Search for lepton-flavour-violating decays of the Higgs and $Z$ bosons with the ATLAS detector

ATLAS Collaboration*

CERN, 1211 Geneva 23, Switzerland

Received: 27 April 2016 / Accepted: 15 January 2017

© CERN for the benefit of the ATLAS collaboration 2017. This article is published with open access at Springerlink.com

1 Introduction

One of the main goals of the Large Hadron Collider (LHC) physics programme at CERN is to discover physics beyond the Standard Model (SM). A possible sign would be the observation of lepton flavour violation (LFV) that could be realised in decays of the Higgs boson or of the $Z$ boson to pairs of leptons with different flavours.

Lepton-flavour-violating decays of the Higgs boson can occur naturally in models with more than one Higgs doublet [1–4], composite Higgs models [5,6], models with flavour symmetries [7], Randall–Sundrum models [8] and many others [9–16]. LFV $Z$ boson decays are predicted in models with heavy neutrinos [17], extended gauge models [18] and supersymmetry [19].

The most stringent bounds on the LFV decays of the Higgs and $Z$ bosons other than $H \rightarrow \mu \nu$ are derived from direct searches [20]. The CMS Collaboration has performed the first direct search for LFV $H \rightarrow \mu \tau$ decays [21] and reported a small excess (2.4 standard deviations) of data over the predicted background. Their results give a 1.51% upper limit on $\text{Br}(H \rightarrow \mu \tau)$ at the 95% confidence level (CL). The ATLAS Collaboration has also performed a search [22] for the LFV $H \rightarrow \mu \tau$ decays in the final state with one muon and one hadronically decaying $\tau$-lepton, $\tau_{\text{had}}$, and reported a 1.85% upper limit on $\text{Br}(H \rightarrow \mu \tau)$ at the 95% CL. The most stringent indirect constraint on $H \rightarrow e\mu$ decays is derived from the results of searches for $\mu \rightarrow e\gamma$ decays [23], and a bound of $\text{Br}(H \rightarrow e\mu) < 10^{-7}$ is obtained [24,25]. The bound on $\mu \rightarrow e\gamma$ decays suggests that the presence of a $H \rightarrow \mu\tau$ signal would exclude the presence of a $H \rightarrow e\tau$ signal, and vice versa, at an experimentally observable level at the LHC [25]. It is also important to note that a relatively large $\text{Br}(H \rightarrow \mu\tau)$ can be achieved without any particular tuning of the effective couplings, while a large $\text{Br}(H \rightarrow e\tau)$ is possible only at the cost of some fine-tuning of the corresponding couplings [25].

Upper bounds on the LFV $Z \rightarrow e\mu$, $Z \rightarrow \mu\tau$ and $Z \rightarrow e\tau$ decays were set by the LEP experiments [26,27]: $\text{Br}(Z \rightarrow e\mu) < 1.7 \times 10^{-6}$, $\text{Br}(Z \rightarrow e\tau) < 9.8 \times 10^{-6}$, and $\text{Br}(Z \rightarrow \mu\tau) < 1.2 \times 10^{-5}$ at the 95% CL. The ATLAS experiment set the most stringent upper bound on the LFV $Z \rightarrow e\mu$ decays [28]: $\text{Br}(Z \rightarrow e\mu) < 7.5 \times 10^{-7}$ at 95% CL.

This paper describes three new searches for LFV decays of the Higgs and $Z$ bosons. The first study is a search for $H \rightarrow e\tau$ decays in the final state with one electron and one hadronically decaying $\tau$-lepton, $\tau_{\text{had}}$. The second analysis is a simultaneous search for the LFV $H \rightarrow e\tau$ and $H \rightarrow \mu\tau$ decays in the final state with a leptonically decaying $\tau$-lepton, $\tau_{\text{lep}}$. A combination of results of the earlier ATLAS search for the LFV $H \rightarrow \mu\tau_{\text{had}}$ decays [22] and the two searches described in this paper is also presented. The third study constitutes the first ATLAS search for LFV decays of the $Z$ boson with hadronic $\tau$-lepton decays in the channel $Z \rightarrow \mu\tau_{\text{had}}$. The search for LFV decays in the $\tau_{\text{lep}}$ analysis is based on the novel method introduced in Ref. [29]; the searches in the $\tau_{\text{had}}$ analyses are based on the techniques developed for the SM $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ search. All three searches are based on the data sample of $pp$ collisions collected at a centre-of-mass energy of $\sqrt{s} = 8$ TeV and corresponding to an integrated luminosity of 20.3 fb$^{-1}$. Given the overlap between the analysis techniques used in the $H \rightarrow e\tau_{\text{had}}$ search and in the $Z \rightarrow \mu\tau_{\text{had}}$ search, from here on they are referred to as the

* e-mail: atlas.publications@cern.ch
\( \tau_{\text{had}} \) channels; the \( H \rightarrow \ell \tau_{\text{lep}} \) search is referred to as the \( \tau_{\text{lep}} \) channel, where \( \ell = e, \mu \).

### 2 The ATLAS detector and object reconstruction

The ATLAS detector\(^1\) is described in detail in Ref. [30]. ATLAS consists of an inner tracking detector (ID) covering the range \(|\eta| < 2.5\), surrounded by a superconducting solenoid providing a 2 T axial magnetic field, a high-granularity electromagnetic (\(|\eta| < 3.2\)) calorimeter, a hadronic calorimeter (\(|\eta| < 4.9\)), and a muon spectrometer (MS) (\(|\eta| < 2.7\)) with a toroidal magnetic field.

The signatures of LFV searches reported here are characterised by the presence of an energetic lepton originating directly from the boson decay and carrying roughly half of its energy, and the hadronic or leptonic decay products of a \( \tau \)-lepton. The data in the \( \tau_{\text{had}} \) channels were collected with single-lepton triggers: a single-muon trigger with the threshold of \( p_T = 24 \text{ GeV} \) and a single-electron trigger with the threshold \( E_T = 24 \text{ GeV} \). The data in the \( \tau_{\text{lep}} \) channel were collected using asymmetric electron-muon triggers with \((p_T^e, E_T^\tau) > (18, 8) \text{ GeV} \) and \((E_T^\mu, p_T^\mu) > (14, 8) \text{ GeV} \) thresholds. The \( p_T \) and \( E_T \) requirements on the objects in the presented analyses are at least 2 GeV higher than the trigger requirements.

A brief description of the object definitions is provided below. The primary vertex is chosen as the proton–proton collision vertex candidate with the highest sum of the squared transverse momenta of all associated tracks [31].

Muons are reconstructed using an algorithm that combines information from the ID and the MS [32]. Muon quality criteria such as inner-detector hit requirements are applied to achieve a precise measurement of the muon momentum and to reduce the misidentification rate. Muons are required to have \( p_T > 10 \text{ GeV} \) and to be within \(|\eta| < 2.5\). The distance between the \( z \)-position of the point of closest approach of the muon inner-detector track to the beamline and the \( z \)-coordinate of the primary vertex is required to be less than 1 cm. In the \( \tau_{\text{lep}} \) channel, there is an additional cut on the transverse impact parameter significance, defined as the transverse impact parameter divided by its uncertainty: \(|d_0|/\sigma_{d_0} < 3\). These requirements reduce the contamination due to cosmic-ray muons and beam-induced backgrounds. Typical reconstruction and identification efficiencies for muons meeting these selection criteria are above 95% [32].

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeters matched to tracks in the ID. They are required to have transverse energy \( E_T > 15(12) \text{ GeV} \) in the \( \tau_{\text{had}} \) (\( \tau_{\text{lep}} \)) channel, to be within the pseudorapidity range \(|\eta| < 2.47\), and to satisfy the medium shower shape and track selection criteria defined in Ref. [33]. Candidates found in the transition region between the barrel and end-cap calorimeters \([1.37 < |\eta| < 1.52]\) are not considered in the \( \tau_{\text{had}} \) channel. Typical reconstruction and identification efficiencies for electrons satisfying these selection criteria range between 80 and 90%, depending on \( E_T \) and \( \eta \).

Exactly one lepton (electron or muon) satisfying the above identification requirements is allowed in the \( \tau_{\text{had}} \) channels. In the \( \tau_{\text{lep}} \) channel, only events with exactly one identified muon and one identified electron are retained. All lepton (electron or muon) candidates must be matched to the corresponding trigger objects and satisfy additional isolation criteria, based on tracking and calorimeter information, in order to suppress the background from misidentified jets or from semileptonic decays of charm and bottom hadrons. The calorimeter isolation variable \( I(E_T, \Delta R) \) is defined as the sum of the total transverse energy in the calorimeter in a cone of size \( \Delta R \) around the electron cluster or the muon track, divided by the \( E_T \) of the electron cluster or the \( p_T \) of the muon, respectively. The track-based isolation \( I(p_T, \Delta R) \) is defined as the scalar sum of the transverse momenta of tracks within a cone of size \( \Delta R \) around the electron or muon track, divided by the \( E_T \) of the electron cluster or the muon \( p_T \), respectively. The contribution due to the lepton itself is not included in either sum. The isolation requirements used in the \( \tau_{\text{had}} \) and \( \tau_{\text{lep}} \) channels, optimised to reduce the contamination from non-prompt leptons, are listed in Table 1.

Hadronically decaying \( \tau \)-leptons are identified by means of a multivariate analysis technique [34] based on boosted decision trees, which exploits information about ID tracks and clusters in the electromagnetic and hadronic calorimeters. The \( \tau_{\text{had}} \) candidates are required to have \(+1\) or \(−1\) net charge in units of electron charge, and must be 1- or 3-track (1- or 3-prong) candidates. Events with exactly one \( \tau_{\text{had}} \) can-

---

\(^{1}\) 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 upward. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \( \phi \) being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle \( \theta \) as \(|\eta| = −\ln \tan(\theta/2)\). The transverse momentum and the transverse energy are defined as \( p_T = p \times \sin(\theta) \) and \( E_T = E \times \sin(\theta) \), respectively. The distance \( \Delta R \) in \( \eta-\phi \) space is defined as \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \).

| Table 1 Summary of isolation requirements applied for the selection of isolated electrons and muons. The isolation variables are defined in the text |
|---|---|
| \( \tau_{\text{lep}} \) channels | \( \tau_{\text{had}} \) channels |
| Electrons | \( I(E_T, 0.3) < 0.13 \) | \( I(E_T, 0.2) < 0.06 \) |
| \( I(p_T, 0.3) < 0.07 \) | \( I(p_T, 0.4) < 0.06 \) |
| Muons | \( I(E_T, 0.3) < 0.14 \) | \( I(E_T, 0.2) < 0.06 \) |
| \( I(p_T, 0.3) < 0.06 \) | \( I(p_T, 0.4) < 0.06 \) |
didate satisfying the medium identification criteria [34] with \( p_T > 20 \text{ GeV} \) and \( |\eta| < 2.47 \) are considered in the \( \tau_{\text{had}} \) channels. In the \( \tau_{\text{lep}} \) channel, events with identified \( \tau_{\text{had}} \) candidates are rejected to avoid overlap between \( H \to \ell \tau_{\text{had}} \) and \( H \to \ell \tau_{\text{lep}} \). The identification efficiency for \( \tau_{\text{had}} \) candidates satisfying these requirements is (55–60)%. Dedicated criteria [34] to separate \( \tau_{\text{had}} \) candidates from misidentified electrons are also applied, with a selection efficiency for true \( \tau_{\text{had}} \) decays (that pass the \( \tau_{\text{had}} \) identification requirements described above) of 95%. To reduce the contamination due to backgrounds where a muon mimics a \( \tau_{\text{had}} \) signature, events in which an identified muon with \( p_T(\mu) > 4 \text{ GeV} \) overlaps with an identified \( \tau_{\text{had}} \) are rejected [35]. The probability to misidentify a jet with \( p_T > 20 \text{ GeV} \) as a \( \tau_{\text{had}} \) candidate is typically (1–2)% [34].

