Search for the lepton flavour violating decay $\tau^(-) \rightarrow \mu^(-)\mu^+(+\mu^(-)$

Alexander, M.; Ali, S.

Published in:
Journal of High Energy Physics

DOI:
10.1007/JHEP02(2015)121

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Search for the lepton flavour violating decay
\( \tau^- \rightarrow \mu^- \mu^+ \mu^- \)

The LHCb collaboration

E-mail: jonathan.harrison@manchester.ac.uk

ABSTRACT: A search for the lepton flavour violating decay \( \tau^- \rightarrow \mu^- \mu^+ \mu^- \) is performed with the LHCb experiment. The data sample corresponds to an integrated luminosity of 1.0 fb\(^{-1}\) of proton-proton collisions at a centre-of-mass energy of 7 TeV and 2.0 fb\(^{-1}\) at 8 TeV. No evidence is found for a signal, and a limit is set at 90% confidence level on the branching fraction, \( \mathcal{B}(\tau^- \rightarrow \mu^- \mu^+ \mu^-) < 4.6 \times 10^{-8} \).

KEYWORDS: Rare decay, Tau Physics, Hadron-Hadron Scattering

ArXiv ePrint: 1409.8548
1 Introduction

Lepton flavour violating processes are allowed within the context of the Standard Model (SM) with massive neutrinos, but their branching fractions are of order $10^{-40}$ [1, 2] or smaller, and are beyond the reach of any currently conceivable experiment. Observation of charged lepton flavour violation (LFV) would therefore be an unambiguous signature of physics beyond the Standard Model (BSM), but no such process has been observed to date [3].

A number of BSM scenarios predict LFV at branching fractions approaching current experimental sensitivities [4], with LFV in $\tau^-$ decays often enhanced with respect to $\mu^-$ decays due to the large difference in mass between the two leptons (the inclusion of charge-conjugate processes is implied throughout). If charged LFV were to be discovered, measurements of the branching fractions for a number of channels would be required to determine the nature of the BSM physics. In the absence of such a discovery, improving the experimental constraints on the branching fractions for LFV decays would help to constrain the parameter spaces of BSM models.

This paper reports on an updated search for the LFV decay $\tau^- \to \mu^- \mu^+ \mu^-$ with the LHCb experiment [5] at the CERN LHC. The previous LHCb analysis of this channel produced the first result on a search for LFV $\tau^-$ decays at a hadron collider [6]. Using 1.0 fb$^{-1}$ of proton-proton collision data collected at a centre-of-mass energy of 7 TeV, a limit was set on the branching fraction, $B(\tau^- \to \mu^- \mu^+ \mu^-) < 8.0 \times 10^{-8}$ at 90% confidence
level (CL). The current best experimental upper limits are $B(\tau^- \rightarrow \mu^- \mu^+ \mu^-) < 2.1 \times 10^{-8}$ at 90% CL from Belle [7] and $B(\tau^- \rightarrow \mu^- \mu^+ \mu^-) < 3.3 \times 10^{-8}$ at 90% CL from BaBar [8]. In the analysis presented here, an additional LHCb data set, corresponding to 2.0 fb$^{-1}$ of integrated luminosity collected at 8 TeV, is added to the previous data set, and a number of new analysis techniques are introduced.

The search for LFV in $\tau^-$ decays at LHCb takes advantage of the large inclusive $\tau^-$ production cross-section at the LHC, where $\tau^-$ leptons are produced almost entirely from the decays of $b$ and $c$ hadrons. Using the $b\bar{b}$ and $c\bar{c}$ cross-sections measured by LHCb [9, 10] and the inclusive $b \rightarrow \tau$ and $c \rightarrow \tau$ branching fractions [3], the inclusive $\tau^-$ cross-section is estimated to be 85 $\mu$b at 7 TeV.

Selection criteria are implemented for the signal mode, $\tau^- \rightarrow \mu^- \mu^+ \mu^-$, and for the calibration and normalisation channel, which is $D_s^- \rightarrow \phi \pi^-$ with $\phi \rightarrow \mu^+ \mu^-$. To avoid potential bias, $\mu^- \mu^+ \mu^-$ candidates with mass within $\pm 30$ MeV/$c^2$ (approximately three times the expected mass resolution) of the known $\tau^-$ mass are initially excluded from the analysis. Discrimination between a potential signal and the background is performed using a three-dimensional binned distribution in two multivariate classifiers and the mass of the $\tau^-$ candidate. One classifier is based on the three-body decay topology and the other on muon identification.

2 Detector and triggers

The LHCb detector [5] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of momentum, $p$, with a relative uncertainty that varies from 0.4% at low momentum to 0.6% at 100 GeV/$c$. The minimum distance of a track to a primary vertex, the impact parameter (IP), is measured with a resolution of $(15 + 29/p_T)$ $\mu$m, where $p_T$ is the component of $p$ transverse to the beam, in GeV$/c$. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors (RICH) [11]. Photon, electron and hadron candidates are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers [12].

