cross-section are derived for each of the signal scenarios considered in this analysis. For a given signal scenario, values of the production cross-section (parameterised by \( \mu \)) yielding \( \text{CL}_s < 0.05 \), where \( \text{CL}_s \) is computed using the asymptotic approximation [115], are excluded at \( \geq 95 \% \) confidence level (CL).

9 Results

For each NN training, a separate fit to the \( NN_{out} \) distribution in the signal region and the overall number of events in the low-\( NN_{out} \), \( W+\text{jets} \), and single-top control regions is performed, with free normalisation parameters for the \( t\bar{t} \), single-top, and \( W+\text{jets} \) background processes. The normalisation parameters obtained from fits to data using the background-only hypothesis are consistent across all trainings. They are always applied in the following and vary between 1.09 ± 0.22 and 1.29 ± 0.23 for \( t\bar{t} \), between 0.84 ± 0.12 and 0.93 ± 0.12 for \( W+\text{jets} \), and between 0.46 ± 0.27 and 0.54 ± 0.26 for single top. The normalisation parameter for the single-top process reduces the event yield by approximately a factor of two; however, the expected yield from the alternative scheme to model the interference between the \( t\bar{t} \) and \( tW \) processes leads to an even smaller yield. Observed and expected event yields after the background-only fit are listed in table 4 for one NN training.

The \( NN_{out} \) distributions after the background-only fit are validated with data–MC comparisons in the control regions, as shown in figure 5 for one particular NN training. In general, good agreement is found for all trainings in all control regions, although the single-top CR is typically more problematic due to the interference effects mentioned above [102]. NPs corresponding to the systematic uncertainties covering the observed differences between data and MC simulation are not constrained significantly because only the overall number of events in each CR enters the fit. The systematic uncertainties are therefore fully propagated to the SR.

A comparison between data and background expectation in the signal region is shown in figure 6 after the background-only fit for three different NN trainings. Good agreement is found for all trainings. The largest discrepancies at high values of \( NN_{out} \) are observed for the LQ^u_{mix} model with \( B = 0.0 \), i.e. for the decay into top-quarks and neutrinos as shown in figure 6(c).
Figure 5. Data and background expectation in (a) the $W$+jets CR, (b) the single-top CR, and (c) the low-NN$_{out}$ CR after the simultaneous background-only fit for a training with $vLQ^{YM} \rightarrow b\mu/t\nu$ and $B = 0.5$. The hatched band indicates the total post-fit uncertainty. The ratios of data to background expectation are shown in the bottom panels.

Figure 6. Data and background expectation in the signal region after the simultaneous background-only fit to data for (a) a training with $vLQ^{YM}_{mix} \rightarrow b\mu/t\nu$ and $B = 0.5$, (b) a training with $LQ^{d}_{mix} \rightarrow t\bar{c}/b\nu$ with $B = 0.5$, and (c) a training with $LQ^{u}_{mix} \rightarrow t\nu$, i.e. $B = 0.0$. Minor background contributions from $t\bar{t}+H$ and $Z$+jets are combined into ‘other’. Expected pre-fit signal distributions with $B$ corresponding to the respective training are added on top of the background expectation, using a mass of 1700 GeV for vector LQs and 1300 GeV for scalar LQs. The hatched band indicates the total post-fit uncertainty. The ratios of data to background expectation are shown in the bottom panels.
Table 4. Observed and expected event yields in the control and signal regions for a training for \( vLQ^{YM} \to b\mu/t\nu \) and \( B = 0.5 \) after the background-only fit. The uncertainties in the background predictions include both the statistical and systematic components. For comparison, expected event yields are shown for a \( vLQ^{YM} \) signal at a mass point of 1700 GeV and \( B = 0.5 \) including its pre-fit uncertainties.

