Search for lepton-flavor violation in different-flavor, high-mass final states in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

M. Aaboud et al. (ATLAS Collaboration)

(Received 18 July 2018; published 13 November 2018)

A search is performed for a heavy particle decaying into different-flavor, dilepton pairs ($ee$, $\mu\mu$, or $\mu\tau$), using 36.1 fb$^{-1}$ of proton-proton collision data at $\sqrt{s} = 13$ TeV collected in 2015–2016 by the ATLAS detector at the Large Hadron Collider. No excesses over the Standard Model predictions are observed. Bayesian lower limits at the 95% credibility level are placed on the mass of a $Z'$ boson, the mass of a supersymmetric $\tau$-sneutrino, and on the threshold mass for quantum black-hole production. For the $Z'$ and sneutrino models, upper cross-section limits are converted to upper limits on couplings, which are compared with similar limits from low-energy experiments and which are more stringent for the $\mu\tau$ modes.

DOI: 10.1103/PhysRevD.98.092008

I. INTRODUCTION

Lepton flavor violation (LFV) is forbidden in the Standard Model (SM) of particle physics but is allowed in many extensions of the SM. Many such models predict new particles with LFV decays, such as $Z'$ bosons [1], scalar neutrinos in $R$-parity-violating (RPV) [2,3] supersymmetry (SUSY) and quantum black holes (QBH) in low-scale gravity [4]. Processes with flavor-violating dilepton decays are expected to produce pairs of prompt, different-flavor leptons, a final state with a clear experimental signature and a low background from SM processes. The $Z/\gamma^* \to \ell\ell$ process is an irreducible background for same-flavor lepton searches but in different-flavor searches is limited to the production and decay of a $\tau\tau$ pair. This paper describes a search for new phenomena in final states with two leptons of different flavor using 36.1 fb$^{-1}$ of data from proton-proton ($pp$) collisions at $\sqrt{s} = 13$ TeV at the Large Hadron Collider (LHC).

A common extension of the SM is the addition of an extra $U(1)$ gauge symmetry resulting in a massive neutral vector boson known as a $Z'$ boson [1]. The search presented in this paper assumes a $Z'$ boson with the same quark couplings and chiral structure as the SM $Z$ boson but only leptonic decays that violate lepton flavor conservation are allowed. The parameter $Q_{ij}$, where $i,j = 1...3$ represent the three lepton generations, gives the strength of the LFV couplings relative to the SM $\ell\ell$ couplings. The ATLAS Collaboration placed lower limits of 3.3, 2.9, and 2.7 TeV on the mass of a $Z'$ boson decaying into $ee$, $\mu\mu$, and $\mu\tau$ pairs, respectively, using 3.2 fb$^{-1}$ of the 13 TeV data [5], while the CMS Collaboration has placed limits up to 4.4 TeV on a $Z'$ boson decaying into an $ee$ final state using 35.9 fb$^{-1}$ [6]. Following the same methodology as in Ref. [5], this paper assumes only one LFV coupling is different from zero at any time for the purpose of setting limits on the cross section times branching of each final state considered. Polarization of $\tau$-leptons is not included in the model, but its impact on the $\tau$-lepton acceptance is found to be negligible, and it does not impact the sensitivity to a possible signal.

In RPV SUSY, the superpotential terms allowing LFV are expressed as $\frac{1}{2} \lambda_{ijk} L_i E^c_j + \lambda^\prime_{ijk} L_i Q_j D^c_k$, where $L$ and $Q$ are the $SU(2)$ doublet superfields of leptons and quarks, $E$ and $D$ are the $SU(2)$ singlet superfields of charged leptons and down-type quarks, $\lambda$ and $\lambda^\prime$ are Yukawa couplings, and the indices $i,j,$ and $k$ denote generations. A $\tau$-sneutrino ($\tilde{\nu}_\tau$) may be produced in $pp$ collisions by $d\bar{d}$ annihilation and subsequently decay into $ee$, $\mu\mu$, or $\mu\tau$. Although only $\tilde{\nu}_e$ is considered in this paper, results apply to any sneutrino flavor. For the theoretical prediction of the cross section times branching ratio, the $\tilde{\nu}_e$ coupling to first-generation quarks ($\lambda_{311}^\prime$) is assumed to be 0.11 for all channels. As in the $Z'$ model, each lepton-flavor-violating final state is considered separately. It is assumed that $\lambda_{312}^\prime = \lambda_{321}^\prime = 0.07$ for the $ee$ final state, $\lambda_{313}^\prime = \lambda_{331}^\prime = 0.07$ for the $\mu\mu$ final state, and $\lambda_{323}^\prime = \lambda_{332}^\prime = 0.07$ for the $\mu\tau$ final state. These values are chosen such that the signal width is narrow, and allow for comparisons with previous ATLAS and CMS searches [5,7,8]. The CMS Collaboration has recently excluded $R$-parity-violating supersymmetric models below 1.7 TeV for $\lambda_{132}^\prime = \lambda_{231}^\prime = \lambda_{311}^\prime = 0.01$ [6].

---

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3.
Various models introduce extra spatial dimensions to reduce the value of the Planck mass and resolve the hierarchy problem. The search described in this paper presents interpretations based on the Arkani-Hamed–Dimopoulos–Dvali (ADD) model [9], assuming \( n = 6 \), where \( n \) is the number of extra dimensions, and on the Randall-Sundrum (RS) model [10] with one extra dimension. Due to the increased strength of gravity at short distances in these models, \( pp \) collisions at the LHC could produce states exceeding the threshold mass (\( m_\text{th} \)) and form black holes. For the models considered, \( m_\text{th} \) is assumed to be equivalent to the extra-dimensional Planck scale. For masses beyond 3–5 \( m_\text{th} \), it is expected that thermal black holes would be produced [11,12], characterized by high-multiplicity final states. The search presented in this paper focuses on the mass region below 3–5 \( m_\text{th} \), known as the quantum gravity regime [13–15], where production of nonthermal (or quantum) black holes is expected and these black holes could decay into two-particle final states, producing the topology investigated in this paper. Nonthermal quantum black holes would have a continuum mass distribution from \( m_\text{th} \) up to the beginning of the thermal regime. For the models considered in this paper, the thermal regime is assumed to start at 3\( m_\text{th} \). The decay of quantum black holes would be governed by a yet unknown theory of quantum gravity. The two main assumptions of the extra-dimensions models considered [4] in this paper are (a) gravity couples with equal strength to all SM particle degrees of freedom and (b) gravity conserves local symmetries (color and electric charge) but can violate global symmetries such as lepton-flavor and baryon-number conservation. Following these assumptions, the branching ratio to each final state is calculated. QBHs decaying into different-flavor, opposite-charge lepton pairs are created via \( q\bar{q} \) or \( gg \) annihilation. The branching ratio to \( \ell^+\ell^- \) is 0.87\% (0.34\%) for a \( q\bar{q} \) (\( gg \)) initial state [4]. These models were used in previous ATLAS and CMS searches for quantum black holes in dijet [16–18], lepton + jet [19], photon + jet [20], \( e\mu [6] \), and same-flavor dilepton [21] final states.

II. THE ATLAS DETECTOR

The ATLAS detector [22] is a general-purpose particle detector with approximately forward-backward symmetric cylindrical geometry.\(^1\) It is composed of four main components, each responsible for identifying and reconstructing different types of particles: the inner detector (ID), the electromagnetic and hadronic calorimeters, and the muon spectrometer (MS). Each of the subdetectors is divided into two components, barrel and end cap, to provide coverage close to 4\( \pi \) in solid angle. In addition, two magnet systems allow charge and momentum measurements: an axial magnetic field of 2.0 T provided by a solenoid surrounding the ID and a toroidal magnetic field for the MS. The ID, the closest component to the interaction point, is used to reconstruct the trajectories of charged particles in the region \( |\eta| < 2.5 \) and measure their momenta. It is composed of three subsystems: (1) the silicon pixel detector, including an additional inner layer at a radius of 3.2 cm added in 2015 [23,24], (2) the semiconductor tracker, used in conjunction with the silicon pixel detector to determine primary and secondary vertices with high precision thanks to their high granularity, (3) the transition radiation tracker, providing additional tracking in the region \( |\eta| < 2.0 \) and electron identification through the detection of transition radiation x-ray photons.

