Discovering True Muonium in $K_L \rightarrow (\mu^+\mu^-)\gamma$

Yao Ji

Institut für Theoretische Physik, Universität Regensburg, Regensburg 93040, Germany

Department of Physics, University of Maryland, College Park, MD 20742

(Dated: May 19, 2018)

Theoretical and phenomenological predictions of $\mathcal{BR}(K_L \rightarrow (\mu^+\mu^-)\gamma) \sim 7 \times 10^{-13}$ are presented for different model form factors $F_{K_L\gamma\gamma^*}(Q^2)$. These rates are comparable to existing and near-term rare $K_L$ decay searches at J-PARC and CERN, indicating a discovery of true muonium is possible.

The model uncertainties are sufficiently small that detection of true muonium could discriminate between the form factor models. Further discussion of potential backgrounds is made.

Lepton universality predicts differences in electron and muon observables should occur only due to their mass difference. Measurements of $(g-2)_\mu$ [1], nuclear charge radii [2, 3], and rare meson decays [4] have shown hints of violations to this universality. The bound state of $(\mu^+\mu^-)$, true muonium, presents a unique opportunity to study lepton universality in and beyond the Standard Model [5]. To facilitate these studies, efforts are on-going to improve theoretical predictions [6].

Lepton universality predicts differences in electron and muon observables should occur only due to their mass difference. Measurements of $(g-2)_\mu$ [1], nuclear charge radii [2, 3], and rare meson decays [4] have shown hints of violations to this universality. The bound state of $(\mu^+\mu^-)$, true muonium, presents a unique opportunity to study lepton universality in and beyond the Standard Model [5]. To facilitate these studies, efforts are on-going to improve theoretical predictions [6].

In recent years, a strong focus on rare kaon decays has developed in the search for new physics. The existing KOTO experiment at J-PARC [14] and proposed NA62-KLEVER at CERN [15] hope to achieve sensitivities of $\mathcal{BR} \sim 10^{-13}$ allowing a 1% measurement of $\mathcal{BR}(K_L \rightarrow \pi^0\nu\bar{\nu}) \sim 10^{-11}$. Malenfant was the first to propose $K_L$ as a source of $(\mu^+\mu^-)$ [9]. He estimated $\mathcal{BR}(K_L \rightarrow (\mu^+\mu^-)) \sim 5 \times 10^{-13}$ by approximating $F_{K_L\gamma\gamma^*}(Q^2 = 4M_{\mu^+\mu^-}^2) \sim F_{K_L\gamma\gamma^*}(0)$ where $Q^2$ is the off-shell photon invariant mass squared. This two-body decay is the reach of rare kaon decay searches and is an attractive process for discovering $(\mu^+\mu^-)$. The decay has simple kinematics with a single, monochromatic photon (of $E_\gamma = 203.6$ MeV if the $K_L$ is at rest) plus $(\mu^+\mu^-)$ which could undergo a two-body dissociate or decay into two electrons (with $M_{\mu^+\mu^-}^2 \sim 4M_e^2$). Another motivation for the search for this rare decay is its unique dependence on the form factor. Previous extractions of the form factor relied upon radiative Dalitz decays, $K_L \rightarrow \ell^+\ell^-\gamma$, the most recent being from the KTEV collaboration [16]. In these analyses, the phenomenological form factor is integrated over bins in $Q^2$, and fits to differential cross section data. In contrast, the $(\mu^+\mu^-)$ branching ratio gives the form factor at one $Q^2$ and fixes one-parameter form factors. Further, a measurement of $(\mu^+\mu^-)$ would help to better understand the kaon form factor through a completely different set of systematic and statistical uncertainties to the existing measurements.

