Search for lepton-flavour-violating $H \to \mu\tau$ decays of the Higgs boson with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: A direct search for lepton-flavour-violating $H \to \mu\tau$ decays of the recently discovered Higgs boson with the ATLAS detector at the LHC is presented. The analysis is performed in the $H \to \mu\gamma_{\text{had}}$ channel, where $\gamma_{\text{had}}$ is a hadronically decaying $\tau$-lepton. The search is based on the data sample of proton-proton collisions collected by the ATLAS experiment corresponding to an integrated luminosity of 20.3 fb$^{-1}$ at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. No statistically significant excess of data over the predicted background is observed. The observed (expected) 95% confidence-level upper limit on the branching fraction, $\text{Br}(H \to \mu\tau)$, is 1.85% (1.24%).

KEYWORDS: Hadron-Hadron Scattering, Beyond Standard Model, Higgs physics, Lepton production

ArXiv ePrint: 1508.03372
1 Introduction

The observation of the Higgs boson [1, 2] with a mass of about 125 GeV [3] by the ATLAS and CMS experiments is a great success of the Large Hadron Collider (LHC) physics program at CERN. The next important step in this program are searches for signs of new physics beyond the Standard Model (SM) and detailed studies of the Higgs boson properties. Direct evidence for physics beyond the SM could be indicated via lepton-flavour-violating (LFV) Higgs boson decays. If the SM is replaced with an effective field theory, which has a single Higgs boson and is required to be renormalizable only to a finite mass scale, then LFV couplings may be introduced [4]. LFV decays can also occur naturally in models with more than one Higgs doublet [5–8], composite Higgs models [9, 10], models with flavour symmetries [11], Randall-Sundrum models [12] and many others [13–18].

There are three possibilities for LFV effects mediated via virtual Higgs bosons: \( \mu \rightarrow e \), \( \tau \rightarrow \mu \), and \( \tau \rightarrow e \) transitions. Indirect experimental constraints are reviewed and translated into constraints on \( \text{Br}(H \rightarrow e\mu, \mu\tau, e\tau) \) in recent papers [4, 19]. Searches for \( \mu \rightarrow e\gamma \) [20] place a very stringent constraint on \( H \rightarrow e\mu \) decays: \( \text{Br}(H \rightarrow e\mu) < O(10^{-8}) \) [4, 19]. The indirect constraints on \( H \rightarrow \mu\tau, e\tau \) decays mostly come from searches for \( \tau \rightarrow \mu\gamma, e\gamma \) [21–23] or other rare \( \tau \)-lepton decays [24], as well as from measurements of the anomalous magnetic moment of the muon and the electron [25] and are much less stringent: \( \text{Br}(H \rightarrow \mu\tau, e\tau) < O(10\%) \) [4, 19]. 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 [19]. It is also important to note that the presence of a \( H \rightarrow \mu\tau \) signal would essentially exclude the presence of a
$H \rightarrow e\tau$ signal, and vice versa, at an experimentally observable level at the LHC due to strong experimental bounds on $\mu \rightarrow e\gamma$ decays [19]. The CMS Collaboration has recently performed the first direct search for LFV $H \rightarrow \mu\tau$ decays [26] and reported a slight excess (2.4 standard deviations) of data over the predicted background. Their results give a 1.57% upper limit on $\text{Br}(H \rightarrow \mu\tau)$ at the 95% confidence level (CL).

This paper presents a search for lepton-flavour-violating $H \rightarrow \mu\tau$ decays of the recently discovered Higgs boson with one muon and one hadronically decaying $\tau$-lepton ($\tau_{\text{had}}$) in the final state. All Higgs boson production processes are considered. The analysis is based on the data sample of $pp$ collisions which was collected at a centre-of-mass energy of $\sqrt{s} = 8\text{ TeV}$ and corresponds to an integrated luminosity of $20.3\text{ fb}^{-1}$.

2 The ATLAS detector and object reconstruction

The ATLAS detector\(^1\) is described in detail in ref. [27]. ATLAS consists of an inner tracking detector (ID) covering the range $|\eta| < 2.5$, surrounded by a superconducting solenoid providing a 2 T magnetic field, electromagnetic ($|\eta| < 3.2$) and hadronic calorimeters ($|\eta| < 4.9$) and a muon spectrometer (MS) ($|\eta| < 2.7$) with a toroidal magnetic field.

The signature of LFV $H \rightarrow \mu\tau$ decays used in this search is characterised by the presence of an energetic muon, originating directly from a Higgs boson decay and carrying roughly half of its energy, and the hadronic decay products of a $\tau$-lepton. The data were collected with a single-muon trigger with a transverse momentum, $p_T = p \sin \theta$, threshold of $p_T = 24\text{ GeV}$. The $H \rightarrow \tau\tau$ and the LFV $H \rightarrow \mu\tau$ signatures with a muon and $\tau_{\text{had}}$ in the final state share many common features. Therefore, the object definitions and data quality cuts used in this analysis are the same as those in the recently published ATLAS search for $H \rightarrow \tau\tau$ decays [28]. A brief description of the object definitions is provided below.

Muon candidates are reconstructed using an algorithm [29] that combines information from the ID and the MS. 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 $|\eta| < 2.5$. Typical reconstruction and identification efficiencies for muons satisfying these selection criteria are above 95% [29]. Exactly one identified muon is required in this analysis.

