

New physics with the lepton flavor violating decay $\tau \rightarrow 3\mu$

Zaineb Calcuttawala,^{1,*} Anirban Kundu,^{1,†} Soumitra Nandi,^{2,‡} and Sunando Kumar Patra^{2,3,§}

¹*Department of Physics, University of Calcutta, 92 Acharya Prafulla Chandra Road, Kolkata 700009, India*

²*Department of Physics, Indian Institute of Technology, Guwahati 781039, India*

³*Department of Physics, Indian Institute of Technology, Kanpur 208016, India*



(Received 6 March 2018; published 7 May 2018)

Lepton flavor violating (LFV) processes are a smoking gun signal of new physics (NP). If the semileptonic B decay anomalies are indeed due to some NP, such operators can potentially lead to LFV decays involving the second and the third generation leptons, like $\tau \rightarrow 3\mu$. In this paper, we explore how far the nature of NP can be unraveled at the next generation B -factories like Belle-II, provided the decay $\tau \rightarrow 3\mu$ has been observed. We use four observables with which the differentiation among NP operators may be achieved to a high confidence level. Possible presence of multiple NP operators are also analyzed with the optimal observable technique. While the analysis can be improved even further if the final state muon polarizations are measured, we present this work as a motivational tool for the experimentalists, as well as a template for the analysis of similar processes.

DOI: [10.1103/PhysRevD.97.095009](https://doi.org/10.1103/PhysRevD.97.095009)

I. INTRODUCTION

Lepton flavor is an accidental symmetry of the standard model (SM), and there are many extensions of the SM, like the seesaw models, supersymmetric SM, flavor-changing Z' or scalars, leptoquarks, or left-right symmetric models, that can naturally break this symmetry. Even within the ambit of SM, neutrino mixing provides a source of leptonic flavor violation (LFV), but the rates are too small to be observed in any near future [1]. Thus, observation of any LFV decay is a smoking gun signal of new physics (NP). In general four types of LFV processes have been looked for: (i) leptonic decays ($\tau \rightarrow 3e$, $\tau \rightarrow 3\mu$, $\mu \rightarrow 3e$, $\tau \rightarrow 1e + 2\mu$, $\tau \rightarrow 1\mu + 2e$), (ii) radiative decays ($\tau \rightarrow e\gamma$, $\tau \rightarrow \mu\gamma$, $\mu \rightarrow e\gamma$), (iii) semileptonic decays ($\ell_1 \rightarrow \ell_2 M$, where M is some meson), and (iv) conversion (like $\mu \rightarrow e$). They are not all independent, e.g., a flavor-changing electromagnetic penguin can also give rise to leptonic LFV decays. Most of the decays, however, have very stringent limits [1], the branching ratios (BR) being typically of the order of 10^{-8} or even smaller.

The interest in such LFV decays have recently been rekindled from the observation that some of the semileptonic B decay modes show anomalous deviations from the SM expectations, which may possibly be explained by lepton flavor nonuniversality (LFNU) as well as LFV. As an example, let us refer the reader to a recent attempt in Refs. [2,3], where the authors have shown that both R_K , R_{K^*} and $R(D)$, $R(D^*)$ anomalies can be explained satisfactorily with only two new operators, if the weak and mass bases of the charged leptons $\{\mu, \tau\}$ are related by a field rotation [4]. The apparent excess in the LFV decay channel of the Higgs boson, $h \rightarrow \mu\tau$, as was once reported by the CMS Collaboration [5], coupled with such B decay anomalies, could lead to some well-motivated and fairly constrained models of LFV [6]. The LFV operators with only leptonic fields can also be induced by renormalization group (RG) running of semileptonic operators [7]. Thus, there is enough motivation to seriously look into such LFV channels; they might be observed at the Large Hadron Collider (LHC) [8], or dedicated super-B factories like Belle-II. Implications of such LFV decays may also be found in Ref. [9].

In this paper, we will focus solely on the leptonic channel $\tau \rightarrow 3\mu$. The channel has already been studied in detail in the literature; there are both model-independent [10–12] as well as partially model-dependent [13,14] studies. The Belle Collaboration has an upper bound on the BR [15]

$$\text{BR}(\tau \rightarrow 3\mu) < 2.1 \times 10^{-8} \quad (1)$$

at 90% confidence level (CL). One reason for looking at this particular channel is the possibility of leptonic rotation

* zaineb.calcuttawala@gmail.com

† anirban.kundu.cu@gmail.com

‡ soumitra.nandi@gmail.com

§ sunando.patra@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

in the $\{\mu, \tau\}$ sector as mentioned above, which invariably leads to such LFV channels out of lepton flavor conserving operators in the weak basis. Another reason, of course, is the relative ease with which the final state muons can be detected in both hadronic and e^+e^- colliders.

Here we would like to push the studies on LFV in $\tau \rightarrow 3\mu$ further by asking and answering a few questions. Observation of even a single $\tau \rightarrow 3\mu$ event is a definite signal for NP. Assuming that one observes, possibly at a super-B factory like Belle-II, a few events for the LFV decay in question, will one be able to unearth the nature of the possible operators that can lead to such a decay? It has been shown [10,13] that there can be six independent LFV operators in the chiral basis that lead to $\tau \rightarrow 3\mu$. If the final state muon polarizations are not measured, all the operators are *a priori* equally probable, and obviously only the number of events will not tell us anything about the presence or absence of any of these operators; it can only yield some estimate on the respective Wilson coefficients (WC). So, are there observables which will help us to differentiate between these operators? We will show that this is indeed possible, without using higher-order differential cross sections like Ref. [10] which may have very few number of events in each bin. At this point, we also note that some more operators can be generated through Fierz reordering, but obviously they are not independent of the first six, and therefore we will not consider them any further.

The second question that we would like to ask is whether the existence of more than one such operators can be disentangled from the data. Here the answer will be partially positive, unless, again, the muon polarizations are measured. If one can have a sizeable number of events, and measure the muon polarizations too, one may have in principle further observables, but we would like to be conservative. Anyway, as we will show, one does not

expect more than 70 events or so at the most with an integrated luminosity $\mathcal{L}_{\text{int}} = 50 \text{ ab}^{-1}$ at Belle-II.

We will use the method of optimal observables (OO), which has already been used in different areas of particle physics [16,17], and in particular, for flavor physics [18,19]. This method displays the amount of significance level (“how many sigma” in standard parlance) by which one point in the allowed parameter space can be separated from another point. This is the only way to approach the question of model differentiation before the arrival of the data. Once one has the data, other methods, like the unbinned multivariate maximum likelihood, may be employed.

A related question is the number of events with which one can have a successful differentiation among models, where a model is specified by its operator structure and WCs. As expected, if the number of events is too small, it will be harder to differentiate among various models, or in other words, the significance level will be lower. We will quantify this statement subsequently.

The paper is arranged as follows. In Sec. II, we enlist all the possible NP operators that can give rise to the $\tau \rightarrow 3\mu$ decay, and the observables that we deal with are discussed in Sec. III. In Sec. IV, we show the differentiation among models with only one NP operator. Section V deals with models with two such NP operators, and we discuss how well the presence of the second operator can be found out from various observables. Section VI summarizes and concludes the paper.

II. THE NEW PHYSICS OPERATORS

For this section, we will follow the notation and convention of Ref. [13]. The most general LFV Lagrangian can be written as

$$\begin{aligned} \mathcal{L} = & \frac{1}{\Lambda^2} [g_{LL}^S(\bar{\mu}_L\mu_R)(\bar{\mu}_R\tau_L) + g_{LR}^S(\bar{\mu}_L\mu_R)(\bar{\mu}_L\tau_R) + g_{RL}^S(\bar{\mu}_R\mu_L)(\bar{\mu}_R\tau_L) + g_{RR}^S(\bar{\mu}_R\mu_L)(\bar{\mu}_L\tau_R) \\ & + g_{LL}^V(\bar{\mu}_R\gamma^\alpha\mu_R)(\bar{\mu}_L\gamma_\alpha\tau_L) + g_{LR}^V(\bar{\mu}_R\gamma^\alpha\mu_R)(\bar{\mu}_R\gamma_\alpha\tau_R) + g_{RL}^V(\bar{\mu}_L\gamma^\alpha\mu_L)(\bar{\mu}_L\gamma_\alpha\tau_L) + g_{RR}^V(\bar{\mu}_L\gamma^\alpha\mu_L)(\bar{\mu}_R\gamma_\alpha\tau_R) \\ & + \frac{1}{2}g_{LR}^T(\bar{\mu}_L\sigma^{\alpha\beta}\mu_R)(\bar{\mu}_L\sigma_{\alpha\beta}\tau_R) + \frac{1}{2}g_{RL}^T(\bar{\mu}_R\sigma^{\alpha\beta}\mu_L)(\bar{\mu}_R\sigma_{\alpha\beta}\tau_L)], \end{aligned} \quad (2)$$

and we will denote the operator accompanying g_{IJ}^X ($X = S, V, T$, and $I, J = L, R$) as O_{IJ}^X . Λ is the cutoff scale, which we will set at 5 TeV for our analysis. We separate the operators into three major classes: **S** (operators of the form O_{IJ}^S), **V** (the O_{IJ}^V operators) and **T** (the tensor operators O_{IJ}^T). Thus, the effective Lagrangian is of the form

$$\mathcal{L} = \frac{1}{\Lambda^2} \left[\sum_{I,J=L,R} (g_{IJ}^S O_{IJ}^S + g_{IJ}^V O_{IJ}^V) + \sum_{I \neq J} g_{IJ}^T O_{IJ}^T \right]. \quad (3)$$

In the above mentioned basis, not all the ten operators are independent; Fierz transformation relates the two tensor operators with the rest, and the pairs $O_{LL}^S - O_{LL}^V$ and $O_{RR}^S - O_{RR}^V$ are also related [10]. Thus, only four scalar and two vector operators are enough to span the operator basis. However, we keep all of them for the time being, as the mediator that has been integrated out may give rise to operators that are linear combinations of the six independent ones, like the tensor operators that can be generated from some hypothetical spin-2 mediators.