No significant deviations of the data from the expected SM background are observed. Upper 95\% CL limits on the cross-sections of pair-produced LQs can be calculated in simultaneous signal-plus-background fits to the CRs and the SR, in which the background normalisations and possible signal contributions are determined. The largest uncertainty in each of the resulting signal strengths is statistical in nature. For the three signal hypotheses shown in figure 6, the statistical uncertainty exceeds 85\% of the total uncertainty for the two scalar LQ models at \( m_{LQ} = 1.3 \) TeV and rises to nearly 100\% for the vector LQ case at \( m_{LQ} = 1.7 \) TeV. The resulting limits on the cross-section for the four scalar LQ models are shown in figure 7 as a function of the LQ mass for a fixed \( B = 0.5 \). Corresponding limits for the four vector LQ models are shown in figure 8.

These cross-section limits are compared with the theoretical cross-section predictions, shown in blue, resulting in lower limits on the signal mass for \( B = 0.5 \). The uncertainty band around the theory prediction includes PDF, \( \alpha_s \), and renormalisation and factorisation scale uncertainties. The expected and observed limits for \( B = 0.5 \) are summarised in table 5 for the eight LQ models considered in this analysis. The total impact of systematic uncertainties on the cross-section limits reaches 15\% for LQ masses above 1 TeV, corresponding to 20 GeV in the expected mass limit.

Limits on LQ pair-production are also evaluated across a wide range of values for the branching ratio of LQs into charged leptons. For that, the statistical interpretation is performed in steps of 0.05 in \( B \) between 0.0 and 0.95 for up-type scalar and vector LQs and between 0.05 and 0.95 for down-type scalar LQs. For up-type LQs, for which NNs have been trained at four different values of \( B \), the NN resulting in the best expected cross-section limit is chosen at each step. The analysis is not sensitive to final states with zero or two leptons; therefore, \( B = 1.0 \) is omitted for all LQs and so is \( B = 0.0 \) for down-type

<table>
<thead>
<tr>
<th></th>
<th>( W+jets ) CR</th>
<th>( \text{Single-top} ) CR</th>
<th>( \text{Low-}N\text{N}_{\text{out}} ) CR</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t\bar{t} )</td>
<td>860 ± 140</td>
<td>186 ± 35</td>
<td>1370 ± 150</td>
<td>53 ± 10</td>
</tr>
<tr>
<td>Single top</td>
<td>103 ± 87</td>
<td>131 ± 47</td>
<td>200 ± 110</td>
<td>36 ± 14</td>
</tr>
<tr>
<td>( W+jets )</td>
<td>1240 ± 130</td>
<td>101 ± 28</td>
<td>265 ± 55</td>
<td>32.4 ± 6.9</td>
</tr>
<tr>
<td>( t\bar{t}+V )</td>
<td>11.0 ± 1.8</td>
<td>4.47 ± 0.79</td>
<td>180 ± 28</td>
<td>16.7 ± 2.6</td>
</tr>
<tr>
<td>Diboson</td>
<td>94.3 ± 9.8</td>
<td>7.6 ± 1.9</td>
<td>94 ± 11</td>
<td>11.1 ± 1.2</td>
</tr>
<tr>
<td>( t\bar{t}+H )</td>
<td>1.27 ± 0.16</td>
<td>1.00 ± 0.12</td>
<td>14.4 ± 1.7</td>
<td>1.34 ± 0.18</td>
</tr>
<tr>
<td>( Z+jets )</td>
<td>6.46 ± 0.32</td>
<td>2.18 ± 0.11</td>
<td>7.20 ± 0.36</td>
<td>1.31 ± 0.07</td>
</tr>
<tr>
<td>Total background</td>
<td>2308 ± 48</td>
<td>433 ± 21</td>
<td>2126 ± 46</td>
<td>152 ± 13</td>
</tr>
<tr>
<td>Observed events</td>
<td>2310</td>
<td>430</td>
<td>2124</td>
<td>157</td>
</tr>
<tr>
<td>( vLQ^{YM}_{\text{mix}}(1.7 \text{TeV}, B = 0.5) )</td>
<td>0.109 ± 0.022</td>
<td>0.097 ± 0.016</td>
<td>1.57 ± 0.10</td>
<td>38.9 ± 2.6</td>
</tr>
</tbody>
</table>
Figure 7. Expected (dashed black) and observed (solid black) 95% CL upper limits on the cross-section of pair-produced scalar LQs, assuming $B = 0.5$. The green (yellow) band shows the $\pm 1\sigma (\pm 2\sigma)$ uncertainty region around the expected limit. The theoretical prediction and its $\pm 1\sigma$ uncertainty band are shown in blue. Limits are presented for (a) up-type scalar LQs decaying into muons, (b) up-type scalar LQs decaying into electrons, (c) down-type scalar LQs decaying into muons, and (d) down-type scalar LQs decaying into electrons.

