

# A precision relation between $\Gamma(K \rightarrow \mu^+ \mu^-)(t)$ and $\mathcal{B}(K_L \rightarrow \mu^+ \mu^-)/\mathcal{B}(K_L \rightarrow \gamma\gamma)$

Avital Dery,<sup>a</sup> Mitrajyoti Ghosh,<sup>a</sup> Yuval Grossman,<sup>a</sup> Teppei Kitahara<sup>b,c,d,e</sup>  
and Stefan Schacht<sup>f</sup>

<sup>a</sup>*Department of Physics, LEPP, Cornell University,  
Ithaca, NY 14853, U.S.A.*

<sup>b</sup>*Institute for Advanced Research, Nagoya University,  
Nagoya 464-8601, Japan*

<sup>c</sup>*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University,  
Nagoya 464-8602, Japan*

<sup>d</sup>*KEK Theory Center, IPNS, KEK,  
Tsukuba 305-0801, Japan*

<sup>e</sup>*CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics,  
Chinese Academy of Sciences,  
Beijing 100190, China*

<sup>f</sup>*Department of Physics and Astronomy, University of Manchester,  
Manchester, M13 9PL, United Kingdom*

E-mail: [avital.dery@cornell.edu](mailto:avital.dery@cornell.edu), [mg2338@cornell.edu](mailto:mg2338@cornell.edu), [yg73@cornell.edu](mailto:yg73@cornell.edu),  
[teppeik@mki.nagoya-u.ac.jp](mailto:teppeik@mki.nagoya-u.ac.jp), [stefan.schacht@manchester.ac.uk](mailto:stefan.schacht@manchester.ac.uk)

**ABSTRACT:** We find that the phase appearing in the unitarity relation between  $\mathcal{B}(K_L \rightarrow \mu^+ \mu^-)$  and  $\mathcal{B}(K_L \rightarrow \gamma\gamma)$  is equal to the phase shift in the interference term of the time-dependent  $K \rightarrow \mu^+ \mu^-$  decay. A probe of this relation at future kaon facilities constitutes a Standard Model test with a theory precision of about 2%. The phase has further importance for sensitivity studies regarding the measurement of the time-dependent  $K \rightarrow \mu^+ \mu^-$  decay rate to extract the CKM matrix element combination  $|V_{ts}V_{td}\sin(\beta+\beta_s)| \approx A^2\lambda^5\bar{\eta}$ . We find a model-independent theoretically clean prediction,  $\cos^2\varphi_0 = 0.96 \pm 0.03$ . The quoted error is a combination of the theoretical and experimental errors, and both of them are expected to shrink in the future. Using input from the large- $N_C$  limit within chiral perturbation theory, we find a theory preference towards solutions with negative  $\cos\varphi_0$ , reducing a four-fold ambiguity in the angle  $\varphi_0$  to a two-fold one.

**KEYWORDS:** Kaons, Rare Decays, CP Violation

**ARXIV EPRINT:** [2211.03804](https://arxiv.org/abs/2211.03804)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Setup and notation</b>	<b>3</b>
<b>3</b>	<b>Determination of <math>\cos^2 \varphi_0</math>: model-independent</b>	<b>5</b>
<b>4</b>	<b>Determination of <math>\cos \varphi_0</math>: model dependent</b>	<b>7</b>
4.1	Chiral perturbation theory	8
4.2	Detailed assumptions within ChPT	10
4.3	Lattice QCD	11
4.4	SM prediction for the short-distance physics	11
<b>5</b>	<b>Going beyond the two-fold ambiguity</b>	<b>12</b>
<b>6</b>	<b>Discussion and conclusions</b>	<b>13</b>
<b>A</b>	<b>Update of the theory estimate for <math>\chi_{\gamma\gamma}(\mu)</math></b>	<b>15</b>

---

## 1 Introduction

A recent proposal has shown that short-distance parameters of the decay  $K \rightarrow \mu^+ \mu^-$  can be cleanly extracted from a measurement of the  $K_L$ – $K_S$  time-dependent rate [1–3]. The time-dependent rate for a beam of initial  $K^0$  particles can be written as

$$\frac{1}{\mathcal{N}} \frac{d\Gamma(K^0 \rightarrow \mu^+ \mu^-)}{dt} = f(t) \equiv C_L e^{-\Gamma_L t} + C_S e^{-\Gamma_S t} + 2 C_{\text{Int.}} \cos(\Delta M_K t - \varphi_0) e^{-\frac{\Gamma_L + \Gamma_S}{2} t}, \quad (1.1)$$

where  $\mathcal{N}$  is a normalization factor,  $\Gamma_L$  ( $\Gamma_S$ ) is the  $K_L$  ( $K_S$ ) decay width, and  $\Delta M_K$  is the  $K_L$ – $K_S$  mass difference. Then, the four experimental parameters characterizing the time dependence,

$$\{C_L, C_S, C_{\text{Int.}}, \varphi_0\}, \quad (1.2)$$

are directly related to the four theory parameters describing the system [2],

$$\{|A(K_S)_{\ell=0}|, |A(K_L)_{\ell=0}|, |A(K_S)_{\ell=1}|, \arg[A(K_S)_{\ell=0}^* A(K_L)_{\ell=0}]\}, \quad (1.3)$$

where the subscripts  $\ell = 0$  ( $s$ -wave symmetric wave function) and  $\ell = 1$  ( $p$ -wave anti-symmetric wave function) correspond to the CP-odd and -even ( $\mu^+ \mu^-$ ) final states, respec-

tively. The relations between the experimental and theory parameters are given by

$$\begin{aligned}
C_L &= |A(K_L)_{\ell=0}|^2, \\
C_S &= |A(K_S)_{\ell=0}|^2 + \beta_\mu^2 |A(K_S)_{\ell=1}|^2, \\
C_{\text{Int.}} &= |A(K_S)_{\ell=0}^* A(K_L)_{\ell=0}| = |A(K_S)_{\ell=0}| |A(K_L)_{\ell=0}|, \\
\varphi_0 &= \arg [A(K_S)_{\ell=0}^* A(K_L)_{\ell=0}],
\end{aligned}
\tag{1.4}$$

with

$$\beta_\mu = \sqrt{1 - \frac{4m_\mu^2}{m_{K^0}^2}}.
\tag{1.5}$$

The experimental parameter  $\varphi_0$ , which is the phase shift of the oscillating rate in eq. (1.1), is a combination of the relative weak and strong phases between the  $K_S$  and  $K_L$  amplitudes to the CP-odd final state. In this paper, we demonstrate that this phase shift is closely related to the proportionality coefficient in the ratio between the rates of  $K_L \rightarrow \mu^+ \mu^-$  and  $K_L \rightarrow \gamma \gamma$ .

The ratio between the rates of  $K_L \rightarrow \mu^+ \mu^-$  and  $K_L \rightarrow \gamma \gamma$  is of historical significance. Using CPT invariance, unitarity, and the well-motivated assumption that the absorptive part of the  $K_L \rightarrow \mu^+ \mu^-$  amplitude is dominated by the two-photon intermediate state, the ratio between the rates of  $K_L \rightarrow \mu^+ \mu^-$  and  $K_L \rightarrow \gamma \gamma$  is bounded by the lower limit [4–7]

$$R_{K_L} \equiv \frac{\Gamma(K_L \rightarrow \mu^+ \mu^-)}{\Gamma(K_L \rightarrow \gamma \gamma)} \geq 1.195 \times 10^{-5}.
\tag{1.6}$$

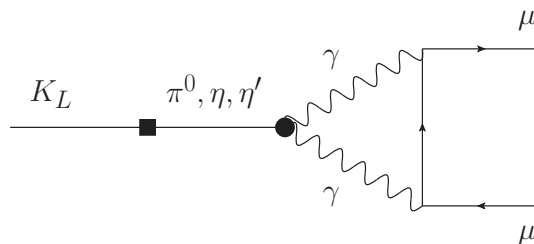
However, back in the 1970’s, this conflicted with the contemporary experimental upper bound of  $0.4 \times 10^{-5}$  [8, 9], leading to the so-called “ $K_L \rightarrow \mu^+ \mu^-$  puzzle” [5, 10] that gained much attention. But today, this ratio is measured in the experiment as [11]

$$R_{K_L}^{\text{exp}} = (1.250 \pm 0.022) \times 10^{-5}.
\tag{1.7}$$

This means that the observed branching ratio  $\mathcal{B}(K_L \rightarrow \mu^+ \mu^-) = (6.84 \pm 0.11) \times 10^{-9}$  [11], along with  $\mathcal{B}(K_L \rightarrow \gamma \gamma) = (5.47 \pm 0.04) \times 10^{-4}$  [11] known today obey the lower limit prescribed by CPT invariance and unitarity.