Jets are reconstructed using the anti-\( k_T \) jet clustering algorithm [36] with a radius parameter \( R = 0.4 \), taking the deposited energy in clusters of calorimeter cells as inputs. Fully calibrated jets [37] are required to be reconstructed in the range \(|\eta| < 4.5\) and to have \( p_T > 30 \text{ GeV} \). To suppress jets from multiple proton–proton collisions in the same or nearby beam bunch crossings, tracking information is used for central jets with \(|\eta| < 2.4\) and \( p_T < 50 \text{ GeV} \). In the \( \tau_{\text{lep}} \) channel, these central jets are required to have at least one track originating from the primary vertex. In the \( \tau_{\text{had}} \) channel, tracks originating from the primary vertex must contribute more than half of the jet \( p_T \) when summing the scalar \( p_T \) of all tracks in the jet; jets with no associated tracks are retained.

In the pseudorapidity range \(|\eta| < 2.5\), jets containing \( b \)-hadrons (\( b \)-jets) are selected using a tagging algorithm [38]. These jets are required to have \( p_T > 30 \text{ GeV} \) in the \( \tau_{\text{had}} \) channel, and \( p_T > 20 \text{ GeV} \) in the \( \tau_{\text{lep}} \) channel. Two different working points with \( \sim 70 \) and \( \sim 80\% \) \( b \)-tagging efficiencies for \( b \)-jets in simulated \( t\bar{t} \) events are used in the \( \tau_{\text{had}} \) and \( \tau_{\text{lep}} \) channels, respectively. The corresponding light-flavour jet misidentification probability is (0.1–1)%, depending on the \( p_T \) and \( \eta \) of the jet. Only a very small fraction of signal events have \( b \)-jets, therefore events with identified \( b \)-jets are vetoed in the selection of signal events.

Some objects might be reconstructed as more than one candidate. Overlapping candidates, separated by \( \Delta R < 0.2 \), are resolved by discarding one object and selecting the other one in the following order of priority (from highest to lowest): muons, electrons, \( \tau_{\text{had}} \), and jet candidates [35].

The missing transverse momentum (with magnitude \( E_T^{\text{miss}} \)) is reconstructed using the energy deposits in calorimeter cells calibrated according to the reconstructed physics objects (\( e, \gamma, \tau_{\text{had}}, \text{jets} \), and \( \mu \)) with which they are associated [39]. In the \( \tau_{\text{had}} \) channels, the energy from calorimeter cells not associated with any physics object is included in the \( E_T^{\text{miss}} \) calculation. It is scaled by the scalar sum of \( p_T \) of tracks which originate from the primary vertex but are not associated with any objects divided by the scalar sum of \( p_T \) of all tracks in the event which are not associated with objects. The scaling procedure achieves a more accurate reconstruction of \( E_T^{\text{miss}} \) under high pile-up conditions.

### 3 Signal and background samples

The LFV signal is estimated from simulation. The major Higgs boson production processes (gluon fusion \( ggH \), vector-boson fusion VBF, and associated production \( WH/ZH \)) are considered in the reported searches for LFV \( H \to \ell \tau \) and \( H \to \mu \tau \) decays. In the \( \tau_{\text{lep}} \) channel, all backgrounds are estimated from data. In the \( \tau_{\text{had}} \) channels, the \( Z/\gamma^* \to \tau \tau \) and multi-jet backgrounds are estimated from data, while the other remaining backgrounds are estimated from simulation, as described below.

The largely irreducible \( Z/\gamma^* \to \tau \tau \) background is modelled by \( Z/\gamma^* \to \mu \mu \) data events, where the muon tracks and associated energy deposits in the calorimeters are replaced by the corresponding simulated signatures of the final-state particles of the \( \tau \)-lepton decay. In this approach, essential features such as the modelling of the kinematics of the produced boson, the modelling of the hadronic activity of the event (jets and underlying event) as well as contributions from pile-up are taken from data. Therefore, the dependence on the simulation is minimised and only the \( \tau \)-lepton decays and the detector response to the \( \tau \)-lepton decay products are based on simulation. This hybrid sample is referred to as embedded data in the following. A detailed description of the embedding procedure can be found in Ref. [40].

The \( W+jets \), \( Z/\gamma^* \to \mu \mu \) and \( Z/\gamma^* \to ee \) backgrounds are modelled by the ALPGEN [41] event generator interfaced with PYTHIA8 [42] to provide the parton showering, hadronisation and the modelling of the underlying event. The backgrounds with top quarks are modelled by the POWHEG [43–45] (for \( t\bar{t} \), \( Wt \) and \( s \)-channel single-top production) and AcerMC [46] (\( t \)-channel single-top production) event generators interfaced with PYTHIA8. The ALPGEN event generator interfaced with HERWIG [47] is used to model the \( WW \) process, and HERWIG is used for the \( ZZ \) and \( WZ \) processes.

The events with Higgs bosons produced via \( ggH \) or VBF processes are generated at next-to-leading-order (NLO) accuracy in QCD with the POWHEG [48] event generator interfaced with PYTHIA8 to provide the parton showering, hadronisation and the modelling of the underlying event. The associated production (\( ZH \) and \( WH \)) samples are simulated using PYTHIA8. All events with Higgs bosons are produced with a mass of \( m_H = 125 \text{ GeV} \) assuming the narrow width approximation and normalised to cross sections calculated at next-to-next-to-leading order (NNLO) in QCD [49–51]. The SM \( H \to \tau \tau \) decays are simulated by PYTHIA8; the other SM decays of the Higgs boson are negligible. The LFV Higgs
boson decays are modelled by the EvtGen [52] event generator according to the phase-space model. In the $H \rightarrow \mu\tau$ and $H \rightarrow e\tau$ decays, the $\tau$-lepton decays are treated as unpolarised because the left- and right-handed $\tau$-lepton polarisation states are produced at equal rates. Finally, the LFV $Z$ boson decays are simulated with PYTHIA8 assuming an isotropic decay. The width of the $Z$ boson is set to its measured value [20].

For all simulated samples, the decays of $\tau$-leptons are modelled with TAUOLA [53] and the propagation of particles through the ATLAS detector is simulated with GEANT4 [54,55]. The effect of multiple proton–proton collisions in the same or nearby beam bunch crossings is accounted for by overlaying additional minimum-bias events. Simulated events are weighted so that the distribution of the average number of interactions per bunch crossing matches that observed in data.

Background contributions due to non-prompt leptons in the $\tau_{\text{lep}}$ channel and multi-jet events in the $\tau_{\text{had}}$ channel are estimated using data-driven techniques described in Sects. 4.2 and 5.2.

4 Search for $H \rightarrow e\tau$ decays in the $\tau_{\text{had}}$ channel

The search for the LFV $H \rightarrow e\tau$ decays in the $\tau_{\text{had}}$ channel follows exactly the same analysis strategy and utilises the same background estimation techniques as those used in the ATLAS search for the LFV $H \rightarrow \mu\tau$ decays in the $\tau_{\text{had}}$ channel [22]. The only major difference is that a high-$E_T$ electron is required in the final state instead of a muon. A detailed description of the $H \rightarrow e\tau_{\text{had}}$ analysis is provided in the following sections.

4.1 Event selection and categorisation

Signal $H \rightarrow e\tau$ events in the $e\tau_{\text{had}}$ final state are characterised by the presence of exactly one energetic electron and one $\tau_{\text{had}}$ of opposite-sign (OS) charge as well as moderate $E_T^{\text{miss}}$, which tends to be aligned with the $\tau_{\text{had}}$ direction. Same-sign (SS) charge events are used to control the rates of background contributions. Events with identified muons are rejected. Backgrounds for this signature can be broadly classified into two major categories:

- Events with a misidentified $\tau_{\text{had}}$ signature. These are dominated by $W+\text{jets}$ events with some contribution from multi-jet (many of which have genuine electrons from semileptonic decays of heavy-flavour hadrons), diboson ($VV$), $t\bar{t}$ and single-top events with $N_{\text{OS}}>N_{\text{SS}}$. Additional contributions to this category arise from $Z(\rightarrow ee)+\text{jets}$ events, where a $\tau_{\text{had}}$ signature can be mimicked by either a jet (no charge correlation) or an electron (strong charge anti-correlation).

Events with a misidentified $\tau_{\text{had}}$ tend to have a much softer $p_T(\tau_{\text{had}})$ spectrum and a larger angular separation between the $\tau_{\text{had}}$ and $E_T^{\text{miss}}$ directions. These properties are exploited to suppress backgrounds and define signal and control regions. Events with exactly one electron and exactly one $\tau_{\text{had}}$ with $E_T(e)>26$ GeV, $p_T(\tau_{\text{had}})>45$ GeV and $|\eta(e)-\eta(\tau_{\text{had}})|<2$ form a baseline sample as it represents a common selection for both the signal and control regions. The $|\eta(e)-\eta(\tau_{\text{had}})|$ cut has $\sim 99\%$ efficiency for signal and rejects a considerable fraction of multi-jet and $W+\text{jets}$ events. Similarly as done in Ref. [22], two signal regions are defined using the transverse mass $m_T$, of the $e-E_T^{\text{miss}}$ and $\tau_{\text{had}}-E_T^{\text{miss}}$ systems: OS events with $m_T^{e,E_T^{\text{miss}}}>40$ GeV and $m_T^{\tau_{\text{had}},E_T^{\text{miss}}}>30$ GeV form the signal region-1 (SR1), while OS events with $m_T^{e,E_T^{\text{miss}}}<40$ GeV and $m_T^{\tau_{\text{had}},E_T^{\text{miss}}}<60$ GeV form the signal region-2 (SR2). Both regions have similar sensitivity to the signal (see Sect. 4.4). The dominant background in SR1 is $W+\text{jets}$, while the $Z/\gamma^*$ leads to $\tau\tau$ and $Z\rightarrow ee+\text{jets}$ backgrounds dominate in SR2. The modelling of the $W+\text{jets}$ background is checked in a dedicated control region (WCR) formed by events with $m_T^{e,E_T^{\text{miss}}}>60$ GeV and $m_T^{\tau_{\text{had}},E_T^{\text{miss}}}>40$ GeV. As discussed in detail in Sect. 4.2, the modelling of the $Z/\gamma^*\rightarrow\tau\tau$ and $Z\rightarrow ee+\text{jets}$ backgrounds is checked in SR2. The choice of $m_T$ cuts to define SR1, SR2 and WCR is motivated by correlations between $m_T^{e,E_T^{\text{miss}}}$ and $m_T^{\tau_{\text{had}},E_T^{\text{miss}}}$ in $H \rightarrow e\tau$ signal and major background ($W+\text{jets}$ and $Z/\gamma^*\rightarrow\tau\tau$) events, as illustrated in Fig. 1. No events with identified $b$-jets are allowed in SR1, SR2 and WCR. The modelling of the $t\bar{t}$ and single-top backgrounds is checked in a dedicated control region (TCR), formed by events that satisfy the baseline selection and have at least two jets, with at least one being $b$-tagged. Table 2 provides a summary of the event selection criteria used to define the signal and control regions.