The calorimeter system covers the pseudorapidity range \( |\eta| < 4.9 \). Within the region \( |\eta| < 3.2 \), electromagnetic calorimetry is provided by barrel and end cap high-granularity lead/liquid-argon (LAr) electromagnetic calorimeters, with an additional thin LAr presampler covering \( |\eta| < 1.8 \), to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillating-tile calorimeter, segmented into three barrel structures within \( |\eta| < 1.7 \), and two copper/LAr hadronic end cap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimized for electromagnetic and hadronic measurements, respectively.

Surrounding the calorimeter system, the MS is the subdetector furthest from the interaction point. It consists of three layers of precision tracking chambers and fast detectors for triggering on muons. Tracking coverage is provided for \( |\eta| < 2.7 \) by three layers of precision drift tube chambers, with cathode strip chambers in the innermost layer for \( |\eta| > 2.0 \), while trigger coverage is provided by resistive plate and thin gap chambers for \( |\eta| < 2.4 \).

The trigger and data-acquisition system is based on two levels of online event selection [25]: the level-1 trigger and the high-level trigger. The level-1 trigger is hardware based and uses a subset of detector information to provide quick trigger decisions and reduce the accepted rate to 100 kHz. The high-level trigger is software based and exploits the full detector information to further reduce the acceptance rate to about one kHz.

III. DATA AND SIMULATED SAMPLES

The data sample used for this analysis was collected during 2015 and 2016 from \( pp \) collisions at a center-of-mass energy of 13 TeV. After selecting periods with stable
beams and applying data-quality requirements, the total integrated luminosity is 36.1 fb$^{-1}$ with an uncertainty of 2.1%, derived following a methodology similar to that detailed in Ref. [26] from a calibration of the luminosity scale using $x$-$y$ beam-separation scans.

The $pp \rightarrow Z' \rightarrow \ell^+\ell^-$ samples were generated at leading order (LO) using the generator PYTHIA 8.186 [27] with the NNPDF23LO [28] parton distribution function (PDF) set and the A14 [29] set of tuned parameters. Signal samples for 25 mass points ranging from 0.5 to 5 TeV were generated in 0.1 TeV steps from 0.5 to 2 TeV, 0.2 TeV steps from 2 to 3 TeV, and 0.5 TeV steps from 3 to 5 TeV. The production cross section was calculated with the same generator used for simulation. The cross section and signal shape in the dilepton invariant mass was calculated with the same generator used for simulation. The cross section was calculated at LO (order (LO)) using the generator PYTHIA 8.186 [27] with the NNPDF23LO PDF set and the A14 tune. The signal samples were generated at NLO using the POWHEG-BOX [36–38] generator (v2 for $t\bar{t}$ and v1 for single-top) with the CT10 [39] PDF set used in the matrix-element calculations. PYTHIA 6.4.28 [40] and the corresponding Perugia 2012 tune [41] were used to simulate the parton shower, hadronization, and the underlying event. Top quarks were decayed using MADSpin [42], preserving all spin correlations. The $h_{\text{damp}}$ parameter, which controls the $p_T$ of the first emission beyond the Born configuration in POWHEG-BOX, was set to the mass of the top quark. The main effect of this parameter is to regulate the high-$p_T$ emission against which the $t\bar{t}$ system recoils. The mass was set to the top quark mass of 172.5 GeV. The EVTGEN 1.2.0 program [43] was used for the properties of $b$- and $c$-hadron decays. A value of $831^{+20}_{-29}$ (scale) $^{+35}_{-33}$ (PDF + $\alpha_S$) $^{+23}_{-22}$ (mass uncertainty) pb is used for the $\prod$ production cross section, computed with Top++ [44], incorporating NNLO QCD corrections, including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms. A $W\ell$ production cross section of 71.7 ± 3.8 pb is used, as computed in Ref. [45] to approximately NNLO (NNLL + NLO) accuracy.

Diboson processes producing at least two charged leptons were simulated using the SHERPA 2.2.2 generator [46]. The matrix elements contain all diagrams with four electroweak vertices. Fully leptonic decays were calculated for up to one parton (four leptons, or two leptons and two neutrinos) or zero partons (three leptons and one neutrino) at NLO and up to three partons at LO using the COMIX [47] and OpenLOOPS [48] matrix-element generators and merged with the SHERPA parton shower [49] using the ME+PS@NLO prescription [50]. The CT10 PDF set was used in conjunction with the default parton shower tuning provided by the authors of SHERPA. Inclusive cross-section values of 1.28, 4.51, and 10.64 pb are used for ZZ, WZ, and WW production, respectively.

Events containing $W$ or $Z$ bosons are generated using POWHEG-Box v2 interfaced to the PYTHIA 8.186 parton shower model. The CT10 PDF set is used in the matrix element. The AZNLO set of tuned parameters [51] is used, with PDF set CTEQ6L1, for the modeling of nonperturbative effects. The EVTGEN 1.2.0 program is used for the properties of $b$- and $c$-hadron decays. PHOTOS++ [52] is used for QED emissions from electroweak vertices and charged leptons. The $W/Z$ samples are normalized with the NNLO cross sections. This background contribution is normalized to an inclusive cross section of 1.9 nb, calculated for $m_{\ell\ell} > 60$ GeV.

Processes such as $W$ + jets and multijet production with jets that are misidentified as leptons were estimated through a combination of data-driven methods and simulation,
IV. EVENT RECONSTRUCTION AND SELECTION

This search is optimized to look for new phenomena in the high mass range. Events are selected if they satisfy a single-muon or electron trigger with a $p_T$ threshold of 50 GeV for muons and 60 or 120 GeV for electrons. The single-electron trigger with higher $p_T$ threshold has a looser identification requirement, resulting in an increased trigger efficiency at high $p_T$.

Electron candidates are formed by associating the energy in clusters of cells in the electromagnetic calorimeter with a track in the ID [56]. A likelihood discriminant suppresses contributions from hadronic jets, photon conversions, Dalitz decays, and semileptonic heavy-flavor hadron. The likelihood discriminant utilizes lateral and longitudinal calorimeter shower shapes plus tracking and cluster-track matching quantities. The discriminant criterion is a function of the $p_T$ and $|\eta|$ of the electron candidate. Two operating points are used in this analysis, as defined in Ref. [57]: medium and tight. The tight working point (85% efficient at $p_T = 65$ GeV determined with $Z \rightarrow ee$ events) is required for electron candidates. Electron candidates must have $p_T > 65$ GeV and $|\eta| < 2.47$, excluding the region $1.37 < |\eta| < 1.52$, where the energy reconstruction performance is degraded due to the presence of extra inactive material. Further requirements are made on the transverse and longitudinal impact parameters of the track, which is the distance between the $z$-position of the point of closest approach of the track in the ID to the beam line and the $z$-coordinate of the primary vertex relative to the primary vertex of the event ($d_0$ and $\Delta z_0$). The requirements are the following: $|d_0/\sigma_d| < 5$ and $|\Delta z_0 \sin \theta| < 0.5$ mm. Candidates are required to satisfy relative track-based (as defined above for muon candidates) and calorimeter-based isolation requirements with an efficiency of 99% to suppress background from nonprompt electrons originating from heavy-flavor semileptonic decays, charged hadrons, and photon conversions from $\pi^0$ decays. The sum of the calorimeter transverse energy deposits (excluding the electron itself) in an isolation cone of size $\Delta R = 0.2$ divided by the electron $p_T$ is the discriminant used in the calorimeter-based isolation criterion. For the reducible background estimation, electron candidates passing the medium working point (95% efficient at $p_T = 65$ GeV determined with $Z/\gamma' \rightarrow ee$ events) are referred to as “loose electrons.”

Candidate muon tracks are initially reconstructed independently in the ID and the MS. The two tracks are input to a combined fit which takes into account the energy loss in the calorimeter and multiple scattering. Muon identification is based on information from both the ID and MS to ensure that muons are reconstructed with the optimal momentum resolution up to very high $p_T$ using the high-$p_T$ operating point [58]. Only tracks with hits in each of the three stations of the muon spectrometer are considered. This provides a muon $p_T$ resolution of about 10% at 1 TeV. Moreover, muon candidates are required to be within the ID acceptance region of $|\eta| < 2.5$, fulfill $|d_0/\sigma_d| < 3$ and $|\Delta z_0 \sin \theta| < 0.5$ mm, have a $p_T$ larger than 65 GeV, and fulfill a track-based isolation criterion with an efficiency of 99% over the full range of muon momenta to further reduce contamination from non-prompt muons. The scalar sum of the transverse momenta of tracks (excluding the muon itself) in an isolation cone of size $\Delta R = 0.2$ divided by the muon $p_T$ is the discriminant used in the track-based isolation criterion. For the reducible background estimation, muon candidates fulfilling all selection criteria except the isolation criterion are called “loose muons.”