In this paper, we show that the phase  $\varphi_0$  is cleanly predicted in the Standard Model (SM), up to a four-fold discrete ambiguity, making its measurement a potent test of the SM. The discrete ambiguity can be partially resolved by using further theory input from the literature in the large- $N_C$  limit of chiral perturbation theory (ChPT). This result is additionally significant for sensitivity estimations of a future measurement of the short-distance parameters.

Leptonic kaon decays have been a field that received a lot of attention in the literature recently. Effects from CPV in kaon mixing on  $K \rightarrow \mu^+ \mu^-$  have been taken into account in refs. [1, 3], and implications for physics beyond the SM have been studied in refs. [12–14]. Another future high precision test of the SM employing the ratio  $\mathcal{B}(K_S \rightarrow \mu^+ \mu^-)_{\ell=0} / \mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu})$  has been identified in ref. [15]. Advances in calculating



**Figure 1.** Leading order Feynman diagram of the long-distance two-photon contribution to  $K_L \rightarrow \mu^+ \mu^-$ .

$K_L \rightarrow \mu^+ \mu^-$  and  $K_L \rightarrow \gamma \gamma$  on the lattice can be found in refs. [16–18]. On the experimental side, the LHCb collaboration recently found an improved bound on  $K_S \rightarrow \mu^+ \mu^-$  [19] and  $K_{S,L} \rightarrow 2(\mu^+ \mu^-)$  [20].

In section 2 we introduce our notation and summarize key results from the literature. In section 3 we determine  $\cos^2 \varphi_0$ , which predicts  $\varphi_0$  up to a four-fold ambiguity, in a model-independent way, only assuming that the long-distance contributions are SM-like. In section 4 we reduce this ambiguity to a two-fold one by using the large- $N_C$  limit and assuming that the short-distance physics is known. We demonstrate that the remaining ambiguity cannot be resolved using current knowledge in section 5. We conclude in section 6.

## 2 Setup and notation

We work within a framework defined by the following approximations, as detailed in ref. [2]:

- (i) We neglect CPV in mixing, which is a sub-dominant effect for our purposes. Note that when considering higher order corrections, this effect can be taken into account consistently [3].
- (ii) We neglect CPV in the long-distance contribution.
- (iii) We assume that the leptonic current is (axial-)vectorial, i.e., given by  $\bar{\mu} \gamma^\mu (a + b \gamma^5) \mu$ .

In the SM, all three approximations are fulfilled within the precision relevant for our findings. In addition, in the following we assume that

- (iv) The long-distance contribution to the  $K_L \rightarrow \mu^+ \mu^-$  amplitude is SM-like. That is, the only non-negligible intermediate state is the di-photon state.

Regarding the short-distance contribution, we explicitly state whenever our results are relevant regardless of any assumption on the nature of the short-distance physics, and when SM input is used.

Adopting this setup, it has been shown that the decay of  $K_S$  to the CP-odd final state,  $(\mu^+ \mu^-)_{\ell=0}$  involves only short-distance physics. Additionally, the  $K_L \rightarrow \mu^+ \mu^-$  decay proceeds only to the CP-odd final state,  $(\mu^+ \mu^-)_{\ell=0}$ . However, the latter involves two contributions of different underlying physics:

1. A short-distance (SD) contribution, arising to leading order from box and electroweak penguin diagrams, for which the ingredients for a precise SM prediction are straightforward to derive.
2. A long-distance (LD) contribution, strongly dominated by the on-shell two photon intermediate state, see figure 1.

The on-shell two photon contribution is known to be significantly larger in magnitude than both the SD contribution and the off-shell part of the LD contribution.

The  $K^0 \rightarrow (\mu^+ \mu^-)_{\ell=0}$  and  $\bar{K}^0 \rightarrow (\mu^+ \mu^-)_{\ell=0}$  amplitudes can be written as a sum of two general contributions, with corresponding weak and strong phases,

$$\begin{aligned} A_{\ell=0} &= |A_{SD}| e^{i\theta_{SD}} e^{i\delta_{SD}} + |A_{LD}| e^{i\theta_{LD}} e^{i\delta_{LD}}, \\ \bar{A}_{\ell=0} &= - \left( |A_{SD}| e^{-i\theta_{SD}} e^{i\delta_{SD}} + |A_{LD}| e^{-i\theta_{LD}} e^{i\delta_{LD}} \right), \end{aligned} \quad (2.1)$$

where the overall minus sign for  $\bar{A}_{\ell=0}$  is due to the CP nature of the final state. Using the convention

$$|K_S\rangle = p|K^0\rangle + q|\bar{K}^0\rangle, \quad |K_L\rangle = p|K^0\rangle - q|\bar{K}^0\rangle, \quad (2.2)$$

the mass eigenstate amplitudes are related to  $A_{\ell=0}$  and  $\bar{A}_{\ell=0}$  by [2]

$$\begin{aligned} A(K_S)_{\ell=0} &= \frac{1}{\sqrt{2}} [A_{SD}(1 + \lambda_{SD}) + A_{LD}(1 + \lambda_{LD})], \\ A(K_L)_{\ell=0} &= \frac{1}{\sqrt{2}} [A_{SD}(1 - \lambda_{SD}) + A_{LD}(1 - \lambda_{LD})], \end{aligned} \quad (2.3)$$

where

$$A_{SD} \equiv |A_{SD}| e^{i\theta_{SD}} e^{i\delta_{SD}}, \quad A_{LD} \equiv |A_{LD}| e^{i\theta_{LD}} e^{i\delta_{LD}}, \quad (2.4)$$

$$\bar{A}_{SD} \equiv -|A_{SD}| e^{-i\theta_{SD}} e^{i\delta_{SD}}, \quad \bar{A}_{LD} \equiv -|A_{LD}| e^{-i\theta_{LD}} e^{i\delta_{LD}}, \quad (2.5)$$

and

$$\lambda_{SD} \equiv \frac{q}{p} \frac{\bar{A}_{SD}}{A_{SD}}, \quad \lambda_{LD} \equiv \frac{q}{p} \frac{\bar{A}_{LD}}{A_{LD}}. \quad (2.6)$$

As we focus here mainly on amplitudes with  $\ell = 0$  final states, compared to the notation of ref. [2] we use for brevity the notation

$$\lambda_{SD} \equiv \lambda_0^{SD}, \quad \lambda_{LD} \equiv \lambda_0^{LD}, \quad (2.7)$$

$$A(K_S) \equiv A(K_S)_{\ell=0}, \quad A(K_L) \equiv A(K_L)_{\ell=0}. \quad (2.8)$$

The amplitudes are normalized such that

$$\mathcal{B}(K_{S,L} \rightarrow \mu^+ \mu^-)_{\ell=0} = \frac{\tau_K \beta_\mu}{16\pi m_K} |A(K_{S,L})|^2. \quad (2.9)$$

We also use

$$\mathcal{B}(K_L \rightarrow \gamma\gamma) = \frac{\tau_K}{32\pi m_K} |A(K_L \rightarrow \gamma\gamma)|^2. \quad (2.10)$$

Up to this point, the labels LD and SD are just naming. However, in the following we will treat them as corresponding to what we think of as long-distance and short-distance amplitudes. It is important to note that the separation into LD and SD contributions is not well-defined. We think of short-distance physics as having no sources for a strong phase, while long-distance physics can go on-shell. However, any on-shell intermediate state can also be considered to contribute off-shell. Therefore there is no way to unambiguously define the separation. In the following we keep the strong phases general, while we insert knowledge of the SM weak phase of the long-distance contribution.