The trigger [13] consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction. Candidate events are first required to pass the hardware trigger, which selects muons with a transverse momentum $p_T > 1.48$ GeV/$c$ in the 7 TeV data or $p_T > 1.76$ GeV/$c$ in the 8 TeV data. In the software trigger, at least one of the final-state particles is required to have both $p_T > 0.8$ GeV/$c$ and IP $> 100 \mu$m with respect to all of the primary $pp$ interaction...
vertices (PVs) in the event. Finally, the tracks of two or more of the final-state particles are required to form a vertex that is significantly displaced from the PVs.

3 Monte Carlo simulation

In the simulation, $pp$ collisions are generated using PYTHIA [14] with a specific LHCb configuration [15]. Decays of hadronic particles are described by EvtGen [16], in which final-state radiation is generated using PHOTOS [17]. For the $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ signal channel, the final-state particles are distributed according to three-body phase-space. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [18, 19] as described in ref. [20].

As the $\tau^-$ leptons produced in the LHCb acceptance originate almost exclusively from heavy quark decays, they can be classified in one of five categories according to the parent particle. The parent particle can be the following: a $b$ hadron; a $D^-_s$ or $D^-$ meson that is produced directly in a proton-proton collision or via the decay of an excited charm meson; or a $D^+_s$ or $D^-$ meson resulting from the decay of a $b$ hadron. Events from each category are generated separately and are combined in accordance with the measured cross-sections and branching fractions. Variations of the cross-sections and branching fractions within their uncertainties are considered as sources of systematic uncertainty.

4 Event selection

Candidate $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ decays are selected by requiring three tracks that combine to give a mass close to that of the $\tau^-$ lepton, and that form a vertex that is displaced from the PV. The tracks are required to be well-reconstructed muon candidates with $p_T > 300$ MeV/$c$ that have a significant separation from the PV. There must be a good fit to the three-track vertex, and the decay time of the candidate forming the vertex has to satisfy $ct > 100$ µm. As the $\tau^-$ leptons are produced predominantly in the decays of charm mesons, where the $Q$-values are relatively small (and so the charm meson and the $\tau^-$ are almost collinear in the laboratory frame), a requirement on the pointing angle, $\theta$, between the momentum vector of the three-track system and the vector joining the primary and secondary vertices is used to remove poorly reconstructed candidates ($\cos \theta > 0.99$). Contamination from pairs of tracks originating from the same particle is reduced by removing same-sign muon pairs with mass lower than $250$ MeV/$c^2$.

The decay $D^-_s \rightarrow \eta (\mu^+ \mu^- \gamma) \mu^- \bar{\nu}_\mu$ is a source of irreducible background near the signal region, and therefore candidates with a $\mu^+ \mu^-$ invariant mass below $450$ MeV/$c^2$ are removed. Signal candidates containing muons that result from the decay of the $\phi(1020)$ meson are removed by excluding $\mu^+ \mu^-$ masses within $\pm 20$ MeV/$c^2$ of the known $\phi(1020)$ meson mass.

The signal region is defined by a $\pm 20$ MeV/$c^2$ window (approximately two times the expected mass resolution) around the known $\tau^-$ mass. Candidates with $\mu^- \mu^+ \mu^-$ invariant mass between $1600$ and $1950$ MeV/$c^2$ are kept to allow evaluation of the background contributions in the signal region. In the following, the wide mass windows on either side of the
signal region are referred to as the data sidebands. The signal region for the normalisation channel, $D_{s}^{-} \to \phi (\mu^{+}\mu^{-}) \pi^{-}$, which has a similar topology to that of the $\tau^{-} \to \mu^{-}\mu^{+}\mu^{-}$ decay, is defined by a $\pm 20$ MeV/$c^{2}$ window around the $D_{s}^{-}$ mass, with the $\mu^{+}\mu^{-}$ mass required to be within $\pm 20$ MeV/$c^{2}$ of the $\phi(1020)$ meson mass. Where appropriate, the rest of the selection criteria are identical to those for the signal channel, with one of the muon candidates replaced by a pion candidate.

5 Signal and background discrimination

Three classifiers are used to discriminate between signal and background: an invariant mass classifier that uses the reconstructed mass of the $\tau^{-}$ candidate; a geometric classifier, $M_{3\text{body}}$; and a particle identification classifier, $M_{\text{PID}}$.

