Testing lepton flavor universality in $\Upsilon(4S)$ decays

Sébastien Descotes-Genon[*] and Martín Novoa-Brunet[†]

Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France

Svjetlana Fajfer[‡] and Jernej F. Kamenik[§]

Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
and Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia

We propose a novel method to probe the persistent hints of lepton flavor universality violation observed in semileptonic B decays. Relying on the specific properties of the Belle II experiment, it consists in comparing the inclusive rates of $\Upsilon(4S) \to e^{\pm}\mu^{\mp}X$, $\Upsilon(4S) \to \mu^{\pm}\tau_{\text{had}}^{\mp}X$ and $\Upsilon(4S) \to e^{\pm}\tau_{\text{had}}^{\mp}X$. We show that such a measurement can be directly related to the ratio $R(X)_{\tau\ell} \equiv \Gamma(B \to X\tau\nu)/\Gamma(B \to X\ell\nu)$ ($\ell = e$ or μ), once appropriate experimental cuts are applied to suppress the effects of neutral B mixing and leptons emitted through charm or tau decays. Such a measurement would thus constitute an additional and potentially competitive probe of lepton flavor universality in $b \to c\ell\nu$ transitions, complementary to existing exclusive measurements, accessible in the Belle II environment.

I. INTRODUCTION

The universality of electroweak interactions of leptons is a direct consequence of the gauge structure of the Standard Model (SM). Within the SM, lepton flavor universality (LFU) is violated solely due to the different masses of the charged leptons. In the past several years a number of experiments observed the violation of LFU in the $b \to c\tau\nu_\tau$ decays, represented by the ratios $R_{D^{(*)}} = \mathcal{B}(B \to D^{(*)}\tau\nu_\tau)/\mathcal{B}(B \to D^{(*)}l\nu_l)$ [1–10]. Heavy flavor averaging group [1] currently reports $R_{D^*} = 0.295 \pm 0.011 \pm 0.008$ and $R_D = 0.340 \pm 0.027 \pm 0.013$, which differ from the SM predictions $R_{D^*}^{\text{SM}} = 0.258 \pm 0.0005$ and $R_D^{\text{SM}} = 0.299 \pm 0.003$ on the level of $\sim 3\sigma$. A related 1.8σ tension has also been reported in the $B_c \to J/\psi\tau\nu$ decay branching ratio [11].

In light of these intriguing results, several associated tests of LFU have been proposed in processes involving third generation quarks and leptons, including the rare B meson decays $B_c \to \tau\nu$ [12,13], $B \to K^{(*)}\nu\bar{\nu}$ [14], $B \to K^{(*)}\tau^+\tau^-$ [15,16], as well as high-p_T mono-tau [17] or tau-pair production [18] at the LHC. Finally, following the proposal in Ref. [19], the *BABAR* collaboration has recently measured the LFU ratio in $\Upsilon(3S)$ decays [20]

$$R_{\tau/\mu}^{\Upsilon(3S)} \equiv \frac{\mathcal{B}(\Upsilon(3S) \to \tau^+\tau^-)}{\mathcal{B}(\Upsilon(3S) \to \mu^+\mu^-)} = 0.966 \pm 0.008 \pm 0.014, \tag{1}$$

where the first (second) uncertainty estimate is due to statistics (systematics). The measured value is within 1.8σ of the SM prediction $[R_{\tau/\mu}^{\Upsilon(3S)}]_{\text{SM}} = 0.9948(1)$ [19]. This measurement is probing the $R_{D^{(*)}}$ LFU anomaly through $b\bar{b} \to \ell^+\ell^-$ transitions [21].

In the following we propose a related but potentially more direct test through inclusive dileptonic $\Upsilon(4S)$ decays by defining

$$\mathcal{B}_{\ell\ell'}^{\Upsilon(4S)} \equiv \mathcal{B}(\Upsilon(4S) \to \ell^+\ell'^-X), \tag{2}$$

as the inclusive dileptonic branching fraction for $\Upsilon(4S)$ decays to a pair of opposite charged leptons of different flavors, where X denotes all the other (hadronic) activity and missing momentum in the event.

This fully inclusive measurement exploits several key capabilities of the Belle II experiment as well as some specific features of $\Upsilon(4S)$ and b hadron decays. On the experimental side, the excellent beam energy calibration of SuperKEKB can ensure that the $\Upsilon(4S)$ resonance is produced on shell even if its invariant mass is not reconstructed explicitly from the final state. This also allows the nonresonant background to be well estimated

[*]sebastien.descotes-genon@ijclab.in2p3.fr
[†]martin.novoa@ijclab.in2p3.fr
[‡]svjetlana.fajfer@ijs.si
[§]jernej.kamenik@cern.ch

from sideband measurements. On the theory side, this inclusive decay is almost entirely saturated by decays into $B\bar{B}$ final states. Moreover, one can analyse the production of the leptons either from an initial b-quark decay or from subsequent parts of the decay chain in detail. All in all, ratios of the form

$$R_{\ell\ell'}^{\Upsilon(4S)} \equiv \frac{\mathcal{B}_{\ell''\ell}^{\Upsilon(4S)}}{\mathcal{B}_{\ell''\ell'}^{\Upsilon(4S)}}, \tag{3}$$

where ℓ, ℓ', ℓ'' are three different flavors of leptons e, μ, τ, provide a very interesting ground to probe lepton flavor universality with an inclusive measurement at Belle II,[1] complementary to exclusive measurements accessible to both Belle II and LHCb experiments. In particular, under suitable experimental conditions one can relate

$$R_{\tau\ell}^{\Upsilon(4S)} = R(X)_{\tau\ell} + \dots, \tag{4}$$

where $\ell = e, \mu$ and $R(X)_{\tau\ell} \equiv \Gamma(B \to X\tau\nu)/\Gamma(B \to X\ell\nu)$ is the inclusive B decay LFU ratio, which can be precisely computed in the SM as $\mathcal{R}(X)_{\tau\ell} = 0.223(4)$ [23]. The dots denote corrections due to neutral B meson mixing effects and charm pollution. In the following we discuss both effects and estimate the accuracy with which this ratio can be measured and compared with the SM expectation, in order to extract potential violations of LFU.