We then have [2]

$$\lambda_{LD} = - \left( \frac{V_{cd}V_{cs}^*}{V_{cd}^*V_{cs}} \right) \left( \frac{V_{ud}^*V_{us}}{V_{ud}V_{us}^*} \right) = -e^{-2i\theta_{uc}}. \quad (2.11)$$

Equation (2.11) corrects a typo in eq. (42) of ref. [2]. Here,

$$\theta_{uc} \equiv \arg \left( -\frac{V_{cd}V_{cs}^*}{V_{ud}V_{us}^*} \right) = \mathcal{O}(\lambda^4), \quad (2.12)$$

such that, to  $\mathcal{O}(\lambda^4)$ ,

$$\arg \lambda_{LD} = -2\theta_{LD} + \pi = \pi, \quad (2.13)$$

that is

$$\theta_{LD} = 0, \quad \lambda_{LD} = -1. \quad (2.14)$$

No assumptions are made for the short-distance phase,

$$\arg \lambda_{SD} = -2\theta_{SD} + \pi. \quad (2.15)$$

We have therefore

$$A(K_S) = \frac{1}{\sqrt{2}} A_{SD} (1 + \lambda_{SD}) = i\sqrt{2} |A_{SD}| \sin \theta_{SD} e^{i\delta_{SD}}, \quad (2.16)$$

$$A(K_L) = \frac{1}{\sqrt{2}} [A_{SD} (1 - \lambda_{SD}) + 2A_{LD}] = \sqrt{2} (|A_{LD}| e^{i\delta_{LD}} + |A_{SD}| \cos \theta_{SD} e^{i\delta_{SD}}). \quad (2.17)$$

Note that  $A(K_S)$  is pure short-distance and is manifestly CP-odd. The oscillation term in the rate is then controlled by the interference term:

$$A(K_S)^* A(K_L) = -2i |A_{SD}| \sin \theta_{SD} (|A_{LD}| e^{i\Delta\delta} + |A_{SD}| \cos \theta_{SD}), \quad (2.18)$$

where  $\Delta\delta \equiv \delta_{LD} - \delta_{SD}$ .

### 3 Determination of $\cos^2 \varphi_0$ : model-independent

From eqs. (1.4) and (2.18), we have

$$\cos^2 \varphi_0 = \frac{\text{Re}[A(K_S)^* A(K_L)]^2}{|A(K_S)A(K_L)|^2} = \frac{(\sqrt{2}|A_{LD}| \sin \Delta\delta)^2}{|A(K_L)|^2} = \frac{[A(K_L)_{\text{absorptive}}]^2}{|A(K_L)|^2}, \quad (3.1)$$

where

$$A(K_L)_{\text{absorptive}} \equiv \text{Im}[A(K_L)], \quad (3.2)$$

and we define the imaginary part relative to the strong phase of the SD amplitude (or in the basis in which the SD amplitude carries no strong phase, hence  $\delta_{LD} = \Delta\delta$ ). This convention corresponds to choosing  $C_{\text{had.}}$ , introduced in section 4 below, to be real without loss of generality.

The numerator is simply the on-shell long-distance contribution to the  $K_L$  amplitude. Under the assumption that the only non-negligible intermediate state is the di-photon state [21], this absorptive part is equal to the discontinuity of the three-point diagram (see figure 1),

$$\left| \sqrt{2} |A_{LD}| \sin \Delta\delta \right| = |A(K_L)_{\text{absorptive}}| = \left| \frac{1}{2i} \text{Disc}(K_L \rightarrow \gamma\gamma \rightarrow \mu^+\mu^-) \right|. \quad (3.3)$$

This discontinuity can be computed in a straightforward way using Cutkosky rules, which entail considering the cut intersecting both photon propagators. In this way the left part of the cut diagram is directly related to the measured rate of  $K_L \rightarrow \gamma\gamma$ , and the right  $\gamma\gamma \rightarrow \mu^+\mu^-$  part of the diagram requires QED only, allowing to extract the magnitude of the discontinuity completely model independently. This leads to

$$\cos^2 \varphi_0 = \frac{[A(K_L)_{\text{absorptive}}]^2}{|A(K_L)|^2} = C_{\text{QED}}^2 \frac{\Gamma(K_L \rightarrow \gamma\gamma)}{\Gamma(K_L \rightarrow \mu^+\mu^-)}, \quad (3.4)$$

i.e.,

$$\cos^2 \varphi_0 \mathcal{B}(K_L \rightarrow \mu^+\mu^-) = C_{\text{QED}}^2 \mathcal{B}(K_L \rightarrow \gamma\gamma), \quad (3.5)$$

where  $C_{\text{QED}}$  describes the  $\gamma\gamma \rightarrow \mu^+\mu^-$  transition and is given as [21]

$$C_{\text{QED}} = \frac{\alpha_{em} m_\mu}{\sqrt{2} \beta_\mu m_K} \log \left( \frac{1 - \beta_\mu}{1 + \beta_\mu} \right) + \mathcal{O}(\alpha_{em}^2). \quad (3.6)$$

Equation (3.5) demonstrates that  $\cos^2 \varphi_0$  is simply the proportionality factor parameterizing to what extent is the  $K_L \rightarrow \mu^+\mu^-$  rate saturated by the absorptive contribution from the intermediate  $\gamma\gamma$  state. Using the measured ratio of rates  $R_{K_L}$ , see eq. (1.6), we have a clean SM prediction for the phase  $\varphi_0$ , up to a four-fold discrete ambiguity,

$$\cos^2 \varphi_0 = \frac{C_{\text{QED}}^2}{R_{K_L}^{\text{exp}}}. \quad (3.7)$$

This SM prediction is dependent only on

1. The measurement of  $R_{K_L}$ , currently with an uncertainty of  $\mathcal{O}(2\%)$ ;
2. A QED calculation, here taken up to relative corrections of  $\mathcal{O}(\alpha_{em})$ ;
3. The assumption that other intermediate on-shell contributions ( $3\pi$ ,  $\pi\pi\gamma$ ) are negligible [21].

Inserting

$$C_{\text{QED}}^2 = 1.195 \times 10^{-5} [1 + \mathcal{O}(\alpha_{em})], \quad (3.8)$$

and

$$R_{K_L}^{\text{exp}} = (1.250 \pm 0.022) \times 10^{-5}, \quad (3.9)$$

which we obtain, not taking into account any correlations, from [11]

$$\mathcal{B}(K_L \rightarrow \mu^+ \mu^-) = (6.84 \pm 0.11) \times 10^{-9}, \quad (3.10)$$

$$\mathcal{B}(K_L \rightarrow \gamma\gamma) = (5.47 \pm 0.04) \times 10^{-4}, \quad (3.11)$$

using Gaussian error propagation, we arrive at  $\cos^2 \varphi_0 = 0.96 \pm 0.02$ , where the quoted error reflects only the experimental error on  $R_{K_L}^{\text{exp}}$ .

There are two sources of theoretical errors. One of them is the higher order QED calculation, resulting in an error of order  $\alpha_{em} \sim 1\%$ . This error is reducible, that is, if needed the calculation of the higher-order corrections can be done. The other source of uncertainty are the intermediate states that we neglected, such as  $3\pi$  and  $\pi\pi\gamma$  [21]. These contributions are estimated to be at most 1% of the two-photon state that we considered. Each of the two effects results in about 1% error, and thus we conservatively add them linearly resulting in a total theory error of 0.02 to arrive at our final estimate,

$$\cos^2 \varphi_0 = 0.96 \pm 0.02_{\text{exp}} \pm 0.02_{\text{th}}. \quad (3.12)$$

Combining the experimental and theoretical errors in quadrature we then have

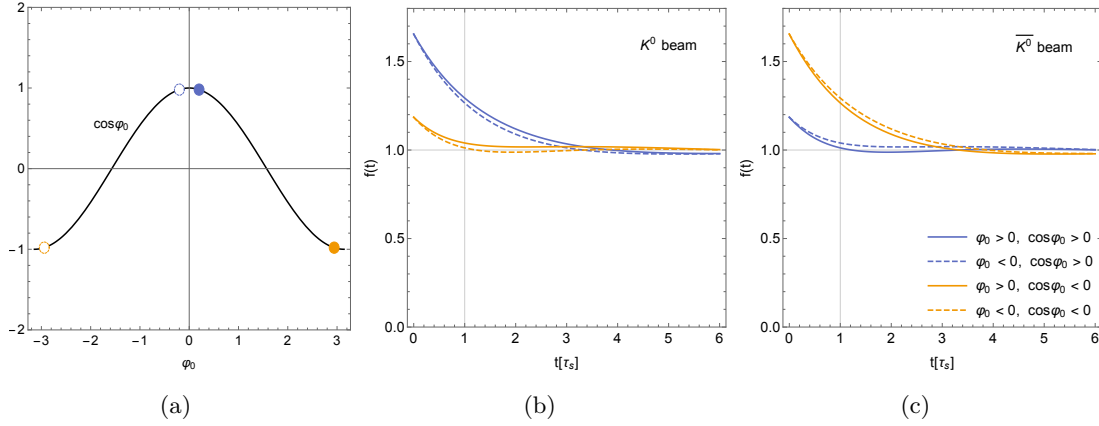
$$\cos^2 \varphi_0 = 0.96 \pm 0.03. \quad (3.13)$$

Note that the error quoted in eq. (3.13) is therefore to be interpreted as an estimate of the total uncertainty, rather than as a statistical error only.