The multivariate classifier $M_{3\text{body}}$ is based on the geometry and kinematic properties of the final-state tracks and the reconstructed $\tau^{-}$ candidate. It aims to reject backgrounds from combinations of tracks that do not share a common vertex and those from multi-body decays with more than three final-state particles. The variables used in the classifier include the vertex fit quality, the displacement of the vertex from the PV, the pointing angle $\theta$, and the IP and fit $\chi^{2}$ of the tracks. An ensemble-selected (blended) [21], custom boosted decision tree (BDT) classifier is used [22, 23], as described in the following. In the blending method the input variables are combined [24] into one BDT, two Fisher discriminants [25], four neural networks [26], one function-discriminant analysis [27] and one linear discriminant [28]. Each classifier is trained using simulated signal and background samples, where the composition of the background is a mixture of $b\bar{b} \to \mu\mu X$ and $c\bar{c} \to \mu\mu X$ processes according to their relative abundances as measured in data. As each category of simulated signal events has different kinematic properties, a separate set of classifiers is trained for each. One third of the available signal sample is used at this stage, along with one half of the background sample. The classifier responses, along with the original input variables, are then used as input to the custom BDT classifier, which is trained on the remaining half of the background sample and a third of the signal sample, with the five categories combined, to give the final classifier response. The responses of the classifier on the training and the test samples are found to be in good agreement, suggesting no overtraining of the classifier is present. As the responses of the individual classifiers are not fully correlated, blending the output of the classifiers improves the sensitivity of the analysis in our data sample by 6% with respect to that achievable by using the best single classifier. The $M_{3\text{body}}$ classifier response is calibrated using the $D_{s}^{-} \to \phi (\mu^{+}\mu^{-}) \pi^{-}$ control channel to correct for differences in response between data and simulation. Figure 1 shows good agreement between $D_{s}^{-} \to \phi (\mu^{+}\mu^{-}) \pi^{-}$ data and simulation for one of the input variables to $M_{3\text{body}}$ and for the classifier response. A systematic uncertainty of 2% is assigned to account for any remaining differences. The classifier response is found to be uncorrelated with mass for both the signal sample and the data sidebands.

The multivariate classifier $M_{\text{PID}}$ uses information from the RICH detectors, the calorimeters and the muon detectors to obtain the likelihood that each of the three final-state particles is compatible with the muon hypothesis. The value of the $M_{\text{PID}}$ response
is taken as the smallest likelihood of the three muon candidates. The $M_{\text{PID}}$ classifier uses a neural network that is trained on simulated events to discriminate muons from other charged particles. The $M_{\text{PID}}$ classifier response is calibrated using muons from $J/\psi \rightarrow \mu^+\mu^-$ decays in data.

For the $M_{3\text{body}}$ and $M_{\text{PID}}$ responses, a binning is chosen via the CLs method [30, 31] by maximising the difference between the median CLs values under the background-only hypothesis and the signal-plus-background hypothesis, whilst minimising the number of bins. The binning optimisation is performed separately for the 7 TeV and 8 TeV data sets, because there are small differences in event topology with changes of centre-of-mass energy. The optimisation does not depend on the signal branching fraction. The bins at lowest values of $M_{3\text{body}}$ and $M_{\text{PID}}$ response do not contribute to the sensitivity and are excluded from the analysis. The distributions of the responses of the two classifiers, along with their binning schemes, are shown in figure 2.

The expected shapes of the invariant mass spectra for the $\tau^- \rightarrow \mu^-\mu^+\mu^-$ signal in the 7 TeV and 8 TeV data sets are taken from fits to the $D_s^- \rightarrow \phi (\mu^+\mu^-) \pi^-$ control channel in data. Figure 3 shows the fit to the 8 TeV data. No particle identification requirements are applied to the pion. The signal distribution is modelled with the sum of two Gaussian functions with a common mean, where the narrower Gaussian contributes 70% of the total signal yield, while the combinatorial background is modelled with an exponential function. The expected width of the $\tau^-$ signal in data is taken from simulation, scaled by the ratio of the widths of the $D_s^-$ peaks in data and simulation.

6 Backgrounds

The background processes for the $\tau^- \rightarrow \mu^-\mu^+\mu^-$ decay consist mainly of heavy meson decays yielding three muons in the final state, or one or two muons in combination with two or one misidentified particles. There are also a large number of events with one or two muons from heavy meson decays combined with two or one muons from elsewhere in the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Distribution of (a) $D_s^-$ flight distance and (b) $M_{3\text{body}}$ response for $D_s^- \rightarrow \phi (\mu^+\mu^-) \pi^-$ candidates at 8 TeV. The dashed (red) lines indicate the data and the solid (black) lines indicate the simulation. The data are background-subtracted using the sPlot technique [29].}
\end{figure}
event. Decays containing undetected final-state particles, such as $K^0_L$ mesons, neutrinos or photons, can give large backgrounds, which vary smoothly in the signal region. The most important background channel of this type is found to be $D^-_s \rightarrow \eta (\mu^+ \mu^- \gamma) \mu^- \bar{\nu}_\mu$, about 90% of which is removed by the requirement on the dimuon mass. The small remaining contribution from this process has a mass distribution similar to that of the other backgrounds in the mass range considered in the fit. The dominant contributions to the background from misidentified particles are from $D^-_s \rightarrow K^+ \pi^- \pi^-$ and $D^-_s \rightarrow \pi^+ \pi^- \pi^-$ decays. However, these events populate mainly the region of low $M_{\text{PID}}$ response and are reduced to a negligible level by the exclusion of the first bin.