<table>
<thead>
<tr>
<th>Exp. limit [GeV]</th>
<th>Obs. limit [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LQ_{u}^{\text{mix}} \rightarrow t\nu/b\mu$</td>
<td>$1440^{+50}_{-60}$</td>
</tr>
<tr>
<td>$LQ_{u}^{\text{mix}} \rightarrow t\nu/bc$</td>
<td>$1440^{+50}_{-60}$</td>
</tr>
<tr>
<td>$LQ_{d}^{\text{mix}} \rightarrow t\mu/b\nu$</td>
<td>$1380^{+50}_{-60}$</td>
</tr>
<tr>
<td>$LQ_{d}^{\text{mix}} \rightarrow t\tau/b\tau$</td>
<td>$1410^{+50}_{-60}$</td>
</tr>
<tr>
<td>$vLQ_{\text{YM}}^{\text{mix}} \rightarrow t\nu/b\mu$</td>
<td>$1930^{+50}_{-60}$</td>
</tr>
<tr>
<td>$vLQ_{\text{YM}}^{\text{mix}} \rightarrow t\nu/bc$</td>
<td>$1930^{+50}_{-60}$</td>
</tr>
<tr>
<td>$vLQ_{\text{min}}^{\text{mix}} \rightarrow t\nu/b\mu$</td>
<td>$1660^{+50}_{-60}$</td>
</tr>
<tr>
<td>$vLQ_{\text{min}}^{\text{mix}} \rightarrow t\nu/bc$</td>
<td>$1650^{+50}_{-60}$</td>
</tr>
</tbody>
</table>

Table 5. Expected and observed 95% CL lower limits on the LQ mass at $B = 0.5$ for the eight signal hypotheses considered in this analysis.
LQs. The cross-section upper limits and the mass exclusion curves across the $B$ plane are shown in figure 9 for scalar LQs and in figure 10 for vector LQs.

Expected and observed limits on the leptoquark mass as a function of $B$ agree well everywhere, except for a small deviation for the $LQ_{\text{mix}}^{u}$ model at $B = 0.0$, as already discussed above in the context of figure 6. Differences between up- and down-type LQs can be observed especially for low values of $B$, where events with $LQLQ \rightarrow t_{\text{lep}} \nu_{\text{had}} \nu$ increase the sensitivity. The shapes of the exclusion limits for vector LQs with the Yang–Mills and the minimal coupling are very similar, since no significant kinematic differences exist at these higher masses. When comparing the limits for up-type scalar LQs with those for vector LQs, the kinematic differences due to spin correlations in $LQLQ \rightarrow t_{\text{lep}} \nu_{\text{had}} \nu$ become relevant, i.e. when approaching $B = 0.0$, the expected lower limit on the mass decreases faster for the vector LQ, as expected from figure 4(b).
Figure 9. Expected (solid white, ±1σ ranges dashed) and observed (solid orange) exclusion limits on the leptoquark mass as a function of the branching ratio into charged leptons at 95% CL. The observed upper limit on the signal cross-section in each bin is shown on the z-axis. Limits are presented for (a) up-type scalar LQs decaying into muons, (b) up-type scalar LQs decaying into electrons, (c) down-type scalar LQs decaying into muons, and (d) down-type scalar LQs decaying into electrons. For up-type LQs the range in $B$ is 0–0.95, for down-type it is 0.05–0.95.

10 Conclusion

Results of a search for pair-produced scalar and vector leptoquarks decaying into quarks of the third generation and charged or neutral leptons of the first or second generation are presented, targeting the single-lepton final state. The analysis is based on data collected by the ATLAS experiment in $\sqrt{s} = 13$ TeV proton–proton collisions, corresponding to an integrated luminosity of 139 fb$^{-1}$. Several neural networks are trained for various signal hypotheses, covering a wide range of parameters. No significant deviations from the Standard Model expectation are observed and upper limits on the production cross-section are derived for eight models as a function of leptoquark mass and branching ratio into the charged lepton.