Writing the BR in terms of the new WCs [13], one may easily show that the decay $\tau \rightarrow 3\mu$ has maximal sensitivity to **V** or **T** operators. As an example, the present bound on $\text{BR}(\tau \rightarrow 3\mu) < 2.1 \times 10^{-8}$ translates to $|g_{RL}^S| \approx 1$, while $|g_{RL}^T|, |g_{RL}^V| \sim \mathcal{O}(0.1)$. Thus, for a given number of events, the reach for **V** or **T** WCs is better than that for the **S** ones.

The left-chiral fields being $SU(2)$ doublets, one can also get a neutrino-antineutrino pair out of the operators O_{RR}^S and O_{RR}^V , which technically gives an extra contribution to the SM τ decay channel $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$. However, the couplings will turn out to be so constrained as not to affect this channel in any significant amount. Similar LFV operators for $\mu \rightarrow 3e$ may affect the extraction of the Fermi coupling G_F from the muon lifetime in a measurable way.

One may try to look for $\tau \rightarrow \mu\gamma$ by contracting a pair of muons and taking the photon with momentum q^μ out from the loop. For the scalar operators, the contribution vanishes in the $q^2 = 0$ limit, and for the vector operators, this amounts to charge and not the transition magnetic moment renormalization.

To begin with, we will consider the presence of one operator at a time. This generates six independent models

spanning over the **S** and **V** classes. Next, we will consider the presence of two operators at a time, which include the well-motivated combinations like

$$\begin{aligned} O_{9L}^S &= O_{LL}^S + O_{RL}^S, & O_{10L}^S &= O_{LL}^S - O_{RL}^S, \\ O_{9R}^S &= O_{LR}^S + O_{RR}^S, & O_{10R}^S &= O_{LR}^S - O_{RR}^S, \end{aligned} \quad (4)$$

and similarly for the **V** class. Our goal will be to pinpoint whether or not these two coupling scenarios can be differentiated from those involving only one coupling at a time.

III. OBSERVABLES

In this section, we will define the observables which we have used in our analysis to differentiate the effects of different NP operators. As an example, we consider the **S** class of operators, O_{IJ}^S , taken from Eq. (3), and consider the decay $\tau^- \rightarrow \mu^- \mu^+ \mu^-$. The double differential cross section for the antimuon is given, after integrating out the phase space for the two muons, by

$$\begin{aligned} \frac{dB_\tau}{dx d(\cos\theta)} &= \frac{T_\tau m_\tau^5}{128 \times 48\pi^3 \Lambda^4} [3x^2(|g_{RL}^S|^2 + |g_{LL}^S|^2 + |g_{RR}^S|^2 + |g_{LR}^S|^2) - x^3(3|g_{RL}^S|^2 + 2|g_{LL}^S|^2 + 2|g_{RR}^S|^2 + 3|g_{LR}^S|^2) \\ &\quad + x^2 \cos\theta(3|g_{RL}^S|^2 + |g_{LL}^S|^2 - |g_{RR}^S|^2 - 3|g_{LR}^S|^2) - x^3 \cos\theta(3|g_{RL}^S|^2 + 2|g_{LL}^S|^2 - 2|g_{RR}^S|^2 - 3|g_{LR}^S|^2)], \end{aligned} \quad (5)$$

where we take the muons to be massless, and use the notation

$$B_\tau \equiv \text{BR}(\tau \rightarrow 3\mu). \quad (6)$$

Here, $T_\tau = 1/\Gamma$ is the lifetime of the τ lepton, $x = 2E_{\bar{\mu}}/m_\tau$ is the reduced energy of the antimuon, and θ is angle between the polarization of the τ and the momentum of the antimuon, following the convention of Ref. [11]. For further discussion, let us define

$$\begin{aligned} g_1 &\equiv |g_{RL}^S|^2 + |g_{LL}^S|^2 + |g_{RR}^S|^2 + |g_{LR}^S|^2, \\ g_2 &\equiv 3|g_{RL}^S|^2 + 2|g_{LL}^S|^2 + 2|g_{RR}^S|^2 + 3|g_{LR}^S|^2, \\ g_3 &\equiv 3|g_{RL}^S|^2 + |g_{LL}^S|^2 - |g_{RR}^S|^2 - 3|g_{LR}^S|^2, \\ g_4 &\equiv 3|g_{RL}^S|^2 + 2|g_{LL}^S|^2 - 2|g_{RR}^S|^2 - 3|g_{LR}^S|^2. \end{aligned} \quad (7)$$

Thus

$$B_\tau = \frac{T_\tau m_\tau^5}{128 \times 24\pi^3 \Lambda^4} \left[g_1 - \frac{1}{4} g_2 \right]. \quad (8)$$

The number of events gives the information only on the combination $g_1 - \frac{1}{4} g_2$.

Another observable is the observed integrated forward-backward asymmetry A_{FB} , defined as

$$A_{FB} = \frac{N_F - N_B}{N_F + N_B}, \quad (9)$$

with

$$\begin{aligned} N_F &= \sigma_{\text{Prod}} \mathcal{L}_{\text{int}} \epsilon \int_0^1 dx \int_0^1 d(\cos\theta) \frac{dB_\tau}{dx d(\cos\theta)}, \\ N_B &= \sigma_{\text{Prod}} \mathcal{L}_{\text{int}} \epsilon \int_0^1 dx \int_{-1}^0 d(\cos\theta) \frac{dB_\tau}{dx d(\cos\theta)}, \end{aligned} \quad (10)$$

where σ_{Prod} is the τ production cross section, \mathcal{L}_{int} is the integrated luminosity, and ϵ is the combined detection efficiency in the $\tau \rightarrow 3\mu$ channel.

We will also define the x -dependent asymmetry, normalized to the total decay width, as

$$\begin{aligned} \frac{dA_{FB}}{dx} &\equiv A'_{FB}(x) \\ &= \sigma_{\text{Prod}} \mathcal{L}_{\text{int}} \epsilon \frac{\int_0^1 d(\cos\theta) \frac{dB_\tau}{dx d(\cos\theta)} - \int_{-1}^0 d(\cos\theta) \frac{dB_\tau}{dx d(\cos\theta)}}{N_B + N_F} \\ &\equiv \frac{N_F(x) - N_B(x)}{N}, \end{aligned} \quad (11)$$

where $N = N_F + N_B$ gives the total number of signal events.

Instead of the antimuon, one can play an identical game with one of the same-sign muons (i.e., the one with the same sign as the decaying τ), say the more energetic of the two. Let us define

$$\begin{aligned} g'_1 &\equiv |g_{RL}^S|^2 + 3|g_{LL}^S|^2 + 3|g_{RR}^S|^2 + |g_{LR}^S|^2, \\ g'_2 &\equiv |g_{RL}^S|^2 + 4|g_{LL}^S|^2 + 4|g_{RR}^S|^2 + |g_{LR}^S|^2, \\ g'_3 &\equiv |g_{RL}^S|^2 + 7|g_{LL}^S|^2 - 7|g_{RR}^S|^2 - |g_{LR}^S|^2, \\ g'_4 &\equiv |g_{RL}^S|^2 + 4|g_{LL}^S|^2 - 4|g_{RR}^S|^2 - |g_{LR}^S|^2, \end{aligned} \quad (12)$$

so that

$$\begin{aligned} \frac{dB_\tau}{dyd(\cos\alpha)} &= \frac{T_\tau m_\tau^5}{128 \times 96\pi^3 \Lambda^4} \\ &\times [3y^2 g'_1 - 2y^3 g'_2 + y^2 \cos\alpha g'_3 - 2y^3 \cos\alpha g'_4], \end{aligned} \quad (13)$$

where $y = 2E_\mu/m_\tau$ is the reduced energy of the more energetic same-sign muon, and α is angle between its direction and the polarization of the τ lepton.

In an analogous way to Eq. (11), one can define

$$\begin{aligned} \mathcal{A}'_{FB}(y) &= \frac{N_F(y) - N_B(y)}{N} \\ &= \sigma_{\text{Prod}} \mathcal{L}_{\text{int}} \epsilon \frac{\int_0^1 d(\cos\alpha) \frac{dB_\tau}{dyd(\cos\alpha)} - \int_{-1}^0 d(\cos\alpha) \frac{dB_\tau}{dyd(\cos\alpha)}}{N}. \end{aligned} \quad (14)$$

As will be shown later, the observables \mathcal{A}'_{FB} and \mathcal{A}_{FB} are useful to differentiate the sensitivities of the subtypes of operators within a particular type, say **S** or **V**.

Similarly, for the **V** class of models, one obtains, in an analogous way to Eq. (5),

$$\begin{aligned} \frac{dB_\tau}{dx d(\cos\theta)} &= \frac{T_\tau m_\tau^5}{128 \times 12\pi^3 \Lambda^4} [3x^2(4|g_{RL}^V|^2 + |g_{LL}^V|^2 + |g_{RR}^V|^2 + 4|g_{LR}^V|^2) - 2x^3(6|g_{RL}^V|^2 + |g_{LL}^V|^2 + |g_{RR}^V|^2 + 6|g_{LR}^V|^2) \\ &+ x^2 \cos\theta(12|g_{RL}^V|^2 + |g_{LL}^V|^2 - |g_{RR}^V|^2 - 12|g_{LR}^V|^2) - 2x^3 \cos\theta(6|g_{RL}^V|^2 + |g_{LL}^V|^2 - |g_{RR}^V|^2 - 6|g_{LR}^V|^2)]. \end{aligned} \quad (15)$$

The corresponding BR is

$$B_\tau = \frac{T_\tau m_\tau^5}{128 \times 12\pi^3 \Lambda^4} [2g_1 - g_2], \quad (16)$$

where,

$$\begin{aligned} g_1 &\equiv 4|g_{RL}^V|^2 + |g_{LL}^V|^2 + |g_{RR}^V|^2 + 4|g_{LR}^V|^2, \\ g_2 &\equiv 6|g_{RL}^V|^2 + |g_{LL}^V|^2 + |g_{RR}^V|^2 + 6|g_{LR}^V|^2, \\ g_3 &\equiv 12|g_{RL}^V|^2 + |g_{LL}^V|^2 - |g_{RR}^V|^2 - 12|g_{LR}^V|^2, \\ g_4 &\equiv 6|g_{RL}^V|^2 + |g_{LL}^V|^2 - |g_{RR}^V|^2 - 6|g_{LR}^V|^2. \end{aligned} \quad (17)$$

From Eqs. (5), (8), (15), and (16), one finds that **V**-type operators generate more events than **S**-type operators, if the orders of magnitude of their WCs are similar. The number of events as well as the angular distribution depend on the model subtype. We refer the reader to Fig. 1, which shows this explicitly.