The LFV signal is searched for by performing a fit to the mass distribution in data, $n_{ee}^{\text{MMC}}$, reconstructed from

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

where $\ell = e, \tau_{\text{had}}$ and $\Delta\phi$ is the azimuthal separation between the directions of the lepton ($e$ or $\tau_{\text{had}}$) and $E_T^{\text{miss}}$ vectors.
The main improvement comes from requiring that the relative orientations of the neutrino and other lepton decay products are consistent with the mass and kinematics of a τ-lepton decay. This is achieved by maximising a probability defined in the kinematically allowed phase-space region. The MMC used in the $H \rightarrow \tau \tau$ analysis [35] is modified to take into account that there is only one neutrino from a hadronic τ-lepton decay in LFV $H \rightarrow \tau \tau$ events. For a Higgs boson with $m_H = 125$ GeV, the reconstructed $m_{\tau\tau}^{\text{MMC}}$ distribution has a roughly Gaussian shape with a full width at half maximum of $\sim 19$ GeV. The analysis is performed “blinded” in the 110 GeV $< m_{\tau\tau}^{\text{MMC}} <$ 150 GeV regions of SR1 and SR2, which contain 93.5 and 95% of the expected signal events in SR1 and SR2, respectively. The event selection and the analysis strategy are defined without looking at the data in these blinded regions.

4.2 Background estimation

The background estimation method takes into account the background properties and composition discussed in

---

**Fig. 1** Two-dimensional distributions of the transverse mass of the $e$-$E_T^{\text{miss}}$ system, $m_{\text{miss}}^{E_T}$, and that of the $\tau_{\text{had}}$-$E_T^{\text{miss}}$ system, $m_{\tau\text{had}}^{E_T}$, in simulated $Z/\gamma^{*} \rightarrow \tau \tau$ (top left plot), W+jets (top right plot), $H \rightarrow \tau\tau e$ signal (bottom left plot) and data (bottom right plot) events. Magenta, red and yellow boxes on the bottom right plot illustrate SR1, SR2, and WCR, respectively. All events used for these distributions are required to have a well-identified electron and $\tau_{\text{had}}$ (as described in text) of opposite charge with $p_T(\tau_{\text{had}}) > 20$ GeV and $E_T(e) > 26$ GeV.

**Table 2** Summary of the event selection criteria used to define the signal and control regions (see text)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>SR1</th>
<th>SR2</th>
<th>WCR</th>
<th>TCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_T(e)$</td>
<td>$&gt;26$ GeV</td>
<td>$&gt;26$ GeV</td>
<td>$&gt;26$ GeV</td>
<td>$&gt;26$ GeV</td>
</tr>
<tr>
<td>$p_T(\tau_{\text{had}})$</td>
<td>$&gt;45$ GeV</td>
<td>$&gt;45$ GeV</td>
<td>$&gt;45$ GeV</td>
<td>$&gt;45$ GeV</td>
</tr>
<tr>
<td>$</td>
<td>\eta(e) - \eta(\tau_{\text{had}})</td>
<td>$</td>
<td>$&lt;2$</td>
<td>$&lt;2$</td>
</tr>
<tr>
<td>$m_{\tau\tau}^{E_T}$</td>
<td>$&gt;40$ GeV</td>
<td>$&lt;40$ GeV</td>
<td>$&gt;60$ GeV</td>
<td>–</td>
</tr>
<tr>
<td>$m_{\tau_{\text{had}}}^{E_T}$</td>
<td>$&lt;30$ GeV</td>
<td>$&lt;60$ GeV</td>
<td>$&gt;40$ GeV</td>
<td>–</td>
</tr>
<tr>
<td>$N_{\text{Jet}}$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>$\geq2$</td>
</tr>
<tr>
<td>$N_{b_{\text{Jet}}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\geq1$</td>
</tr>
</tbody>
</table>
Sect. 4.1. It also relies on the observation that the shape of the $m_{\text{et}}^{\text{MMC}}$ distribution for the multi-jet background is the same for OS and SS events. This observation was made using a dedicated control region, MJCR, with an enhanced contribution from the multi-jet background. Events in this control region are required to meet all criteria for SR1 and SR2 with the exception of the requirement on $|\eta(e) - \eta(\tau_{\text{had}})|$, which is reversed: $|\eta(e) - \eta(\tau_{\text{had}})| > 2$. Therefore, the total number of OS background events, $N_{\text{bkg}}^{\text{OS}}$, in each bin of the $m_{\text{et}}^{\text{MMC}}$ (or any other) distribution in SR1 and SR2 can be obtained according to the following formula:

$$N_{\text{OS}}^{\text{bkg}} = r_{QCD} \cdot N_{\text{data}}^{\text{SS}} + \sum_{i} N_{\text{bkg-i}}^{\text{OS-SS}},$$

where the individual terms are described below. $N_{\text{data}}^{\text{SS}}$ is the number of SS data events, which contains significant contributions from $W$+jets events, multi-jet and other backgrounds.

The fractions of multi-jet background in SS data events inside the $110 \text{ GeV} < m_{\text{et}}^{\text{MMC}} < 150 \text{ GeV}$ mass window are $\sim 27$ and $\sim 64\%$ in SR1 and SR2, respectively. The contributions $N_{\text{bkg-i}}^{\text{OS-SS}} = N_{\text{bkg-i}}^{\text{SS}} - r_{QCD} \cdot N_{\text{bkg-i}}^{\text{OS}}$ are add-on terms for the different background components (where bgk-i indicates the $i^{th}$ background source: $Z \rightarrow \tau\tau$, $Z \rightarrow ee$, $W$+jets, $VV$, $H \rightarrow \tau\tau$ and events with $t$-quarks), which also account for components of these backgrounds already included in SS data events.\footnote{The $r_{QCD} \cdot N_{\text{bkg-i}}^{\text{SS}}$ correction in the add-on term is needed because same-sign data events include multi-jet as well as electroweak events ($Z \rightarrow \tau\tau$, $Z \rightarrow ee$, $W$+jets, $VV$, $H \rightarrow \tau\tau$ and events with $t$-quarks) and their contributions cannot be separated.}

The factor $r_{QCD} = N_{\text{multi-jet}}^{\text{SS}} / N_{\text{multi-jet}}^{\text{OS}}$ accounts for potential differences in flavour composition and, as a consequence, in jet $\rightarrow \tau_{\text{had}}$ misidentification rates of final-state jets introduced by the same-sign or opposite-sign charge requirements. The value of $r_{QCD} = 1.0 \pm 0.13$ is obtained from a multi-jet enriched control region in data using a method discussed in Ref. [58]. This sample is obtained by selecting events with $E_{\text{Tmiss}} < 15 \text{ GeV}$, $m_{\text{T}}^{\text{Emiss}} < 30 \text{ GeV}$, removing the isolation criteria of the electron candidate and using the loose identification criteria for the $\tau_{\text{had}}$ candidate [34]. The systematic uncertainty on $r_{QCD}$ is estimated by varying the selection cuts described above. The obtained value of $r_{QCD}$ is also verified in the MJCR region, which has a smaller number of events but where the electron and $\tau_{\text{had}}$ candidates pass the same identification requirements as events in SR1 and SR2.

The data and simulation samples used for the modelling of background processes are described in Sect. 3. A discussion of each background source is provided below.

The largely irreducible $Z/\gamma^* \rightarrow \tau\tau$ background is modelled by the embedded data sample described in Sect. 3. The $Z/\gamma^* \rightarrow \tau\tau$ normalisation is a free parameter in the final fit to data and it is mainly constrained by events with $60 \text{ GeV} < m_{\text{et}}^{\text{MMC}} < 90 \text{ GeV}$ in SR2.

Events due to the $W$+jets background are mostly selected when the $\tau_{\text{had}}$ signature is mimicked by jets. This background is estimated from simulation, and the WCR region is used to check the modelling of the $W$+jets kinematics and to obtain separate normalisations for OS and SS $W$+jets events. The difference in these two normalisations happens to be statistically significant. An additional overall normalisation factor for the $N_{\text{OS-SS}}^{W+\text{jets}}$ term in Eq. (1) is introduced as a free parameter in the final fit in SR1. By studying WCR events and SR1 events with $m_{\text{et}}^{\text{MMC}} > 150 \text{ GeV}$ (dominated by $W$+jets background), it is also found that an $m_{\text{et}}^{\text{MMC}}$ shape correction, which depends on the number of jets, $P_T(\tau_{\text{had}})$ and $|\eta(e) - \eta(\tau_{\text{had}})|$, needs to be applied in SR1. This correction is derived from SR1 events with $m_{\text{et}}^{\text{MMC}} > 150 \text{ GeV}$ and it is applied to events with any value of $m_{\text{et}}^{\text{MMC}}$. The corresponding modelling uncertainty is set to be $50\%$ of the difference of the $m_{\text{et}}^{\text{MMC}}$ shapes obtained after applying the SR1-based and WCR-based shape corrections. The size of this uncertainty depends on $m_{\text{et}}^{\text{MMC}}$ and it is as large as $\pm 10\%$ for $W$+jets events with $m_{\text{et}}^{\text{MMC}} < 150 \text{ GeV}$. In the case of SR2, good modelling of the $N_{\text{jet}}, P_T(\tau_{\text{had}})$ and $|\eta(e) - \eta(\tau_{\text{had}})|$ distributions suggests that such a correction is not needed. However, a modelling uncertainty in the $m_{\text{et}}^{\text{MMC}}$ shape of the $W$+jets background in SR2 is set to be $50\%$ of the difference between the $m_{\text{et}}^{\text{MMC}}$ shape obtained without any correction and the one obtained after applying the correction derived for SR1 events. The size of this uncertainty is below $10\%$ in the $110 \text{ GeV} < m_{\text{et}}^{\text{MMC}} < 150 \text{ GeV}$ region, which contains most of the signal events. It was also checked that applying the same correction in SR2 as in SR1 would affect the final result by less than $4\%$ (see Sect. 6). The modelling of jet fragmentation and the underlying event has a significant effect on the estimate of the jet $\rightarrow \tau_{\text{had}}$ misidentification rate in different regions of the phase space and has to be accounted for with a corresponding systematic uncertainty. To estimate this effect, the analysis was repeated using a sample of $W$+jets events modelled by ALPGEN interfaced with the HERWIG event generator. Differences in the $W$+jets predictions in SR1 and SR2 are found to be $\pm 12$ and $\pm 15\%$, respectively, and are taken as corresponding systematic uncertainties.

