# Explaining $\boldsymbol{h} \rightarrow \boldsymbol{\mu}^{ \pm} \boldsymbol{\tau}^{\mp}, \boldsymbol{B} \rightarrow \boldsymbol{K}^{*} \boldsymbol{\mu}^{+} \boldsymbol{\mu}^{-}$, and $\boldsymbol{B} \rightarrow \boldsymbol{K} \boldsymbol{\mu}^{+} \boldsymbol{\mu}^{-} / \boldsymbol{B} \rightarrow \boldsymbol{K} \boldsymbol{e}^{+} \boldsymbol{e}^{-}$in a Two-Higgs-Doublet Model with Gauged $L_{\mu}-L_{\tau}$ 

Andreas Crivellin, ${ }^{1}$ Giancarlo D'Ambrosio, ${ }^{1,2}$ and Julian Heeck ${ }^{3}$<br>${ }^{1}$ CERN Theory Division, CH-1211 Geneva 23, Switzerland<br>${ }^{2}$ INFN-Sezione di Napoli, Via Cintia, 80126 Napoli, Italy<br>${ }^{3}$ Service de Physique Théorique, Université Libre de Bruxelles, Boulevard du Triomphe, CP225, 1050 Brussels, Belgium (Received 13 January 2015; published 14 April 2015)


#### Abstract

The LHC has observed, so far, 3 deviations from the Standard Model (SM) predictions in flavor observables: LHCb reported anomalies in $B \rightarrow K^{*} \mu^{+} \mu^{-}$and $R(K)=B \rightarrow K \mu^{+} \mu^{-} / B \rightarrow K e^{+} e^{-}$, while CMS found an excess in $h \rightarrow \mu \tau$. 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 $\tau \rightarrow \mu \mu \mu$ and $B_{s}-\bar{B}_{s}$ mixing, one can explain $h \rightarrow \mu \tau$, $B \rightarrow K^{*} \mu^{+} \mu^{-}$and $R(K)$ simultaneously, obtaining interesting correlations among the observables.


DOI: 10.1103/PhysRevLett.114.151801
PACS numbers: $12.60 . \mathrm{Fr}, 13.20 . \mathrm{He}, 13.35 . \mathrm{Dx}, 14.70 . \mathrm{Pw}$

Introduction.-So far, the LHC has completed the Standard Model (SM) by discovering the last missing piece, the Brout-Englert-Higgs particle [1,2]. 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: $h \rightarrow \mu \tau, B \rightarrow K^{*} \mu^{+} \mu^{-}$, and $R(K)=B \rightarrow K \mu^{+} \mu^{-} /$ $B \rightarrow K e^{+} e^{-}$. 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 $P_{5}^{\prime}$ [5]) in $B \rightarrow K^{*} \mu^{+} \mu^{-}$[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 $\left(\bar{s} \gamma_{\alpha} P_{L} b\right)\left(\bar{\mu} \gamma^{\alpha} \mu\right)$, which can be achieved in models with an additional heavy neutral $Z^{\prime}$ gauge boson [13-15]. Furthermore, LHCb [16] recently found indications for the violation of lepton-flavor universality in

$$
\begin{equation*}
R(K)=\frac{B \rightarrow K \mu^{+} \mu^{-}}{B \rightarrow K e^{+} e^{-}}=0.745_{-0.074}^{+0.090} \pm 0.036 \tag{1}
\end{equation*}
$$

which disagrees from the theoretically rather clean SM prediction $R_{K}^{S M}=1.0003 \pm 0.0001$ [17] by $2.6 \sigma$. A possible explanation is again a NP contributing to $C_{9}^{\mu \mu}$ involving muons, but not electrons [18-20]. Interestingly, the value for

[^0]$C_{9}$ required to explain $R(K)$ is of the same order as the one required by $B \rightarrow K^{*} \mu^{+} \mu^{-}$[8,21]. In Ref. [15], a model with gauged muon minus tauon number $\left(L_{\mu}-L_{\tau}\right)$ was proposed in order to explain the $B \rightarrow K^{*} \mu^{+} \mu^{-}$anomaly.

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

$$
\begin{equation*}
\operatorname{Br}[h \rightarrow \mu \tau]=\left(0.89_{-0.37}^{+0.40}\right) \% \tag{2}
\end{equation*}
$$

which differs 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 $\operatorname{Br}[h \rightarrow \mu \tau]$ can easily be up to $10 \%$, taking into account this operator only [23-28]. However, it is, in general, difficult to get dominant contributions to this operator in a UV complete model, as, for example, in models with vectorlike leptons [29]. Therefore, among the several attempts to explain this $h \rightarrow \mu \tau$ observation [30-34], most of them rely 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 $\mathrm{U}(1)_{L_{\mu^{-}}-L_{\tau}}$ is interesting, in general: Not only is this an anomaly-free global symmetry within the SM [36-38], it is also a good zeroth-order approximation for neutrino mixing with a quasidegenerate 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 $\tau \rightarrow 3 \mu[42,43]$ and $h \rightarrow \mu \tau$ [35].