Given the value of  $\cos^2 \varphi_0$  there are four possible values for the phase shift,  $\varphi_0$ . Two of them correspond to overall constructive interference in the time-dependent rate and two correspond to destructive interference, depending on the sign of  $\cos \varphi_0$ . We plot the time-dependent rate for the four possibilities in figure 2 for both a  $K^0$  or a  $\bar{K}^0$  beam. Note that, for a  $K^0$  beam, a positive (negative)  $\cos \varphi_0$  results in constructive (destructive) interference. As pointed out in ref. [1], the situation is reversed for a  $\bar{K}^0$  beam where a negative (positive)  $\cos \varphi_0$  results in constructive (destructive) interference.

#### 4 Determination of $\cos \varphi_0$ : model dependent

The four-fold ambiguity cannot be resolved in a model-independent way. Thus, in this section we use input from theory in order to try to reduce it. We first consider the sign of  $\cos \varphi_0$ . We discuss below how chiral perturbation theory (ChPT) and lattice QCD can help in this regard.



**Figure 2.** (a) Illustration of the four-fold ambiguity arising from  $\cos^2 \varphi_0 = 0.96$ . (b) The time dependence of the  $K^0(t) \rightarrow \mu^+\mu^-$  rate,  $f(t)$ , as defined in eq. (1.1), for the four values of  $\varphi_0$ . We have used a normalization in which  $C_L$  of eq. (1.1) is set to unity. (c) The same for an initial pure  $\bar{K}^0$  beam.

#### 4.1 Chiral perturbation theory

Within chiral perturbation theory and using the large- $N_C$  limit, it has been shown that the sign of the absorptive amplitude relative to the short-distance contribution, which determines the sign of  $\cos \varphi_0$ , can be determined [22, 23].

In the following, we relate our notation for the long-distance contribution to that appearing in the literature in order to apply existing results, primarily of refs. [22–25]. We rewrite the long-distance amplitude for  $K_L \rightarrow \mu^+\mu^-$  as

$$\begin{aligned} \sqrt{2}|A_{LD}|e^{i\Delta\delta} &= [A_{LD}]_{\text{dispersive}} + i[A_{LD}]_{\text{absorptive}} \\ &\equiv C_{\text{had.}} \left[ (A_{LD}^{\text{local}} + \text{Re}A_{LD}^{\gamma\gamma}) + i\text{Im}A_{LD}^{\gamma\gamma} \right], \end{aligned} \quad (4.1)$$

where  $C_{\text{had.}}$  encodes the hadronic behavior of the effective  $K_L\gamma\gamma$  vertex. Without loss of generality, we take  $C_{\text{had.}}$  to be real, consistent with eq. (3.2). The dispersive part is split in two parts to be consistent with the literature,

$$[A_{LD}]_{\text{dispersive}} = C_{\text{had.}} \left( A_{LD}^{\text{local}} + \text{Re}A_{LD}^{\gamma\gamma} \right), \quad (4.2)$$

$$[A_{LD}]_{\text{absorptive}} = C_{\text{had.}} \text{Im}A_{LD}^{\gamma\gamma}, \quad (4.3)$$

where  $A_{LD}^{\text{local}}$  denotes the local counterterm, which is real [23]. Using the fact that the decay  $K_L \rightarrow \gamma\gamma$  has the same hadronic behavior, we define

$$A(K_L \rightarrow \gamma\gamma) \equiv C_{\text{had.}} A_{\gamma\gamma}. \quad (4.4)$$

Note that  $A_{\gamma\gamma} \neq A_{LD}^{\gamma\gamma}$ . The former is part of the  $K_L \rightarrow \gamma\gamma$  amplitude, while the latter is

the two-photon contribution to  $K_L \rightarrow \mu^+ \mu^-$ . Then we have

$$\begin{aligned} \frac{\Gamma(K_L \rightarrow \mu^+ \mu^-)}{\Gamma(K_L \rightarrow \gamma\gamma)} &= 2\beta_\mu \left| \frac{\sqrt{2}|A_{LD}|e^{i\Delta\delta} + \sqrt{2}|A_{SD}|\cos\theta_{SD}}{A(K_L \rightarrow \gamma\gamma)} \right|^2 \\ &= 2\beta_\mu \frac{\left( A_{LD}^{\text{local}} + \text{Re}A_{LD}^{\gamma\gamma} + \sqrt{2}|A_{SD}|\cos\theta_{SD}/C_{\text{had.}} \right)^2 + (\text{Im}A_{LD}^{\gamma\gamma})^2}{|A_{\gamma\gamma}|^2}. \end{aligned} \quad (4.5)$$

Note that  $C_{\text{had.}}$  now appears as a factor accompanying the short-distance contribution. We can now easily relate to the notations of ref. [23], with (the superscript “IU” denotes the initials of the authors of ref. [23])

$$\frac{A_{LD}^{\text{local}}}{|A_{\gamma\gamma}|} = \frac{\alpha_{em}m_\mu}{\pi m_K} [\chi_{\gamma\gamma}(\mu)]^{\text{IU}}, \quad (4.6)$$

$$\frac{\text{Re}A_{LD}^{\gamma\gamma}}{|A_{\gamma\gamma}|} = \frac{\alpha_{em}m_\mu}{\pi m_K} \left[ \text{Re}C_{\gamma\gamma} - \frac{5}{2} + \frac{3}{2} \log \left( \frac{m_\mu^2}{\mu^2} \right) \right]^{\text{IU}}, \quad (4.7)$$

$$\frac{\sqrt{2}|A_{SD}|\cos\theta_{SD}}{C_{\text{had.}}|A_{\gamma\gamma}|} = \frac{\alpha_{em}m_\mu}{\pi m_K} [\chi_{\text{short}}]^{\text{IU}}, \quad (4.8)$$

$$\frac{\text{Im}A_{LD}^{\gamma\gamma}}{|A_{\gamma\gamma}|} = \frac{\alpha_{em}m_\mu}{\pi m_K} [\text{Im}C_{\gamma\gamma}]^{\text{IU}} = \frac{1}{\sqrt{2}\beta_\mu} C_{\text{QED}}. \quad (4.9)$$

The assumptions and findings of the literature, as conveyed in refs. [22, 23], can now be summarized as the following:

1. Using results in the large- $N_C$  limit, refs. [22, 23] find destructive interference between the short-distance and the local long-distance contributions,

$$\frac{\cos\theta_{SD}}{C_{\text{had.}}A_{LD}^{\text{local}}} < 0. \quad (4.10)$$

2. ref. [23] uses phenomenological analyses of the form factor in  $K_L \rightarrow \gamma e^+ e^-$ ,  $K_L \rightarrow \gamma \mu^+ \mu^-$  and  $K_L \rightarrow e^+ e^- \mu^+ \mu^-$  from data, together with theory considerations, to estimate the local counter term. Using up-to-date inputs, we update their estimation (see appendix A) and find

$$[\chi_{\gamma\gamma}(m_\rho)]^{\text{IU}} = \frac{\pi m_K}{\alpha_{em}m_\mu} \frac{A_{LD}^{\text{local}}}{|A_{\gamma\gamma}|} = (6.10 \pm 1.01) > 0, \quad (4.11)$$

where we have set  $\mu^2 = m_\rho^2$  here and in the following whenever we make use of specific numerical estimates. Hence, using eq. (4.10),

$$\text{sgn}[C_{\text{had.}}] = -\text{sgn}[\cos\theta_{SD}]. \quad (4.12)$$

3. This, in turn, determines the sign of the absorptive long-distance amplitude relative to the short-distance contribution,

$$\text{sgn}([A_{LD}]_{\text{absorptive}}) = \text{sgn}[\cos\theta_{SD}], \quad (4.13)$$

where we used the fact that  $\text{sgn}[\text{Im}A_{LD}^{\gamma\gamma}] = \text{sgn}[C_{\text{QED}}] = -1$ .