The expected numbers of background events within the signal region, for each bin in $M_{\text{3body}}$ and $M_{\text{PID}}$, are evaluated by fitting an exponential function to the candidate mass spectra outside of the signal windows using an extended, unbinned maximum likelihood fit. The parameters of the exponential function are allowed to vary independently in each bin. The small differences obtained if the exponential curves are replaced by straight lines are included as systematic uncertainties. The $\mu^+ \mu^- \mu^-$ mass spectra are fitted over the mass range 1600–1950 MeV/c$^2$, excluding windows of width $\pm 30$ MeV/c$^2$ around the expected signal mass. The resulting fits to the data sidebands for the highest sensitivity bins are shown in figure 4 for 7 and 8 TeV data separately.
Figure 3. Invariant mass distribution of $\phi(\mu^+\mu^-)\pi^-$ candidates in 8 TeV data. The solid (blue) line shows the overall fit, the long-dashed (green) and short-dashed (red) lines show the two Gaussian components of the $D_s^+$ signal and the dot-dashed (black) line shows the combinatorial background contribution.

Figure 4. Invariant mass distributions and fits to the mass sidebands in (a) 7 TeV and (b) 8 TeV data for $\mu^+\mu^-\mu^-$ candidates in the bins of $M_{3\text{body}}$ and $M_{\text{PID}}$ response that contain the highest signal probabilities.
7 Normalisation

The observed number of $\tau^- \to \mu^-\mu^+\mu^-$ candidates is converted into a branching fraction by normalising to the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ calibration channel according to

\[
\mathcal{B}(\tau^- \to \mu^-\mu^+\mu^-) = \frac{\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-)}{\mathcal{B}(D_s^- \to \tau^-\bar{\nu}_\tau)} \times f_{\tau} \times \epsilon_{\text{cal}}^T \times \epsilon_{\text{sig}}^T \times \frac{N_{\text{sig}}}{N_{\text{cal}}} \equiv \alpha N_{\text{sig}} ,
\]

where $\alpha$ is the overall normalisation factor, $N_{\text{sig}}$ is the number of observed signal events and all other terms are described below. Table 1 gives a summary of all contributions to the factor $\alpha$; the uncertainties are taken to be uncorrelated. The branching fraction of the normalisation channel is determined from known branching fractions as

\[
\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-) = \frac{\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} \mathcal{B}(\phi \to \mu^+\mu^-) = (1.32 \pm 0.10) \times 10^{-5},
\]

where $\mathcal{B}(\phi \to K^+K^-)$ and $\mathcal{B}(\phi \to \mu^+\mu^-)$ are taken from ref. [3] and $\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-)$ is taken from ref. [32]. The branching fraction $\mathcal{B}(D_s^- \to \tau^-\bar{\nu}_\tau)$ is taken from refs. [3, 33].

The quantity $f_{\tau}^D$ is the fraction of $\tau^-$ leptons that originate from $D_s^-$ decays. The value of $f_{\tau}^D$ at 7 TeV is calculated using the $b\bar{b}$ and $c\bar{c}$ cross-sections as measured by LHCb [9, 10] at 7 TeV and the inclusive $b \to D_s, c \to D_s, b \to \tau$ and $c \to \tau$ branching fractions [3]. For the value of $f_{\tau}^D$ at 8 TeV the $b\bar{b}$ cross-section is updated to the 8 TeV LHCb measurement [34] and the $c\bar{c}$ cross-section measured at 7 TeV is scaled by a factor of $8/7$, consistent with PYTHIA simulations. The uncertainty on this scaling factor, which is negligible, is found by taking the difference between the value obtained from the nominal parton distribution functions and that from the average of their corresponding error sets [35].

The reconstruction and selection efficiencies, $\epsilon^R$, are products of the detector acceptances for the decay of interest, the muon identification efficiencies and the selection efficiencies. The combined muon identification and selection efficiencies are determined from the yield of simulated events after the full selections are applied. The ratio of efficiencies is corrected to account for the differences between data and simulation in track reconstruction, muon identification, the $\phi(1020)$ mass window requirement in the normalisation channel and the $\tau^-$ mass range. The removal of candidates in the least sensitive bins in the $M_{3\text{body}}$ and $M_{\text{PID}}$ classifier responses is also taken into account.

The trigger efficiencies, $\epsilon^T$, are evaluated from simulation and their systematic uncertainties are determined from the differences between the trigger efficiencies of $B^- \to J/\psi(\mu^+\mu^-)K^- \to \tau^-\mu^-\mu^+$ candidates measured in data and in simulation, using muons with momentum values typical of $\tau^- \to \mu^+\mu^-\mu^-$ signal decays. The trigger efficiency for the 8 TeV data set is corrected to account for differences in trigger conditions across the data taking period, resulting in a relatively large systematic error.