In the case of the $Z \rightarrow ee$ background, there are two components: events in which an electron mimics a $\tau_{\text{had}}$ ($e \rightarrow \tau_{\text{had}}$) and events in which a jet mimics a $\tau_{\text{had}}$ ($\text{jet} \rightarrow \tau_{\text{had}}$). In the first case, the shape of the $Z \rightarrow ee$ background is obtained from simulation. Corrections from data, derived from dedicated tag-and-probe studies [59], are also applied to account for the variation in the $e \rightarrow \tau_{\text{had}}$ misidentification rate as a function of $\eta$. The normalisation of this background component is a free parameter in the final fit to data and it is mainly constrained by events with
Fig. 2 Distributions of the mass reconstructed by the Missing Mass Calculator, \( m_{\text{MMC}}^{e\tau} \), in SR1 (left) and SR2 (right). The background distributions are determined in a global fit (described in Sect. 4.4). The signal distribution corresponds to \( \text{Br}(H \rightarrow e\tau) = 25\% \). The bottom panel of each sub-figure shows the ratio of the observed data to the estimated background. Very small backgrounds due to single top, \( t\bar{t} \), \( \text{VV} \), \( Z \rightarrow ee \) (jet \( \rightarrow \tau \) misid had) and \( H \rightarrow \tau\tau \) events are combined in a single background component labelled as “Other Backgrounds”. The grey band for the ratio illustrates post-fit systematic uncertainties in the background prediction. The statistical uncertainties in the background predictions and data are added in quadrature for the ratios. The last bin in each distribution contains events with \( m_{\text{MMC}}^{e\tau} > 250 \text{ GeV} \).

90 GeV < \( m_{e\tau}^{\text{MMC}} \) < 110 GeV in SR2. For the \( Z \rightarrow ee \) background where a jet is misidentified as a \( \tau \) had candidate and one of the electrons does not pass the electron identification criteria described in Sect. 2, the normalisation factor and shape corrections, which depend on the number of jets, \( p_T(\tau \text{had}) \) and |\( \eta(e) - \eta(\tau \text{had}) \)|, are derived using events with two identified OS electrons with an invariant mass, \( m_{ee} \), in the range of 80–100 GeV. Since this background does not have an OS–SS charge asymmetry, a single correction factor is derived for OS and SS events. Half the difference between the \( m_{e\tau}^{\text{MMC}} \) shape with and without this correction is taken as the corresponding systematic uncertainty.

The TCR is used to check the modelling and to obtain normalisations for OS and SS events with top quarks. The normalisation factors obtained in the TCR are extrapolated into SR1 and SR2, where \( t\bar{t} \) and single-top events may have different properties. To estimate the uncertainty associated with such an extrapolation, the analysis is repeated using the MC@NLO [60] event generator instead of POWHEG for \( t\bar{t} \) production. This uncertainty is found to be ±8% (±14%) for backgrounds with top quarks in SR1 (SR2).

The background due to diboson (\( WW \), \( ZZ \) and \( WZ \)) production is estimated from simulation, normalised to the cross sections calculated at NLO in QCD [49–51]. All other SM Higgs boson decays constitute negligible backgrounds for the LFV signature.

Figure 2 shows the \( m_{e\tau}^{\text{MMC}} \) distributions for data and the predicted backgrounds in each of the signal regions. The backgrounds are estimated using the method described above and their normalisations are obtained in a global fit described in Sect. 4.4. The signal acceptance times efficiencies for passing the SR1 or SR2 selection requirements are 1.8 and 1.4%, respectively, and the combined efficiency is 3.2%. The numbers of observed events in the data as well as the signal and background predictions in the mass region 110 GeV < \( m_{e\tau}^{\text{MMC}} \) < 150 GeV can be found in Table 3.

4.3 Systematic uncertainties

The numbers of signal and background events and the shapes of corresponding \( m_{e\tau}^{\text{MMC}} \) distributions are affected by systematic uncertainties. They are discussed below and changes in event yields are provided for major sources of uncertainties. For all uncertainties, the effects on both the total signal and background predictions and on the shape of the \( m_{e\tau}^{\text{MMC}} \) distribution are evaluated. Unless otherwise mentioned, all sources of experimental uncertainties are treated as fully correlated across signal and control regions in the final fit which is discussed in Sect. 4.4.

The largest systematic uncertainties arise from the normalisation (±12% uncertainty) and modelling of the \( W + \text{jets} \) background. The uncertainties on the \( W + \text{jets} \) normalisa-
tion and $m_{\tau\tau}^{\text{MMC}}$ shape corrections are treated as uncorrelated between SR1 and SR2. The uncertainties in $q_{\text{QCD}}$ ($\pm13\%$) and in the normalisation ($\pm13\%$) and modelling of $Z \rightarrow \tau \tau$ also play an important role. The normalisation uncertainty ($\pm7\%$) for the $Z \rightarrow ee$ (with $e \rightarrow \tau_{\text{had}}$) background has a limited impact on the sensitivity because of a good separation of the signal and $Z \rightarrow ee$ peaks in the $m_{\tau\tau}^{\text{MMC}}$ distribution. The other major sources of experimental uncertainty, affecting both the shape and normalisation of signal and backgrounds, are the uncertainty in the $\tau_{\text{had}}$ energy scale [34], which is measured with $\pm(2\text{-}4\%)$ precision (depending on $\tau_{T}$ and decay mode of the $\tau_{\text{had}}$ candidate), and uncertainties in the embedding method used to model the $Z \rightarrow \tau \tau$ background [35]. Less significant sources of experimental uncertainty, affecting the shape and normalisation of signal and backgrounds, are the uncertainty in the jet energy scale [37,62] and resolution [63]. The uncertainties in the $\tau_{\text{had}}$ energy resolution, the energy scale in calorimeter cells not associated with physics objects are taken into account; however, they are found to be only $\pm(1\text{-}2\%)$. The following experimental uncertainties primarily affect the normalisation of signal and backgrounds: the $\pm2.8\%$ uncertainty in the integrated luminosity [64], the uncertainty in the $\tau_{\text{had}}$ identification efficiency [34], which is measured to be $\pm(2\text{-}3\%)$ for 1-prong and $\pm(3\text{-}5\%)$ for 3-prong decays(where the range reflects the dependence on $p_{T}$ of the $\tau_{\text{had}}$ candidate), the $\pm2.1\%$ uncertainty for triggering, reconstructing and identifying electrons [33], and the $\pm2\%$ uncertainty in the $b$-jet tagging efficiency [38].

Theoretical uncertainties are estimated for the Higgs boson production and for the $VV$ background, which are modelled with the simulation and are not normalised to data in dedicated control regions. Uncertainties due to missing higher-order QCD corrections in the production cross sections are found to be [65] $\pm10.1\%$ ($\pm7.8\%$) for the Higgs boson production via $ggH$ in SR1 (SR2), $\pm1\%$ for the $Z \rightarrow ee$ background and for VBF and VH Higgs boson production, and $\pm5\%$ for the $VV$ background. The systematic uncertainties due to the choice of parton distribution functions used in the simulation are evaluated based on the prescription described in Ref. [65] and the following values are used in this analysis: $\pm7.5\%$ for the Higgs boson production via $ggH$, $\pm2.8\%$ for the VBF and VH Higgs boson production, and $\pm4\%$ for the $VV$ background. Finally, an additional $\pm5.7\%$ systematic uncertainty [65] on $\text{Br}(H \rightarrow \tau \tau)$ is applied to the SM $H \rightarrow \tau \tau$ background.

4.4 Results of the search for LFV $H \rightarrow \tau \tau$ decays in the $\tau_{\text{had}}$ channel

A simultaneous binned maximum-likelihood fit is performed on the $m_{\tau\tau}^{\text{MMC}}$ distributions in SR1 and SR2 and on event yields in WCR and TCR to extract the LFV branching ratio $\text{Br}(H \rightarrow \tau \tau)$. The fit exploits the control regions and the distinct shapes of the $W$+jets, $Z \rightarrow \tau \tau$ and $Z \rightarrow ee$ backgrounds in the signal regions to constrain some of the systematic uncertainties. This increases the sensitivity of the analysis. The post-fit $m_{\tau\tau}^{\text{MMC}}$ distributions in SR1 and SR2 are shown in Fig. 2, and the combined $m_{\tau\tau}^{\text{MMC}}$ distribution for both signal regions is presented in Fig. 3. Figure 2 illustrates the level of agreement between data and background expectations in SR1 and SR2. No statistically significant deviations of the data from the predicted background are observed. An upper limit on the LFV branching ratio $\text{Br}(H \rightarrow \tau \tau)$ for a Higgs boson with $m_{H} = 125$ GeV is set using the CL$_{s}$ modified frequentist formalism [66] with the test statistic based on the profile likelihood ratio [67]. The observed and
5.1 Event selection and signal region definition

Events selected in the $\tau_{lep}$ channel must contain exactly two opposite-sign leptons, one an electron and the other a muon. The lepton with the higher $p_T$ is indicated by $\ell_1$ and the other by $\ell_2$. Additional kinematic criteria, based on the $p_T$ difference between the two leptons and on the angular separations between the leptons and the missing transverse momentum, are applied to suppress the SM background events, which are mainly due to the production of $Z/\gamma^* \to \tau\tau$ and of diboson ($VV$) events. Two mutually exclusive signal regions are defined: one with no central ($|\eta| < 2.4$) light-flavour jets, $\mu_{\ell \ell}$, and the other with one or more central light-flavour jets, $\mu_{\ell \ell, jets}$. The kinematic criteria defining each signal region, summarised in Table 4, are optimised following two guidelines. The first one is to maximise the signal-to-background ratio. The second one is to have, in each signal region, enough events to perform the data-driven background estimation described in Sect. 5.2.