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

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

$$
\begin{equation*}
m_{Z^{\prime}}=\sqrt{2} g^{\prime}\langle\Phi\rangle \equiv g^{\prime} v_{\Phi} \tag{3}
\end{equation*}
$$

and Majorana masses for the right-handed neutrinos. (Active neutrino masses are generated via the seesaw mechanism with close-to-maximal atmospheric mixing and quasidegenerate masses [35].)

Two Higgs doublets are introduced which break the electroweak symmetry: $\Psi_{1}$ with $Q_{L_{\mu}-L_{\tau}}^{\Psi_{1}}=-2$ and $\Psi_{2}$ with $Q_{L_{\mu}-L_{\tau}}^{\Psi_{2}}=0$. Therefore, $\Psi_{2}$ gives masses to quarks and leptons while $\Psi_{1}$ couples only off-diagonally to $\tau \mu$ :

$$
\begin{align*}
\mathcal{L}_{Y} \supset & -\bar{\ell}_{f} Y_{i}^{\ell} \delta_{f i} \Psi_{2} e_{i}-\xi_{\tau \mu} \bar{\ell}_{3} \Psi_{1} e_{2} \\
& -\bar{Q}_{f} Y_{f i}^{u} \tilde{\Psi}_{2} u_{i}-\bar{Q}_{f} Y_{f i}^{d} \Psi_{2} d_{i}+\text { H.c. } \tag{4}
\end{align*}
$$

Here $Q(\ell)$ is the left-handed quark (lepton) doublet, $u(e)$ is the right-handed up quark (charged lepton), and $d$ is the right-handed down quark, while $i$ and $f$ label the three generations. The scalar potential is of a $\mathrm{U}(1)$-invariant 2HDM [44] with additional couplings to the SM singlet $\Phi$, which, most importantly, generates the doublet-mixing term

$$
V\left(\Psi_{1}, \Psi_{2}, \Phi\right) \supset 2 \lambda \Phi^{2} \Psi_{2}^{\dagger} \Psi_{1} \rightarrow \lambda v_{\Phi}^{2} \Psi_{2}^{\dagger} \Psi_{1} \equiv m_{3}^{2} \Psi_{2}^{\dagger} \Psi_{1}
$$

that induces a small vacuum expectation value for $\Psi_{1}$ [35]. We define $\tan \beta=\left\langle\Psi_{2}\right\rangle /\left\langle\Psi_{1}\right\rangle$, and $\alpha$ is the usual mixing angle between the neutral $C P$-even components of $\Psi_{1}$ and $\Psi_{2}$ (see, for example, Ref. [44]). We neglect the additional mixing of the $C P$-even scalars with $\operatorname{Re}[\Phi]$.

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 $Y_{i}^{\ell} \delta_{f i}$ of $\Psi_{2}$ are forced to be diagonal due to the $L_{\mu}-L_{\tau}$ symmetry, $\xi_{\tau \mu}$ gives rise to an off-diagonal entry in the lepton mass matrix:

$$
m_{f i}^{\ell}=\frac{v}{\sqrt{2}}\left(\begin{array}{ccc}
y_{e} \sin \beta & 0 & 0  \tag{5}\\
0 & y_{\mu} \sin \beta & 0 \\
0 & \xi_{\tau \mu} \cos \beta & y_{\tau} \sin \beta
\end{array}\right)
$$

It is this $\tau-\mu$ entry that leads to the LFV couplings of $h$ and $Z^{\prime}$ of interest to this Letter. The lepton mass basis is obtained by simple rotations of $\left(\mu_{R}, \tau_{R}\right)$ and $\left(\mu_{L}, \tau_{L}\right)$ with the angles $\theta_{R}$ and $\theta_{L}$, respectively:

$$
\begin{equation*}
\sin \theta_{R} \simeq \frac{v}{\sqrt{2} m_{\tau}} \xi_{\tau \mu} \cos \beta, \quad \frac{\tan \theta_{L}}{\tan \theta_{R}}=\frac{m_{\mu}}{m_{\tau}} \ll 1 \tag{6}
\end{equation*}
$$

