Explaining $h \to \mu^+\tau^-$, $B \to K^*\mu^+\mu^-$ and $B \to K_\mu^+\mu^-/B \to K_e^+\mu^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$

Andreas Crivellin,1 Giancarlo D’Ambrosio,1,2 and Julian Heeck3

1CERN Theory Division, CH-1211 Geneva 23, Switzerland
2INFN-Sezione di Napoli, Via Cintia, 80126 Napoli, Italy
3Service de Physique Théorique, Université Libre de Bruxelles, Boulevard du Triomphe, CP225, 1050 Brussels, Belgium

The LHC observed so far three deviations from the Standard Model (SM) predictions in flavour observables: LHCb reported anomalies in $B \to K^*\mu^+\mu^-$ and $R(K) = B \to K^*\mu^+\mu^-/B \to K^+\tau^-$ while CMS found an excess in $h \to \mu\tau$. We show, for the first time, how these deviations from the SM can be explained within a single well-motivated model: a two-Higgs-doublet model with gauged $L_\mu - L_\tau$ symmetry. We find that, despite the constraints from $\tau \to \mu\mu\mu$ and $B_\tau$ mixing, one can explain $h \to \mu\tau$, $B \to K^*\mu^+\mu^-$ and $R(K)$ simultaneously, obtaining interesting correlations among the observables.

I. INTRODUCTION

So far, the LHC completed the SM by discovering the last missing piece, the Brout–Englert–Higgs particle [11,2]. Furthermore, no significant direct evidence for physics beyond the SM has been found, i.e. no new particles were discovered. However, the LHC did observe three ‘hints’ for new physics (NP) in the flavor sector, which are sensitive to virtual effects of new particles and can be used as guidelines towards specific NP models: $h \to \mu\tau$, $B \to K^*\mu^+\mu^-$, and $R(K) = B \to K^*\mu^+\mu^-/B \to K^+\tau^-$. It is therefore interesting to examine if a specific NP model can explain these three anomalies simultaneously, predicting correlations among them.

LHCb reported deviations from the SM predictions [3,4] (mainly in an angular observable called $P_Z$ [5]) in $B \to K^*\mu^+\mu^-$ [6] with a significance of 2–3$\sigma$ depending on the assumptions of hadronic uncertainties [7,9]. This discrepancy can be explained in a model independent approach by rather large contributions to the Wilson coefficient $C_9$ [10,12], i.e. an operator $(\bar{\tau}\gamma_\mu b)(\bar{\tau}\gamma^\mu\mu)$, which can be achieved in models with an additional heavy neutral Z’ gauge boson [13,15]. Furthermore, LHCb recently found indications for the violation of lepton flavour universality in

$$R(K) = \frac{B \to K^\*\mu^+\mu^-}{\overline{B} \to K^\*e^+e^-} = 0.745^{+0.099}_{-0.074} \pm 0.036,$$

which disagrees from the theoretically rather clean SM prediction $R^\text{SM}_K = 1.0003 \pm 0.0001$ [17] by 2.6$\sigma$. A possible explanation is again a NP contributing to $C_9^\mu$ involving muons, but not electrons [15,20]. Interestingly, the value for $C_9$ required to explain $R(K)$ is of the same order as the one required by $B \to K^\*\mu^+\mu^-$ [8,21]. In Ref. [15], a model with gauged muon lepton number $L_\mu - L_\tau$ was proposed in order to explain the $B \to K^*\mu^+\mu^-$ anomaly.

Concerning Higgs decays, CMS recently measured a lepton-flavour violating (LFV) channel [22]

$$\text{Br}[h \to \mu\tau] = (0.89^{+0.40}_{-0.37})\%,$$

which disagrees from the SM (where this decay is forbidden) by about 2.4$\sigma$. Such LFV SM Higgs couplings are induced by a single operator up to dim-6 and $\text{Br}[h \to \mu\tau]$ can easily be up to 10% taking into account this operator only [23–25]. However, it is in general difficult to get dominant contributions to this operator in a UV complete model, as for example in models with vector-like leptons [24]. Therefore, among the several attempts to explain this $h \to \mu\tau$ observation [30–34], most of them are relying on models with extended Higgs sectors. One solution employs a two-Higgs-doublet model (2HDM) with gauged $L_\mu - L_\tau$ [35].