The final discriminant used in the $\tau_{lep}$ channel is the collinear mass $m_{coll}$ defined as:

$$m_{coll} = \sqrt{2p_T^{\ell_1}(p_T^{\ell_2} + E_T^{miss})(\cosh |\eta| - \cos \Delta \phi)}.$$  

This quantity is the invariant mass of two massless particles, $\tau$ and $\ell_1$, computed with the approximation that the decay products of the $\tau$ lepton, $\ell_2$ and neutrinos, are collinear to the $\tau$, and that $E_T^{miss}$ originates from the $\nu$. In the $H \to \mu\tau$ ($H \to e\tau$) decay, $\ell_1$ is the muon (electron) and $\ell_2$ is the electron (muon). The differences in rapidity and azimuthal angle between $\ell_1$ and $\ell_2$ are indicated by $\Delta \eta$ and $\Delta \phi$. More sophisticated kinematic variables, such as MMC, do not significantly improve the sensitivity of the $\tau_{lep}$ channel.

5.2 Background estimation

For simplicity, the symmetry method is illustrated here assuming a $H \to \mu\tau$ signal. The same procedure, but with $e$ and $\mu$ exchanged, is valid under the $H \to e\tau$ assumption. The symmetry method is based on the following two premises:

<table>
<thead>
<tr>
<th>$\mu_{\ell \ell}$</th>
<th>$\mu_{\ell \ell, jets}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light leptons</td>
<td>$e^\pm \mu^\mp$</td>
</tr>
<tr>
<td>$\tau_{had}$</td>
<td>veto</td>
</tr>
<tr>
<td>Central jets</td>
<td>0</td>
</tr>
<tr>
<td>$p_T^{\ell_1}$</td>
<td>$\geq 35$ GeV</td>
</tr>
<tr>
<td>$p_T^{\ell_2}$</td>
<td>$\geq 12$ GeV</td>
</tr>
<tr>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>$\Delta \phi$</td>
<td>$\leq 0.7$</td>
</tr>
<tr>
<td>$\Delta \phi$</td>
<td>$\leq 1.0$</td>
</tr>
<tr>
<td>$\Delta \phi$</td>
<td>$\leq 1.0$</td>
</tr>
<tr>
<td>$\Delta \eta$</td>
<td>$\leq 7$ GeV</td>
</tr>
<tr>
<td>$\Delta \phi$</td>
<td>$\geq 1$ GeV</td>
</tr>
</tbody>
</table>
1. SM processes result in data that are symmetric under the exchange of prompt electrons with prompt muons to a good approximation. In other words, the kinematic distributions of prompt electrons and prompt muons are approximately the same.\(^5\)

2. flavour-violating decays of the Higgs boson break this symmetry.

Dilepton events in the dataset are divided into two mutually exclusive samples:

- **\(\mu e\) sample**: \(\ell_1\) is the muon and \(\ell_2\) is the electron (\(p_T^{\mu} \geq p_T^e\))

- **\(ee\) sample**: \(\ell_1\) is the electron and \(\ell_2\) is the muon (\(p_T^e > p_T^\mu\))

With these assumptions, the SM background is split equally between the two samples. The \(H \to \mu \tau\) signal, however, is present only in the \(\mu e\) sample because the \(p_T\) spectrum of electrons from \(H \to \mu \tau\) decays is softer then the muon \(p_T\) spectrum. The number of \(H \to \mu \tau\) events in the \(e\mu\) sample is negligible with the selection criteria described in Sect. 5.1.

For SM events the distributions of kinematic variables in the two samples are the same with good approximation. In particular, the collinear mass distribution differs between the two samples only for the narrow signal peak. The peak, present only in the distribution of the \(\mu e\) sample, is on top of the SM background, which, to a good approximation, can be modelled from the \(e\mu\) collinear mass distribution.

### 5.2.1 Asymmetries in the SM background

Although the \(e\mu-\mu e\) symmetry hypothesis is a good starting assumption, there are effects that can invalidate it and that need to be accounted for. The first effect is due to events containing misidentified and non-prompt leptons, together referred to as *non-prompt* in the following. These leptons originate from misidentified jets or from hadronic decays within jets. They contribute differently to the \(\mu e\) and \(e\mu\) samples because the origin of the non-prompt lepton is different for electrons and for muons. The second effect originates from the different dependencies on \(p_T\) and \(|\eta|\) that the trigger efficiency and reconstruction efficiency can have for electrons and muons. The non-prompt effect is accounted for by estimating the non-prompt background separately from the other backgrounds. The efficiency effect is accounted for by scaling the \(m_{\text{coll}}\) distribution of the \(e\mu\) sample with a scale factor parameterised as a function of the sub-leading lepton \(p_T\), \(p_T^\ell_2\). As shown in Sect. 5.5, the \(e\mu-\mu e\) symmetry is restored when these two effects are taken into account. Smaller effects, which might depend on other parameters such as \(\eta\) or \(p_T^{\ell_1}\), are found to be negligible.

**Events containing non-prompt leptons** The background contribution due to non-prompt leptons is estimated with the matrix method described in Refs. [68,69], which relies on the difference in identification efficiency between prompt and non-prompt leptons. Two lepton categories are defined: tight leptons, which must satisfy all the lepton identification criteria described in Sect. 2, and loose leptons, which are not required to satisfy the primary vertex and isolation criteria. By measuring separately for prompt and non-prompt leptons the tight-to-loose lepton efficiencies, defined as the fraction of loose leptons that are also tight, one can determine the non-prompt background contribution from the number of data events that have two leptons that are either loose or tight. The efficiencies for prompt and non-prompt leptons, parameterised as a function of \(p_T\) and \(|\eta|\), are derived from data with the tag-and-probe method. Prompt efficiencies are derived from an opposite-sign sample enriched in \(e\mu\) events in the same-sign sample (\(\mu^+e^-\) or \(\mu^-e^+\)) where the muon is the tag lepton.

**Asymmetry induced by the different trigger and reconstruction efficiency of electrons and muons** The efficiency to trigger on and reconstruct an \(e\mu\) event, \(\varepsilon_{e\mu}\), is different from the one of a \(\mu e\) event, \(\varepsilon_{\mu e}\). These two efficiencies can be expressed as a function of the \(p_T\) of the two leptons:

\[
\varepsilon_{e\mu} = \varepsilon_{\text{trig.}}^{e\mu} \cdot \varepsilon_{\text{reco.}}^{\ell_1=\mu} \cdot \varepsilon_{\text{reco.}}^{\ell_2=e} (p_T^{\ell_2=e})
\]

\[
\varepsilon_{\mu e} = \varepsilon_{\text{trig.}}^{\ell_1=e} \cdot \varepsilon_{\text{reco.}}^{\ell_1=e} \cdot \varepsilon_{\text{reco.}}^{\ell_2=\mu} (p_T^{\ell_2=\mu}).
\]

In this search, the leading lepton is required to have \(p_T^{\ell_1} > 35\) GeV, which is on the plateau region of the trigger and reconstruction efficiencies. Hence the ratio of the efficiencies can be approximated as:

\[
\frac{\varepsilon_{e\mu}}{\varepsilon_{\mu e}} = \frac{\varepsilon_{\text{trig.}}^{\mu e} \cdot \varepsilon_{\text{reco.}}^{\ell_1=e} \cdot \varepsilon_{\text{reco.}}^{\ell_2=e} (p_T^{\ell_2=e})}{\varepsilon_{\text{trig.}}^{e\mu} \cdot \varepsilon_{\text{reco.}}^{\ell_1=\mu} \cdot \varepsilon_{\text{reco.}}^{\ell_2=e} (p_T^{\ell_2=e})} = \frac{\varepsilon_{\text{trig.}}^{\mu e} \cdot \varepsilon_{\text{reco.}}^{\ell_2=e} (p_T^{\ell_2=e})}{\varepsilon_{\text{trig.}}^{e\mu} \cdot \varepsilon_{\text{reco.}}^{\ell_1=e} (p_T^{\ell_1=e})} \cdot \varepsilon_{\text{reco.}}^{\ell_2=e} (p_T^{\ell_2=e}) \cdot \varepsilon_{\text{trig.}}^{\ell_2=e} (p_T^{\ell_2=e}) \cdot \varepsilon_{\text{reco.}}^{\ell_1=e} (p_T^{\ell_1=e})
\]

\[
= f (p_T^{\ell_2=e}) \cdot \text{Const}.
\]

Therefore, the ratio of the \(e\mu\) and \(\mu e\) event reconstruction efficiencies can be parameterised as a function of the sub-leading lepton \(p_T\), \(f (p_T^{\ell_2=e})\). Using the fit described in...
Sect. 5.4, the parameter \( f \left( p_T^{\ell_2} \right) \) is determined in three \( p_T^{\ell_2} \) bins, 12–20, 20–30, and > 30 GeV.

5.3 Systematic uncertainties

Using the \( e\mu \) asymmetry technique, the only systematic uncertainty associated with the background prediction is due to the non-prompt background modelling. This uncertainty has two components: the first one is the limited number of tag-and-probe events used to extract the prompt and non-prompt efficiencies; the second one is the difference in kinematics, and therefore in sources of non-prompt leptons, between the events used to extract the non-prompt efficiency and the events in the signal regions. This second component is evaluated by measuring the non-prompt efficiencies in subsets of the nominal tag-and-probe sample. The subsets are obtained by applying, one at a time, the kinematic requirements of the signal regions. The ensuing uncertainties in the estimated number of non-prompt events can be as large as 10–50% for the non-prompt efficiency and 3% for the prompt efficiency, depending on the signal region.

Uncertainties related to the signal prediction are the same ones described in Sect. 4.3 with one minor difference in the uncertainty in the signal cross section due to higher-order QCD corrections. This uncertainty is split into two anticoated components: ±12% in SRwithjets and ±20% in SRnojets.

5.4 The statistical model

Assuming that the SM background is completely symmetric when exchanging \( e \leftrightarrow \mu \), the likelihood function for the collinear mass distribution of the \( e\mu \) and \( \mu e \) samples can be written as:

\[
L(b_i, \mu) = \prod_i^{N_m} \text{Pois}(n_i | b_i) \times \text{Pois}(m_i | b_i + \mu s_i),
\]

(3)

where \( n_i \) (\( m_i \)) is the number of \( e\mu \) (\( \mu e \)) events in the \( i \)-th of the \( N_m \) \( m \) bins. The number of background events in the \( i \)-th \( m \) bin is indicated by \( b_i \), and \( s_i \) is the number of \( H \rightarrow \tau\mu \) events in the \( i \)-th mass bin. The number of signal events \( \sum_i s_i \) is normalised to a branching ratio \( \text{Br}(H \rightarrow \tau\mu) = 1\% \), multiplied by a signal strength \( \mu \). The likelihood for the \( m \) distributions with a \( H \rightarrow e\tau \) signal can be defined in a similar way. The contributions due to non-prompt leptons add to the \( e\mu \) and \( \mu e \) terms and are denoted by \( N_{\text{np}} \) and \( M_{\text{np}} \), along with their uncertainties, \( \sigma_{N_{\text{np}}} \) and \( \sigma_{M_{\text{np}}} \). The numbers of non-prompt events in each bin, \( N_{\text{np}} \) and \( M_{\text{np}} \), are treated as Gaussian nuisance parameters.