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeters matched to tracks in the ID. They are required to have a transverse energy, $E_T = E \sin \theta$, greater than 15 GeV, to be within the pseudorapidity range $|\eta| < 2.47$, and to satisfy the medium shower shape and track selection criteria defined in ref. [30]. Candidates found in the transition region between the end-cap and barrel calorimeters ($1.37 < |\eta| < 1.52$) are not considered. Typical reconstruction and identification efficiencies for electrons satisfying these selection criteria range between 80% and 90% depending

---

\(^1\)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 $\eta = -\ln \tan(\theta/2)$. 

---
on $E_T$ and $\eta$. Since electrons do not appear in the $H \to \mu + \tau_{\text{had}}$ decay mode, events with identified electrons are rejected.

Jets are reconstructed using the anti-$k_t$ jet clustering algorithm [31] with a radius parameter $R = 0.4$, taking clusters of calorimeter cells with deposited energy as inputs. Fully calibrated jets [32] are required to be reconstructed in the range $|\eta| < 4.5$ and to have $p_T > 30$ GeV. To reduce the contamination of jets by additional interactions in the same or neighbouring bunch crossings (pile-up), tracks originating from the primary vertex must contribute a large fraction of the jet $p_T$ when summing the scalar $p_T$ of all tracks in the jet. 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. This jet vertex fraction is required to be at least 50% for jets with $|\eta| < 2.4$ and $p_T < 50$ GeV (no cut is applied to jets with $p_T > 50$ GeV). 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 [33], which has an efficiency of $\sim 70\%$ for $b$-jets in $t\bar{t}$ events. 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.

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 charge $q_{\tau} = \pm 1$ in units of electron charge, and must be 1- or 3-track (1- or 3-prong) candidates. Events with exactly one $\tau_{\text{had}}$ candidate satisfying the medium identification criteria [34] with $p_T > 20$ GeV and $|\eta| < 2.47$ are considered in this analysis. 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 of 95%. To reduce the contamination due to backgrounds where a muon fakes a $\tau_{\text{had}}$ signature, events where an identified muon with $p_T(\mu) > 4$ GeV overlaps with an identified $\tau_{\text{had}}$ are rejected [28]. The probability to misidentify a jet with $p_T > 20$ GeV as a $\tau_{\text{had}}$ candidate is typically 1–2%.

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 ($e$, $\gamma$, $\tau_{\text{had}}$, jets, and $\mu$) with which they are associated [35]. 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 ratio of the scalar sum of $p_T$ of tracks which originate from the primary vertex but are not matched to any objects and the scalar sum of $p_T$ of all tracks in the event which are not matched to objects. The scaling procedure achieves a more accurate reconstruction of $E_T^{\text{miss}}$ in high pile-up conditions. In this search, $E_T^{\text{miss}}$ is a signature of neutrinos, and it is used to select and reconstruct signal events, as described below.

3 Event selection and categorization

Signal $H \to \mu \tau$ events in the $\mu \tau_{\text{had}}$ final state are characterised by the presence of an energetic muon and a $\tau_{\text{had}}$ of opposite charge as well as moderate $E_T^{\text{miss}}$, which tends to be
aligned with the $\tau_{\text{had}}$ direction. Backgrounds for this signature can be broadly classified into two major categories:

- Events with true muon and $\tau_{\text{had}}$ signatures, dominated by irreducible $Z/\gamma^* \to \tau\tau$ production with some contributions from the $VV \to \mu\tau + X$ (where $V = W, Z$), $t\bar{t}$, single-top and SM $H \to \tau\tau$ production processes; these events exhibit a very strong charge correlation between the muon and the $\tau_{\text{had}}$; therefore, the expected number of events with opposite-sign (OS) charges ($N_{\text{OS}}$) is much larger than the number of events with same-sign (SS) charges ($N_{\text{SS}}$).

- Events with a fake $\tau_{\text{had}}$ signature, dominated by $W+\text{jets}$ events with some contribution from multi-jet (many multi-jet background events have genuine muons from semileptonic decays of heavy-flavour hadrons), diboson ($VV$), $t\bar{t}$ and single-top events with some charge asymmetry $N_{\text{OS}} > N_{\text{SS}}$; $Z \to \mu\mu+\text{jets}$ events, where a $\tau_{\text{had}}$ signature can be faked by either a jet (no charge correlation) or a muon (strong charge correlation), also contribute to this category.