The abelian symmetry $U(1)_{L_\mu - L_\tau}$ is interesting in general: not only is this an anomaly-free global symmetry within the SM [35,35], it is also a good zeroth-order approximation for neutrino mixing with a quasi-degenerate mass spectrum, predicting a maximal atmospheric and vanishing reactor neutrino mixing angle [39,41]. Breaking $L_\mu - L_\tau$ is mandatory for a realistic neutrino sector, and such a breaking can also induce charged LFV processes, such as $\tau \to 3\mu$ [42,43] and $h \to \mu\mu$ [35].

Supplementing the model of Ref. [35], with the induced Z’ quark couplings of Ref. [15] can resolve all three anomalies from above. Interestingly, the semileptonic B decays imply lower limit on $g'/M_{Z'}$, which allows us to set a lower limit on $\tau \to \mu\mu\mu$, depending on $h \to \mu\tau$.

II. THE MODEL

Our model under consideration is a 2HDM with a gauged $U(1)_{L_\mu - L_\tau}$ symmetry [35]. The $L_\mu - L_\tau$ symmetry with the gauge coupling $g'$ is broken spontaneously by the vacuum expectation value (VEV) of a scalar $\Phi$ with $Q^\Phi_{L_\mu - L_\tau} = 1$, leading to the Z’ mass

$$m_{Z'} = \sqrt{2}g'\langle\Phi\rangle \equiv g'v_\Phi,$$

and Majorana masses for the right-handed neutrinos [1].

Two Higgs doublets are introduced which break the electroweak symmetry: $\Psi_1$ with $Q_{L_\mu - L_\tau} = -2$ and $\Psi_2$ [8,21,44].

1 Active neutrino masses are generated via seesaw with close-to-maximal atmospheric mixing and quasi-degenerate masses [35].
with $Q^2_{L_{e}-L_{τ}} = 0$. Therefore, $Ψ_2$ gives masses to quarks and leptons while $Ψ_1$ couples only off-diagonally to $τ\mu$:

$$L_Y \supset - \overline{\ell_f} Y_{1}^T \delta_{f1} \Psi_1 \ell_i - \xi_{τ\mu} \overline{\ell_f} \Psi_1 c_2$$

$$- \overline{Q_j} Y_{1}^U \Psi_2 u_i - \overline{Q_j} Y_{1}^d \Psi_2 d_i + h.c.$$  \hspace{1cm} (4)

Here $Q (ℓ)$ is the left-handed (quark) doublet, $u (e)$ is the right-handed up-quark (charged-lepton) and $d$ the right-handed down quark while $i$ and $f$ label the three generations. The scalar potential is the one of a $U(1)$-invariant 2HDM [33] with additional couplings to the SM-singlet $Φ$, which most importantly generates the doublet-mixing term

$$V(Ψ_1, Ψ_2, Φ) \supset 2λ_1^2 Ψ_1^2 Ψ_1^2 \equiv m^2_Ψ Ψ_1^2 Ψ_1,$$

that induces a small vacuum expectation value for $Ψ_1$ [35]. We define $tan β = Ψ_2 / Ψ_1$ and $α$ is the usual mixing angle between the neutral CP-even components of $Ψ_1$ and $Ψ_2$ (see for example [41]). We neglect the additional mixing of the CP-even scalars with $R[Φ]$. Quarks and gauge bosons have standard type-I 2HDM couplings to the scalars. The only deviations are in the lepton sector: while the Yukawa couplings $Y_{1}^T \delta_{f1}$ of $Ψ_2$ are forced to be diagonal due to the $L_µ - L_τ$ symmetry, $ξ_{τ\mu}$ gives rise to an off-diagonal entry in the lepton mass matrix:

$$m^ℓ_τ = \frac{v}{\sqrt{2}} \begin{pmatrix} y_τ sin β & 0 & 0 \\ 0 & y_τ sin β & 0 \\ ξ_{τ\mu} cos β & y_τ sin β & 0 \end{pmatrix}. \hspace{1cm} (5)$$

It is this $τ-µ$ entry that leads to the LFV couplings of $h$ and $Z'$ to interest of this letter. The lepton mass basis is obtained by simple rotations of $(µ_R, τ_R)$ and $(µ_L, τ_L)$, determined by the angles $θ_R$ and $θ_L$, respectively:

$$\sin θ_R \approx \frac{v}{\sqrt{2m_τ}} ξ_{τ\mu} cos β, \hspace{1cm} \frac{m_τ}{m_µ} \ll 1. \hspace{1cm} (6)$$