Jets are reconstructed using the anti-$k_t$ algorithm [59] with a radius parameter of 0.4 using energy clusters [60] of calorimeter cells as input. Jet calibrations [61] derived from $\sqrt{s} = 13$ TeV simulated data and from collision data taken at 13 TeV are used to correct the jet energies and directions to those of the particles from the hard-scatter interaction.

Hadronic decays of $\tau$-leptons are composed of a neutrino and a set of visible decay products ($r_{\text{had-vis}}$), typically one or three charged pions and up to two neutral pions. The reconstruction of $\tau$-leptons and their visible hadronic decay products starts with jets reconstructed from topological clusters [62]. The $r_{\text{had-vis}}$ candidates must have $|\eta| < 2.5$ with the transition region between the barrel and end cap calorimeters $(1.37 < |\eta| < 1.52)$ excluded, a $p_T$ greater than 65 GeV, and one or three associated tracks with $\pm 1$ total electric charge. Their identification is performed using a multivariate algorithm that employs boosted decision trees (BDT) using shower shape and tracking information to discriminate against jets. All $r_{\text{had-vis}}$ candidates are...
required to fulfill the “loose” identification requirements of Ref. [63], with an efficiency of 60% (50%) for 1(3)-prong \( \tau \) had-vis. An additional dedicated likelihood-based veto is used to reduce the number of electrons misidentified as \( \tau \) had-vis candidates. The jet to tau fake rate is around 25% (5%) for 1(3)-prong \( \tau \) had-vis decays. For the purpose of the reducible background estimation, \( \tau \) had-vis candidates failing the loose identification requirements are used.

Jets containing \( b \)-hadrons (\( b \)-jets) are identified with a \( b \)-tagging algorithm based on a multivariate technique [64]. Operating points are defined by a single value in the domain of discriminant outputs and are chosen to provide a

<table>
<thead>
<tr>
<th>Object selection</th>
<th>Lepton-pair charges</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_1 )</td>
<td>Nonisolated ( e/\mu ) &amp; ( \tau ) had-vis failing ( \tau ) ID requirements (( p_T^e, p_T^\tau &lt; 200 ) GeV)</td>
</tr>
<tr>
<td>( R_2 )</td>
<td>Isolated ( e/\mu ) &amp; pass ID ( \tau ) had-vis (( p_T^e, p_T^\tau &lt; 200 ) GeV)</td>
</tr>
<tr>
<td>( R_3 )</td>
<td>Nonisolated ( e/\mu ) &amp; ( \tau ) had-vis failing ( \tau ) ID requirements</td>
</tr>
</tbody>
</table>

For the purpose of the reducible background estimation, \( \tau \) had-vis candidates failing the loose identification requirements are used. Jets containing \( b \)-hadrons (\( b \)-jets) are identified with a \( b \)-tagging algorithm based on a multivariate technique [64]. Operating points are defined by a single value in the domain of discriminant outputs and are chosen to provide a

FIG. 1. Distributions in the \( W + \) jets-enriched control region for the \( e\tau \) channel: (a) electron and (b) \( \tau \)-lepton transverse momentum, (c) the \( e\tau \) invariant mass, and (d) the jet multiplicity. No further data points are found in overflow bins. Uncertainty refers to the combination in quadrature of all the statistical and systematic uncertainties considered.
specific $b$-jet efficiency in an inclusive $\bar{t}t$ sample. The employed working point has an efficiency of 77\% and rejection factors of 6 and 134 for charm and light-quark/gluon jets, respectively [64].

Only events with exactly two different-flavor leptons are chosen. As such, there is no overlap between the three channels considered: $e\mu, e\tau, \mu\tau$. They must have a reconstructed primary vertex, defined as the vertex whose constituent tracks have the highest sum of $p_T^2$, and exactly two reconstructed different-flavor lepton candidates meeting the above-mentioned criteria. Events with an additional electron, muon or $\tau_{\text{had-vis}}$ are vetoed. For the $e\mu$ channel only, events with an extra electron or muon fulfilling the loose criteria are also vetoed, including events used for the purpose of the reducible background estimation. For all three channels, the lepton candidates must be back to back in the transverse plane with $\Delta\phi(l, l') > 2.7$. The invariant mass of the dilepton pair is used as the discriminant. No requirement is made on the respective charges of the leptons since it reduces the signal efficiency by as much as 6\% for the highest-mass signals considered due to charge misassignment without a significant effect on the background rejection. To account for differences between data and simulation, corrections are applied to the lepton trigger, reconstruction, identification, and isolation efficiencies as well as the lepton energy/momentum resolution and scale [56–58,63].

FIG. 2. Distributions in the $W$+jets-enriched control region for the $\mu\tau$ channel: (a) the muon and (b) $\tau$-lepton transverse momentum, (c) the $\mu\tau$ invariant mass, and (d) the jet multiplicity. No further data points are found in overflow bins. Uncertainty refers to the combination in quadrature of all the statistical and systematic uncertainties considered.
Double-counting of leptons and jets is avoided by applying an overlap removal algorithm based on the $\Delta R$ distance metric. First, jets within $\Delta R < 0.2$ of any identified isolated muons or electrons are removed. Then, any muons and electrons within $0.2 < \Delta R < 0.4$ from the jet axis are removed.

The missing transverse momentum vector ($E_T^{\text{miss}}$) is defined as the negative vector $p_T$ sum of all identified physics objects and an additional soft term. The soft term is constructed from all tracks that are associated with the primary vertex, but not with any selected physics object. In this way, the $E_T^{\text{miss}}$ incorporates the best calibration of the jets and the other identified physics objects, while maintaining pileup independence in the soft term [65].

The dominant background for the $e\mu$ channel is $t\bar{t}$ production, which can be suppressed by rejecting events that contain one or more $b$-jets ($b$-veto).

<table>
<thead>
<tr>
<th>Source</th>
<th>1 TeV</th>
<th>2 TeV</th>
<th>3 TeV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$e\mu$</td>
<td>$e\mu$</td>
<td>$e\mu$</td>
</tr>
<tr>
<td></td>
<td>$b$-jet veto</td>
<td>$b$-jet veto</td>
<td>$b$-jet veto</td>
</tr>
<tr>
<td>Luminosity</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Top-quark extrapolation</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Top scale</td>
<td>7</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>PDF</td>
<td>16</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Pile up</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dilepton $p_T$ modeling</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Electron iden. and meas.</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Muon iden. and meas.</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$\tau$ iden. and meas.</td>
<td>...</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$\tau$ reconstruction eff.</td>
<td>...</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$\tau$ fake rate</td>
<td>...</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Multijet transf. factor</td>
<td>...</td>
<td>31</td>
<td>2</td>
</tr>
<tr>
<td>Reducible $e\mu$ estimation</td>
<td>2</td>
<td>...</td>
<td>2</td>
</tr>
<tr>
<td>Jet eff. and resol.</td>
<td>1</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>...</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$ resol. and scale</td>
<td>...</td>
<td>...</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>20</td>
<td>37</td>
</tr>
</tbody>
</table>

The “...” sign indicates that the systematic uncertainty is not applicable.
For a $Z'$ boson with a mass of 1.5 TeV, the fractions of events that pass all of the selection requirements are approximately 45%, 45%, 20%, and 15% for the $e\mu$, $\mu\mu$ with $b$-veto, $e\tau$, and $\mu\tau$ final states, respectively.

For the reducible background estimation in the $e\tau$ and $\mu\tau$ channels, the transverse mass $m_T$ of a lepton and the missing transverse momentum is defined as

$$m_T = \sqrt{2p_T E_T^{\text{miss}}[1 - \cos \Delta \phi(\ell', E_T^{\text{miss}})],}$$

where $p_T$ is the transverse momentum of the lepton, $E_T^{\text{miss}}$ is the magnitude of the missing transverse momentum vector, and $\Delta \phi(\ell', E_T^{\text{miss}})$ is the azimuthal angle between the lepton and $E_T^{\text{miss}}$ directions.