For the same-sign muon, one gets

$$\begin{aligned} \frac{dB_\tau}{dyd(\cos\alpha)} &= \frac{T_\tau m_\tau^5}{128 \times 24\pi^3 \Lambda^4} \\ &\times [3y^2 g'_1 - 8y^3 g'_2 + y^2 \cos\alpha g'_3 - 8y^3 \cos\alpha g'_4], \end{aligned} \quad (18)$$

with

$$\begin{aligned} g'_1 &\equiv 4|g_{RL}^V|^2 + 3|g_{LL}^V|^2 + 3|g_{RR}^V|^2 + 4|g_{LR}^V|^2, \\ g'_2 &\equiv |g_{RL}^V|^2 + |g_{LL}^V|^2 + |g_{RR}^V|^2 + |g_{LR}^V|^2, \\ g'_3 &\equiv 4|g_{RL}^V|^2 + 7|g_{LL}^V|^2 - 7|g_{RR}^V|^2 - 4|g_{LR}^V|^2, \\ g'_4 &\equiv |g_{RL}^V|^2 + |g_{LL}^V|^2 - |g_{RR}^V|^2 - |g_{LR}^V|^2. \end{aligned} \quad (19)$$

In an analogous way to Eqs. (11) and (14), one can define $\mathcal{A}'_{FB}(x)$ and $\mathcal{A}_{FB}(y)$.

While we will not discuss the **T**-type operators separately, the double differential decay distribution is given by,

$$\frac{dB_\tau}{dx d(\cos\theta)} = \frac{T_\tau m_\tau^5}{128 \times 4\pi^3 \Lambda^4} 9x^2(1-x)[g_1 + g_2 \cos\theta], \quad (20)$$

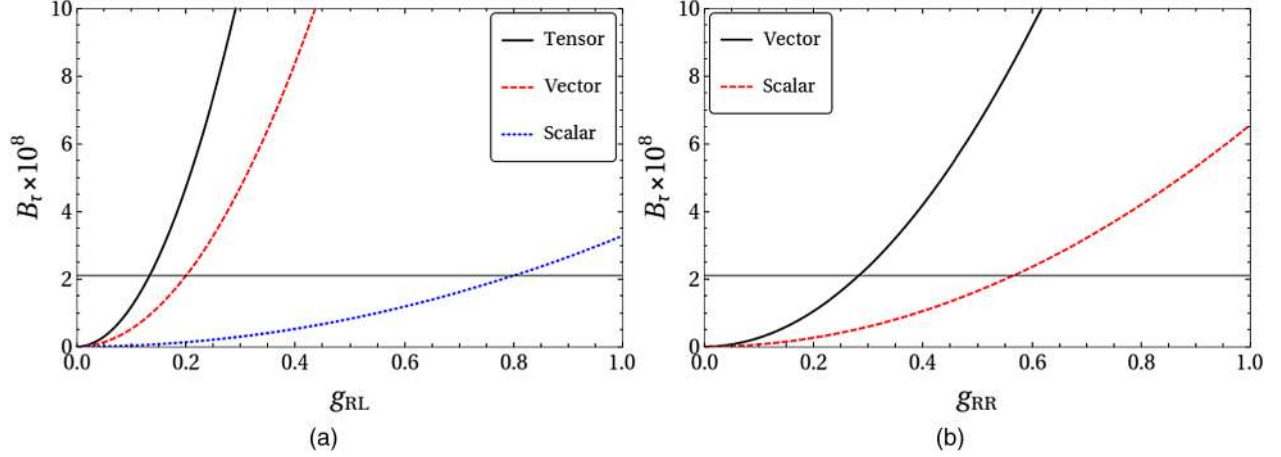


FIG. 1. (a) Variation of $\text{BR}(\tau \rightarrow 3\mu)$ with the WCs g_{RL}^I ($I = S, V, T$). (b) The same for g_{RR}^I ($I = S, V$). The horizontal line shows the present limit. The results for g_{LR}^I are identical to those for g_{RL}^I , and the results for g_{LL}^I are identical to those for g_{RR}^I .

where,

$$g_1 \equiv |g_{RL}^T|^2 + |g_{LR}^T|^2, \quad g_2 \equiv |g_{RL}^T|^2 - |g_{LR}^T|^2. \quad (21)$$

Thus

$$B_\tau = \frac{3T_\tau m_\tau^5}{128 \times 8\pi^3 \Lambda^4} g_1. \quad (22)$$

IV. ANALYSIS

In this section, we discuss the current and future sensitivities of the **S**, **V**, and **T** operators on the observables B_τ , A'_{FB} , and \mathcal{A}'_{FB} . The next subsection deals with the simplified cases where only one operator is considered to be present at a time. While this is instructive and sheds a lot of light on the differentiating power of the observables, a more realistic scenario might involve more than one operators. Thus, in the next section, we discuss the cases where two operators are simultaneously present, and try to see whether those two-operator models can be separated from the single-operator ones.

A. One operator models

Let us assume, to start with, that only one out of the ten possible operators shown in Eq. (2) is present, notwithstanding the fact that not all of them are mutually independent, some being related to the others through Fierz rearrangement. In Fig. 1, we show how the branching ratio B_τ depends on the WCs $g_{RL}^{S,V,T}$ and $g_{RR}^{S,V}$. Identical plots are obtained if one replaces LR with RL , and RR with LL . This $R \leftrightarrow L$ symmetry is true for all subsequent observables and their distributions, which reduces possible independent cases worth discussing by a factor of 2.

Given the combination IJ , if theory tells us the approximate magnitude of the WC g_{IJ}^X , even with the number of events as the sole observable, one can almost immediately

differentiate $X = S$ case with $X = V$ or T . With higher statistics even a differentiation between **V** and **T** may be possible. The present limit on B_τ is translated to $|g_{RL}^T| \leq 0.13$, $|g_{RL}^V| \leq 0.20$, $|g_{RL}^S| \leq 0.80$, $|g_{RR}^V| \leq 0.28$, and $|g_{RR}^S| \leq 0.57$. However, as we do not have any *a priori* knowledge of the magnitudes of the WCs, we have to look for some other observables and use the number of events as a normalization. In other words, we will assume that the total number of events has been given to the community by the experimentalists and see how much extra information we can extract.

In Fig. 2, we show how the differential rate dB_τ/dx varies with the muon energy variable x for a fixed value of the WC, set at 0.1. With the normalization included, the area under the curve gives the total number of events in different x -bins. Note that due to possible paucity of events, one may have only a few bins, 2 or 3, before the data starts thinning out too much to have any statistical significance. Thus, the continuous distribution showed in Fig. 2 is an idealized scenario. Even then, what we find is that the number of events will be markedly different for different classes of operators if the WCs are of the same order, which is very much along the expected line. On the other hand, the asymmetry variables $A'_{FB}(x)$ or $\mathcal{A}'_{FB}(y)$ must show identical pattern for all operators, **S**, **V**, or **T**, with a fixed chirality structure, because the overall normalization cancels in the ratio.

The next task would be to differentiate among the various chiral subclasses of a particular class of model. For illustration, we will take the **S** class of models, and consider the presence of one **S**-type operators at one time. As the sensitivities are higher for **V** and **T** classes, whatever results one has for **S** will only be more enhanced and pronounced for other classes. At the same time, if the underlying theory predicts g_{IJ} values of the order of unity, very small number of events will be harder to sustain under **V** or **T** classes.

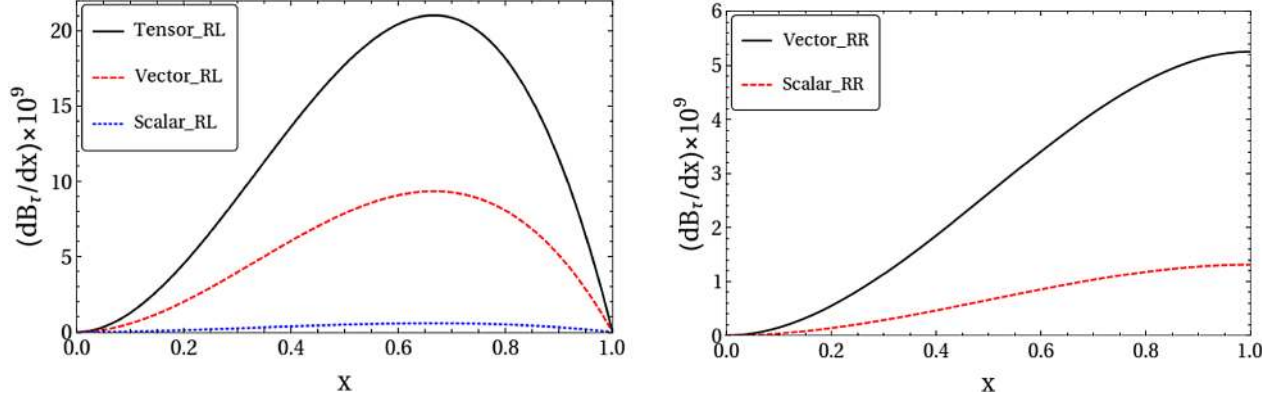


FIG. 2. The decay rate distribution dB_τ/dx for different operators, with the relevant $g_{IJ}^{S,V,T} \approx 0.1$.

In the single-coupling scheme, we consider four different models, depending upon which operator contributes, and denote them as

$$\begin{aligned} \text{Model 1: } & O_{RL}^S, & \text{Model 2: } & O_{LL}^S, \\ \text{Model 3: } & O_{RR}^S, & \text{Model 4: } & O_{LR}^S. \end{aligned} \quad (23)$$

If only one operator contributes, $A'_{FB}(x)$ becomes a function of x only and does not depend on the magnitude of the WCs:

$$\begin{aligned} A'_{FB}(x)_1 &= -A'_{FB}(x)_4 = 6(x^2 - x^3), \\ A'_{FB}(x)_2 &= -A'_{FB}(x)_3 = x^2 - 2x^3. \end{aligned} \quad (24)$$

The integrated asymmetry $A_{FB}(i)$ for the i th model can be obtained by integrating $x \in [0:1]$, and the values are

$$\begin{aligned} A_{FB}(1) &= -A_{FB}(4) = \frac{1}{2}, \\ A_{FB}(2) &= -A_{FB}(3) = -\frac{1}{6}. \end{aligned} \quad (25)$$

There is a zero crossing only for models 2 and 3 at $x = \frac{1}{2}$.