The angle $\theta_{L}$ is automatically small and will be neglected in the following. (Choosing $Q_{L_{\mu}-L_{\tau}}=+2$ for $\Psi_{2}$ would essentially exchange $\theta_{L} \leftrightarrow \theta_{R}$ [35], with little impact on our
study.) A nonvanishing angle $\theta_{R}$ not only gives rise to the LFV decay $h \rightarrow \mu \tau$ due to the coupling

$$
\begin{equation*}
\frac{m_{\tau}}{v} \frac{\cos (\alpha-\beta)}{\cos (\beta) \sin (\beta)} \sin \left(\theta_{R}\right) \cos \left(\theta_{R}\right) \bar{\tau} P_{R} \mu h \equiv \Gamma_{\tau \mu}^{h} \bar{\tau} P_{R} \mu h \tag{7}
\end{equation*}
$$

in the Lagrangian, but also leads to off-diagonal $Z^{\prime}$ couplings to right-handed leptons

$$
g^{\prime} Z_{\nu}^{\prime}(\bar{\mu}, \bar{\tau})\left(\begin{array}{cc}
\cos 2 \theta_{R} & \sin 2 \theta_{R}  \tag{8}\\
\sin 2 \theta_{R} & -\cos 2 \theta_{R}
\end{array}\right) \gamma^{\nu} P_{R}\binom{\mu}{\tau}
$$

while the left-handed couplings are, to a good approximation, flavor conserving. In order to explain the observed anomalies in the $B$ meson decays, a coupling of the $Z^{\prime}$ to quarks is required as well, which is not inherently part of $L_{\mu}-L_{\tau}$ models (aside from the kinetic $Z-Z^{\prime}$ mixing, which is assumed to be small). Following Ref. [15], we introduce heavy vectorlike quarks, i.e., $Q_{L} \equiv\left(U_{L}, D_{L}\right), D_{R}^{c}, U_{R}^{c}$ and their chiral partners $\tilde{Q}_{R} \equiv\left(\tilde{U}_{R}, \tilde{D}_{R}\right), \tilde{D}_{L}^{c}, \tilde{U}_{L}^{c}$, with vectorlike mass terms

$$
\begin{equation*}
m_{Q} \bar{Q}_{L} \tilde{Q}_{R}+m_{D} \overline{\tilde{D}}_{L} D_{R}+m_{U} \overline{\tilde{U}}_{L} U_{R}+\text { H.c. } \tag{9}
\end{equation*}
$$

and $L_{\mu}-L_{\tau}$ charges +1 (i.e., $Q_{L_{\mu}{ }^{-} L_{\tau}}^{D_{R}}=Q_{L_{\mu}-L_{\tau}}^{U_{R}}=-1$ ), coupling them to the $Z^{\prime}$ boson. Yukawa-like couplings involving the heavy vector quarks, the light chiral quarks and $\Phi$,

$$
\begin{align*}
& \Phi \sum_{j=1}^{3}\left(\overline{\tilde{D}}_{R} Y_{j}^{Q} P_{L} d_{j}+\overline{\tilde{U}}_{R} Y_{j}^{Q} P_{L} u_{j}\right) \\
& \quad+\Phi^{\dagger} \sum_{j=1}^{3}\left(\overline{\tilde{D}}_{L} Y_{j}^{D} P_{R} d_{j}+\overline{\tilde{U}}_{L} Y_{j}^{U} P_{R} u_{j}\right)+\text { H.c. } \tag{10}
\end{align*}
$$

then induce couplings of the SM quarks to the $Z^{\prime}$ once $\Phi$ acquires its VEV. Thus, integrating out the heavy vectorlike quarks gives rise to effective $Z^{\prime} \bar{d}_{i} d_{j}$ couplings $[45,46]$ of the form

$$
\begin{equation*}
g^{\prime}\left(\bar{d}_{i} \gamma^{\mu} P_{L} d_{j} Z_{\mu}^{\prime} \Gamma_{i j}^{d L}+\bar{d}_{i} \gamma^{\mu} P_{R} d_{j} Z_{\mu}^{\prime} \Gamma_{i j}^{d R}\right) \tag{11}
\end{equation*}
$$

with Hermitian matrices $\Gamma_{i j}^{d L}$ that are related to the vectorquark masses $m_{Q, D, U}$ and Yukawa couplings $Y^{Q, D, U}$ as follows [15]:
$\Gamma_{i j}^{d R} \simeq-\frac{v_{\Phi}^{2}}{2 m_{D}^{2}}\left(Y_{i}^{D} Y_{j}^{D *}\right), \quad \Gamma_{i j}^{d L} \simeq \frac{v_{\Phi}^{2}}{2 m_{Q}^{2}}\left(Y_{i}^{Q} Y_{j}^{Q *}\right)$,
which holds in the approximation $\left|\Gamma_{i j}^{q R / L}\right| \ll 1$. (Compared to Ref. [15], the vectorlike quarks also have Yukawa couplings $y_{\Psi_{1}}$ to the $\left(L_{\mu}-L_{\tau}\right)$-charged scalar doublet $\Psi_{1}$. This induces a small additional mass mixing among the heavy quarks and also a coupling to $h$ suppressed by $y_{\Psi_{1}} \cos (\alpha-\beta)$. We assume these couplings $y_{\Psi_{1}}$ to be small to avoid large contributions to $g g \rightarrow h$ and $h \rightarrow \gamma \gamma$ ).