Events with a fake $\tau_{\text{had}}$ 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 such backgrounds and define signal and control regions. Events with exactly one muon and exactly one $\tau_{\text{had}}$ with $p_T(\mu) > 26$ GeV, $p_T(\tau_{\text{had}}) > 45$ GeV and $|\eta(\mu) - \eta(\tau_{\text{had}})| < 2$ form a baseline sample. The $|\eta(\mu) - \eta(\tau_{\text{had}})|$ cut has ~99% efficiency for signal and rejects a considerable fraction of multi-jet and $W+\text{jets}$ events. At this stage of the event selection, the identified muon is also required to be isolated in the calorimeters and in the tracking detector in order to reduce contamination from the multi-jet background. Two signal regions are defined using the transverse mass, $m_T^2$, of the $\mu\tau$ and $\tau_{\text{had}}\tau_{\text{had}}$ systems: OS events with $m_T(\mu, E_T^{\text{miss}}) > 40$ GeV and $m_T(\tau_{\text{had}}, E_T^{\text{miss}}) < 30$ GeV form the signal region-1 (SR1), while OS events with $m_T(\mu, E_T^{\text{miss}}) < 40$ GeV and $m_T(\tau_{\text{had}}, E_T^{\text{miss}}) < 60$ GeV form the signal region-2 (SR2). Both signal regions have similar sensitivity to signal (see section 6). The dominant backgrounds in SR1 and SR2 are $W+\text{jets}$ and $Z/\gamma^* \to \tau\tau$ events, respectively. The modelling of the $W+\text{jets}$ background is checked in a dedicated control region (WCR) formed by events with $m_T(\mu, E_T^{\text{miss}}) > 60$ GeV and $m_T(\tau_{\text{had}}, E_T^{\text{miss}}) > 40$ GeV. As it is discussed in detail in section 4, the modelling of the $Z/\gamma^* \to \tau\tau$ background is checked in SR2. The choice of the $m_T$ cuts to define SR1, SR2 and WCR is motivated by correlations between $m_T(\mu, E_T^{\text{miss}})$ and $m_T(\tau_{\text{had}}, E_T^{\text{miss}})$ in $H \to \mu\tau$ signal and major background ($W+\text{jets}$ and $Z/\gamma^* \to \tau\tau$) events, as illustrated in figure 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 1 provides a summary of the event selection cuts used to define the signal and control regions.

The LFV signal is searched for by performing a fit to the mass distribution in data, $m_{\mu\tau}^{\text{MMC}}$, reconstructed from the observed muon, $\tau_{\text{had}}$ and $E_T^{\text{miss}}$ objects by means of the

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

where $\ell = \mu, \tau_{\text{had}}$ and $\Delta \phi$ is the azimuthal separation between the directions of the lepton ($\mu$ or $\tau_{\text{had}}$) and $E_T^{\text{miss}}$ vectors.
Figure 1. Two-dimensional distributions of the transverse mass of the $\mu$-$E_{T}^{miss}$ system, $m_{T}(\mu, E_{T}^{miss})$, and that of the $\tau_{had}$-$E_{T}^{miss}$ system, $m_{T}(\tau_{had}, E_{T}^{miss})$, in simulated $Z/\gamma^{*} \rightarrow \tau\tau$ (top left plot), $W+\text{jets}$ (top right plot), $H \rightarrow \mu\tau$ signal (bottom left plot) and data (bottom right plot) events. Magenta, red and yellow boxes on the bottom right plot illustrate locations of SR1, SR2, and WCR, respectively. All events are required to have a well-identified muon and $\tau_{had}$ of opposite charge with $p_{T}(\tau_{had}) > 20$ GeV and $p_{T}(\mu) > 26$ GeV.

<table>
<thead>
<tr>
<th>Cut</th>
<th>SR1</th>
<th>SR2</th>
<th>WCR</th>
<th>TCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{T}(\tau_{had})$</td>
<td>&gt;26 GeV</td>
<td>&gt;26 GeV</td>
<td>&gt;26 GeV</td>
<td>&gt;26 GeV</td>
</tr>
<tr>
<td>$p_{T}(\mu)$</td>
<td>&gt;45 GeV</td>
<td>&gt;45 GeV</td>
<td>&gt;45 GeV</td>
<td>&gt;45 GeV</td>
</tr>
<tr>
<td>$m_{T}(\mu, E_{T}^{miss})$</td>
<td>&gt;40 GeV</td>
<td>&lt;40 GeV</td>
<td>&gt;60 GeV</td>
<td>–</td>
</tr>
<tr>
<td>$m_{T}(\tau_{had}, E_{T}^{miss})$</td>
<td>&lt;30 GeV</td>
<td>&lt;60 GeV</td>
<td>&gt;40 GeV</td>
<td>–</td>
</tr>
<tr>
<td>$</td>
<td>\eta(\mu) - \eta(\tau_{had})</td>
<td>$</td>
<td>&lt;2</td>
<td>&lt;2</td>
</tr>
<tr>
<td>$N_{\text{Jet}}$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>&gt;1</td>
</tr>
<tr>
<td>$N_{b-\text{Jet}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>&gt;0</td>
</tr>
</tbody>
</table>

Table 1. Summary of the event selection criteria used to define the signal and various control regions (see text).
Missing Mass Calculator [36] (MMC). Conceptually, the MMC is a more sophisticated version of the collinear approximation [37]. The main improvement comes from requiring that the relative orientations of the neutrinos and other \( \tau \)-lepton decay products are consistent with the mass and kinematics of a \( \tau \)-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 [28] is modified to take into account that there is only one neutrino from a hadronic \( \tau \)-lepton decay in LFV \( H \rightarrow \mu \tau \) events. For a Higgs boson with \( m_H = 125 \) GeV, the reconstructed \( m_{\mu \tau}^{\text{MMC}} \) 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 < \( m_{\mu \tau}^{\text{MMC}} \) < 150 GeV regions of SR1 and SR2, which contains \( \sim 94\% \) of the expected signal events. The event selection and the analysis strategy are defined without looking at the data in these blinded regions.