The yields of $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ candidates in data, $N_{\text{cal}}$, are determined from the fits to reconstructed $\phi(\mu^+\mu^-)\pi^-$ mass distributions as shown in figure 3. The variations in the yields when the relative contributions of the two Gaussian components are allowed to vary in the fits are considered as systematic uncertainties.
Table 1. Terms entering into the normalisation factors, $\alpha$, and their combined statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th></th>
<th>7 TeV</th>
<th>8 TeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B(D_s^- \to \phi (\mu^+ \mu^-) \pi^-)$</td>
<td>$(1.32 \pm 0.10) \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td>$B(D_s^- \to \tau^- \bar{\nu}_\tau)$</td>
<td>$(5.61 \pm 0.24) \times 10^{-2}$</td>
<td></td>
</tr>
<tr>
<td>$f_{D_s}^{T}$</td>
<td>$0.78 \pm 0.04$</td>
<td>$0.80 \pm 0.03$</td>
</tr>
<tr>
<td>$\epsilon_{\text{cal}}^R / \epsilon_{\text{sig}}^R$</td>
<td>$0.898 \pm 0.060$</td>
<td>$0.912 \pm 0.054$</td>
</tr>
<tr>
<td>$\epsilon_{\text{cal}}^T / \epsilon_{\text{sig}}^T$</td>
<td>$0.659 \pm 0.006$</td>
<td>$0.525 \pm 0.040$</td>
</tr>
<tr>
<td>$N_{\text{cal}}$</td>
<td>$28,200 \pm 440$</td>
<td>$52,130 \pm 700$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$(7.20 \pm 0.98) \times 10^{-9}$</td>
<td>$(3.37 \pm 0.50) \times 10^{-9}$</td>
</tr>
</tbody>
</table>

8 Results

Tables 2 and 3 give the expected and observed numbers of candidates in the signal region, for each bin of the classifier responses. No significant excess of events over the expected background is observed. Using the CL$_s$ method [30, 31] and eq. (7.1), the observed CL$_s$ value and the expected CL$_s$ distribution are calculated as functions of the assumed branching fraction, as shown in figure 5. The systematic uncertainties on the signal and background estimates, which have a very small effect on the final limits, are included following ref. [30, 31]. The expected limit at 90% (95%) CL for the branching fraction is $B(\tau^- \to \mu^- \mu^+ \mu^-) < 5.0 (6.1) \times 10^{-8}$, while the observed limit at 90% (95%) CL is

$$B(\tau^- \to \mu^- \mu^+ \mu^-) < 4.6 (5.6) \times 10^{-8}.$$  

Whilst the above limits are given for the phase-space model of $\tau^-$ decays, the kinematic properties of the decay would depend on the physical processes that introduce LFV. Reference [36] gives a model-independent analysis of the decay distributions in an effective field-theory approach including BSM operators with different chirality structures. Depending on the choice of operator, the observed limit varies within the range $(4.1 - 6.8) \times 10^{-8}$ at 90% CL. The weakest limit results from an operator that favours low $\mu^+\mu^-$ mass, since the requirement to remove the $D^+_s \to \eta (\mu^+ \mu^- \gamma) \mu^- \bar{\nu}_\mu$ background excludes a large fraction of the relevant phase-space.

In summary, the LHCb search for the LFV decay $\tau^- \to \mu^- \mu^+ \mu^-$ is updated using all data collected during the first run of the LHC, corresponding to an integrated luminosity of $3.0 \text{fb}^{-1}$. No evidence for any signal is found. The measured limits supersede those of ref. [6] and, in combination with results from the $B$ factories, improve the constraints placed on the parameters of a broad class of BSM models [37].
<table>
<thead>
<tr>
<th>$M_{\text{PID}}$ response</th>
<th>$M_{\text{3body}}$ response</th>
<th>Expected</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.28 – 0.32</td>
<td>0.31 – 0.46</td>
<td>3.17 ± 0.66</td>
<td>4</td>
</tr>
<tr>
<td>0.32 – 0.46</td>
<td></td>
<td>9.2 ± 1.1</td>
<td>6</td>
</tr>
<tr>
<td>0.40 – 0.45</td>
<td>0.46 – 0.54</td>
<td>2.89 ± 0.63</td>
<td>6</td>
</tr>
<tr>
<td>0.54 – 0.65</td>
<td>3.17 ± 0.66</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>0.65 – 0.80</td>
<td>3.64 ± 0.72</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>0.80 – 1.00</td>
<td>3.79 ± 0.80</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

| 0.28 – 0.32 | 0.46 – 0.54 | 4.22 ± 0.78 | 6 |
| 0.32 – 0.46 | 8.3 ± 1.1  | 10          |
| 0.45 – 0.54 | 2.3 ± 0.57 | 4           |
| 0.54 – 0.65 | 2.83 ± 0.63 | 8          |
| 0.65 – 0.80 | 2.72 ± 0.69 | 5          |
| 0.80 – 1.00 | 4.83 ± 0.90 | 7          |

| 0.28 – 0.32 | 0.46 – 0.54 | 2.33 ± 0.58 | 6 |
| 0.32 – 0.46 | 8.3 ± 1.1  | 8           |
| 0.54 – 0.63 | 2.07 ± 0.53 | 1          |
| 0.65 – 0.80 | 2.96 ± 0.65 | 4          |
| 0.80 – 1.00 | 3.11 ± 0.69 | 3          |