Flavor observables.-We now recall the necessary formula in the region of interest (i.e., small $\theta_{R}$ ) considering only the processes giving the most relevant bounds on our model, i.e., $B_{s}-\bar{B}_{s}$ mixing, neutrino trident production, and $\tau \rightarrow 3 \mu$.
$h \rightarrow \mu \tau$ : The branching ratio for $h \rightarrow \mu \tau$ reads

$$
\begin{equation*}
\operatorname{Br}[h \rightarrow \mu \tau] \simeq \frac{m_{h}}{8 \pi \Gamma_{\mathrm{SM}}}\left|\Gamma_{\tau \mu}^{h}\right|^{2} \tag{13}
\end{equation*}
$$

where $\Gamma_{\text {SM }} \simeq 4.1 \mathrm{MeV}$ is the decay width in the SM for a 125 GeV Higgs [47] and $\Gamma_{\tau \mu}^{h}$ is defined in Eq. (7). Comparing this to Eq. (2), one sees that both $\sin \theta_{R} \neq 0$ and $\cos (\alpha-\beta) \neq 0$ are required to explain the CMS excess [35].

Lepton decays: While the Higgs contributions to $\tau \rightarrow \mu \mu \mu$ and $\tau \rightarrow \mu \gamma$ turn out to be very small in most regions of parameter space [35] due to the small lepton masses involved, the $Z^{\prime}$ contributions to $\tau \rightarrow 3 \mu$ can be sizable [42] and restrict $\theta_{R}^{2} / v_{\Phi}^{4}$. The branching ratio is given by

$$
\begin{equation*}
\operatorname{Br}[\tau \rightarrow 3 \mu] \simeq \frac{m_{\tau}^{5}}{512 \pi^{3} \Gamma_{\tau}} \frac{g^{\prime 4}}{m_{Z^{\prime}}^{4}} \sin ^{2}\left(2 \theta_{R}\right) \tag{14}
\end{equation*}
$$

which has to be compared to the current upper limit of $2.1 \times 10^{-8}$ at $90 \%$ C.L. [48] obtained by Belle. A combination with data from the $B A B A R$ Collaboration [49] gives an even stronger limit of $1.2 \times 10^{-8}$ at $90 \%$ C.L. [50], to be used in the following. For small $\theta_{R}$, the branching ratio for $\tau \rightarrow \mu \gamma$ is proportional to the same combination $\theta_{R}^{2} / v_{\Phi}^{4}$ but is highly suppressed by $2 \alpha / \pi$, and hence not as restrictive.
$B \rightarrow K^{*} \mu^{+} \mu^{-}$and $B \rightarrow K \mu^{+} \mu^{-} / B \rightarrow K e^{+} e^{-}:$Both $B \rightarrow K^{*} \mu^{+} \mu^{-}$and $R(K)$ are sensitive to the Wilson coefficients $C_{9}^{(1) \mu \mu}$ and $C_{10}^{(1) \mu \mu}$. For conventions see Refs. [8,19]. While in our model the contribution to $C_{10}$ is suppressed by $\sin \left(2 \theta_{R}\right)$ [or even $\sin \left(2 \theta_{L}\right)$ ], the Wilson coefficients $C_{9}^{\mu \mu}$ and $C_{9}^{\prime \mu \mu}$ with muons are generated (as well as $C_{9}^{\tau \tau}$ and the $\theta_{R}$ suppressed $\left.C_{9}^{\mu \tau}\right) . C_{9}^{\prime e e}$ is not affected, which naturally generates violations of lepton-flavor universality in $B \rightarrow K \mu^{+} \mu^{-} / B \rightarrow K e^{+} e^{-}$. We find

$$
\begin{equation*}
C_{9}^{(\prime) \mu \mu} \simeq \frac{g^{\prime 2}}{\sqrt{2} m_{Z^{\prime}}^{2}} \frac{\pi}{\alpha} \frac{1}{G_{F} V_{t b} V_{t s}^{*}} \Gamma_{23}^{d L(R)} \tag{15}
\end{equation*}
$$

where we set $\cos \left(2 \theta_{R}\right)=1$. As already noted in Refs. $[10,51], C_{9}^{\mu \mu}<0$ and $C_{9}^{\prime \mu \mu}=0$ give a good fit to data. Using the global fit of Ref. [8], we see that at the $2 \sigma(1 \sigma)$ level

$$
\begin{equation*}
-0.5(-0.8) \geq \operatorname{ReC}_{9}^{\mu \mu} \geq(-1.6)-2.0 \tag{16}
\end{equation*}
$$