Similarly, for the forward-backward asymmetry $\mathcal{A}'_{FB}(y)$, we find

$$\begin{aligned} \mathcal{A}'_{FB}(y)_1 &= -\mathcal{A}'_{FB}(y)_4 = y^2 - 2y^3, \\ \mathcal{A}'_{FB}(y)_2 &= -\mathcal{A}'_{FB}(y)_3 = \frac{1}{2}(7y^2 - 8y^3). \end{aligned} \quad (26)$$

While all of them show zero-crossing, for the last two models such crossing occurs almost at the end of the kinematic range at $y = 7/8$. The integrated asymmetries are (for $y \in [0:1]$)

$$\mathcal{A}_{FB}(1) = \mathcal{A}_{FB}(3) = -\mathcal{A}_{FB}(2) = -\mathcal{A}_{FB}(4) = -\frac{1}{6}. \quad (27)$$

The $A'_{FB}(x)$ s for different models are shown in Figs. 3(a) and 3(b). Similarly, $\mathcal{A}'_{FB}(y)$ s are shown in Figs. 3(c) and 3(d). In these figures, every theoretical line has broadened out to a thick band, often overlapping with each other. This happens because the number of the events is limited. For every x (y), the error margin in $A'_{FB}(x)$ is approximately given by

$$\delta A'_{FB} = \sqrt{\left(\frac{\partial A'_{FB}(x)}{\partial N_F(x)}\right)^2 (\delta N_F(x))^2 + \left(\frac{\partial A'_{FB}(x)}{\partial N_B(x)}\right)^2 (\delta N_B(x))^2 + \left(\frac{\partial A'_{FB}(x)}{\partial N}\right)^2 (\delta N)^2} \quad (28)$$

where

$$\delta N_F(x) = \sqrt{N_F(x)}, \quad \delta N_B(x) = \sqrt{N_B(x)}, \quad (29)$$

are the statistical errors in the number of events in the forward and backward directions respectively, and $\delta N = \sqrt{N}$.

The expression for $\mathcal{A}'_{FB}(y)$ is analogous. We have not considered the correlation between $N_F(x)$ and $N_B(x)$; depending on the sign of the correlation, the expression can be an overestimation or underestimation, but as we do not have any *a priori* knowledge of the distribution, it is better to stick to zero correlation. The bands in Fig. 3 indicate the 1σ error margins. Clearly, the resolving power is much less for 20 events than with 50 events.

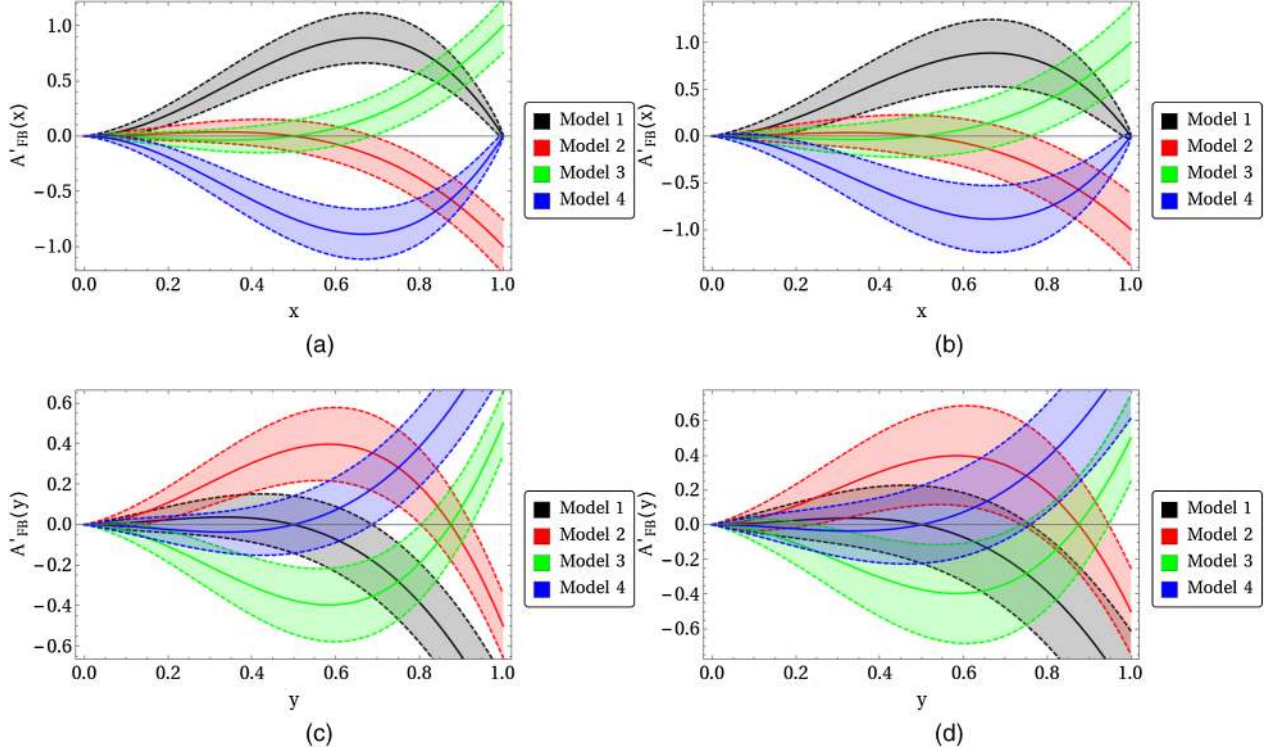


FIG. 3. $A'_{FB}(x)$ for the antimuon with (a) 50 and (b) 20 events for the four single coupling \mathbf{S} class of models. The same for the more energetic of the two muons, $A'_{FB}(y)$, with (c) 50 and (d) 20 events.

Because of the probable paucity of events, the asymmetries may be measured only with a limited number of bins. But even with two bins, low- x ($0 < x < 0.5$) and high- x ($0.5 < x < 1$), one should be able to differentiate between competing models.

The existing bound on $\tau \rightarrow 3\mu$ comes from the analysis of 782 fb^{-1} data from the Belle collaboration [15], and 468 fb^{-1} data from the BABAR collaboration [20]. With a production cross-section of 0.919 nb for the $\tau^+\tau^-$ pairs, one gets 720 million such pairs at Belle and 420 million at BABAR. For 50 ab^{-1} of integrated luminosity at Belle-II, one expects $N_p = 4.6 \times 10^{10}$ $\tau^+\tau^-$ pairs. With a detection efficiency of 7.6% [15], and using the present bound given in Eq. (1), the maximum number of such events is about 73. For our discussion, we will use two scenarios: one with $N = 20$ and the other with $N = 50$. Note that the errors are only statistical in nature. There may be other uncertainties, like fixing the direction of the τ polarization, which will widen up the bands, but that effect is expected to be small with the τ detection ability of Belle-II.

If A_{FB} turns out to be positive (negative), the viable models are 1 or 3 (2 or 4). Similarly, the positive (negative)- A_{FB} models are 2 and 4 (1 and 3). Thus, measurement of only the sign of these asymmetries leave us with a twofold ambiguity. However, $x(y)$ -dependent asymmetry measurements have the ability to resolve the same; \mathbf{S} and \mathbf{V} -type models behave identically. It is enough to measure the asymmetry in two bins: low $x(y)$ bin for $0 \leq x(y) < \frac{1}{2}$ and

high $x(y)$ bin for $\frac{1}{2} \leq x(y) \leq 1$. If these measurements were precise enough, it would have been sufficient not only to pinpoint the model but also to explore whether more than one operators are contributing.¹ Unfortunately, with a limited number of events, the measurements cannot be that precise.

Our results dB_τ/dx and dB_τ/dy in the single-coupling schemes are shown in Fig. 4. The lines broaden out into bands if we take the errors and uncertainties into account. Such broadening, in all probability, will make the lines indistinguishable from one another. However, all these models can be separated from each other from the asymmetry measurements, particularly in the high- $x(y)$ bin.

V. TWO OPERATOR MODELS

Once we establish that given enough events (~ 50), it will be straightforward to differentiate between several one-operator models, the next question is: what if the data is not compatible with any of them? Note that for the single-operator scheme to stand good, all the observables, and not only a few of them, have to be in the right ballpark specified by that model.

¹We, however, show a continuous $x(y)$ -distribution of the asymmetries; one just needs to integrate over the respective bin to have an idea of the relative magnitudes.

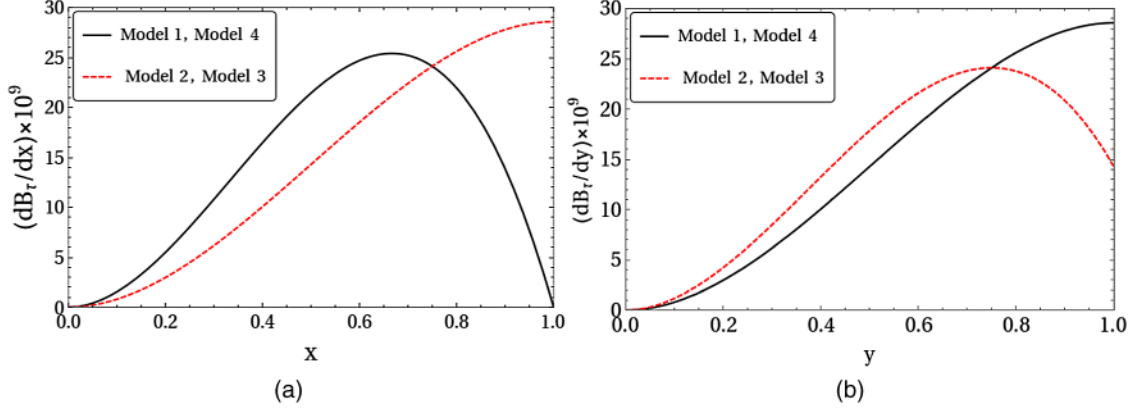


FIG. 4. (a) dB_τ/dx and (b) dB_τ/dy for the four single coupling S class of models.