The \( f \left( p_T^{\ell_2} \right) \) correction, described in Sect. 5.2, is implemented by performing the fit separately in \( N_{p_T^{\ell_2}} = 3 \) \( p_T^{\ell_2} \) bins, labelled with the index \( j \). The corrective scale factor \( A_j \) corresponding to the \( f \left( p_T^{\ell_2} \right) \) value in the \( m \) bin \( i \) and \( p_T^{\ell_2} \) bin \( j \), multiplies the \( e\mu \) yield \( b_{ij} \). These scale factors are treated in the statistical model as unconstrained nuisance parameters.

Adding up the symmetric contribution (\( b_{ij} \)), the non-prompt contributions (\( N_{\text{np}} \) and \( M_{\text{np}} \) ), the \( f \left( p_T^{\ell_2} \right) \) correction, and the signal contribution (\( s_{ij} \)), the likelihood is written as:

\[
L(\mu, b_{ij}, N_{\text{np}}, M_{\text{np}}, s_{ij}) = \prod_i^{N_m} \prod_j^{N_{p_T^{\ell_2}}} \text{Pois}(n_{ij} | A_j b_{ij} + N_{\text{np}}) \times \text{Pois}(m_{ij} | b_{ij} + M_{\text{np}} + \mu s_{ij}) \times \text{Gaus}(n_{ij} | N_{\text{np}}, \sigma_{N_{\text{np}}}) \times \text{Gaus}(m_{ij} | M_{\text{np}}, \sigma_{M_{\text{np}}}).
\]

(4)

5.5 Background model validation

The symmetry-based method is validated with simulation and with data. The validation with simulated samples is performed by comparing the signal strength measured in the SR with background samples, and with signal samples corresponding to several non-zero LFV branching ratios. The validation with data is performed in a validation region (VR) defined as SRnojets, but with at least one angular requirement reversed, \( \Delta \phi(\ell_1, \ell_2) \) or \( \Delta \phi(\ell_1, E_{\text{miss}}) \).

The validation procedure consists of comparing the data, or the sum of the simulated background samples, to the total background estimated from the statistical model. The comparison is done for the \( e\mu \) sample and the \( \mu e \) one. With the simulated samples, it is also verified that the symmetric background and the \( f \left( p_T^{\ell_2} \right) \) do not depend on the presence of an LFV signal.

Generated pseudo-experiments are used to confirm that the statistical model is unbiased. No significant discrepancy was found between the injected signal strength and its fitted value up to LFV branching ratios of 10%.

5.6 Results of the search for LFV \( H \rightarrow e\tau/\mu\tau \) decays in the \( \tau\ell\) channel

Figure 4 compares the observed data to the yields expected from the symmetry-based statistical model. The comparison, combining the different \( p_T^{\ell_2} \) bins, shows the symmetric component of the background (\( b_{ij} \)) as a dashed line, and the total background estimation including the contribution from events containing misidentified and non-prompt leptons as a full line. As can be seen, the background estimation is in good agreement with the data over the full mass range. Table 5 sum-
Fig. 4 Collinear mass distributions in the $\tau_{lep}$ channel: background estimate compared to the events observed in the data in the SR$_{nojets}$ (top) and SR$_{withjets}$ (bottom). Left $e\mu$ channel. Right $\mu e$ channel. In these plots, events from the three $f(p_{T/z})$ bins are combined, although the fit parameters are different in each $f(p_{T/z})$ bin. The signal expected for a $\text{Br}(H \rightarrow \mu \tau) = 1\%$ is shown in the $\mu e$ channel.

marinates the fit results in the data in SR$_{nojets}$ and SR$_{withjets}$: the fitted $f(p_{T/z})$ scale factors, the symmetric background component ($\sum_{i=N_{\text{coll}}^{N_{\text{coll}}}} b_{ij}$) in each $p_{T/z}$ bin, and the non-prompt estimate in the $\mu e$ and the $e\mu$ channels. The excellent level of agreement between the fitted number of events and the observed number is due to the many unconstrained parameters in the fit.

The expected and observed 95% CL upper limits on branching ratios as well as their best fit values are calculated using the statistical model described in Sect. 5.4. Table 6 presents a summary of results for the individual categories.
Table 5 A summary of the fit results in the \( \tau_{\text{lep}} \) channel. The values of the fit parameters \( f \left( p_{T}^{\ell} \right) \), which account for the ratio of the \( e\mu \) and \( \mu\mu \) event reconstruction efficiencies described in Sect. 5.2, are obtained from a background-only fit, and reported for each signal region and for each \( p_{T}^{\ell} \) bin. The expected and observed yields correspond to the number of events used in the fit, representing the 0–300 GeV \( m_{\text{coll}} \) range shown in Fig. 4. The quoted uncertainties in the expected yields represent the statistical (first) and systematic (second) uncertainties, respectively. The post-fit values of systematic uncertainties are provided for the background predictions. The signal predictions are given for \( \text{Br}(H \rightarrow \mu\tau) = 1\% \) in the \( e\mu \) sample and for \( \text{Br}(H \rightarrow \mu\tau) = 1\% \) in the \( \mu\mu \) sample.

<table>
<thead>
<tr>
<th>( p_{T}^{\ell} ) bin (GeV)</th>
<th>( f \left( p_{T}^{\ell} \right) )</th>
<th>LFV Signal, Br = 1%</th>
<th>Total backg.</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_{\text{hadlets}} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12–20</td>
<td>1.11 ± 0.06</td>
<td>( e\mu )</td>
<td>14.9 ± 0.4 ± 2.7</td>
<td>1219 ± 24 ± 27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \mu\mu )</td>
<td>10.7 ± 0.4 ± 2.3</td>
<td>1033 ± 25 ± 20</td>
</tr>
<tr>
<td>20–30</td>
<td>1.07 ± 0.08</td>
<td>( e\mu )</td>
<td>15.1 ± 0.4 ± 2.7</td>
<td>998 ± 22 ± 25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \mu\mu )</td>
<td>12.4 ± 0.4 ± 2.2</td>
<td>950 ± 23 ± 21</td>
</tr>
<tr>
<td>≥30</td>
<td>1.01 ± 0.07</td>
<td>( e\mu )</td>
<td>12.5 ± 0.4 ± 2.2</td>
<td>455 ± 17 ± 16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \mu\mu )</td>
<td>11.4 ± 0.4 ± 2.0</td>
<td>458 ± 16 ± 14</td>
</tr>
<tr>
<td>( \tau_{\text{w/hadlets}} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12–20</td>
<td>1.07 ± 0.10</td>
<td>( e\mu )</td>
<td>5.9 ± 0.3 ± 1.1</td>
<td>222 ± 10 ± 11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \mu\mu )</td>
<td>3.9 ± 0.2 ± 0.9</td>
<td>181 ± 10 ± 9</td>
</tr>
<tr>
<td>20–30</td>
<td>1.24 ± 0.16</td>
<td>( e\mu )</td>
<td>5.4 ± 0.2 ± 1.1</td>
<td>187 ± 9 ± 11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \mu\mu )</td>
<td>4.5 ± 0.2 ± 0.9</td>
<td>161 ± 9 ± 9</td>
</tr>
<tr>
<td>≥30</td>
<td>1.13 ± 0.10</td>
<td>( e\mu )</td>
<td>5.5 ± 0.2 ± 1.0</td>
<td>251 ± 11 ± 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \mu\mu )</td>
<td>4.9 ± 0.2 ± 0.9</td>
<td>229 ± 11 ± 11</td>
</tr>
</tbody>
</table>

and their combination can be found in Table 6 for both the \( H \rightarrow e\tau \) and \( H \rightarrow \mu\tau \) hypotheses.

6 Combined results of the search for LFV \( H \rightarrow e\tau/\mu\tau \) decays

The results of the individual searches for the LFV \( H \rightarrow e\tau \) and \( H \rightarrow \mu\tau \) decays in the \( \tau_{\text{had}} \) (including the result from Ref. [22]) and \( \tau_{\text{lep}} \) channels presented in Sects. 4.4 and 5.6 are statistically combined. The two channels use different background estimation techniques, leading to uncorrelated systematic uncertainties in the background predictions. The systematic uncertainties for the LFV signal are treated as 100% correlated between the two channels. Table 6 presents a summary of results for the expected and observed 95% CL upper limits and the best fit values for the branching ratios for the individual categories and their combination. There is no indication of a signal in the search for the LFV \( H \rightarrow e\tau \) decays. The combined observed, and the median expected, 95% CL upper limits on \( \text{Br}(H \rightarrow e\tau) \) for a Higgs boson with \( m_{H} = 125 \) GeV are 1.04% and 1.21 ± 0.34%, respectively. A small \( \sim 1\sigma \) excess of data over the predicted background is observed in the search for the LFV \( H \rightarrow \mu\tau \) decays. It is mostly driven by a 1.3\( \sigma \) excess in the earlier search in the \( \mu_{\text{had}} \) channel [22]. This corresponds to a best fit value for the branching ratio \( \text{Br}(H \rightarrow \mu\tau) = (0.53 \pm 0.51)\% \). In the absence of any significant signal, an upper limit on the LFV branching ratio \( \text{Br}(H \rightarrow \mu\tau) \) for a Higgs boson with \( m_{H} = 125 \) GeV is set. The corresponding observed, and the median expected, 95% CL upper limits are 1.43% and 1.01 ± 0.29%, respectively. The upper limits on the LFV decays of the Higgs boson are summarised in Fig. 5.

7 Search for \( Z \rightarrow \mu\tau \) using the \( \tau_{\text{had}} \) channel

The search for \( Z \rightarrow \mu\tau \) events is based on \( \mu_{\text{had}} \) final state and utilises the same strategy as the \( H \rightarrow \mu\tau \) analysis documented in Ref. [22], and applied to the \( H \rightarrow e_{\text{had}} \) search described above. The final state is characterised by the presence of a energetic muon and a \( \tau_{\text{had}} \) of opposite charge and the presence of moderate \( E_{\text{miss}}^{\tau} \) aligned with the \( \tau_{\text{had}} \) direction. The typical transverse momenta of the muon and of the \( \tau_{\text{had}} \) are somewhat softer than those expected in Higgs boson LFV decay, due to the lower mass of the \( Z \) boson. The main backgrounds are the same as those observed in \( H \rightarrow \mu_{\text{had}} \) analyses, namely: \( Z \rightarrow \tau\tau \), \( W+\text{jets, multi-jet} \), \( H \rightarrow \tau\tau \), diboson and top backgrounds. The \( m_{\mu\tau}^{\text{MMC}} \) variable is used to extract the signal using the same fit procedure and estimation of systematic uncertainties as for the \( H \rightarrow \mu_{\text{had}} \) search. The corresponding Higgs boson LFV contribution is assumed to be negligible.