Interestingly, the regions for $C_{9}^{\mu \mu}$ required by $R(K)$ and $B \rightarrow K^{*} \mu^{+} \mu^{-}$lie approximately in the same region. Furthermore, a good fit to the current data does not even require $C_{9}^{\prime \mu \mu}$ [8], so we neglect it in the following for simplicity. This can be achieved in the limit $m_{D} \gg m_{Q}$,
resulting in $\Gamma^{d L} \gg \Gamma^{d R}$. We will also assume our $C_{9}^{\mu \mu}$ to be real for simplicity. Note that our model predicts the decay $B \rightarrow K \mu \tau$ (recently discussed in Ref. [52]) to be suppressed by $\theta_{R}^{2}$ compared to $B \rightarrow K \mu \mu$, while $B \rightarrow K \mu e$ and $B \rightarrow K \tau e$ are forbidden.
$B_{s}-\bar{B}_{s}$ mixing: The interactions of $Z^{\prime}$ and $\Phi$ relevant for $B \rightarrow K \mu^{+} \mu^{-}$also contribute to $B_{s}-\bar{B}_{s}$ mixing [15]. For $m_{D} \gg m_{Q}$, we get

$$
\begin{equation*}
\frac{M_{12}}{M_{12}^{\mathrm{SM}}} \simeq 1+\frac{\left(\Gamma_{23}^{d L}\right)^{2}\left(1+\frac{1}{16 \pi^{2}} \frac{g^{\prime 2} m_{Q}^{2}}{m_{Z^{\prime}}^{2}}\right)}{\frac{g_{2}^{4}}{64 \pi^{2}} \frac{m_{Z^{\prime}}^{2}}{m_{W}^{2} g^{\prime 2}}\left(V_{t s}^{*} V_{t b}\right)^{2} S_{0}} \tag{17}
\end{equation*}
$$

We require the NP contribution to be less than $15 \%$ in order to satisfy the experimental bounds [15]. Because of the dominance of the vector quark $Q$, we can express $\Gamma_{23}^{d L}$ directly in terms of $C_{9}^{\mu \mu}$ from Eq. (15) and find the upper bounds
$m_{Z^{\prime}} / g^{\prime}<3.2 \mathrm{TeV} /\left|C_{9}^{\mu \mu}\right|, \quad m_{Q}<41 \mathrm{TeV} /\left|C_{9}^{\mu \mu}\right|$.
Combining Eq. (18) with Eq. (16) then gives an upper bound of $m_{Z^{\prime}} / g^{\prime}<4 \mathrm{TeV}(6.5 \mathrm{TeV})$ at $1 \sigma(2 \sigma)$.

Neutrino trident production: The most stringent bound on flavor-diagonal $Z^{\prime}$ couplings to muons arises from NTP $\nu_{\mu} N \rightarrow \nu_{\mu} N \mu^{+} \mu^{-}[15,53]:$

$$
\begin{equation*}
\frac{\sigma_{\mathrm{NTP}}}{\sigma_{\mathrm{NTP}}^{\mathrm{SM}}} \simeq \frac{1+\left(1+4 s_{W}^{2}+8 \frac{g^{\prime}}{M_{Z^{\prime}}^{2}} \frac{m_{W}^{2}}{g_{2}^{2}}\right)^{2}}{1+\left(1+4 s_{W}^{2}\right)^{2}} . \tag{19}
\end{equation*}
$$

Since 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 [54], we get roughly $m_{Z^{\prime}} / g^{\prime} \gtrsim 550 \mathrm{GeV}$ at $95 \%$ C.L. Compared to $\tau \rightarrow \mu \mu \mu$, 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 $m_{Z^{\prime}}>m_{Z}$, the LHC constraints from the process $p p \rightarrow \mu \mu Z^{\prime} \rightarrow 4 \mu$ (or $3 \mu$ plus missing energy) [55] are currently weaker than NTP [15], but will become competitive with higher luminosities [56-58].

Phenomenological analysis: Concerning the phenomenological consequences of our model, let us first consider the implications of $h \rightarrow \mu \tau$. In the left plot of Fig. 1, we show the regions in the $\cos (\alpha-\beta)-\sin \left(\theta_{R}\right)$ plane which can explain $h \rightarrow \mu \tau$ at the $1 \sigma$ and $2 \sigma$ levels 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 $\tau \rightarrow 3 \mu$ and $\tau \rightarrow \mu \gamma$ are typically negligible [35]. As a side effect, the $h \rightarrow \mu \tau$ rate also implies a change in the $h \rightarrow \tau \tau$ rate, although this is negligible in regions with small $\theta_{R}$. In addition, we show the regions compatible with $\tau \rightarrow 3 \mu$ for various values of $m_{Z^{\prime}} / g^{\prime}$. Note that $g^{\prime} \lesssim 0.3$ in order to avoid a Landau pole below the Planck scale. In summary, small values of $\theta_{R}$ can explain the CMS $h \rightarrow \mu \tau$ excess for moderate to large values of $\tan \beta$ for $\cos (\alpha-\beta) \simeq 0.1$.