4 Background estimation

The background estimation method takes into account the background properties and composition discussed in section 3. It also relies on the assumption that the shape of the \( m_{\mu \tau}^{\text{MMC}} \) distribution for the multi-jet background is the same for OS and SS events. This assumption was verified in the published \( H \rightarrow \tau \tau \) search [38]. In addition, it was confirmed using a dedicated control region, MJCR, with an enhanced contribution from the multi-jet background. Events in this control region are required to pass all criteria for SR1 and SR2 with the exception of the requirement on \( |\eta(\mu) - \eta(\tau_{\text{had}})| \), which is reversed: \( |\eta(\mu) - \eta(\tau_{\text{had}})| > 2 \). Therefore, the number of the total OS background events, \( N_{\text{OS}}^{\text{bkg}} \) in each bin of the \( m_{\mu \tau}^{\text{MMC}} \) (or any other) distribution in SR1 and SR2 can be obtained according to the following formula:

\[
N_{\text{OS}}^{\text{bkg}} = r_{\text{QCD}} \cdot N_{\text{SS}}^{\text{data}} + N_{\text{OS} \rightarrow \tau \tau}^{Z} + N_{\text{OS} \rightarrow \mu \mu}^{Z} + N_{\text{OS} \rightarrow \text{W+jets}}^{W} + N_{\text{OS} \rightarrow \text{SS}}^{\text{top}} + N_{\text{OS} \rightarrow \text{SS}}^{VV} + N_{H \rightarrow \tau \tau}^{H},
\]  

(4.1)

where the individual terms are described below. \( N_{\text{SS}}^{\text{data}} \) is the number of SS data events, which are dominated by \( W+\text{jets} \) events but also contain contributions from multi-jet and other backgrounds. The fractions of multi-jet background in SS data events inside the \( 110 \) GeV < \( m_{\mu \tau}^{\text{MMC}} \) < 150 GeV mass window are \( \sim 17\% \) and \( \sim 44\% \) in SR1 and SR2, respectively. The contributions \( N_{\text{OS} \rightarrow \text{SS}}^{\text{bkg}-i} = N_{\text{OS}}^{\text{bkg}-i} - r_{\text{QCD}} \cdot N_{\text{SS}}^{\text{bkg}-i} \) are \textit{add-on} terms for the different backgrounds components (where \( \text{bkg}-i \) indicates the \( i\)th background source: \( Z \rightarrow \tau \tau, Z \rightarrow \mu \mu, W+\text{jets}, VV, H \rightarrow \tau \tau \) and events with \( t\)-quarks), which also account for components of these backgrounds already included in SS data events.\(^3\) The factor \( r_{\text{QCD}} = N_{\text{OS}}^{\text{multi-jet}}/N_{\text{SS}}^{\text{multi-jet}} \) accounts for potential differences in flavour composition (and, as a consequence, in jet \( \rightarrow \tau_{\text{had}} \) fake rates) of final-state jets introduced by the same-sign or opposite-sign charge requirements. The value of \( r_{\text{QCD}} = 1.10 \pm 0.14 \) is obtained from a

---

\(^3\) The \( r_{\text{QCD}} \cdot N_{\text{SS}}^{\text{bkg}-i} \) correction in the \textit{add-on} term is needed because same-sign data events include multi-jet as well as electroweak events (\( Z \rightarrow \tau \tau, Z \rightarrow \mu \mu, W+\text{jets}, VV, H \rightarrow \tau \tau \) and events with \( t\)-quarks) and their contributions cannot be separated.
multi-jet-enriched control region in data, as discussed in detail in ref. [38]. It was verified in the MJCR in this search.

The largely irreducible $Z/\gamma^* \rightarrow \tau\tau$ background is modelled by $Z/\gamma^* \rightarrow \mu\mu$ 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 minimized 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. [39]. The $Z/\gamma^* \rightarrow \tau\tau$ normalization is a free-floating parameter in the final fit to data and it is mainly constrained by events with 60 GeV < $m_{\text{MMC}}$ < 110 GeV in SR2.