For the dilepton mass calculation in the $e\tau$ and $\mu\tau$ channels, the missing momentum from the neutrino in the hadronic decay of a $\tau$-lepton is estimated and added to the four-momentum of the $\tau_{\text{had-vis}}$ candidate. At the considered momenta, the hadronic decay of the $\tau$-lepton results in the neutrino and the resultant jet being nearly collinear. The neutrino four-momentum is reconstructed from the magnitude of the $E_T^{\text{miss}}$ and the direction of the $\tau_{\text{had-vis}}$ candidate. This technique significantly improves sensitivity by a factor of 2 [7]. For a simulated $Z'$ boson with a mass of 2 TeV, the mass resolution improves from 8% (17%) to 4% (12%) in the $e\tau$ ($\mu\tau$) channel.

V. BACKGROUND ESTIMATION

The background processes for this search can be divided into two categories: reducible and irreducible. The latter is composed of processes which produce two different-flavor prompt leptons, including $Z/\gamma' \rightarrow \tau\tau$, $t\bar{t}$, single-top, and diboson production. These processes are modeled using simulated samples and normalized to their theoretically predicted cross sections. Reducible backgrounds originate from jets misreconstructed as leptons and require the use of data-driven techniques. The contribution from reducible backgrounds is small in the $e\mu$ channel, about 5%, whereas in the $e\tau$ and $\mu\tau$ channels they are the leading components mostly due to jets faking hadronic taus and make up 50%–60% of the total background.

### A. Top-quark background extrapolation

The simulated samples used to estimate single-top-quark and $t\bar{t}$ production are statistically limited for dilepton invariant masses above 1 TeV. Therefore, for $m_{ee} > 700$ GeV, the $t\bar{t}$

<table>
<thead>
<tr>
<th>Process</th>
<th>$2000 &lt; m_{ee} &lt; 3000$ GeV</th>
<th>$m_{ee} &gt; 3000$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>$0.28 \pm 0.09 \pm 0.32$</td>
<td>$(0.16 \pm 0.08 \pm 0.28) \times 10^{-1}$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$0.25 \pm 0.10 \pm 0.11$</td>
<td>$(0.44 \pm 0.01 \pm 0.56) \times 10^{-2}$</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>$&lt; 0.01$</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>$Z/\gamma' \rightarrow \ell\ell$</td>
<td>$(0.48 \pm 0.03 \pm 0.23) \times 10^{-2}$</td>
<td>$(0.16 \pm 0.02 \pm 0.31) \times 10^{-3}$</td>
</tr>
<tr>
<td>Total background</td>
<td>$0.54 \pm 0.13 \pm 0.41$</td>
<td>$(0.21 \pm 0.08 \pm 0.30) \times 10^{-1}$</td>
</tr>
<tr>
<td>Data</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
plus single-top contributions are evaluated using monotonically decreasing functions fitted to the $m_{ee'}$ distribution. Two functional forms are chosen for their stability when varying the fit range and for the quality of the fit:

$$a \cdot m_{ee'}^b \cdot m_{ee'}^{c \ln(m_{ee'})} \quad \text{and} \quad \frac{a}{(m_{ee'} + b)c},$$

where $a$, $b$ and $c$ are free parameters in the fit. To account for fit variations, the lower and upper limits of the fit range were varied in 25 GeV steps between 200–300 GeV and 1000–1200 (800–1000) GeV for $e\mu$ ($e\tau$ and $\mu\tau$). The nominal extrapolation is taken as the average of all the tested fit ranges using both functional forms. The extrapolation is found to agree within statistical uncertainties with the simulated data. For each mass bin, the up and down variations obtained from varying the fit parameters are combined in quadrature with the uncertainty of the fit range variation. This uncertainty is 32% at 2 TeV for the $e\mu$ channel.

**B. Reducible background**

The main reducible backgrounds are $W +$ jets and multijet production. The contribution to the reducible background from muons originating from decays of hadrons in jets is found to be negligible compared with the contribution from fake electrons and $\tau$-leptons. Therefore, in the $e\mu$ channel, where reducible backgrounds are a small contribution to the total, nonprompt muons are neglected and only events with one prompt muon and a jet faking an electron are considered. In channels involving taus, however, reducible backgrounds are more significant, and so contributions from both electrons and muons, primarily nonprompt leptons from heavy flavor decays, are taken into account. However, the dominant source of reducible background in these channels remains fake taus from quark-initiated jets.

**1. $e\mu$ channel: Matrix method**

For the $e\mu$ channel, the matrix method is employed, as detailed in Ref. [21]. The selection criteria are loosened for electron candidates to create a sample of events with a muon and a loose electron as defined in Sec. IV. These events are referred to as loose, while those in which both the electron and muon pass all selections are “tight.” The probability of a loose electron matched to a generated electron to pass the full object selection (the “real efficiency”) is evaluated from $Z \to ee$ simulated events, while the probability that a jet is misidentified as an electron (the “fake rate”) is obtained in a multijet-enriched data

---

**TABLE V.** Expected and observed numbers of $e\mu$ events in the (a) low- and (b) high-mass regions after applying all selection criteria including the $b$-jet veto. The statistical and systematic uncertainties are quoted.

<table>
<thead>
<tr>
<th>Process</th>
<th>$m_{ee'} &lt; 300$ GeV</th>
<th>300 &lt; $m_{ee'}$ &lt; 600 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>1660 ± 20 ± 260</td>
<td>570 ± 10 ± 100</td>
</tr>
<tr>
<td>Diboson</td>
<td>1470 ± 20 ± 130</td>
<td>479 ± 8 ± 55</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>231 ± 18 ± 87</td>
<td>87 ± 8 ± 33</td>
</tr>
<tr>
<td>$Z/\gamma' \to \ell\ell$</td>
<td>86 ± 6 ± 12</td>
<td>18.4 ± 1.3 ± 3.0</td>
</tr>
<tr>
<td>Total background</td>
<td>3450 ± 30 ± 350</td>
<td>1150 ± 20 ± 150</td>
</tr>
<tr>
<td>Data</td>
<td>3411</td>
<td>1082</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>600 &lt; $m_{ee'}$ &lt; 1200 GeV</th>
<th>1200 &lt; $m_{ee'}$ &lt; 2000 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>28.6 ± 1.7 ± 8.9</td>
<td>0.72 ± 0.10 ± 0.85</td>
</tr>
<tr>
<td>Diboson</td>
<td>45.9 ± 1.2 ± 7.7</td>
<td>2.85 ± 0.30 ± 0.76</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>11.0 ± 2.8 ± 4.3</td>
<td>0.1 ± 1.9 ± 0.0</td>
</tr>
<tr>
<td>$Z/\gamma' \to \ell\ell$</td>
<td>1.27 ± 0.06 ± 0.25</td>
<td>(0.70 ± 0.05 ± 0.20) × 10^{-1}</td>
</tr>
<tr>
<td>Total background</td>
<td>87 ± 3 ± 15</td>
<td>3.7 ± 2.0 ± 1.1</td>
</tr>
<tr>
<td>Data</td>
<td>83</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>2000 &lt; $m_{ee'}$ &lt; 3000 GeV</th>
<th>$m_{ee'} &gt; 3000$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>$(2.8 \pm 0.8 \pm 5.5) \times 10^{-2}$</td>
<td>$(0.8 \pm 0.4 \pm 2.3) \times 10^{-3}$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$0.25 \pm 0.10 \pm 0.11$</td>
<td>$(0.42 \pm 0.01 \pm 0.51) \times 10^{-2}$</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>$&lt; 0.001$</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>$Z/\gamma' \to \ell\ell$</td>
<td>$(0.46 \pm 0.03 \pm 0.23) \times 10^{-2}$</td>
<td>$(0.14 \pm 0.02 \pm 0.30) \times 10^{-3}$</td>
</tr>
<tr>
<td>Total background</td>
<td>$0.28 \pm 0.10 \pm 0.14$</td>
<td>$(0.52 \pm 0.04 \pm 0.60) \times 10^{-2}$</td>
</tr>
<tr>
<td>Data</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
sample. To suppress the $W +$ jets contribution to the multijet-enriched sample, its events are required to have $E_T^{miss} < 25$ GeV and $m_T < 50$ GeV as well as to pass the signal region selection outlined in Sec. IV. Both the real efficiency and the fake rate are determined as a function of $p_T$, and they are used to estimate the reducible background contribution. Residual contaminations from $W +$ jets and other SM background processes (top, diboson, and $Z \to \ell\ell$) in the multijet CR are subtracted using simulation. This background is estimated up to around 1.5 TeV, where there are no data. However, at this stage, the expectation of this background is well below one event, and generally negligible compared to diboson and top quark processes.