II. ANALYSIS OF THE DECAY CHAIN

A. $\Upsilon(4S)$ decay

The $\Upsilon(4S)$ resonance overwhelmingly decays into $B\bar{B}$ final states. In particular, there is an experimental bound $\mathcal{B}(\Upsilon(4S) \to B\bar{B}) > 0.96$ [22] but $\mathcal{B}(\Upsilon(4S) \to B\bar{B})$ could actually be even much closer to one. Indeed the dominant non-$B\bar{B}$ final states are expected to consist in light hadrons mediated by $\Upsilon(4S) \to 3g$ decays as well as decays to lighter bottomonium states, in particular, $\Upsilon(4S) \to (\Upsilon(n'S), h_b(n''P))(\pi\pi, \eta, \eta')$ with $n',''< 4$. Experimental measurements already exist for the latter contributions, in particular $\mathcal{B}(\Upsilon(4S) \to \Upsilon(1S) + X) < 4 \times 10^{-3}$, $\mathcal{B}(\Upsilon(4S) \to \Upsilon(2S)\pi^+\pi^-) = 8.2(8) \times 10^{-5}$, $\mathcal{B}(\Upsilon(4S) \to h_b(1P)\eta) = 2.18(21) \times 10^{-3}$ with other known modes below the 10^{-4} level [24]. In total we thus estimate $\mathcal{B}(\Upsilon(4S) \to \text{bottomonia}) < 7 \times 10^{-3}$. On the other hand, the $\Upsilon(4S) \to 3g$ decay width can be estimated in nonrelativistic QCD [25]. At leading order (LO) in velocity and quantum chromodynamics (QCD) expansion, it is given by [26]

$$\Gamma(\Upsilon \to 3g) = 0.0716 \frac{\alpha_s^3 \langle O_1 \rangle_\Upsilon}{m_b^2}. \tag{5}$$

Both leading velocity and QCD corrections are of negative sign and thus serve to decrease the above LO estimate [26]. We thus take it as a conservative upper bound. Using $\alpha_s = 0.22$, $m_b = 4.6$ and the upper estimate on $\langle O_1 \rangle_{\Upsilon(4S)} \lesssim \langle O_1 \rangle_{\Upsilon(3S)} = 1.279$ GeV³ [27] we obtain $\mathcal{B}(\Upsilon(4S) \to 3g) \lesssim 2 \times 10^{-3}$. In total we thus estimate that $\mathcal{B}(\Upsilon(4S) \to B\bar{B}) \gtrsim 0.99$.

B. Lepton production

Having established that $\Upsilon(4S)$ decays almost only into pairs of B mesons, we consider their subsequent decays inclusively, focusing on final states containing leptons $\ell^{(I)} = e, \mu, \tau$ in the final state. We can differentiate between several measurable inclusive dilepton signatures such as

$$\Upsilon(4S) \to e^\pm\mu^\mp X, \qquad \Upsilon(4S) \to \mu^\pm\tau_{\text{had}}^\mp X,$$
$$\Upsilon(4S) \to e^\pm\tau_{\text{had}}^\mp X, \tag{6}$$

where τ_{had} denotes a τ lepton reconstructed from its hadronic decays (e.g., $\tau \to 3\pi\nu$).[2] We will thus define

$$R_{\tau_{\text{had}}e}^{\Upsilon(4S)} \equiv \frac{\mathcal{B}_{\mu\tau_{\text{had}}}^{\Upsilon(4S)}}{\mathcal{B}_{\mu e}^{\Upsilon(4S)}} \quad \text{and} \quad R_{\tau_{\text{had}}\mu}^{\Upsilon(4S)} \equiv \frac{\mathcal{B}_{e\tau_{\text{had}}}^{\Upsilon(4S)}}{\mathcal{B}_{\mu e}^{\Upsilon(4S)}}. \tag{7}$$

To relate these ratios to inclusive B-decay LFU ratios, we need to isolate contributions where each of the two leptons is produced in a separate B-meson decay and suppress backgrounds where one or both leptons do not originate from a direct semileptonic B decay. Requiring different opposite-sign lepton flavor final states removes such contributions from $\Upsilon(4S) \to X + ((\bar{b}b) \to \ell^+\ell^-)$, $b \to q((c\bar{c}) \to \ell^+\ell^-)$ as well as from rare (semileptonic) flavor changing neutral current (FCNC) B and charm decays.[3]

This approach is however not effective against contamination from $b \to q(c \to q'\ell^+\nu)(\bar{c} \to q''\ell'^-\nu)$ and $b \to (c \to q\ell^+\nu)\ell'^-\nu)$ transitions, which we will address in Sec. II E.

For the moment we assume that each of the two different lepton tags originates from a separate B-meson decay pattern. We will focus on $R_{\tau_{\text{had}}\mu}^{\Upsilon(4S)}$ for the time being, but a very similar analysis can be performed for $R_{\tau_{\text{had}}e}^{\Upsilon(4S)}$ swapping muons and electrons in the discussion. The single hadronic tau can be produced in the quark-level transition chains

---

[1] In principle, a similar test could be envisioned using $\psi(3770)$ at BESIII since $\mathcal{B}(\psi(3770) \to D\bar{D}) \sim 93\%$ [22].