The $W+\text{jets}$ and $Z \rightarrow \mu\mu$ backgrounds are modelled by the ALPGEN [40] event generator interfaced with PYTHIA8 [41] to provide the parton showering, hadronization and the modelling of the underlying event. In all $W+\text{jets}$ events, the $\tau_{\text{had}}$ signature is faked by jets. The WCR region is used to check the modelling of the $W+\text{jets}$ kinematics and to obtain normalizations for OS and SS $W+\text{jets}$ events. An additional overall normalization factor for the $N_{OS-SS}^{W+\text{jets}}$ term in eq. (4.1) is introduced as a free-floating parameter in the final fit in SR1. By studying WCR events and SR1 events with $m_{\mu\tau}^{\text{MMC}} > 150$ GeV (dominated by $W+\text{jets}$ background), it is also found that a $m_{\mu\tau}^{\text{MMC}}$ shape correction, which depends on the number of jets, $p_T(\tau_{\text{had}})$ and $|\eta(\mu) - \eta(\tau_{\text{had}})|$, needs to be applied in SR1. This correction is derived from SR1 events with $m_{\mu\tau}^{\text{MMC}} > 150$ GeV and it is applied to events with all values of $m_{\mu\tau}^{\text{MMC}}$. A 50% difference between the SR1-based correction and that obtained in WCR is taken as a corresponding modelling uncertainty on the $m_{\mu\tau}^{\text{MMC}}$ shape for the $W+\text{jets}$ background in SR1. The size of this uncertainty depends on $m_{\mu\tau}^{\text{MMC}}$ and it is as large as ±10% for $W+\text{jets}$ events with $m_{\mu\tau}^{\text{MMC}} < 150$ GeV. In the case of SR2, a good modelling of the $N_{j_{\text{jet}}, p_T(\tau_{\text{had}})}$ and $|\eta(\mu) - \eta(\tau_{\text{had}})|$ distributions suggests that such a correction is not needed. However, a modelling uncertainty on the $m_{\mu\tau}^{\text{MMC}}$ shape of the $W+\text{jets}$ background in SR2 is assigned based on the 50% difference between the default $m_{\mu\tau}^{\text{MMC}}$ shape and the one obtained after applying the correction derived for SR1 events. The size of this uncertainty is below 5% in the 110 GeV < $m_{\mu\tau}^{\text{MMC}}$ < 150 GeV region, that contains most of the signal events. It was also checked that applying the same correction in SR2 as in SR1 had a negligible effect on the final result and the extracted branching ratio $\text{Br}(H \rightarrow \mu\tau)$ (see section 6) would only be affected at a level below 3%. The modelling of jet fragmentation and the underlying event has a significant effect on the estimate of the jet → $\tau_{\text{had}}$ fake 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 $W+\text{jets}$ events modelled by ALPGEN interfaced with the HERWIG [42] event generator. Differences in the $W+\text{jets}$ predictions in SR1 and SR2 are found to be ±9% and ±2%, respectively, and are taken as corresponding systematic uncertainties.

In the case of the $Z \rightarrow \mu\mu$ background, there are two components: events where a muon fakes a $\tau_{\text{had}}$ and events where a jet fakes a $\tau_{\text{had}}$. Predictions for the shape and normalization of the $Z \rightarrow \mu\mu$ (where $\mu \rightarrow \tau_{\text{fake}}$) background are obtained from simulation. For the $Z \rightarrow \mu\mu$ background,
μμ background where a jet is misidentified as a τ_{had} candidate, the normalization factor and shape corrections, which depend on the number of jets, \( p_T(τ_{had}) \) and \(|η(μ) - η(τ_{had})|\), are derived by using events with two identified OS muons with an invariant mass, \( m_{μμ} \), 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. A 50% difference between the \( m_{μμ}^{MMC} \) shape with and without this correction is taken as a corresponding systematic uncertainty.

The backgrounds with top quarks are modelled by the POWHEG \([43–45]\) (for \( tt \), \( Wt \) and \( s \)-channel single-top production) and AcerMC \([46]\) (\( t \)-channel single-top production) event generators interfaced with PYTHIA8 to provide the parton showering, hadronization and the modelling of the underlying event. The TCR is used to check the modelling and to obtain normalizations for OS and SS events with top quarks. The normalization factors obtained in the TCR are extrapolated into SR1 and SR2, where \( tt \) 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 \([47]\) event generator instead of POWHEG for \( tt \) production.\(^4\) This uncertainty is found to be ±7.2% (±3.7%) for backgrounds with top quarks in SR1 (SR2).

The background due to diboson (WW, ZZ and WZ) production is estimated from simulation, normalized to the cross sections calculated at next-to-leading order (NLO) in QCD \([48]\). The ALPGEN event generator interfaced with HERWIG is used to model the WW process, and HERWIG is used for the ZZ and WZ processes.

Finally, events with Higgs bosons produced via gluon fusion or vector-boson fusion (VBF) processes are generated at NLO accuracy with POWHEG \([49]\) event generator interfaced with PYTHIA8 to provide the parton showering, hadronization 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 \( m_H = 125 \) GeV and normalized to cross sections calculated at next-to-next-to-leading order in QCD \([50–52]\). The SM \( H \rightarrow ττ \) decays are simulated by PYTHIA8. The LFV Higgs boson decays are modelled by the EvtGen \([53]\) event generator according to the phase-space model. In the \( H \rightarrow μτ \) decays, the τ-lepton decays are treated as unpolarised because the left- and right-handed τ-lepton polarisation states are produced at equal rates.

All simulated samples are passed through the GEANT4-based ATLAS detector simulation \([54, 55]\). The simulated events are overlaid with additional minimum-bias events to account for the effect of multiple \( pp \) interactions (pile-up) occurring in the same and neighbouring bunch crossings.

Figure 2 shows the \( m_{μτ}^{MMC} \) distributions for data and the predicted backgrounds in each of the signal regions. The backgrounds are estimated using the method described above. The signal efficiencies for passing the SR1 or SR2 selection requirements are 2.1% and 1.5%, respectively, and the combined efficiency is 3.6%. The numbers of observed events in the data as well as the signal and background predictions in the mass region 110 GeV < \( m_{μτ}^{MMC} \) < 150 GeV can be found in table 2.

\(^4\)The same extrapolation uncertainty is assumed for \( tt \) and single-top backgrounds.
Figure 2. Distributions of the mass reconstructed by the Missing Mass Calculator, $m_{\mu\tau}^{\text{MMC}}$, in SR1 (left) and SR2 (right). The background distributions are determined in a global fit. The signal distribution corresponds to $\text{Br}(H \rightarrow \mu\tau)=25\%$. The bottom panel of each sub-figure shows the ratio of the observed data and the estimated background. The grey band for the ratio illustrates post-fit systematic uncertainties on the background prediction. The statistical uncertainties for data and background predictions are added in quadrature for the ratios. The last bin in each distribution contains overflow events.