FIG. 1 (color online). Left panel: Allowed regions in the $\cos (\alpha-\beta)-\sin \left(\theta_{R}\right)$ plane. The blue (light blue) region corresponds to the $1 \sigma$ ( $2 \sigma$ ) region of the CMS measurement of $h \rightarrow \mu \tau$ for $\tan \beta=50$; yellow stands for $\tan \beta=10$. The (dashed) red contours mark deviations of $h \rightarrow \tau \tau$ 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 $\tau \rightarrow 3 \mu$ via $Z^{\prime}$ exchange. Right panel: Allowed regions in the $\Gamma_{23}^{d L}-m_{Z^{\prime}} / g^{\prime}$ plane from $B \rightarrow K^{*} \mu^{+} \mu^{-}$and $R(K)$ (yellow) and $B_{s}$ mixing (blue). For $B_{s}$ mixing, (light) blue corresponds to ( $m_{Q}=15 m_{Z^{\prime}} / g^{\prime}$ ) $m_{Q}=m_{Z^{\prime}} / g^{\prime}$. The horizontal lines denote the lower bounds on $m_{Z^{\prime}} / g^{\prime}$ from $\tau \rightarrow 3 \mu$ for $\sin \left(\theta_{R}\right)=0.005,0.02,0.05$. The gray region is excluded by neutrino trident production (NTP).

In the right plot of Fig. 1, we examine which regions in parameter space can account for $B \rightarrow K^{*} \mu^{+} \mu^{-}$taking into account the constraints from $B_{s}-\bar{B}_{s}$ mixing. Since we focus on the limit $M_{D} \rightarrow \infty$ (i.e., $C_{9}^{\prime} \rightarrow 0$ ), we find that unless $\Gamma_{23}^{d L}$ is rather large, $B \rightarrow K^{*} \mu^{+} \mu^{-}$can be explained without violating bounds from $B_{s}-\bar{B}_{s}$. Only a very small $\Gamma_{23}^{d L}-$ independent region is excluded by NTP. In addition, bounds from $\tau \rightarrow 3 \mu$ depending on $\sin \left(\theta_{R}\right)$ can be obtained.

Concerning $\tau \rightarrow 3 \mu$, future sensitivities down to $\operatorname{Br}[\tau \rightarrow 3 \mu] \simeq 10^{-9}$ seem feasible [61] and will cut deep into our parameter space (see Fig. 2). Using the $1 \sigma$ limits on $h \rightarrow \mu \tau$ to fix $\theta_{R}$ and $B_{s}$ mixing with $C_{9}$ to fix $m_{Z^{\prime}} / g^{\prime}$-as well as the LHC limit $|\cos (\alpha-\beta)|<0.4$-we can obtain a lower limit on the rate $\tau \rightarrow 3 \mu$,

$$
\begin{equation*}
\operatorname{Br}[\tau \rightarrow 3 \mu] \gtrsim 3.8 \times 10^{-8}(10 / \tan \beta)^{2} \tag{20}
\end{equation*}
$$

which implies $\tan \beta \gtrsim 18$ with current data [50] 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 $h \rightarrow \mu \tau, B \rightarrow K^{*} \mu^{+} \mu^{-}$and $R(K)$.

Finally, we remark that a $Z-Z^{\prime}$ mixing angle $\theta_{Z Z^{\prime}}$ [45] is induced by the VEV of $\Psi_{1}$ [35],
$\left|g^{\prime} \theta_{Z Z^{\prime}}\right| \simeq \frac{g_{1} v^{2} \cos ^{2} \beta}{m_{Z^{\prime}}^{2} / g^{\prime 2}} \simeq 10^{-4}\left(\frac{20}{\tan \beta}\right)^{2}\left(\frac{\mathrm{TeV}}{m_{Z^{\prime}} / g^{\prime}}\right)^{2}$,
which leads to small shifts in the vector couplings of $Z$ to muons and tauons,

$$
\begin{equation*}
g_{V}^{Z}(\mu \mu, \tau \tau) \simeq-1 / 2+2 s_{W}^{2} \pm g^{\prime} \theta_{Z Z^{\prime}} /\left(g / c_{W}\right) \tag{22}
\end{equation*}
$$

and thus ultimately to lepton nonuniversality [43]. For the values of interest to our study (see Fig. 2), and in the limit
$m_{Z} \ll m_{Z^{\prime}}$, the shift is automatically small enough to satisfy experimental bounds and leads to tiny branching ratios $Z \rightarrow \mu \tau$ below $10^{-8}$ (for $\theta_{R}<0.1$ ). Note that the couplings to electrons and quarks remain unaffected. For $m_{Z^{\prime}} \gg m_{Z}$, the $\rho$ parameter is enhanced by [45]

$$
\begin{equation*}
\rho-1 \simeq 1.2 \times 10^{-4}\left(\frac{\theta_{Z Z^{\prime}}}{10^{-3}}\right)^{2}\left(\frac{m_{Z^{\prime}}}{\mathrm{TeV}}\right)^{2} \tag{23}
\end{equation*}
$$



FIG. 2 (color online). Allowed regions in the $m_{Z^{\prime}} / g^{\prime}-\sin \left(\theta_{R}\right)$ plane: The horizontal stripes correspond to $h \rightarrow \mu \tau(1 \sigma)$ for $\tan \beta=85,50,25$ and $\cos (\alpha-\beta)=0.2$; (light) blue stands for (future) $\tau \rightarrow 3 \mu$ limits at $90 \%$ C.L. The gray regions are excluded by NTP or $B_{s}-\bar{B}_{s}$ mixing in combination with the $1 \sigma$ range for $C_{9}$ [see Eq. (18)].
and can therefore be compatible with electroweak precision data $\left(\rho-1<9 \times 10^{-4}\right.$ at $2 \sigma$ [62]).