The uncertainties associated with the matrix method are evaluated by considering systematic effects on the electron fake rate. Uncertainties of the real electron efficiency have a negligible impact on the estimation and are not considered. The systematic uncertainties in the fake rate include:

1. the choice of multijet-enriched region,
2. uncertainties on the Monte Carlo subtraction in the multijet-enriched region, and
3. the difference in the fake rates obtained using this method and those obtained from simulated $W +$ jets events.

The overall uncertainty of the $e\mu$ reducible background is about 30%. Given that in the $e\mu$ channel this contribution is about 7% of the total background over the invariant mass range considered, the uncertainties in the estimation method have a small impact on the results.

2. $e\tau$ and $\mu\tau$ channels: $W +$ jets estimate

The dominant background for the $e\tau$ and $\mu\tau$ channels is the $W +$ jets process, where a jet is misidentified as a $r_{\text{had-vis}}$ candidate. It is estimated using simulated events with each jet weighted by its probability to pass the $\tau$-lepton identification as measured in data. This not only ensures the correct fake rate but also improves the statistical precision of the estimate, since events failing the $\tau$-lepton identification requirements are not discarded. The $r_{\text{had-vis}}$ fake rate is measured in a $W \to e/\mu +$ jets control region as a function of the $p_T$, $\eta$, and number of tracks of the $r_{\text{had-vis}}$ candidates. The $W +$ jets-enriched control region uses the same selection as the signal region, but reverses the back-to-back criterion to $\Delta \phi(\ell, \ell') < 2.7$, and uses $\tau$-leptons fulfilling all requirements except identification, although a minimum requirement on the BDT discriminant is retained. Only events with exactly one electron or muon fulfilling all selection criteria, as well as $m_T > 80$ GeV to enrich the $W +$ jets contribution, are used.

---

<table>
<thead>
<tr>
<th>TABLE VI: Expected and observed numbers of $e\tau$ events in the (a) low- and (b) high-mass regions after applying all selection criteria. The statistical and systematic uncertainties are quoted.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) $N_{\text{events}}$ with $m_{\tau\tau} &lt; 600$ GeV</strong></td>
</tr>
<tr>
<td>Process</td>
</tr>
<tr>
<td>Top</td>
</tr>
<tr>
<td>Diboson</td>
</tr>
<tr>
<td>Multijet and $W +$ jets</td>
</tr>
<tr>
<td>$Z/\gamma^* \to \ell\ell$</td>
</tr>
<tr>
<td>Total background</td>
</tr>
<tr>
<td>Data</td>
</tr>
<tr>
<td><strong>(b) $N_{\text{events}}$ with $m_{\tau\tau} &gt; 600$ GeV</strong></td>
</tr>
<tr>
<td>Process</td>
</tr>
<tr>
<td>Top</td>
</tr>
<tr>
<td>Diboson</td>
</tr>
<tr>
<td>Multijet and $W +$ jets</td>
</tr>
<tr>
<td>$Z/\gamma^* \to \ell\ell$</td>
</tr>
<tr>
<td>Total background</td>
</tr>
<tr>
<td>Data</td>
</tr>
<tr>
<td>Process</td>
</tr>
<tr>
<td>Top</td>
</tr>
<tr>
<td>Diboson</td>
</tr>
<tr>
<td>Multijet and $W +$ jets</td>
</tr>
<tr>
<td>$Z/\gamma^* \to \ell\ell$</td>
</tr>
<tr>
<td>Total background</td>
</tr>
<tr>
<td>Data</td>
</tr>
</tbody>
</table>
Events where the invariant mass of the e or μ and the \( \tau_{\text{had-vis}} \) candidate is between 80 and 110 GeV are vetoed to reduce contamination from Z boson decays. Contributions from non-W + jets processes are subtracted using simulation. The \( \tau_{\text{had-vis}} \) candidates present in the remaining events are dominated by jets. The contribution of events with nonprompt electrons is estimated from simulation and found to be less than 1%. The \( \tau_{\text{had-vis}} \) fake rate is defined as the fraction of \( \tau_{\text{had-vis}} \) candidates in the sample that also pass the \( \tau_{\text{had-vis}} \) identification. This rate is used to weight simulated \( W + \) jets events. The resulting distribution obtained for the \( W + \) jets is validated in the \( W + \) jets-enriched control region, where good agreement is found between data and the expected SM background processes.

The uncertainties in the \( \tau_{\text{had-vis}} \) fake rate are evaluated from

1. the modeling of the loose \( \tau_{\text{had-vis}} \) identification requirement in simulation,

2. the statistical uncertainty of the data-driven estimation of the \( \tau \)-lepton fake rate, and

3. the differences in \( \tau \)-lepton fake rate between signal and control regions.

These errors are detailed in the following paragraphs.

The \( \tau_{\text{had-vis}} \) fake rate is reevaluated when removing the \( m_T > 80 \) GeV requirement to check the contamination from non-\( W + \) jets processes. The effect on the fake rate and the final estimation of the \( W + \) jets background is about 2%.

The statistical uncertainty of the fake rate in the control regions is propagated through the estimate. The impact is small at low \( m_{\ell\ell} \) but is the leading uncertainty of the fake rate in the range \( m_{\ell\ell} > 1 \) TeV.

The jet composition of the fake \( \tau_{\text{had-vis}} \) background is evaluated from simulated \( W + \) jets events. The control region where the \( \tau_{\text{had-vis}} \) fake rate is evaluated should have a jet composition similar to that in the signal region. Therefore, \( W + \) jets simulated events are used to investigate the difference between the fake rates measured in the \( W + \) jets control and signal regions. The comparison reveals a slightly higher fake rate in the signal region, consistent with the lower expected gluon contribution. The relative difference between these fake rates is assigned as a systematic uncertainty, which contributes an uncertainty of about 8% to the total background at \( m_{\ell\ell} = 1 \) TeV.

### 3. eτ and \( \mu \tau \) channels: Multijet estimate

The multijet background contributions in the e\( \tau \) and \( \mu \tau \) channels are evaluated using events in three control regions \( (R_1, R_2, R_3) \). The events must pass the selection for the signal region, except that in \( R_1 \) and \( R_3 \) the electron/muon