Table 2. Data yields, signal and post-fit OS-SS background predictions (see eq. (4.1)) for the 110 GeV < $m_{\mu\tau}^{\text{MMC}}$ < 150 GeV region. The signal predictions are given for $\text{Br}(H \rightarrow \mu\tau)=0.77\%$. The background predictions are obtained from the combined fit to SR1, SR2, WCR and TCR. The post-fit values of systematic uncertainties are provided for the background predictions. For the total background, all correlations between various sources of systematic uncertainties and backgrounds are taken into account. The quoted uncertainties represent the statistical (first) and systematic (second) uncertainties, respectively.
5 Systematic uncertainties

The largest systematic uncertainties arise from the normalization (±10% uncertainty) and modelling\textsuperscript{5} of the W+jets background. The uncertainties on $r_{\text{QCD}}$ (±12.7%) and on the normalization (±6% uncertainty) and modelling of $Z \rightarrow \tau\tau$ also play an important role. The other major sources of experimental uncertainty, affecting both the shape and normalization of signal and backgrounds, are the uncertainty on the $\tau_{\text{had}}$ energy scale [34] (measured with ±(2–4)% precision) and uncertainties on the embedding method used to model the $Z \rightarrow \tau\tau$ background [28]. Less significant sources of experimental uncertainty, affecting the shape and normalization of signal and backgrounds, are the uncertainty on the jet energy scale [32, 56] and resolution [57]. The uncertainties in the $\tau_{\text{had}}$ energy resolution, the momentum scale and resolution of muons, and the scale uncertainty on $E_{\text{T}}^{\text{miss}}$ due to the energy in calorimeter cells not associated with physics objects are taken into account, however, they are found to be relatively small. The following experimental uncertainties primarily affect the normalization of signal and backgrounds: the ±2.8% uncertainty on the integrated luminosity [58], the uncertainty on the $\tau_{\text{had}}$ identification efficiency [34], which is measured to be ±(2–3)% for 1-prong and ±(3–5)% for 3-prong decays, the ±2.1% uncertainty for triggering, reconstructing and identifying muons [29, 59], and the ±2% uncertainty on the $b$-jet tagging efficiency [33].

Theoretical uncertainties are estimated for the signal and for the $H \rightarrow \tau\tau$, $VV$ and $Z \rightarrow \mu\mu$ (with $\mu \rightarrow \tau_{\text{had}}^{\text{fake}}$) backgrounds, which are modelled with the simulation and are not normalized to data in dedicated control regions. Uncertainties due to missing higher-order QCD corrections on the production cross sections are found to be [60] ±10.1% (±7.8%) for the Higgs boson production via gluon fusion in SR1 (SR2), ±1% for the $Z \rightarrow \mu\mu$ background and for VBF and $VH$ Higgs boson production, and ±5% 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. [60] and the following values are used in this analysis: ±7.5% for the Higgs boson production via gluon fusion, ±2.8% for the VBF and $VH$ Higgs boson production, and ±4% for the $Z \rightarrow \mu\mu$ and $VV$ backgrounds. Finally, an additional ±5.7% systematic uncertainty on Br($H \rightarrow \tau\tau$) is applied to the SM $H \rightarrow \tau\tau$ background.

6 Results

A simultaneous binned maximum-likelihood fit is performed on the $m_{\mu\tau}^{\text{MMC}}$ distributions in SR1 and SR2 and on event yields in WCR and TCR to extract the LFV branching ratio Br($H \rightarrow \tau\tau$). The fit exploits the control regions and the distinct shapes of the $W+\text{jets}$ and $Z \rightarrow \tau\tau$ backgrounds in the signal regions to constrain some of the systematic uncertainties. This leads to an improved sensitivity of the analysis. The post-fit $m_{\mu\tau}^{\text{MMC}}$ distributions in SR1 and SR2 are shown in figure 2, and the combined $m_{\mu\tau}^{\text{MMC}}$ distribution for both signal regions is presented in figure 3. Figure 2 illustrates good agreement between

\textsuperscript{5}Some of these uncertainties (e.g., uncertainties due to $m_{\mu\tau}^{\text{MMC}}$ shape corrections and extrapolation uncertainties) are discussed in the text above.
Figure 3. Post-fit combined $m_{\mu\tau}^{\text{MMC}}$ distribution obtained by adding individual distributions in SR1 and SR2. In the lower part of the figure, the data are shown after subtraction of the estimated backgrounds. The grey band in the bottom panel illustrates the post-fit systematic uncertainties on the background prediction. The statistical uncertainties for data and background predictions are added in quadrature on the bottom part of the figure. The signal is shown assuming $\text{Br}(H \to \mu\tau)=0.77\%$, the central value of the best fit to $\text{Br}(H \to \mu\tau)$. The last bin of the distribution contains overflow events.