[2] These hadronic tau lepton decays need to be efficiently disentangled from backgrounds like hadronic B decays involving three or more charged pions.

[3] The exceptions with $\ell^\pm = \tau^\pm$ where one of the taus decays leptonically and the other hadronically, leading to a final state with a hadronic tau and a lighter lepton, turn out to be numerically negligible as they are suppressed by small $\Upsilon(4S) \to X + $ bottomonium, $B \to X + $ charmonium [24], and $B \to X\tau^+\tau^-$ [15] branching ratios, respectively.

$$b \to q\tau\nu,$$
$$b \to q\bar{q}'(c \to q''\tau\nu). \tag{8}$$

On the other hand a single muon (or equivalently electron) can originate from

$$b \to q\mu\nu,$$
$$b \to q\bar{q}'(c \to q''\mu\nu),$$
$$b \to q(\tau \to \mu\nu\nu)\nu,$$
$$b \to q\bar{q}'(c \to q''(\tau \to \mu\nu\nu)\nu). \tag{9}$$

Inclusive semileptonic b hadron decays (i.e., $b \to q\ell\nu$) are well under theoretical control and thus the associated rates can be well predicted, including possible effects of LFU violation [23]. The same cannot necessarily be said for inclusive semileptonic charm decays [28], which thus represent a challenging background. One could imagine that the charge of the leptons could help us to disentangle the origin of the lepton, either from a b or from a c-quark. However, one should take into account that in approximately half of the cases, the $\Upsilon(4S)$ decays into neutral B mesons, which can oscillate and spoil the identification between the charge of initial quark and that of the lepton. We discuss strategies of how to mitigate this effect next.

C. Mixing effects

We first define the amplitudes $A(B^0 \to \ell^-X), A(B^0 \to \ell^+X), A(\bar{B}^0 \to \ell^+X), A(\bar{B}^0 \to \ell^-X)$ embedding the complete meson decay chains (for instance it may contain $B \to D\pi$ followed by $D \to \ell X$), so that the lepton is not necessarily produced by the decay of the b-quark. However, it is not produced by the decay of the light quark in the B, which means that in the isospin limit, we have equalities of the type

$$A(B^0 \to \ell^-X) = A(B^+ \to \ell^-X) = A_{\ell^-}$$
$$A(\bar{B}^0 \to \ell^-X) = A(B^- \to \ell^-X) = \bar{A}_{\ell^-} \quad \dots, \tag{10}$$

where the presence/absence of the bar indicates the charge of the b-quark inside the B-meson and the subscript denotes the charge and flavor of the lepton.

If we look for $\Upsilon(4S) \to \ell_1\ell_2X$ (with one and two being different, either by flavor or charge) through an intermediate $B^0\bar{B}^0$ state, we can use the description introduced for the study of charge conjugation parity symmetry (CP) violation from the production of an intricated B-meson pair (sec 1.2.3 in Ref. [29]), leading to the time-dependent rate where one of the two B mesons decay into a state containing ℓ_1 at a time t_1 and the other one into a state containing ℓ_2 at a time t_2, leading to

$$R(t_1, t_2) = Ce^{-\Gamma(t_1+t_2)}[\mathcal{I} - \cos(\Delta m(t_1 - t_2))\mathcal{C}$$
$$+ 2\sin(\Delta m(t_1 - t_2))\mathcal{S}], \tag{11}$$

where C is a normalization coming from angular integration, Δm is the difference of mass between the two mass eigenstates, Γ is their average width, the approximations $|q/p| = 1$ and $\Delta\Gamma = 0$ have been used, and we have

$$\mathcal{I} = \left[(|A_1|^2 + |\bar{A}_1|^2)(|A_2|^2 + |\bar{A}_2|^2) \right.$$
$$\left. - 4\text{Re}\left(\frac{q}{p}A_1^*\bar{A}_1\right)\text{Re}\left(\frac{q}{p}A_2^*\bar{A}_2\right) \right], \tag{12}$$

$$\mathcal{C} = \left[(|A_1|^2 - |\bar{A}_1|^2)(|A_2|^2 - |\bar{A}_2|^2) \right.$$
$$\left. + 4\text{Im}\left(\frac{q}{p}A_1^*\bar{A}_1\right)\text{Im}\left(\frac{q}{p}A_2^*\bar{A}_2\right) \right], \tag{13}$$

$$\mathcal{S} = \left[\text{Im}\left(\frac{q}{p}A_1^*\bar{A}_1\right)(|A_2|^2 - |\bar{A}_2|^2) \right.$$
$$\left. + (|A_1|^2 - |\bar{A}_1|^2)\text{Im}\left(\frac{q}{p}A_2^*\bar{A}_2\right) \right]. \tag{14}$$

In order to prevent too large effects from mixing, one could consider cutting too large time differences $|t_1 - t_2|$, so that there has not been enough time for the evolution to take place. Cutting at $|t_1 - t_2| = \alpha/\Delta m$ (where α is the cut parameter) leads to

$$R_{B^0\bar{B}^0}^\alpha \equiv \int_{-\alpha/\Delta m}^{\alpha/\Delta m} R(t')\text{d}t'$$
$$= \frac{2C}{\Gamma^2}\left[(1 - e^{-\frac{\alpha}{x}})\mathcal{I} - \frac{1 - e^{-\frac{\alpha}{x}}(\cos\alpha - x\sin\alpha)}{1 + x^2}\mathcal{C} \right], \tag{15}$$

where $x = \Delta m/\Gamma \simeq 0.769$ [1]. In the case of B^+B^-, where no mixing is involved, we have

$$R_{B^+B^-}^\alpha \equiv \int_{-\alpha/\Delta m}^{\alpha/\Delta m} R(t')|_{\Delta m=0}\text{d}t' = \frac{2C}{\Gamma^2}(1 - e^{-\frac{\alpha}{x}})(\mathcal{I} - \mathcal{C}). \tag{16}$$