Conclusions.-In this Letter, we showed for the first time that all three LHC anomalies in the flavor sector can be explained within a single well-motivated model: a 2 HDM with a gauged $L_{\mu}-L_{\tau}$ symmetry and effective $Z^{\prime} \bar{s} b$ couplings induced by heavy vectorlike quarks. Except for the $\tau-\mu$ couplings, the Higgs sector resembles the type-I 2HDM. Therefore, the constraints from $h$ decays or LHC searches for $A^{0} \rightarrow \tau^{+} \tau^{-}$are rather weak, and $h \rightarrow \mu \tau$ can be easily explained in a wide parameter space. The model can also account for the deviations from the SM in $B \rightarrow K^{*} \mu^{+} \mu^{-}$ and naturally leads to the right amount of lepton-flavoruniversality violating effects in $R(K)$. Because of the small values of the $\tau-\mu$ mixing angle $\theta_{R}$, sufficient to account for $h \rightarrow \mu \tau$, the $Z^{\prime}$ contributions to $\tau \rightarrow 3 \mu$ are not in conflict with present bounds for large $\tan \beta$ in wide ranges of parameter space. Interestingly, $B \rightarrow K^{*} \mu^{+} \mu^{-}$and $R(K)$ combined with $B_{s}-\bar{B}_{s}$ put an upper limit on $m_{Z^{\prime}} / g^{\prime}$ resulting in a lower limit on $\tau \rightarrow 3 \mu$ if $\operatorname{Br}[h \rightarrow \mu \tau] \neq 0$ : 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 $L_{\mu}-L_{\tau}$ breaking scale further implies the masses of the $Z^{\prime}$ and the right-handed neutrinos to be at the TeV scale, potentially testable at the LHC with interesting additional consequences for LFV observables.
A. C. is supported by a Marie Curie Intra-European Fellowship of the European Community's 7th Framework Programme under Contract No. PIEF-GA-2012-326948. G. D. acknowledges partial support by MIUR under Project No. 2010YJ2NYW. The work of J. H. 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 $C_{9}^{\mu \mu}$ and $C_{9}^{\prime \mu \mu}$.