In this letter, we present the $\mathcal{BR}(K_L \rightarrow (\mu^+\mu^-)\gamma)$ including full $\mathcal{O}(\alpha)$ radiative corrections and four different treatments of the form factor $F_{K_L\gamma\gamma^*}(Q^2)$, thereby avoiding Malenfant’s approximation. It is shown that the approximation underestimates the branching ratio by a model-dependent 15-60%. Possible discovery channels are discussed and brief comments on important backgrounds are made.

Following previous calculations for atomic decays of mesons [8, 9], the branching ratio can be computed

$$\frac{\mathcal{BR}(K_L \rightarrow (\mu^+\mu^-)\gamma)}{\mathcal{BR}(K_L \rightarrow \gamma\gamma)} = \frac{\alpha^4\zeta(3)}{2} \left(1 - z_{TM}\right)^3 \left[1 - \frac{0.439\alpha}{\pi}\right] |f(z_{TM})|^2,$$

(1)

where $\zeta(3) = \sum_{n} 1/n^3$ arising from the sum over all allowed $(\mu^+\mu^-)$ states, $z_{TM} = M_{TM}^2/M_K^2 \approx 4M_\mu^2/M_K^2$, and $f(z) = F_{K_L\gamma\gamma^*}(z)/F_{K_L\gamma\gamma^*}(0)$. Previous computation of radiative corrections considered only the vacuum polarization from the flavor found in the final state [8]. We have computed the full $(\mu^+\mu^-)$ results including the electronic, muonic, and hadronic vacuum polarization [6] as well as the QED process $K_L \rightarrow \gamma^*(k) + \gamma^*(P_{K_L} - k) \rightarrow \gamma + TM$ demonstrated by Fig. 1 where $P_{K_L}$ is the four-momentum of the $K_L$. In this contribution, one should take the convolution of the QED amplitude with double-virtual-photon form factor $F_{K_L\gamma\gamma^*}(k^2/M_K^2)$, $(P_{\gamma} - k)^2/M_K^2$. For our purpose, however, taking the form factor to be $F_{\gamma\gamma^*}(0, z_{TM})$ is a sufficient approximation as shown in [17]. A similar calculation for positronium, where other lepton flavors and hadronic loop corrections are negligible, finds the $\alpha/\pi$ coefficient is $-52/9$. 

1 discover true muonium via Heavy Photon Search (HPS) [10] experiment is search-
The theoretical predictions for $f(z)$ are computed as a series expansion to first order in $z$ with slope $b$. It is typically decomposed into $b = b_V + b_D$. $b_V$ arises from a weak transition from $K_L \to P$ followed by a strong-interaction vector interchange $P \to V\gamma$ and concluding with the vector meson mixing with the off-shell photon. Here, we denote with $P$ the pseudoscalars ($\pi^0, \eta, \eta'$) and with $V$ the vector mesons ($\rho, \omega, \phi$). The second term, $b_D$, arises from the direct weak vertex $K_L \to V\gamma$ which then mixes with $\gamma + \gamma'$ which requires modeling. Following [19], the predictions of $b_V$ and $b_D$ are divided into whether nonet or octet symmetry in the light mesons is assumed.

To compute $b_V$, one integrates out the vector mesons from the $P \to V\gamma$ vertex and assuming a particular pseudoscalar symmetry, the effective Lagrangian is derived and low energy constants can be used. $b_V^{\text{tet}} = 0$ at leading order due to the cancellation between $\pi^0$ and $\eta$ in the Gell-Mann-Okubo relation [20, 21]. In the nonet realization, a nonzero contribution coming from $\eta'$ yields $b_V^{\text{nonet}} = r_V M_K^2 / M_\rho^2 \sim 0.46$ [22], where $r_V$ is a model-independent parameter depending on the couplings of each decomposed meson field in the effective Lagrangian and are ultimately determined by experimental data.