Denoting the result without cut in the time difference as $R_{BB} \equiv R_{BB}^\infty$, we see then that the effect of mixing corresponds to

$$R_{B^0\bar{B}^0}^\alpha - R_{B^+B^-}^\alpha = \left(1 - \frac{e^{-\frac{\alpha}{x}}(1 + x^2 - \cos\alpha + x\sin\alpha)}{x^2} \right)$$
$$\times (R_{B^0\bar{B}^0} - R_{B^+B^-}), \tag{17}$$

whereas we have

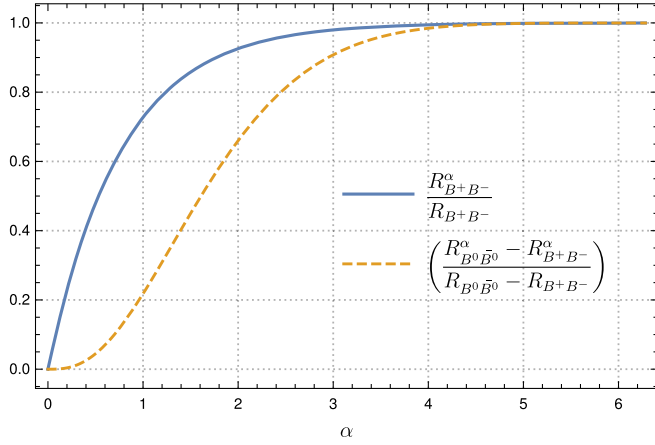$$R_{B^+B^-}^\alpha = (1 - e^{-\frac{\alpha}{x}})R_{B^+B^-}, \tag{18}$$

FIG. 1. Relative branching fraction (blue solid) and mixing effect (yellow dashed) as a function of the cut parameter α.

In Fig. 1, we illustrate the impact of the cut on the branching fraction of the decay into B^+B^- and on the mixing effect. We see that this cut can efficiently suppress the impact of mixing while keeping a large fraction of the B^+B^- signal.

We can study the impact of this cut on the $\Upsilon(4S)$ decay rate. Introducing $\rho = \mathcal{B}(\Upsilon(4S) \to B^+B^-) = 0.514 \pm 0.006$ [1], we have $\mathcal{B}(\Upsilon(4S) \to B^0\bar{B}^0) = 0.486 \pm 0.006 = 1 - \rho - \epsilon$ with $\epsilon < 0.01$ according to our estimates. We have then the total rate $R = \mathcal{B}(\Upsilon(4S) \to BB)$

$$
\begin{aligned}
R &= \rho R_{B^+B^-} + (1 - \rho - \epsilon)R_{B^0\bar{B}^0} \\
&= (1 - \epsilon)R_{B^+B^-} + (1 - \rho - \epsilon)(R_{B^0\bar{B}^0} - R_{B^+B^-}),
\end{aligned} \tag{19}
$$

where the first term corresponds to the rate without mixing (in the isospin limit) and the second term to the contamination due to mixing. Cutting $|t_1 - t_2| \le \alpha/\Delta m$, we have

$$
R^\alpha = (1 - \epsilon)R^\alpha_{B^+B^-} + (1 - \rho - \epsilon)(R^\alpha_{B^0\bar{B}^0} - R^\alpha_{B^+B^-}) \tag{20}
$$

$$
\begin{aligned}
&= (1 - \epsilon)(1 - e^{-\frac{\alpha}{x}})R_{B^+B^-} \\
&\quad + (1 - \rho - \epsilon)\left(1 - \frac{e^{-\frac{\alpha}{x}}(1 + x^2 - \cos\alpha + x\sin\alpha)}{x^2}\right) \\
&\quad \times (R_{B^0\bar{B}^0} - R_{B^+B^-}).
\end{aligned} \tag{21}
$$

The first term in Eq. (21) goes like $O(\alpha)$ whereas the second term goes like $O(\alpha^3)$. Moreover, $(R_{B^0\bar{B}^0} - R_{B^+B^-})$ is equal to

$$
R_{B^0\bar{B}^0} - R_{B^+B^-} = \frac{2C}{\Gamma^2}\frac{x^2}{1 + x^2}\mathcal{C}, \tag{22}
$$

so it goes like $O(x^2)$ and it is isospin suppressed. Then the second term in Eq. (21), corresponding to the mixing effects, is suppressed significantly.

From Fig. 1, we see that $\alpha = 0.53$ would ensure that the second contribution is $O(1\%)$ of $R_{B^+B^-}$, taking into account the suppressions by the α-dependent factor, by $x^2/(1 + x^2)$ and by $1 - \rho$ (but not taking into account the isospin suppression, which would further suppress this term). On the other hand, the first contribution in Eq. (21) would be half of $R_{B^+B^-}$ (essentially R without the effect of mixing in the isospin limit). More generally, a fit to R^α as a function of α would allow one to put a bound on $(R_{B^0\bar{B}^0} - R_{B^+B^-})$ and to extract $R_{B^+B^-}$ directly.