The \( Z \rightarrow \mu\tau \) analysis differs from the \( H \rightarrow \mu_{\text{had}} \) one as follows:
computed under the assumption that either $\text{Br}(H \rightarrow \mu \tau) = 0$ or $\text{Br}(H \rightarrow \mu \tau) = 0$. The expected and observed 95% confidence level (CL) upper limits and the best fit values for the branching ratios for the individual categories and their combination. The $\mu \tau_{\text{had}}$ channel is from Ref. [22].

Table 6 Results of the search for the LFV $H \rightarrow e \tau$ and $H \rightarrow \mu \tau$ decays. The limits are computed under the assumption that either $\text{Br}(H \rightarrow \mu \tau) = 0$ or $\text{Br}(H \rightarrow e \tau) = 0$. The expected and observed 95% confidence level (CL) upper limits and the best fit values for the branching ratios for the individual categories and their combination. The $\mu \tau_{\text{had}}$ channel is from Ref. [22].

<table>
<thead>
<tr>
<th>Channel</th>
<th>Category</th>
<th>Expected limit (%)</th>
<th>Observed limit (%)</th>
<th>Best fit Br (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H \rightarrow e \tau_{\text{had}}$</td>
<td>SR1</td>
<td>2.81$^{+0.06}_{-0.79}$</td>
<td>3.0</td>
<td>0.33$^{+1.48}_{-1.59}$</td>
</tr>
<tr>
<td></td>
<td>SR2</td>
<td>2.95$^{+1.16}_{-0.82}$</td>
<td>2.24</td>
<td>$-1.33^{+1.56}_{-1.80}$</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>2.07$^{+0.82}_{-0.58}$</td>
<td>1.81</td>
<td>$-0.47^{+1.08}_{-1.18}$</td>
</tr>
<tr>
<td></td>
<td>$\mu \tau_{\text{had}}$</td>
<td>1.66$^{+0.72}_{-0.46}$</td>
<td>1.45</td>
<td>$-0.45^{+0.89}_{-0.97}$</td>
</tr>
<tr>
<td>$H \rightarrow e \tau_{\text{lep}}$</td>
<td>$\text{SR}_{\text{with}}$</td>
<td>3.33$^{+1.60}_{-0.93}$</td>
<td>3.99</td>
<td>0.74$^{+1.59}_{-1.62}$</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>1.48$^{+0.60}_{-0.42}$</td>
<td>1.36</td>
<td>$-0.26^{+0.82}_{-0.79}$</td>
</tr>
<tr>
<td>$H \rightarrow \mu \tau$</td>
<td>Combined</td>
<td>1.21$^{+0.49}_{-0.34}$</td>
<td>1.04</td>
<td>$-0.34^{+0.64}_{-0.66}$</td>
</tr>
</tbody>
</table>

\begin{align*}
\text{SR1} & : & 1.60 & \pm & 0.45 & & 1.55 & & -0.07 & \pm & 0.81 & \pm & 0.86 \\
\text{SR2} & : & 1.75 & \pm & 0.49 & & 3.51 & & 1.94 & \pm & 0.89 & \\
\text{Combined} & : & 1.24 & \pm & 0.50 & & 1.85 & & 0.77 & \pm & 0.62 & \\
\text{SR}_{\text{with}} & : & 2.03 & \pm & 0.93 & & 2.38 & & 0.31 & \pm & 0.99 & \\
\text{Combined} & : & 1.73 & \pm & 0.74 & & 1.79 & & 0.03 & \pm & 0.88 & \pm & 0.86
\end{align*}

Fig. 5 Upper limits on LFV decays of the Higgs boson in the $H \rightarrow e \tau$ hypothesis (left) and $H \rightarrow \mu \tau$ hypothesis (right). The limits are computed under the assumption that either $\text{Br}(H \rightarrow \mu \tau) = 0$ or $\text{Br}(H \rightarrow e \tau) = 0$. The $\mu \tau_{\text{had}}$ channel is from Ref. [22].

Table 7 Summary of the $Z \rightarrow \mu \tau_{\text{had}}$ event selection criteria used to define the signal and control regions (see text)

<table>
<thead>
<tr>
<th>Cut</th>
<th>SR1</th>
<th>SR2</th>
<th>WCR</th>
<th>TCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T(\mu)$</td>
<td>$&gt;30$ GeV</td>
<td>$&gt;30$ GeV</td>
<td>$&gt;30$ GeV</td>
<td>$&gt;30$ GeV</td>
</tr>
<tr>
<td>$p_T(\tau_{\text{had}})$</td>
<td>$&gt;30$ GeV</td>
<td>$&gt;30$ GeV</td>
<td>$&gt;30$ GeV</td>
<td>$&gt;30$ GeV</td>
</tr>
<tr>
<td>$</td>
<td>y(\mu) - y(\tau_{\text{had}})</td>
<td>$</td>
<td>$&lt;2$</td>
<td>$&lt;2$</td>
</tr>
<tr>
<td>$m_T^{\text{missing}}$</td>
<td>$&gt;30$ and $&lt;75$ GeV</td>
<td>$&lt;30$ GeV</td>
<td>$&gt;60$ GeV</td>
<td>--</td>
</tr>
<tr>
<td>$m_T^{\text{had,missing}}$</td>
<td>$&lt;20$ GeV</td>
<td>$&lt;45$ GeV</td>
<td>$&gt;40$ GeV</td>
<td>--</td>
</tr>
<tr>
<td>$N_{\text{jet}}$</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>$&gt;1$</td>
</tr>
<tr>
<td>$N_{\text{b-jet}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$&gt;0$</td>
</tr>
</tbody>
</table>
The signal and control regions are defined in the same way as in the $H \rightarrow \mu\tau_{\text{had}}$ analysis, but the cut values are lowered to match the kinematics of $Z$ boson decay products. The exact definition is given in Table 7.

### Table 8  Data yields, signal and post-fit OS–SS background predictions (see Eq. (1)) for the $Z \rightarrow \mu\tau_{\text{had}}$ 80 GeV < $m_{\mu\tau}^\text{MMC}$ <115 GeV region.

<table>
<thead>
<tr>
<th></th>
<th>SR1</th>
<th>SR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>86 ±2</td>
<td>56 ±2</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>3260 ±30</td>
<td>7060 ±40</td>
</tr>
<tr>
<td>$W$+jets</td>
<td>1350 ±70</td>
<td>7060 ±40</td>
</tr>
<tr>
<td>Same-Sign events</td>
<td>1110 ±40</td>
<td>930 ±30</td>
</tr>
<tr>
<td>$VV + Z \rightarrow \mu\mu$</td>
<td>410 ±60</td>
<td>240 ±60</td>
</tr>
<tr>
<td>$H \rightarrow \tau\tau$</td>
<td>25.1 ±0.5</td>
<td>41 ±1</td>
</tr>
<tr>
<td>Top</td>
<td>22 ±4</td>
<td>15 ±4</td>
</tr>
<tr>
<td>Total background</td>
<td>6170 ±100</td>
<td>8880 ±100</td>
</tr>
<tr>
<td>Data</td>
<td>6134</td>
<td>8982</td>
</tr>
</tbody>
</table>

- The signal and control regions are defined in the same way as in the $H \rightarrow \mu\tau_{\text{had}}$ analysis, but the cut values are lowered to match the kinematics of $Z$ boson decay products. The exact definition is given in Table 7.

- The LFV $H \rightarrow \mu\tau_{\text{had}}$ signal sample is replaced with a LFV $Z \rightarrow \mu\tau$ signal sample.

- The shape correction for $W$+jets in SR1 is obtained from the $m_{\mu\tau}^\text{MMC}$ > 110 GeV sideband in SR1.

- Due to larger $W$+jets contribution in SR1 and SR2, the shape corrections for the $W$+jets samples are calculated using a three-dimensional binning scheme in $p_T(\tau_{\text{had}})$, $|\eta(\mu)| - |\eta(\tau_{\text{had}})|$ and $N_jet$.

- The $W$+jets extrapolation uncertainty, which accounts for the difference between the $W$+jets ALPGEN PYTHIA and HERWIG samples, is also included as a shape uncertainty.

The numbers of observed events and background in each of the regions are given in Table 8. The efficiencies for simulated $Z \rightarrow \mu\tau$ signal events to pass the SR1 and SR2 selections are 1.2 and 0.8%, respectively. Figure 6 shows the $m_{\mu\tau}^\text{MMC}$ distribution for data and predicted background in each of the signal regions. The discrepancy observed in the $m_{\mu\tau}^\text{MMC}$ range 80–100 GeV of SR1 was studied carefully. All the other SR1 distributions, including lepton momenta, transverse masses, and missing transverse momentum, are in excellent agreement with the predictions, and the background shapes are constrained in the control regions as well as in SR2. This discrepancy is hence attributed to a statistical fluctuation.

No excess of data is observed and the CL$_s$ limit-setting technique is used to calculate the observed and expected lim-
its on the branching ratio for $Z \to \mu\tau$ decays. The observed 95% CL limit on $\text{Br}(Z \to \mu\tau)$ is $1.7 \times 10^{-5}$, which is lower than the expected upper limit of $\text{Br}(Z \to \mu\tau) = 2.6 \times 10^{-5}$, but still within the 2σ band. This corresponds to a best fit value for the branching ratio $\text{Br}(Z \to \mu\tau) = -1.6^{+1.3}_{-1.0} \times 10^{-5}$. The results for the different signal regions are summarised in Table 9.

### 8 Summary

Searches for lepton-flavour-violating decays of the $Z$ and Higgs bosons are performed using a data sample of proton-proton collisions recorded by the ATLAS detector at the LHC corresponding to an integrated luminosity of 20.3 fb$^{-1}$ at $\sqrt{s} = 8$ TeV. Three LFV decays are considered: $H \to e\tau$, $H \to \mu\tau$, and $Z \to \mu\tau$. The search for the Higgs boson decays is performed in the final states where the $\tau$-lepton decays either to hadrons or to leptons (electron or muon). The search for the $Z$ boson decays is performed in the final state with the $\tau$-lepton decaying into hadrons. No significant excess is observed, and upper limits on the LFV branching ratios are set. The observed and the median expected 95% CL upper limits on $\text{Br}(H \to e\tau)$ are 1.04% and $1.21^{+0.49}_{-0.34}$%, respectively. This direct search for the $H \to e\tau$ decays places significantly more stringent constraints on $\text{Br}(H \to e\tau)$ than earlier indirect estimates. In the search for the $H \to \mu\tau$ decays, the observed and the median expected 95% CL upper limits on $\text{Br}(H \to \mu\tau)$ are 1.43% and $1.01^{+0.40}_{-0.29}$%, respectively. A small deficit of data compared to the predicted background is observed in the search for the LFV $Z \to \mu\tau$ decays. The observed and the median expected 95% CL upper limits on $\text{Br}(Z \to \mu\tau)$ are $1.69 \times 10^{-5}$ and $2.58 \times 10^{-5}$, respectively.