For $b_D$, the derivation is more complicated and relies on models. In the naive factorization model (FM) [23], the dominant contribution to the weak vertex is assumed to be factorized current\times current operators which neglect the chiral structure of QCD. A free parameter, $k_F$, is introduced that is related to goodness of the factorized current approximation. If this factorization was exact, $k_F = 1$. In this scheme, $b_D^{\text{nonet}} = 2 b_D^{\text{tet}} = 1.41 k_F$. This model predicts the process $K_L \to n^0\gamma\gamma$ as well, and we use the unweighted average of the two most recent measurements of this process to fix $k_F = 0.55(6)$ [24, 25].

In the Bergström-Massó-Singer (BMS) model [26], the direct transition is instead assumed to be dominated by a weak vector-vector interaction ($K_L \to \gamma + K^* \to \gamma + \rho, \omega, \phi \to \gamma + \gamma'$). BMS further assumes that no $\Delta I = \frac{1}{2}$ enhancement occurs. This model produces a complete form factor:

$$f_{\gamma^*,\text{BMS}}(z) = \frac{1}{1 - M_{K^*}^2 z} + \frac{C}{1 - M_{K^*}^2 z} \left( \frac{4}{3} - \frac{1}{1 - M_{K^*}^2 z} - \frac{1}{9} - \frac{1}{1 - M_{K^*}^2 z} \right).$$

The two terms correspond to the vector interchange and direct transition, respectively. Expanding this expression in powers of $z$, we find the BMS model predicts

$$b_{\text{BMS}} = M_{K^*}^2 - \frac{1}{9} C \alpha_{K^*} \left( \frac{M_K^2}{M_{\rho}^2} + \frac{2 M_{K^*}^2}{M_{\rho}^2} + \frac{M_{\phi}^2}{M_{\rho}^2} \right)$$

$$= \alpha_{\text{BMS}} + b_{\text{D,BMS}} \quad (3)$$

Under the model assumptions, $-\alpha_{K^*}$ is theoretically estimated to be $\sim 0.2 - 0.3$ [26]. $C = 2.7(4)$ depends on a number of other mesonic decay rates [16, 27], and we used the modern values [18]. The error comes from the experimental uncertainty which is dominated by the two $K^*$ measurements. $\BR(K^* \to K^0\gamma)$ contributes $\Delta C \sim 13\%$ and $\Gamma_{K^*,\text{tot}}$ contributes $\Delta C \sim 4\%$ due to a disagreement between decay modes. This choice of $C$ and $\alpha_{K^*}$ is consistent with the measured rates for $K_L \to \ell^+\ell^-\gamma$.

D’Ambrosio et. al. advocates the view that $b_{\text{D,BMS}}$ is one of a series of contributions to $b_D$, which should be summed together with the model-independent $b_V$ [19]. They construct another contribution by factorizing the vector coupling (FMV) similar to FM but first restricting the Lagrangian to left-handed currents. For the different symmetry realizations, $b^{\text{nonet}}_D = 3.14 \eta \sim 0.66$ and $b^{\text{tet}}_D = 2.42 \eta \sim 0.51$ where $\eta$ is a coefficient multiplying the naive weak coupling $G_\ell$ and like $k_F$ is related to the quality of the factorization assumption. We use their value of $\eta = g_W^{\text{Wilson}} / |g_{K^*\tau\pi}\ell_{LO} = 0.21$. Our theoretical results are compiled in Table I. These values disagree outside their error, and a 10% precision measurement would be able to discriminate between them. This is in contrast to the radiative Dalitz decays, where the theoretical values are consistent.