| 0.28 – 0.32 | 0.46 – 0.54 | 2.69 ± 0.62 | 1 |
| 0.32 – 0.46 | 7.5 ± 1.0  | 5           |
| 0.63 – 0.75 | 2.06 ± 0.53 | 3          |
| 0.65 – 0.80 | 3.16 ± 0.66 | 2          |
| 0.80 – 1.00 | 4.67 ± 0.84 | 2          |

| 0.28 – 0.32 | 0.46 – 0.54 | 2.19 ± 0.55 | 2 |
| 0.32 – 0.46 | 3.38 ± 0.76 | 5           |
| 0.63 – 0.75 | 1.52 ± 0.46 | 3           |
| 0.65 – 0.80 | 1.28 ± 0.47 | 1           |
| 0.80 – 1.00 | 4.42 ± 0.83 | 7           |

Table 2. Expected background candidate yields in the 7 TeV data set, with their uncertainties, and observed candidate yields within the $\tau^-$ signal window in the different bins of classifier response. The classifier responses range from 0 (most background-like) to +1 (most signal-like). The first bin in each classifier response is excluded from the analysis.
<table>
<thead>
<tr>
<th>$\mathcal{M}_{\text{PID}}$ response</th>
<th>$\mathcal{M}_{\text{3body}}$ response</th>
<th>Expected</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.26 - 0.34$</td>
<td>$39.6 \pm 2.3$</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>$0.34 - 0.45$</td>
<td>$32.2 \pm 2.1$</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>$0.45 - 0.61$</td>
<td>$28.7 \pm 2.0$</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>0.54 – 0.61</td>
<td>$0.61 - 0.70$</td>
<td>$9.7 \pm 1.2$</td>
<td>5</td>
</tr>
<tr>
<td>$0.61 - 0.70$</td>
<td>$11.4 \pm 1.3$</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>$0.70 - 0.83$</td>
<td>$7.3 \pm 1.1$</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>$0.83 - 0.94$</td>
<td>$6.0 \pm 1.0$</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.94 – 1.00</td>
<td>$13.6 \pm 1.4$</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>$0.40 - 0.54$</td>
<td>$12.1 \pm 1.3$</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>$0.54 - 0.61$</td>
<td>$8.3 \pm 1.0$</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>$0.61 - 0.70$</td>
<td>$2.60 \pm 0.62$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$0.70 - 0.83$</td>
<td>$1.83 \pm 0.60$</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>$0.83 - 0.94$</td>
<td>$2.93 \pm 0.72$</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>$0.94 - 1.00$</td>
<td>$2.69 \pm 0.63$</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>0.61 – 0.71</td>
<td>$13.5 \pm 1.4$</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>$0.61 - 0.70$</td>
<td>$10.9 \pm 1.2$</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>$0.70 - 0.83$</td>
<td>$9.7 \pm 1.2$</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>$0.83 - 0.94$</td>
<td>$3.35 \pm 0.69$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$0.94 - 1.00$</td>
<td>$4.60 \pm 0.89$</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0.71 – 0.80</td>
<td>$4.09 \pm 0.81$</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>$0.61 - 0.70$</td>
<td>$2.78 \pm 0.68$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$0.70 - 0.83$</td>
<td>$2.99 \pm 0.72$</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$0.83 - 0.94$</td>
<td>$3.93 \pm 0.81$</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$0.94 - 1.00$</td>
<td>$3.22 \pm 0.68$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$0.80 - 1.00$</td>
<td>$5.12 \pm 0.86$</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>$0.80 - 1.00$</td>
<td>$4.44 \pm 0.79$</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>$0.94 - 1.00$</td>
<td>$3.80 \pm 0.78$</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0.80 – 1.00</td>
<td>$2.65 \pm 0.68$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$0.70 - 0.83$</td>
<td>$3.05 \pm 0.67$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$0.83 - 0.94$</td>
<td>$1.74 \pm 0.54$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$0.94 - 1.00$</td>
<td>$3.36 \pm 0.70$</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Expected background candidate yields in the 8 TeV data set, with their uncertainties, and observed candidate yields within the $\tau^-$ signal window in the different bins of classifier response. The classifier responses range from 0 (most background-like) to +1 (most signal-like). The first bin in each classifier response is excluded from the analysis.
Figure 5. Distribution of CL$_s$ values as a function of the assumed branching fraction for $\tau^- \rightarrow \mu^-\mu^+\mu^-$, under the hypothesis to observe background events only. The dashed line indicates the expected limit and the solid line the observed one. The light (yellow) and dark (green) bands cover the regions of 68% and 95% confidence for the expected limit.