<table>
<thead>
<tr>
<th></th>
<th>SR1</th>
<th>SR2</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected limit on Br($H \to \mu\tau$) [%]</td>
<td>1.60$^{+0.64}_{-0.45}$</td>
<td>1.75$^{+0.71}_{-0.49}$</td>
<td>1.24$^{+0.50}_{-0.35}$</td>
</tr>
<tr>
<td>Observed limit on Br($H \to \mu\tau$) [%]</td>
<td>1.55</td>
<td>3.51</td>
<td>1.85</td>
</tr>
<tr>
<td>Best fit Br($H \to \mu\tau$) [%]</td>
<td>$-0.07^{+0.81}_{-0.86}$</td>
<td>$1.94^{+0.92}_{-0.89}$</td>
<td>0.77$^{+0.62}_{-0.62}$</td>
</tr>
</tbody>
</table>

Table 3. The expected and observed 95% confidence level (CL) upper limits and the best fit values for the branching fractions for the two signal regions and their combination.

data and background expectations in SR1. A small excess of the data over the predicted background is observed in the 120 GeV< $m_{\mu\tau}^{\text{MMC}}$ <140 GeV region in SR2. This small excess in SR2 has a local significance of 2.2 standard deviations and a combined significance for both signal regions of 1.3 standard deviations. This corresponds to a best fit value for the branching fraction of $\text{Br}(H \to \mu\tau)=(0.77 \pm 0.62)\%$. Due to the low significance of the observed excess, an upper limit on the LFV branching ratio $\text{Br}(H \to \mu\tau)$ for a Higgs boson with $m_H = 125$ GeV is set using the CL$_s$ modified frequentist formalism [61] with the profile likelihood-ratio test statistics [62]. The observed and the median expected 95% CL upper limits are 1.85% and 1.24$^{+0.50}_{-0.35}$%, respectively. Table 3 provides a summary of all results.
7 Summary

A direct search for lepton-flavour-violating $H \rightarrow \mu \tau$ decays of the recently discovered Higgs boson is performed in the $\tau_{\text{had}}$ decay mode of $\tau$-leptons 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 a centre-of-mass energy of $\sqrt{s} = 8$ TeV. The observed and the median expected upper limits at 95\% CL on the branching fraction, $\text{Br}(H \rightarrow \mu \tau)$, are 1.85\% and $1.24^{+0.50}_{-0.35}$, respectively. This search places significantly more stringent constraints on $\text{Br}(H \rightarrow \mu \tau)$ compared to earlier indirect estimates. The result of this analysis is found to be consistent with the one published by the CMS Collaboration [26].

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; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF, DNSRC and Lundbeck Foundation, Denmark; EPLANET, ERC and NSRF, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, DFG, HGF, MPG and AvH Foundation, Germany; GSRT and NSRF, Greece; RGC, Hong Kong SAR, China; ISF, MINERVA, GIF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNiSW and NCN, Poland; GRICES and FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.) and in the Tier-2 facilities worldwide.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


[21] BABAR collaboration, B. Aubert et al., Searches for Lepton Flavor Violation in the Decays $\tau^+ \to e^+\gamma$ and $\tau^+ \to \mu^+\gamma$, *Phys. Rev. Lett.* **104** (2010) 021802 [arXiv:0908.2381] [inSPIRE].


ATLAS collaboration, *Modelling $Z \rightarrow \tau\tau$ processes in ATLAS with $\tau$-embedded $Z \rightarrow \mu\mu$ data*, 2015 *JINST* **10** P09018 [arXiv:1506.05623] [SPIRE].