We conclude that within a model for the short-distance contribution, and adopting the assumptions in the literature regarding the long-distance physics, i.e., the large- $N_C$  limit, the sign of  $\cos \varphi_0$  is determined,

$$\begin{aligned} \text{sgn}[\cos \varphi_0] &= \text{sgn} \left[ \frac{\text{Re}[A(K_S)^* A(K_L)]}{|A(K_S)A(K_L)|} \right] \\ &= \text{sgn} \left[ \frac{[A(K_L)]_{\text{absorptive}} \text{sgn}[\sin \theta_{SD}]}{|A(K_L)|} \right] \\ &= \text{sgn}[\tan \theta_{SD}]. \end{aligned} \quad (4.14)$$

## 4.2 Detailed assumptions within ChPT

The considerations leading to the assumption of destructive interference between the short-distance and the local long-distance contributions, as conveyed in eq. (4.10), involve some details of the structure of the  $K_L \rightarrow \gamma\gamma$  amplitude within the ChPT. According to ref. [26], the on-shell tensor amplitude for  $K_L \rightarrow \gamma\gamma$  starts from at  $\mathcal{O}(p^6)$  in the ChPT as

$$-iA(K_L \rightarrow \gamma\gamma) = \varepsilon^{\mu\nu\rho\sigma} \epsilon_{1\mu}(q_1) \epsilon_{2\nu}(q_2) q_{1\rho} q_{2\sigma} c^{(6)}(0,0), \quad (4.15)$$

where  $c^{(6)}(q_1^2, q_2^2)$  is an amplitude for  $K_L \rightarrow \gamma^*(q_1) \gamma^*(q_2)$ . We factorize the  $\pi^0$  exchange contribution in  $c^{(6)}(0,0)$  and define a dimensionless reduced amplitude as  $c_{\text{red}}^{(6)}$ ,

$$c^{(6)}(0,0) = -\frac{2}{\pi} \alpha_{em} F_0 (G_8 - G_{27}) \frac{1}{1 - r_\pi^2} c_{\text{red}}^{(6)}, \quad (4.16)$$

and  $c_{\text{red}}^{(6)}$  can be expanded in the ChPT as

$$\begin{aligned} c_{\text{red}}^{(6)} &= 1 + \frac{1 - r_\pi^2}{3(1 - r_\eta^2)} \left[ (1 + \xi) c_\theta + 2\sqrt{2} \hat{\rho} s_\theta \right] \left( \frac{F_\pi}{F_{\eta_8}} c_\theta - 2\sqrt{2} \frac{F_\pi}{F_{\eta_1}} s_\theta \right) \\ &\quad - \frac{1 - r_\pi^2}{3(1 - r_{\eta'}^2)} \left[ 2\sqrt{2} \hat{\rho} c_\theta - (1 + \xi) s_\theta \right] \left( \frac{F_\pi}{F_{\eta_8}} s_\theta + 2\sqrt{2} \frac{F_\pi}{F_{\eta_1}} c_\theta \right), \end{aligned} \quad (4.17)$$

with  $\sqrt{2}F_0 = \sqrt{2}F_\pi = f_\pi = (130.2 \pm 0.8) \text{ MeV}$  [27],  $F_{\eta_8} = (1.27 \pm 0.02) F_\pi$ ,  $F_{\eta_1} = (1.14 \pm 0.05) F_\pi$  [28, 29], and  $r_P \equiv m_P/m_K$ . Combining the above formulae gives

$$-iA(K_L \rightarrow \gamma\gamma) = -\frac{\sqrt{2}\alpha_{em}f_\pi (G_8 - G_{27})}{\pi(1 - r_\pi^2)} c_{\text{red}}^{(6)} \varepsilon^{\mu\nu\rho\sigma} \epsilon_{1\mu}(q_1) \epsilon_{2\nu}(q_2) q_{1\rho} q_{2\sigma}. \quad (4.18)$$

Here, the point is that  $\mathcal{O}(p^4)$  contributions vanish within the ChPT, which are proportional to  $c_{\text{red}}^{(4)}$  and

$$c_{\text{red}}^{(4)} = 1 + \frac{1 - r_\pi^2}{3(1 - r_{\eta_8}^2)} = \frac{4 - 3r_{\eta_8}^2 - r_\pi^2}{3(1 - r_{\eta_8}^2)} = 0, \quad (4.19)$$

where the Gell-Mann-Okubo mass formula,  $4m_K^2 = 3m_{\eta_8}^2 + m_\pi^2$ , is used. Therefore,  $c_{\text{red}}^{(6)}$  amplitudes correspond to the violation of the Gell-Mann-Okubo formula implying that the

sign of  $c_{\text{red}}^{(6)}$  is sensitive to the  $\eta$ - $\eta'$  mixing angle  $\theta$  in the octet-singlet basis, the  $SU(3)_F$  breaking  $\xi$  [30, 31], the nonet symmetry breaking  $\hat{\rho}$  [32–34], and their higher-order corrections.  $(1 + \xi)$  is proportional to the  $K_L \rightarrow \eta_8$  form factor, while  $\hat{\rho}$  is proportional to the  $K_L \rightarrow \eta_1$  one. However, by considering the typical parameter regions;  $\theta \approx -20^\circ$ ,  $\xi \sim 0.0$ – $0.2$ , and  $\hat{\rho} \approx 0.8$  [26], one can predict  $\text{sgn}[c_{\text{red}}^{(6)}] > 0$ , which leads to

$$\text{sgn}[A(K_L \rightarrow \gamma\gamma)] = \text{sgn}[A(K_L \rightarrow \pi^0 \rightarrow \gamma\gamma)]. \quad (4.20)$$

Combining this relation with  $\text{sgn}[G_8 - G_{27}]$  which can be extracted from the  $\Delta S = 1$  effective Lagrangian in the large- $N_C$  limit [23, 35, 36],  $\text{sgn}[C_{\text{had.}}] > 0$  can be predicted, see eq. (4.12), where, in the SM we have  $\cos \theta_{SD}^{\text{SM}} < 0$ .

### 4.3 Lattice QCD

In the last decade, lattice QCD made paramount progress in the treatment of  $K \rightarrow \pi\pi$  [37–40]. Moreover, recently, lattice QCD made advances in the calculation of  $K_L \rightarrow \mu^+\mu^-$  and  $K_L \rightarrow \gamma\gamma$  [16–18]. ChPT parameters like  $G_8$  can now also be extracted from fits to lattice results, as shown in ref. [41].

While it seems to us that the data is available to extract the sign, we were unable to find it from the available publications. It would be interesting to use the available lattice data to obtain it. Such an extraction would be interesting to confront the ChPT results.

### 4.4 SM prediction for the short-distance physics

Within the SM, the short distance contribution arises from the following effective Hamiltonian [42]

$$\mathcal{H}_{\text{eff}} = -\frac{G_F}{\sqrt{2}} \frac{\alpha_{em}}{2\pi \sin^2 \theta_W} [V_{ts}^* V_{td} Y(x_t) + V_{cs}^* V_{cd} Y_{NL}] [(\bar{s}d)_{V-A} (\bar{\mu}\mu)_{V-A}] + \text{h.c.} \quad (4.21)$$

We can then write (in the basis where  $\delta_{SD} = 0$ )

$$\begin{aligned} A(K_L)_{SD}^{\text{SM}} &= (\sqrt{2}|A_{SD}| \cos \theta_{SD})^{\text{SM}} \\ &= \frac{\sqrt{2}G_F \alpha_{em}}{\pi \sin^2 \theta_W} |V_{ts}^* V_{td} Y(x_t) + V_{cs}^* V_{cd} Y_{NL}| f_K m_\mu m_K \cos \theta_{SD}^{\text{SM}}, \end{aligned} \quad (4.22)$$

where we identify

$$\theta_{SD}^{\text{SM}} = \arg \left( -\frac{V_{ts}^* V_{td} + V_{cs}^* V_{cd} Y_{NL}/Y(x_t)}{V_{cs}^* V_{cd}} \right). \quad (4.23)$$

We therefore find,

$$\tan \theta_{SD}^{\text{SM}} = -\frac{\eta}{(1 - \rho) + \frac{1}{A^2 \lambda^4} \frac{Y_{NL}}{Y(x_t)}} + \mathcal{O}(\lambda^6) < 0. \quad (4.24)$$

Hence, from eq. (4.14), within the SM and under the aforementioned model-dependent assumptions, we have

$$[\cos \varphi_0]_{\text{large } N_C}^{\text{SM}} < 0. \quad (4.25)$$

Together with the result of section 3,  $\cos^2 \varphi_0 = 0.96 \pm 0.03$ , we have

$$[\cos \varphi_0]_{\text{large } N_C}^{\text{SM}} = -0.98 \pm 0.02. \quad (4.26)$$

Note that the error combines both a statistical error from experiment as well as a theory component, i.e. is to be interpreted as an estimate of the total uncertainty.