<table>
<thead>
<tr>
<th>Process</th>
<th>( m_{\ell\ell} &lt; 300 ) GeV</th>
<th>( 300 &lt; m_{\ell\ell} &lt; 600 ) GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>1380 ± 20 ± 300</td>
<td>1160 ± 20 ± 250</td>
</tr>
<tr>
<td>Diboson</td>
<td>318 ± 8 ± 55</td>
<td>225 ± 6 ± 42</td>
</tr>
<tr>
<td>Multijet and ( W + ) jets</td>
<td>6900 ± 200 ± 1400</td>
<td>1650 ± 50 ± 380</td>
</tr>
<tr>
<td>( Z/\gamma^* \to \ell\ell )</td>
<td>1650 ± 40 ± 270</td>
<td>339 ± 14 ± 71</td>
</tr>
<tr>
<td>Total background</td>
<td>10300 ± 200 ± 1700</td>
<td>3380 ± 60 ± 550</td>
</tr>
<tr>
<td>Data</td>
<td>10525</td>
<td>3378</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>( 600 &lt; m_{\ell\ell} &lt; 1200 ) GeV</th>
<th>( 1200 &lt; m_{\ell\ell} &lt; 2000 ) GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>95 ± 1 ± 26</td>
<td>3.5 ± 0.2 ± 2.7</td>
</tr>
<tr>
<td>Diboson</td>
<td>33.3 ± 1.7 ± 9.2</td>
<td>2.9 ± 0.5 ± 1.5</td>
</tr>
<tr>
<td>Multijet and ( W + ) jets</td>
<td>140 ± 10 ± 43</td>
<td>6.4 ± 1.0 ± 2.4</td>
</tr>
<tr>
<td>( Z/\gamma^* \to \ell\ell )</td>
<td>14.4 ± 1.2 ± 4.3</td>
<td>0.88 ± 0.07 ± 0.32</td>
</tr>
<tr>
<td>Total background</td>
<td>282 ± 10 ± 61</td>
<td>13.7 ± 1.1 ± 5.0</td>
</tr>
<tr>
<td>Data</td>
<td>255</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>( 2000 &lt; m_{\ell\ell} &lt; 3000 ) GeV</th>
<th>( m_{\ell\ell} &gt; 3000 ) GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>0.17 ± 0.03 ± 0.21</td>
<td>(0.56 ± 0.19 ± 0.94) \times 10^{-2}</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.54 ± 0.30 ± 0.38</td>
<td>(0.62 ± 0.09 ± 0.95) \times 10^{-2}</td>
</tr>
<tr>
<td>Multijet and ( W + ) jets</td>
<td>0.87 ± 0.35 ± 0.89</td>
<td>(0.87 ± 0.70 ± 0.60) \times 10^{-2}</td>
</tr>
<tr>
<td>( Z/\gamma^* \to \ell\ell )</td>
<td>(0.78 ± 0.03 ± 0.42) \times 10^{-1}</td>
<td>(0.63 ± 0.04 ± 0.84) \times 10^{-2}</td>
</tr>
<tr>
<td>Total background</td>
<td>1.65 ± 0.46 ± 1.30</td>
<td>(0.27 ± 0.07 ± 0.33) \times 10^{-1}</td>
</tr>
<tr>
<td>Data</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
must fail isolation and the $\tau_{\text{had-vis}}$ candidate must fail identification, and in $R_1$ and $R_2$ the leptons must have the same electric charge and the electron/muon $p_T$ must be less than 200 GeV to avoid signal contamination. For a $Z_0$ boson with a mass of 500 GeV, the lowest signal mass considered in this paper, the contamination from the signal process in $R_2$ is found to be below 1%. The region definitions are listed in Table I. The background contribution is estimated as $N_{MJ} = N_{R_3} \times N_{R_2}/N_{R_1}$. The transfer factor $N_{R_2}/N_{R_1}$ is calculated as a function of the dilepton mass to encapsulate correlations between $m_{\tau\tau}$ and the isolation and identification requirements, and it is fitted with a polynomial. In each of the regions defined, the contributions from other SM processes, such as $W +$ jets, $Z +$ jets, $Z/\gamma' \rightarrow \ell\ell'$, diboson, and top-quark production, are subtracted using simulation. The contribution from the multijet background is $\sim60\%$ ($\sim20\%$) of the $W +$ jets background for the $e\tau$ ($\mu\tau$) channel, corresponding to $\sim25\%$ ($\sim10\%$) of the total expected background.

The multijet background is estimated using a transfer factor obtained using same-charge lepton pairs and applied to opposite-charge plus same-charge lepton pairs. To check the validity of this procedure, the multijet background is also estimated using a transfer factor obtained with
opposite-charge pairs. The difference between the resulting shape and transfer factors is assigned as a systematic uncertainty. The impact of this uncertainty is about 7% at 1 TeV.

The statistical uncertainties in the $m_{\ell^+\ell^-}$-dependent transfer factor and the subtraction of simulated events are propagated to the final estimate and assigned as a systematic uncertainty. The overall uncertainty is 50% (15%) at 1 TeV for the $e\tau$ ($\mu\tau$) channel. The uncertainty in the $\mu\tau$ channel is smaller because the transfer factor is found to have a negligible effect on the dilepton invariant mass, and the transfer-factor fit uncertainties are reduced.

C. Reducible background validation

The validity of the background estimation is checked in the $W +$ jets control region. Figures 1 and 2 show the electron, muon, and $\tau$-lepton $p_T$, $\ell\tau$ invariant mass and jet multiplicity distributions for the $e\tau$ and $\mu\tau$ channels, respectively, in the $W +$ jets control region. Good agreement is observed between the data and the background prediction. The contribution from each SM background for each of the final states in the signal region is given in Sec.VII.

VI. SYSTEMATIC UNCERTAINTEES

The sources of experimental uncertainty considered are pileup effects; lepton efficiencies due to triggering, identification, reconstruction, isolation, energy scale, and resolution [56–58,63,66]; jet energy scale and resolution [61]; $b$-tagging [64]; and $E_T^{miss}$ [65]. Sources of uncertainty are considered for both the simulated background and signal processes.
Mismodeling of the muon momentum resolution at the TeV scale from residual misalignment of the muon precision chambers can alter the signal and background shapes. A corresponding uncertainty is obtained from studies performed in dedicated data-taking periods with no magnetic field in the MS. The muon reconstruction efficiency is affected at high $p_T$ by possible large energy losses in the calorimeter. The associated uncertainty is estimated by comparing studies of $Z \rightarrow \mu \mu$ data events extrapolated to high $p_T$ with the results predicted by simulation [67]. The effect on the muon reconstruction efficiency was found to be approximately 3% per TeV as a function of muon $p_T$.

The uncertainty of the electron identification efficiency extrapolation is determined from the differences in the electron shower shapes in the EM calorimeters between data and simulation in the $Z \rightarrow ee$ peak, which are propagated to the high $p_T$ electron sample. The effect on the electron identification efficiency is 2% and is independent of $p_T$ for electrons with $p_T > 150$ GeV [67].

The treatment of systematic uncertainties for $\tau$-leptons with $p_T$ up to 100 GeV is detailed in Ref. [62]. An additional uncertainty of 20% per TeV is assigned to the reconstruction efficiency of $\tau$-leptons with $p_T > 100$ GeV to account for the degradation of the modeling and reconstruction efficiency from track merging, derived from studies in simulation and in dijet data events at 8 TeV [68].

The $E_{\text{miss}}$ uncertainty is derived from the uncertainties of the momenta of physics objects and uncertainties of the soft term determined by comparisons with simulation.

A mismodeling of the dilepton $p_T$ variable is found in the $\tau\tau$ simulation. After reweighting to data in a $\tau\tau$ control region, an uncertainty is assigned to account for the effect on the dilepton invariant mass spectrum.

M. AABOUD et al. PHYS. REV. D 98, 092008 (2018)
A pile-up modeling uncertainty is estimated by varying the distribution of pile-up events in the reweighting of the Monte Carlo, to cover the uncertainty on the ratio between the predicted and measured inelastic cross section in the fiducial volume defined by \( M_X > 13 \text{ GeV} \) where \( M_X \) is the mass of the hadronic system [69].

The uncertainty of 2.1% in the luminosity applies to the signal and to backgrounds derived from simulations.

The uncertainties of the reducible background estimation in the \( e\mu \) channel, and the \( \tau \)-lepton fake rate, the multijet transfer factor calculation, and the top-quark extrapolation are presented in Sec. V.

The PDF uncertainties are the dominant systematic uncertainties affecting the background estimates, together with the uncertainty on the extrapolation to estimate the expected limits shown with the ±1 and ±2 standard deviation uncertainty bands.