The angle $θ_L$ is automatically small and will be neglected in the following.\footnote{Choosing $Q^2_{L_{e}-L_{τ}} = 4$ for $Ψ_2$ would essentially exchange $θ_L \leftrightarrow θ_R$ [35], with little impact on our study.} A non-vanishing angle $θ_R$ not only gives rise to the LFV decay $h \to µτ$ due to the coupling

$$\frac{m_τ}{v} \frac{cos(α - β)}{cos(β)} sin(θ_R) cos(θ_R) P_R µ h \equiv Γ^{h}_{µτ} P_R µ h, \hspace{1cm} (7)$$

in the Lagrangian, but also leads to off-diagonal $Z'$ couplings to right-handed leptons

$$g'Z'_{µ} (ψ, τ) = \begin{pmatrix} cos θ_R \sin θ_R \sin θ_R \cos θ_R \end{pmatrix} γ_{µ} P_R \left( µ \right), \hspace{1cm} (8)$$

while the left-handed couplings are to a good approximation flavour conserving. In order to explain the observed anomalies in the $B$ meson decays, a coupling of the $Z'$ to quarks is required as well, not inherently part of $L_µ - L_τ$ models (aside from the kinetic $Z-Z'$ mixing, which is assumed to be small). Following Ref. [35], we introduce heavy vector-like quarks, i.e. $Q_L\equiv (U_L,D_L), D'_R, U'_R$, and their chiral partners $Q_R\equiv (U_R,D_R), D'_L, U'_L$, with vector-like mass terms

$$m_Q \overline{Q}_L Q_R + m_D \overline{D}_L D_R + m_U \overline{U}_L U_R + h.c., \hspace{1cm} (9)$$

and $L_µ - L_τ$ charges +1 (i.e. $Q^u_{L_{e}-L_{τ}} = Q^µ_{L_{e}-L_{τ}} = -1$), coupling them to the $Z'$ boson. Yukawa-like couplings involving the heavy vector-quarks, the light chiral quarks and $Φ$

$$Φ \sum_{j=1}^{3} \left( \overline{D}_j Y_j^D P_R d_j + \overline{U}_j Y_j^U P_R u_j \right) + h.c. \hspace{1cm} (10)$$

then induce couplings of the SM quarks to the $Z'$ once $Φ$ acquires its VEV. Thus, integrating out the heavy vector-like quarks gives rise to effective $Z'd_d d_j$ couplings [35][40] of the form

$$g' \left( \overline{d}_i γ^µ P_R d_j Z'_{ij} Γ^{dL}_{ij} + \overline{d}_i γ^µ P_R d_j Z'_{ij} Γ^{dU}_{ij} \right), \hspace{1cm} (11)$$

with hermitian matrices $Γ^{dL}_{ij}$ that are related to the vector quark masses $m_{Q,D,U}$ and Yukawa couplings $Y_{Q,D,U}$ as follows [15]:

$$Γ^{dL}_{ij} \approx -\frac{v_2^φ}{2m^2_D} (Y_i^D Y_j^{D*}), \hspace{1cm} Γ^{dU}_{ij} \approx \frac{v_2^φ}{2m^2_Q} (Y_i^Q Y_j^{Q*}) \hspace{1cm} (12)$$

which holds in the approximation $|Γ^{dL/L}_{ij}| \ll 1.3$. \hspace{1.5cm} III. FLAVOUR OBSERVABLES

We will now recall the necessary formula in the region of interest (i.e. small $θ_R$) considering only the processes giving to most relevant bounds on our model, i.e. $B_s-B_s$ mixing, neutrino trident production and $τ \rightarrow 3µ$.

A. $h \rightarrow µτ$

The branching ratio for $h \rightarrow µτ$ reads

$$Br \left[ h \rightarrow µτ \right] \approx \frac{m_h}{8πΓ^{h}_{µτ}} |Γ^{h}_{µτ}|^2, \hspace{1cm} (13)$$

where $Γ_{SM} \approx 4.1$ MeV is the decay width in the SM for a 125 GeV Higgs [17] and $Γ^{h}_{µτ}$ is defined in Eq. (7). Comparing this to Eq. (2), one sees that both $sin θ_R$ and $cos(α - β)$ are required to explain the CMS excess [35].