## 5 Going beyond the two-fold ambiguity

In order to determine the sign of  $\sin \varphi_0$  and get rid of the remaining ambiguity, we would need to determine the signs and magnitudes of the competing short-distance and long-distance dispersive contributions,

$$\begin{aligned} \sin \varphi_0 &= \frac{\text{Im}[A(K_S)^* A(K_L)]}{|A(K_S) A(K_L)|} \\ &= - \frac{\left([A_{LD}]_{\text{dispersive}} + \sqrt{2}|A_{SD}| \cos \theta_{SD}\right) \text{sgn}[\sin(\theta_{SD})]}{|A(K_L)|}. \end{aligned} \quad (5.1)$$

We recall that  $[A_{LD}]_{\text{dispersive}}$  can be written as

$$[A_{LD}]_{\text{dispersive}} = \sqrt{2}|A_{LD}| \cos \Delta\delta = C_{\text{had.}} \left(A_{LD}^{\text{local}} + \text{Re} A_{LD}^{\gamma\gamma}\right), \quad (5.2)$$

see eq. (4.1). Existing semi-phenomenological theory estimations of  $[A_{LD}]_{\text{dispersive}}$  come with large theory uncertainties. Using the estimate of ref. [23] as in eq. (4.11), updated with existing data (see appendix A), we have for the long-distance contribution,

$$\begin{aligned} \frac{[A_{LD}]_{\text{dispersive}}}{|A(K_L \rightarrow \mu^+ \mu^-)|} &= \frac{C_{\text{had.}}}{|A(K_L \rightarrow \mu^+ \mu^-)|} \left(A_{LD}^{\text{local}} + \text{Re} A_{LD}^{\gamma\gamma}\right) \\ &= \frac{C_{\text{had.}} |A_{\gamma\gamma}|}{|A(K_L \rightarrow \mu^+ \mu^-)|} \frac{\alpha_{em} m_\mu}{\pi m_K} \left( [\chi_{\gamma\gamma}(m_\rho)]^{\text{IU}} + \left[ \text{Re} C_{\gamma\gamma} - \frac{5}{2} + \frac{3}{2} \log \left( \frac{m_\mu^2}{m_\rho^2} \right) \right]^{\text{IU}} \right) \\ &= \text{sgn}[C_{\text{had.}}] \sqrt{\frac{2\beta_\mu}{R_{K_L}}} \frac{\alpha_{em} m_\mu}{\pi m_K} [(6.10 \pm 1.01) - 5.14] \in [-0.009, 0.37], \end{aligned} \quad (5.3)$$

where  $\text{sgn}[C_{\text{had.}}] > 0$  derived in the previous section is used.

For the short-distance SM contribution, we have (see eq. (4.22))

$$\begin{aligned} \frac{A(K_L)_{SD}^{\text{SM}}}{|A(K_L \rightarrow \mu^+ \mu^-)|} &= \frac{\sqrt{2} G_F \alpha_{em}(m_Z)}{\pi \sin^2 \theta_W} \frac{|V_{ts}^* V_{td} Y(x_t) + V_{cs}^* V_{cd} Y_{NL}| f_K m_\mu m_K \cos \theta_{SD}^{\text{SM}}}{|A(K_L \rightarrow \mu^+ \mu^-)|} \\ &= -0.331 \pm 0.008, \end{aligned} \quad (5.4)$$

where we use the following inputs,

$$\begin{aligned}
 Y(x_t) &= 0.931 \pm 0.005 \text{ [3]}, & Y_{\text{NL}} &= (2.95 \pm 0.46) \times 10^{-4} \text{ [43]}, \\
 A &= 0.790^{+0.017}_{-0.012}, & \lambda &= 0.22650 \pm 0.00048, \\
 \bar{\rho} &= 0.141^{+0.016}_{-0.017}, & \bar{\eta} &= 0.357 \pm 0.011, \\
 m_K &= 497.61 \text{ MeV}, & m_\mu &= 105.658 \text{ MeV}, \\
 G_F &= 1.166378 \times 10^{-5} \text{ GeV}^{-2}, & f_K &= 155.7 \text{ MeV}, \\
 \alpha_{em} &= 1/137, & \alpha_{em}(m_Z) &= 1/129, \\
 \sin^2 \theta_W &= 0.23, & m_\rho &= 775.26 \text{ MeV}.
 \end{aligned} \tag{5.5}$$

In particular, the CKM input results in  $\cos \theta_{SD}^{\text{SM}} = -0.94$ .

Hence, the large theory uncertainty on the dispersive long-distance contribution, as reflected in the range given in eq. (5.3), does not allow to determine if it is larger or smaller than the short-distance SM contribution, eq. (5.4). We conclude that with current knowledge on the dispersive long-distance contribution the sign of  $\sin \varphi_0$  cannot be determined.

Therefore, we have, using Gaussian error propagation,

$$[\cos \varphi_0]_{\text{large } N_C}^{\text{SM}} = -0.98 \pm 0.02, \quad [\sin \varphi_0]_{\text{large } N_C, \text{theory}}^{\text{SM}} = \pm(0.21 \pm 0.07). \tag{5.6}$$

We note that due to the nature of the theoretical error, the error should not be interpreted as a statistical error but rather as an estimate of the uncertainty.

## 6 Discussion and conclusions

In this work we have related the phase shift,  $\varphi_0$ , appearing in the time-dependent decay rate of a neutral kaon to a dimuon pair to the ratio of integrated rates of  $K_L \rightarrow \mu^+ \mu^-$  and  $K_L \rightarrow \gamma \gamma$ . This relation holds to an excellent approximation under the well-motivated assumption that the two-photon intermediate state dominates the absorptive contribution, and within any model in which the short-distance leptonic current is axial or vectorial, as in the SM. The only input required other than the ratio of integrated rates is a coefficient calculated within QED. We find that  $\cos^2 \varphi_0$  is precisely predicted model independently, and given by

$$\cos^2 \varphi_0 = 0.96 \pm 0.02_{\text{exp}} \pm 0.02_{\text{th}}. \tag{6.1}$$

The experimental error comes from the error of  $R_{K_L}$  and the theory error is our estimate of the size of higher order QED corrections and the contribution from other intermediate states beside the di-photon state. The result leaves a four-fold ambiguity in  $\varphi_0$ .

The phase shift,  $\varphi_0$ , is also of experimental significance since it controls the integrated number of interference events. For a  $K^0$  beam a positive value of  $\cos \varphi_0$  is preferred as it enhances the interference and improves the feasibility of extracting clean short-distance information from the interference term. For a  $\bar{K}^0$  beam the situation is reversed, and a negative value of  $\cos \varphi_0$  is preferred. Thus, for experiments employing a proton beam on target, where the number of  $K^0$  particles is expected to exceed that of  $\bar{K}^0$  particles, a positive  $\cos \varphi_0$  would be preferred.

We were unable to determine the sign of  $\cos \varphi_0$  completely model independently. Thus, with the use of several assumptions, that is, within the framework of ChPT and using a typical parameter region (motivated by the large- $N_C$  limit, as detailed in section 4.2), the sign of  $\cos \varphi_0$ , relative to the short-distance contribution, can be predicted. We find that within this framework, and assuming that the short-distance contribution is SM-like, there is a theory preference towards  $\cos \varphi_0 < 0$ . New Physics can potentially yield a different sign for  $\cos \varphi_0$ . A measurement of the angle  $\varphi_0$  is therefore a test of the validity of several assumptions, pertaining to both the short-distance physics and the ChPT description of the dispersive long-distance physics.