However, even the principle of Occam’s razor may not be enough reason for not invoking the double-operator scheme. We will, as before, be confined within the S class of models, and consider the cases where two WCs are nonzero at a time. As we have shown, this will be the case if the underlying theory forces the muon current to be pure S or P .² Thus, the question we ask is: If the new physics is described by two $\tau \rightarrow 3\mu$ operators, with what confidence level can we differentiate that from those cases where only one of them is present? Note that the number of events will act as the tightest constraint on the parameter space. We will try to differentiate these models, hereafter called $O2$ for two effective operators, from that with only one operator, which we call the “seed” model. For example, we consider the following $O2$ models:

$$\begin{aligned}
 \text{Model A: } & O_{RL}^S \quad \text{and} \quad O_{LL}^S, \\
 \text{Model B: } & O_{RL}^S \quad \text{and} \quad O_{RR}^S, \\
 \text{Model C: } & O_{RL}^S \quad \text{and} \quad O_{LR}^S.
 \end{aligned} \tag{30}$$

We will compare the differentiability of these models (A, B, and C) with the seed model with one operator O_{RL}^S . To achieve this goal, we need to find the parameter space spanned by g_{RL}^S and another g_{JJ}^S , which depends on what $O2$ model we consider. We further check with what confidence level the allowed regions for models A, B, and C can be separated from the single-operator model given by O_{RL}^S . To complete the study, we include three further $O2$ models, namely,

$$\begin{aligned}
 \text{Model D: } & O_{RR}^S \quad \text{and} \quad O_{RL}^S, \\
 \text{Model E: } & O_{RR}^S \quad \text{and} \quad O_{LL}^S, \\
 \text{Model F: } & O_{RR}^S \quad \text{and} \quad O_{LR}^S.
 \end{aligned} \tag{31}$$

where the first operator is treated as the seed.

Models B and D are different, because of different seeds. The seed models are chosen in such a way as to have

²And pure V or A for the V class of models.

positive A_{FB} for the opposite-sign muon for them; the negative A_{FB} models will have a corresponding relationship, which can be obtained by flipping L and R , $L \leftrightarrow R$.³ Let us mention here that the confidence interval contours will depend on the type of seed operators being considered.

For this part of the analysis, we will use the optimal observable (OO) technique. For a detailed discussion on this technique, we refer the reader to Refs. [16,17]. In the context of B decays, this method has been applied in Refs. [18,19]. The essential point of the OO technique is that this gives the optimal set of observables (which are in general functions of experimental observables) with which two points in the parameter space of different models can be differentiated with maximum efficiency. In other words, this gives the maximum possible separation, in terms of confidence level, between two points in the parameter space as a function of experimental observables. In practice, the systematic errors reduce the confidence intervals.

As has been shown in Refs. [18,19], this method is all the more useful when one does not have any experimental data; in the presence of data, one can do a maximum likelihood analysis. This also means that not all the systematic uncertainties are taken into account. Thus, OO acts more as a motivational tool to the experimentalists than as an instrument for detailed quantitative theoretical studies.

Even with only the S class of operators, the parameter space of WCs is four-dimensional. A complete analysis is not only cumbersome but also of very little help in the real-life scenario where the number of events will definitely be below 100 and therefore a fine scan of the parameter space, with a two-dimensional binning on $x(y)$ and $\cos\theta(\cos\alpha)$, will have so few events per bin as to make the analysis meaningless. The only constraint on the WCs comes from the nonobservability of the decay.

In the OO technique, one writes any observable \mathcal{O} , depending on a variable ϕ , as

³If we consider the asymmetry based on the more energetic like-sign muon, models A-F have negative asymmetry, while the corresponding $L \leftrightarrow R$ models have positive asymmetry.

$$\mathcal{O}(\phi) = \sum_i C_i f_i(\phi), \quad (32)$$

which can be generalized to a set of variables denoted by ϕ . Here, all the f_i s are independent, and C_i s are some constants. The major goal of this technique is to extract C_i s. In our case, C_i s will be functions of g_i s and g'_i s defined earlier. Our analysis can be done by defining a quantity analogous to χ^2 , such as

$$\chi^2 = \sum_{i,j} (C_i - C_i^0) V_{ij}^{-1} (C_j - C_j^0). \quad (33)$$

The C_i^0 s are called the seed values, which can be considered as model inputs. The covariance matrix V_{ij} s are defined as

$$V_{ij} = \langle \Delta C_i \Delta C_j \rangle = \frac{(M^{-1})_{ij} \sigma_T}{N},$$

with $M_{ij} = \int \frac{f_i(\phi) f_j(\phi)}{O(\phi)} d\phi.$ (34)

In the above expression of V_{ij} , $\sigma_T = \int O(\phi) d\phi$ and $N = \sigma_{\text{Prod}} \mathcal{L}_{\text{int}} \epsilon B_\tau$, as defined earlier. For a specific model, χ^2 gives the confidence level separation between the seed value C_i^0 (seed model) and the model under consideration, parametrized by C_i .

Looking at Eq. (33), it is clear that the shape of the fixed χ^2 hypersurface depends on V_{ij}^{-1} , and the centroid of that (where $\chi^2 = \chi_{\text{min}}^2 = 0$) changes with the seed values. These fixed χ^2 surfaces are what determines the separation between models essentially. Thus separation between any two models 1 and 2, with seed at 1, will in general be not equal to the separation when the seed is at 2. This is the reason for treating Models B and D separately in Eqs. (30) and (31).

In the case of single operator model, the seed values of the WCs corresponding to the 50 events are obtained as

$$|g_{RL}^S|^2 = 0.44 \text{ (A, B, C)}, \quad |g_{RR}^S|^2 = 0.22 \text{ (D, E, F)}. \quad (35)$$

For the negative A_{FB} models, one may take $|g_{LR}^S|^2 = 0.44$ and $|g_{LL}^S|^2 = 0.22$. Our additional inputs are

$$m_\tau = 1.78 \text{ GeV}, \quad T_\tau = 290.3 \text{ fs}, \quad \Lambda = 5 \text{ TeV}. \quad (36)$$

We show our results for the **S** class of models; **V** class of models will show identical results. The observables that we use are $A'_{FB}(x)$, $\mathcal{A}'_{FB}(y)$ (both defined in the previous section), as well as dB_τ/dx and dB_τ/dy , the expressions for which can be obtained from Eqs. (5) and (13):

$$\frac{dB_\tau}{dx} = \frac{m_\tau^5 T_\tau}{128 \times 24\pi^3 \Lambda^4} (3x^2 g_1 - x^3 g_2),$$

$$\frac{dB_\tau}{dy} = \frac{m_\tau^5 T_\tau}{128 \times 48\pi^3 \Lambda^4} (3y^2 g'_1 - 2y^3 g'_2). \quad (37)$$

From Eq. (11), one can write

$$A'_{FB}(x) = \frac{1}{B_\tau} \times T_\tau \frac{m_\tau^5}{128 \times 48\pi^3 \Lambda^4} (x^2 g_3 - x^3 g_4). \quad (38)$$

For 50 events, $B_\tau = 1.43 \times 10^{-8}$. Similarly,

$$\mathcal{A}'_{FB}(y) = \frac{1}{B_\tau} \times T_\tau \frac{m_\tau^5}{128 \times 96\pi^3 \Lambda^4} (y^2 g'_3 - 2y^3 g'_4). \quad (39)$$

Before we show our results for all the 6 models, let us mention a few important points here.

- (1) The determination of χ^2 involves an integration over the variable ϕ of Eq. (32). If over the region of integration the observable for the seed model becomes zero for any value of ϕ , the integration diverges. Thus, one has to cut off such badly behaving regions. For example, if the observable for the seed model becomes zero at the end points, say a and b , one has to perform the integral between $a + \epsilon$ and $b - \epsilon$, where ϵ is taken to be so small as not to affect the observable (like, say, the number of events). More concrete examples are given below.
- (2) One may ask why we do not use a two-variable analysis and use the double differential cross section as the observable. This would have certainly been useful, and more powerful as an analytical tool, if we could manage a large number of events so that even the two-dimensional bins have enough number of events. With a small number of events, such an analysis would not give much useful information.

All our observables depend on only two functions, f_1 and f_2 , with the argument being x for the opposite-sign muon and y for the like-sign more energetic muon. Depending on the observables, the combinations C_1 and C_2 are as follows.

A. Observable: $A'_{FB}(x)$

We show the coefficients C_1 and C_2 in Table I, and Fig. 5 displays our results for Models A-F. As the plots are not self-explanatory, let us clearly specify what they mean. The diagonal band with negative slope, in each of the plots,

TABLE I. C_1 and C_2 for different models. The observable is $A'_{FB}(x)$.

Model	Seed	Second operator	C_1	C_2
A	RL	LL	$3 g_{RL}^S ^2 + g_{LL}^S ^2$	$3 g_{RL}^S ^2 + 2 g_{LL}^S ^2$
B	RL	RR	$3 g_{RL}^S ^2 - g_{RR}^S ^2$	$3 g_{RL}^S ^2 - 2 g_{RR}^S ^2$
C	RL	LR	$3 g_{RL}^S ^2 - 3 g_{LR}^S ^2$	$3 g_{RL}^S ^2 - 3 g_{LR}^S ^2$
D	RR	RL	$- g_{RR}^S ^2 + 3 g_{RL}^S ^2$	$-2 g_{RR}^S ^2 + 3 g_{RL}^S ^2$
E	RR	LL	$- g_{RR}^S ^2 + g_{LL}^S ^2$	$-2 g_{RR}^S ^2 + 2 g_{LL}^S ^2$
F	RR	LR	$- g_{RR}^S ^2 - 3 g_{LR}^S ^2$	$-2 g_{RR}^S ^2 - 3 g_{LR}^S ^2$