\footnote{Compared to Ref. [15], the vector-like quarks also have Yukawa couplings $y_{ψ_1}$ to the $L_µ - L_τ$-charged scalar doublet $Ψ_1$. This induces a small additional mass mixing among the heavy quarks, and also a coupling to $h$ suppressed by $y_{ψ_1} cos(α - β)$. We assume these couplings $y_{ψ_1}$ to be small to avoid large contributions to $gg \rightarrow h$ and $h \rightarrow γγ$.}
FIG. 1: Left: Allowed regions in the $\cos(\alpha - \beta)\sin(\theta_R)$ plane. The blue (light blue) region corresponds to the 1σ (2σ) region of the CMS measurement of $h \to \mu\tau$ for $\tan \beta = 50$; yellow stands for $\tan \beta = 10$. The (dashed) red contours mark deviations of $h \to \tau\tau$ by 10% compared to the SM for $\tan \beta = 50$ (10). The vertical green lines illustrate the naive LHC limit $|\cos(\alpha - \beta)| \lesssim 0.4$, horizontal lines denote the 90% C.L. limit on $\tau \to 3\mu$ via $Z'$ exchange.

Right: Allowed regions in the $\Gamma_{23}^{\text{DL}} - m_{Z'}/g'$ plane from $B \to K'\mu^+\mu^-$ and $R(K)$ (yellow) and $B_s$ mixing (blue). For $B_s$ mixing (light) blue corresponds to $(m_Q = 15m_{Z'}/g')$, $m_{Z'}/g'$. The horizontal lines denote the lower bounds on $m_{Z'}/g'$ from $\tau \to 3\mu$ for $\sin(\theta_R) = 0.05, 0.02, 0.005$. The gray region is excluded by NTP.

B. Lepton decays

While the Higgs contributions to $\tau \to \mu\mu$ and $\tau \to \mu\gamma$ turn out to be very small in most regions of parameter space due to the small lepton masses involved, the $Z'$ contributions to $\tau \to 3\mu$ can be sizable and restrict $\theta_R^2/v^4_f$, but highly suppressed by $2\alpha/\pi$, and hence not as restrictive.

C. $B \to K^*\mu^+\mu^-$ and $B \to K_{\mu4}\mu^-/B \to K^{\pm}\mu^-$

Both $B \to K^*\mu^+\mu^-$ and $R(K)$ are sensitive to the Wilson coefficients $C_9^{(\text{CQ})}$ and $C_9^{(\text{CQ})}$. While in our model the contribution to $C_{10}$ is suppressed by $\sin(2\theta_R)$ (or even $\sin(2\theta_L)$), the Wilson coefficients $C_9^{(\text{CQ})}$ and $C_9^{(\text{CQ})}$ with muons are generated (as well as $C_7^{(\text{CQ})}$ and the $\theta_R$ suppressed $C_9\tau^\mu$). $C_9^{(\text{CQ})}$ is not affected, which naturally generates violations of lepton flavour universality in $B \to K_{\mu4}\mu^-/B \to K^{\pm}\mu^-$. We find

$$C_9^{(\text{CQ})} \sim \frac{g^2}{\sqrt{2m_{Z'}^2},} \sin(2\theta_R)^4 \sin^2(2\theta_R)$$

which has to be compared to the current upper limit of $2.1 \times 10^{-8}$ at 90% C.L. obtained by Belle. A combination with data from BaBar gives an even stronger limit of $1.2 \times 10^{-8}$ at 90% C.L., to be used in the following. For small $\theta_R$, the branching ratio for $\tau \to \mu\gamma$ is proportional to the same combination $\theta_R^2/v^4_f$, but highly suppressed by $2\alpha/\pi$, and hence not as restrictive.