### Table VIII

<table>
<thead>
<tr>
<th>Model (b-veto)</th>
<th>Expected limit [TeV]</th>
<th>Observed limit [TeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>LFV ( Z' )</td>
<td>4.3 4.3 3.7 3.5</td>
<td>4.5 4.4 3.7 3.5</td>
</tr>
<tr>
<td>RPV SUSY ( \tilde{\nu}_\tau )</td>
<td>3.4 3.4 2.9 2.6</td>
<td>3.4 3.4 2.9 2.6</td>
</tr>
<tr>
<td>QBH ADD ( n = 6 )</td>
<td>5.6 5.5 4.9 4.5</td>
<td>5.6 5.5 4.9 4.5</td>
</tr>
<tr>
<td>QBH RS ( n = 1 )</td>
<td>3.3 3.4 2.8 2.7</td>
<td>3.4 3.4 2.9 2.6</td>
</tr>
</tbody>
</table>

FIG. 6. The observed and expected 95% credibility-level upper limits on the (a) \( Z^0 \) boson, (b) \( \tau \)-sneutrino (\( \tilde{\nu}_\tau \)), and (c) QBH ADD and RS production cross section times branching ratio for decays into an \( \mu \tau \) final state. The signal theoretical cross section times branching ratio lines for the \( Z^0 \) model, the QBH ADD model assuming six extra dimensions, and the RS model with one extra dimension are obtained from the simulation of each process, while the RPV SUSY \( \tilde{\nu}_\tau \) includes the NLO \( K \)-factor calculated using Looptools [33]. The acceptance times efficiency of the ADD and RS QBH models agree within 1%, and the same curve is used for limit extraction. The expected limits are shown with the ±1 and ±2 standard deviation uncertainty bands.
The contribution from PDF uncertainties is estimated using different PDF sets and eigenvector variations within a particular PDF set for the top-quark, diboson, and $W^+\text{jets}$ backgrounds. The CT10 PDF uncertainty due to eigenvector variations is evaluated through the use of LHAPDF \[70\] following the prescriptions of Ref.\[71\]. The uncertainty related to the choice of PDF is evaluated by comparing the results with those from the central value of other PDF sets: MMHT2014\[72\], NNPDF3.0 \[73\], and CT14 \[31\]. PDF-related uncertainties in the signal shape are not considered. The uncertainties of the $m_{Z'}$ modeling in $t\bar{t}$ events are obtained using separate simulated samples generated with the renormalization scale and $h_{\text{damp}}$ parameter varied by factors of 2 and $1/2$, and are referred to as"Top scale" in Table II. These uncertainties for $W + \text{jets}$ are not considered as they are found to be small, given that this background is mainly composed of real lepton ($\epsilon$ or $\mu$) and fake $\tau$ pairs. For the diboson background prediction, the PDF systematic is the leading uncertainty.

Experimental systematic uncertainties common to signal and background processes are assumed to be correlated. The systematic uncertainties of the estimated SM background and signal yields are summarized in Tables II and III. For signal processes, only experimental systematic uncertainties are considered. The simulated samples contribute a 3% statistical uncertainty to the overall signal acceptance times efficiency.

VII. RESULTS

Tables IV–VII show the expected and observed numbers of events in the low and high mass regions for each channel. The $e\mu$ background is dominated by $t\bar{t}$ and diboson events, while $W + \text{jets}$ events are dominant for the $e\tau$ and $\mu\tau$ final states.

Figure 3 shows the dilepton invariant mass distributions for the $e\mu$, $e\mu$ with a $b$-veto, $e\tau$, and $\mu\tau$ channels. The largest deviation found in the data is a deficit in the 1.1–1.4 TeV range of the $e\mu$ channel, with a global significance of 1.8 standard deviations, obtained using the BUMPHER program \[74\]. Due to the parton luminosity tail in the LFV $Z'$ model, the impact of the deficit in the 1.1–1.4 TeV range is also seen in the observed limit for $Z'$ boson masses up to 4–5 TeV. No significant excess is found in any channel.

The electron-muon event with an invariant mass of 2.1 TeV found in the previous version of this analysis \[5\] no longer satisfies the event selection, since the previously selected muon candidate is found to overlap....
with a jet using the criteria of this paper and is no longer classified as a prompt muon.

Since no deviations from the SM prediction are observed, model-dependent exclusion limits are extracted using a Bayesian method implemented with the Bayesian analysis toolkit [75]. A binned likelihood function is constructed from the product of the Poisson probabilities of the observed and expected numbers of events in each \( m_{\ell'\ell'} \) mass bin as in Ref. [5]. A 95% credibility level (CL) Bayesian upper limit is placed on the signal cross section times branching ratio.

Expected exclusion limits are obtained by generating 1000 pseudo-experiments for each signal mass point. The median value of the pseudo-experiment distribution of the 95% C.L. Bayesian upper limit is taken as the expected limit. The one- and two-standard deviation intervals of the expected limit are obtained by finding the 68% and 95% intervals of the pseudo-experiment upper limit distribution.

The invariant mass spectrum for each final state is analyzed in 60 bins from 120 GeV to 10 TeV. The bin width is around 7% of the dilepton mass throughout the whole range. The predicted width of the \( Z' \) boson, 3% for \( m_{Z'} = 2 \) TeV, is smaller than the detector resolutions for the \( e\mu \) and the \( \mu\tau \) channels, which are approximately 8% and 12%, respectively, at the same \( Z' \) boson mass. For the \( e\tau \) final state the detector resolution is 4% at \( m_{Z'} = 2 \) TeV, comparable to the \( Z' \) boson width. The width of the \( \tilde{\nu}_e \) is below 1%, and hence the resolution of the detector is larger than the width for each of the final states investigated.

Figures 4–6 show the observed and expected 95% C.L. upper limits on the production cross section times branching ratio of the \( Z' \), RPV SUSY \( \tilde{\nu}_e \) and QBH models for each of the final states considered. The extracted limits are not as strong for signal masses above about 2.5 TeV due to a decrease in acceptance at very high \( p_T \) and, specifically to the LFV \( Z' \) model, low-mass signal production due to PDF suppression. The results are summarized in Table VIII. The acceptance times efficiency of the ADD and RS QBH models agree within 1%, and the same prediction is used for the limit extraction.

Results expressed in terms of the coupling limits can be directly compared to those obtained from precision low-energy experiments [76–79]. For the \( Z' \) model the cross...
section times branching ratio is proportional to $Q_{\ell\ell'}^2$, and
the same quark couplings as the SM Z boson are used. The
limits on $Q_{\ell\ell'}$ are shown in Fig. 7 as a function of $m_\chi$ for
the three channels. The most stringent coupling limits from
low-energy experiments are from $\mu$-to-$e$ conversion and
$\mu \to eee$ for the $e\mu$ channel, from $\tau \to eee$ and $\tau \to e\mu\mu$ for
the $e\tau$ channel, and from $\tau \to \mu\mu\mu$ and $\tau \to e\mu\mu$ for the $\mu\tau$
channel. The current experimental limits on these processes
are converted to coupling limits using the formulae of
Ref. [3] and are shown in Fig. 8. For the $e\tau$ and $\mu\tau$
channels, the observed limit is restricted up to the $Z'$ mass
point of 4 TeV. This because for the higher mass points, the
limit on $Q_{\ell\ell'}$ becomes sufficiently large that the total width
of the $Z'$ would be significantly larger than the experi-
mental resolution and violate our assumptions on that. For
the $e\mu$ channel, the coupling limits in this paper do not
compete with those from low-energy experiments, but for
the $e\tau$ and $\mu\tau$ channels, the coupling limits in this paper are
more stringent, though they require additional assumptions
on the quark couplings.

For the $\tilde{\nu}_e$ model, the dependence on the couplings is
more complicated because both the production and the
decay violate lepton-flavor conservation. Assuming only
the $d\bar{d}$ and $\ell\ell'$ couplings, the cross section times branching
ratios are proportional to the Yukawa couplings $|\lambda_{31i}\tilde{\lambda}_{3ij}|^2/(3|\lambda_{31i}|^2 + 2|\lambda_{3ij}|^2)$, where $ij = 12, 13,$ and
23 for the $e\mu$ channel, and $\mu\tau$ channels, respectively. The factor
3 in the denominator accounts for color and the factor 2 is
because both final-state charge combinations are allowed
($\ell^-\bar{\ell}'^+$). The limits on $|\lambda_{3ij}|$ versus $|\lambda_{31i}|$ are shown in
Fig. 8 for $\tilde{\nu}_e$, masses of 1 TeV, 2 TeV, and 3 TeV. The most
stringent coupling limits set by low-energy experiments
derive from $\mu$-to-$e$ conversion for the $e\mu$ channel, from $\tau \to e\eta$
for the $e\tau$ channel, and from $\tau \to \mu\eta$ for the $\mu\tau$ channel.
The coupling limits in Ref. [3] are scaled to current
experimental limits on these processes [76] and are shown
in Fig. 8. For the $e\mu$ channel, the coupling limits in this
paper do not compete with those from low-energy experi-
ments, but for $e\tau$ and $\mu\tau$ channels, the coupling limits in this
paper are more stringent.