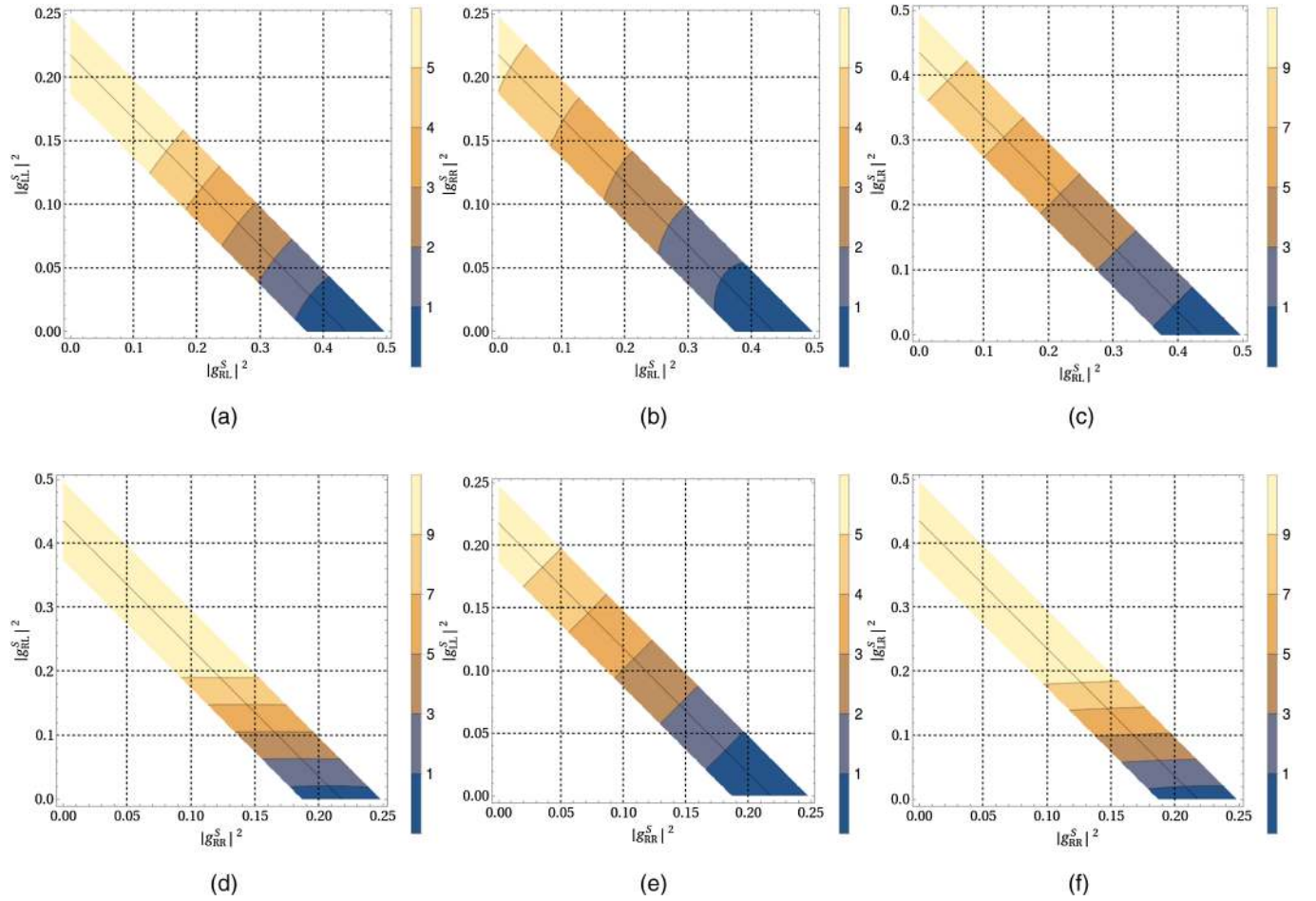


FIG. 5. The differentiability of the models A-F, shown in (a)-(f) respectively, from the “seed” model, with $A'_{FB}(x)$ as the observable.

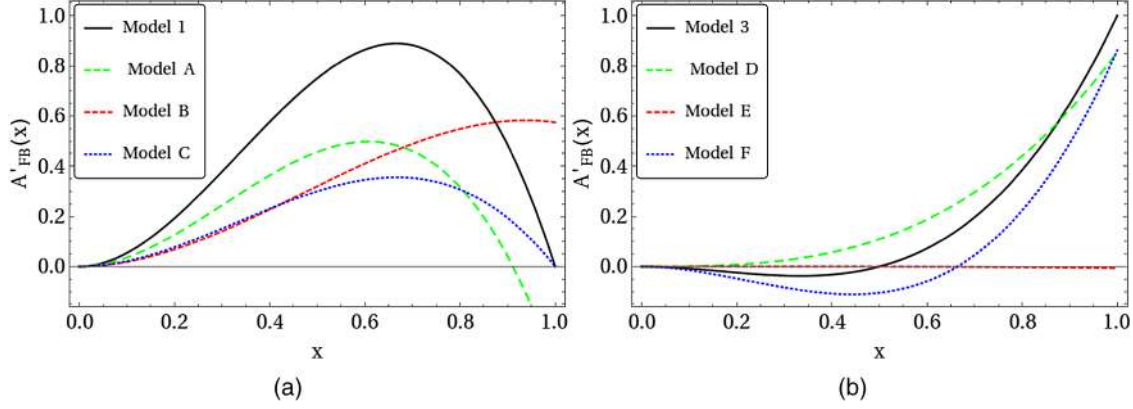
represents the allowed region in the parameter space of the various $O2$ models. Only the two relevant WCs are taken to be nonzero, keeping the others fixed at zero. Once the experimentalists obtain a certain number of events, this will specify a line in the two-dimensional parameter space over which the allowed models, each of them specified by some WCs, may lie. The exact position of the line will depend on what model one chooses, but the analysis must take into account the constraint imposed by this line.⁴ The uncertainties in the data will broaden the line to a band, whose width will ultimately depend on the number of events as well as the detector parameters. As a very rough guess, we take \sqrt{N} to be the width for the line with N events. The plots are drawn for $N = 50$; thus, the band includes all the points for which the number of events lie approximately between 43 and 57. The separation contours are drawn on these bands only. We expect the bands to be narrower in actual experiments.

Let us consider Fig. 5(a). This takes $|g_{RL}^S|^2 = 0.435$ as the seed value. The plot tells us that this one coupling

⁴That is why even for a two-parameter model the degree of freedom is only 1.

model can be differentiated from the one with $|g_{RL}^S|^2 = 0.2$ and $|g_{LL}^S|^2 = 0.1$ at more than 3σ if we have approximately 50 events and use $A'_{FB}(x)$ as our observable. Similarly, the model with $|g_{RL}^S|^2 = 0.35$ and $|g_{LL}^S|^2 = 0.05$ can be separated from the above mentioned seed model by less than 2σ . The actual numbers should be even worse as the systematic uncertainties will also creep in. Similar conclusions hold for Models B, C, D, E, and F, for which the results are shown from Figs. 5(b)–(f), respectively. As we mentioned before, contours for Models B and D are not the same, although they involve the same set of operators. This is because the seed is different, which ultimately controls the correlation matrix.

For models D-F, the seed model has a zero crossing for $A'_{FB}(x)$ at $x = \frac{1}{2}$. Unlike in the case of the differential decay distribution and observables proportional to it, the integrated observable in this case may become negative in different parts of the parameter space. This makes the covariance matrix V_{ij} not positive definite. We note here that for our purpose, i.e., to construct χ^2 , the integrated observables serve only as the normalization of V_{ij}^{-1} ($=M_{ij}$). We have taken the modulus of the integrand for each value of $x(y)$ for this reason. This, while keeping the nature of the


 FIG. 6. Comparison of $A'_{FB}(x)$ for the $O2$ Models A-F with the respective seed model. The relevant WCs are taken from Fig. 5.

error ellipsoids intact, will always keep the covariance matrix positive definite. On the flip-side, this makes the integral diverge at the zero crossing point. Thus, to evaluate the correlation matrix V_{ij}^{-1} and to cancel this divergence, one has to remove the tiny patch $0.495 < x < 0.505$ from integration. This has only a negligible effect on the number of events, but keeps the necessary integrations convergent.

In Fig. 6, we show, as an illustration, the behavior of $A'_{FB}(x)$ for Models A-C vis-a-vis the seed Model 1 and Models D-F with seed Model 2, for which the differentiability is at the 3σ level. The corresponding WCs, extracted from Fig. 5, are displayed in Table II. We note that Models A-C can be differentiated from the seed Model 1 with only $|g_{RL}^S|^2$ for all values of x . On the other hand, Models D and F can be differentiated from seed Model 2 with only $|g_{RR}^S|^2$ (Model 3) only for medium values of x , and zero-crossing of $A'_{FB}(x)$ plays a crucial role.

B. Observable: $\mathcal{A}'_{FB}(y)$

In an analogous way, one can use the more energetic of the like-sign muons, and the corresponding asymmetry $\mathcal{A}'_{FB}(y)$. The coefficients C_1 and C_2 , from Eqs. (13) and (14), are shown in Table III.

In Fig. 8, we show the distribution of $\mathcal{A}'_{FB}(y)$ for Models A-F, comparing A-C with Model 1 as seed and D-F with

Model 3 as seed respectively. The corresponding WCs are given in Table IV. While a 3σ separation between the models is possible, one notes that the differentiation works best in the middle- y region, rather than at the endpoints.

C. Observable: dB_τ/dx and dB_τ/dy

Study of the differential BRs is instructive. First, let us refer the reader to Tables V and VI respectively for the coefficients C_1 and C_2 in all the models considered. For both dB_τ/dx and dB_τ/dy , this shows immediately that Models A and B must yield identical distributions; same is true for the pair D and F. This is because the BR does not

 TABLE III. C_1 and C_2 for different models, with $\mathcal{A}'_{FB}(y)$ as the observable.

Model	Seed	Second operator	C_1	C_2
A	RL	LL	$ g_{RL}^S ^2 + 7 g_{LL}^S ^2$	$ g_{RL}^S ^2 + 4 g_{LL}^S ^2$
B	RL	RR	$ g_{RL}^S ^2 - 7 g_{RR}^S ^2$	$ g_{RL}^S ^2 - 4 g_{RR}^S ^2$
C	RL	LR	$ g_{RL}^S ^2 - g_{LR}^S ^2$	$ g_{RL}^S ^2 - g_{LR}^S ^2$
D	RR	RL	$-7 g_{RR}^S ^2 + g_{RL}^S ^2$	$-4 g_{RR}^S ^2 + g_{RL}^S ^2$
E	RR	LL	$-7 g_{RR}^S ^2 + 7 g_{LL}^S ^2$	$-4 g_{RR}^S ^2 + 4 g_{LL}^S ^2$
F	RR	LR	$-7 g_{RR}^S ^2 - g_{LR}^S ^2$	$-4 g_{RR}^S ^2 - g_{LR}^S ^2$

 TABLE II. The WCs, as obtained from Fig. 5, for which Models A-F are separable from the respective seed models at 3σ level. The integrated asymmetries are also shown.