Acknowledgments

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FANO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (U.S.A.). The Tier1 computing centres are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom). We are indebted to the communities behind the multiple open source software packages on which we depend. We are also thankful for the computing resources and the access to software R&D tools provided by Yandex LLC (Russia). Individual groups or members have received support from EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), Conseil général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR (Russia), XuntaGal and GENCAT (Spain), Royal Society and Royal Commission for the Exhibition of 1851 (United Kingdom).
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


[33] BELLE collaboration, A. Zupanc et al., Measurements of branching fractions of leptonic and hadronic $D_s^+$ meson decays and extraction of the $D_s^+$ meson decay constant, JHEP 09 (2013) 139 [arXiv:1307.6240] [inSPIRE].

[34] LHCb collaboration, Production of $J/\psi$ and $\Upsilon$ mesons in pp collisions at $\sqrt{s} = 8$ TeV, JHEP 06 (2013) 064 [arXiv:1304.6977] [inSPIRE].


V. Shevchenko\textsuperscript{64}, A. Shires\textsuperscript{9}, R. Silva Coutinho\textsuperscript{48}, G. Simi\textsuperscript{22}, M. Sirendi\textsuperscript{47}, N. Skidmore\textsuperscript{46}, I. Skillcorn\textsuperscript{34}, T. Skwarnicki\textsuperscript{59}, N.A. Smith\textsuperscript{52}, E. Smith\textsuperscript{55,49}, E. Smith\textsuperscript{53}, J. Smith\textsuperscript{47}, M. Smith\textsuperscript{54}, H. Snoek\textsuperscript{41}, M.D. Sokoloff\textsuperscript{57}, F.J.P. Soler\textsuperscript{51}, F. Soomro\textsuperscript{39}, D. Souza\textsuperscript{48}, B. Souza De Paula\textsuperscript{2}, B. Spaan\textsuperscript{8}, P. Spradlin\textsuperscript{51}, S. Sridharan\textsuperscript{38}, F. Stagni\textsuperscript{38}, M. Stahl\textsuperscript{11}, S. Stahl\textsuperscript{11}, O. Steinkamp\textsuperscript{40}, O. Stenyakin\textsuperscript{35}, S. Stevenson\textsuperscript{55}, S. Stoica\textsuperscript{29}, S. Stone\textsuperscript{59}, B. Storacci\textsuperscript{40}, S. Straticiuc\textsuperscript{29}, U. Straumann\textsuperscript{40}, R. Strollo\textsuperscript{22}, V.K. Subbah\textsuperscript{38}, L. Sun\textsuperscript{57}, W. Sutcliffe\textsuperscript{53}, K. Swientek\textsuperscript{27}, S. Swientek\textsuperscript{2}, V. Syropoulos\textsuperscript{42}, M. Szczekowski\textsuperscript{28}, P. Szczypka\textsuperscript{39,38}, T. Szumlak\textsuperscript{27}, S. T’Jampens\textsuperscript{4}, M. Teklishyn\textsuperscript{7}, G. Tellarini\textsuperscript{16,9}, F. Teubert\textsuperscript{38}, C. Thomas\textsuperscript{55}, E. Thomas\textsuperscript{38}, J. van Tilburg\textsuperscript{41}, V. Tisserand\textsuperscript{4}, M. Tobin\textsuperscript{39}, J. Todd\textsuperscript{37}, S. Tolk\textsuperscript{42}, L. Tomassetti\textsuperscript{16,7}, D. Tonelli\textsuperscript{38}, S. Topp-Joergensen\textsuperscript{55}, N. Torr\textsuperscript{55}, E. Tourneur\textsuperscript{4}, S. Tourneur\textsuperscript{39}, M.T. Tran\textsuperscript{39}, M. Tresch\textsuperscript{10}, A. Trisovic\textsuperscript{38}, A. Tsaregorodtsev\textsuperscript{6}, P. Tsopelas\textsuperscript{41}, N. Tuning\textsuperscript{41}, M. Ubeda Garcia\textsuperscript{38}, A. Ukleja\textsuperscript{28}, A. Ustyuzhanin\textsuperscript{64}, U. Uwer\textsuperscript{11}, C. Vacca\textsuperscript{15}, V. Vagnoni\textsuperscript{14}, G. Valenti\textsuperscript{14}, A. Vallier\textsuperscript{7}, R. Vazquez Gomez\textsuperscript{18}, P. Vazquez Regueiro\textsuperscript{37}, C. Vázquez Sierra\textsuperscript{37}, S. Vecchi\textsuperscript{16}, J.J. Velthuis\textsuperscript{46}, M. Veltri\textsuperscript{17,6}, G. Veneziano\textsuperscript{29}, M. Vesterinen\textsuperscript{11}, B. Viaud\textsuperscript{7}, D. Vieira\textsuperscript{2}, M. Viteles Diaz\textsuperscript{37}, X. Vilasis-Cardona\textsuperscript{36-9}, A. Vollhardt\textsuperscript{40}, D. Volyanskyy\textsuperscript{10}, D. Voong\textsuperscript{46}, A. Vorobiev\textsuperscript{30}, V. Vorobyev\textsuperscript{34}, C. Voß\textsuperscript{63}, J.A. de Vries\textsuperscript{41}, R. Waldi\textsuperscript{63}, C. Wallace\textsuperscript{48}, R. Wallace\textsuperscript{12}, J. Walsh\textsuperscript{23}, S. Wandernoth\textsuperscript{11}, J. Wang\textsuperscript{59}, D.R. Ward\textsuperscript{47}, N.K. Watson\textsuperscript{45}, D. Websdale\textsuperscript{53}, M. Whitehead\textsuperscript{48}, J. Wicht\textsuperscript{38,9}, D. Wiedner\textsuperscript{11}, G. Wilkinson\textsuperscript{55,38}, M.P. Williams\textsuperscript{45}, M. Williams\textsuperscript{56}, H.W. Wilschut\textsuperscript{66}, F.F. Wilson\textsuperscript{49}, J. Wimberley\textsuperscript{58}, J. Wishahi\textsuperscript{9}, W. Wislicki\textsuperscript{28}, M. Witek\textsuperscript{26}, G. Wormser\textsuperscript{7}, S.A. Wotton\textsuperscript{47}, S. Wright\textsuperscript{47}, K. Wylie\textsuperscript{38}, Y. Xie\textsuperscript{61}, Z. Xing\textsuperscript{39}, Z. Xu\textsuperscript{39}, Z. Yang\textsuperscript{9}, X. Yuan\textsuperscript{3}, O. Yushchenko\textsuperscript{35}, M. Zangoli\textsuperscript{14}, M. Zavertyaev\textsuperscript{10,9}, L. Zhang\textsuperscript{59}, W.C. Zhang\textsuperscript{12}, Y. Zhang\textsuperscript{1}, A. Zhelyozov\textsuperscript{41}, A. Zhokhov\textsuperscript{31} and L. Zhong\textsuperscript{9}.