Given the assumptions that were made to arrive to the conclusion about the sign of  $\cos \varphi_0$ , and the fact that the prediction can be modified for models beyond the SM, we conclude that neither solution is unequivocally theoretically favored. That is, we cannot conclude that we know the sign of  $\cos \varphi_0$  to high confidence. Thus, when planing to perform the experiment we encourage the experimental collaborations to consider both possible signs for  $\cos \varphi_0$  for purposes of sensitivity estimations.

We therefore conclude by emphasizing that the time dependence of the kaon decay rate to two muons provides two very clean SM predictions:

1. The coefficient of the interference term allows the extraction of the theoretically clean decay rate  $\mathcal{B}(K_S \rightarrow \mu^+ \mu^-)_{\ell=0}$ . In the SM, this observable is proportional to the CKM combination  $|V_{ts} V_{td} \sin(\beta + \beta_s)|$ .
2. Although it includes long-distance as well as short-distance physics, the phase shift in the interference term,  $\varphi_0$ , is predicted cleanly up to a four-fold ambiguity.

Thus, an experiment that performs the time-dependence studies of the  $K \rightarrow \mu^+ \mu^-$  decay rate provides two independent tests of the SM from the same measurement. While the phase  $\varphi_0$  is not determined model independently, it directly impacts the measurement of the  $K \rightarrow \mu^+ \mu^-$  time dependent rate as it affects the interference between the  $K_S$  and  $K_L$  amplitudes in the total rate. The phase shift  $\varphi_0$  is therefore a quantity of critical importance in kaon physics, as a way to test the SM, and in extension, as a probe of new physics beyond the SM because of its sensitivity to short distance effects.

## Acknowledgments

We thank Joachim Brod, Hector Gisbert, Antonio Pich, and Emmanuel Stamou for useful discussions. Y.G. is supported in part by the NSF grant PHY1316222. S.S. is supported by a Stephen Hawking Fellowship from UKRI under reference EP/T01623X/1 and the Lancaster-Manchester-Sheffield Consortium for Fundamental Physics, under STFC research grant ST/T001038/1. The work of T.K. is supported by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Early-Career Scientists (Grant No. 19K14706) and the JSPS Core-to-Core Program (Grant No. JPJSCCA20200002). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising. This work uses existing data which is available at locations cited in the bibliography.

## A Update of the theory estimate for $\chi_{\gamma\gamma}(\mu)$

In this section we update the theory estimate of ref. [23]. The experimental input that goes into this estimate comes from analyses of the form factor in  $K_L \rightarrow \gamma e^+ e^-$ ,  $K_L \rightarrow \gamma \mu^+ \mu^-$  and  $K_L \rightarrow e^+ e^- \mu^+ \mu^-$ , and can be summarized by the parameter  $\alpha_{\text{exp.}}$ , defined by

$$\alpha_{\text{exp.}} = -m_\rho^2 \frac{d}{dq^2} f(q^2, 0) \Big|_{q^2=0}, \quad (\text{A.1})$$

where  $f(q_1^2, q_2^2)$  is the  $K_L \rightarrow \gamma\gamma$  form factor. The experimental value derived from  $K_L \rightarrow \gamma e^+ e^-$  has been updated by the KTeV collaboration after ref. [23] was published. We take [44]

$$\alpha_{\text{exp.}}|_{ee} = -1.73 \pm 0.05, \quad (\text{A.2})$$

combined with [45, 46]

$$\begin{aligned} \alpha_{\text{exp.}}|_{\mu\mu} &= -1.54 \pm 0.10, \\ \alpha_{\text{exp.}}|_{ee\mu\mu} &= -1.59 \pm 0.37, \end{aligned} \quad (\text{A.3})$$

and arrive at the weighted average

$$\alpha_{\text{exp.}} = -1.691 \pm 0.044. \quad (\text{A.4})$$

Comparing with ref. [23] (who quote  $\alpha_{\text{exp.}} = -1.611 \pm 0.044$ ), the central value has gone up by  $\sim 5\%$  while the relative error remains the same.

Inserting this into eq. (22) of ref. [23], we have

$$\chi_{\gamma\gamma}(m_\rho) = (6.10 \pm 0.16_{\text{exp.}}) - \Delta_\Lambda, \quad (\text{A.5})$$

(updated from  $(5.83 \pm 0.15_{\text{exp.}}) - \Delta_\Lambda$ ). Following ref. [23] we take

$$|\Delta_\Lambda| \leq 1.0, \quad (\text{A.6})$$

and reach the result

$$\left[ \chi_{\gamma\gamma}(m_\rho) \right]_{\text{IU}} = 6.10 \pm 0.16_{\text{exp.}} \pm 1.0_{\text{th.}}. \quad (\text{A.7})$$

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP<sup>3</sup> supports the goals of the International Year of Basic Sciences for Sustainable Development.

## References

- [1] G. D'Ambrosio and T. Kitahara, *Direct CP Violation in  $K \rightarrow \mu^+ \mu^-$* , *Phys. Rev. Lett.* **119** (2017) 201802 [[arXiv:1707.06999](https://arxiv.org/abs/1707.06999)] [[INSPIRE](#)].
- [2] A. Dery, M. Ghosh, Y. Grossman and S. Schacht,  *$K \rightarrow \mu^+ \mu^-$  as a clean probe of short-distance physics*, *JHEP* **07** (2021) 103 [[arXiv:2104.06427](https://arxiv.org/abs/2104.06427)] [[INSPIRE](#)].

- [3] J. Brod and E. Stamou, *Impact of indirect CP violation on  $Br(K_S \rightarrow \mu^+ \mu^-)_{\ell=0}$* , DO-TH 22/24 (2022), [arXiv:2209.07445](#) [INSPIRE].
- [4] N.H. Christ and T.D. Lee, *CP Nonconservation and Inequalities Between  $\mu^+ \mu^-$  and  $2 \gamma$  Decay Rates of  $K_S^0$  and  $K_L^0$* , *Phys. Rev. D* **4** (1971) 209 [INSPIRE].
- [5] H. Stern and M.K. Gaillard, *Review of the  $K_L \rightarrow \mu^+ \mu^-$  puzzle*, *Annals Phys.* **76** (1973) 580 [INSPIRE].
- [6] L.M. Sehgal,  *$K_2^0 \rightarrow \mu^+ + \mu^-$  as a test of neutral lepton currents*, *Nuovo Cim. A* **45** (1966) 785 [INSPIRE].
- [7] L.M. Sehgal, *Electromagnetic contribution to the decays  $K_S \rightarrow l \bar{l}$  and  $K_L \rightarrow l \bar{l}$* , *Phys. Rev.* **183** (1969) 1511 [INSPIRE].
- [8] A.R. Clark et al., *Experimental Limits on the Decays  $K_L^0 \rightarrow \mu^+ \mu^-$ ,  $e^+ e^-$ , and  $\mu^\pm e^\mp$* , *Phys. Rev. Lett.* **26** (1971) 1667 [INSPIRE].
- [9] R.G. Arnold et al., *A Measurement of the Ratio of the Decay Rates  $K_L^0 \rightarrow 2\gamma/K_L^0 \rightarrow 3\pi^0$* , *Phys. Lett. B* **28** (1968) 56 [INSPIRE].
- [10] H.H. Chen, K. Kawarabayashi and G.L. Shaw, *Comments on the unitarity bound in  $K_L^0 \rightarrow \mu^+ \mu^-$* , *Phys. Rev. D* **4** (1971) 3514 [INSPIRE].
- [11] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01 [INSPIRE].
- [12] V. Chobanova et al., *Probing SUSY effects in  $K_S^0 \rightarrow \mu^+ \mu^-$* , *JHEP* **05** (2018) 024 [[arXiv:1711.11030](#)] [INSPIRE].
- [13] M. Endo, T. Goto, T. Kitahara, S. Mishima, D. Ueda and K. Yamamoto, *Gluino-mediated electroweak penguin with flavor-violating trilinear couplings*, *JHEP* **04** (2018) 019 [[arXiv:1712.04959](#)] [INSPIRE].
- [14] A. Dery and M. Ghosh,  *$K \rightarrow \mu^+ \mu^-$  beyond the standard model*, *JHEP* **03** (2022) 048 [[arXiv:2112.05801](#)] [INSPIRE].
- [15] A.J. Buras and E. Venturini, *Searching for New Physics in Rare K and B Decays without  $|V_{cb}|$  and  $|V_{ub}|$  Uncertainties*, *Acta Phys. Polon. B* **53** (2021) A1 [[arXiv:2109.11032](#)] [INSPIRE].
- [16] N.H. Christ, X. Feng, L. Jin, C. Tu and Y. Zhao, *Lattice QCD calculation of the two-photon contributions to  $K_L \rightarrow \mu^+ \mu^-$  and  $\pi^0 \rightarrow e^+ e^-$  decays*, *PoS LATTICE2019* (2020) 128 [INSPIRE].
- [17] Y. Zhao and N.H. Christ, *Calculating  $K \rightarrow \gamma\gamma$  using lattice QCD*, *PoS LATTICE2021* (2022) 451 [INSPIRE].
- [18] N. Christ, X. Feng, L. Jin, C. Tu and Y. Zhao, *Lattice QCD calculation of  $\pi^0 \rightarrow e^+ e^-$  decay*, [arXiv:2208.03834](#) [INSPIRE].
- [19] LHCb collaboration, *Constraints on the  $K_S^0 \rightarrow \mu^+ \mu^-$  Branching Fraction*, *Phys. Rev. Lett.* **125** (2020) 231801 [[arXiv:2001.10354](#)] [INSPIRE].
- [20] M. Fernandez Gomez on behalf of the LHCb Collaboration, *Search for  $K^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$* , talk given at the *International Conference on Kaon Physics (KAON 2022)*, Osaka, Japan, 13–16 September 2022, <https://conference-indico.kek.jp/event/169/contributions/3464/>.
- [21] B.R. Martin, E. De Rafael and J. Smith, *Neutral kaon decays into lepton pairs*, *Phys. Rev. D* **2** (1970) 179 [INSPIRE].