A. Wildauer\textsuperscript{101}, H.G. Wilkens\textsuperscript{30}, H.H. Williams\textsuperscript{122}, S. Williams\textsuperscript{107}, C. Willis\textsuperscript{90}, S. Willocq\textsuperscript{86}, A. Wilson\textsuperscript{89}, J.A. Wilson\textsuperscript{18}, I. Wingerter-Seez\textsuperscript{5}, F. Winklmeier\textsuperscript{116}, B.T. Winter\textsuperscript{211}, M. Wittgen\textsuperscript{143}, J. Wittkowski\textsuperscript{60}, S.J. Wollstadt\textsuperscript{83}, M.W. Wolter\textsuperscript{39}, H. Wolter\textsuperscript{126a,126b}, B.K. Wosiek\textsuperscript{39}, J. Wotschack\textsuperscript{30}, M.J. Woudstra\textsuperscript{84}, K.W. Wozniak\textsuperscript{39}, M. Wu\textsuperscript{55}, M. Wu\textsuperscript{31}, S.L. Wu\textsuperscript{173}, X. Wu\textsuperscript{49}, Y. Wu\textsuperscript{89}, T.R. Wyatt\textsuperscript{84}, B.M. Wynne\textsuperscript{46}, S. Xella\textsuperscript{36}, L. Xu\textsuperscript{33a,33b}, L. Xu\textsuperscript{25}, B. Yabsley\textsuperscript{150}, S. Yacoobi\textsuperscript{145a}, R. Yakabe\textsuperscript{67}, M. Yamada\textsuperscript{66}, D. Yamaguchi\textsuperscript{157}, Y. Yamaguchi\textsuperscript{118}, A. Yamamoto\textsuperscript{66}, S. Yamamoto\textsuperscript{155}, T. Yamanaka\textsuperscript{152}, K. Yamauchi\textsuperscript{103}, Y. Yamazaki\textsuperscript{67}, Z. Yan\textsuperscript{22}, H. Yang\textsuperscript{33e}, H. Yang\textsuperscript{173}, Y. Yang\textsuperscript{151}, W-M. Yao\textsuperscript{15}, Y. Yasu\textsuperscript{66}, E. Yatsenko\textsuperscript{9}, K.H. Yan\textsuperscript{Wong}\textsuperscript{41}, J. Ye\textsuperscript{40}, S. Ye\textsuperscript{25}, I. Yeletskikh\textsuperscript{65}, A.L. Yen\textsuperscript{57}, E. Yildirim\textsuperscript{42}, K. Yorita\textsuperscript{171}, R. Yoshida\textsuperscript{6}, K. Yoshihara\textsuperscript{122}, C. Young\textsuperscript{143}, C.J.S. Young\textsuperscript{30}, S. Youssef\textsuperscript{22}, D.R. Yu\textsuperscript{15}, J. Yu\textsuperscript{8}, J.M. Yu\textsuperscript{89}, J. Yu\textsuperscript{114}, L. Yuan\textsuperscript{67}, S.P.Y. Yuen\textsuperscript{21}, A. Yurkewicz\textsuperscript{108}, I. Yusuf\textsuperscript{28,ak}, B. Zabinski\textsuperscript{39}, R. Zaidan\textsuperscript{63}, A.M. Zaitsev\textsuperscript{130,ab}, J. Zalieckas\textsuperscript{14}, A. Zaman\textsuperscript{148}, S. Zambito\textsuperscript{57}, L. Zanello\textsuperscript{132a,132b}, D. Zani\textsuperscript{88}, C. Zeitnitz\textsuperscript{175}, M. Zeman\textsuperscript{128}, A. Zemla\textsuperscript{39a}, Q. Zeng\textsuperscript{143}, K. Zengel\textsuperscript{23}, O. Zenin\textsuperscript{130}, T. Ženiš\textsuperscript{144a}, D. Zerwas\textsuperscript{117}, D. Zhang\textsuperscript{89}, F. Zhang\textsuperscript{173}, H. Zhang\textsuperscript{33e}, J. Zhang\textsuperscript{6}, L. Zhang\textsuperscript{48}, R. Zhang\textsuperscript{33b}, X. Zhang\textsuperscript{33d}, Z. Zhang\textsuperscript{117}, X. Zhao\textsuperscript{33d,117}, Z. Zhao\textsuperscript{33b}, A. Zhenechugov\textsuperscript{65}, J. Zhong\textsuperscript{120}, B. Zhou\textsuperscript{89}, C. Zhou\textsuperscript{45}, L. Zhou\textsuperscript{40}, M. Zhou\textsuperscript{148}, N. Zhou\textsuperscript{33l}, C.G. Zhu\textsuperscript{33d}, H. Zhu\textsuperscript{33a}, J. Zhu\textsuperscript{89}, Y. Zhu\textsuperscript{33b}, X. Zhuang\textsuperscript{33a}, K. Zhukov\textsuperscript{96}, A. Zibell\textsuperscript{174}, D. Zieminska\textsuperscript{61}, N.I. Zimine\textsuperscript{65}, C. Zimmermann\textsuperscript{83}, S. Zimmermann\textsuperscript{48}, Z. Zinonos\textsuperscript{54}, M. Zinner\textsuperscript{83}, M. Ziolkowski\textsuperscript{41}, L. Živković\textsuperscript{13}, G. Zobernig\textsuperscript{173}, A. Zoccoli\textsuperscript{20a,20b}, M. zur Nedden\textsuperscript{16}, G. Zurzolo\textsuperscript{104a,104b}, L. Zwalinski\textsuperscript{30}.

\textsuperscript{1} Department of Physics, University of Adelaide, Adelaide, Australia
\textsuperscript{2} Physics Department, SUNY Albany, Albany NY, United States of America
\textsuperscript{3} Department of Physics, University of Alberta, Edmonton AB, Canada
\textsuperscript{4} Department of Physics, Ankara University, Ankara\textsuperscript{(a)}; Istanbul Aydin University, Istanbul\textsuperscript{(b)}; Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey\textsuperscript{(c)}
\textsuperscript{5} LAPP, CNRS/IN2P3 and Université Savoie Mont Blanc, Annecy-le-Vieux, France
\textsuperscript{6} High Energy Physics Division, Argonne National Laboratory, Argonne IL, United States of America
\textsuperscript{7} Department of Physics, University of Arizona, Tucson AZ, United States of America
\textsuperscript{8} Department of Physics, Brandeis University, Waltham MA, United States of America
\textsuperscript{9} Department of Physics, Boston University, Boston MA, United States of America
\textsuperscript{10} Department of Physics, Boston University, Boston MA, United States of America
\textsuperscript{11} Department of Physics, Boston University, Boston MA, United States of America
\textsuperscript{12} Department of Physics, Brandeis University, Waltham MA, United States of America
\textsuperscript{13} Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
\textsuperscript{14} School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
\textsuperscript{15} Department of Mathematics and Physics, Bogazici University, Istanbul\textsuperscript{(a)}; Department of Physics Engineering, Gaziantep University, Gaziantep\textsuperscript{(b)}; Department of Physics, Dopus University, Istanbul, Turkey\textsuperscript{(c)}
\textsuperscript{16} INFN Sezione di Bologna\textsuperscript{(a)}; Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy

JHEP11(2015)211

- 27 -