### Acknowledgements

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 thank Avital Dery and Aielet Efrati for their significant contribution and dedication to this study. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; EMS, MPG and MPT, Germany; SNSF, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRS, Morocco; FOM and NWO, The Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MINE/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, SNF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, UK; DOE and NSF, USA. 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, FP7, Horizon2020 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; Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, UK. 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), Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, UK. 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), HERA-B, MINEPI (Italy), NL-T1 (The Netherlands), PIC (Spain), 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. [70].

### Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Funded by SCOAP3.

### References


Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil; Electrical Circuits Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora, Brazil; Federal University of Sao Joao del Rei (UFSJ), Sao Joao del Rei, Brazil; Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil

Physics Department, Brookhaven National Laboratory, Upton, NY, USA

Transilvania University of Brasov, Brasov, Romania; National Institute of Physics and Nuclear Engineering, Bucharest, Romania; Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj Napoca, Romania; University Politehnica Bucharest, Bucharest, Romania; West University in Timisoara, Timisoara, Romania

Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina

Cavendish Laboratory, University of Cambridge, Cambridge, UK

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

CERN, Geneva, Switzerland

Enrico Fermi Institute, University of Chicago, Chicago, IL, USA

Departamento de Física, Pontificia Universidad Católica de Chile, Santiago, Chile; Departamento de Física, Universidad Técnica Federico Santa María, Valparaíso, Chile

Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; Department of Modern Physics, University of Science and Technology of China, Anhui, China; Department of Physics, Nanjing University, Jiangsu, China; School of Physics, Shandong University, Shandong, China; Department of Physics and Astronomy, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University (also affiliated with PKU-CHEP), Shanghai, China; Physics Department, Tsinghua University, Beijing 100084, China

Laboratoire de Physique Corpusculaire, Clermont Université and Université Blaise Pascal and CNRS/IN2P3, Clermont-Ferrand, France

Nevis Laboratory, Columbia University, Irvington, NY, USA

Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy; Dipartimento di Fisica, Università della Calabria, Rende, Italy

Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Krakow, Poland; Marian Smoluchowski Institute of Physics, Jagiellonian University, Kraków, Poland

Institute of Nuclear Physics, Polish Academy of Sciences, Kraków, Poland

Physics Department, Southern Methodist University, Dallas, TX, USA

Physics Department, University of Texas at Dallas, Richardson, TX, USA

DESY, Hamburg and Zeuthen, Germany

Institut für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany

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

Department of Physics, Duke University, Durham, NC, USA

SUPA-School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK

INFN Laboratori Nazionali di Frascati, Frascati, Italy

Fakultät für Mathematik und Physik, Albert-Ludwigs-Universität, Freiburg, Germany

Section de Physique, Université de Genève, Geneva, Switzerland

INFN Sezione di Genova, Genoa, Italy; Dipartimento di Fisica, Università di Genova, Genoa, Italy

E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi, Georgia; High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia

II Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany

SUPA-School of Physics and Astronomy, University of Glasgow, Glasgow, UK

II Physikalisches Institut, Georg-August-Universität, Göttingen, Germany

Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France

Department of Physics, Hampton University, Hampton, VA, USA

Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, MA, USA

Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; ZITI Institut für technische Informatik, Ruprecht-Karls-Universität Heidelberg, Mannheim, Germany

Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima, Japan
(a) Department of Physics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong; (b) Department of Physics, The University of Hong Kong, Hong Kong, China; (c) Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

Department of Physics, Indiana University, Bloomington, IN, USA

Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria

University of Iowa, Iowa City, IA, USA

Department of Physics and Astronomy, Iowa State University, Ames, IA, USA

Joint Institute for Nuclear Research, JINR Dubna, Dubna, Russia

KEK, High Energy Accelerator Research Organization, Tsukuba, Japan

Graduate School of Science, Kobe University, Kobe, Japan

Faculty of Science, Kyoto University, Kyoto, Japan

Kyoto University of Education, Kyoto, Japan

Department of Physics, Kyushu University, Fukuoka, Japan

Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina

Physics Department, Lancaster University, Lancaster, UK

(a) INFN Sezione di Lecce, Lecce, Italy; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy

Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK

Department of Physics, Jožef Stefan Institute, University of Ljubljana, Ljubljana, Slovenia

Schule of Physics and Astronomy, Queen Mary University of London, London, UK

Department of Physics, Royal Holloway University of London, Surrey, UK

Department of Physics and Astronomy, University College London, London, UK

Louisiana Tech University, Ruston, LA, USA

Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France

Fysiska institutionen, Lunds universitet, Lund, Sweden

(a) Departamento de Fisica Teorica C-15, Universidad Autonoma de Madrid, Madrid, Spain

Institut für Physik, Universität Mainz, Mainz, Germany

School of Physics and Astronomy, University of Manchester, Manchester, UK

CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France

Department of Physics, University of Massachusetts, Amherst, MA, USA

Department of Physics, McGill University, Montreal, QC, Canada

School of Physics, University of Melbourne, Melbourne, VIC, Australia

Department of Physics, The University of Michigan, Ann Arbor, MI, USA

Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA

(a) INFN Sezione di Milano, Milan, Italy; (b) Dipartimento di Fisica, Università di Milano, Milan, Italy

B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Republic of Belarus

National Scientific and Educational Centre for Particle and High Energy Physics, Minsk, Republic of Belarus

Group of Particle Physics, University of Montreal, Montreal, QC, 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, Munich, Germany

Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Munich, Germany

Nagasaki Institute of Applied Science, Nagasaki, Japan

Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan

(a) INFN Sezione di Napoli, Naples, Italy; (b) Dipartimento di Fisica, Università di Napoli, Naples, Italy

Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA

Institute of Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands

Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, The Netherlands

Department of Physics, Northern Illinois University, DeKalb, IL, USA

Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia
Department of Physics and Astronomy, University of Sussex, Brighton, UK
School of Physics, University of Sydney, Sydney, NSW, Australia
Institute of Physics, Academia Sinica, Taipei, Taiwan
Department of Physics, Technion: Israel Institute of Technology, Haifa, Israel
Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
International Center for Elementary Particle Physics and Department of Physics, The University of Tokyo, Tokyo, Japan
Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo, Japan
Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
Department of Physics, University of Toronto, Toronto, ON, Canada
\((a)\)TRIUMF, Vancouver, BC, Canada; \((b)\)Department of Physics and Astronomy, York University, Toronto, ON, Canada
Faculty of Pure and Applied Sciences, and Center for Integrated Research in Fundamental Science and Engineering, University of Tsukuba, Tsukuba, Japan
Department of Physics and Astronomy, Tufts University, Medford, MA, USA
Department of Physics and Astronomy, University of California Irvine, Irvine, CA, USA
\((a)\)INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy; \((b)\)ICTP, Trieste, Italy; \((c)\)Dipartimento di Chimica Fisica e Ambiente, Università di Udine, Udine, Italy
Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
Department of Physics, University of Illinois, Urbana, IL, USA
Instituto de Fisica Corpuscular (IFIC) and Departamento de Fisica Atomica, Molecular y Nuclear and Departamento de Ingeniería Electrónica and Instituto de Microelectrónica de Barcelona (IMB-CNM), University of Valencia and CSIC, Valencia, Spain
Department of Physics, University of British Columbia, Vancouver, BC, Canada
Department of Physics and Astronomy, University of Victoria, Victoria, BC, Canada
Department of Physics, University of Warwick, Coventry, UK
Waseda University, Tokyo, Japan
Department of Particle Physics, The Weizmann Institute of Science, Rehovot, Israel
Department of Physics, University of Wisconsin, Madison, WI, USA
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität, Würzburg, Germany
Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal, Germany
Department of Physics, Yale University, New Haven, CT, USA
Yerevan Physics Institute, Yerevan, Armenia
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France
\(a\) Also at Department of Physics, King’s College London, London, UK
\(b\) Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
\(c\) Also at Novosibirsk State University, Novosibirsk, Russia
\(d\) Also at TRIUMF, Vancouver, BC, Canada
\(e\) Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY, USA
\(f\) Also at Department of Physics, California State University, Fresno, CA, USA
\(g\) Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
\(h\) Also at Departament de Fisica de la Universitat Autonoma de Barcelona, Barcelona, Spain
\(i\) Also at Departamento de Física e Astronomia, Faculdade de Ciencias, Universidade do Porto, Porto, Portugal
\(j\) Also at Tomsk State University, Tomsk, Russia
\(k\) Also at Universita di Napoli Parthenope, Naples, Italy
\(l\) Also at Institute of Particle Physics (IPP), Victoria, BC, Canada
\(m\) Also at National Institute of Physics and Nuclear Engineering, Bucharest, Romania
\(n\) Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia
\(o\) Also at Department of Physics, The University of Michigan, Ann Arbor, MI, USA
\(p\) Also at Centre for High Performance Computing, CSIR Campus, Rosebank, Cape Town, South Africa
\(q\) Also at Louisiana Tech University, Ruston, LA, USA
\(r\) Also at Institucio Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
Also at Graduate School of Science, Osaka University, Osaka, Japan
Also at Department of Physics, National Tsing Hua University, Hsinchu, Taiwan
Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
Also at Department of Physics, The University of Texas at Austin, Austin, TX, USA
Also at Institute of Theoretical Physics, Ilija State University, Tbilisi, Georgia
Also at CERN, Geneva, Switzerland
Also at Georgian Technical University (GTU), Tbilisi, Georgia
Also at Ochadai Academic Production, Ochanomizu University, Tokyo, Japan
Also at Manhattan College, New York, NY, USA
Also at Hellenic Open University, Patras, Greece
Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan
Also at School of Physics, Shandong University, Shandong, China
Also at Moscow Institute of Physics and Technology, State University, Dolgoprudny, Russia
Also at Section de Physique, Université de Genève, Geneva, Switzerland
Also at Eotvos Lorand University, Budapest, Hungary
Also at International School for Advanced Studies (SISSA), Trieste, Italy
Also at Department of Physics and Astronomy, University of South Carolina, Columbia, SC, USA
Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China
Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria
Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
Also at National Research Nuclear University MEPhI, Moscow, Russia
Also at Department of Physics, Stanford University, Stanford, CA, USA
Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
Also at Flensburg University of Applied Sciences, Flensburg, Germany
Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia
Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
Deceased