The BMS form factor also has been used to phenomenologically fit $K_L \to \ell^+\ell^-\gamma$ for both $\ell =$
TABLE I. Theoretical values of $b$ and $BR(K_L \to (\mu^+\mu^-)\gamma)$ for the models considered in this paper.

| Model          | $b_{\text{theory}}$ | $BR_{TM} \times 10^{13}$ |
|----------------|----------------------|---------------------------|
| (FM)$_{\text{octet}}$ | 0.40(4)$^a$          | 5.90(9)                   |
| (FM)$_{\text{nonet}}$ | 1.24(6)$^a$          | 7.68(15)                  |
| (BMS)$_{\text{nonet}}$ | 0.76(9)              | 6.63(20)                  |
| (BMS+FM)$_{\text{octet}}$ | 0.85(10)            | 6.82(22)                  |
| (BMS+FM)$_{\text{nonet}}$ | 1.45(10)            | 8.16(25)                  |

$^a$ Using value of $k_F = 0.55(6)$ derived from $K_L \to \pi^0\gamma\gamma$[24, 25]

$e, \mu$, and $C\alpha_{K^*}$ is derived from the differential cross sections of these processes; yielding $(C\alpha_{K^*})_e = -0.517(30)_{\text{stat}}(22)_{\text{sys}}$[16] and $(C\alpha_{K^*})_\mu = -0.37(7)$[16], which are each input into our prediction for $(\mu^+\mu^-)$.

We also consider the D’Ambrosio-Isidori-Portolés (DIP) phenomenological $F_{\gamma\gamma\gamma}(z_1, z_2)$[28]:

$$f_{\gamma\gamma\gamma,\text{DIP}}(z_1, z_2) = 1 + \alpha_{\text{DIP}} \left( \frac{z_1}{z_1 - M_K^2} + \frac{z_2}{z_2 - M_K^2} \right) + \beta_{\text{DIP}} \left( \frac{z_1}{z_1 - M_K^2} \right) \left( \frac{z_2}{z_2 - M_K^2} \right),$$

(4)

where $z_1 = z_{TM}, z_2 = 0$ for $(\mu^+\mu^-)$ production. To set $\alpha_{\text{DIP}}$, we take the values from $K_L \to e^+e^-\gamma$, $\alpha_{\text{DIP}, e} = -1.729(43)_{\text{stat}}(28)_{\text{sys}}$[16], and from $K_L \to \mu^+\mu^-\gamma$, $\alpha_{\text{DIP}, \mu} = -1.54(10)$[16]. Our phenomenological results are compiled in Table II. Comparing the phenomenological form factors, they are indistinguishable within uncertainty in $(\mu^+\mu^-)$ production. This is perhaps unsurprising because they arise from the same underlying data, but the difference in functional forms could be discriminated by higher precision data.

TABLE II. Values of $|f(z_{TM})|$ and $BR(K_L \to (\mu^+\mu^-)\gamma)$ computed using the phenomenological form factors with parameters set by either radiative $K_L$ decay to $e$ or $\mu$.

| Model          | $|f(z_{TM})|$ | $BR_{TM} \times 10^{13}$ |
|----------------|--------------|---------------------------|
| BMS$_{ee}$     | 1.134(6)$^a$ | 6.60(10)                  |
| BMS$_{e\gamma}$ | 1.119(8)     | 6.42(11)                  |
| DIP$_{ee}$     | 1.139(6)$^a$ | 6.66(10)                  |
| DIP$_{e\gamma}$ | 1.124(9)     | 6.48(12)                  |

$^a$ The systematic and statistical errors have been summed

Due to the small value of $z_{PS} \approx 4M_K^2/M_K^2$, the branching ratio to positronium, $BR(K_L \to (e^+e^-)\gamma) = 9.31(5) \times 10^{-13}$, is independent of the form factor within the error of $BR(K_L \to \gamma\gamma)$ and slightly larger than $(\mu^+\mu^-)$. While this branching ratio also has not been measured, one can construct a ratio

$$R = \frac{BR(K_L \to (\mu^+\mu^-)\gamma)}{BR(K_L \to (e^+e^-)\gamma)} = \frac{(1 - z_{TM})^3 (1 - 0.439 \pm 0.12)}{(1 - z_{PS})^3 (1 - \frac{z_{PS}}{2})} |f(z_{TM})|^2 |f(z_{PS})|^2$$

$$= 0.55767(2) \frac{|f(z_{TM})|^2}{|f(z_{PS})|^2},$$

(5)

which is independent of the $BR(K_L \to \gamma\gamma)$ uncertainty and directly measures lepton universality without an uncertainty due to $Q^2$ binning. By taking the largest and smallest theoretical values of $b$ to give a gross range, we predict $R = 0.76(14)$. Applying the same procedure to the phenomenological form factors yields $R = 0.707(9)$.