1 Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil
2 Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
3 Center for High Energy Physics, Tsinghua University, Beijing, China
4 LAPP, Université de Savoie, CNRS/IN2P3, Annecy-le-Vieux, France
5 Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, Clermont-Ferrand, France
6 CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
7 LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
8 LPNHE, Université Pierre et Marie Curie, Université Paris Diderot, CNRS/IN2P3, Paris, France
9 Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany
10 Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany
11 Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
12 School of Physics, University College Dublin, Dublin, Ireland
13 Sezione INFN di Bari, Bari, Italy
14 Sezione INFN di Bologna, Bologna, Italy
15 Sezione INFN di Cagliari, Cagliari, Italy
16 Sezione INFN di Ferrara, Ferrara, Italy
17 Sezione INFN di Firenze, Firenze, Italy
18 Laboratori Nazionali dell’INFN di Frascati, Frascati, Italy
19 Sezione INFN di Genova, Genova, Italy
20 Sezione INFN di Milano Bicocca, Milano, Italy
21 Sezione INFN di Milano, Milano, Italy
22 Sezione INFN di Padova, Padova, Italy
23 Sezione INFN di Pisa, Pisa, Italy
24 Sezione INFN di Roma Tor Vergata, Roma, Italy
25 Sezione INFN di Roma La Sapienza, Roma, Italy
26 Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
27 AGH – University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
28 National Center for Nuclear Research (NCBJ), Warsaw, Poland
29 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
30 Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
31 Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia
32 Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
33 Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia
34 Budker Institute of Nuclear Physics (SB RAS) and Novosibirsk State University, Novosibirsk, Russia
35 Institute for High Energy Physics (IHEP), Protvino, Russia
36 Universitat de Barcelona, Barcelona, Spain
37 Universidad de Santiago de Compostela, Santiago de Compostela, Spain
38 European Organization for Nuclear Research (CERN), Geneva, Switzerland
39 Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
40 Physik-Institut, Université Zürich, Zürich, Switzerland
41 Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands
42 Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, The Netherlands
43 NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
44 Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
45 University of Birmingham, Birmingham, United Kingdom
46 H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
47 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
48 Department of Physics, University of Warwick, Coventry, United Kingdom
49 STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
50 School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
51 School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
52 Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
53 Imperial College London, London, United Kingdom
54 School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
55 Department of Physics, University of Oxford, Oxford, United Kingdom
56 Massachusetts Institute of Technology, Cambridge, MA, United States
57 University of Cincinnati, Cincinnati, OH, United States
58 University of Maryland, College Park, MD, United States
59 Syracuse University, Syracuse, NY, United States
60 Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to 2
61 Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China, associated to 3
62 Departamento de Fisica, Universidad Nacional de Colombia, Bogota, Colombia, associated to 8
63 Institut für Physik, Universität Rostock, Rostock, Germany, associated to 11
64 National Research Centre Kurchatov Institute, Moscow, Russia, associated to 31
65 Instituto de Fisica Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain, associated to 36
66 Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to 41
67 Celal Bayar University, Manisa, Turkey, associated to 38

a Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
b P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
c Università di Bari, Bari, Italy
d Università di Bologna, Bologna, Italy
e Università di Cagliari, Cagliari, Italy
f Università di Ferrara, Ferrara, Italy