D. $B_s$-$\bar{B}_s$ mixing

The interactions of $Z'$ and $\Phi$ relevant for $B \to K_{\mu4}\mu^-$ also contribute to $B_s$-$\bar{B}_s$ mixing. For $m_D \gg m_Q$,
we get
\[
\frac{M_{12}}{M_{13}^2} \simeq 1 + \left( \frac{\Gamma_{12}^D}{\Gamma_{13}^D} \right)^2 \left( \frac{4}{g_1^2 \sin^2 \beta + m_{W_1}^2} \right)^2 \left( \frac{V_{13}^* V_{13}}{g_2} \right)^2 S_0.
\] (17)

We require the NP contribution to be less than 15% in order to satisfy the experimental bounds [15]. Due to the dominance of the vector-quark Q we can express \( \Gamma_{12}^D \) directly in terms of \( C_9^{\mu\mu} \) from Eq. (15) and find the upper bounds
\[
m_{Z'}/g' < 3.2 \text{ TeV}/|C_9^{\mu\mu}|, \quad m_Q < 41 \text{ TeV}/|C_9^{\mu\mu}|.
\] (18)

Combining Eq. (18) with Eq. (16) then gives an upper bound of \( m_{Z'}/g' < 4 \text{ TeV} \) (6.5 TeV) at 1 \( \sigma \) (2 \( \sigma \)).

**E. Neutrino trident production**

The most stringent bound on flavour-diagonal \( Z' \) couplings to muons arises from neutrino trident production (NTP) \( \nu_\mu N \to \nu_\mu N \mu^+ \mu^- \) [15, 53]:
\[
\sigma_{\text{NTP}} \approx \frac{1 + \left( 1 + 4 s_W^2 + 8 \frac{s_W^2 m_{V_1}^2}{m_{Z'}^2 g_2^2} \right)^2}{1 + (1 + 4 s_W^2)^2} \cdot \frac{\sigma_{\text{SM NTP}}}{\Gamma_{12}^D} \cdot \frac{\Gamma_{13}^D}{\Gamma_{13}^D}.
\] (19)

Seeing as our region of interest is in the small \( \theta_R \) regime, the NTP bound is basically independent of the angle \( \theta_R \). Taking only the CCFR data [24], we get roughly \( m_{Z'}/g' \lesssim 550 \text{ GeV} \) at 95\% C.L. Compared to \( \tau \to \mu \mu \mu \) the trident neutrino bound only dominates for very small values of \( \theta_R \), roughly when \( \theta_R \lesssim 10^{-3} \) (see Fig. 2 (right)).

For \( m_{Z'} > m_Z \), the LHC constraints from the process \( pp \to \mu \mu Z' \to 4 \mu \) (or 3\( \mu \) plus missing energy) [35] are currently weaker than NTP [15], but will become competitive with higher luminosities [56, 58].

**F. Phenomenological analysis**

Concerning the phenomenological consequences of our model, let us first consider the implications of \( h \to \mu \tau \). In the left plot of Fig. 1 we show the regions in the \( \cos(\alpha - \beta) \)–\( \sin(\theta_R) \) plane which can explain \( h \to \mu \tau \) at the 1\( \sigma \) and 2\( \sigma \) level for different values of \( \tan \beta \). Measurements of the h couplings to vector bosons require \( |\cos(\alpha - \beta)| \lesssim 0.4 \) [59, 60] while the Higgs effects in \( \tau \to 3 \mu \) and \( \tau \to \mu \gamma \) are typically negligible [35]. As a side effect, the \( h \to \mu \tau \) rate also implies a change in the \( h \to \tau \tau \) rate, although this is negligible in regions with small \( \theta_R \). In addition we show the regions compatible with \( \tau \to 3 \mu \) for various values of \( m_{Z'}/g' \). Note that \( g' \lesssim 0.3 \) in order to avoid a Landau pole below the Planck scale. In summary, small values of \( \theta_R \) can explain the CMS \( h \to \mu \tau \) excess for moderate to large values of \( \tan \beta \) for \( \cos(\alpha - \beta) \approx 0.1 \).

In the right plot of Fig. 1 we examine which regions in parameter space can account for \( B \to K^\ast \mu^+ \mu^- \) taking into account the constraints from \( B_s - \bar{B}_s \) mixing. Since we focus on the limit \( M_D \to \infty \) (i.e. \( C_9^{\mu \mu} \to 0 \)) we find that unless \( \Gamma_{13}^D \) is rather large, \( B \to K^\ast \mu^+ \mu^- \) can be explained without violating bounds from \( B_s - \bar{B}_s \). Only a very small \( \Gamma_{25}^D \) independent region is excluded by NTP. In addition, bounds from \( \tau \to 3 \mu \) depending on \( \sin(\theta_R) \) can be obtained.