**VIII. CONCLUSIONS**

A search for a heavy particle decaying into an $e\mu$, $e\tau$, or
$\mu\tau$ final state is conducted using 36.1 fb$^{-1}$ of proton-proton
collision data at $\sqrt{s} = 13$ TeV recorded by the ATLAS
detector at the Large Hadron Collider. The Standard Model
predictions are consistent with the data. From the $e\mu$, $e\tau$,
and $\mu\tau$ final states, Bayesian lower limits at 95% credibility
level are set on the mass of a $Z'$ vector boson with lepton-
flavor-violating couplings at 4.5, 3.7, and 3.5 TeV, respec-
tively; on the mass of a supersymmetric $\tau$-sneutrino with
R-parity-violating couplings at 3.4, 2.9, and 2.6 TeV; and
on the threshold mass for quantum black-hole production in
the context of the Arkani-Hamed–Dimopoulos–Dvali
(Randall-Sundrum) model at 5.6 (3.4), 4.9 (2.9), and 4.5
(2.6) TeV. The quantum black hole limits extracted are
below those extracted in dijet searches, since the branching
ratio to dijet is expected to be much larger than to dilepton.
Coupling limits for the lepton-flavor-violating $Z'$ boson and
$\tilde{\nu}_e$ models are more stringent than those from low-energy
experiments for the $e\tau$ and $\mu\tau$ modes.

**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; STScI, Belarus; CNPq and
FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN;
CONICYT, Chile; CAS, MOST and NSFC, China;
COLCIENCIAS, Colombia; MSMT CR, MPO CR and
VSC CR, Czech Republic; DNRF and DNSRC, Denmark;
IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG,
Georgia; BMBF, HGF, and MPG, Germany; GSRT,
Greece; RGC, Hong Kong SAR, China; ISF and
Benzoziyo Center, Israel; INFN, Italy; MEXT and JSPS,
Japan; CNRST, Morocco; NWO, Netherlands; RCN,
Norway; MNiSW and NCN, Poland; FCT, Portugal;
MNE/IFA, Romania; MES of Russia and NRC KI,
Russian Federation; JINR; MESTD, Serbia; MSSR,
Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South
Africa; MINECO, Spain; SRC and Wallenberg Foundation,
Sweden; SERI, SNSF and Cantons of Bern and Geneva,
Switzerland; MOST, Taiwan; TAEK, Turkey; STFC,
United Kingdom; DOE and NSF, United States of
America. In addition, individual groups and members have
received support from BCKDF, the Canada Council,
CANARIE, CRC, Compute Canada, FQRNT, and the
Ontario Innovation Trust, Canada; EPLANET, ERC,
ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie
Actions, European Union; Investissements d’Avenir Labex
and Idex, ANR, Région Auvergne and Fondation Partager
le Savoir, France; DFG and AvH Foundation, Germany;
Herakleitos, Thales and Aristeia programmes co-financed
by EU-ESF and the Greek NSRF; BSF, GIF and Minerva,
Israel; BRF, Norway; CERCA Programme Generalitat de
Catalunya, Generalitat Valenciana, Spain; 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. [81].
[7] ATLAS Collaboration, Search for a Heavy Neutral Particle Decaying to $e\mu$, $e\tau$, or $\mu\tau$ in $pp$ Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector, Phys. Rev. Lett. 115, 031801 (2015).


[76] Belle Collaboration, Search for lepton flavor violating $\tau^-$ decays into $\ell^-'\eta$, $\ell^-'\eta'$, and $\ell^-'\pi^{\pm}$, Phys. Lett. B 648, 341 (2007).


SEARCH FOR LEPTON-FLAVOR VIOLATION IN …

PHYS. REV. D 98, 092008 (2018)
SEARCH FOR LEPTON-FLAVOR VIOLATION IN …

PHYS. REV. D 98, 092008 (2018)


(ATLAS Collaboration)

1Department of Physics, University of Adelaide, Adelaide, Australia
2Physics Department, SUNY Albany, Albany, New York, USA
3Department of Physics, University of Alberta, Edmonton, Alberta, Canada
4aDepartment of Physics, Ankara University, Ankara, Turkey
4bIstanbul Aydin University, Istanbul, Turkey
4cDivision of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5LAPP, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, Annecy, France
6High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois, USA
7Department of Physics, University of Arizona, Tucson, Arizona, USA
8Department of Physics, University of Texas at Arlington, Arlington, Texas, USA
9Physics Department, National and Kapodistrian University of Athens, Athens, Greece
10Department of Physics, National Technical University of Athens, Zografou, Greece
11Department of Physics, University of Texas at Austin, Austin, Texas, USA
12Bahcesehir University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey
12Bilkent University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey
13Department of Physics, Bogazici University, Istanbul, Turkey
14Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey
15Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
15bPhysics Department, Tsinghua University, Beijing, China
15cDepartment of Physics, Nanjing University, Nanjing, China
15dUniversity of Chinese Academy of Science (UCAS), Beijing, China
16Institute of Physics, University of Belgrade, Belgrade, Serbia
17Department for Physics and Technology, University of Bergen, Bergen, Norway
18Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, California, USA
19Institut für Physik, Humboldt Universität zu Berlin, Berlin, Germany
20Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
21School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
21aCentro de Investigaciones, Universidad Antonio Nariño, Bogota, Colombia
23aDepartment of Physics, Brandeis University, Waltham, Massachusetts, USA
27Department of Physics, Alexandru Ioan Cuca University of Iasi, Iasi, Romania
27aUniversity of Chinese Academy of Sciences, Beijing, China
27bInstitute of Physics, Chinese Academy of Sciences, Beijing, China
27cInstitute of Physics, University of Belgrade, Belgrade, Serbia
SEARCH FOR LEPTON-FLAVOR VIOLATION IN …

PHYS. REV. D 98, 092008 (2018)

092008-31
Department of Physics, University of Massachusetts, Amherst, Massachusetts, USA
Department of Physics, McGill University, Montreal, Quebec, Canada
School of Physics, University of Melbourne, Victoria, Australia
Department of Physics, University of Michigan, Ann Arbor, Michigan, USA
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan, USA
B. I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Belarus
Research Institute for Nuclear Problems of Byelorussian State University, Minsk, Belarus
Group of Particle Physics, University of Montreal, Montreal, Quebec, Canada
P. N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia
Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia
National Research Nuclear University MEPhI, Moscow, Russia
D. V. Skobeltsyn Institute of Nuclear Physics, M. V. Lomonosov Moscow State University, Moscow, Russia
Fakultät für Physik, Ludwig-Maximilians-Universität München, München, Germany
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany
Nagasaki Institute of Applied Science, Nagasaki, Japan
Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico, USA
Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, Netherlands
Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, Netherlands
Department of Physics, Northern Illinois University, DeKalb, Illinois, USA
Budker Institute of Nuclear Physics and NSU, SB RAS, Novosibirsk, Russia
Novosibirsk State University Novosibirsk, Russia
Institute for High Energy Physics of the National Research Centre Kurchatov Institute, Protvino, Russia
Department of Physics, New York University, New York, New York, USA
The Ohio State University, Columbus, Ohio, USA
Faculty of Science, Okayama University, Okayama, Japan
Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, Oklahoma, USA
Department of Physics, Oklahoma State University, Stillwater, Oklahoma, USA
Palacký University, RCPTM, Joint Laboratory of Optics, Olomouc, Czech Republic
Center for High Energy Physics, University of Oregon, Eugene, Oregon, USA
LAL, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
Graduate School of Science, Osaka University, Osaka, Japan
Department of Physics, University of Oslo, Oslo, Norway
Department of Physics, Oxford University, Oxford, United Kingdom
LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris, France
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
Konstantinov Nuclear Physics Institute of National Research Centre “Kurchatov Institute”, PNPI, St. Petersburg, Russia
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
Laboratório de Instrumentação e Física Experimental de Partículas—LIP, Portugal
Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
Departamento de Física, Universidade de Coimbra, Coimbra, Portugal
Centro de Física Nuclear da Universidade de Lisboa, Lisboa, Portugal
Departamento de Física, Universidade do Minho, Braga, Portugal
Departamento de Física Teórica del Cosmos, Universidad de Granada, Granada (Spain), Spain
Dep Física and CEFITEC of Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
Czech Technical University in Prague, Prague, Czech Republic
Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic
Particle Physics Department, Rutherford Appleton Laboratory, Didcot, United Kingdom
IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz, California, USA
Departamento de Física, Pontificia Universidad Católica de Chile, Santiago, Chile
Departamento de Física, Universidad Técnica Federico Santa María, Valparaíso, Chile
Department of Physics, University of Washington, Seattle, Washington, USA