Model	$ g_{RL}^S ^2$	$ g_{LL}^S ^2$	$ g_{RR}^S ^2$	$ g_{LR}^S ^2$	A_{FB} (int.)
1(Seed)	0.435	0.5
A	0.265	0.085	0.239
B	0.186	...	0.125	...	0.309
C	0.305	0.131	0.2
3(Seed)	0.218	...	0.167
D	0.063	...	0.186	...	0.215
E	...	0.110	0.108	...	-0.001
F	0.188	0.061	0.073

 TABLE IV. The WCs, as obtained from Fig. 7, for which Models A-F are separable from the respective seed models at 3σ level. The integrated asymmetries are also shown.

Model	$ g_{RL}^S ^2$	$ g_{LL}^S ^2$	$ g_{RR}^S ^2$	$ g_{LR}^S ^2$	A_{FB} (int.)
1(Seed)	0.435	-0.167
A	0.314	0.061	-0.074
B	0.313	...	0.061	...	-0.167
C	0.216	0.219	0.001
3(Seed)	0.218	...	-0.167
D	0.187	...	0.124	...	-0.167
E	...	0.099	0.119	...	-0.016
F	0.121	0.194	-0.018

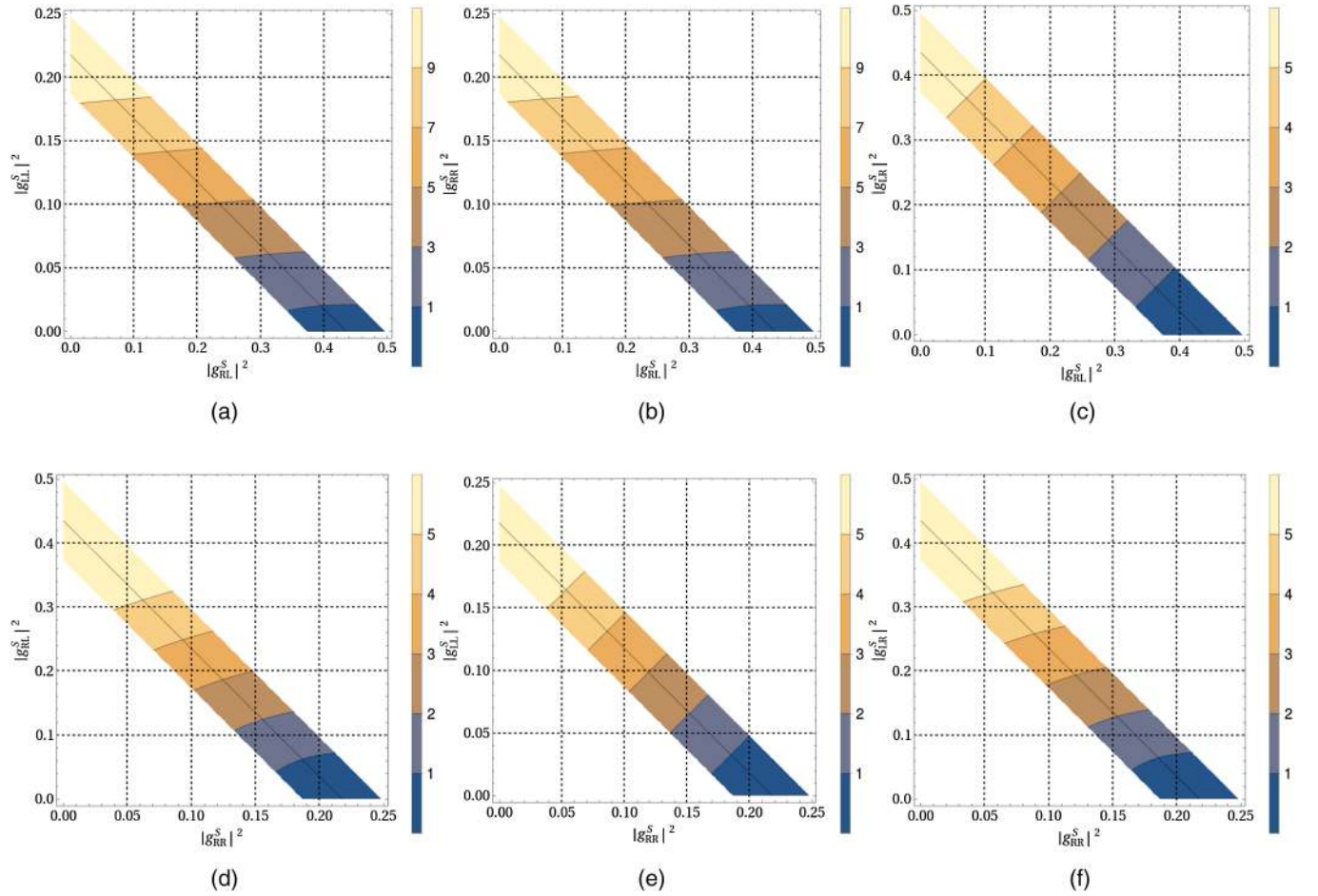


FIG. 7. The differentiability of the models A-F, shown in (a)-(f) respectively, from the seed model, with $\mathcal{A}'_{FB}(y)$ as the observable.

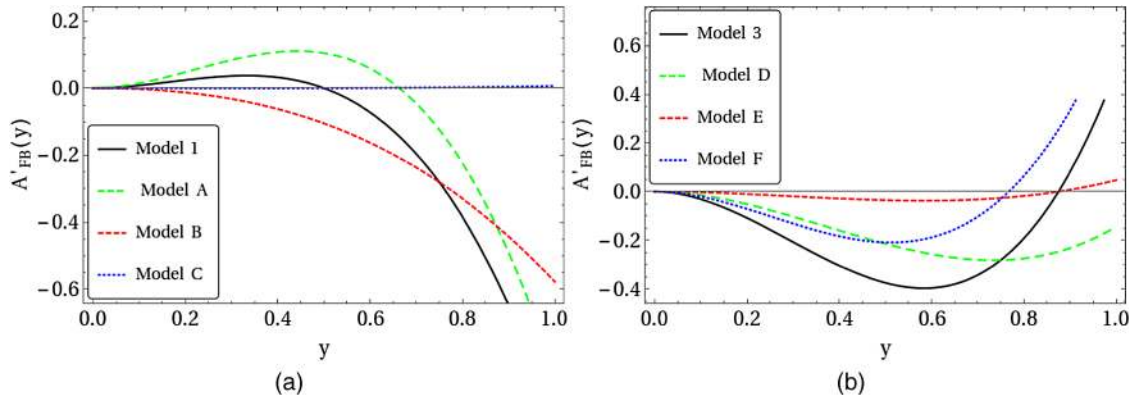


FIG. 8. Comparison of the distributions of the asymmetry $\mathcal{A}'_{FB}(y)$ for the models A-F from the respective seed (y) model. Here the relevant WCs are taken from the plots in Fig. 7.

depend on the change $R \leftrightarrow L$. Models C and E are very poorly differentiable from their respective seeds (at less than 1σ) and so we do not discuss them any further, neither do we show the corresponding separation plots.

Even though the pattern seems similar, there is an important difference. With dB_τ/dx as the observable, we can separate Models A(B) or D(F) from the corresponding

seed models at 3σ or more, depending on the respective WCs. This can be seen from Fig. 9, as well as Table VII. With dB_τ/dy as the observable, there is no available parameter space with 50 events where any model can be separated at more than 2σ from the seed models. This is why we do not show the corresponding plots for dB_τ/dy . Thus, as far as the measurement of the number of events in

TABLE V. C_1 and C_2 for different models. The observable is dB_τ/dx .

Model	Seed	Second operator	C_1	C_2
A	RL	LL	$ g_{RL}^S ^2 + g_{LL}^S ^2$	$3 g_{RL}^S ^2 + 2 g_{LL}^S ^2$
B	RL	RR	$ g_{RL}^S ^2 + g_{RR}^S ^2$	$3 g_{RL}^S ^2 + 2 g_{RR}^S ^2$
C	RL	LR	$ g_{RL}^S ^2 + g_{LR}^S ^2$	$3 g_{RL}^S ^2 + 3 g_{LR}^S ^2$
D	RR	RL	$ g_{RR}^S ^2 + g_{RL}^S ^2$	$2 g_{RR}^S ^2 + 3 g_{RL}^S ^2$
E	RR	LL	$ g_{RR}^S ^2 + g_{LL}^S ^2$	$2 g_{RR}^S ^2 + 2 g_{LL}^S ^2$
F	RR	LR	$ g_{RR}^S ^2 + g_{LR}^S ^2$	$2 g_{RR}^S ^2 + 3 g_{LR}^S ^2$

TABLE VI. C_1 and C_2 for different models. The observable is dB_τ/dy .

Model	Seed	Second operator	C_1	C_2
A	RL	LL	$ g_{RL}^S ^2 + 3 g_{LL}^S ^2$	$ g_{RL}^S ^2 + 4 g_{LL}^S ^2$
B	RL	RR	$ g_{RL}^S ^2 + 3 g_{RR}^S ^2$	$ g_{RL}^S ^2 + 4 g_{RR}^S ^2$
C	RL	LR	$ g_{RL}^S ^2 + g_{LR}^S ^2$	$ g_{RL}^S ^2 + g_{LR}^S ^2$
D	RR	RL	$3 g_{RR}^S ^2 + g_{RL}^S ^2$	$4 g_{RR}^S ^2 + g_{RL}^S ^2$
E	RR	LL	$3 g_{RR}^S ^2 + 3 g_{LL}^S ^2$	$4 g_{RR}^S ^2 + 4 g_{LL}^S ^2$
F	RR	LR	$3 g_{RR}^S ^2 + g_{LR}^S ^2$	$4 g_{RR}^S ^2 + g_{LR}^S ^2$

different energy bins goes, it is preferable to detect the unlike-sign muon, than one of the like-sign muons.