Note added.-Recently, CMS released its final analysis of the $h \rightarrow \mu \tau$ search as a preprint [63], resulting in slightly changed values of $\operatorname{Br}[h \rightarrow \mu \tau]=\left(0.84_{-0.37}^{+0.39}\right) \%$, which, however, have no impact on our study.
[1] G. Aad et al. (ATLAS Collaboration), Phys. Lett. B 716, 1 (2012).
[2] S. Chatrchyan et al. (CMS Collaboration), Phys. Lett. B 716, 30 (2012).
[3] U. Egede, T. Hurth, J. Matias, M. Ramon, and W. Reece, J. High Energy Phys. 11 (2008) 032.
[4] W. Altmannshofer, P. Ball, A. Bharucha, A. J. Buras, D. M. Straub, and M. Wick, J. High Energy Phys. 01 (2009) 019.
[5] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, J. High Energy Phys. 05 (2013) 137.
[6] R. Aaij et al. (LHCb Collaboration), Phys. Rev. Lett. 111, 191801 (2013).
[7] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, J. High Energy Phys. 12 (2014) 125.
[8] W. Altmannshofer and D. M. Straub, arXiv:1411.3161.
[9] S. Jäger and J. Martin Camalich, arXiv:1412.3183.
[10] S. Descotes-Genon, J. Matias, and J. Virto, Phys. Rev. D 88, 074002 (2013).
[11] W. Altmannshofer and D. M. Straub, Eur. Phys. J. C 73, 2646 (2013).
[12] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, Phys. Rev. Lett. 112, 212003 (2014).
[13] R. Gauld, F. Goertz, and U. Haisch, Phys. Rev. D 89, 015005 (2014).
[14] A. J. Buras, F. De Fazio, and J. Girrbach, J. High Energy Phys. 02 (2014) 112.
[15] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Phys. Rev. D 89, 095033 (2014).
[16] R. Aaij et al. (LHCb Collaboration), Phys. Rev. Lett. 113, 151601 (2014).
[17] C. Bobeth, G. Hiller, and G. Piranishvili, J. High Energy Phys. 12 (2007) 040.
[18] R. Alonso, B. Grinstein, and J. Martin Camalich, Phys. Rev. Lett. 113, 241802 (2014).
[19] G. Hiller and M. Schmaltz, Phys. Rev. D 90, 054014 (2014).
[20] D. Ghosh, M. Nardecchia, and S. Renner, J. High Energy Phys. 12 (2014) 131.
[21] T. Hurth, F. Mahmoudi, and S. Neshatpour, J. High Energy Phys. 12 (2014) 053.
[22] CMS Collaboration, Report No. CMS-PAS-HIG-14-005.
[23] R. Harnik, J. Kopp, and J. Zupan, J. High Energy Phys. 03 (2013) 026.
[24] G. Blankenburg, J. Ellis, and G. Isidori, Phys. Lett. B 712, 386 (2012).
[25] S. Davidson and P. Verdier, Phys. Rev. D 86, 111701 (2012).
[26] A. Arhrib, Y. Cheng, and O. C. Kong, Phys. Rev. D 87, 015025 (2013).
[27] A. Arhrib, Y. Cheng, and O. C. Kong, Europhys. Lett. 101, 31003 (2013).
[28] J. Kopp and M. Nardecchia, J. High Energy Phys. 10 (2014) 156.
[29] A. Falkowski, D. M. Straub, and A. Vicente, J. High Energy Phys. 05 (2014) 092.
[30] A. Dery, A. Efrati, Y. Nir, Y. Soreq, and V. Susi, Phys. Rev. D 90, 115022 (2014).
[31] M. D. Campos, A.E.C. Hernández, H. Päs, and E. Schumacher, arXiv:1408.1652.
[32] A. Celis, V. Cirigliano, and E. Passemar, arXiv:1409.4439.
[33] D. Aristizabal Sierra and A. Vicente, Phys. Rev. D 90, 115004 (2014).
[34] C.-J. Lee and J. Tandean, arXiv:1410.6803.
[35] J. Heeck, M. Holthausen, W. Rodejohann, and Y. Shimizu, arXiv:1412.3671.
[36] X. He, G. C. Joshi, H. Lew, and R. Volkas, Phys. Rev. D 43, R22 (1991).
[37] R. Foot, Mod. Phys. Lett. A 06, 527 (1991).
[38] X.-G. He, G. C. Joshi, H. Lew, and R. Volkas, Phys. Rev. D 44, 2118 (1991).
[39] P. Binetruy, S. Lavignac, S. T. Petcov, and P. Ramond, Nucl. Phys. B496, 3 (1997).
[40] N. F. Bell and R. R. Volkas, Phys. Rev. D 63, 013006 (2000).
[41] S. Choubey and W. Rodejohann, Eur. Phys. J. C 40, 259 (2005).
[42] G. Dutta, A. S. Joshipura, and K. Vijaykumar, Phys. Rev. D 50, 2109 (1994).
[43] J. Heeck and W. Rodejohann, Phys. Rev. D 84, 075007 (2011).
[44] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, Phys. Rep. 516, 1 (2012).
[45] P. Langacker, Rev. Mod. Phys. 81, 1199 (2009).
[46] A. J. Buras, F. De Fazio, and J. Girrbach, J. High Energy Phys. 02 (2013) 116.
[47] S. Dittmaier, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka et al., arXiv:1201.3084.
[48] B. Aubert et al. (BABAR Collaboration), Phys. Rev. Lett. 104, 021802 (2010).
[49] J. Lees et al. (BABAR Collaboration), Phys. Rev. D 81, 111101 (2010).
[50] Y. Amhis et al. (Heavy Flavor Averaging Group), arXiv:1412.7515.
[51] S. Descotes-Genon, J. Matias, and J. Virto, Proc. Sci., EPSHEP2013, 361 (2013).
[52] S. L. Glashow, D. Guadagnoli, and K. Lane, Phys. Rev. Lett. 114, 091801 (2015).
[53] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Phys. Rev. Lett. 113, 091801 (2014).
[54] S. Mishra et al. (CCFR Collaboration), Phys. Rev. Lett. 66, 3117 (1991).
[55] E. Ma, D. Roy, and S. Roy, Phys. Lett. B 525, 101 (2002).
[56] K. Harigaya, T. Igari, M. M. Nojiri, M. Takeuchi, and K. Tobe, J. High Energy Phys. 03 (2014) 105.
[57] N. F. Bell, Y. Cai, R. K. Leane, and A. D. Medina, Phys. Rev. D 90, 035027 (2014).
[58] F. del Aguila, M. Chala, J. Santiago, and Y. Yamamoto, arXiv:1411.7394.
[59] B. Dumont, J. F. Gunion, Y. Jiang, and S. Kraml, Phys. Rev. D 90, 035021 (2014).
[60] B. Dumont, J. F. Gunion, Y. Jiang, and S. Kraml, arXiv: 1409.4088.
[61] T. Aushev, W. Bartel, A. Bondar, J. Brodzicka, T. Browder et al., arXiv:1002.5012.
[62] K. Olive et al. (Particle Data Group), Chin. Phys. C38, 090001 (2014).
[63] V. Khachatryan et al. (CMS Collaboration), arXiv: 1502.07400 .


[^0]:    Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