In addition, lower limits on the leptoquark mass are set across a range of branching ratios for all models. At a branching ratio of 0.5 they reach values of 1460 GeV (1440 GeV)
Figure 10. Expected (solid white, ±1σ ranges dashed) and observed (solid orange) exclusion limits on the leptoquark mass as a function of the branching ratio into charged leptons at 95% CL. The observed upper limit on the signal cross-section in each bin is shown on the z-axis. Limits are presented for (a) vector LQs in the Yang–Mills coupling scenario decaying into muons, (b) vector LQs in the Yang–Mills coupling scenario decaying into electrons, (c) vector LQs in the minimal coupling scenario decaying into muons, and (d) vector LQs in the minimal coupling scenario decaying into electrons.

for up-type scalar leptoquarks decaying into muons (electrons) and 1370 GeV (1390 GeV) for down-type scalar leptoquarks decaying into muons (electrons). For the first time, dedicated neural networks are used to search for $U_1$ vector leptoquarks. At a branching ratio of 0.5 the resulting lower limits on the mass for the decay into muons are 1980 GeV and 1710 GeV for the Yang–Mills and the minimal coupling scenario, respectively. The decay into electrons is also probed and limits of 1900 GeV (1620 GeV) for the Yang–Mills (minimal) coupling scenario are derived.

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; BMWFW and FWF, Austria; ANAS, Azerbaijan; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN: ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MEiN, Poland; FCT, Portugal; MNE/IFA, Romania; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DSI/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TENMAK, Türkiye; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; PRIMUS 21/SCI/017 and UNCE SCI/013, Czech Republic; COST, ERC, ERDF, Horizon 2020 and Marie Sklodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Harkleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and MINERVA, Israel; Norwegian Financial Mechanism 2014-2021, Norway; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, 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), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [116].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

References


[31] LHCb collaboration, *Measurement of lepton universality parameters in $B^+ \to K^+ℓ^+ℓ^−$ and $B^0 \to K^0ℓ^+ℓ^−$ decays*, arXiv:2212.09153 [SPIRE].


---


[100] M. Feindt, A neural bayesian estimator for conditional probability densities, physics/0402093.


[102] ATLAS collaboration, Studies of $t\bar{t}/tW$ interference effects in $b\bar{b}\ell^+\ell^-'\nu\bar{\nu}'$ final states with POWHEG and MadGraph5_AMC@NLO setups, ATL-PHYS-PUB-2021-042 (2021).


Gaziantep University, Gaziantep; (c) Department of Physics, Istanbul University, Istanbul; (d) Istinye University, Sarıyer, Istanbul; Türkiye

22 (c) Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá; (b) Departamento de Física, Universidad Nacional de Colombia, Bogotá; Colombia

23 (a) Dipartimento di Fisica e Astronomia A. Righi, Università di Bologna, Bologna; (b) INFN Sezione di Bologna; Italy

24 Physikalisches Institut, Universität Bonn, Bonn; Germany

25 Department of Physics, Boston University, Boston MA; United States of America

26 Department of Physics, Brandeis University, Waltham MA; United States of America

27 (a) Transilvania University of Brasov, Brasov; (b) Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest; (c) Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi; (d) National Institute for Research and Development of Isotopic and Molecular Technologies, Physics Department, Cluj-Napoca; (e) University Politehnica Bucharest, Bucharest; (f) West University in Timisoara, Timisoara; Romania

28 (a) Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Košice; Slovak Republic

29 Physics Department, Brookhaven National Laboratory, Upton NY; United States of America

30 Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Física, y CONICET, Instituto de Física de Buenos Aires (IFIBA), Buenos Aires; Argentina

31 California State University, CA; United States of America

32 Cavendish Laboratory, University of Cambridge, Cambridge; United Kingdom

33 (a) Department of Physics, University of Cape Town, Cape Town; (b) Theba Labs, Western Cape; (c) Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg; (d) National Institute of Physics, University of the Philippines Diliman (Philippines); (e) University of South Africa, Department of Physics, Pretoria; (f) University of Zululand, KwaDlangezwa; (g) School of Physics, University of the Witwatersrand, Johannesburg; South Africa