D. Charm pollution

As shown in the previous section, cutting on the time difference of the two decaying B mesons can suppress mixing effects and allow one to distinguish leptons originating from B and charm decays by charge. However, we still need to quantify the expected initial amount of charm contamination. Under the assumption that each tagged lepton originates from a separate B decay chain (which we will relax in the next section), the ratio $R^{\Upsilon(4S)}_{\tau_{\mathrm{had}}\mu}$ can be conveniently expressed in terms of

$$
\begin{aligned}
[R^{\Upsilon(4S)}_{\tau_{\mathrm{had}}\mu}]^{-1} &= \frac{\mathcal{B}(B \to X\mu\nu) + \mathcal{B}(B \to X(h_c \to X'\mu\nu))}{\mathcal{B}(B \to X\tau_{\mathrm{had}}\nu) + \mathcal{B}(B \to X(h_c \to X'\tau_{\mathrm{had}}\nu))} \\
&\quad + \frac{\mathcal{B}(\tau \to \mu\nu\bar{\nu})}{\mathcal{B}(\tau \to \tau_{\mathrm{had}})},
\end{aligned} \tag{23}
$$

where h_c denotes any weakly decaying charmed hadron, i.e., D^+, D^0, D_s, Λ_c and their charge conjugates. As discussed above, using charge identification (ID), but also possibly a cut on leptons not originating from the secondary vertex (i.e., from B decays), it should be possible to suppress contributions where the leptons originate from secondary charm or, in the case of muons, tau decays, by efficiency factors $\epsilon^{(i)} \ll 1$. This allows us to simplify the above expression and write it in terms of the inverse of the inclusive ratio $R(X)_{\tau\mu}$. We obtain

$$
\begin{aligned}
&[R^{\Upsilon(4S)}_{\tau_{\mathrm{had}}\mu_b}]^{-1} \\
&= \left\{ [R(X)_{\tau\mu}]^{-1}\left[1 - \epsilon^{(3)}\frac{\mathcal{B}(B \to X(h_c \to X'\tau\nu))}{\mathcal{B}(B \to X\tau\nu)}\right] \right. \\
&\quad \left. + \epsilon^{(1)}\mathcal{B}(\tau \to \mu\nu\bar{\nu}) + \epsilon^{(2)}\frac{\mathcal{B}(B \to X(h_c \to X'\mu\nu))}{\mathcal{B}(B \to X\tau\nu))} \right\} \\
&\quad \times [\mathcal{B}(\tau \to \tau_{\mathrm{had}})]^{-1},
\end{aligned} \tag{24}
$$

where μ_b denotes muons consistent with originating from the secondary (i.e., b-decay) vertex and we have already used the fact that $\mathcal{B}(B \to X\tau\nu) \gg \mathcal{B}(B \to X(h_c \to X'\tau\nu))$, which we verify below.

We can estimate the size of all three corrections on the right-hand side of Eq. (24) (up to the $\epsilon^{(i)}$ efficiencies) based almost purely on experimental information. Starting with

the $\epsilon^{(1)}$ term, $\mathcal{B}(\tau \to \mu\bar{\nu}\nu) = (17.39 \pm 0.04)\%$ [24] we see, that even without cuts (for $\epsilon^{(1)} \simeq 1$) it leads to an order 4% (computable) systematic effect in $\mathcal{R}(X)_{\tau\mu}$.

We estimate the second and third term thanks to the identity

$$\mathcal{B}(B \to X(h_c \to X'\ell\nu))$$
$$= \mathcal{B}(B \to X_c)\sum_i f(c \to h_c^{(i)})\mathcal{B}(h_c^{(i)} \to X\ell\nu), \quad (25)$$

where the sum runs over all weakly decaying charmed hadrons and $f(c \to h_c^{(i)})$ are the corresponding fragmentation functions. We use Ref. [30] for the charm-inclusive decay branching ratio $\mathcal{B}(B \to X_c) = (97 \pm 4)\%$ and

Ref. [31] for the charm fragmentation functions. Note that the above estimate relies on factorization of the inclusive B-decay amplitudes and is thus subject to related theoretical uncertainties. In addition, the application of charm fragmentation functions extracted from high energy e^+e^- and ep collision data to B decays carries further systematic errors. Consequently, our background evaluations should be taken as order-of-magnitude estimates, which are however sufficient for our purpose. For $\mathcal{B}(D_s \to X\mu\nu)$ we use values measured by CLEO for the electron in the final state [32] $\mathcal{B}(D_s \to Xe\nu) = (6.52 \pm 0.39 \pm 0.15)\%$ which can serve as an effective upper bound on $\mathcal{B}(D_s \to X\mu\nu)$ assuming $e - \mu$ LFU in charm decays. We also use $\mathcal{B}(D^+ \to Xe\nu) = 0.1607 \pm 0.0030$ and $\mathcal{B}(D^0 \to Xe\nu) = 0.0649 \pm 0.0011$ [24]. Finally, we obtain

$$\mathcal{B}(B \to X(h_c \to X'\mu\nu)) = \mathcal{B}(b \to X_c)\{f(c \to D^0)\mathcal{B}(D^0 \to X\mu\nu) + f(c \to D^+)\mathcal{B}(D^+ \to X\mu\nu)$$
$$+ f(c \to D_s)\mathcal{B}(D_s \to X\mu\nu) + f(c \to \Lambda_c)\mathcal{B}(\Lambda_c \to X\mu\nu) + ...\} \lesssim 0.088,$$
$$\mathcal{B}(B \to X(h_c \to X'\tau\nu)) = \mathcal{B}(b \to X_c)\{f(c \to D^+)\mathcal{B}(D^+ \to \tau\nu) + f(c \to D_s)\mathcal{B}(D_s \to \tau\nu) + ...\} \simeq 0.0067. \quad (26)$$

These values are to be compared with the large electron positron collider (LEP) experimental determination of $\mathcal{B}(b \to q\tau\nu) \simeq \mathcal{B}(B \to X\tau\nu) = (2.41 \pm 0.23)\%$ [23]. In particular, before cuts and without lepton charge ID (for $\epsilon^{(2)} \simeq 1$) the second term in Eq. (24) would represent a dominant 80% systematic effect in the determination of $\mathcal{R}(X)_{\tau\mu}$. Finally, the effect of the $\epsilon^{(3)}$ term before cuts (for $\epsilon^{(3)} \simeq 1$) represents a relative 28% systematic effect on the determination of $\mathcal{R}(X)_{\tau\mu}$.