Concerning \( \tau \to 3 \mu \), future sensitivities down to \( Br[\tau \to 3 \mu] \approx 10^{-9} \) seem feasible [61] and will cut deep into our parameter space (see Fig. 2). Using the 1\( \sigma \) limits on \( h \to \mu \tau \) to fix \( \theta_R \) and \( B_s \) mixing with \( C_9^{\mu \mu} \) to fix \( m_{Z'}/g' \) as well as the LHC limit \( |\cos(\alpha - \beta)| < 0.4 \) – we can obtain a lower limit on the rate \( \tau \to 3 \mu \)
\[
Br[\tau \to 3 \mu] \gtrsim 3 \times 10^{-8} \left( \frac{10}{\tan \beta} \right)^2,
\] (20)
which implies \( \tan \beta \gtrsim 18 \) with current data [59] and \( \tan \beta \gtrsim 61 \) if branching ratios down to \( 10^{-9} \) can be probed in the future. This is the main prediction of our simultaneous explanation of \( h \to \mu \tau \), \( B \to K^\ast \mu^+ \mu^- \) and \( R(K) \).

Finally, we remark that a \( Z' \)–\( Z' \) mixing angle \( \theta_{ZZ'} \) [15] is induced by the VEV of \( \Psi_1 \) [35]
\[
|g' \theta_{ZZ'}| \approx \frac{g_1 v^2 \sin \beta}{m_{Z'}/g'^2} \approx 10^{-4} \left( \frac{20}{\tan \beta} \right)^2 \left( \frac{\text{TeV}}{m_{Z'}/g'} \right)^2,
\] (21)
which leads to small shifts in the vector couplings of \( Z \) to muons and taus
\[
g'^2(\mu\mu, \tau\tau) \approx -1/2 + 2s_W^2 \pm g' \theta_{ZZ'}/(g/cW),
\] (22)
and thus ultimately to lepton non-universality [49]. For the values of interest to our study (see Fig. 2), and in
the limit $m_Z \ll m_{Z'}$, the shift is automatically small enough to satisfy experimental bounds and leads to tiny branching ratios $Z \to \mu\tau$ below $10^{-8}$ (for $\tan \beta < 0.1$). Note that the couplings to electrons and quarks remain unaffected. For $m_{Z'} > m_Z$, the $\rho$ parameter is enhanced by \cite{45}

$$\rho - 1 \simeq 1.2 \times 10^{-4} \left( \frac{\theta_{ZZ'}}{10^{-3}} \right)^2 \left( \frac{m_{Z'}}{10 \text{ TeV}} \right)^2,$$

and is therefore compatible with electroweak precision data ($\rho - 1 < 9 \times 10^{-4}$ at 2$\sigma$ \cite{62}) for the parameter space studied in this letter.

IV. CONCLUSIONS

In this letter we showed for the first time that all three LHC anomalies in the flavour sector can be explained within a single well-motivated model: A 2HDM with a gauged $L_u - L_s$ symmetry and effective $Z'sb$ couplings induced by heavy vector-like quarks. Except for the $\tau - \mu$ couplings, the Higgs sector resembles the one of a 2HDM of type-I. Therefore, the constraints from $h$ decays or LHC searches for $\lambda^0 \to \tau^+\tau^-$ are rather weak and $h \to \mu\tau$ can be easily explained in a wide parameter space. The model can also account for the deviations from the SM in $B \to K\mu^+\mu^-$ and naturally leads to the right amount of lepton-flavour-universality violating effects in $R(K)$. Due to the small values of the $\tau - \mu$ mixing angle $\theta_R$, sufficient to account for $h \to \mu\tau$, the $Z'$ contributions to $\tau \to 3\mu$ are not in conflict with present bounds for large $\tan \beta$ in wide ranges of parameter space. Interestingly, $B \to K^{*}\mu^+\mu^-$ and $R(K)$ combined with $B_s\bar{B}_s$ put a upper limit on $m_{Z'}/y'$ resulting in a lower limit on $\tau \to 3\mu$ if $\text{Br}[h \to \mu\tau] \neq 0$: for lower values of $\tan \beta$ the current experimental bounds are reached and future sensitivities will allow for a more detailed exploration of the allowed parameter space. The possible range for the $L_u - L_s$ breaking scale further implies the masses of the $Z'$ and the right-handed neutrinos to be at the TeV scale, potentially testable at the LHC with interesting additional consequences for LFV observables.