34 Department of Physics, Carleton University, Ottawa ON; Canada

35 (a) Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies — Université Hassan II, Casablanca; (b) Faculté des Sciences, Université Ibn-Tofail, Kénitra; (c) Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech; (d) LPMR, Faculté des Sciences, Université Mohamed Premier, Oujda; (e) Faculté des sciences, Université Mohammed V, Rabat; (f) Institute of Applied Physics, Mohammed VI Polytechnic University, Ben Guerir; Morocco

36 CERN, Geneva; Switzerland

37 Affiliated with an institute covered by a cooperation agreement with CERN

38 Affiliated with an international laboratory covered by a cooperation agreement with CERN

39 Enrico Fermi Institute, University of Chicago, Chicago IL; United States of America

40 LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand; France

41 Nevis Laboratory, Columbia University, Irvington NY; United States of America

42 Niels Bohr Institute, University of Copenhagen, Copenhagen; Denmark

43 (a) Dipartimento di Fisica, Università della Calabria, Rende; (b) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati; Italy

44 Physics Department, Southern Methodist University, Dallas TX; United States of America

45 Physics Department, University of Texas at Dallas, Richardson TX; United States of America

46 National Centre for Scientific Research "Demokritos", Agia Paraskevi; Greece

47 (a) Department of Physics, Stockholm University; (b) Oskar Klein Centre, Stockholm; Sweden

48 Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen; Germany

49 Fakultät Physik, Technische Universität Dortmund, Dortmund; Germany

50 Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden; Germany

51 Department of Physics, Duke University, Durham NC; United States of America

52 SUFA — School of Physics and Astronomy, University of Edinburgh, Edinburgh; United Kingdom

53 INFN e Laboratori Nazionali di Frascati, Frascati; Italy
Also Affiliated with an institute covered by a cooperation agreement with CERN
b Also at Borough of Manhattan Community College, City University of New York, New York NY; United States of America
c Also at Bruno Kessler Foundation, Trento; Italy
d Also at Center for High Energy Physics, Peking University; China
e Also at Centro Studi e Ricerche Enrico Fermi; Italy
f Also at CERN, Geneva; Switzerland
g Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland
h Also at Departament de Física de la Universitat Autònoma de Barcelona, Barcelona; Spain
i Also at Department of Financial and Management Engineering, University of the Aegean, Chios; Greece
j Also at Department of Physics and Astronomy, Michigan State University, East Lansing MI; United States of America
k Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY; United States of America
l Also at Department of Physics, Ben Gurion University of the Negev, Beer Sheva; Israel
m Also at Department of Physics, California State University, East Bay; United States of America
n Also at Department of Physics, California State University, Sacramento; United States of America
o Also at Department of Physics, King’s College London, London; United Kingdom
p Also at Department of Physics, University of Fribourg, Fribourg; Switzerland
q Also at Department of Physics, University of Thessaly; Greece
r Also at Department of Physics, Westmont College, Santa Barbara; United States of America
s Also at Hellenic Open University, Patras; Greece
t Also at Institut Català de Recerca i Estudis Avançats, ICREA, Barcelona; Spain
u Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg; Germany
v Also at Institute of Particle Physics (IPP); Canada
w Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku; Azerbaijan
x Also at Institute of Theoretical Physics, Ilia State University, Tbilisi; Georgia
y Also at Lawrence Livermore National Laboratory, Livermore; United States of America
z Also at Physics Department, An-Najah National University, Nablus; Palestine
aa Also at The City College of New York, New York NY; United States of America
ab Also at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing; China
ac Also at TRIUMF, Vancouver BC; Canada
ad Also at Università di Napoli Parthenope, Napoli; Italy
ae Also at University of Chinese Academy of Sciences (UCAS), Beijing; China
af Also at University of Colorado Boulder, Department of Physics, Colorado; United States of America
ag Also at Yeditepe University, Physics Department, Istanbul; Türkiye

* Deceased