In summary, the term with $\epsilon^{(1)}$ is small thanks to the low value of $\mathcal{B}(\tau \to \mu\bar{\nu}\nu)$, whereas the factors of $\epsilon^{(2)}$ and $\epsilon^{(3)}$ have large values but are related to charm pollution, which (hopefully) can be reduced thanks to charge ID leading to small efficiencies $\epsilon^{(2,3)}$.

E. Leptons emitted from the same B-meson

Lastly we need to consider backgrounds where both leptons are of different charge and flavor, but originate from the same B-decay chain, corresponding to the parton-level chain

$$b \to q(c \to q'\ell^+\nu)(\bar{c} \to q''\ell'^-\nu) \quad \text{and}$$
$$b \to (c \to q\ell^+\nu)\ell'^-\nu. \quad (27)$$

Denoting these processes collectively as $B \to X\ell\ell'$, and assuming they can be suppressed by cutting on leptons not originating from the secondary vertex (i.e., from b decays), we can again write the relative correction to Eq. (24) due to these contributions expanded to leading order in all $\epsilon^{(i)}$ as

$$[R_{\tau_{\mathrm{had}}\mu_b}^{\Upsilon(4S)}]^{-1} = [R(X)_{\tau\mu}\mathcal{B}(\tau \to \tau_{\mathrm{had}})]^{-1}\left[1 - \epsilon^{(4)}\frac{\mathcal{B}(B \to X)\mathcal{B}(\bar{B} \to X\tau e)}{\mathcal{B}(B \to X\tau\nu)\mathcal{B}(\bar{B} \to Xe\nu)} + \epsilon^{(5)}\frac{\mathcal{B}(B \to X)\mathcal{B}(\bar{B} \to X\mu e)}{\mathcal{B}(B \to X\mu\nu)\mathcal{B}(\bar{B} \to Xe\nu)}\right] + ..., \quad (28)$$

where the inclusive hadronic B-decay branching ratio is denoted as $\mathcal{B}(B \to X) \lesssim 1 - \sum_\ell \mathcal{B}(B \to X\ell\nu) \simeq 0.76$, we take $\mathcal{B}(B \to X_c e\nu) \simeq \mathcal{B}(B \to X_c \mu\nu) \simeq 0.11$ [24], and the ellipsis denotes the remaining corrections on the right-hand side of Eq. (24). Using the numerical values given above we obtain for the relevant $b \to (c \to q\ell^+\nu)\ell'^-\nu$ transitions

$$\mathcal{B}(B \to X(h_c \to X'e\nu)\tau\nu) = \mathcal{B}(B \to X_c\tau\nu)[f(c \to D^0)\mathcal{B}(D^0 \to Xe\nu) + f(c \to D^+)\mathcal{B}(D^+ \to Xe\nu)$$
$$+ f(c \to D_s)\mathcal{B}(D_s \to Xe\nu) + f(c \to \Lambda_c)\mathcal{B}(\Lambda_c \to Xe\nu) + ...] \simeq 0.0021,$$
$$\mathcal{B}(B \to X(h_c \to X'\tau\nu)e\nu) = \mathcal{B}(B \to X_c e\nu)f(c \to D_s)\mathcal{B}(D_s \to \tau\nu) \simeq 0.00042,$$
$$\mathcal{B}(B \to X(h_c \to X'e\nu)\mu\nu) = \mathcal{B}(B \to X_c\mu\nu)[f(c \to D^0)\mathcal{B}(D^0 \to Xe\nu) + f(c \to D^+)\mathcal{B}(D^+ \to Xe\nu)$$
$$+ f(c \to D_s)\mathcal{B}(D_s \to Xe\nu) + f(c \to \Lambda_c)\mathcal{B}(\Lambda_c \to Xe\nu) + ...] \simeq 0.0095$$

$$\mathcal{B}(B \to X(h_c \to X'\mu\nu)e\nu) = \mathcal{B}(B \to X_c e\nu)[f(c \to D^0)\mathcal{B}(D^0 \to X\mu\nu) + f(c \to D^+)\mathcal{B}(D^+ \to X\mu\nu)$$

$$+ f(c \to D_s)\mathcal{B}(D_s \to X\mu\nu) + f(c \to \Lambda_c)\mathcal{B}(\Lambda_c \to X\mu\nu) + \ldots] \lesssim 0.010. \quad (29)$$

Finally, for the decay chain $b \to qc\bar{c}(c \to q\ell\nu)(\bar{c} \to q'\ell'\nu)$, using $\mathcal{B}(B \to X_{c\bar{c}}) \simeq 22\%$ [33] and after including $c \to q\ell\nu$ and $\bar{c} \to q'\ell'\nu$ transition rates, we find

$$\mathcal{B}(B \to X(h_c \to X'e\nu)(h_{\bar{c}} \to X''\tau\nu)) = \mathcal{B}(B \to X_{c\bar{c}})f(\bar{c} \to \bar{D}_s)\mathcal{B}(\bar{D}_s \to \tau\nu)$$

$$\times [f(c \to D^0)\mathcal{B}(D^0 \to Xe\nu) + f(c \to D^+)\mathcal{B}(D^+ \to Xe\nu)$$

$$+ f(c \to D_s)\mathcal{B}(D_s \to Xe\nu) + f(c \to \Lambda_c)\mathcal{B}(\Lambda_c \to Xe\nu)] \simeq 0.0001,$$