As the V class of models show an identical behavior, we conclude that based on only the data and without any *a priori* knowledge of the WCs, it is impossible to differentiate between the classes, but within a particular class, it is possible to differentiate among the various Lorentz structures of the effective operators. With enough events, one should be able to differentiate single-operator models from the double-operator models, like those with pure S or V (O_9), or with pure P or A (O_{10}) muon current. If we have approximately 50 events, $A'_{FB}(x)$ may help us differentiate

TABLE VII. The representative values of the WCs, obtained from Fig. 9, for which Models A and D can be differentiated from the respective seed models at 3σ level.

Model	$ g_{RL}^S ^2$	$ g_{LL}^S ^2$	$ g_{RR}^S ^2$	$ g_{LR}^S ^2$
1(Seed)	0.435
A	0.202	0.117
3(Seed)	0.218	...
D	0.372	...	0.032	...

O_{9L} or O_{10L} models from RL by about 5σ , and O_{9R} and O_{10R} models from RR by about 7σ . With dB_τ/dx , the former set is differentiable at about 3σ , while the latter is at less than 2σ . (The 9(10)L(R) models are specified by equal magnitudes of the two WCs.) As C_1 and C_2 involve $|g_{IJ}^S|^2$, and hence it is insensitive to the sign or phase of the WCs.

In general, LFV models can also involve electrons in the final state, from operators leading to $\tau^- \rightarrow e^-e^+e^-$ or $\tau^- \rightarrow e^-e^+\mu^-$. The BRs of these channels have bounds comparable to that of $\tau \rightarrow 3\mu$, so we may expect a similar number of events at Belle-II, and a similar analysis will work. However, detection of final state electrons in an e^-e^+ machine will have lesser efficiency than that of final state muons.

VI. CONCLUSION

In this paper, we focus on the LFV decay $\tau \rightarrow 3\mu$. This is of crucial importance in the light of semileptonic B-decay anomalies, which hint at some new physics involving second and third generation leptons, probably a mixing among the charged leptons. The present limit on this mode translates to ~ 70 events at the most at Belle-II with 50 ab^{-1} integrated luminosity. While even a single event will unequivocally indicate new physics, we try to answer a more ambitious question: is it possible to say anything about the underlying operators from the observables? Needless to say, the answer will be vital for model builders.

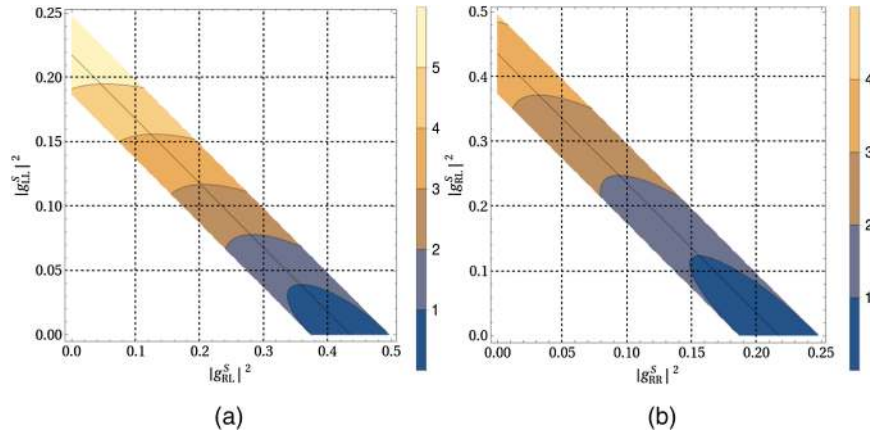


FIG. 9. The differentiability of the Models A and D, shown in (a) and (b) respectively, from the seed model, with dB_τ/dx as the observable. Model B is identical with Model A, and Model F is identical with Model D.

The answer to this question would have been much easier if final state muon polarizations could have been measured. As far as the present technologies go, this is not easily attainable. However, as we show, one can form other observables, which are relatively clean and at the same time can yield significant information. One of the observables is the asymmetry of either the unlike-sign muon, or the like-sign more energetic muon, which is to be measured with respect to the initial τ -polarization direction. If one can measure the asymmetries, even with the associated error margins, in different energy bins, this can differentiate between the different types of operators in a particular class (scalar, vector, or tensor).

Another important observable, as expected, is the number of events in different energy bins, either the unlike-sign or the like-sign muons. Just like the asymmetries, it can potentially differentiate among the different chiral structures of the operators, although to a lesser extent. Given the total number of events, one can also have an idea of the magnitude of the relevant WCs. We expect more events for V or T type operators, so their WCs, g_{IJ}^V or g_{IJ}^T , can be probed better.

It may so happen that there are more than one NP operators. A typical case is when the muon current is purely vector or axial-vector in nature. If we have enough number of events (~ 50), we should be able to say whether there is only one underlying operator or two. Asymmetries are the better observables, but the distribution of the number of events can also help and act as complementary ones.

One must, however, remember that such an analysis involves the risk of underestimating the errors by neglecting the systematic uncertainties. Thus, this is to be seen more as a motivation to the experimentalists. Once the data is available, other powerful analysis methods, like the maximum likelihood, can be applied.

ACKNOWLEDGMENTS

A. K. thanks the Science and Engineering Research Board (SERB), Government of India, for a research grant. S. K. P. is supported by the Grants No. IFA12-PH-34 and No. SERB/PHY/2016348 by Government of India.

-
- [1] A comprehensive list of all LFV processes and their limits in present and future experiments can be found in, e.g., M. Lindner, M. Platscher, and F. S. Queiroz, A call for new physics: The muon anomalous magnetic moment and lepton flavor violation, [arXiv:1610.06587](https://arxiv.org/abs/1610.06587); J. Heeck, Interpretation of lepton flavor violation, *Phys. Rev. D* **95**, 015022 (2017).
- [2] D. Choudhury, A. Kundu, R. Mandal, and R. Sinha, Minimal Unified Resolution to $R_{K^{(*)}}$ and $R(D^{(*)})$ Anomalies with Lepton Mixing, *Phys. Rev. Lett.* **119**, 151801 (2017).
- [3] D. Choudhury, A. Kundu, R. Mandal, and R. Sinha, $R_{K^{(*)}}$ and $R(D^{(*)})$ anomalies resolved with lepton mixing, [arXiv:1712.01593](https://arxiv.org/abs/1712.01593).
- [4] S. L. Glashow, D. Guadagnoli, and K. Lane, Lepton Flavor Violation in B Decays?, *Phys. Rev. Lett.* **114**, 091801 (2015).
- [5] V. Khachatryan *et al.* (CMS Collaboration), Search for lepton-flavour-violating decays of the Higgs boson, *Phys. Lett. B* **749**, 337 (2015).
- [6] D. Choudhury, A. Kundu, S. Nandi, and S. K. Patra, Unified resolution of the $R(D)$ and $R(D^*)$ anomalies and the lepton flavor violating decay $h \rightarrow \mu\tau$, *Phys. Rev. D* **95**, 035021 (2017).
- [7] F. Feruglio, P. Paradisi, and A. Pattori, Revisiting Lepton Flavor Universality in B Decays, *Phys. Rev. Lett.* **118**, 011801 (2017).
- [8] B. Bhattacharya, R. Morgan, J. Osborne, and A. A. Petrov, Studies of lepton flavor violation at the LHC, [arXiv:1802.06082](https://arxiv.org/abs/1802.06082).
- [9] A. G. Akeroyd, M. Aoki, and Y. Okada, Lepton flavour violating tau decays in the left-right symmetric model, *Phys. Rev. D* **76**, 013004 (2007); A. Crivellin, S. Najjari, and J. Rosiek, Lepton flavor violation in the standard model with general dimension-six operators, *J. High Energy Phys.* **04** (2014) 167; S. M. Boucenna, J. W. F. Valle, and A. Vicente, Are the B decay anomalies related to neutrino oscillations?, *Phys. Lett. B* **750**, 367 (2015).
- [10] A. Matsuzaki and A. I. Sanda, Analysis of lepton flavor violating $\tau^\pm \rightarrow \mu^\pm \mu^\pm \mu^\mp$ decays, *Phys. Rev. D* **77**, 073003 (2008).
- [11] R. Brüser, T. Feldmann, B. O. Lange, T. Mannel, and S. Turczyk, Angular analysis of new physics operators in polarized $\tau \rightarrow 3\ell$ decays, *J. High Energy Phys.* **10** (2015) 082.
- [12] A. Falkowski and K. Mimouni, Model independent constraints on four-lepton operators, *J. High Energy Phys.* **02** (2016) 086.
- [13] M. Giffels, J. Kallarackal, M. Kramer, B. O'Leary, and A. Stahl, The lepton-flavour violating decay $\tau \rightarrow \mu\mu\bar{\mu}$ at the CERN LHC, *Phys. Rev. D* **77**, 073010 (2008).
- [14] C. Hays, M. Mitra, M. Spannowsky, and P. Waite, Prospects for new physics in $\tau \rightarrow l\mu\mu$ at current and future colliders, *J. High Energy Phys.* **05** (2017) 014.
- [15] K. Hayasaka *et al.*, Search for lepton flavor violating τ decays into three leptons with 719 million produced $\tau^+\tau^-$ pairs, *Phys. Lett. B* **687**, 139 (2010).
- [16] J. F. Gunion, B. Grzadkowski, and X. G. He, Determining the Top-Antitop and Z Z Couplings of a Neutral Higgs Boson of Arbitrary CP Nature at the NLC, *Phys. Rev. Lett.* **77**, 5172 (1996).

- [17] D. Atwood and A. Soni, Analysis for magnetic moment and electric dipole moment form-factors of the top quark via $e^+e^- \rightarrow t\bar{t}$, *Phys. Rev. D* **45**, 2405 (1992).
- [18] S. Bhattacharya, S. Nandi, and S.K. Patra, Optimal-observable analysis of possible new physics in $B \rightarrow D^{(*)}\tau\nu_\tau$, *Phys. Rev. D* **93**, 034011 (2016).
- [19] Z. Calcuttawala, A. Kundu, S. Nandi, and S.K. Patra, Optimal observable analysis for the decay $b \rightarrow s$ plus missing energy, *Eur. Phys. J. C* **77**, 650 (2017).
- [20] J.P. Lees *et al.* (BABAR Collaboration), Limits on tau lepton-flavor violating decays in three charged leptons, *Phys. Rev. D* **81**, 111101 (2010).