Acknowledgments

A. Crivellin is supported by a Marie Curie Intra-European Fellowship of the European Community’s 7th Framework Programme under contract number PIEF-GA-2012-326948. G. D’Ambrosio acknowledges the partial support my MIUR under the project number 2010YJ2NYW. The work of J. Heeck is funded in part by IISN and by Belgian Science Policy (IAP VII/37). We thank Gian Giudice for useful discussions. We are grateful to Wolfgang Altmannshofer and David Straub for useful discussions and additional information concerning the model-independent fit to $C_{\mu\mu}^{\mu\mu}$ and $C_{\mu\mu'}^{\mu\mu'}$.

Note added: during the publication process of this letter, CMS has released its final analysis of the $h \to \mu\tau$ search as a preprint \cite{84}, resulting in slightly changed values – $\text{Br}[h \to \mu\tau] = (0.84_{-0.39}^{+0.39})\%$ – which have however only a small impact on our study.

\begin{thebibliography}{99}
[1] G. Aad et al. (ATLAS Collaboration), Phys.Lett. B716, 1 (2012), 1207.7214.
[2] S. Chatrchyan et al. (CMS Collaboration), Phys.Lett. B716, 30 (2012), 1207.7235.
[3] U. Egede, T. Hurth, J. Matias, M. Ramon, and W. Reece, JHEP 0811, 032 (2008), 0807.2589.
[4] W. Altmannshofer, P. Ball, A. Bharucha, A. J. Buras, D. M. Straub, et al., JHEP 0901, 019 (2009), 0811.1214.
[5] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, JHEP 1305, 137 (2013), 1303.5794.
[6] R. Aaij et al. (LHCb collaboration), Phys.Rev.Lett. 111, 191801 (2013), 1308.1707.
[7] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, JHEP 1412, 125 (2014), 1407.8526.
[8] W. Altmannshofer and D. M. Straub (2014), 1411.3161.
[9] S. J¨ager and J. Martin Camalich (2014), 1412.3183.
[10] S. Descotes-Genon, J. Matias, and J. Virto, Phys.Rev. D88, 074002 (2013), 1307.5683.
[11] W. Altmannshofer and D. M. Straub, Eur.Phys.J. C73, 2646 (2013), 1308.1501.
[12] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, Phys.Rev.Lett. 112, 212003 (2014), 1310.3887.
[13] R. Gauld, F. Goertz, and U. Haisch, Phys.Rev. D89, 015005 (2014), 1308.1599.
[14] A. J. Buras, F. De Fazio, and J. Girrbach, JHEP 1402, 112 (2014), 1311.6729.
[15] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Phys.Rev. D89, 095033 (2014), 1403.1269.
[16] R. Aaij et al. (LHCb collaboration), Phys.Rev.Lett. 113, 151601 (2014), 1406.6482.
[17] C. Bobeth, G. Hiller, and G. Piranishvili, JHEP 0712, 015025 (2013), 1210.8241.
[18] R. Alonso, B. Grinstein, and J. Martin Camalich, Phys.Rev.Lett. 113, 241802 (2014), 1407.7044.
[19] G. Hiller and M. Schmaltz, Phys.Rev. D90, 054014 (2014), 1408.1627.
[20] D. Ghosh, M. Nardecchia, and S. Renner, JHEP 1412, 131 (2014), 1408.4097.
[21] T. Hurth, F. Mahmoudi, and S. Neshatpour, JHEP 1412, 053 (2014), 1410.4545.
[22] CMS (CMS Collaboration) (2014), CMS-PAS-HIG-14-005.
[23] R. Harnik, J. Kopp, and J. Zupan, JHEP 1303, 026 (2013), 1209.1397.
[24] G. Blankenburg, J. Ellis, and G. Isidori, Phys.Lett. B712, 386 (2012), 1209.5704.
[25] S. Davidson and P. Verdier, Phys.Rev. D86, 111701 (2012), 1211.1248.
[26] A. Arhrib, Y. Cheng, and O. C. Kong, Phys.Rev. D87, 015025 (2013), 1210.8241.