$$\mathcal{B}(B \to X(h_c \to X'e\nu)(h_{\bar{c}} \to X''\mu\nu)) = \mathcal{B}(B \to X_{c\bar{c}}) \times [f(c \to D^0)\mathcal{B}(D^0 \to X\mu\nu) + f(c \to D^+)\mathcal{B}(D^+ \to X\mu\nu)$$

$$+ f(c \to D_s)\mathcal{B}(D_s \to X\mu\nu) + f(c \to \Lambda_c)\mathcal{B}(\Lambda_c \to X\mu\nu)]$$

$$\times [f(c \to D^0)\mathcal{B}(D^0 \to Xe\nu) + f(c \to D^+)\mathcal{B}(D^+ \to Xe\nu)$$

$$+ f(c \to D_s)\mathcal{B}(D_s \to Xe\nu) + f(c \to \Lambda_c)\mathcal{B}(\Lambda_c \to Xe\nu)] \simeq 0.0018. \quad (30)$$

Putting these values together we observe that these backgrounds are individually comparable in size to the signal (i.e., they would represent approximately 80% and 150% relative corrections, respectively) in absence of cuts to suppress them (for $\epsilon^{(4,5)} \simeq 1$). While they are similar in magnitude, they are highly correlated and contribute with opposite signs, so that they tend to cancel to a degree for $\epsilon^{(4)} \simeq \epsilon^{(5)}$. In fact, the two terms become exactly equal in the limit where one can neglect charm decays to muons and taus (in the ratio $R_{\tau e}^{\Upsilon(4S)}$ these would be charm decays to electrons and taus). On the other hand, contrary to the corrections outlined in Eq. (24), the corrections considered in this section cannot be suppressed using only lepton charge ID. This highlights the crucial importance of discriminating against leptons originating from the same B-decay chain, for instance through geometrical considerations. An alternative strategy could consist in discriminating leptons arising from the secondary (B-decay) vertices from those arising further down in the decay chains. A quantitative assessment of the feasibility of either of the two approaches through an appropriate experimental analysis would require a dedicated experimental study and is beyond the scope of this work.

III. CONCLUSIONS

Relying on the specific properties of B-factories and in particular the Belle II experiment, we have proposed to compare the inclusive rates of $\Upsilon(4S) \to e^{\pm}\mu^{\mp}X$, $\Upsilon(4S) \to \mu^{\pm}\tau_{\text{had}}^{\mp}X$ and $\Upsilon(4S) \to e^{\pm}\tau_{\text{had}}^{\mp}X$. This measurement can be related to the ratio $R(X)_{\tau\ell} \equiv \Gamma(b \to X\tau\nu)/\Gamma(b \to X\ell\nu)$ ($\ell = e$ or μ), once appropriate experimental cuts are applied to suppress the effects of neutral B mixing and leptons emitted from rare FCNC (semileptonic) B decays, as well as secondary charmonium, charm and tau decays. The feasibility of our proposal crucially assumes that

hadronically decaying tau leptons originating from the B decay vertices can be efficiently disentangled from backgrounds (e.g., from hadronic B decays involving three or more charged pions) at Belle II. A dedicated experimental study of this is however beyond the scope of the present paper.

We have focused on the case of $R_{\tau\mu}^{\Upsilon(4S)} \simeq R(X)_{\tau\mu}$, but our discussion applies equally well to the tau-electron combination, swapping the roles played by electrons and muons. The current deviations in $B \to D^*\ell\nu$ and $B \to D\ell\nu$ when τ channels are compared to electronic or muonic modes are at the level of 10% (for the LFU ratios of branching ratios) and provide a benchmark for the target sensitivity of our proposal. This is illustrated by the very simple case where new physics mimics the $V - A$ structure of $b \to c\tau\nu$ currents in the SM, leading to a universal rescaling of all $b \to c\tau\nu$ branching ratios.

Given our estimates, the systematic uncertainties in the determination of $R(X)_{\tau\ell}$ from a measurement of $R_{\tau\ell}^{\Upsilon(4S)}$ could be brought below a given value (ϵ_{sys}) provided that (1) cuts on the $B - \bar{B}$ impact parameter difference can suppress the neutral B meson mixing effects below ϵ_{sys} combined with an efficient lepton charge ID to suppress semileptonic charm-decay contamination; (2a) multiple leptons originating from the same B decay chain can be suppressed to better than ϵ_{sys} or alternatively (2b) leptons arising from the secondary (B-decay) vertices can be discriminated against those arising further down in the decay chains to roughly better than ϵ_{sys}. Further dedicated experimental studies are needed to establish the actually attainable precision by Belle II.

In summary, we have proposed a novel method to test the persistent hints of violation of LFU observed in semileptonic B decays. This measurement would constitute an additional and potentially competitive probe of LFU violations in $b \to c\ell\nu$ transitions, complementary to

exclusive measurements and accessible in the Belle II environment.

ACKNOWLEDGMENTS

[1] Y. S. Amhis *et al.*, Averages of *b*-hadron, *c*-hadron, and τ-lepton properties as of 2018, Eur. Phys. J. C **81,** 226 (2021).

[2] J. P. Lees *et al.*, Evidence for an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays, Phys. Rev. Lett. **109,** 101802 (2012).

[3] J. P. Lees *et al.*, Measurement of an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays and implications for charged Higgs bosons, Phys. Rev. D **88,** 072012 (2013).

[4] M. Huschle *et al.*, Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle, Phys. Rev. D **92,** 072014 (2015).

[5] R. Aaij *et al.*, Measurement of the Ratio of Branching Fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$, Phys. Rev. Lett. **115,** 111803 (2015); **115,** 159901(E) (2015).