- [22] D. Gomez Dumm and A. Pich, *Long distance contributions to the  $K_L \rightarrow \mu^+ \mu^-$  decay width*, *Phys. Rev. Lett.* **80** (1998) 4633 [[hep-ph/9801298](#)] [[INSPIRE](#)].
- [23] G. Isidori and R. Unterdorfer, *On the short distance constraints from  $K_{L,S} \rightarrow \mu^+ \mu^-$* , *JHEP* **01** (2004) 009 [[hep-ph/0311084](#)] [[INSPIRE](#)].
- [24] M. Knecht, S. Peris, M. Perrottet and E. de Rafael, *Decay of pseudoscalars into lepton pairs and large- $N_c$  QCD*, *Phys. Rev. Lett.* **83** (1999) 5230 [[hep-ph/9908283](#)] [[INSPIRE](#)].
- [25] G. D'Ambrosio, G. Isidori and J. Portoles, *Can we extract short distance information from  $B(K_L \rightarrow \mu^+ \mu^-)$ ?*, *Phys. Lett. B* **423** (1998) 385 [[hep-ph/9708326](#)] [[INSPIRE](#)].
- [26] V. Cirigliano, G. Ecker, H. Neufeld, A. Pich and J. Portoles, *Kaon Decays in the Standard Model*, *Rev. Mod. Phys.* **84** (2012) 399 [[arXiv:1107.6001](#)] [[INSPIRE](#)].
- [27] FLAVOUR LATTICE AVERAGING GROUP (FLAG) collaboration, *FLAG Review 2021*, *Eur. Phys. J. C* **82** (2022) 869 [[arXiv:2111.09849](#)] [[INSPIRE](#)].
- [28] R. Escribano, S. González-Solís, P. Masjuan and P. Sanchez-Puertas,  *$\eta'$  transition form factor from space- and timelike experimental data*, *Phys. Rev. D* **94** (2016) 054033 [[arXiv:1512.07520](#)] [[INSPIRE](#)].
- [29] L. Gan, B. Kubis, E. Passemar and S. Tulin, *Precision tests of fundamental physics with  $\eta$  and  $\eta'$  mesons*, *Phys. Rept.* **945** (2022) 1 [[arXiv:2007.00664](#)] [[INSPIRE](#)].
- [30] J.F. Donoghue, B.R. Holstein and Y.C.R. Lin, *Chiral perturbation theory corrections to  $K_L \rightarrow \gamma\gamma$* , *Nucl. Phys. B* **277** (1986) 651 [[INSPIRE](#)].
- [31] X.G. He, C.S. Huang and X.Q. Li,  *$SU(3)$  and nonet breaking effects in  $K_L \rightarrow \gamma\gamma$  induced by  $s \rightarrow d$  gluon due to anomaly*, *Phys. Rev. D* **67** (2003) 096005 [[hep-ph/0210435](#)] [[INSPIRE](#)].
- [32] J.F. Donoghue, E. Golowich and B.R. Holstein, *Long Distance Chiral Contributions to the  $K_L K_S$  Mass Difference*, *Phys. Lett. B* **135** (1984) 481 [[INSPIRE](#)].
- [33] H.-Y. Cheng, *How Good Is Nonet Symmetry for Pseudoscalar Mesons?*, *Phys. Lett. B* **245** (1990) 122 [[INSPIRE](#)].
- [34] G. D'Ambrosio and J. Portoles, *Analysis of  $K_L \rightarrow \pi^+ \pi^- \gamma$  in chiral perturbation theory*, *Nucl. Phys. B* **533** (1998) 523 [[hep-ph/9711210](#)] [[INSPIRE](#)].
- [35] A. Pich and E. de Rafael, *Weak  $K$  amplitudes in the chiral and  $1/N_c$  expansions*, *Phys. Lett. B* **374** (1996) 186 [[hep-ph/9511465](#)] [[INSPIRE](#)].
- [36] J.-M. Gerard, C. Smith and S. Trine, *Radiative kaon decays and the penguin contribution to the  $\Delta I = 1/2$  rule*, *Nucl. Phys. B* **730** (2005) 1 [[hep-ph/0508189](#)] [[INSPIRE](#)].
- [37] RBC, UKQCD collaborations, *Direct CP violation and the  $\Delta I = 1/2$  rule in  $K \rightarrow \pi\pi$  decay from the standard model*, *Phys. Rev. D* **102** (2020) 054509 [[arXiv:2004.09440](#)] [[INSPIRE](#)].
- [38] RBC, UKQCD collaborations, *Standard Model Prediction for Direct CP Violation in  $K \rightarrow \pi\pi$  Decay*, *Phys. Rev. Lett.* **115** (2015) 212001 [[arXiv:1505.07863](#)] [[INSPIRE](#)].
- [39] T. Blum et al.,  *$K \rightarrow \pi\pi$   $\Delta I = 3/2$  decay amplitude in the continuum limit*, *Phys. Rev. D* **91** (2015) 074502 [[arXiv:1502.00263](#)] [[INSPIRE](#)].
- [40] T. Blum et al., *The  $K \rightarrow (\pi\pi)_{I=2}$  Decay Amplitude from Lattice QCD*, *Phys. Rev. Lett.* **108** (2012) 141601 [[arXiv:1111.1699](#)] [[INSPIRE](#)].
- [41] A. Pich and A. Rodríguez-Sánchez,  *$SU(3)$  analysis of four-quark operators:  $K \rightarrow \pi\pi$  and vacuum matrix elements*, *JHEP* **06** (2021) 005 [[arXiv:2102.09308](#)] [[INSPIRE](#)].

- [42] G. Buchalla, A.J. Buras and M.E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev. Mod. Phys.* **68** (1996) 1125 [[hep-ph/9512380](#)] [[INSPIRE](#)].
- [43] M. Gorbahn and U. Haisch, *Charm Quark Contribution to  $K(L) \rightarrow \mu^+ \mu^-$  at Next-to-Next-to-Leading*, *Phys. Rev. Lett.* **97** (2006) 122002 [[hep-ph/0605203](#)] [[INSPIRE](#)].
- [44] KTeV collaboration, *Measurements of the Decay  $K_L \rightarrow e^+ e^- \gamma$* , *Phys. Rev. Lett.* **99** (2007) 051804 [[hep-ex/0702039](#)] [[INSPIRE](#)].
- [45] KTeV collaboration, *Measurement of the branching ratio and form-factor of  $K_L \rightarrow \mu^+ \mu^- \gamma$* , *Phys. Rev. Lett.* **87** (2001) 071801 [[INSPIRE](#)].
- [46] KTeV collaboration, *Measurements of the Decay  $K_L \rightarrow e^+ e^- \mu^+ \mu^-$* , *Phys. Rev. Lett.* **90** (2003) 141801 [[hep-ex/0212002](#)] [[INSPIRE](#)].