We now focus upon the experimental situation. Throughout, we assume a 10% acceptance. The largest previous experimental data set that could be used to study $BR(K_L \to (\mu^+\mu^-)\gamma)$ is KTEV. We estimate from the number of events reported for $BR(K_L \to \ell^+\ell^-\gamma)$[16] that at least 1000 times the luminosity would be required for just one $(\mu^+\mu^-)$ event. From the existing data, one might expect to place a limit on the order of $BR(K_L \to (\mu^+\mu^-)\gamma) \lesssim 10^{-9}$.

The KOTO experiment at J-PARC has reported $3.560(0.013) \times 10^7 K_L$ per $2 \times 10^{14}$ protons on target (POT)[29]. Their 2013 physics run accumulated $1.6 \times 10^{18}$ POT[14] which would correspond to 0.015 $(\mu^+\mu^-)$ events. Through their 2015 physics run, 20 times the $K_L$ decays have been recorded[14], indicating 0.3 produced $(\mu^+\mu^-)$ events and a limit of $\lesssim 10^{-11}$. Unfortunately, the KOTO experiment is designed to detect only photons, and detecting purely photon decay products of $(\mu^+\mu^-)$ would be difficult. The J-PARC kaon beam hopes to run into the 2020s with an additional flux upgrade so a discovery is quite possible in an experiment with lepton identification. The NA62-KLEVER proposal[15] for a rare $K_L$ beam at CERN hopes to start by 2026 and accumulate $3 \times 10^{13} K_L$ over 5 years, which would also be nearly sufficient for single-event sensitivity.

A few channels are available to measure the branching ratio of true muonium: dissociated $(\mu^+\mu^-)$ with or without $\gamma$, decayed $e^+e^-$ with or without $\gamma$, or $\ell^+\ell^-$ similar to SUSY searches with invisible decays[30]. The decay to $\pi^0\gamma$ is suppressed by $10^{-5}$ but KOTO can search for it without modification[31].

For each channel, different backgrounds matter. The dominant backgrounds will arise from the free decays $K_L \to \ell^+\ell^-\gamma$. We compute the branching ratio for this by integrating the differential cross section in an invariant mass bin, $M_{bin}$, around the $(\mu^+\mu^-)$ peak to obtain a background estimate. In the case of electrons, the bin is centered around the $(\mu^+\mu^-)$ peak; for muon final states it is defined as $[2m_\mu, 2m_\mu + M_{bin}]$. This difference in binning reflects that the muons are
above threshold. For bin size similar to KTEV, the values are $\mathcal{BR}(K_L \to e^+e^-\gamma)_{\text{bin}} = 1.2 \times 10^{-8}M_{\text{bin}}$, and $\mathcal{BR}(K_L \to \mu^+\mu^-\gamma)_{\text{bin}} = 5.0 \times 10^{-9}M_{\text{bin}}$ where $M_{\text{bin}}$ is in MeV. This large raw background ($\sim 10^4 \times$ the signal) will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

The smoothness of the background differential cross section around the $(\mu^+\mu^-)$ peak should allow accurate modeling from the sidebands. Reconstruction of the $K_L$ allows the energy of the $K$ to be used to cut on the $\gamma$ and leptonic energies. The two two-body decay topology suggests cuts on momenta and angular distribution will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.

This suggests the higher energy of the proposed CERN facility would be desirable. Additionally, vertex cuts will have to be reduced, but it has distinct features compared to true muonium decays which can be leveraged.