[6] S. Hirose *et al.*, Measurement of the τ Lepton Polarization and $R(D^*)$ in the Decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$, Phys. Rev. Lett. **118,** 211801 (2017).

[7] S. Hirose *et al.*, Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ with one-prong hadronic τ decays at Belle, Phys. Rev. D **97,** 012004 (2018).

[8] R. Aaij *et al.*, Measurement of the Ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ Branching Fractions Using Three-Prong τ-Lepton Decays, Phys. Rev. Lett. **120,** 171802 (2018).

[9] R. Aaij *et al.*, Test of Lepton Flavor Universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays, Phys. Rev. D **97,** 072013 (2018).

[10] A. Abdesselam *et al.*, Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method, arXiv:1904.08794.

[11] R. Aaij *et al.*, Measurement of the Ratio of Branching Fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$, Phys. Rev. Lett. **120,** 121801 (2018).

[12] D. Bečirević, S. Fajfer, N. Košnik, and O. Sumensari, Leptoquark model to explain the *B*-physics anomalies, R_K and R_D, Phys. Rev. D **94,** 115021 (2016).

[13] R. Alonso, B. Grinstein, and J. M. Camalich, Lifetime of B_c^- Constrains Explanations for Anomalies in $B \to D^{(*)}\tau\nu$, Phys. Rev. Lett. **118,** 081802 (2017).

[14] S. Descotes-Genon, S. Fajfer, J. F. Kamenik, and M. Novoa-Brunet, Implications of $b \to s\mu\mu$ anomalies for future measurements of $B \to K^{(*)}\nu\bar{\nu}$ and $K \to \pi\nu\bar{\nu}$, Phys. Lett. B **809,** 135769 (2020).

[15] J. F. Kamenik, S. Monteil, A. Semkiv, and L. V. Silva, Lepton polarization asymmetries in rare semi-tauonic $b \to s$ exclusive decays at FCC-*ee*, Eur. Phys. J. C **77,** 701 (2017).

[16] B. Capdevila, A. Crivellin, S. Descotes-Genon, L. Hofer, and J. Matias, Searching for New Physics with $b \to s\tau^+\tau^-$ Processes, Phys. Rev. Lett. **120,** 181802 (2018).

[17] A. Greljo, J. M. Camalich, and J. D. Ruiz-Álvarez, Mono-τ Signatures at the LHC Constrain Explanations of *B*-decay Anomalies, Phys. Rev. Lett. **122,** 131803 (2019).

[18] D. A. Faroughy, A. Greljo, and J. F. Kamenik, Confronting lepton flavor universality violation in B decays with high-p_T tau lepton searches at LHC, Phys. Lett. B **764,** 126 (2017).

[19] D. Aloni, A. Efrati, Y. Grossman, and Y. Nir, Υ and ψ leptonic decays as probes of solutions to the $R_D^{(*)}$ puzzle, J. High Energy Phys. 06 (2017) 019.

[20] J. Lees *et al.*, Precision Measurement of the $\mathcal{B}(\Upsilon(3S) \to \tau^+\tau^-)/\mathcal{B}(\Upsilon(3S) \to \mu^+\mu^-)$ Ratio, Phys. Rev. Lett. **125,** 241801 (2020).

[21] C. H. García-Duque, J. H. Muñoz, N. Quintero, and E. Rojas, Extra gauge bosons and lepton flavor universality violation in Υ and *B* meson decays, Phys. Rev. D **103,** 073003 (2021).

[22] P. D. Group, Review of particle physics, Prog. Theor. Exp. Phys. **2020,** 083C01 (2020).

[23] M. Freytsis, Z. Ligeti, and J. T. Ruderman, Flavor models for $\bar{B} \to D^{(*)}\tau\bar{\nu}$, Phys. Rev. D **92,** 054018 (2015).

[24] P. Zyla *et al.*, Review of particle physics, Prog. Theor. Exp. Phys. **2020,** 083C01 (2020).

[25] G. T. Bodwin, E. Braaten, and G. Lepage, Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium, Phys. Rev. D **51,** 1125 (1995); **55,** 5853(E) (1997).

[26] W.-L. Sang, F. Feng, and Y. Jia, $\mathcal{O}(\alpha_s v^2)$ corrections to hadronic decay of vector quarkonia, Phys. Rev. D **102,** 094021 (2020).

[27] H. S. Chung, J. Lee, and C. Yu, NRQCD matrix elements for *S*-wave bottomonia and $\Gamma[\eta_b(nS) \to \gamma\gamma]$ with relativistic corrections, Phys. Lett. B **697,** 48 (2011).

[28] P. Gambino and J. F. Kamenik, Lepton energy moments in semileptonic charm decays, Nucl. Phys. **B840,** 424 (2010).

[29] D. Boutigny *et al.*, The *BABAR* physics book: Physics at an asymmetric *B* factory, 1998, https://slac.stanford.edu/pubs/slacreports/reports19/slac-r-504.pdf.

[30] B. Aubert *et al.*, Study of inclusive B- and anti-B0 decays to flavor-tagged D, D(s) and Lambda+(c), Phys. Rev. D **75,** 072002 (2007).

[31] M. Lisovyi, A. Verbytskyi, and O. Zenaiev, Combined analysis of charm-quark fragmentation-fraction measurements, Eur. Phys. J. C **76**, 397 (2016).

[32] D. Asner *et al.*, Measurement of absolute branching fractions of inclusive semileptonic decays of charm and charmed-strange mesons, Phys. Rev. D **81**, 052007 (2010).

[33] D. Abbaneo *et al.*, Combined results on *b* hadron production rates and decay properties, arXiv:hep-ex/0112028.