[27] A. Arhrib, Y. Cheng, and O. C. Kong, Europhys.Lett. 101, 31003 (2013), 1208.4669.
[28] J. Kopp and M. Nardecchia, JHEP 1410, 156 (2014), 1406.5303.
\end{thebibliography}
[29] A. Falkowski, D. M. Straub, and A. Vicente, JHEP 1405, 092 (2014), 1312.5329.
[30] A. Dery, A. Efrati, Y. Nir, Y. Soreq, and V. Susi, Phys.Rev. D90, 115022 (2014), 1408.1371.
[31] M. D. Campos, A. E. C. Hernández, H. Pás, and E. Schumacher (2014), 1408.1652.
[32] A. Celis, V. Cirigliano, and E. Passemar (2014), 1409.4439.
[33] D. Aristizabal Sierra and A. Vicente, Phys.Rev. D90, 115004 (2014), 1409.7690.
[34] C.-J. Lee and J. Tandean (2014), 1410.6803.
[35] J. Heeck, M. Holthausen, W. Rodejohann, and Y. Shimizu (2014), 1412.3671.
[36] X. He, G. C. Joshi, H. Lew, and R. Volkas, Phys.Rev. D43, 22 (1991).
[37] R. Foot, Mod.Phys.Lett. A6, 527 (1991).
[38] X.-G. He, G. C. Joshi, H. Lew, and R. Volkas, Phys.Rev. D44, 2118 (1991).
[39] P. Binetruy, S. Lavignac, S. T. Petcov, and P. Ramond, Nucl.Phys. B496, 3 (1997), hep-ph/9610481.
[40] N. F. Bell and R. R. Volkas, Phys.Rev. D63, 013006 (2001), hep-ph/0008177.
[41] S. Choubey and W. Rodejohann, Eur.Phys.J. C40, 259 (2005), hep-ph/0411190.
[42] G. Dutta, A. S. Joshipura, and K. Vijaykumar, Phys.Rev. D50, 2109 (1994), hep-ph/9405292.
[43] J. Heeck and W. Rodejohann, Phys.Rev. D84, 075007 (2011), 1107.5238.
[44] G. Branco, P. Ferreira, L. Lavoura, M. Rebelo, M. Sher, et al., Phys.Rept. 516, 1 (2012), 1106.0034.
[45] P. Langacker, Rev.Mod.Phys. 81, 1199 (2009), 0801.1345.
[46] A. J. Buras, F. De Fazio, and J. Girrbach, JHEP 1302, 116 (2013), 1211.1896.
[47] S. Dittmaier, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka, et al (2012), 1201.3084.
[48] B. Aubert et al. (BaBar Collaboration), Phys.Rev.Lett. 104, 021802 (2010), 0908.2381.
[49] J. Lees et al. (BaBar Collaboration), Phys.Rev. D81, 111101 (2010), 1002.4550.
[50] Y. Amhis et al. (Heavy Flavor Averaging Group (HFAG)) (2014), 1412.7515.
[51] S. Descotes-Genon, J. Matias, and J. Virto, PoS EPS-HEP2013, 361 (2013), 1311.3876.
[52] S. L. Glashow, D. Guadagnoli, and K. Lane (2014), 1411.0565.
[53] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Phys.Rev.Lett. 113, 091801 (2014), 1406.2332.
[54] S. Mishra et al. (CCFR Collaboration), Phys.Rev.Lett. 66, 3117 (1991).
[55] E. Ma, D. Roy, and S. Roy, Phys.Lett. B525, 101 (2002), hep-ph/0110146.
[56] K. Harigaya, T. Igar, M. M. Nojiri, M. Takeuchi, and K. Tobe, JHEP 1403, 105 (2014), 1311.0870.
[57] N. F. Bell, Y. Cai, R. K. Leane, and A. D. Medina, Phys.Rev. D90, 035027 (2014), 1407.3001.
[58] F. del Aguila, M. Chala, J. Santiago, and Y. Yamamoto (2014), 1411.7394.
[59] B. Dumont, J. F. Gunion, Y. Jiang, and S. Kraml, Phys.Rev. D90, 035021 (2014), 1405.3584.
[60] B. Dumont, J. F. Gunion, Y. Jiang, and S. Kraml (2014), 1409.4088.
[61] T. Asheev, W. Bartel, A. Bondar, J. Brodzicka, T. Browder, et al. (2010), 1002.5012.
[62] K. Olive et al. (Particle Data Group), Chin.Phys. C38, 090001 (2014).
[63] V. Khachatryan et al. (CMS Collaboration) (2015), 1502.07400.